IRVING FISHER COMMITTEE
ON CENTRAL-BANK STATISTICS

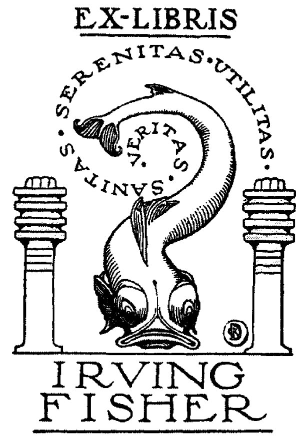
ifc Bulletin

No. 20 • April 2005

Contents

Proceedings
IFC Conference Basel, 2004

Documents Session 2
Workshops C, D and E



Irving Fisher Committee on Central-Bank Statistics

Executive Body:
Paul Van den Bergh (Chair)
Almut Steger
Rudi Acx
Radha Binod Barman
Kenneth Coates
Bart Meganck
Marius van Nieuwkerk
João Cadete de Matos

IFC Secretariat:
Rudi Acx
National Bank of Belgium
De Berlaimontlaan 14
1000 Brussels, Belgium
Tel.: +32-2-221 2403
Fax: +32-2-221 3230
E-mail: rudi.acx@nbb.be

IFC-Bulletin

Editor:
David S. Bieri

Editorial Address:
Bank for International Settlements
MED Publications
Centralbahnplatz 2
CH–4002 Basel
Switzerland
E-mail: service.MED-publications@bis.org

The IFC-Bulletin is published at irregular intervals. Subscriptions are available free of charge. Send requests to the IFC Office.

ifc Bulletin

No 20 – April 2005

Contents

What is the IFC?

The Irving Fisher Committee (IFC) is a forum for discussion on statistical issues that are of interest to central banks. The Committee, which derives its name from the great American economist and statistician Irving Fisher, is part of the International Statistical Institute (ISI).

Objectives

By providing a forum for discussion, the IFC aims at:

- participating in the discussion on adapting statistical systems to changing requirements;
- promoting the adoption of international statistical standards and methodologies;
- sharing experience on the development of new statistics and the implementation of new methods of collecting, compiling and disseminating statistical information;
- exchanging views between central bankers and academics on statistical methods and techniques;
- facilitating personal contacts between central-bank statisticians.

Strategy

To achieve its objectives, the IFC organizes conferences, which take place both inside and outside the framework of the ISI's biennial Sessions. The first "outside" conference – on the challenges to central bank statistical activities – is scheduled for summer 2002 at the Bank for International Settlements in Basle.

The conferences are supported by the publication of the IFC Bulletin, which contains the conference papers and other articles.

The IFC has a Web site (http://www.ifcommittee.org), on which an electronic version of the IFC Bulletin can be found.

What kind of topics are discussed?

Any kind of theoretical or practical statistical subject that has a relationship with the activities of central banks can be considered for discussion. The subjects will mostly be in the area of monetary, financial and balance of payments statistics.

Membership and Structure

Central banks and other institutions interested in statistical systems and statistical techniques that have a bearing on the collection, compilation and distribution of central-bank statistics can become members by simple application. Members are entitled to appoint delegates to participate in the IFC's activities and to contribute to its conferences by presenting papers.

The prime decision-taking body is the assembly of members' delegates at the "administrative meetings" that are organized during the conferences. Here the IFC's strategy is determined. At these meetings an Executive Body is elected, which is charged with the committee's day-to-day business and with the preparation of the "administrative meetings". Likewise, at the "administrative meetings" topics are proposed for future conferences.

A Short History

The Irving Fisher Committee (IFC) was established on the initiative of a number of central banks statisticians who were attending the ISI Corporate Members Meeting at the 1995 ISI Session in Beijing.

In 1997, during the 51st ISI Sessions in Istanbul, the IFC held its inaugural meeting. At the "administrative meeting" held during that Session an Executive Body was established and it was decided to start publishing the IFC Bulletin devoted to the activities of the IFC. Two years later, at the 52nd ISI Session in Helsinki, the IFC's presence was further strengthened. In 2001, at the 53rd ISI Session in Seoul, the IFC presented a programme comprising an invited papers meeting on "Financial Stability Statistics" and several contributed papers meetings.

In 2002, a conference on "Challenges to Central Bank Statistical Activities" was organised in co-operation with the Bank for International Settlements (BIS), which hosted it at its premises in Basle. 160 statisticians representing 73 countries participated. Some 50 papers were presented. In 2004, another IFC "Basel Conference" was held in cooperation with the BIS. Some 150 statisticians, mainly central bankers, and originating from 65 countries, discussed "Central Bank issues regarding Financial and National Accounts" in three sessions and eight workshops.

In 2003, at the 54^{th} ISI Session in Berlin, the IFC participated with nearly 40 papers, presented in two Invited Papers Meetings and three Contributed Papers Meetings. The General Assembly of the ISI accorded to the IFC the Status of an independent ISI Section on a provisional basis up to the 55^{th} ISI Session (Sydney, April 2005).

IFC Bulletin

The IFC Bulletin is the official periodical of the Irving Fisher Committee. The Bulletin contains articles and the text of papers presented within the framework of the ISI Conferences. Institutions and individuals active in the field of central-bank statistics can subscribe to the Bulletin free of charge.

Overview of past, current and future developments of the IFC

Summary by the IFC Chair, March 2005

The Irving Fisher Committee on Central Bank Statistics (IFC) was established in 1997 on the initiative of a number of central bank statisticians attending the 1995 Session of the International Statistical Institute (ISI) in Beijing. Its objective was to promote cooperation amongst central banks on statistical methodological issues through the sponsoring of meetings in the context of biennial ISI sessions (Istanbul 1997; Helsinki, 1999; Korea, 2001; Berlin 2003). In 2002 the BIS organised the first "independent" conference of the IFC on the theme of "Challenges to Central Bank Statistical Activities". 160 statisticians representing 73 countries participated in this conference and in all more than 50 papers were presented. The latter were published in the IFC Bulletin, the official publication of the Committee. Information on the IFC structure and activities can be found on www.ifcommittee.org.

IFC Conference 2004

In 2004, the BIS supported the IFC in the organisation of a second independent conference on "Central Bank Issues Regarding National and Financial Accounts". The discussion focused on three major themes: (i) price indices, (ii) output, output gaps and productivity and (iii) financial accounts. This time 150 statisticians and economists from 65 central banks attended and 58 papers were presented. The proceedings of the conference have been, or will be published, in the IFC Bulletins No 19, 20 and 21.

The debates at the second independent IFC conference confirmed that there are a number of clear policy issues that emerge from discussions amongst central bank experts on statistical methodological questions. Regarding price indices, for instance, a number of different indices are now being used which each suppose to measure inflation in one way or another. This could potentially lead to confusion, particularly at the current relatively low levels of overall inflation, and make the task of communicating monetary policy targets and outcomes more difficult. With regard to measures of real variables in national accounts, it is clear that data in volume are based on a number of (sometimes very rough) estimates and assumptions which, moreover, differ considerably from one country to another. Statistics derived from these real variables, such as output gaps and productivity have thus to be interpreted with the necessary care. Finally, though a lot of efforts are put into the development of a coherent set of financial accounts with breakdowns for various sectors and instruments, the data are not always sufficient or timely for the purpose of analysing financial stability issues. More detailed information is required, for instance, on balance sheet positions and exposures of all sectors, on asset prices including those on real estate, and on indicators of liquidity in financial markets.

There was a general consensus at the conference that in these and other statistical areas of key importance to central banks, international methodologies were being developed or implemented without an adequate input of central bank compilers and analysts. It was noted on a number of occasions that the IFC could play a useful role in this respect, though care had to be taken not to duplicate efforts being made in existing international expert groups.

Future of the IFC

At the 54th ISI Session in Berlin in 2003, the IFC was granted provisional recognised section status of the ISI on a number of conditions. This included refocusing the domain of the Committee on monetary and financial statistics, opening up its membership to a broader group of experts from the private sector and the academic community and adapting the name of the

Committee to reflect these changes. Other ISI requirements would also apply, such as drafting formal statutes for the Committee and introducing a membership fee.

During and following the Berlin meeting, various IFC members expressed reservations about changing the direction of the IFC as proposed by the ISI. Many central bank members seemed to be in favour of a much more active and dynamic IFC, if possible with strong support from the BIS. Strong cooperation amongst central banks was seen to be the priority and a prerequisite for a sustained interaction with non-central bank statisticians. The IFC Executive Body therefore decided to consult the whole IFC membership before going forward with changing the objectives and governance of the IFC in order to become a recognised official ISI section. A survey was carried out in July–August 2004 on the possible future direction and activities of the IFC. 62 replies were received out of the 80 listed institutional members and 300 individual names on the IFC contact list. Of these, 34 responses represented "institutional" views of the respective central banks.

The results of the survey[1] indicates that the central bank institutional members prefer to continue to operate in the current informal way as an ISI committee, preferably also under the umbrella of the BIS. The individual members are more in favour of a formal ISI section status on the lines proposed by the ISI. Both institutional and individual respondents express a preference for formal statutes. All respondents consider the discussion of statistical issues of interest to central banks as the main objective of the IFC. These include the collection, compilation, analysis and dissemination across a broad range of statistical domains of interest to central banks, with special attention to key areas for which central banks typically have a direct responsibility. Regarding the possible membership structure, both institutional and individual respondents agree that central banks and central bank staff as individuals should constitute the core membership and there is resistance to admitting experts from the private sector as core members. Regarding the openness of the IFC to the academic world, institutional respondents see academics as non-core members, while individual respondents would prefer academics as core members.

The survey results were discussed with key BIS member central banks in the margin of the September 2004 IFC conference. The support provided by the BIS for the IFC activities was very much appreciated and was seen to be essential to ensure that the IFC could continue to operate as an active international discussion forum for central bank economists and statisticians. The hope was expressed that the IFC Executive Body could negotiate a compromise solution with the ISI that would take into account the interests of the key central bank institutional members of the IFC as well as the broader objectives of the ISI and the non-institutional IFC members.

The IFC Executive Body has prepared draft statutes for the IFC as the basis for the negotiation with the ISI. The draft defines the IFC as a focus group of the central bank statistical community, comprising both compilers and users of economic, monetary and financial statistics. However, it clearly expresses the interest of this community in working closely with other sections and committees of the ISI. A simple membership structure is proposed consisting of institutional and associate members. The governance structure would likewise be simple and transparent with a Council representing the institutional members, a small Executive consisting of a Chair and four Vice-Chairs and a secretariat (the Chairs and Vice-Chairs would be senior central bank officials and would rotate on a two-three year basis). The fee structure would differentiate between institutional and associate members. The IFC would use the services of the ISI Permanent Office for the collection of the fees and for the communication with the ISI and its different bodies.

The draft statutes were presented to, and discussed by, the ISI Executive Committee in December 2004. The ISI expressed a clear interest in associating the central banking community with the ISI. It understood that the central banks have an interest in a wide range of statistical domains and activities and that they would like to cooperate with various ISI Sections and Committees. The ISI was pleased with the IFC's enthusiastic contributions to ISI activities and recommended that the IFC continue developing its plans, and that ISI would examine how the IFC could best fit within the ISI organisational structure. The latter was under review as part of the strategic thinking currently taking place on the future of the ISI. During the Sydney ISI Session, the ISI structure would be discussed after which the EC would be better able to review the proposed statutes of the IFC. A clear plan regarding the IFC's proposed future direction should then be developed during the next two years. For the present the IFC should retain its provisional section status.

1 *The survey results are available from the Chair or the Secretariat of the IFC.*

In the context of these ongoing negotiations, senior officials of the National Bank of Belgium and the BIS have held exploratory discussions on the future governance structure of the IFC and the support that the BIS could provide to the IFC in the longer run. The National Bank, as one of the founding members of the IFC and one of its most active contributors[2], would be willing to take over the chair of the Committee after Sydney for a two-three year period. In that capacity it would try to rally the support of the key BIS member central banks for a more permanent and formal IFC, operating under the umbrella of both the ISI and the BIS. At some point the BIS would be approached with a request to take on the Secretariat function of the IFC.

The 55th ISI Session in Sydney in 2005

The IFC is sponsoring the organisation of different meetings in the context of the forthcoming 55th Session of the ISI in Sydney (5–12 April 2005). The meetings will cover costs, quality and relevance of financial statistics; financial soundness indicators; accounting standards and their impact on financial statistics; optimal methods for data quality improvements in financial statistics; and the use of surveys in financial statistics. A number of the meetings will be part of a special theme day on Finance and Statistics, during which the Deputy Governor of the Reserve Bank of Australia will deliver a keynote speech. The Reserve Bank of Australia will also organise a social event for the central bank statisticians attending the ISI Session. The latter have also been invited to attend the Fourth International Symposium on Business and Industrial Statistics held in Cairns immediately after the Sydney Session. This symposium is sponsored by the Association of Statistics on Business and Industry, one of the ISI groups with which the IFC would like to cooperate in the future.

Joint Bank of Canada/IFC Workshop in 2005

Following the IFC Conference in September 2004, the Bank of Canada has proposed to sponsor a joint workshop with the IFC on Data Requirements for Analysing the Stability and Vulnerabilities of Mature Financial Systems. This workshop will take place in Ottawa from 20–21 June. Participation will be by invitation only but will include, apart from central banks from the major financial centres, representatives from national institutes, international organisations, the academic community and the private sector. The background documents and summaries of the discussions will be published in the IFC Bulletin.

Independent conference in 2006

Informal discussions are taking place to determine possible topics for the independent conference of the IFC in 2006, which is expected to take place again at the BIS. A list of possible topics is attached. It is based on various suggestions made by IFC members in the recent past.

The 56th ISI Session in Lisbon, Portugal in 2007

The 56th ISI Biennial Session will be held in Lisboa, Portugal from 22–27 August 2007. As usual, the IFC is ready to organise several Invited Paper Meetings and Contributed Paper Meetings of the event. In particular, it would like to use the ISI sessions to co-sponsor meetings with other ISI groups such as the International Organisation of Official Statistics, the International Association of Survey Statisticians or the Association of Statistics on Business and Industry (SBI). The IFC has submitted the following proposals for Invited paper Meetings to the Programme Committee:

2 *All institutional and practical arrangements concerning the IFC are currently informal. The National Bank of Belgium currently provides the Secretariat for the Committee and prints and distributes the IFC Bulletin. Various NBB officials have chaired the IFC in the past or have been members of its Executive Body. There is no specific mechanism for selecting members of the Executive Body. The Head of Information, Statistics and Administration at the BIS is currently the chair of the IFC. The BIS has taken on the commitment to edit the forthcoming editions of the IFC Bulletin and to maintain the IFC website.*

- Measuring Productivity, organiser: Bart Meganck (to be co-sponsored, if possible, with the IAOS)
- Measures of Output and Prices of Financial Services, organiser: Rodha Binod Barman
- Measures of Flows and Stocks in Financial Accounts, organiser: Rudi Acx (to be co-sponsored, if possible with the IAOS)
- Statistical Tools used in Financial Risk Management, organiser: Paul Van den Bergh (to be co-sponsored, if possible, with the SBI).

Apart from these IPM's the IFC would also propose a number of Contributed Paper Meetings, which are more open for participation. IFC members are invited to make proposals in this respect or to volunteer for the organisation of an event.

In the past, the IFC has tried to associate the central bank of the host country with the ISI Session and the IFC activities on that occasion. The Executive Body therefore agreed to invite Joao Cadete de Matos to join the IFC Executive. Joao has been an enthusiastic supporter of the IFC and the Executive Body looks forward to working with him.

I would like to encourage IFC members and supporters to provide me with any feedback they may have on this report. In particular, my colleagues on the Executive Body and I would welcome views as to possible future activities for the Committee and topics for discussion. More detailed information on the future governance of the IFC will be made available in the Bulletin as and when it becomes available.

Paul Van den Bergh

Annex

Possible list of Topics for Future IFC discussions

General methodological issues

Core concepts in statistical nomenclatures across statistical domains
Statistical revisions: policies and practices
International comparability of national economic and financial statistics
Usefulness of private sector statistical information/value of commercial databases
Impact of changes in accounting frameworks

National and Financial Accounts

Impact of planned changes to SNA and BOP Manuals
The residency criterion in a globalised world
Usefulness of regional economic data
Monthly GDP statistics, flash estimates of GDP
Productivity measures
Measures of capacity, capacity utilisations, output gaps
Measures of the informal economy
Measurement, valuation and reporting of tangible and intangible assets
Improving statistics of the government's sector fiscal positions (deficit/surplus, debt)
Measures of private sector wealth

Price statistics

Use of hedonic methods for price statistics
Calculation of price indices for the services sector, in particular financial services
Improving data on import prices
Price data needed to analyse exchange-rate pass-through effects
Statistics on the housing markets: prices, turnover, financing
Statistics from payment and settlement systems

Financial statistics (macro)

Financial soundness (macro-prudential) indicators
Improving data on non-bank financial institutions
Usefulness of real-time data from financial markets
Statistical requirements for analysing financial market activity
International comparability of financial statistics

Financial statistics (micro)

Statistical impact of Basel II
Statistical tools used by financial institutions for coping with Basel II
Statistical tools for measuring credit, market and operational risk

Survey Statistics

Use of surveys by central banks across statistical domains
Use of surveys of market participants/private sector expectations on economic and financial developments

Data collection and dissemination

Statistical quality assessment frameworks
Use of microdata
Respondents' fatigue, data relationship programmes
Data protection issues
Dissemination of statistics over the Internet

A word of thanks to Hans Van Wijk

As one of the founding fathers of the Irving Fisher Committee (IFC), Hans has been very dedicated to the activities of the IFC. Together with a handful of central bank statisticians, Hans participated in the ISI-Session in Bejing in 1995, where the idea was born to establish a group, under the umbrella of the ISI, of central banks statisticians in order to focus on specific issues of interest to central bank statistical experts. At the next ISI-Session in Istanbul in 1997, Hans presided, as first chairman of the IFC, the inaugural administrative meeting in presence of Dr. Zoltan Kenessy. That meeting kicked off the activities of the IFC and the number of institutional members increased rapidly afterwards. Very quickly the activities of IFC at the biannual ISI Sessions became well-known and respected.

The IFC-Bulletin, created shortly after the first administrative meeting in 1997, became the living link between the IFC members. Hans has been the Editor of the Bulletin since the very beginning and has remained very passionate about this publication: texts were presented in correct language, with attractive graphics and tables. Many authors can testify that Hans pressed them to respect submissions deadlines. Hans himself contributed to the Bulletin by writing articles on IFC activities. When the IFC-Bulletin went on the Internet, Hans also took charge of this. With the organisation of the independent conferences of the IFC in 2002 and 2004, the number of papers published in the Bulletin increased substantially. Moreover, as the IFC was starting discussions with the ISI about permanent session status, the Bulletin was also used to communicate ongoing developments to members on a regular basis.

Hans combined his function as Editor with that of member of the IFC Executive Body. In this capacity he launched many new ideas and showed a lot of initiative, which contributed further to the development of the IFC. Hans has always been preoccupied with, and contributed strongly to, the internationalisation of the IFC. He used his contacts in international statistical committees to increase the attention of professional statisticians in central banks and outside to the activities of the IFC.

At the 2004 independent conference in Basel, Hans announced that he would step down as Editor and Member of the IFC Executive Body. He was honoured at the conference for his outstanding contribution to central bank cooperation in the statistical area and for coaching the committee towards becoming a truly recognised and reputable international body involved in central bank statistical issues. On behalf of my colleagues of the Executive Body, I would like to take this opportunity to thank Hans again for his many efforts. We will ensure that his legacy is safeguarded and herewith nominate him as Honorary Member of the IFC. We are convinced that Hans will remain connected, and hopefully not only by reading the Bulletin, with the Irving Fisher Committee.

Paul Van den Bergh

SESSION 2

Output, capacity and productivity: key issues

Chair: Steven Keuning (European Central Bank)

Papers:

National accounts statistics: some key issues
R. Acx (National Bank of Belgium)

Comparing growth in GDP and labour productivity: measurement issues
Nadim Ahmad, François Lequiller, Pascal Marianna, Dirk Pilat, Paul Schreyer and Anita Wölfl (OECD)

Estimating the output gap in the Polish economy: the VECM approach
Michał Gradzewicz and Marcin Kolasa (National Bank of Poland)

Estimating capacity utilization from survey data
Norman Morin and John Stevens (Federal Reserve Board)

National accounts statistics: some key issues[1]

R. Acx (National Bank of Belgium)

1. Introduction

National accounts statistics do form one of the most comprehensive and challenging economic statistical products, as it tries to incorporate all relevant economic data and tendencies. It stems from its nature that a long list of key issues for national accounts statistics can be established.

However, many of these key issues can be regrouped under two headings:
- the *internal coherence* in the statistical system;
- the *international comparability* of the data presented.

2. The need for internal coherence

On the one hand there is the internal coherence, or consistency within national accounts, and on the other hand the coherence of national accounts data with other sets of macroeconomic statistics, such as the balance payments statistics, the statistics on government operations and the monetary and financial statistics, is very high ranked on the wish list of the users.

The internal coherence is of the utmost importance for the study and analysis of the development of productivity in an economy. The usefulness of indicators on productivity and competitiveness does heavily depend on the consistency of data within the national accounts. Especially reliable data on the volume of labour are at the centre of interest, and more precisely the number of hours worked, and reliable data on the labour costs for the calculation of parameters such as the unit labour cost and labour cost per hour worked.

It is obvious that internal coherence and the increasing demand for more detailed information and more functional breakdowns in the data do have their limits.

The SNA is primarily intended to provide data to meet the needs of analysts and policy makers interested in the behavior of the economy, and forms inevitably a compromise intended to yield the maximum benefits to different kinds of users and may therefore not be optimal for any one single purpose.

The SNA presents different ways in which the system may be adapted to meet differing circumstances and needs. Flexibility regarding the use of classifications, and the development of satellite or thematic accounts do offer a way out and will gain importance in the coming years. These satellite accounts (e.g. satellite accounts for education, health services, R&D, the non profit sector, tourism, environment, social accounting matrix) are closely linked to the main central system but are not bound to employ exactly all the same concepts. The setting of priorities in developing these thematic accounts, as they are not yet compulsory, can be left to the countries in function of their specific needs.

As regards classification for example, in the actual SNA the household sector is further sub sectored on the basis of the household's principal source of income. For some purposes it might also be appropriate to sub sector according to socio-economic or even to geographical criteria leading to a SAM (Social Accounting Matrix). It might also be useful to have separate data for public, national private and foreign controlled corporations. All of this additional information provides an analytical approach very suitable for researchers and decision makers.

1 *Paper for the 2nd IF4C-Conference, Basle, September 9–10 2004.*

3. Further increasing international comparability

Under the aegis of the international organisations the system of national accounts has substantially improved the international comparability. Filling out the system with data, remains however the main challenge for the national compilers of the accounts.

Data deficiencies for purposes of economic analysis are not necessarily due to deficiencies in the *system* of national accounts, but may result from lack of information to compile variables for all levels of detail of the classifications foreseen in the system.

Fruitful application of national accounting data would therefore be enhanced by improved availability and quality of basic data to compile the accounts, close adherence of statistical offices to the rules of the system, and more knowledge by economists of institutional differences between countries.

National accounts compilers can contribute in still another way to inform the user on the degree of international comparability, or in other words on the degree to which their national accounts do comply with the international agreed standards. Metadata should be presented in full detail as regarding the description of the data sources and the operational methods used in national accounts. Transparency in metadata is an essential element for the credibility of statisticians. This crucial task for statisticians in different domains has been stressed in the morning sessions and in the key notes addressed by Mr. Bill White.

At the level of international comparability we are currently in a very challenging period with the first revision of the SNA 1993. A whole machinery with think tanks is currently active in order to modernise the system.

The management of the updating process belongs to the responsibility of the Inter-Secretariat Working Group on National Accounts (ISWGNA). It co-ordinates the input from the Advisory Expert Group on National Accounts, the electronic discussion groups, the task forces, the working groups on national accounts at the level of the UN, OECD and EU, and the Canberra II group on the measurement of non-financial assets.

The Advisory Expert Group on National Accounts has defined operational criteria for the selection of those issues which might be candidate for the revision. These criteria are:

- there should not be fundamental or comprehensive changes to the 1993 SNA that would impede the process of its implementation, which in many countries has not yet been achieved;
- candidates for updating should be issues that are emerging in the new economic environment;
- candidates for updating should be issues that are widely demanded by users;
- any recommendation for change should have its internal consistency and consistency with related manuals;
- any recommendation for change should be tested for feasibility of implementation in countries with developed and less developed statistical systems.

At present, reflections, analyses and discussions are organized on the selected issues for revision at different levels of groups. For some issues proposals have already been forwarded to the ISWGNA. The aim is to finalize and release the SNA 1993 revision 1 by 2008. At the level of the European Union working groups on financial accounts and non-financial accounts are set up and the EU has already contributed in many respects to the revision via other channels.

Notwithstanding the different fora – including electronic discussion groups – and the widespread posting of documents on the internet, it seems that the final user finds it difficult to get involved in the whole revision process. Ignorance of the importance of these revisions can also be at the origin of the lack of involvement on the side of the final user. In the framework of a contribution to the IARIW-conference of August 2004, a survey among users (mainly economic analysts) has however been conducted. It can be observed that the awareness of the ongoing SNA-revision is still not very high within the central bank community, which is in fact a user-group with increasing demands for national accounts data.

The most important issues that have been selected for review of the system of non financial national accounts can be regrouped into categories according to main the scope of the revision and are illustrated with some examples.

a. Firstly, proposals for revision which provide more precise instructions on elements which were already part of the SNA 1993.

   Currently only expenditure on large databases generating services for at least one year is capitalized in the SNA. This principle will now be applied to expenditure on all databases, leading to an increase of the gross fixed capital formation.

   The "Build – Own – Operate – Transfer" schemes are quite new and the SNA could provide recommendations on how to treat these in the national accounts. These schemes become quite popular now in Europe, due to the fact that the European Commission is stimulating very much the Public Private Partnerships. Quite some analysis has been carried out by the European Commission on the institutional sector classification of these partnerships and is based on the criterion of who is bearing the – different types of – risks in the partnership.

b. Secondly, proposals for revision which are clearly related to changes in the classification of transactions.

Currently, expenditure in research and development is classified as intermediate consumption. There is however a strong tendency to consider these kind of expenditures as capital formation. Some test calculations have been carried out by a few counties and the current consultation procedure has recently been extended by the Canberra II group till December 2004. There remains still quite a lot of work in this area to ensure a consistent treatment in related economic activities and across different economic agents in relation with other proposals for revision.

Expenditure on weapons that can only be used for military purposes are currently considered as intermediate consumption. Here too there is support to consider all expenditure on military weapons as gross fixed capital formation.

The cost of ownership transfer has been discussed in detail by groups involved in the revision procedure. These costs will continue to be booked as capital formation and written off over the expected period of detention. The expected costs of ownership transfer in case of disposal of the asset is also to be written down over the expected period of detention, while the so called terminal costs (for instances cost related to the dismantling of nuclear plants) should be written off over the whole life of the asset.

c. And finally, a third group of proposals for revision which introduce new concepts into the national accounts framework or change the boundary of the system.

The SNA 1993 only recognizes “goodwill” when it is purchased by incorporated business. The question has been raised if the revised SNA should not also recognize the internally generated goodwill and apply the same rules for all types of business. There are up till now not yet clear indications on the possible outcome as much of the analysis is still underway.

A fairly new concept to the SNA is the proposal to make explicit estimates on the cost of capital services in the production account. The idea behind the proposal is to provide more suitable data for the analysis of productivity, as the concept is seen as the cost for using a unit of capital during a period, which does in fact form the cost of input of capital similar to the one on the input of labor. If accepted, this revision will imply new models for estimating the rate of return, or the profit rate, of a capital good, which in turn accentuates the need for very reliable data on the capital stock (with quite a lot of details of the composition on the capital stock). That proposal for revision is very reasonable, but might conflict with one of the criteria for revision, namely the one stating that new recommendations will only be done if they can be implemented by countries with developed and less developed statistical systems.

In summing up the issues for revisions, I hope to contribute in stimulating the central bank community to participate actively with national accounts compilers in the SNA 1993 revision and to take up their duties in consultations which should best be more oriented to the users.

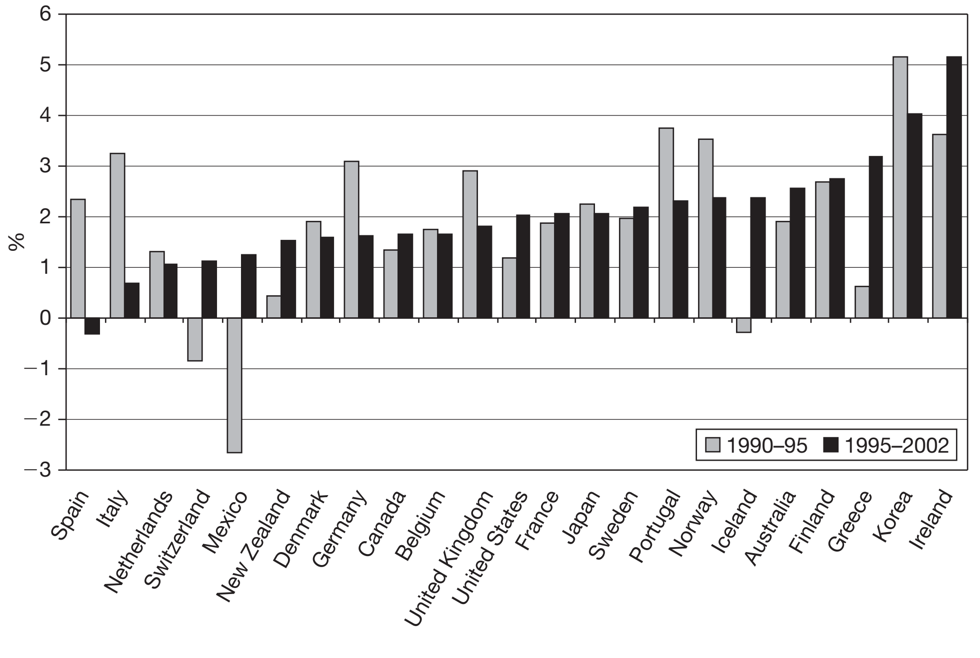
R. Acx (National Bank of Belgium)

Comparing growth in GDP and labour productivity: measurement issues

Nadim Ahmad, François Lequiller, Pascal Marianna, Dirk Pilat, Paul Schreyer and Anita Wölfl (OECD)

Growth and productivity are on the policy agenda in most OECD countries. Recent OECD work has highlighted large diversities in growth and productivity as well as a range of policies that could enhance them (OECD, 2001*a*, 2003*a*, 2003*b*). In the United States, Gross Domestic Product (GDP) growth has increased substantially faster than in large European countries or Japan, partly because the US population expanded rapidly during the 1990s. Moreover, estimates of labour productivity growth, measured as GDP per hour worked, suggest that US labour productivity has grown faster than that of some large European Union countries such as Italy and Germany (see Chart 1). Also, US GDP per capita has grown more in comparative terms, since strong labour productivity growth was combined with increased labour utilisation over the 1990s, in contrast with several European countries.

Chart 1 – Growth of GDP per hour worked; 1990–95 and 1995–2002
Annual average growth rate



Source: OECD estimates based on the OECD productivity database.

This Statistics Brief highlights measurement issues that can affect international comparisons of GDP and productivity growth and therefore the validity of cross-country analysis.[1] These measurement issues do not undermine the strong performance of the United States compared to the large EU countries and Japan for real GDP growth during the period 1995–2002. However, differences in annual average growth of GDP per capita and labour productivity

1 *For more detail on this subject, see Ahmad, et al., 2003 at http://www.oecd.org/findDocument/0,2350,en_2649 _33715_1_119684_1_1_1,00.html.*

between these countries for the same period are small enough to fall within the range of statistical uncertainty.

Measuring nominal GDP

Comparability of nominal GDP is significantly dependent on the use of a common conceptual framework. The current framework is the 1993 version of the international "System of National Accounts" (SNA), which nearly all OECD member countries now use as the basis of their national accounts. Despite this convergence, however, some differences still exist between countries regarding the degree to which the manual has been implemented.[2]

Military expenditures

The coverage of government investment in the US National Income and Product Accounts (NIPA) is wider than that recommended by the SNA, since it includes expenditures on military equipment that are not considered assets by the SNA. The other OECD countries strictly follow the SNA in this matter. As the amount of public investment affects the level of GDP, this results in a statistical difference in the GDP measurement. Fortunately, the impact on GDP *growth* tends to be relatively small. Over the past decade, this methodological difference has only reduced annual US GDP growth by 0.03% on average. Recent planned increases in US military expenditure may reverse this effect, however. Convergence on this issue is expected in the next edition of the SNA, in 2008. In the meantime, the OECD publishes data for the United States in its Annual National Accounts Database, which adjusts for this difference.

Financial intermediation services

Most banking services are not explicitly charged. Thus, in the SNA, the production of banks is estimated using the difference between interests received and paid, known as "Financial Intermediation Service Indirectly Measured" (FISIM). All OECD member countries estimate total FISIM. While it is relatively straightforward to estimate FISIM, the key problem is breaking it down between final consumers (households) and intermediate consumers (businesses). Only the former has a direct impact on GDP. In the United States, Canada and Australia, such a breakdown has been estimated in the national accounts for some time, in accordance with the SNA. In Europe and Japan, however, the implementation of a breakdown between final and intermediate consumers has been delayed until 2005.

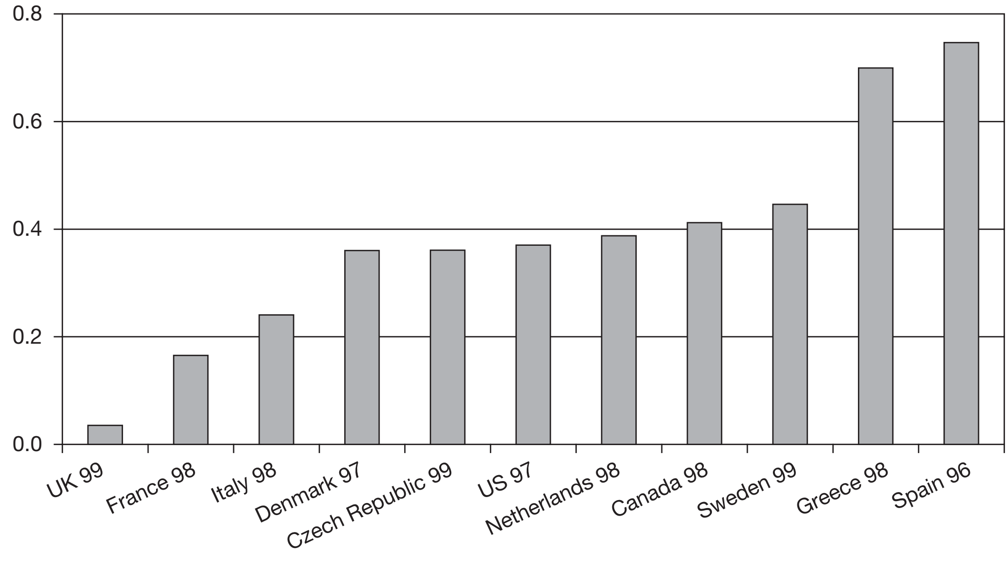
Before the NIPA was comprehensively revised in December 2003, imputed household consumption of financial services accounted for 2.3% of US GDP compared to zero in Europe and Japan. Fortunately, the impact on GDP *growth* is limited to less than 0.1% per year, the effect being positive in some years and negative in others. Furthermore, the recent revision of the US accounts has significantly reduced the difference in levels to probably just over 1% of GDP, thus roughly halving the impact on growth. Preliminary estimates suggest that with the implementation of the allocation of FISIM between both sectors in 2005, GDP levels would increase by approximately 1.3% in European countries and by nearly 3% in Japan. With these changes, diversities arising from methodological differences should be mostly eliminated.

Investment in software

Measurement of software investment is another significant issue in the comparability of GDP. The SNA recommends that software expenditures be treated as investment once the acquisition satisfies conventional asset requirements. When introduced, this change added nearly 2% to GDP for the United States, around 0.7% for Italy and France, 0.5% for the United Kingdom. Doubts on the comparability of these data were raised when comparing "investment ratios", which are defined as the share of total software expenditures that are recorded as investments. These ratios range from under 4% in the United Kingdom to over 70% in Spain (see Chart 2). One would have expected them to be roughly the same across OECD countries.

2 *An important issue not considered in this paper relates to valuing activities in the non-observed economy. This is the subject of Statistics Brief No. 5.*

Chart 2 – Investment ratios for purchased software



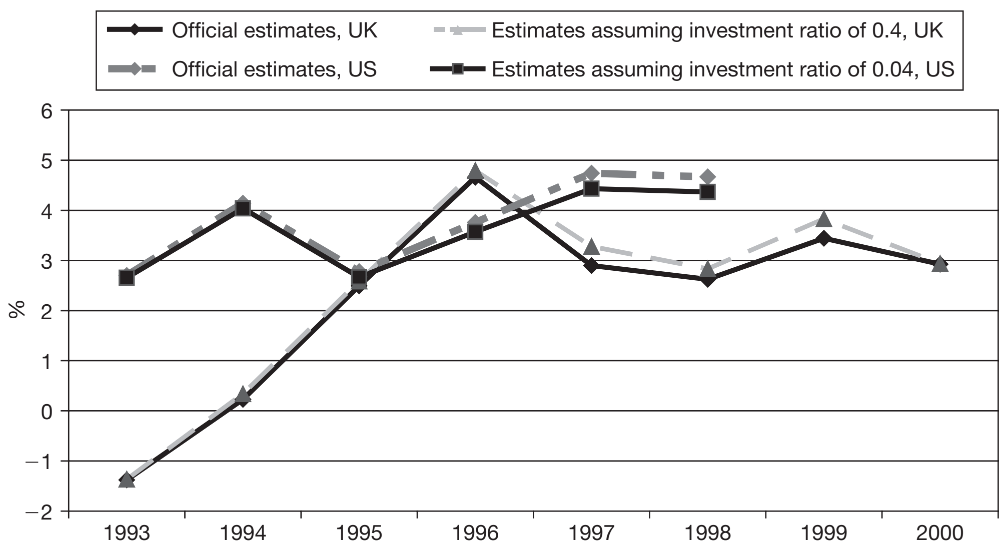
Source: OECD.

An OECD/Eurostat Task Force established in October 2001 confirmed that different estimation procedures contributed significantly to the differences in software capitalisation rates, and a set of recommendations describing a harmonised method for estimating software was formulated (see Lequiller, et al., 2003; Ahmad, 2003). Most of these recommendations eventually will be implemented by countries but until then differences in software and GDP measures will persist.

In practice, National Statistical Offices use one of two distinct methods to estimate software investment. The first derives data from business accounts. The second disregards business accounts, measures the total supply of computer services in an economy and estimates directly the amount of software with asset characteristics. The first approach tends to produce systematically lower estimates of investment than the second. This is mainly because businesses tend to use very prudent criteria when capitalising software, particularly if there are no tax incentives for doing so. Countries using the second approach, such as the United States, had a higher measured investment than countries such as France, the United Kingdom and Japan where statistical methods were more inspired by the first approach. As a result, while the amount of total software

Chart 3 – Sensitivity of GDP growth rates to different investment ratios for purchased software

Annual growth rate



Source: OECD.

expenditures may be more or less similar, the amount of software expenditures recorded as investment, *and thus included in GDP*, is significantly higher in the United States than in France, the United Kingdom or Japan for purely methodological reasons.

The impact of these methodological differences on GDP growth can be substantial. Chart 3 shows the estimated impact on GDP growth in the United States, assuming an investment ratio of 0.04 (which is the one currently used in the United Kingdom), and for the United Kingdom, assuming an investment ratio of 0.4 (which is close to that used in the United States and some other countries). In both cases, the OECD Task Force procedure for own-account production is applied and a number of assumptions are made.

The results show that the impact on UK GDP growth can reach 0.2%, and even 0.4%, in some years. Similar results are likely to occur depending on the size of the investment ratio in each country and the approach used. Changes of between ±0.25% of GDP should thus be expected. However, the variations in growth arising from the different methodologies is unlikely to be as large from 2000 onwards, since expenditure on software before then was exceptionally high to address the Y2K problem.

Measuring real GDP

Measurement becomes even more complicated when price and quality changes have to be accounted for. In this paper only temporal price indices are discussed, not spatial price indices (see Box "Purchasing power parities in comparisons of labour productivity growth").

Purchasing power parities in comparisons of labour productivity growth

Purchasing Power Parities (PPPs) for GDP are spatial price indices that compare levels of real GDP and its components internationally. As they are associated with *level* comparisons, there is normally no need to invoke PPPs for comparisons of *growth* in GDP and productivity. Thus, the method used by countries to compile the data discussed in the main text does not involve PPP. However, in principle, it is possible to construct an index of relative output growth by using a time series of PPPs, and applying them to one country's current-price GDP. The resulting GDP level, expressed in current international prices, can then be related to the GDP level of another country to form an index of relative GDP growth between two countries. This method is based on different weighting schemes and indeed, empirical differences can be sizable as recently shown by Callow (2003). One might argue that the comparison of the two methods should in itself be interesting because it reveals effects of different weighting schemes. This is true in a world of complete and high-quality statistics. Practically, however, PPPs are based on a smaller sample of prices and on less detailed weights than national price indices. For purposes of comparing relative output and productivity *growth*, the comparison based on constant national prices is preferred. PPPs should be used when output and productivity *levels* are the object of comparison across countries. For more detailed discussion of PPPs, see Statistics Brief No.3.

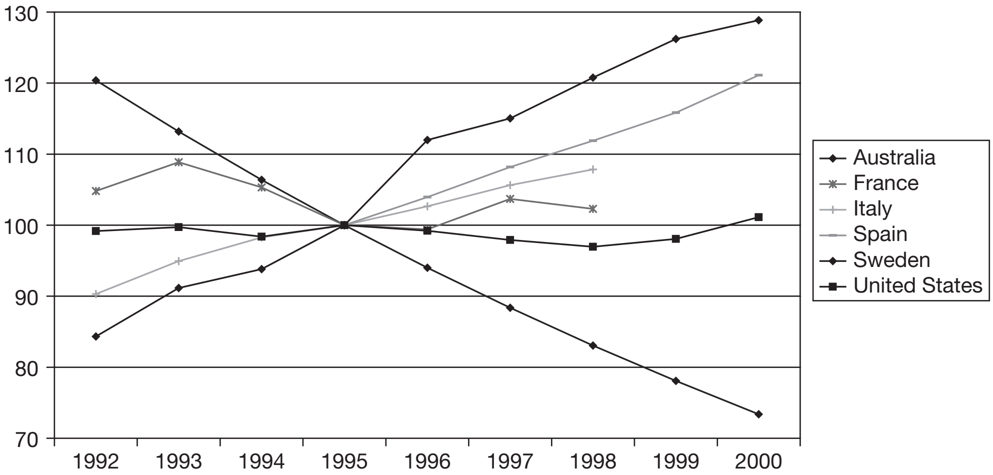
Adjusting for quality change: the role of hedonic price indexes

A widely discussed issue at the height of the "new economy" debate was the international comparability of rates of economic growth, given that the United States and some European countries apply very different statistical methodologies to the computation of price indices for information and communication technology (ICT) products. Because an alternative price index translates into a different measure of volume growth, the question has been posed whether some, or all, of the measured variation in growth between countries is a statistical illusion rather than reality.

The main challenge is to accurately account for quality changes in these high-technology goods, for example computers. The necessary quality adjustments are not standardised across countries. Consequently, between 1995 and 1999, the US price index of office accounting and photocopying equipment (which includes computers) dropped by more than 20% annually, compared with 13% in the United Kingdom and – at that time – a mere 7% in Germany.[3] Because computers are internationally traded, their price changes should be similar across countries.

3 *Germany has recently introduced hedonic methods for IT products. The first publication of these data took place in 2002.*

Chart 4 – Price indices for software investment
Indices, 1995 = 100



Source: OECD.

Another illustration of the difference in price indices is evident in Chart 4, which shows deflators for software investment. For example, the price index for Australia fell by about 30% between 1995 and 2000, while the price index for Sweden rose by about the same amount. This does not reflect real price movements but rather the dearth of price information available in this area.

Thus, at least part of the differences in measured price changes appears to be due to methodological differences. The natural question is: how would GDP growth in Germany, the United Kingdom or any other country change if US methods were applied? Clearly, if the US price index for computers is applied to Italy's or the United Kingdom's investment expenditure, their investment volume will show more rapid growth, as will the volume measures of the computer industry's output. However, the direct effect on GDP growth of different price indices is limited owing to three factors.

First, only final products have an impact on GDP. Thus, errors on the price index of an intermediate good such as semiconductors will only affect the contribution of the semiconductor industry to total GDP growth but not GDP growth itself. A second distinction is that, even for final demand products such as personal computers, the impact of an error on the price index will only affect GDP if the product is manufactured in the country. Third, if imported products are used as intermediate inputs then the absence of hedonic deflators in a country's national accounts will lead to an *overstatement* of real GDP growth (assuming that hedonic deflators represent the preferred measure) because imports will be lower and imports have a negative impact on GDP.

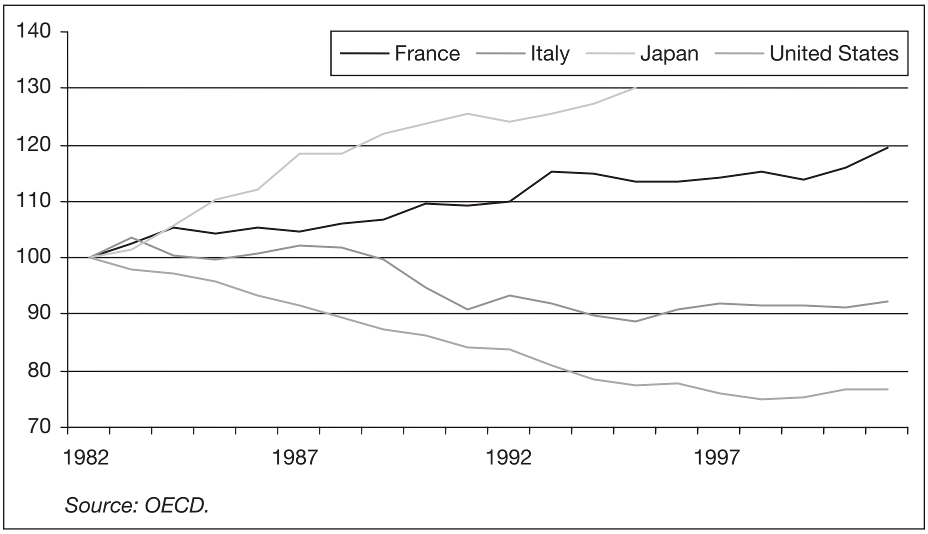
This is why simulations to obtain an order of magnitude for the impact of price adjustments of ICT products on the rate of change of real GDP lead, in general, to only modest effects, around +0.1% (Lequiller, 2001; Deutsche Bundesbank, 2001; Schreyer, 2002). A review of the impact of hedonic price indices on aggregate volume growth in the United States found that the quality change in personal computers added 0.25% to the estimate of annual real GDP growth over the period 1995–99 (Landefeld and Grimm, 2000). While quite high, this has to be put in proportion to a rate of real GDP growth of 4.15% per year in the United States during this period.

Measuring real output in services

The service sector now accounts for about 70 to 80% of aggregate production and employment in OECD economies and continues to grow. But measuring output and productivity growth in many services is not straightforward. The measurement of *non-market services* is even more difficult, because, by definition, there is no output price and thus no deflator, other than costs. Health and education are the main non-market services and they have a significant impact on GDP because they contribute to final demand. This difficulty may raise doubts regarding the international comparability of the volume estimate of production of these sectors. In this paper we have focused on the health and social services sector.

Currently, a vast majority of OECD countries measure volumes of health services as the sum of deflated costs. However, such input-based methods fail to reflect the quantity and quality of output, thereby mis-measuring productivity growth. Some countries have therefore tried to implement what are called output measures. Chart 5 shows that measured productivity in health and social services is in secular decline in the United States and Italy, while increasing slightly in France and steadily in Japan. Given the nature of the industry and given the difficulty in obtaining appropriate price and volume measures, it would seem that at least some of the cross-country differences in productivity growth are due to measurement differences. However, it is very difficult to quantify these effects.

Chart 5 – Real value added per employed person in the health and social work industry
Indices, 1982 = 100



Zero productivity simulations

Noting that certain service industries have been characterised by prolonged periods of negative measures of productivity growth, one could conclude that poor measurement may provide an explanation. Wölfl (2003) simulated what would happen if productivity growth in these services for France, the United States and Germany had not been negative but zero and further simulated the impacts on measured overall labour productivity growth. The empirical results suggest strong negative indirect effects on measured productivity growth for industries that use these services, partly outweighing the direct positive effect of the adjustment on productivity growth for the services themselves. The effect on aggregate productivity growth may be very small, depending on the type of measurement problem and the importance of the adjusted services for other industries and for the total economy. Setting negative productivity growth rates at zero, would lead, for instance, to a 0.19 percentage point change in measured aggregate productivity growth for France and a 0.08 percentage point change for the United States over the period 1990–2000. While these effects would directly translate into measured labour productivity growth, they are comparatively small.

The choice of index numbers

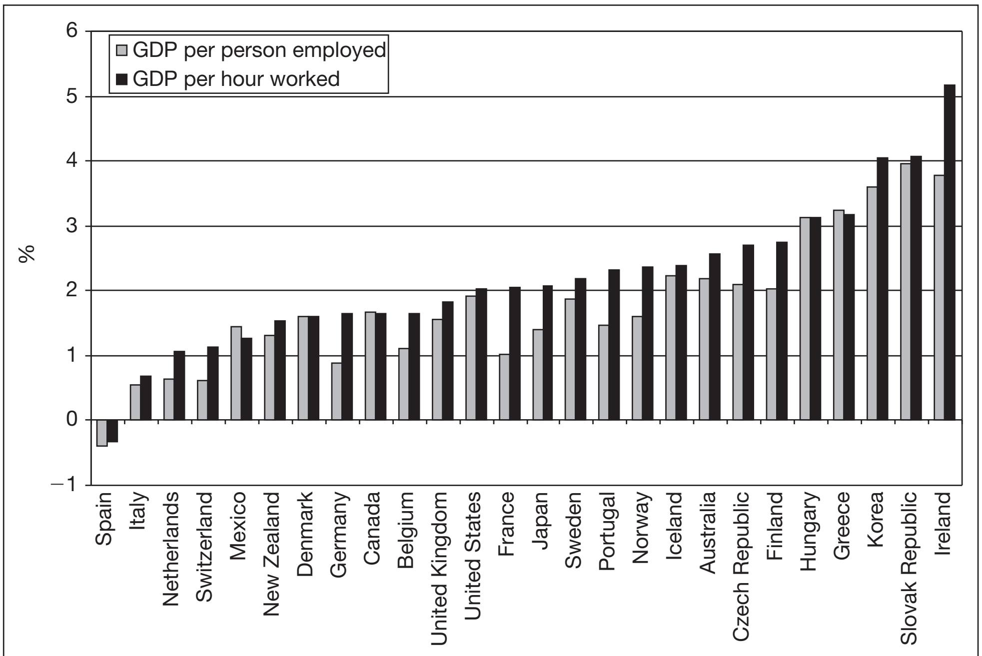
GDP growth is a single number, drawn up as a combination of the change in volumes of the several hundred goods and services categories that constitute the classifications of national accounts. To compile this number, countries use different formulae in practice, even if the trend is toward using chained rather than fixed formulae. "Fixed base" Laspeyres volume indices are currently still in use in some OECD countries (e.g. Japan, Germany, Finland, Spain, Sweden and Switzerland). Annually chained Laspeyres indices are the formula recommended by Eurostat for its Member countries. About half of the EU countries have changed their national accounts to a chained annually re-based Laspeyres method, and the remainder will change by 2005. The Australian and New Zealand accounts are also based on this index number formula. Annually weighted Fisher indices are currently used in the national accounts of the United States and Canada.

With regard to international comparability, does it matter which index number formula is chosen to compute volume GDP growth, given the same set of prices and quantities for GDP components in two countries? Schreyer (2001) used detailed final expenditure statistics to assess the effects of choosing fixed Laspeyres index numbers over chained Fisher index numbers, in the presence of significant relative price changes. The results confirmed that chained Fisher indices tend to produce systematically lower figures for GDP growth than fixed Laspeyres, ranging from −0.26% per year in Japan to −0.06% per year in Canada, the US being at −0.15%. In this case, the statistical methods implemented in the United States have thus decreased the estimate of GDP growth compared to other methods.

Measuring labour input

Labour input can be measured using total employment, total hours worked, or a quality adjusted measure of labour input. Considering the importance of the change in the number of hours worked, adjusting for this is particularly important for cross country comparisons of labour productivity. Chart 6 illustrates that in most OECD countries, while productivity growth over the 1995–2002 period was much more rapid after adjustment for hours worked, this difference is not similar for all countries. It is negligible for the United States, but important for Japan, France and Germany. Unfortunately, adjustments for the composition of labour are not available.

Chart 6 – Comparison of growth in GDP per hour worked and GDP per person employed, 1995–2002



Source: OECD.

Data on total hours worked (computed using surveys carried out on households or enterprises) are often not consistent with National Accounts. Some uncertainty remains regarding the comparability of measures of hours worked in OECD countries although this uncertainty is greater for the level of hours worked than for its growth rate. Nevertheless, comprehensive estimates of annual working time currently exist only for a limited number of OECD countries and the quality of data in this area therefore differs across Member countries, introducing some uncertainty in measures of labour productivity growth. Total hours worked can be derived by combining estimates of annual hours worked per person employed with average employment levels over the year that are in accordance with National Accounts production boundaries. Such employment series have not been collected systematically or examined very closely in the past.

The quality of the labour productivity measures has relied on the vigilance of the analysts and their understanding of the annual hours and employment series. Work is currently underway at the OECD to develop a series of estimates of total hours worked that ensure the consistency over time of employment and annual hours worked in calculations of labour productivity.

Conclusion and future work

This paper cannot be considered to fully cover all statistical differences in GDP and productivity growth between countries. In particular, the comparability of imputed rents has not been explored and some methodological differences remain probably yet to be discovered. However, the impression is that the known differences remain small when compared to total growth differentials. Therefore, the assessment of a substantially more rapid growth in the US cannot be undermined by statistical defects. However, growth differentials for GDP per capita or labour productivity between the US and other countries have been smaller on average during 1995–2002. The diagnosis remains therefore more fragile for these variables, in particular considering the difficulties regarding the measurement of labour input. The OECD is highly committed to working with Member countries towards maximising convergence on statistical methodologies and to provide a better statistical base for such analyses. In this regard the organisation is currently developing a reference database on productivity at the aggregate level.

Glossary

ICT: This stands for Information and Communication Technologies. It covers the range of new information goods and services, from software and computers to mobile phones, including semi-conductors.

Hedonic pricing: This refers to a technique which consists of using econometrics to price the different characteristics of a product, thus allowing price index compilers to take better account of the differences in the quality of the product. A good synonym would be "fully quality adjusted price indices". In general, hedonic pricing leads to price indices which grow slower (or decrease more) than non-hedonic pricing.

Further Information

Ahmad, N, F. Lequiller, P. Marianna, D. Pilat, P. Schreyer and A. Wölfl (2003), "Comparing labour productivity growth in the OECD area: the role of measurement", *Statistics Directorate Working Paper 2003/5*, OECD, Paris.

Ahmad, N. (2003), "Measuring Investment in Software", *STI Working Paper 2003–6*, OECD, Paris.

Callow, Julien (2003); "Getting Europe to work", *Credit Suisse First Boston Fixed Income Research*; 10 April.

Deutsche Bundesbank (2001), "Problems of international comparisons of growth – a supplementary analysis", Appendix to *Monthly Report*, May.

Landefeld, Steven J. and Bruce T. Grimm (2000), "A Note on the Impact of Hedonics and Computers on Real GDP", *Survey of Current Business*, December.

Lequiller, François (2001), "La nouvelle économie et la mesure de la croissance du PIB"; *Série des documents du travail de la Direction des Etudes et Synthèses Economiques G2001/01*; INSEE.

Lequiller, F., N. Ahmad, S. Varjonen, W. Cave and K.H. Ahn (2003), "Report of the OECD Task Force on Software Measurement in the National Accounts", *Statistics Directorate Working Paper 2003/1*, OECD, Paris.

OECD (2001*a*), *The New Economy: Beyond the Hype*, Paris.

OECD (2003*a*), *The Sources of Economic Growth in OECD Countries*, Paris.

OECD (2003*b*), *ICT and Economic Growth – Evidence from OECD countries, Industries and Firms*, Paris, forthcoming.

Schreyer, Paul (2001), "Information and Communication Technology and the Measurement of Real Output, Final Demand and Productivity", *Economics of Innovation and New Technology*, Vol. 10. 2001.

Schreyer , Paul (2002), "Computer price indices and international growth and productivity comparisons", *Review of Income and Wealth*.

Wölfl , Anita (2003), "Productivity Growth in Service Industries: an Assessment of Recent Patterns and the Role of Measurement", *STI Working Papers 2003–6*, OECD, Paris.

Previous issues of Statistics Brief

www.oecd.org/statistics
No.1, October 2001 – Trade in Goods and Services: Statistical Trends and Measurement Challenges
No.2, February 2002 – Creation of the Euro area: Implications for Economic Statistics
No.3, March 2002 – Purchasing power parities – measurement and uses
No.4, June 2002 – Measuring Student Knowledge and Skills: The PISA 2000 Experience
No.5, November 2002 – Measuring the Non-Observed Economy
No.6, October 2003 – Measuring Regional Economies

The Statistics Brief is published by the Statistics Directorate of the OECD
Editor in chief: Enrico Giovannini
Editor: Brian Finn
Editorial team: Benoit Arnaud, Eileen Capponi

For further information contact:
The Editor
STATISTICS BRIEF
Statistics Directorate
OECD
2, rue André-Pascal
75775 Paris France
email STD.STATBRIEF@oecd.org

Do you want to subscribe to *Statistics Brief* ? Register now at
www.oecd.org/oecddirect **and choose Statistics Brief or email/write to the Editor.**

Nadim Ahmad, François Lequiller, Pascal Marianna, Dirk Pilat, Paul Schreyer and Anita Wölfl
Statistics Directorate, OECD
francois.lequiller@oecd.org

Estimating the output gap in the Polish economy: the VECM approach[1]

Michał Gradzewicz and Marcin Kolasa (National Bank of Poland)

1. Introduction

The aim of the paper is to present the results of research on calculating the output gap in the Polish economy over the period 1996–2002. The analysis is based on selected methods.

The output gap, defined as a relative discrepancy between the actual GDP and the potential GDP (relative to the level of potential GDP), is an indicator of disequilibrium in the real economy. Thus, the concept of output gap is based on the definition of unobservable potential.

One can find various definitions of potential output in literature. From the monetary authority's point of view the definition proposed by Okun (1962) is still the most influential. He described potential GDP as the maximum quantity of output the economy can produce under the conditions of full employment (which is understood as the maximum level of employment not generating inflationary pressure[2]). Later refinements of the definition stressed alternative aspects of the above-mentioned qualification, e.g. the intensity of use of labour and capital (Artus (1977)). Output gap is stressed in the neo-keynesian models of the DSGE type (*Dynamic Stochastic General Equilibrium*) with sticky prices (cf. Clarida, Gali, Gertler (2000)). Here the concept of potential output is different – the equilibrium level reached without nominal rigidities, that is, with fully flexible prices and wages.

Due to the unobservable character of potential product, there is no fully accepted method of estimation. Even the selection of one definition of potential often implies few empirical methods of estimation of output gap. Methods based directly or indirectly on Okun's definition usually use the production function. On the other hand, when taking into consideration neo-keynesian theory, the potential product is achieved when there are no rigidities, which is true in the long run. Thus, the potential is assumed to be the long-run growth path of the product. Methodologies that decompose the product into permanent and transitory components make use of this interpretation of potential output, but without direct connections with neo-keynesian definition.

Apart from discrepancy of inflation from the inflation target, the output gap is a relevant component of the so called "Taylor rule" (Taylor (1993)), which is a premise of decisions made by monetary authorities. Moreover, the concept of output gap is used to decompose the budget deficit into cyclical and structural components. The lack of a commonly accepted estimation method of potential output implies a high uncertainty of macroeconomic policy. The uncertainty is also strengthened by errors connected with real-time estimations of economic aggregates or revisions of the business cycle phase and different interpretations of economic processes due to the inflow of new data (Orphanides, van Norden (2002)).

The authors decided to present the results of output gap calculations using two methods:

- a method based on a two-factor dynamic production function (estimated in the co-integrated VECM system, in which the potential GDP is calculated as the product resulting from maximum (in the Okun sense) level of production inputs (Chapter 2);

1 *The authors would like to thank R. Sawiński, W. Rogowski and Z. Żółkiewski for their assistance and all colleagues from DAMS that actively participated in making this paper. The views expressed in this paper are those of the authors and do not necessarily reflect those of the National Bank of Poland.*

2 *The equilibrium unemployment, not generating inflationary pressure, is called NAIRU (Non-Accelerating Inflation Rate of Unemployment). With different definitions of labor market equilibrium, one can also distinguish the NAWRU (Non-Accelerating Wage rate of Unemployment) and the natural rate of unemployment (Kwiatkowski (2002), p.154). These rates of unemployment are of course different, but taking into consideration their non-observable character, they are often exchanged in applications.*

- the GDP Permanent-Transitory Decomposition, using long-term restrictions in the vector error correction model (VECM) imposed in an endogenous way by co-integrating relationships; Chapter 3 contains the results for two models: one based on the long-term production function and the other on the permanent income hypothesis.

The selected methods are connected by making use of cointegration analysis and the system estimation of Johansen. Their common feature is also the fact, that they have origins in economics, in contrast to the mechanical methods of potential estimation (e.g. deterministic trend, Hodrick-Prescott filter, Band-Pass filter[3]). This is why the applied methods are said to be structural (Chagny, Döpke (2001)), although the approximation of some variables with filters means that they should be called rather semistructural approaches.

On the basis of Phillips curve, research on the influence of obtained output gaps on inflation in the Polish economy was conducted, though the selected definitions of the potential output do not necessarily have to lead to such a relation. Section 4 analyses time relationship between the estimated time-series of output gaps and various inflation measures (CPI, PPI, GDP deflator and one of the core inflation indexes). This analysis was made using cross-correlations of the inflation and the gap, implied by the covariance generating function of the VAR stochastic process.

2. Production function approach

Estimating the production function

In order to estimate the potential GDP of the Polish economy, the dynamic Cobb-Douglas function was selected as the production function. It's one of the simplest ways of describing the transformation process of inputs into output. Regardless of the functional form, on the assumption that production inputs are paid proportionally to their marginal products, a constant share of their income in GDP implies that elasticity of substitution equals one. It suggests the choice (within the family of functions with constant elasticity of substitution – CES) of a Cobb-Douglas production function. Stability of factor income shares in GDP seems to hold in the Polish economy – during 1995–2001 the share of labour compensation of employees in gross added value oscillated between 50% and 53%, and it was almost identical in 1995 and 2001. Stationarity of this series was confirmed by the ADF-GLS test (at 10% significance level) and the complementary KPSS test (at 5% significance level). The reliability of these tests is low due to the short sample used in the analysis.

Assuming Cobb-Douglas technology, constant returns to scale and Hicks neutral technological progress, the production function can be written as follows (cf. ŻółTowska (1997)):

$$Y_t = A(\alpha_t L_t)^{\alpha}(\beta_t K_t)^{1-\alpha},$$

where Y_t is the output (GDP), L_t and K_t are the inputs of labour and capital, while α_t and β_t are technical progress functions in the meaning of Harrod and Solow, respectively. As it is difficult to separate labour productivity growth from capital productivity growth, the production function is often presented as follows:

$$Y_t = TFP_t L_t^{\alpha} K_t^{1-\alpha},$$

where TFP_t is the total factor productivity and reflects technical progress increasing the productivity of both labour – by improving the qualifications of employees, and capital – with introduction of advanced technologies. Having this structure, the TFP_t variable makes it possible to introduce variations in the A factor, and thus take into account factors which cannot be explained by technical development. One of such factors is the effectiveness of social resistance to the introduction of new technologies (cf. Prescott (1997)).

The direct estimation of the production function using the OLS (in which the simultaneous-equation model is reduced to one-equation) does not seem to be a proper method for at least two reasons. Firstly, the output and inputs of labour and capital cannot be treated as independent, so the assumption that explanatory variables are exogenous does not hold (cf. Griliches, Mairess (1995)). Thus, the assumption of the exogenity of explanatory variables is violated. Secondly, according to economic theory, at least GDP and capital are generated by non-stationary stochastic

3 *One should also bear in mind that even pure mechanical methods imply some economic features of the filtered variables.*

processes (King, Plosser, Stock, Watson (1991)), so the use of OLS may lead to spurious regressions, which is the pure statistical relation between time series, without any economic meaning.

One of the methods allowing to avoid the above mentioned methodological errors is the multidimensional cointegration analysis. In this paper, the system was estimated in the form of a vector error correction model (VECM), according to the Johansen (1991) procedure. The cointegration relationship between production and inputs of labour and capital estimated in the above way can be considered as a well-estimated production function provided that the model has been correctly specified.

The quarterly data[4] covering the period 1995–2002 was used for empirical analysis. To eliminate the impact of seasonal factors on the results, all variables were seasonally adjusted using the multiplicative moving average method. The labour input (L_t) was assumed to be equal to the number of employed persons according to the Labour Force Survey (LFS). The variable describing the capital in the Polish economy (K_t) was assumed to be equal to the gross value of fixed assets in the national economy.[5] The data on real GDP (Y_t) is taken from publications of the Central Statistical Office.

The results of unit root testing based on the augmented Dickey-Fuller test with the generalised least square method (GLS) indicates that the logarithms of seasonally-adjusted Y_t, L_t and K_t variables are integrated of order one (Table 1). As the power of ADF-type tests is quite low, the complementary KPSS test was also conducted indicating the same integration order. Economic assumptions about the integration order of GDP and capital have been empirically confirmed. Thus, there are grounds for looking for cointegrating relationships between these variables.

Table 1 – Results of the unit root test

Variable	ADF-GLS statistics	KPSS statistics	Conclusion at the 0.1 significance level
$\Delta\log(Y_t)$	−1.99**	0.09**	I(1)
$\Delta\log(L_t)$	−2.00**	0.08**	I(1)
$\Delta\log(K_t)$	−1.91*	0.16*	I(1)

**significant at the 0.1 significance level.*
***significant at the 0.05 significance level.*
Source: Own calculations.

In the first stage of the estimation process it was assumed that the TFP_t variable can be approximated by the exponential trend, i.e. a linear trend after calculating the logarithm. The assumption was made in order to make use of standard Johansen (1991) procedure and calculate factor elasticises of the product.

Cointegration tests indicate that there is one cointegrating relationship between Y_t, L_t and K_t at the 1% significance level. This is so both in the case of the trace test and the test of maximum eigen-value (Table 2). Consequently, the VECM system was estimated with the assumption that there is one cointegrating relationship.

Table 2 – Johansen cointegration test

Hypothesis: number of cointegrating relationships	Eigen-value	Trace statistic	Maximum eigen-value statistic
None	0.74	53.07***	37.38***
No more than one	0.28	15.68	9.06

****hypothesis rejected at a 0.01 level of significance.*
Source: Own calculations.

4 *Cointegration analysis is generally combined with a long-run approach, based on annual data. Due to a short sample problem (the analysis covers 8 years), interpretation of obtained relations as long-run relations should be cautious.*
5 *According to calculations done by Marcin Kolasa and Roman Sawiński, DAMS, NBP.*

The estimation yields the following assessment of the long-term relationship between GDP, labour input and capital input:

$$\hat{Y}_t = L_t^{0.57} K_t^{0.43} e^{0.29+0.01t},$$

The estimated long-term relationship parameters have the expected signs and are statistically significant. The same can be said of the α adjustment matrix (for one cointegrating relationship it has the form of a vector), corresponding to the equations describing labour and capital dynamics. The adjustment parameters indicate that the equilibrium state is regained quicker through labour input (half of the adjustments lasted no more than 5 quarters) than through capital input (half of the adjustments after 5 years). This seems intuitively right and is consistent with observations of how the economy works.

The adjustment parameter in the equation describing GDP dynamics is statistically not different from zero. This characteristic has important implications for the entire system and is related to the idea of weak exogeneity (cf. S. Johansen (1991)). The weak exogenous character of GDP with respect to the cointegrating relationship can be interpreted by saying that adjustments stem only from production factors, and not from the production itself.[6]

For the estimation, two restrictions were imposed on the VECM system. The first has a normalising nature, the second is connected with the assumption of constant economies of scale. The quotient test of the likelihood ratio imposed on the restricted system indicates that there is no reason for rejecting the hypothesis of their legitimacy (with the probability of 0.52). The system describes the dynamics of endogenous variables well, and this is proven by R^2 (equal 0.42 for the GDP equation, 0.40 for the employment and 0.95 for the capital equation) and the tested lack of autocorrelation of the random component. The roots of the characteristic polynomial corresponding to the system indicate that it is stable.

The last stage in estimating the production function is to replace the exponential deterministic trend approximating the total factor productivity with the TFP_t^* estimation. For this purpose, after estimating the production function using the exponential trend, residual values of the production function without the trend were calculated and smoothed using the Hodrick-Prescott filter. At the end, the process of estimating the production function was repeated in an analogous VECM system, in which the dynamic structure and restrictions of the original system were maintained, but the exponential trend was replaced with the estimated TFP_t^*. The calculated elasticities of the product with respect to the input of production factors differ from those previously calculated only at the third decimal place. Consequently, the final estimation of the production function can be described by the following formula:

$$\hat{Y}_t = TFP_t^* L_t^{0.55} K_t^{0.45},$$

where TFP_t^* is the series of residuals after smoothing them with the HP40 filter.

The estimated production function was then used to calculate the potential output.

Potential labour input

According to the methods used by OECD (cf. C. Giorno et al. (1995)), the potential labour input used to calculate the potential output is obtained from the following formula:

$$L^* = LF(1 - NAWRU),$$

where *LF* is the labour force, while *NAWRU* is the non-accelerating wage rate of unemployment. Hence L^* is the number of employed persons at the natural unemployment rate, consistent with the concept of the production function.

Labour force was assumed to be the number of professionally active persons according to LFS, but it is more difficult to estimate the *NAWRU*. The starting point was the method proposed by Elmeskov (1993), in which the change in the rate of the wage inflation is proportional to the difference between the actual unemployment rate and *NAWRU*, which can be expressed

6 *It could suggest the demand characted of observed GDP, but it's not the case when one takes into consideration the fact, that innovations in the VECM system are the combination of short- and long-run disturbances.*

as follows

$$\Delta^2 \log W = -a(U - NAWRU), a > 0$$

where *W* – wage level, *U* – unemployment rate. The underlying methodology is thus consistent with the Phillips curve supplemented with adaptative expectations, according to which the expected wage inflation in the current period is equal to the rise in wages in the previous period (cf. Staiger, Stock, Watson (1996)). Assuming that the *NAWRU* is constant between any two consecutive quarters[7] allows one to calculate the *a* parameter for subsequent periods and, as a result, to calculate the *NAWRU* series.

The method proposed by Elmeskov is used for annual data, so it seems right to modify it for calculations based on quarterly data. In this study, the modification consists in replacing the current value of the *U* variable with a distribution of its lags, i.e.

$$\Delta^2 \log W = -a(\varphi(L)U - NAWRU), a > 0$$

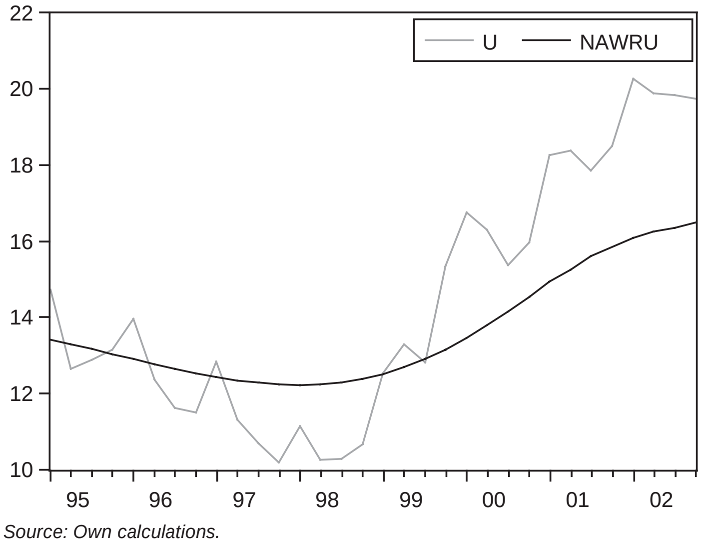
where $\varphi(L)$ is a lag polynomial of order four, while *L* is the lag operator. A hypergeometrical distribution of $\varphi(L)$ polynomial coefficients was assumed to take into account the delay with which the labour market situation affects rises in wages. The resulting formula is:

$$NAWRU = \varphi(L)U - \frac{\Delta\varphi(L)U - \Delta^3 \log W}{\Delta^2 \log W}.$$

The drawback of the Elmeskov approach is that the short-term *NAWRU* generated by this method varies significantly with unemployment changes, which can be explained by the impact of not only the unemployment level, but also its changes, on wage inflation. This problem is usually reduced by smoothing the series using the Hodrick-Prescott filter (cf. Giorno et al. (1995)).

The use of the HP filter gives rise to the generally known doubts connected with the choice of a right smoothing parameter and the end of sample bias. As there were no other premises, the smoothing parameter of 1600 was adopted, which is the same value as that chosen by Hodrick and Prescott (1980) and also a standard parameter used for quarterly data.[8] The problem of the beginning of the series can be easily eliminated by adding data from 1992–1994 to the *NAWRU* calculation. The end of the series bias is far more problematic. Thus the most recent elements of sample have been adjusted so that the average *NAWRU* level in 2002 is approximately 16%.[9]

Chart 1 – NAWRU in Poland, 1995–2002



Source: Own calculations.

7 *The assumption that NAWRU is constant between two consecutive quarters seems to be only technical. However, in many interpretations NAWRU is dependent on institutional and structural factors (por. Layard, Nickel, Jackman (1991)), which in the short run can be assumed constant. Hence, the assumption made is to some extent theory-driven.*

8 *Discussion of the effect of using smoothing parameters in the HP filter in: Canova (1993).*

9 *In comparison, the NAIRU rate for Poland estimated using the SVAR method was 15% in mid-2002 and grew by about 1 percent annually. Cf. BRE Bank S.A. (2002).*

Since the *NAWRU* value for the end of the sample is assumed arbitrarily, the estimate of the potential GDP and the output gap should also include a sensitivity analysis taking account of a possibly wrong estimate of the equilibrium unemployment. This analysis is presented later on in the text. The estimate obtained on the basis of the actual unemployment level according to BAEL is presented in chart 1.

Estimating the potential production and the output gap

According to the methodology developed by OECD (cf. Giorno et al. (1995)), the following formula is used to calculate the potential GDP:

$$Y_t^* = f(K_t, L_t^*, TFP_t^*),$$

where Y_t^* is the potential GDP and *f* is the estimated production function.

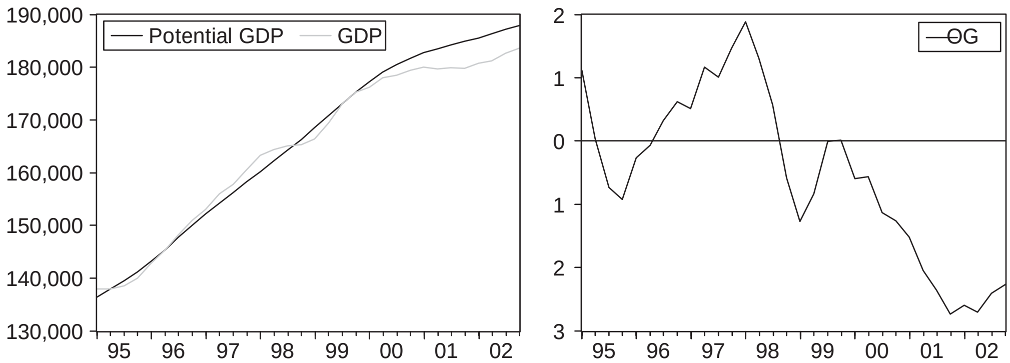
If the estimated Cobb-Douglas production function is used, the potential GDP for the Polish economy can be calculated using the formula below

$$Y_t^* = TFP_t^* L_t^{*0.57} K_t^{0.43}.$$

The estimated potential output can be used to calculate the contribution of particular factors to potential GDP growth. For a description, calculations and results see appendix A.

After the potential GDP has been estimated, it is possible to calculate the output gap as the difference between the actual and potential GDP level. The results for Poland are presented in chart 2.

Chart 2 – GDP, potential GDP and the output gap as a percentage of the potential GDP in Poland (seasonally adjusted data)



Source: Own calculations.

The results of output gap estimation using a method based on the production function show that this gap was positive until the third quarter of 1998, and then it fell below 1%. In the second half of 1999, GDP was almost equal to its potential level. As from the beginning of 2000, the output gap was negative and kept worsening until the end of 2001. Since 2002, one can observe a gradual decrease of output gap.

It has already been mentioned that the estimation of the potential GDP and the output gap depends on the estimation of the NAWRU unemployment rate. The elasticity of GDP in relation to labour input indicates that the underestimation (overestimation) of NAWRU by one percentage point decreases (increases) the potential GDP by 0.6–0.7%. As a result, the output gap decreases (increases) by 0.6–0.7 of a percentage point.

3. Permanent-transitory decomposition of the GDP

An alternative approach to determining the potential output, used both in economic theory and empirical research, is to treat it as the long-term GDP trend. The disputed issue is the method of determining the trend, as a theoretical concept, on the basis of a time series of a relatively high frequency.

Econometric research based on the stochastic description of economic phenomena concentrates on the permanent-transitory (PT) decomposition of the product time series. In this approach, the potential product is treated as the permanent part of the output constructed by eliminating the effect of transitory disturbances.

The starting point of the analysis is a VAR-type dynamic econometric model composed of variables integrated of order one. There should be long-term relationships between the variables making up this system, consistent with the selected economic theory.[10] This condition imposes recursive cross-restrictions on the parameters on the moving average representation of the original model, leading to the representation of the vector error correction (VECM). The next step in the analysis is the PT decomposition of disturbances affecting the system. The decomposition itself and identification of both types of disturbances is made endogenously in the system, using the long-run restrictions imposed not implicitly on the VAR system by VECM. The last element of the identification process is the assumption that short- and long-run factors are independent (uncorrelated). The number and type of both sorts of disturbances is implied by the number of cointegrating relationships in the system (permanent disturbances are called common trends). The mathematical details of the described decomposition, developed by Yang (1998), are presented in appendix B.

PT decomposition methods are often based on strong assumptions. Multivariate Beveridge-Nelson decompositions (Evans, Reichlin (1994)) or unobservable factor methods (usually solved using the Kalman filter) assume that the trend of product is a random walk process and ignore the adjustments of potential after the occurrence of a permanent shock. The potential is then an imagined level of the product achieved after all transitory adjustments have died out. However, research, using Real Business Cycle models, indicates that the transitory dynamics of permanent productivity shocks also influence potential output (Lippi, Reichlin (1994)). The method of trend determination used in this study takes into account not only the long-term impact of permanent disturbances, but also the accompanying transitory adjustments. Compared to other methods based on cointegrating relationships (incl. Cochrane (1994), Duspaquier, Guay, St-Amant (1999)), which take transitory adjustments into account, this methodology developed by Yang properly identifies the number of permanent disturbances affecting the system (which is equal to the number of common stochastic trends governing the behaviour of the whole system). In addition, this procedure does not require multiple estimations (as do the above mentioned methods), thus its relative efficiency.

The basis for the decomposition is the estimation of the VECM, in which the permanent relationship is advocated by a certain economic theory. A decision was made to estimate two systems:

- The PT-PF model, based on the long-term production function hypothesis;
- The PT-PIH model, based on the permanent income hypothesis.

The PT-PF model

The analysed dynamic system is composed of 3 variables: real GDP, the number of employed persons (data from LFS survey) and capital.[11] All variables are measured quarterly and were adjusted seasonally using the TRAMO/SEATS[12] method. The sample range is 1995–2002.

The variables making up the system are integrated of order 1 (cf. table 1) and the economic theory defines the cointegration relationship in this system as the Cobb-Douglas production function (discussed in more detail in Chapter 2). This justifies estimating the system as a vector error correction mechanism (VECM).

A restriction of constant returns to scale was imposed on the cointegration relationship parameters (validated successfully by the likelihood ratio test at the probability level of around 0.82). The estimated GDP elasticity of labour input is 0.493 (with an error of 0.09), and

10 *In the econometric sense, the existence of long-term relationships means that there is a stationary linear combination of I(1) variables. A relationship like this is called a cointegrating relationship.*

11 *Calculations made by Marcin Kolasa and Roman Sawinski, DAMS, NBP.*

12 *The choice of different seasonal adjustment was an effect of the desire to get more smooth data without random deviations, which is important when using methods making use of long-run properties of time series. On the other hand, with the PF approach, the emphasis was on leaving the annual dynamics of the data intact. The different character of employed methodologies is thus the reason for estimating two production functions using different data sets.*

product elasticity of capital input is 0.507 (with an error of 0.09). Johansen tests of rank of cointegration (trace and the maximum eigenvalue) confirm, just as the economic theory suggests, that there is one cointegrating vector for the described system (with a slope and trend in the cointegration relationship) – see table 3. The existence of one cointegration relationship was also confirmed by eigenvalues of the VAR(4) system corresponding to the analysed VECM system: two of them reached values close to 1, while the remaining eigenvalues were definitely lower, implying the existence of two common stochastic trends, which, in a three-dimensional system, means one cointegration relationship. The precise form of the long-run relationship is as follows:

$$\hat{Y}_t = L_t^{0.493} K_t^{0.507} e^{-0.055+0.009t}$$

Values of adjustment coefficients to the previous period's disequilibria (the error correction term) for labour and capital inputs have the appropriate signs, implying convergence to the long-run path, and the following respective values: −0.20 (meaning that 50% of adjustments are done after 6 quarters) and −0.05 (one half of adjustments take place in 6 years). The speed of adjustments of labour is greater than that of capital, which is consistent with the theory. Estimation results imply that the parameter measuring the strength of output adjustments to the long-run level is statistically insignificant, which is referred to as the weak exogenity of the output with respect to the cointegration vector, so the results are analogous to that obtained in the Chapter 2.

Table 3 – Johansen cointegration test for the PT-PF model

Hypothesis: number of cointegrating relationships	Eigenvalue	Trace statistic	Maximum eigen-value statistic
None	0.73	51.58***	37.06***
No more than one	0.33	14.52	11.36

****hypothesis rejected at a 0.01 significance level.*
Source: Own calculations.

Because of the quarterly frequency of data used for estimation, the short-term dynamic is reflected by rates of growth to the 3rd order (this corresponds to a four lag VAR in levels). A zero restriction on lags of the 2nd order was imposed on the short-term dynamics (this was successfully validated by the Wald significance test of the regressor group). This restriction was imposed on the short-term dynamics to gain degrees of freedom while assuring the description of the seasonal pattern. All free (not implied by the restrictions of common trends) roots of the characteristic polynomial are located within a unit circle, which implies system stability.

Chart 3 – Output gap in the PT-PF model as a percentage of the potential

Source: own calculations.

Lagrange multiplier tests of the auto-correlation indicate that there are no reasons for rejecting the hypothesis of the zero autocorrelation of disturbances up to the 12^{th} order.

The course of the output gap in 1996–2002 is shown on chart 3. In accordance with the assumptions for the decomposition, the output gap is stationary in the econometric sense. Both the ADF non-stationary test and the complementary KPSS stationary test successfully validated the appropriate hypothesis at the 0.05 significance level.

In 1996–1998 the gap is positive and reaches a peak equal to 1.15% of the potential GDP, later the gap became negative, with a local minimum (−1.8%) in the third quarter of 2000. In 2001, the tendency of the economy to regain equilibrium was interrupted again, with another minimum at −0.8%. However, recent observations again show signs of the economy returning to equilibrium.

The comparison of systems based on production function (PF and PT-PF models)

Despite the different method of filtering data from seasonal factors, both estimated VECM systems, based on production function (PF and PT-PF models) have very similar properties. Statistical tests confirmed the equality of fundamental parameters of both models (see Table 4).

Table 4 – Equality tests of models PF and PT-PF

Parameter	Trace statistics	Probability
Elasticity (α)	0.423	0.679
TFP dynamice	0.772	0.455
Employment equation ECT coefficient	0.001	0.999
Capital equation ECT coefficient	0.471	0.646

Source: own calculations.

Elasticity of production with respect to the labour input was 0.55 in the PF model and 0.49 in the PT-PF model. More importantly, obtained estimates are not significantly different from the share of labour costs in the gross value added (ranging from 0.5 and 0.53), used often as an approximation of this elasticity. Quarterly growth rates of TFP (assuming constant rates in order to make use of Johansen procedure), equal in both cases 1% and 0.9%, are not also statistically different. The same results arise from the analysis of the speed of equilibrium correction mechanism. In both systems product proved to be weakly exogenous.

In spite of the similarity of both systems the resulting output gaps (see chart 2 and 3) are different, especially in the period 2000–2002. This is of course the result of different methods and assumptions applied to calculate the potential. The full comparison of output gaps should be extended with another dimension – time relations with inflation, which will be shown in Chapter 4.

The PT-PIH model

The authors decided to make the Yang decomposition based on a model with different economic assumptions. The construction of this model and the character of cointegration restrictions are in line with the consumption theory, particularly with Permanent Income Hypothesis (PIH), though the model is not an empirical verification of this hypothesis (especially in its strong version developed by Hall (1978)).

The system consists of three variables, measured quarterly (the sample range is 1995–2002) and seasonally adjusted using the TRAMO/SEATS method: the real GDP, consumption and the short-term real interest rate. These variables are integrated of order one (cf. table 5). For GDP and consumption, the appearance of the unit root in time series is economically justified (cf. King, Plosser, Stock, Watson (1991)), but there is no clear agreement among economists as to the real interest rate.

In this system, the long-term relationship is defined as a stationary consumption/GDP ratio, independent on the interest rate. Imposing a restriction on the system, that eliminates the

Table 5 – Results of the unit root test

Variable	DF-GLS statistic	KPSS statistics	Conclusion at 0.1 significance level
$\Delta\log(Y_t)$	−1.99**	0.09**	I(1)
$\Delta(r_t)$	−2.58**	0.25*	I(1)
$\Delta\log(C_t)$	−3.59** (trend)	0.12**	I(1)

***significant at a 0.05 significance level.*
Source: Own calculations.

interest rate from the long-run relationship (successfully validated by the Wald test at the 0.4 probability level), led to the estimated elasticity of consumption in relation to the product equal to 1.0076. This confirms the hypothesis of the stationary proportion of consumption to the GDP.

The results of the Johansen cointegration rank test (see Table 6) indicate that there is one cointegration relationship at the 0.08 significance level. The existence of one long-term relationship is proven by an analysis of the eigenvalues of a corresponding VAR. Two eigenvalues are close to one (equal in module to 0.96), while the other are lower, which implies that there are two common stochastic trends in the system, and hence one cointegrating relationship.

Table 6 – The Johansen cointegration test for the PT-PIH model

Hypothesis: number of cointegrating relationships	Eigenvalue	Trace statistic	Maximum eigenvalue statistic
None	0.43	22.38*	15.802*
No more than one	0.207	6.58	6.51

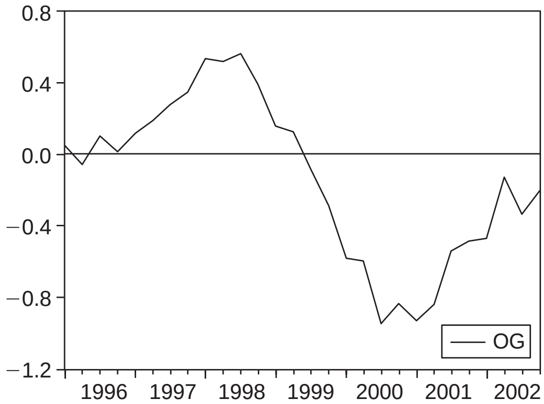
**hypothesis rejected at a 0.1 significance level.*
Source: Own calculations.

The analysis of coefficients of the model at the error correction mechanism leads to the following conclusions:

- consumption adjusts very quickly to the equilibrium level (coefficient equal to −0.35) – more than half of the adjustments are finished after 2 quarters and 75% of the adjustments take place after 3 quarters;
- as in the previous system, the product is weakly exogenous (statistically insignificant in terms of adjustment coefficient);
- there are also strong movements in the interest rate, which does not directly participate in the error correction mechanism, but its effect is visible in the short-term dynamics of the system.

The short-term dynamics of the system were approximated by rates of growth of variables up to the 3rd order, which is consistent with the quarterly frequency of the data. On the basis of tests of the Lagrange multiplier of the joint residual auto-correlation, one can say that the adopted scheme of transitional dynamics correctly approximated the dynamics of this phenomenon, eliminating auto-correlation from disturbances (up to the 12th order). All free eigenvalues of the system are smaller than one, so the described system is stable (the variables in the system return to sustainable growth paths). The R^2 calculated for equations are as follows: 0.85 for the consumption equation, 0,58 for interest rate and 0.77 for GDP. Taking into account the fact, that the model describes dynamics of variables, one can say, that the fit is reasonable.

Chart 4 – PT-PIH model output gap in percentages of the potential



Source: own calculations.

The output gap generated by the PT-PIH model (see chart 4) is stationary (both the ADF and the KPSS tests at the 0.05 significance level successfully validate the hypothesis of the stationary nature of the gap). Its course in time is similar to the PF model (cf. chart 3), but the scale of imbalance is smaller (the maximum deviation of about 1% of the potential GDP). It is worth noting that the moment when the gap changes the sign is in the second quarter of 1999, it reaches minimum in 2000 and the break in the growth trend in 2002. Recent observations confirm that economy is heading towards an equilibrium path.

4. Links between the calculated output gaps and inflation

One of the applications of the output gap is the evaluation of inflationary pressure in the economy. Modelling the relationship between inflation and output gap can be based on the Phillips curve. This relation exists provided firms set prices in the Calvo (1983) style in the monopolistic competition environment, which has not been tested for the Polish economy. The following analysis assumes such a dependency exists. Moreover, definitions of the potential assumed in calculating output gaps are not theoretically the same as in the neo-keynesian Phillips curve (which is the product achieved without ridigities). The authors assumed that despite lack of strict theoretical consistency, the obtained output gap measures may influence inflation. Due to these remarks, one should treat the results carefully.

The impact of the output gap on inflation can be measured using several tools. One of them is direct estimation of the Phillips curve. However, it imposes *a priori* restrictions on the distribution of lags in the equation (in order to approximate the *forward-looking* part of the model). One way to avoid this problem is to calculate the cross-correlations between the current inflation and the lagged gap, ordered according to the increasing time-interval between the two series.

In the case of the short time series, an important limitation on using cross-correlations to assess the strength and time lag of the gap effect on inflation is the significant decrease in the precision of estimators as the delay grows. This problem can be solved by constructing a two-dimensional stochastic process generating output gap and inflation data, and then determining its covariance-generating function (cf. Hamilton (1994), pp. 261–268). Normalisation of the elements of this function generates theoretical correlation relationships in time between the analysed time series, implied by the features of the stochastic process describing the joint development of these series.

In other words, one assumes that there exists a bivariate stochastic process, describing the dynamics of output gap and inflation (VAR type). In case of many stochastic processes, also in case of VAR process, one can analytically obtain moments of the process. The moments of the 2nd order (off-diagonal elements of covariance matrix) are thus the theoretical counterparts of empirical cross-correlations

$$\Gamma_i = E[(\mathbf{y}_t - E(\mathbf{y}_t))(\mathbf{y}_{t-i} - E(\mathbf{y}_{t-i}))'],$$

where E is an expectations operator and $\mathbf{y}_t$ is a vector containing output gap and inflation.

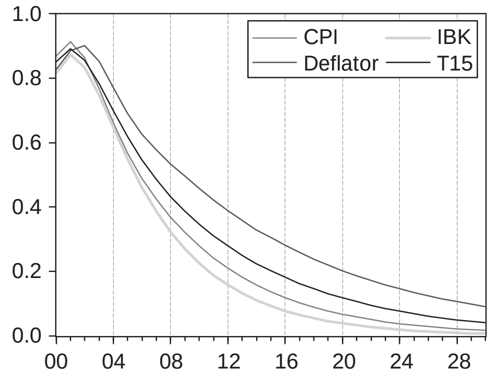
Charts 5, 6 and 7 (the horizontal axis shows the quarterly lag of the gap in comparison to the specific inflation rate) present the theoretical coefficients of correlation between the estimated output gaps and different price change measures, obtained by assuming a VAR-type data generation process. All estimated VAR systems are stable, disturbances are uncorrelated (up to 8th order) and describe the variables with at least 90% fit.

The following four measures of annual inflation were used in the analysis:[13]

- consumer price index (CPI),
- GDP deflator,
- core inflation (CPI excluded regulated prices - IBK);
- T15 – truncated distribution of prices according to volatility (15% of the most and the least volatile prices).

Theoretical correlations connected with PF model show strong connections with analysed inflation measures (see chart 5). The maximum of correlation is in the first quarter in case of CPI, IBK and T15 and at half year in case of the GDP deflator. A strong correlation (above 0.5) persists within a 1.5 year period, but in case of the GDP deflator it lasts about 2 years. This indicates a high persistence of demand factors generating inflationary pressure in the economy.

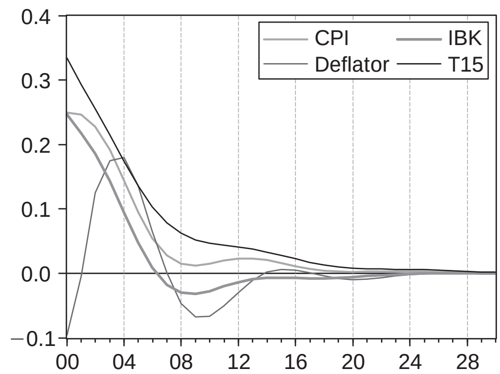
Chart 5 – Correlations – the PF model



Source: own calculations.

Interactions between the PT-PF output gap and inflation are weaker. The time structure of relations is also different (see chart 6). In case of inflation measured by T15 and core inflation IBK the maximum of correlation is in the current quarter and then gradually vanishes. The maximum influence on CPI is reached after one quarter and then quickly disappears. The time structure of the relationship between PT-PF output gap and the GDP deflator is different but the obtained correlations are quite small and probably statistically insignificant. It is worth noting that correlations with other measures are also quite low and inference based on this model could be inappropriate.

Chart 6 – Correlations – the PT-PF model

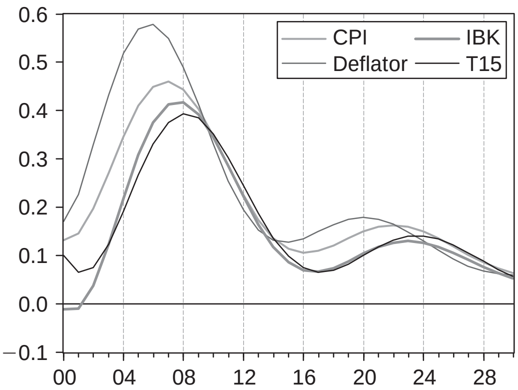


Source: own calculations.

13 *One should mention that it is hard to establish empirically the measure of inflation consistent with the Phillips curve (in particular, some volatility of price indices is due to changes in relative process, which the Phillips curve does not describe).*

In case of the PT-PIH model (see chart 7) the strength of correlation is moderate, but the time structure of relation is different. The strongest and fastest relation occurs with the GDP deflator. In all cases the maxima of correlation occurs after 5–8 quarters. The time path of correlations with the other three measures of inflation is similar, with the maximum of about 0.4.

Chart 7 – Correlations – the PT-PIH model



Source: own calculations.

The obtained lag structure in case of output gap calculated using PT-PIH decomposition is not in line with EU countries, in which the lag is not greater than one year (see Coenen, Wieland (2000)). The reason (mentioned before) could be that the definition of the potential is not in line with the potential used in Phillips curve. But when one assumes that such a great lag between output gap and inflation is correct, the changes in output gap resulting from the PT-PIH model influences inflation after over a year.

5. Summary

The alternative methods of estimating the output gap presented above differ with regard to the concept and method of calculation. The potential GDP estimated using the production function approach can be viewed as the supply-side of the economy, i.e. the GDP level corresponding to long-term inputs of production factors. On the other hand, methods based on the permanent-transitory decomposition of GDP use long-term relationships between macroeconomic aggregates and yield potential GDP which is a product of accumulated shocks.

Consequently, it is not surprising that these alternative methods lead to different estimates of the output gap. Moreover, the series obtained are not theoretically equivalent, hence the differences in their time relationships with inflation. The development of the gaps and the analysis of their impact on inflation (conditional on the existence of the relationship described by the Phillips curve and the possibility of using the gaps obtained in it) show the lack of any inflationary pressure from the demand side, which may be the case till the end of 2003. In view of relatively strong assumptions made during the estimation process and time relationships analysis, caution is recommended while drawing any conclusions.

Bibliography

Almon, S., (1965), *The Distributed Lag Between Capital Appropriations and Expenditures*, Econometrica 33

Artus, P., (1977), *Measures of potential output in manufacturing for eight industrial countries, 1955–78*, IMF Staff Papers No. 24

Beveridge, S., Nelson, C.R., (1981), *A new approach to Decomposition of Economic Time Series into Permanent and Transitory Component with Particular Attention to Measurement of the "Business Cycle"*, Journal of Monetary Economics 7, s. 151–174

Blanchard, O.J., Quah, D., (1989), *The Dynamic Effects of Aggregate Demand and Supply Disturbances*, The American Economic Review 79(4), s.655–673

BRE Bank S.A., (2002), *Ile bezrobocia bez przyspieszania inflacji, czyli próba oszacowania NAIRU dla polskiej gospodarki*, Miesięczny Przegląd Makroekonomiczny, kwiecień

Calvo, G.A., (1983), *Staggered Contracts in a Utility-Maximizing Framework*, Journal of Monetarny Economics No. 12, s. 383–98

Canova, F., (1993), *Detrending and business cycle facts*, CEPR Discussion Paper No. 782

Chagny, O., Döpke, J., (2001), *Measures of the Output Gap in the Euro-Zone: An Empirical Assessment of Selected Methods*, Working Paper No. 1053, Kiel Institute of World Economics

Clarida, R., Gali, J., Gertler, M., (2000), *Monetary Policy Rules and Macroeconomic Stability: Evidence and some Theory*, Quarterly Journal of Economics 115, s. 147–180

Cochrane, J.H., (1994), *Permanent and Transitory Components of GNP and Stock Prices*, Quarterly Journal of Economics 109, s. 241–265

Coenen, G., Wieland, V., (2000), *A Small Estimated Euro Area Model with Rational Expectations and Nominal Rigidities*, ECB Working Paper No. 30

Duspaquier C., Guay, A., St-Amant, P., (1999), *A Survey of Alternative Methodologies for Estimating Potential Output and the Output Gap*, Journal of Macroeconomics 21, s.577–595

Elmeskov, J., (1993), *High and persistent unemployment: assessment of the problem and its causes*, OECD Economics Department Working Paper No. 132

Engle, R., Granger, C., (1987), *Co-integration and Error Correction: Representation, Estimation and Testing*, Econometrica 35, s. 251–276

Evans, G., Reichlin, L., (1994), *Information, forecasts and measurement of the business cycle*, Journal of Monetary Economics 33, s. 233–254

Giorno, C., Richardson, P., Roseveare, D., van den Noord, P., (1995), *Estimating potential output, output gaps structural budget balances*, OECD Economics Department Working Paper No. 152

Gordon, R.J., (1984), *Unemployment and the Growth of Potential Output in the 1980s*, Brookings Papers on Economic Activity No.15(2)

Griliches, Z., Mairess, J., (1995), *Production Functions: The Search for Identification*, NBER Working Paper No. 5067

Hall, R.E., (1978), *Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence*, Journal of Political Economy 86

Hamilton, J.D., (1994), *Time Series Analysis*, Princeton University Press

Hodrick, R., Prescott, E., (1980), *Post-war US business cycles: an empirical investigation*, Working Paper, Carnegie Mellon University

Johansen, S., (1988), *Statistical analysis of cointegrating vectors*, Journal of Economic Dynamics and Control 12, s. 231–254

Johansen, S., (1991), *Likelihood-based inference in cointegrated vector autoregressions*, Econometrica 59

King, R.G., Plosser, C.I., Stock, J.H., Watson, M.W., (1991), *Stochastic Trends and Economic Fluctuations*, The American Economic Review 81(4), s. 819–839

Koyck, L.M., (1954), *Distributed Lags and Investment Analysis*, North-Holland, Amsterdam

Kwiatkowski, E., (2002), *Bezrobocie. Podstawy teoretyczne*, PWN, Warszawa

Layard R., Nickell S., Jackman R., (1991), *Unemployment: Macroeconomic Performance and the Labour Market*, Oxford University Press

Lippi, M., Reichlin, L., (1994), *Diffusion of technical change and the decomposition of output into trend and cycle*, Review of Economic Studies 61, s. 19–30

Okun A. (1962), *Potential GNP: Its Measurement and Significance*, American Statistical Association, Proceedings of the Business and Economic Statistics Section, Washington

Orphanides, A., van Norden, S., (2002), *The Unreliability of Output-gap Estimates in Real Time*, The Review of Economics and Statistics 84 (4), s. 569–583

Prescott, E.C., (1997), *Needed: A Theory of Total Factor Productivity*, Federal Reserve Bank of Minneapolis

Rocznik Statystyczny Rzeczpospolitej Polskiej 2002, GUS, Warszawa

Rõõm, M., (2001), *Potential Output Estimates for Central and East European Countries Using Production Function Method*, Working Papers of Eesti Pank No. 2

Schumacher, Ch., (2000), *Forecasting Trend Output in the Euro Area*, HWWA Discussion Paper 109/2000, Hamburg

Staiger, D., Stock, J.H., Watson, M.W., (1996), *How precise are estimates of the natural rate of unemployment?*, NBER Working Paper No. 5477

Taylor, J.B., (1993), *Discretion versus Policy Rules In Practice*, Carnegie-Rochester Series on Public Policy 39, s. 195–214

Welfe, W., Welfe, A. (1996), *Ekonometria stosowana*, PWE, Warszawa

Yang, M., (1998), *On identifying permanent and transitory shocks in VAR models*, Economic Letters 58, s. 171–175

Zięba, J. (2003), *Długo- i krótkookresowa składowa Produktu Krajowego Brutto wedlug metody Blancharda i Quaha*, Materialy i Studia nr 155, NBP

Żółtowska, E., (1997), *Funkcje produkcji, teoria, estymacja, zastosowania*, Uniwersytet Łódzki

Appendix

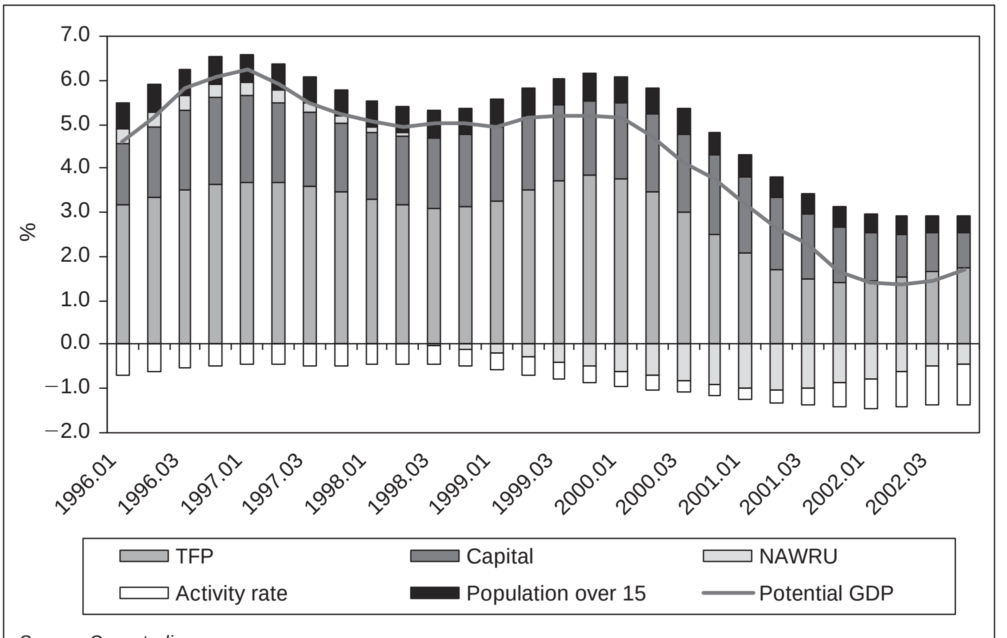
A. Decomposition of the potential GDP growth in Poland

One of the advantages of the method based on the production function is that it can be used to calculate the contribution of particular factors to potential GDP growth. If the professional activity rate is defined as the proportion of the labour force to the number of inhabitants older than 15 and if the estimated production function is used, the rate of potential GDP growth can be broken down into five factors using the following formula:

$$\Delta \log(PKB*) = \Delta \log(TFP*) + 0.57\Delta \log(\overline{L}) + 0.57\Delta \log(s) + 0.57\Delta \log(1 - NAWRU) + 0.43\Delta \log(K)$$

where: $\overline{L}$ is the number of inhabitants older than 15 and s is the professional activity rate defined above.

Chart 8 – Decomposition of the potential GDP growth (year on year) in Poland



Source: Own studies.

Chart 8 shows that long-term GDP growth in Poland in 1995–2002 was mainly driven by the increasing total factor productivity. In every quarter of this period, the share of TFP in the potential GDP growth exceeded 60%, or even 90% in 2002. Between 1995–2002, the potential GDP increased on average by 2.9 percentage points a year due to TFP growth. Over the entire period analysed, the contribution of capital growth and growth of the population older than 15 were also positive, and amounted, respectively, to 1.6 and 0.6 of a percentage point. The breakdown reveals unfavourable trends on the Polish labour market. As a result of the decreasing professional activity rate, the long-term GDP growth rate was 0.5 percentage points lower every year. Since the end of 1998, the NAWRU unemployment level has been growing in Poland, slowing the potential GDP growth in this period by 0.6 percentage points a year on the average. The growing NAWRU can be interpreted as a symptom of unemployment hysteresis.

B. Permanent-transitory decomposition (Yang (1998))

Consider an n-dimensional stochastic process $\{x_t\}$ of integrated variables and its error correction (VECM) representation:

$$\Delta x_t = \delta + \theta x_{t-1} + \sum_{i=1}^{p} \Pi_i \Delta x_{t-i} + \varepsilon_t \qquad (1)$$

where $\Delta = 1 - L$ is the difference operator and L is the lag operator, δ is deterministic part of the model, θ and Π_i are matrices of parameters. ε_t is a vector of random disturbances from individual equations with variance-covariance matrix Ω. The system (1) could be estimated by the ML procedure developed by Johansen (1988). In case of r cointegrating relations, tested by Johansen procedures, θ has a reduced rank and can be decomposed as $\theta = \alpha\beta'$ where α and β are full-rank $(n \times r)$ matrices. Columns of β are cointegrating relations between the variables of the system and elements of α are system reactions to previous period disequilibria. The system should converge monotonically to the long-run relationship $\beta' x_t$ with an adjustment rate given in in α. Defining $\Pi(L) = I_n - \sum_{i=1}^{p} \Pi_i L^i$ and $A(L) = \Pi(L)(1 - L) + \theta L$ one can convert VECM into a corresponding VAR representation: $A(L)x_t = \delta + \varepsilon_t$.

As elements of x_t are I(1), the Wold theorem assures that its first differences have an infinite Vector Moving Average representation (VMA), showing the way disturbances of previous periods affect the current value of variables:

$$\Delta x_t = C(L)(\delta + \epsilon_t) = \mu + \epsilon_t + C_1\epsilon_{t-1} + C_2\epsilon_{t-2} + \cdots \tag{2}$$

where $\mu = C(1)\delta$ is a deterministic part, $C(1) = \sum_{i=0}^{\infty} C_i$ is a sum of all short-run multipliers and the matrix polynomial is of the form: $C(L) = \sum_{i=0}^{\infty} C_i L^i$ with normalization $C(0) = I_n$.

Engle and Granger (1987) showed that by defining $C^*(L) = (C(L) - C(1))(1-L)^{-1}$ equation (2) can be represented as:

$$\Delta x_t = C(L)(\delta + \epsilon_t) = \mu + C(1)\epsilon_t + C^*(L)\Delta\epsilon_t \tag{3}$$

In case of C(1) being of reduced rank k ($k < n$), there are $r = n - k$ cointegrating relations. C(1) can be decomposed as: $C(1) = hg'$ where h and g are $(n \times k)$ matrices. Equation (3) shows the decomposition of the matrix polynomial C(L) into a permanent part C(1) and a transitory lag distribution $C^*(L)\Delta\varepsilon_t$. It is clear that there are only k linear combinations of disturbances permanently affecting x_t – they are of the form $g'\varepsilon_t$.

Johansen (1992) showed that $C(1) = \beta_+(\alpha_+^T(I - \sum_{i=1}^{p}\Pi_i)\beta_+)^{-1}\alpha_+^T$ where α_+ and β_+ are orthogonal complements[14] to corresponding matrices. Equation (3) can be then presented as:

$$\Delta x_t = \mu + \beta_+(\alpha_+^T(I_n - \sum_{i=1}^{p}\Pi_i)\beta_+)^{-1}\alpha_+^T\epsilon_t + C^*(L)\Delta\epsilon_t \tag{4}$$

with $\alpha_+^T\varepsilon_t$ being a group of permanent shocks and matrix $\beta_+(\alpha_+^T(I - \sum_{i=1}^{p}\Pi_i)\beta_+)^{-1}$ showing their long-run impact on the variables creating the system.

In order to obtain the components of x_t, that are created by permanent disturbances, taking into account not only their long-run, but also short-run impact, one should find a connection between long-run shocks $\alpha_+^T\varepsilon_t$ and the MA representation of (2). The stochastic component $C(L)\varepsilon_t$ can be divided into 2 components, of which one is the permanent one $\alpha_+^T\varepsilon_t$:

$$C(L)\epsilon_t = C(L)\bar{\alpha}_+ \cdot \alpha_+^T\epsilon_t + C(L)\bar{\gamma}\cdot\gamma^T\epsilon_t \tag{5}$$

where γ is any $(n \times r)$ matrix chosen to assure that $[\alpha_+ \; \gamma]$ is invertible and satisfying the equation:[15] $[\bar{\alpha}_+ \; \bar{\gamma}] = \{[\alpha_+ \; \gamma]^T\}^{-1}$. Representation showed in (5) assumes that multipliers of permanent and transitory disturbances are linear combinations of multipliers of model (3). The dynamic influence of shocks on variables forming x_t was then divided into disturbances coming from the long-run and short-run. To finish the decomposition one must choose matrix γ. Assuming that permanent disturbances $\alpha_+^T\varepsilon_t$ and transitory ones $\gamma^T\varepsilon_t$ are orthogonal (independent), which is desired in impulse-response analysis, one can get the following relation defining γ:

$$\gamma = \alpha - \alpha_+(\alpha_+^T\Omega\alpha_+)^{-1}\alpha_+^T\Omega\alpha \tag{6}$$

14 *Orthogonal complement of* $(n \times r)$ *matrix* α *is a* $(n \times n - r)$ *matrix* α_+ *that satisfy the relation* $\alpha^T\alpha_+ = 0$. *It's a matrix which columns form a basis of subspace, which is orthogonal to the subspace generated by columns of* α *matrix.*

15 *It is the solution to the equation of the form:* $\bar{\alpha}_+ \cdot \alpha_+^T + \bar{\gamma} \cdot \gamma^T = I_n$

Permanent components of vector x_t could be obtained as a deterministic part of equation (3) with the whole stochastic component of permanent disturbances:

$$x_t^* = \mu + (C(L)\bar{\alpha}_+)\alpha_+^T \varepsilon_t$$

Thus x_t^* is defined to be that part of x_t,that was generated by only permanent disturbances, so according to our definition, it could be treated as trends of elements of vector x_t.

Abstract

This article presents three estimates of the output gap, one using the production function method, and the other two by assessing the long-term product using cointegration relationships (based on the production function and on the hypothesis of permanent income). It also presents an analysis of time-relationships between the estimated output gaps and selected measures of inflation using the covariance of a VAR-type stochastic process. The methods employed yield different estimates of the output gap. The time paths of calculated gaps and the analysis of time relationships (conditional on the existence of relationship described by the Phillips curve and the possibility of using obtained gaps in it) allow the authors to conclude that there's no inflationary pressure from the aggregate demand in the Polish economy, at least till the end of 2003.

Key words: Output gap, VECM, production function, Permanent-Transitory Decomposition

JEL classification: E32, C32.

Michał Gradzewicz
e-mail: Michal.Gradzewicz@mail.nbp.pl

Marcin Kolasa
e-mail: Marcin.Kolasa@mail.nbp.pl

National Bank of Poland, DAMS

Estimating capacity utilization from survey data

Norman Morin and John Stevens (Federal Reserve Board) [1]

1. Introduction

The amount of resource slack in the economy is closely watched by policymakers, academics, and industry analysts. The Federal Reserve publishes monthly estimates of capacity utilization that have long been used to help analyze developments in the industrial sector. Aggregate measures of utilization are constructed from detailed industry-level utilization rates that are themselves often used to reveal potential industry-level bottlenecks. These measures of capacity utilization are used to signal emerging supply chain problems, to forecast investment by manufacturers, and to assess the likelihood of an acceleration or deceleration in inflation.

The estimates produced by the Federal Reserve reflect a methodology that has been continuously refined over nearly 50 years. Since 1990, the principal data source used by the Federal Reserve to construct estimates of manufacturing capacity has been the Survey of Plant Capacity (SPC) conducted by the United States Census Bureau.[2] The Census Bureau and the Federal Reserve have adopted an economic definition of capacity that assumes the full employment of all variable factors of production and the use of only the equipment in place and ready to operate.[3] This definition, made more precise later, captures the key aspects of capacity utilization that are thought to be useful for analyzing prices, capital spending, and industry bottlenecks.

The Federal Reserve's measures build upon the SPC rates in two key ways. First, the Federal Reserve attempts to remove statistical noise that arises from sampling error. For a given industry, the Federal Reserve calculates an initial capacity index by dividing the Federal Reserve's index of industrial production (IP) by the corresponding SPC utilization rate. Because these series are from different data sources, whose coverage and construction may differ, part of the annual movement of their ratio may simply reflect measurement error. The Federal Reserve's final capacity indexes combine these initial capacity indexes with other indicators of capacity that are suggested by economic theory and consistent with the definition of capacity described above. Only that part of the initial capacity indexes related to the other measures of capacity expansion is incorporated into the final estimates, and as a result, the noise in the initial capacity indexes is removed. Second, the Federal Reserve sharpens the signal from the SPC measures by bringing other information to bear, such as a knowledge of changes in the SPC survey questions; changes in the sample construction for the SPC; and capacity information in physical units from other sources. Thus, the Federal Reserve's estimation procedures in no way reflect shortcomings of the SPC, but rather represent an analytical exercise that is not possible when simply tabulating survey responses.

In the past, some observers have argued that survey-based utilization rates tend to have less cyclical amplitude than would be suggested by other, more direct, estimates of capacity, including surveys that directly ask about capacity rather than utilization (demonstrated initially by Perry, 1973).[4] Consequently, if utilization rates vary less over the business cycle, then a capacity index directly constructed from those rates will exhibit more cyclical variability than would be suggested by other indicators of capacity, such as industry capital spending patterns or physical estimates of capacity (as utilization rates appear in the denominator). We test the cyclicality of

1 *We would like to thank Carol Corrado and Joyce Zickler for helpful comments and suggestions. The opinions expressed here are those of the authors and not necessarily those of the Board of Governors of the Federal Reserve System or its staff.*

2 *The Survey of Plant Capacity began in 1974 and is jointly funded by the Federal Reserve Board and the Department of Defense.*

3 *See Corrado and Mattey (1997) and Forest (1979) for more detailed discussions of the economic underpinnings of the various definitions of capacity.*

4 *Also see Raddock and Forest (1976), Christiano (1981), and Schnader (1984).*

capacity indexes constructed using the SPC and the Federal Reserve Board (FRB) utilization rates over the past 30 years, and find that neither measure is excessively cyclical at the aggregate level. At the detailed industry level, however, we find that for a several industries, capacity indexes constructed directly from SPC utilization rates exhibit excess cyclicality, but this excess movement is removed by the Federal Reserve's methodology. The risk, of course, is that the Federal Reserve removes too much information. We cannot test this possibility directly, as the Federal Reserve also adds information, but we can test the *net* effect of the Federal Reserve's methodology by examining the ability of both sets of utilization rates to predict capital spending, future capacity expansions, and prices.

A brief history of capacity measurement at the Federal Reserve is provided in the next section. In section 3 we provide details on the SPC, and in section 4 we walk through the methodology used by the Federal Reserve to combine the SPC data with other indicators of capacity change. Section 5 analyzes the cyclical properties of the SPC and FRB utilization rates, and tests whether the refinements of and additions to the SPC rates results in any net loss of useful information.

2. A brief history of the Federal Reserve capacity and capacity utilization measures

Indexes of output and capacity were first developed by the Board's staff during the economic expansion in the mid-1950s (see Raddock, 1987). These early estimates covered several major manufactured materials. The major materials indexes were based on measures of physical volume from government and trade sources, and were used internally to analyze current business conditions, primarily inflationary pressures and the demand for capital goods. In the 1960s, the Federal Reserve maintained separate measures of capacity and utilization for manufacturing and for selected industrial materials. Unlike the unpublished major materials index, however, the published estimates for manufacturing were not constructed from physical volume data. The manufacturing capacity indexes were instead based on end-of-year utilization rates from the McGraw-Hill survey of capacity utilization that were divided into December values of the Federal Reserve's indexes of production.[5] The year-to-year changes in these implied capacity estimates were then refined using alternative indicators of capacity expansion, such as a measure of gross capital stocks, and linearly interpolated to the quarterly frequency.[6]

Periodically throughout the 1970s and 1980s, the detail covered by the manufacturing indexes and the materials indexes was expanded, and in 1983, the scope of coverage was widened to include mining and utilities.[7] Utilization rates from the new Survey of Plant Capacity (SPC) from the Bureau of the Census, which started in 1974, began to be incorporated into the Federal Reserve estimates.

In 1990, the publication of the capacity and capacity utilization figures were combined into a single statistical release with the industrial production indexes. The 1990 capacity revision, described in the June 1990 *Federal Reserve Bulletin* (Raddock, 1990), created an integrated and more detailed system of output, capacity, and capacity utilization measures for total industry and for a variety of industry sub-aggregates. This move resulted in several changes to the overall capacity system. Most importantly, the materials system was discontinued as a separate entity; the primary source of utilization rates for manufacturing industries became the Bureau of the Census's Survey of Plant Capacity; and the number of detailed industry-level measures in manufacturing was more than doubled to 54 individual series.

Further revisions to the capacity system in the 1990s maintained the structure introduced in the 1990 revision. Individual series were occasionally added or eliminated to reflect changes in the related production indexes, and several technical improvements were introduced.

5 *McGraw-Hill, Inc. (later its DRI subsidiary) collected annual data on both utilization rates and on capacity expansion from a sample of large companies each December from 1954 to 1988.*

6 *Typically, the logarithm of the ratio of an industry's implied capacity to its capital stock – a capital productivity measure – was regressed on a series of deterministic trends and dummy variables; implicitly, the model assumed a unit elasticity of capacity with respect to the stock of capital, an assumption that was relaxed in 1997. A similar model was run using the logarithm of the ratio of implied capacity to the capacity index directly asked about in the McGraw-Hill survey. The annual estimate of Federal Reserve capacity was the average of the fitted values from the two models.*

7 *The new monthly Federal Reserve statistical release (G.3), "Capacity Utilization," began in January 1977 and included monthly utilization rates, as well as quarterly data for output, capacity, and utilization for manufacturing and industrial materials (and their major component series).*

The capital measures that are used as alternative measures of capacity expansion were further refined in 1995 to reflect the flow of services derived from the net stocks of productive assets (Raddock, 1996). The measures – known as capital input or capital services – are rental-price-, or user-cost-, weighted indexes of the asset-level net capital stocks; that is, the indexes weight growth rates in the net stocks of individual assets by an estimate of that asset's share of the aggregate marginal product of the industry's capital. In 1997, the regression models that relate SPC-based implied capacities to alternative indicators of capacity expansion (see footnote 4) were made more flexible by relaxing the restriction of a unit elasticity on the capital measure (Gilbert, Morin, and Raddock; 2000).[8]

In December of 2002, the Federal Reserve issued a comprehensive revision of industrial production, capacity, and capacity utilization whose primary purpose was to reclassify the detailed industry structure of production and capacity from the Standard Industrial Classification (SIC) system to the North American Industry Classification System (NAICS).[9] The reclassification changed industry details a bit but left the overall industry coverage of the capacity system essentially unchanged. As of the 2002 comprehensive revision, the capacity system included 85 individual series – a mix of 3-, 4- and 6-digit NAICS industries – of which 67 are in manufacturing, 16 in mining, and 2 in utilities.

In addition to the Federal Reserve, a number of groups have, at various points in time, published indexes of capacity and capacity utilization.[10] In the early 1980s there were seven separate capacity and utilization measures that covered the manufacturing sector. These measures included the annual estimates from McGraw-Hill and the fourth-quarter SPC measures from the Census Bureau, both of which served as source data for the FRB estimates. Quarterly estimates were published by the BEA, Wharton Econometric Forecasting Associates (WEFA), and Rinfret Associates (see Schnader, 1984 for a detailed discussion of the various measures). Currently, the only long-running survey-based measure of capacity utilization with broad industry coverage is the SPC; the Institute for Supply Management (or ISM, formerly the National Association of Purchasing Management, or NAPM) has published semiannual estimates of manufacturing utilization rates since December 1985.[11] Narrower measures for specific industries are produced by various government and trade groups, often in terms of physical units (such raw steel capability from the American Iron and Steel Institute).

3. The Survey of Plant Capacity

The Survey of Plant Capacity (SPC) is now the primary source for the annual utilization rates used to construct the monthly capacity and capacity utilization rate measures for manufacturing industries. The SPC, a mandatory survey that is jointly sponsored by the Federal Reserve Board and the Department of Defense, measures fourth-quarter rates of capacity utilization for about 17,000 manufacturing plants at the 6-digit NAICS industry level.[12] The SPC began with

8 *Two additional refinements to the construction of capacity were introduced in 1999 and described in the March 2000* Bulletin. *First, a new interpolation procedure was introduced to form monthly time series of capacity based on the fourth-quarter baseline capacity estimates produced by the regression models. The new procedure allowed capacity growth rates to change smoothly over time instead of imposing a constant growth rate throughout the year, while maintaining the same fourth-quarter to fourth-quarter growth rates calculated under the old procedure. Second, the models that relate SPC-based implied capacity to alternative indicators of capacity were expanded to include variables that capture the age profile of the capital stock. In several studies, age variables have been used to better capture the effect of technological change.*

9 *This undertaking, which involved mapping numerous data sources (including plant-level micro-data) from SIC to NAICS, is documented in Corrado (2003) as well as a series of papers: Bayard and Klimek (2003), Morin (2003), and Stevens (2003).*

10 *Even considering just a single plant, there exists a range of definitions of capacity. "Engineering capacity" is the most uncomplicated notion of capacity – the maximum level of output when operating the existing machinery at the peak possible linespeed nearly 24 hours per day, 7 days per week with only minimal downtime. Conversely, capacity may refer the plant manager's "preferred capacity," a unit cost-minimizing level or profit-maximizing level of output, which, except for a few specific industries, is likely to be noticeably lower than the engineering maximum. In between is the concept of "full production" capacity, in which capacity is the level at which all variable inputs are used at the maximum level, without consideration of the rising materials, labor, and other costs that would undoubtedly be present as output exceeded preferred capacity. Furthermore, one set of assumptions may be appropriate to answer questions about inflationary pressures or industry bottlenecks and another for mobilizing resources in wartime.*

11 *The Institute for Supply Management (ISM) surveys about 400 firms and is not a statistical sample, although the distribution of companies surveyed roughly corresponds to the relative composition of two-digit SIC industries.*

12 *The 2002 Census of Manufacturers recorded 344,000 plants (down about 20,000 from the 1997 Census), so the SPC sample represents about 5 percent of manufacturing establishments. However, because large plants are included in the sample with certainty, the establishments surveyed by the SPC account for a touch over 50 percent of manufacturing shipments.*

a "mini-survey" in 1973 of 4000 plants that covered the fourth quarters of 1973 and 1972. The full survey began in 1974 with a sample of about 9000 plants, and was conducted annually until 1988. From 1990 to 1996, the survey was conducted biannually, but each survey collected two years worth of information. For the 1995–1996 survey, the sample size was expanded to the current size of about 17,000 plants. Since 1997, the survey has been conducted annually with this larger sample size.

In the 2002 SPC survey form, the instructions to plant managers for estimating full production capability were:

Full Production Capability – The maximum level of production that this establishment could reasonably expect to attain under **normal** and **realistic** operating conditions fully utilizing the machinery and equipment in place. In estimating market value at full production capability, consider the following

- Assume **only** the machinery and equipment in place and ready to operate will be utilized. Do not include facilities or equipment that would require extensive reconditioning before they can be made operable.
- Assume **normal** downtime, maintenance, repair, and cleanup. If full production requires additional shifts or hours of operation, then appropriate downtime should be considered in the estimate.
- Assume labor, materials, utilities, etc. are fully available.
- Assume number of shifts, hours of plant operations, and overtime pay that can be sustained under **normal** conditions and a **realistic** work schedule.
- Assume a product mix that was **typical** or representative of your production during the fourth quarter. If your plant is subject to short-run variation assume the product mix of the current period.
- Do not assume increased use of productive facilities outside the plant for services (such as contracting out subassembly work) in excess of the proportion that would be normal during the fourth quarter.

The SPC is a statistical survey. A new probability sample for the SPC is drawn every five years from the Census of Manufactures; the 2004 SPC will be the first year of a new sample drawn from the 2002 Census. Each industry is treated independently, and, based on the Tillé sampling procedure (Slanta, 2003), establishments are selected with a "probability proportionate to size;" industries denoted as "priority industries" by the Department of Defense are sampled more heavily. Census staff follow up on non-responses with additional mailings or phone calls, and final response rates are around 80 percent. Preliminary survey results are published around October, and final SPC results are normally published in January; revisions are usually minimal. The sample is augmented annually to reflect new plant births.

In 1999, the SPC began to be published on a NAICS basis, and this transition in classification schemes affected the SPC sample in two significant ways: Both logging and the publishing piece of "printing and publishing" left the NAICS-based manufacturing sector. However, the Federal Reserve's definition of the industrial sector did not change, and at the request of Federal Reserve Board, the Census Bureau agreed to continue sampling publishers; because presently the sample is drawn from the Census of Service Industries. Unfortunately, because logging is now under the purview of the Department of Agriculture, the Census Bureau was not able to continue sampling firms in this industry.

The SPC has implemented three significant changes since 1974. First, in 1982 respondents were requested to complete the survey form even if the plant was idle (but not permanently closed) during the fourth quarter. Presumably, before 1982, the SPC undercounted idled plants, and, consequently, reported industry-level utilization rates that were higher in downturns than would otherwise have been the case (although this has been difficult to detect statistically).

Second, before the 1989–1990 survey, plant managers were asked about their "preferred level of operation" and "practical capacity"; now the survey asks for the level of "full production" and "national emergency production." However, the definitions of "preferred level of operation" and "full production" appear close enough that they are treated as a single time series without any ad hoc adjustments.[13]

Finally, in addition to the large sample expansion with the 1995–1996 survey, the survey implemented a change to the assumptions about plant shifts. In surveys before the 1995–1996

13 *Doyle (2000) showed that, using plant level data, the Federal Reserve's assumption overlaying of preferred utilization rates from the pre-1989 surveys and full production rates from the surveys for the 1989–1994 period is not rejected by the data.*

SPC, the respondent was instructed to determine capacity hours and shifts by using the maximum level attained in the last five years; since then, the respondent is allowed to assume extra shifts at capacity: "***If full production requires additional shifts or hours of operation***, then appropriate downtime should be considered in the estimate" (bold italics added). This suggests that if the five year moving maximum of shifts and hours that was used before 1995 is less than what a plant manger would have chosen unconstrained, then the reported level of capacity before 1995 was lower than it would have been under the newer instructions; therefore, all else equal, one would expect that utilization rates should have exhibit a discrete downward shift in 1995. This shift is observed in the data; at the manufacturing level, the discrete shift appears about 4 percentage points.[14]

In addition to answering questions about actual production, full production capability, and national emergency production capability, managers report other useful information. For example, managers report reasons for changes in full production capability relative to the fourth quarter a year earlier. Reasons include:

- capital expenditures
- capital retirements
- price changed but product mix is the same
- change in method of operation
- change in product mix or product specifications
- change in material input

Managers also provide reasons for operating at less than 100 percent of their full production capability in the fourth quarter. Reasons include:

- Not most profitable to operate at full production capability
- Insufficient supply of materials
- Insufficient orders
- Insufficient supply of local labor force/skills
- Lack of sufficient fuel or electric energy
- Equipment limitations
- Storage limitations
- Logistics/transportation constraints

Managers indicate the minimum time that would be required to ramp up production to both full and national emergency production levels.

The survey also collects information on shift-work patterns. The data on the workweek of capital has been used by researchers to investigate the procyclicality of productivity and capital utilization.[15] The survey asks plant managers for

- Days per week in operation
- Plant hours per week in operation
- Weeks in operation in the quarter
- Total number of production workers
- Hours worked by temporary production workers
- Overtime hours worked by production workers
- Total number of temporary production workers
- Total hours worked by production workers

The SPC and other surveys of utilization rates may yield significantly different utilization rates for a given industry as a result of important differences in the degree of the specificity of the survey's definition of capacity, the sampling unit (plant or firm), the sample size, and the industry composition of the sample. For example, the establishment-based rates in the SPC are substantially lower than the rates for the same industry in the ISM survey of firms, whose sample includes companies with multiple establishments. Between 1990 and 2002 the aggregate SPC rates for manufacturing averaged about 10 percentage points lower than the operating rates from the ISM. Multi-establishment companies presumably take into account intrafirm bottlenecks that limit the overall capacity of the firm, whereas the respondents to the Survey of Plant

14 *In principle, the downward shift in rates can be decomposed into an effect due to the wording change and to the sample expansion. We plan to explore this issue further using the microdata from the SPC at the Census Bureau's Center for Economic Studies.*

15 *See, for example, Mattey and Strongin (1995), Beaulieu and Mattey (1996), and Shaprio (1996).*

Capacity only consider the maximum output of their own establishment (see Bureau of the Census, 1983).[16]

4. Constructing the Federal Reserve estimates of capacity and capacity utilization

The Federal Reserve constructs capacity indexes and utilization rates that completely cover the industrial sector (manufacturing, mining, and electric and gas utilities) and that are consistent with the Federal Reserve measures of industrial production. Estimates for industry aggregates, such as manufacturing, are constructed by combining the individual series. Six general steps are involved in calculating the utilization rates published by the Federal Reserve:

Step 1: Construct preliminary implied capacity indexes

The first step in producing a capacity index is to divide the Federal Reserve production index for the industry by a benchmark utilization rate – both are typically either fourth-quarter or end-of-year estimates. The implied capacity index ($ICAP_t$) for period t is:

$$ICAP_t = IP_t / U_t, \quad (1.1)$$

[17]

and, like the production index, is expressed as a percentage of output in a base year. For about 90 percent for manufacturing capacity, the Survey of Plant Capacity provides the utilization rate for denominator of (1.1).

The implied capacity estimates in (1.1) provide the general trend movements of capacity as well as initial estimates of the levels that are consistent with the Federal Reserve production indexes. For example, if the production index for an industry has been roughly constant while the survey-based utilization rates have risen, then the implied capacity index would trend down.

Step 2: Relate the implied capacity estimates to alternative indicators of capacity

Although a capacity index published by the Federal Reserve derives its level and trend movements from the implied capacity index, the annual changes in capacity are determined by additional information on the economic determinants of capacity expansion. The Federal Reserve uses regression-based procedures to combine these additional measures with the SPC utilization rates. The purpose of the regressions is to ensure that the year-to-year changes in the published estimates of capacity conform to movements in the alternative determinants of capacity change. For about 90 percent of manufacturing industries, the principal alternative indicator is a measure of industry capital input. Relating the implied capacity indexes to these other measures removes from the implied capacity index the part of the year-to-year movements that appears to be measurement or sampling error-related noise and that does not appear to represent actual changes in an industry's productive capacity.

The refined estimates of annual capacity are the fitted values of the regression of implied capacity on industry capital input (K_t); a deterministic trend (t); dummy variables for outliers, level shifts and trend breaks (D_{it}); and on a variable related to the average age of the capital stock, A_t.[18]

$$\log\left(ICAP_t\right) = \alpha_0 + \alpha_1 t + \beta \log\left(K_t\right) + \gamma \log\left(A_t\right) + \sum_i \delta_i D_{it} + \upsilon_t \quad (1.2)$$

16 *Moreover, once one moves beyond a single plant, the practical maximum output of an industry is typically less than the sum of the individual capacities of the constituent plants; although an individual plant may produce at its maximum possible rate, all plants attempting simultaneously to produce at capacity may induce supply bottlenecks for critical inputs, and factor price increases might make producing at capacity infeasible. By similar reasoning, the capacity for the manufacturing sector is clearly less than the sum of industry-level capacities.*

17 *For example, if the production index in the fourth quarter of 2002 is 120 (120 percent of the average of 1997 production) and the related utilization rate is 80 percent, the implied capacity index for 2002 is 120/0.8 = 150.0.*

18 *The age variable is the ratio of the age of an industry's capital stock relative to its expected service life, given the mix of assets that compose the stock. This measure represents the portion of the aggregate life of a given mix of assets that has been used up. In several studies, age variables have been used to capture the effect of embodied technological change – the idea that productivity augmenting technological change is vintage specific, that is, it is embodied in the design of new equipment and structures, rather than affecting all existing inputs in the production process.*

or, where the lower case letters represent natural logarithms,

$$icap_t = \alpha_0 + \alpha_1 t + \beta k_t + \gamma a_t + \sum_i \delta_i D_{it} + \upsilon_t. \tag{1.3}$$

A principal result from the regression-based procedure is that the coefficients of the capacity regression determine the relationship between capital productivity – the ratio of capacity and capital input – and the determinants of capacity. Rewriting (1.2) in terms of capital productivity yields

$$icap_t - k_t = \alpha_0 + \alpha_1 t + (\beta - 1)\, k_t + \gamma a_t + \sum_i \delta_i D_{it} + \upsilon_t. \tag{1.4}$$

Equation (1.4) shows that one can represent the model-based capacity estimates as the sum of the contributions of capital input and capital productivity, where capital productivity embodies the combined effects of total factor productivity, labor at capacity (such as the work period of capital at capacity), and capital deepening.[19]

In short, the *trend* in a published industry-level capacity index is derived primarily from the trend in the industry's implied capacity index, and the *annual changes* in the capacity index reflect changes in the flow of services derived from the industry's stock of capital. Although the capacity indexes that are the fitted values of (1.2) are generally procyclical – following the cycles in capital spending – they do not fluctuate as much as the preliminary implied capacity indexes, either at an annual frequency or at a business-cycle frequency.

Step 3: Interpolate the annual estimates to a monthly frequency

The end-of-year or fourth-quarter capacity estimates (depending on data source) for the 85 individual component series are interpolated to a monthly frequency. Given fourth-quarter target levels for each year, monthly rates of change are constructed via a cubic interpolation that allows monthly rates of change to evolve smoothly.

Step 4: Apply annual capability adjustments

The Federal Reserve Board's estimates of capacity attempt both to capture the concept of sustainable maximum output and to produce estimates of capacity utilization that are historically consistent, so that a given utilization rate in the present implies about the same degree of slack as in the past. The other government sources or private trade groups from which capacity estimates are derived, however, do not necessarily use a uniform definition of capacity, and their figures may be based on a different concepts of capacity. The Federal Reserve produces a correction factor – the annual capability adjustment – to minimize the effects of these definitional differences. In particular, this correction factor reduces the level of capacity for industries whose estimates appear to be based on short-term peaks or on an engineering concept rather than on an estimate of sustainable maximum output.[20]

Second, an adjustment is made for historical continuity. Most utilization rates for the manufacturing sector were based on the McGraw-Hill utilization rate survey, which, after 34 years, was discontinued in 1988. In the years that the company-based McGraw-Hill and establishment-based Survey of Plant Capacity overlapped (1974–1988), the McGraw-Hill utilization rate for an industry generally possessed a significantly higher mean than the operating rate for the same industry from the Survey of Plant Capacity. After the demise of the McGraw-Hill survey, the Survey of Plant Capacity became the principal source of manufacturing utilization rates, and the annual capability adjustments to capacity were adjusted to maintain roughly the same average utilization rate over the period in which the two surveys overlapped.

19 *In a simple constant returns to scale model of capacity as a function of capital (k), labor at capacity (L_C), and total factor productivity (A), $Q_{C,t} = A_t K_t^{\alpha} L_{C,t}^{1-\alpha}$, or, where lower case letters represent the natural logarithms of the variables, the model is $q_{C,t} = a_t + ak_t + (1 - \alpha)l_{C,t}$. The log of capital productivity is then $\pi_t = q_{C,t} - k_t = a_t + (\alpha - 1)k_t + (1 - \alpha)l_{C,t}$, which can be rewritten as $\pi_t = a_t - (1 - \alpha)(l_{C,t} - k_t)$. This represents capital productivity as proportional to total factor productivity and inversely related to capital deepening, which is determined by labor at capacity and capital.*

20 *The adjustment is particularly large for electricity generation, where up to one third of generating capacity is reserved to meet peak summer demand, and, in addition, where a considerable amount of capacity is kept as a safety margin. Much of this capacity is not sufficiently efficient to run on a consistent basis, and is excluded from the Federal Reserve estimate of sustainable output.*

Step 5: Construct aggregate series

The aggregation of capacity and capacity utilization rates presents distinct issues compared with the aggregation of individual production indexes, as capacity and utilization are constructed and defined in relation to industrial production.[21] An annual utilization aggregate is calculated as

$$U_{At} = \frac{\sum_{i \in A} P_{it} I_{it}}{\sum_{i \in A} P_{it} C_{it}} = \frac{\sum_{i \in A} \left(P_{it} C_{it} \right) U_{it}}{\sum_{i \in A} P_{it} C_{it}}, \tag{1.5}$$

where I is the industry-level production index, P is industry-level unit value-added, C is the capacity index, and U is the annual utilization rate. Thus, the aggregate annual utilization rates are equivalent to capacity-weighted aggregates of individual utilization rates; that is, they are a combination of the individual utilization rates weighted by proportions that reflect the individual's share in the aggregate current value of production at capacity.

Monthly capacity aggregates are constructed in three steps:

- Utilization aggregates are calculated on an annual basis through the most recent full year as in (1.5).
- The *annual* aggregate capacity index is derived by dividing the corresponding production index by the utilization aggregate.
- The *monthly* aggregate capacity index is obtained by interpolating the annual capacity index from the previous step with a Fisher index of its constituent monthly capacity series. For the very recent period, since the most recent full year, each monthly capacity aggregate is extrapolated by this same Fisher index, adjusted by a factor that accounts for the differences in their relative growth rates.[22]

Step 6: Construct aggregate utilization rates

Aggregate utilization rates are calculated by dividing the appropriate production index by the related capacity index.

5. The cyclical and explanatory properties of Federal Reserve capacity indexes

The Federal Reserve Board (FRB) and the Census Bureau's Survey of Plant Capacity (SPC) are currently the only sources for lengthy time series of detailed industry-level utilization rates.[23] As discussed earlier, utilization rates from previous long-running surveys – specifically, from the BEA and McGraw Hill – possessed less cyclical amplitude than other, more direct measures of capacity would imply (Perry, 1973; Christiano, 1981). Given an index of production, damped cyclical amplitude for utilization rates mechanically implies capacity indexes that exhibit greater cyclical movements than could be explained by capital spending patterns and changes in the capital stock. Indeed, at the industry level, the survey-based implied capacity measures often implied contractions of capacity in recessions that appeared implausible, and this "lost" capacity was soon "found" as the economy recovered.[24] The excess cyclicality may represent either a cyclical bias in the implied capacity indexes or a cyclical bias in the alternative indicators of capacity expansion. As discussed below, both forms of this bias will result in a positive correlation between

21 *The* Federal Reserve Bulletin *article by Corrado, Gilbert, and Raddock (1997) describes the aggregation of capacity and capacity utilization in much greater detail.*

22 *As shown by the steps above, capacity aggregates are not simply annually weighted Fisher indexes of the individual capacity series. If a capacity aggregate were to be formulated in a way similar to that of a production aggregate and if a utilization aggregate were calculated as a ratio of the two separately aggregated series, then a noticeable distortion in this utilization aggregate would occur if: (1) the relative price of a component industry changes significantly, and (2) the utilization rate of the component differs from the average of the group.*

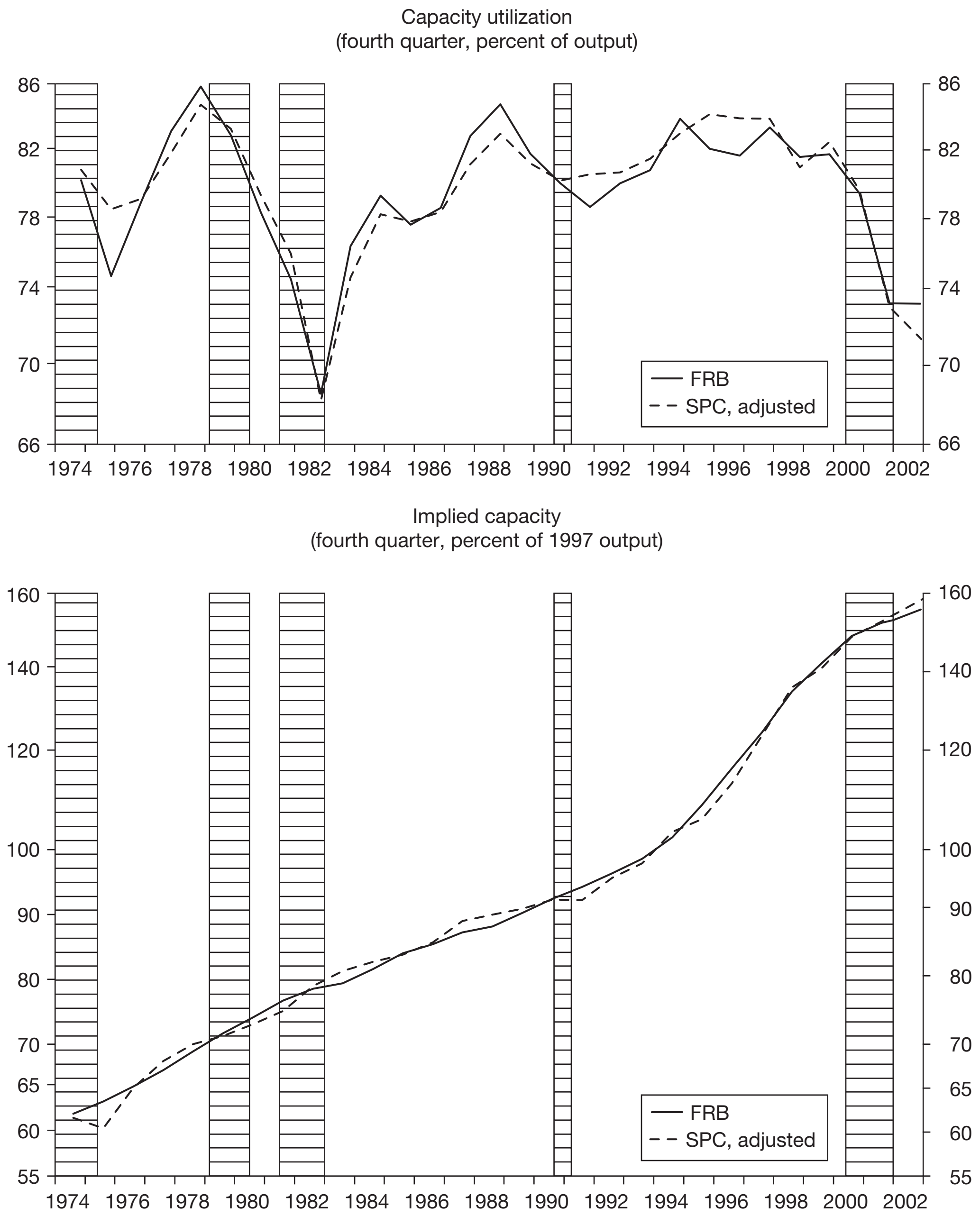
23 *The industry-level FRB rates generally begin in 1972, with many series extending back considerably farther, and the SPC rates begin in 1974. The Institute for Supply Management has published an aggregate manufacturing utilization rate semiannually since November 1990.*

24 *This lost-and-found capacity may be the result of the respondents being more likely in a downturn than in an upturn to exclude the marginal plant and equipment from their appraisal of capacity or, if the survey is at the firm level, more likely to exclude temporarily idled facilities from the calculation of capacity.*

changes in capacity and changes in production (even after conditioning on alternative measures of capacity expansion and controlling for endogeneity problems).

We look at the cyclicality of the FRB and SPC capacity indexes for both aggregate manufacturing and for the 21 3-digit NAICS manufacturing industries. The fourth-quarter manufacturing utilization rates and corresponding capacity indexes from the FRB and SPC are shown in figure 1. At the aggregate manufacturing level, we find no statistical evidence of a cyclical bias in either the FRB or SPC capacity indexes. At the detailed industry level, several SPC indexes appear to exhibit a cyclical bias, but this apparent bias does not show through to the FRB indexes. If the bias is due to mismeasurement of the alternative indicators of capacity expansion, then removing it results in a loss of information content. However, because the Federal Reserve's capacity methodology also brings other information to bear – such as knowledge about changes in SPC methodology – the net effect on the relative information content in the FRB indexes is ultimately an empirical question. Our results show that the net effect is to boost the power of the FRB indexes (relative to the SPC

Figure 1 – Manufacturing capacity utilization and capacity



Adjusted: mean difference between the SPC and FRB rates removed, and the level shift after 1994 dummied out.
The shaded areas represent periods between the peaks and toughs in total industrial production.

indexes) to predict series that capacity utilization is usually expected to influence – industry investment spending, capacity growth, and industry prices. The data appendix includes details about the sources and construction of all the data used in the models below.

Cycles in FRB rates and SPC rates

Table 1 shows that aggregate SPC utilization rates possess less cyclical amplitude than the FRB utilization rates. In the six trough-to-peak and peak-to-trough episodes since 1974, only in the 1994 to 2002 period did the SPC rates move more – both in terms of percentage points or in terms of standard deviations. In the other five episodes, the SPC rates moved at least 0.9 percentage points less than the FRB rates; and in those episodes the average difference was 2.1 percentage points, or 0.7 standard deviation, smaller. In the most recent period, the difference between FRB and SPC rates is greatly reduced if the estimated combined effect of the 1995–1996 sample expansion and change in the SPC instructions is removed from the SPC utilization rates.

Table 1 – Peak-to-trough and trough-to-peak changes in manufacturing utilization rates

	Change in utilization rates		Change in standardized utilization rates	
	FRB	SPC	FRB	SPC
1974–1978	5.7	4.0	1.5	0.9
1978–1982	−17.4	−16.5	−4.4	−3.8
1982–1988	16.2	14.7	4.1	3.4
1988–1991	−6.2	−2.4	−1.6	−0.6
1991–1994	5.3	2.5	1.3	0.6
1994–2002	−10.7	−16.0	−2.7	−3.7
1994–2002	−10.7	−11.3**	−2.7	−2.6**

Notes:

The first two columns show the change between the fourth quarters of the years indicated in percentage points. The two columns on the right show the change between the fourth quarters of the years indicated in terms of standard deviations of the respective utilization rates.

The starred entries remove the estimated effect of the sample expansion and change in instructions that began in the 1995–1996 SPC survey. The effect was estimated by a regression of the SPC utilization rates on a constant, change in manufacturing IP, a dummy variable that was 1 from 1995 through 2002, and an AR(1) error; the effect of the dummy variable was removed from the series.

Below, we investigate the cyclical properties of the FRB and SPC utilization rates and implied capacity indexes using the basic procedures employed by Perry (1973).

The difference between FRB and SPC capacity indexes

The FRB capacity index for most industries is derived from the fitted values of (1.2), where the implied capacity index is the ratio of the Federal Reserve production index for the industry divided by the SPC utilization rate. Apart from level differences (for historical continuity), the difference between the logarithms of the FRB and SPC capacity indexes should, therefore, roughly be the residuals from the regression in (1.2).[25] The difference series, then, should embody the information contained in the SPC that is discarded by the FRB capacity indexes as a result of the modeling procedure. For each industry in the table, the model used is

$$c_t^{FRB} - c_t^{SPC} = \alpha_0 + \alpha_1 \Delta q_t + \alpha_2 DUM95^{+} + \upsilon_t \quad (1.6)$$

25 *The difference, $c_t^{FRB} - c_t^{SPC}$, will not be precisely the residuals from the actual capacity models employed in the construction of the FRB capacity indexes because the published FRB capacity indexes are constructed at a significantly finer level of detail (65 NAICS manufacturing industries) and the capacity indexes for about 10 percent of manufacturing capacity are based on data in physical units from trade sources.*

where c^i is the logarithm of capacity and i = FRB or SPC; Δq is a fourth-quarter over fourth-quarter measure of instrumented output (the fitted values of the differenced-log of the industry's production index regressed on the differenced logarithm of production worker hours, the unemployment rate, and the diff-log of real GDP); and υ is an error that follows an AR(1) process. The regressions also include a level-shift dummy variable to account for the 1995–1996 change in the Census survey. The production measure is instrumented because implied capacity is defined as production divided by utilization, and as a result, regressing the changes in implied capacity on changes in production would very likely suffer from the production index being correlated with the error in the implied capacity index.

If an FRB capacity index does not exhibit excess cyclicality, but the SPC measure does show more cyclicality than expected, then the difference between the FRB and SPC measures should be significantly negatively related to current output.

The results are displayed in table 2. The difference between the FRB and SPC capacity indexes is negatively related to output in all but three cases. At the aggregate manufacturing level, however, the cyclical measure is insignificant, which implies that the difference between FRB and SPC capacity indexes is likely merely noise (apart from a positive constant related to the FRB measures retaining historical continuity with the McGraw Hill survey). However, at the detailed industry level, the difference between the FRB capacity indexes and the SPC capacity indexes is significant for 5 of 21 industries (at a 5 percent significance level).

Table 2 – Explaining the movements of the ratio of FRB capacity and SPC capacity with production and capital input

NAICS	Industry	Coefficient on Δq
	Manufacturing	−0.08
	Excluding high-tech industries	−0.08
311	Food	−0.53
312	Beverage and tobacco	−0.64*
313	Textile mills	0.03
314	Textile product mills	−0.29**
315	Apparel	−0.67***
316	Leather	−0.27*
321	Wood products	−0.01
322	Paper	0.00
323	Printing	−0.41*
324	Petroleum and coal products	−0.19
325	Chemicals	−0.32*
326	Plastics and rubber products	−0.32***
327	Nonmetallic minerals	−0.15
331	Primary metals	−0.19***
332	Fabricated metal products	−0.02
333	Machinery	−0.11**
334	Computer and electronic product	−0.11
335	Electrical equipment and appliances	−0.12*
336	Transportation equipment	0.06
337	Furniture	−0.02
339	Miscellaneous manufacturing	−0.40

Notes:

Regressions run from 1974 to 2002.

**Significant at the 10 percent level, **significant at the 5 percent level, ***significant at the 1 percent level.*

The high-technology industries include computers, communications equipment, and semiconductors (NAICS 3341, 3342, 334412-9).

The results suggest that a handful of SPC implied capacity indexes display more cyclicality than the corresponding FRB capacity indexes. However, these results are silent on whether the explanation is that the FRB indexes show no excess cyclicality, while a subset of SPC capacity indexes possess excess cyclicality; or if nearly all SPC indexes are cyclically biased, but the

FRB indexes, which are derived, in part from SPC utilization rates, inherit the cyclicality of the SPC rates. These possibilities are investigated below.

Cycles in FRB and SPC capacity indexes

A capacity index is considered to exhibit cyclical bias if there is a statistically significant positive relationship between capacity and output after controlling for the relationship between capacity and capital input.

The model used for examining the cyclicality of the capacity indexes is:

$$\Delta c_t^i = \alpha_0 + \alpha_1 \Delta q_t + \alpha_2 \Delta k_t + \upsilon_t, \tag{1.7}$$

where, in addition to the variables defined above, k is a measure of year-end industry capital input. As shown in table 3, at the aggregate manufacturing level, neither the FRB nor the SPC measures exhibit a significant and positive relationship between the change in capacity and the change in output. Differences emerge, however, at a more detailed industry level. Although FRB capacity indexes show a significant relationship with output in only two industries (one at the 1 percent level and one at just the 10 percent level), the SPC implied capacity indexes show a statistically significant and positive relationship between changes in capacity growth and changes in output (at the 5 percent level) in just over half of the industries. As Perry (1973) wrote in the context of the McGraw-Hill survey, “since the variation in the capital stock should capture much of the true variation in capacity, it is extremely doubtful that this relation between output and capacity represents a genuine case of rising output inducing capacity growth.” Instead, the most likely explanation, under Perry’s interpretation, is a cyclical bias yielding “lost-and-found” capacity in some of the detailed SPC-based indexes.

Table 3 – Explaining the change in capacity with the change in production

NAICS	Industry	Coefficient on Δq	
		FRB	SPC
	Manufacturing	−0.01	0.07
	Excluding high-tech industries	−0.02	0.03
311	Food	−0.00	1.14*
312	Beverage and tobacco	−0.02	0.81
313	Textile mills	0.08***	0.35*
314	Textile product mills	0.02	0.49**
315	Apparel	−0.01	0.89**
316	Leather	0.00	0.65***
321	Wood products	0.01	0.09
322	Paper	0.10	0.13
323	Printing	0.11	0.74**
324	Petroleum and coal products	0.21	0.86**
325	Chemicals	0.02	0.20
326	Plastics and rubber products	−0.06	0.31**
327	Nonmetallic minerals	−0.01	0.16*
331	Primary metals	−0.03	0.22**
332	Fabricated metal products	0.00	0.02
333	Machinery	−0.02	0.22**
334	Computer and electronic product	0.03	0.34***
335	Electrical equip. and appliances	−0.01	0.21**
336	Transportation equipment	−0.01	0.09
337	Furniture	0.00	0.31**
339	Miscellaneous manufacturing	−0.23*	0.19

Notes:

Regressions run from 1974 to 2002.

**Significant at the 10 percent level, **significant at the 5 percent level, ***significant at the 1 percent level.*

The high-technology industries include computers, communications equipment, and semiconductors (NAICS 3341, 3342, 334412-9).

An alternative explanation lies in the mismeasurement of changes in capital input. If the measured percent change in capital input is too large in recessions (due, for instance, to a counter-cyclical scrappage rate that is not accounted for in the construction of the underlying capital stock measures) and too low in expansions (due to the *level* of the capital stock being too high at the end of a recession from mismeasured scrappage), then we might find a spurious relationship between changes in capacity and changes in production. In effect, production changes proxy for the countercyclical scrappage. Note, however, that mismeasurement of capital input can also work in the other direction if asset depreciation rates are procyclical (i.e., when output levels are high, the equipment is worked more intensively and depreciates more rapidly, and therefore measures of capital with a fixed age-efficiency profile would be too procyclical); the FRB methodology assumes a depreciation rate that is independent of the business cycle. The net effect of cyclicality in scrappage or depreciation is not known, so we can only raise these possibilities as caveats to keep in mind when interpreting our results on cyclical bias.

In summary, neither the capacity index for aggregate manufacturing utilization based on SPC rates nor the published FRB manufacturing capacity index exhibit a significant degree of excess cyclicality. Moreover, at the detailed industry level, while a handful of SPC utilization rates appear to imply capacity indexes that possess excess cyclicality, the corresponding FRB capacity indexes, derived by combining data from the SPC with information on industry capital spending, do not exhibit excess cyclicality. If the excess cyclicality is due to mismeasurement of capital input, then removing this cyclicality from the SPC also removes valuable information from the FRB indexes. However, the FRB indexes incorporate other information beyond the SPC utilization rates, including the measures of capital; information on changes in SPC sample construction; information on changes in the SPC questionnaire; data on capacity in physical units from trade and government sources; and dummy variables to account for outliers and level- or trend-shifts. The net impact of removing from the implied capacity indexes cyclicality in those individual series in which it exists; of removing what would appear to be measurement error from the SPC; and including in the FRB measures the additional information related to capital input, survey changes, and so on, is an empirical question that hinges whether the ability of the FRB utilization rates to predict movements in series that are of interest to policymakers and analysts – such as future industry capital spending, price inflation, and capacity expansion – has been augmented or reduced relative to the utilization rates from the SPC.

Predicting industry investment

To describe the value of utilization rates as measures of slack, economists point to their ability to help predict capital spending and explain price pressures. We now investigate whether the FRB methodology adds to the ability of SPC rates to explain movements in these variables.

First, we compare the power of FRB and SPC utilization rates to explain changes in industry investment. The model is

$$\Delta I_t = \alpha_0 + \alpha_1 \Delta I_{t-1} + \alpha_2 IK_{t-1} + \beta U^i_{t-1} + \varepsilon_t \qquad (1.8)$$

where I is industry investment spending (chain-weighted, annual average), IK is the ratio of current dollar industry investment to the lag of industry current cost capital stock (annual average investment divided by end-of-year capital from the prior year), and U^i is the FRB or SPC utilization rate for the fourth quarter. The Census Bureau's Census of Manufactures and Annual Survey of Manufactures are the sources for the annual current-dollar investment data, which are chain aggregated using BEA investment deflators. See the data appendix for more detail on the construction of the investment and capital series. The investment and capital variables are included to control for investment relative to a long-run investment/capital ratio. The model is initially estimated excluding the utilization measures, and the first two columns of numbers in table 4 show the increase in the R-squared obtained by including the lagged utilization measures. The two columns on the right display the t-statistics on the lagged utilization rate measures.

After controlling for the investment/capital ratio and lagged investment, the lagged FRB and SPC utilization rates possess significant explanatory power for well over half of the industries. At the manufacturing level, the utilization rates are significant at the 1 percent level, and, for both the FRB and SPC rates, a one percent increase in manufacturing utilization rates, all else equals, leads to a 1.5 percent increase in capital spending the following year. Comparing the FRB and SPC results, the FRB rates are significant in every case in which SPC rates are significant, and

Table 4 – Predicting industry-level investment by lagged utilization rates

NAICS	Industry	Increment to R-squared		t-statistic on lagged utilization	
		FRB	SPC	FRB	SPC
	Manufacturing	32.4	29.1	3.6***	3.3***
	Excluding high-tech industries	34.1	32.9	3.7***	3.6***
311	Food	1.0	3.1	0.5	0.9
312	Beverage and tobacco	0.0	2.0	0.0	−0.7
313	Textile mills	23.0	11.2	3.1***	2.0**
314	Textile product mills	1.8	0.7	0.8	0.5
315	Apparel	7.8	5.7	1.6	1.3
316	Leather	24.5	20.5	2.9***	2.6**
321	Wood products	28.3	20.6	3.6***	2.9***
322	Paper	32.0	40.7	3.6***	4.4***
323	Printing	6.8	6.2	1.5	1.4
324	Petroleum and coal products	2.1	3.4	−0.8	1.1
325	Chemicals	24.3	19.5	3.1***	2.6**
326	Plastics and rubber products	31.6	24.0	4.7***	3.7***
327	Nonmetallic minerals	32.2	17.1	3.6***	2.3**
331	Primary metals	25.8	19.3	3.0***	2.5**
332	Fabricated metal products	23.9	22.4	2.8***	2.7**
333	Machinery	54.4	36.4	5.6***	3.8***
334	Computer and electronic product	11.8	6.8	1.9*	1.4
335	Electrical equipment and appliances	27.8	15.8	3.5***	2.4**
336	Transportation equipment	6.0	6.2	1.5	1.5
337	Furniture	6.1	4.2	1.6	1.3
339	Miscellaneous manufacturing	11.6	0.3	2.0*	−0.3

Notes:
Regressions run from 1974 to 2001.
**Significant at the 10 percent level, **significant at the 5 percent level, ***significant at the 1 percent level.*
The high-technology industries include computers, communications equipment, and semiconductors (NAICS 3341, 3342, 334412-9).

the increment to the R-squared is greater for the SPC rate in only 5 out of the 21 industries. The coefficient estimates possess the wrong sign in only a couple instances.

Thus, for predicting the change in industry investment, the construction of the FRB measures has not discarded important information contained in the SPC utilization rates. If anything, the net effect of the Federal Reserve's methodology is to add information to the measures of utilization.

The exercise is repeated for the investment/capital ratio (which is more directly related to the change in the stock of capital) using the same framework:

$$IK_t = \alpha_0 + \alpha_1 \Delta I_{t-1} + \alpha_2 IK_{t-1} + \beta U^i_{t-1} + \varepsilon_t . \qquad (1.9)$$

As shown in table 5, once again both sets of utilization rates possess significant explanatory power for capital spending at the manufacturing level; they are significant at the 1 percent level. At the industry level, both sets of rates are significant in nearly all industries, and both sets rarely have the wrong sign. In all but four cases, the FRB rates increase the R-squared measures relative to regressions excluding utilization rates by more than the SPC rates. Again, the FRB rates do not appear to discard important information contained in the SPC utilization rates relevant for explaining movements in capital spending.

Predicting capacity growth

The largest difference between the FRB measures of capacity and utilization and the SPC-based measures is in their ability to explain future changes in capacity. One would expect, all else equal, that high utilization rates would be a signal to increase capacity. Table 6 displays the

Table 5 – Predicting industry-level investment/capital ratios by lagged utilization rates

NAICS	Industry	Increment to R-squared		*t*-statistic on lagged utilization	
		FRB	SPC	FRB	SPC
	Manufacturing	16.9	16.3	4.1***	4.0***
	Excluding high-tech industries	15.2	14.9	4.3***	4.2***
311	Food	2.1	2.4	1.0	1.1
312	Beverage and tobacco	0.0	0.2	−0.1	−0.6
313	Textile mills	22.6	12.3	3.5***	2.3**
314	Textile product mills	6.7	4.4	1.4	1.1
315	Apparel	4.5	5.8	1.5	1.7
316	Leather	11.2	7.5	2.9***	2.2**
321	Wood products	15.1	10.7	4.2***	3.2***
322	Paper	10.0	12.3	4.0***	4.9***
323	Printing	10.5	6.1	2.1**	1.6
324	Petroleum and coal products	0.9	0.9	−0.9	0.8
325	Chemicals	4.6	4.1	3.4***	3.2***
326	Plastics and rubber products	36.3	32.1	4.8***	4.3***
327	Nonmetallic minerals	11.9	6.6	3.6***	2.4***
331	Primary metals	7.9	6.1	3.2***	2.7***
332	Fabricated metal products	9.5	8.4	3.4***	3.1***
333	Machinery	10.0	6.5	5.1***	3.5***
334	Computer and electronic product	10.1	6.1	2.8***	2.1**
335	Electrical equipment and appliances	27.0	17.2	3.9***	2.8***
336	Transportation equipment	7.4	5.9	1.7*	1.5
337	Furniture	7.7	4.4	2.0*	1.5
339	Miscellaneous manufacturing	10.0	0.0	2.0*	0.1

Notes:
Regressions run from 1974 to 2001.
**Significant at the 10 percent level, **significant at the 5 percent level, ***significant at the 1 percent level.*
The high-technology industries include computers, communications equipment, and semiconductors (NAICS 3341, 3342, 334412-9).

results of regressing the change in FRB capacity and SPC-based implied capacity on lagged utilization rates and lagged production increases. The model estimated is:

$$\Delta c_t^i = \alpha_0 + \alpha_1 \Delta ip_{t-1} + \alpha_2 U_{t-1}^i + \varepsilon_t \qquad (1.10)$$

where *ip* is the Federal Reserve production index for the industry, and all the variables are fourth-quarter values.

In nearly every case, lagged utilization rates are significant predictors of future additions to capacity for the FRB measures. For overall manufacturing and manufacturing excluding high-tech industries, a one percentage point increase in fourth-quarter FRB utilization rates leads to about a 0.15 percent increase in capacity the following year. Conversely, the lagged SPC-based utilization rates are significant in fewer than one-half of the industries. The dramatically reduced significance in the SPC-based models likely arises from the combined effects of noisier dependent variables (the SPC-based implied capacity indexes) and less cyclically sensitive regressors (the SPC operating rates). One exception is beverage and tobacco products, where the SPC rate is significant at the 10 percent level, while the FRB rate is insignificant.

Predicting industry prices

Finally, lagged FRB and SPC utilization rates prove reasonably useful as predictors of changes in industry-specific price inflation. *Changes* in inflation are examined rather than *levels* of inflation, as Phillips curve-type models that are estimated with changes in inflation yield an estimate of the non-accelerating inflation capacity utilization (NAICU) rate.

Table 6 – Predicting the change in capacity by lagged utilization rates

NAICS	Industry	*t*-statistic on lagged utilization	
		FRB	SPC
	Manufacturing	1.7*	−0.9
	Excluding high-tech industries	3.3***	−0.0
311	Food	2.2**	1.9*
312	Beverage and tobacco	−0.5	2.1**
313	Textile mills	2.7**	1.6
314	Textile product mills	2.9***	1.8*
315	Apparel	2.2***	3.1***
316	Leather	2.4***	0.6
321	Wood products	4.1***	2.1**
322	Paper	3.1***	2.2**
323	Printing	0.8	1.7*
324	Petroleum and coal products	2.6**	2.9**
325	Chemicals	1.7*	−0.5
326	Plastics and rubber products	4.2***	0.8
327	Nonmetallic minerals	4.6***	1.5
331	Primary metals	7.7***	2.6**
332	Fabricated metal products	4.2***	1.3
333	Machinery	4.5***	0.7
334	Computer and electronic product	−1.1	−2.7**
335	Electrical equipment and appliances	3.5***	0.2
336	Transportation equipment	2.0*	1.0
337	Furniture	1.4	−0.5
339	Miscellaneous manufacturing	1.9*	−0.9

Notes:
Regressions run from 1974 to 2002.
**Significant at the 10 percent level, **significant at the 5 percent level, ***significant at the 1 percent level.*
The high-technology industries include computers, communications equipment, and semiconductors (NAICS 3341, 3342, 334412-9).

The hurdle is fairly high for utilization rates to be useful predictors of prices, as the model also includes momentum terms (lagged changes in industry price inflation), proxies for supply shocks (changes in energy price inflation), and changes in industry wage inflation. The regression model is:

$$\Delta\pi_t = \alpha_0 + \alpha_1 \Delta\pi_{t-1} + \alpha_2 \Delta\pi_{t-1}^{energy} + \alpha_3 \Delta\pi_{t-1}^{wages} + \beta U_{t-1}^{i} + \varepsilon_t \qquad (1.11)$$

where π is the rate of change of the price of industry output (therefore $\Delta\pi$ is the change in inflation), π^{energy} is the producer price index for energy, and π^{wages} is the rate of change in industry wages; all are annual averages. The data sources and methods are described in the data appendix.

As shown in table 7, lagged FRB utilization rates are significant at least at the 10 percent level in more than half of the industries, although utilization rates register the wrong sign in 6 cases (and in one, the coefficient is significant at the 10 percent level). SPC utilization rates are significant at the 10 percent level in only 4 cases, and the estimated coefficients have the wrong sign in more cases than the FRB rates. SPC rates, however, perform better in the model for nonmetallic minerals, where the FRB utilization rates are not significant. The FRB modeling procedure that combines SPC utilization rates and information on industry capital spending significantly improves the explanatory power of utilization rates in these simple price equations.

Table 7 – Predicting the change in industry price inflation by lagged utilization rates

NAICS	Industry	*t*-statistic on lagged utilization	
		FRB	SPC
	Manufacturing	2.2**	2.2**
311	Food	0.3	−0.8
312	Beverage and tobacco	2.6**	−0.6
313	Textile mills	3.0***	3.1***
314	Textile product mills	1.8*	1.5
315	Apparel	−0.3	−0.5
316	Leather	−1.8*	−0.5
321	Wood products	−0.9	−1.3
322	Paper	1.9*	1.1
323	Printing	2.7**	0.6
324	Petroleum and coal products	1.1	1.0
325	Chemicals	0.9	−0.2
326	Plastics and rubber products	2.2**	1.1
327	Nonmetallic minerals	0.2	1.7*
331	Primary metals	−0.1	−0.1
332	Fabricated metal products	2.7**	2.2**
333	Machinery	1.9*	1.3
334	Computer and electronic product	−1.0	0.8
335	Electrical equipment and appliances	2.5**	2.6**
336	Transportation equipment	−0.2	−0.1
337	Furniture	2.4**	1.4
339	Miscellaneous manufacturing	2.0*	1.3

Notes:
Regressions run from 1974 to 2002.
**Significant at the 10 percent level, **significant at the 5 percent level, ***significant at the 1 percent level.*

6. Conclusion

This paper reviewed the concepts underlying the Federal Reserve measures of capacity and capacity utilization, their history, and the methods used to construct them. The Census Bureau's Survey of Plant Capacity (SPC), the only current long-running and broadly based survey of utilization rates, was discussed in detail. The aggregate manufacturing utilization rates from the Census Bureau do not appear to be insufficiently cyclical, and therefore a capacity index derived by dividing a manufacturing production measure by the aggregate SPC utilization series does not possess what Perry (1973) called a cyclical bias. Cyclical bias had previously been shown to be a feature of survey-based rates from earlier vintages of government and industry utilization surveys; capacity tended to be "lost" in recessions and "found" quickly as industry recovered.

At the detailed industry-level, however, a cyclical bias is found in the SPC data for several industries. In contrast, Federal Reserve capacity measures, which combine survey-based information from the Census Bureau and from alternative indicators of capacity, such as measures of industry capital input, are shown not to possess a cyclical bias.

Utilization rates from both the Census Bureau and the Federal Reserve are shown to be excellent predictors of industry capital spending. The Census Bureau measures are generally less successful at predicting future capacity expansion and changes in industry price inflation, while the Federal Reserve measures perform reasonably well in both cases. In sum, while the Federal Reserve's estimation method successfully removes the cyclical bias found in the implied capacity indexes for several industries, it does so without removing from the Census measures useful information for explaining movements in industry capital spending, capacity expansion, and changes in industry price inflation. Moreover, the Federal Reserve measures typically perform better than the SPC measures in these exercises. As a result, the regression-based procedure employed by the Federal Reserve to combine the SPC-based utilization rates with other information, principally measures of industry capital, appears, on net, to add information content to the measures capacity utilization published in the very useful Survey of Plant Capacity.

References

Bacharach, M.B. (1965), "Estimating Nonnegative Matrices from Marginal Data," *International Economic Review*, Vol. 6, pp. 294–310.

Bayard, Kimberly, and Shawn Klimek (2003), "Reclassifying the Census of Manufactures from the Standard North American Classification System to the North American Classification System, 1963 to 1992," *2003 Proceedings of the American Statistical Association*, Business and Economics Statistics Section, Alexandria, VA.

Beaulieu, J. Joseph, and Joe Mattey (1998), "The Workweek of Capital and Capital Utilization in Manufacturing," *Journal of Productivity Analysis*, Vol. 10, pp. 199–223.

Bureau of the Census (1982), "Survey of Plant Capacity, 1981," Current Industrial Reports, MQ-Cl(81)-1.

Bureau of the Census (1983), "Survey of Plant Capacity, 1982," Current Industrial Reports, MQ-Cl(82)-1.

Bureau of the Census (1992), "Survey of Plant Capacity, 1990," Current Industrial Reports, MQ-Cl(90)-1.

Bureau of the Census (1997), "Survey of Plant Capacity, 1996," Current Industrial Reports, MQ-Cl(96).

Bureau of Labor Statistics (1983), "Trends in Multifactor Productivity, 1948–1981," BLS Bulletin No. 2178.

Christiano, Lawrence J. (1981), "A Survey of Measures of Capacity Utilization," International Monetary Fund Staff Papers, pp. 144–199.

Corrado, Carol A., Charles E. Gilbert, and Richard D. Raddock (1997), "Industrial Production and Capacity Utilization: Historical Revision and Recent Developments," *Federal Reserve Bulletin*, v89, pp. 67–92.

Corrado, Carol A. (2003), "Industrial Production and Capacity Utilization: The 2002 Historical and Annual Revision," *Federal Reserve Bulletin*, v83, pp. 151–176.

Corrado, Carol, and Joe Mattey (1997), "Capacity Utilization," *Journal of Economic Perspectives*, Vol. 11, No. 1 (Winter), pp. 151–167.

De Leeauw, Frank (1979), "Why Capacity Utilization Rates Differ," Measures of Capacity Utilization: Problems and Tasks, Federal Reserve Board of Governors, Staff Study 105.

Doyle, Maura P. (2000), "The 1989 Change in the Definition of Capacity: A Plant-Level Perspective," U.S. Census Bureau, Center for Economic Studies Working Paper CES-WP-00–09.

Forest, Lawrence R., Jr. (1979), "Capacity Utilization: Concepts and Measurement," Measures of Capacity Utilization: Problems and Tasks, Federal Reserve Board of Governors, Staff Study 105.

Gilbert, Charles C., and Michael F. Mohr (1996), "Capital Stock Estimates for Manufacturing Industries: Methods and Data," Board of Governors of the Federal Reserve System working paper.

Gilbert, Charles C., Norman J. Morin, and Richard D. Raddock (2000), "Industrial Production and Capacity Utilization: Recent Developments and the 1999 Annual Revision," *Federal Reserve Bulletin*, v86, pp. 188–205.

Morin, Norman J. (2003), "NAICS and the 2002 Historical Revision of Industrial Production, Capacity, and Capacity Utilization," *2003 Proceedings of the American Statistical Association*, Business and Economics Statistics Section, Alexandria, VA.

Raddock, Richard D. (1987), "Federal Reserve Estimates of Capacity and Capacity Utilization," *Industrial Capacity: Major Issues and National Security Implications*, conference volume.

Raddock, Richard D. (1990), "Recent Developments in Industrial Capacity and Utilization," *Federal Reserve Bulletin*, v76, pp. 411–435.

Raddock, Richard D. (1995), "A Revision to Industrial Production and Capacity Utilization, 1991–1995," *Federal Reserve Bulletin*, v81, pp. 16–25.

Raddock, Richard D., and Lawrence R. Forest (1976), "New Estimates of Capacity Utilization: Manufacturing and Materials," *Federal Reserve Bulletin*, v62, pp. 892–905.

Perry, George L. (1973), "Capacity in Manufacturing," Brookings Papers on Economic Activity, Vol. 3, 701–742.

Schnader, Marjorie H. (1984), "Capacity Utilization," *Handbook of Economic and Financial Measures*, Dow Jones-Irwin: Homewood IL, 74–104.

Shapiro, Matthew D. (1996), "Macroeconomic Implications of Variations in the Workweek of Capital," Brookings Papers on Economic Activity, Vol. 2, pp. 79–119.

Slanta, John (2003), "Tillé Sampling in a Nutshell," Census Bureau Memorandum, Manufacturing Programs Methodology Branch.

Stevens, John J. (2003), "Overhauling Industrial Production: The 2002 Annual Revision," *2003 Proceedings of the American Statistical Association*, Business and Economics Statistics Section, Alexandria, VA.

Abstract

In this paper, we review the history and concepts behind the Federal Reserve's measures of capacity and capacity utilization, summarize the methods used to construct the measures, and describe the principal source data for these measures – the Census Bureau's Survey of Plant Capacity. We show that the aggregate manufacturing utilization rate from the Survey of Plant Capacity does not exhibit the "cyclical bias" possessed by utilization rates from the less statistically rigorous utilization rate surveys previously used to estimate the Federal Reserve's measures. At the detailed industry level, utilization rates from the Survey of Plant Capacity for several industries do appear to possess a cyclical bias, but we demonstrate that this bias is removed in the construction of the Federal Reserve capacity measures. We further show that the Federal Reserve measures, by combining the Census survey utilization rates with other indicators of capacity, do not discard significant information contained in the Census rates. In fact, the Federal Reserve procedures add to the predictive content of the Census utilization rates in models of capital spending, capacity expansion, and changes in price inflation.

JEL codes: D24, E22, E31.

Norman Morin and John Stevens (Federal Reserve Board)
August 2004

Data appendix

Utilization rates: The Survey of Plant Capacity (SPC) from the Bureau of the Census collects utilization rate data at the 4-digit SIC (from 1974 to 1996) and 6-digit NAICS level (from 1997 on). The SPC utilization rate data on a 6-digit NAICS basis were aggregated to the 3-digit NAICS level using value-added weights from the Annual Survey of Manufacturers (ASM) and Census of Manufactures (COM). The SPC data on a 4-digit SIC basis were converted to the 6-digit NAICS level as shown in Morin (2003) using the variable share mapping from Bayard and Klimek (2003). The resulting 6-digit NAICS data were aggregated to the 3-digit NAICS level as above.

Nominal investment: ASM/COM data on capital expenditures on new equipment and structures are compiled at the 4-digit SIC through 1996. From 1997, data were compiled on total capital expenditures on equipment and on structures at the 6-digit NAICS industry levels. The SIC-based capital data on a 4-digit SIC basis were converted to the 6-digit NAICS level using the variable share mapping from Bayard and Klimek (2003).

Real investment: Real investment measures require estimating real industry-by-asset investment and aggregating these data to the industry level with asset-specific price deflators (see Mohr and Gilbert, 1996, for details). This is performed in four steps. First, US-level asset totals are taken from the BEA NIPA data. Second, industry-level investment totals are taken from the ASM/COM; Third, given the estimates of total investment by each manufacturing industry (and total US excluding manufacturing) and the total US investment in each asset category, industry-by-asset investment is estimated using the biproportional matrix balancing (or RASing) technique of Bacharach (1965); the initial estimates of the asset distribution of industry investment were taken from the roughly quinquennial Capital Flows Tables (CFT) of the BEA.[26] The industry-level real investment measures are Fisher chain-weighted aggregates of the asset-level investment flows.

Capital stocks: Industry-level net capital stocks are constructed as a Fisher index of the industry-by-asset capital stocks, where the weights are the asset-specific prices (see BLS, 1983). Industry-by-asset capital stocks are constructed using the perpetual inventory model system (PIMS) methodology (see BLS, 1983, and Mohr and Gilbert, 1996). Each asset is assigned a specific age-efficiency profile that describes the proportion of its original efficiency that remains in each period as the asset ages.[27] For a given industry, the capital stock in a particular asset category is a weighted sum of all past investment flows, where the weights are given by the age-efficiency profile.

Current-cost capital stocks: The replacement cost, in current dollars, of the net capital stock is constructed by taking the real capital stock levels for each asset category, multiplying them by the asset price deflators for that year, and summing to the industry level.

Capital input: Industry-level capital input measures estimate the potential flow of services derived from the net capital stocks in the various asset categories. They are constructed as a Tornqvist index of the industry-by-asset capital stocks where the weights are the asset-specific rental prices or user costs (see BLS, 1983). The rental price for a particular asset, $p(r + \delta - \dot{p}/p)\tau$, is the marginal product of that asset, where p is the asset price, r is a required rate of return, δ is a depreciation rate, and τ is a tax term (see BLS, 1983).

Industry wages: Industry wages are constructed by dividing the industry wage bill for production workers by production worker hours for the industry, both taken from the ASM/COM. The data were collected at the 4-digit SIC through 1996. From 1997, data were compiled at the 6-digit NAICS levels. The SIC-based data were converted to the 6-digit NAICS level using the variable share mapping from Bayard and Klimek (2003).

Industry output prices: The industry prices are derived by chain-aggregating detailed shipments deflators from the BEA gross output by industry data system. The detailed SIC-based data were classified on a NAICS basis using the shares derived by Bayard and Klimek (2003).

26 *Given row (asset investment) and column (industry investment) totals that sum to the same value; non-negativity constraints on investment; and an initial guess on the asset allocation of industry investment, the RASing procedure converges to a unique industry-by-asset investment flow. For the years a CFT exists, it is used as the initial guess for the RASing procedure; for years between CFTs, a linear interpolation of the adjacent CFTs are used; for years after the most recent CFT, the final allocation from the previous year is used as the initial guess for the current year; for years before the first CFT, the final allocation from the following year is used as the initial guess.*

27 *The age efficiency profile is based on integrating over all possible asset service lives given a stochastic mean service life and standard deviation (for asset discards) and a hyperbolic beta-decay function (for asset decay). See Mohr and Gilbert (1996) for details.*

WORKSHOP C

Productivity statistics

Chair: Jean Cordier (Banque de France)

Papers:

Revisiting recent productivity developments across OECD countries
Les Skoczylas and Bruno Tissot (Bank for International Settlements)

Labour productivity vs. total factor productivity
Ulrich Kohli (Swiss National Bank)

Measuring total factor productivity for the United Kingdom
Charlotta Groth, Maria Gutierrez-Domenech and Sylaja Srinivasan
(Bank of England)

Information technology and productivity changes in the banking industry
Luca Casolaro and Giorgio Gobbi (Bank of Italy)

Discussant's summary on Workshop C: Productivity Statistics
Hlabi Morudu (South African Reserve Bank)

Revisiting recent productivity developments across OECD countries[1]

Les Skoczylas and Bruno Tissot (Bank for International Settlements)[2]

1. A growing interest in productivity developments

Some definitions

The issue of productivity, and, in particular, its different developments observed across the main industrial countries, have raised considerable interest in recent public debates (see, for instance, Bank for International Settlements (2004)). In general, the generic term of "productivity" refers to labour productivity, defined as real output per unit of labour.

The concepts at stake are, however, slightly more complicated:

- The definition of productivity can be much wider, since labour is not the sole input used when producing one unit of GDP; for instance, Section VI below emphasises the usefulness of considering the productivity of capital and therefore total factor productivity (TFP).
- Even if restricted to the input of labour, productivity can have different meanings, depending, for instance, on the data available or the country considered. Labour productivity is traditionally calculated by dividing the level of output (in volumes) by the number of people employed ("output per person"). But the exact definition of the numerator may vary, and can be GDP, value added in the business sector, or manufacturing output … Moreover, the denominator can also be expressed as the number of hours worked ("output per hour"), being evidently the number of people employed times the number of average hours worked per person in the sector considered (whole economy, business sector, manufacturing, …).
- These definitions do matter. If average hours worked per person experience a dramatic change, following for instance the introduction of new legislation or a shift in the share of part-time jobs in the labour force, reasoning in terms of GDP per person employed or per hours worked can make a big difference. This might even be complicated by the fact that any productivity effect from a change in working hours might depend on its cause.[3]
- While productivity is constructed as a ratio, comparisons across countries only rarely deal with its levels. In fact, when they refer to "productivity", most observers focus on its growth, i.e. on the changes in labour productivity that are due to movements in real output or labour input or both.

In the present paper, we stick to the most common definitions and define *productivity* (resp *productivity per hour*) as the level of output per person (resp per hour worked) and *productivity gain* (resp *productivity gain per hour*) as the change in this ratio. *Total factor productivity* (*TFP*, resp *TFP per hour*) refers to the combined productivity of capital and labour, the latter being expressed in terms of number of people employed (resp of hours worked). Finally, we mainly consider the *output of the business sector* for the practical reasons detailed below, with the exception of Section II since the most widely watched international comparisons of productivity levels refer to the whole economy.

1 *Paper presented at the Irving Fisher Committee conference on "Central bank issues regarding national and financial accounts", Basel 9–10 September 2004.*

2 *The authors are members of the Bank for International Settlements, Basel, Switzerland. They have greatly benefited from the help of Philippe Hainaut and the extended comments of Palle Andersen as well as from Claudio Borio, Andy Filardo, Marjorie Santos, Philip Turner and Bill White. The usual disclaimer nonetheless applies: views expressed in this paper are those of the authors, and remaining errors theirs alone.*

3 *By construction, total output per person tends to fall if the share of part-time workers rises while output per hour should remain unchanged. But output per hour can also be negatively affected if the share of low-skilled workers (often the case in part-time jobs) in employment increases.*

Productivity is a key issue at the country level …

One obvious reason for the widespread interest in productivity is that it is actually hard to overstate its macroeconomic significance. At the country level, productivity plays a major role in determining living standards as it is a key variable shaping potential supply (see, for instance, Oesterreichische Nationalbank (2004)). To put it simply, higher rates of productivity growth will sustain larger real changes in profits and wages in the long run – GDP will double in around 25 years if productivity rises by 3% per year, and in 70 years if the annual rate is only 1%.

For policy-makers, estimates of productivity growth and thus potential output are a key element in ascertaining the state of the output gap. The derived prospects for capacity and inflationary pressures have large implications for interest rates. In the longer-run, the potential growth rate is also a key variable driving real equilibrium interest rates, at least at the level of theory – see for instance Burda and Wyplosz (2001) for a brief overview.[4] Turning to the fiscal side, the level of potential output is of particular importance in judging the stance of fiscal policy as well as the sustainability of budget positions. In particular, a key issue regarding the fiscal consequences of ageing are long-term developments in potential growth, because they determine the pool of resources (i.e. national incomes) from which future social benefits will be drawn.

Developments in productivity have also significant implications for financial markets. For instance, if trend productivity is driving potential growth higher, this should sustain investors' expectations of earnings and in turn support share prices. By contrast, if estimates of productivity growth and thus long-run output are overstated (e.g. if the country is experiencing a steady but undetected slowdown), expectations of future earnings' growth are likely to be disappointed, which, eventually, may trigger a downward correction in equities.

… as well as at the global level

At the global level, different productivity growth rates condition relative economic performance. They appear to have been a key factor behind the diverging GDP growth rates observed in the main industrial countries in the past few decades, though other structural factors such as the dynamism of the labour force have also been in play. In turn, diverging developments in domestic spending can have a meaningful impact on external imbalances, as observed in the past few years.

The implications of productivity for relative expected returns are also a powerful driver of international capital flows, with important consequences for global financial markets. For instance, the reported improvement in US productivity gains during 1995–2000 raised the expected rate of return of US capital, relative to other countries. As a result, the United States became particularly attractive for foreign investors during that period, allowing the financing of a large current account deficit without apparent pressure on exchange rates and interest rates – in fact, the dollar appreciated sharply over this period (regarding this discussion, see Bailey et al. (2001)).

Finally, the degree to which higher productivity levels or gains in some countries can be replicated in others influences global growth prospects and have important policy implications. For instance, there has been a growing belief in recent years that the use of information technology has positively affected the acceleration of US productivity growth, prompting several countries to promote the acquisition of IT equipment.[5] Such policies might however not yield the expected results should the US performance be attributed mainly to structural factors (such as a rise in the rate of technological progress) rather than to the sole impact of IT capital deepening. In this event, other policy actions might be required in those countries willing to replicate the US experience, for instance in order to improve the functioning of their product and labour markets.

The basic picture: productivity levels are the highest in the United States and appear to have grown more rapidly than in other OECD countries in recent years

The first thing to note is the large differences in GDP per capita across industrial countries (PPS-adjusted; Table 1). The *United States* tops the league, with a lead of around 30% over other main industrial economies. Compared to this measure, the *euro area* is better placed in terms of output per worker, with the gap with the United States being significantly reduced. The main

4 *And also ECB (2004a) for a discussion on the relationship between productivity and natural real interest rates.*

5 *For a cross-country overview of productivity developments during the IT accumulation period of the 1990s see for instance Gust and Marquez (2000).*

reason is that employment rates are lower, reflecting social choices (e.g. early retirement) but also higher structural unemployment. When output is measured per hour worked, the euro area disadvantage narrows further to around 10% relative to the United States, as euro area employees spend less time working. In other words, the euro area lags the United States in output per capita partly because its citizens are slightly less productive, but chiefly because of structural rigidities and more leisure (Blanchard (2004)). The relative position of *Japan* is still less favourable in terms of labour productivity.

Another important feature is that, for most of the postwar period, Europe and Japan steadily raised their productivity levels towards the US level. By the early 1990s, however, this convergence seemed to have halted, and subsequently might even have reversed (for an assessment of the recent European economic fortune, see Denis et al. (2004)). As a result, the US productivity performance has recently improved again in relative terms: the US advance over the euro area as a whole (but substantial differences exist among countries within the zone) and Japan has slightly increased since 1995, in terms both of GDP per person employed and of GDP per hours worked. These differences have been particularly evident since the latest downturn, with US productivity rising markedly in 2002 and 2003.

Table 1 – Productivity levels[1]

United States = 100

	GDP per capita		Labour productivity			
			per person employed		per hour worked	
	1995	2003	1995	2003	1995	2003
United States	100	100	100	100	100	100
Euro area	72	70	84	77	95	89
Of which: Germany	77	70	81	73	97	90
France	75	74	93	88	108	107
Italy	75	70	93	80	104	88
Spain	57	62	78	73	83	75
Netherlands	78	78	80	73	107	98
Belgium	78	76	98	92	111	106
Austria	84	79	81	74	96	87
Greece	47	52	64	70	61	64
Portugal	47	49	47	49	47	51
Finland	69	72	81	76	87	80
Ireland	64	87	86	92	86	99
Japan	81	74	72	69	71	69
United Kingdom	72	77	76	79	81	83
Canada	80	87	89	86	92	86
Sweden	77	75	79	74	89	85
Denmark	81	80	76	75	92	89
Norway	86	96	84	92	110	123
Iceland	81	76	83	74	84	73

[1] *Whole economy; calculations made using purchasing power standards (PPS).*

Source: European Commission, Eurostat (2004a&b).

2. Comparisons across countries are difficult

Instead of being genuine, the apparent reversal of trends noted above might however just reflect measurement uncertainties.

Comparisons of productivity levels are notoriously imprecise

A widely-shared view is that *levels* of productivity are not internationally comparable. The reason why is that serious measurement problems are related to both the numerator (output) and the denominator (the labour factor) used in estimating productivity (for the whole issue of measuring productivity, see Maddison and van Ark (1994) as well as OECD (2001a)).

Indeed, international comparisons of the level of output are very sensitive to the exchange rates used (Magnien et al. (2002)). If current exchange rates are considered, then a country with a depreciating exchange rate will see its productivity decline, all other things being equal. One solution, adopted in the Table 1 presented above, is to deal with fixed exchange rates, for instance using purchasing power parity indexes (Schreyer and Koechlin (2002)); but significant difficulties surround these calculations (Richardson (2001)).

A second difficulty is that economic concepts still differ across countries, despite ongoing progress in harmonising local practices. Informal activity is not taken into consideration homogeneously across countries and total GDP numbers might include a larger part of it in one country than in another (Blades and Roberts (2002)). Similar difficulties exist with respect to employment, which can be defined in different ways, depending on a country or a sector. Furthermore, data on hours worked are notoriously more difficult to obtain and to compare than data on persons employed (OECD 2004a).

A third problem is that even if economic concepts were identically defined across countries, the way they are statistically measured might still differ because of inherent uncertainty. Steel production is easy enough to measure, but the real value of lawyer services for instance is harder to pin down. In addition, measurement procedures cannot be fixed over time. Statistics do change, because of the introduction of new techniques (hedonic prices), the need to reflect a moving reality (declining importance of the mining sector in total output), or the limited availability of some data (leading to the continuous implementation/improvement of statistical surveys). These uncertainties might even have increased in recent years since it could have become harder to collect information and to accurately measure economic activity within and outside national boundaries (developments of multinational companies, increased trade and financial integration, …; see for instance in the US case Hatzius (2004)).

Certainly, some studies have tried to resolve this issue by confining comparisons to specific sectors,[6] that are considered to be easy to measure and thus less prone to cross-country discrepancies. A widely held view is that such data in manufacturing are fairly reliable while measuring output in the service sectors is more difficult – it is less tangible than that of the traditional goods industries. Nonetheless, sectoral comparisons also have shortcomings. For instance, the focus on manufacturing is misleading since this sector differs in size across countries and represents only a minor (and declining) part of today's economies. In addition, empirical evidence suggests that even in the sole manufacturing sector, productivity levels might not be comparable and can in particular depend of the choice of the base year (Sorensen and Schjerning (2003)). Finally, measurement problems could be more serious at the sectoral level than at the macro level. The reason why is that some measurement errors at the sectoral levels could "wash out" through aggregation, since the output from some sectors is used elsewhere as inputs.[7]

Table 2 illustrates the sheer size of these uncertainties, by showing that various international comparisons of productivity levels display significant differences. Indeed, alternative estimates would change the respective ranking of industrial countries. For instance, GDP per hour would be lower (by around 4%) in France than in the United States according to one estimate, and significantly higher (by 6%) according to another.

Comparing productivity changes is also misleading

The reasons detailed above explain why most international comparisons have focused on productivity changes. The basic assumption is that even if the measured levels of total output and/or of labour inputs differ between two countries, changes in these levels are likely to be more comparable. In addition, comparing growth rates (in volumes) does note require using a common exchange rate.

6 *By using industry-specific conversion factors to calculate productivity levels. See Pilat (1996) and the work conducted at the Groningen Growth and Development Centre, in particular van Ark (1993). Denis et al. (2004) have in addition pinpointed the industrial sectors that appear to have driven the EU-US productivity differentials over recent decades. Some research is also conducted on firm-specific data (O'Mahony and van Ark (2003)).*

7 *Schreyer (2001) observed that, for instance, the impact of using different set of ICT deflators is likely to be small when looking at aggregate measures of GDP volume growth but much higher when looking at disaggregated measures of outputs, inputs and productivity.*

Table 2 – Levels of labour productivity, total economy, 1999 (United Kingdom=100): ranges for alternative estimates

	GDP per worker	GDP per hour
United States	141–145	118–126
Japan	93–107	88–93
France	114–119	113–133
Germany	105–107	107–116
Italy	117–130	123–132
United Kingdom	100	100
Canada	113–118	99–114

Sources: Drew et al. (2001).

However, even productivity changes are not free of measurement problems. There are many types of reasons (see, for an extended review, O'Mahony and van Ark (2003)):

- It is well known that substantial uncertainty surrounds the measurement of employment growth, because of difficulties in tracking new forms of jobs or newly created firms. In several countries (United States, Canada and Switzerland) concurrent surveys (payroll survey versus household survey) have presented clearly diverging pictures of job creations in recent years. Similar difficulties may surround measures of hours worked.[8] A recent example of the uncertainty in measuring employment has reflected the expansion in "mini-jobs" in Germany over the past few years (subsidised low-paid jobs free of some social security charges and taxes).
- A second difficulty is linked to the uncertainties already mentioned regarding output levels. Different methods of calculating value added in some sectors can influence their weight in GDP and thus their contribution to output growth (regarding the important issue of the services sector, see Wölfl (2003)). The fact that trade, where value added might be harder to measure, has been a major contributor to the rise in US productivity in recent years is one example of this (Table 3). Another and related issue is the measurement of spending on software. These expenditures, which have grown rapidly in the past few years, have been treated as investment (thus positively contributing to GDP growth, in contrast to intermediate consumption expenditures) to a much larger extent in the United States than in the other main industrial countries (Ahmad et al. (2003)).
- A third set of problems is related to the measurement of output deflators. It is well known that correctly measuring the price of one service may be particularly difficult. But finding sound measures of real output or reliable deflators is challenging in several large sectors of the economy. Moreover, as emphasised by Griliches (1994), the share in GDP of "reasonably measurable" sectors, which include agriculture and manufacturing activities, tends to decline over time as services expand. Another widely noted problem is the use of hedonic price indices, which allow better account to be taken of quality improvements, especially, but not only, in IT products (see Schreyer (2001) for the issue of ICT deflators). Roughly spoken, real volume growth is much lower when using traditional deflators than hedonic price indexes. For instance, the price of a laptop can be estimated to have risen by 10% using traditional statistical techniques (i.e. by measuring the price of the "average" laptop sold in a store) but declined by 10% if its quality (speed, memory …) has improved by 20%.[9] The consequence is that applying quality-adjusted deflators can affect aggregate GDP growth, depending on the size of the sector considered, and often lead to higher productivity growth rates than previously assumed.[10]
- A final issue regarding how to deflate nominal growth rates is the use of chained- rather than fixed-indices. This last difficulty has been recently highlighted in Japan. The adoption in 2004 of a chain-type index for calculating output led to a sharp revision in real GDP growth, by more than one percentage point downwards for the 2003 fiscal year.

8 *See Eldridge et al. (2004) for an analysis of different estimates in average hours worked in the United States.*

9 *See Congressional Budget Office (2002) for a short presentation of these techniques and the surrounding issues in measuring US productivity.*

10 *These problems are compounded by the fact that measurements of productivity in IT-producing and in IT-using industries differ across countries (Pilat and Wölfl (2004)).*

Table 3 – US labour productivity gains, by sector

	Employment share[1] 1987–2002	Productivity gains[2]	
		1988–95	1996–2002
Private industries	85	1.4	2.4
Agriculture and mining	2	0.7	–0.5
Construction	6	–0.8	–0.3
Manufacturing	14	2.9	4.4
Durable goods	9	3.9	5.5
Non-durable goods	5	1.7	2.9
Transportation and utilities	4	1.7	1.1
Wholesale trade	5	2.9	3.5
Retail trade	11	2.8	4.1
Finance, insurance and real estate	6	2.3	2.9
Services	37	0.2	2.1
Government	15	0.7	1.2

Note: Break in series in 2000 due to a new breakdown in industry branches.
[1]*As a percentage of persons engaged in domestic production.*
[2]*BIS calculations using quantity indices for gross output and hours worked by full-time and part-time employees, by industry; annual rates in per cent.*

Source: US Bureau of Economic Analysis.

All in all, several estimates suggest that measured GDP growth (and thus productivity gains) in Europe would be higher, perhaps by almost half a percentage point annually, if statistical methods were more similar to those used in the United States.[11]

Regarding the measurement of employment growth, some estimates show that this could influence calculations of labour productivity gains by up to another half a percentage point annually in some countries (Ahmad et al. (2003)).

These difficulties should not be overlooked. Cross-country differences in measured yearly productivity gains are not that large and are often comparable to the degree of uncertainty reported above, or even lower. Hence, apparent discrepancies in productivity gains would perhaps narrow sharply or even be reversed should same statistical methodologies be applied across national borders.

The basic conclusion is thus that international comparisons of productivity gains might be misleading. It should be noted, however, that many observers – especially in the financial markets – do not appear to be very concerned by these difficulties.[12]

3. The issue of interest: have trend-productivity gains changed?

Changes in productivity gains …

If statistical measurements are time-consistent, the most severe distortions in the estimation of growth in both output and labour should remain relatively unchanged over time. Hence, they would tend to disappear when looking at changes in productivity gains (i.e. the second derivative). For instance, the importance of a rapidly-growing sector might be overstated in one country (when compared to the situation in other countries) because of the statistical methods used, but these methods are likely to remain unchanged over time. As a result, productivity gains for the whole economy would be persistently overestimated but this overestimation would remain constant. This means that measures of changes in the rates of productivity growth (i.e. on the *acceleration* of productivity) should be more comparable across countries.

In this context, we used the OECD Economic Outlook database (OECD (2004b)) with the view to work on a relatively homogeneous statistical source (compared to using national data for

11 *See Ahmad et al. (2003) for a detailed discussion of these effects. Other estimates give the same order of magnitude. For instance, Sakellaris and Vijselaar (2004) find that quality-adjusted output would grow almost 0.5 percentage point faster in the euro area. Research at the German central bank (Deutsche Bundesbank (2001); Scheuer (2001)) leads to the same kind of estimate – the difference in growth between Germany and the United States due to measurement issues was around 0.4 percentage point annually in the second half of the nineties, close to the bias calculated for the United Kingdom at the Bank of England (Wadhwani (2000)) and for France at the statistical office (Lequiller (2001)).*

12 *For a reflection of mainstream markets' views on relative productivity performances, see for instance Levy (2003). From the same community, an opposite and less widely-shared view is given by Daly (2004).*

each country). We also restricted comparisons to the business sector in order to avoid the special difficulties involved in measuring output in the government sector and non-market production.[13]

These data are presented in Table 4. They broadly show that labour productivity gains in the *OECD area* have decreased over the past four decades, in terms of both output per employee and output per hour worked. This decline was shared by almost all economies, though some countries (especially the *United States*) managed to reverse this trend since around the mid-1990s. Turning to the productivity of capital, a global feature is that it has declined over the past four decades, though less rapidly over time: it has almost stabilised on average in the OECD and has even been rising in the most recent decades in several countries (here also, especially in the United States).

… should be more comparable across countries

There are two possible caveats when comparing changes in productivity gains across countries.

The first is that methodological changes are implemented from time to time and this could lead to sudden changes in measured productivity gains: say, for instance if the size of a rapidly-growing sector is suddenly revised upwards. However, these sorts of statistical revisions are

Table 4 – Productivity in the business sector

Average annual percentage changes[1]

	Output per person				Output per hour worked				Output per capital			
	1966–1975	1976–1985	1986–1995	1996–2004	1966–1975	1976–1985	1986–1995	1996–2004	1966–1975	1976–1985	1986–1995	1996–2004
Australia	2.6	2.0	1.1	2.1	1.6	2.4	1.0	2.3	−1.0	−1.1	−0.1	0.2
Austria	5.0	2.8	2.5	1.8	–	–	–	–	−2.2	−2.3	−1.3	−1.7
Belgium	3.7	2.9	1.8	1.3	4.8	3.4	2.3	2.0	−0.0	−0.8	−0.5	−0.6
Canada	2.1	1.0	1.0	1.5	2.9	1.4	1.1	1.4	−0.2	−0.5	0.1	1.0
Denmark	2.2	1.7	1.8	2.0	3.5	3.0	2.1	2.3	−3.2	−1.1	−1.3	−1.3
Finland	4.7	3.0	3.6	2.3	5.8	3.5	3.7	2.6	…	−0.3	−0.9	2.9
France	4.4	2.7	2.1	1.2	4.5	3.7	2.6	2.0	−0.8	−3.1	−1.8	−0.5
Germany	3.8	2.0	2.0	1.0	5.1	2.6	2.7	1.6	−1.7	−0.6	0.1	−0.2
Iceland	3.7	2.4	1.2	2.6	5.2	3.2	1.3	2.6	−0.4	0.4	0.1	0.0
Ireland	5.5	3.8	3.7	3.6	6.1	4.6	4.1	4.7	3.4	0.4	3.2	3.2
Italy	5.2	2.7	2.4	0.5	5.9	3.4	2.6	0.8	−0.0	−0.2	−0.7	−1.4
Japan	7.1	2.7	2.2	1.7	5.4	2.8	3.2	2.1	−3.6	–2.2	−2.4	−2.0
Netherlands	4.0	2.1	1.4	0.9	4.1	2.8	3.3	1.4	0.6	−0.0	0.6	−0.4
New Zealand	1.0	0.7	0.9	1.3	1.8	0.9	0.9	1.5	−1.9	−0.8	−0.4	0.1
Norway	3.7	2.1	1.8	2.4	5.1	3.4	2.2	2.9	1.6	−0.2	0.8	1.4
Spain	5.4	3.3	1.6	0.7	5.2	4.3	1.8	0.7	−0.6	−2.6	−1.3	−0.5
Sweden	3.2	1.5	2.7	2.2	4.6	1.9	2.1	2.7	−1.3	−1.7	−0.7	0.2
Switzerland	2.2	0.8	−0.1	0.7	1.9	1.6	0.2	1.0	−2.4	−0.9	−1.4	−1.1
United Kingdom	3.3	2.4	1.6	1.6	3.3	3.1	1.7	1.8	−0.4	0.5	−0.0	−1.0
United States	1.6	1.2	1.3	2.6	2.5	1.3	1.2	2.8	−1.0	0.2	0.8	0.8
Euro area[2]	3.4	2.5	2.0	0.9	4.5	3.3	2.6	1.4	−2.1	−1.2	−0.7	−0.6
OECD ex. US[3]	3.3	2.4	1.9	1.3	4.4	2.9	2.5	1.7	−2.2	−1.1	−0.9	−0.9
OECD[4]	2.8	2.0	1.7	1.9	3.7	2.3	2.0	2.1	−1.5	−0.5	−0.2	−0.2

[1]*Cross-country comparisons for 1965–1975 might be misleading given that data for some countries were not available for the entire period (see annex A).*

[2]*Weighted average of Belgium, France, Germany, Italy, Netherlands and Spain, based on 2000 GDP and PPP exchange rates.*

[3]*Weighted average of Australia, Belgium, Canada, France, Germany, Italy, Japan, Netherlands, Spain, Sweden, United Kingdom.*

[4]*Weighted average of the group defined in footnote 3 plus the United States.*

Sources: OECD; national data; BIS calculations; –: data not available.

13 *Difficulties that might severely influence the acceleration in measured productivity gains. For instance, Nordhaus (2002) found that the rebound in US productivity growth from the 1978–95 period to the 1996–2000 period was between 1.04 and 1.61 percentage points depending on the definitions of output used (i.e. an uncertainty of roughly ½ percentage point).*

usually time-consistent, the general practice being to retropolate the data – though this retropolation might not be fully possible if insufficient information is available for the past years.

The second and perhaps most important difficulty is that discrepancies in measuring productivity gains might well widen over time. Attention has mainly focussed on possible sector-related biases. This could appear, for instance, if a sector is characterised worldwide by a growing size (relative to other sectors, in all countries) and, solely because of measurement issues, higher productivity gains in a specific country (relatively to the other countries[14]). The increasing importance of the IT industry and of spending on software over the past decade could be good examples of these difficulties.

This second difficulty should not be overstated as it is unlikely to affect cross-country comparisons that much. First, divergences among countries and sectors do not only reflect measurement differences and are often "real", i.e. due to the intrinsic economic performance of the countries considered – for instance any specialisation in a rapidly growing sector.[15] Second, even if there might be significant statistical biases in the way the economic structures are measured, they would have to change rapidly in order to have a meaningful influence on the change in the relative growth rates of GDP. For instance, *annual GDP growth rates* in Europe and the United States might differ by up to half a percentage point because of statistical biases, as noted above. This spread in favour of the United States has perhaps increased over time, but by less than its present size – i.e. by less than half of a percentage point, by construction. Hence, differences in *GDP growth rates* are unlikely to have changed sharply over time. Indeed, several estimates suggest that the contribution of statistical discrepancies to measured differentials in annual productivity growth rates have represented much less than half a percentage point,[16] i.e. well below the relative changes in trend-productivity gains observed across the main industrial economies over the past few decades (see below).

Measuring structural changes in productivity gains

The developments above suggest that international comparisons should focus on whether productivity gains in a specific economy have, over time, improved or not in relative terms (i.e. compared to other countries). This requires comparing trend-developments in productivity gains between countries and can be done using a various set of statistical techniques. In this respect, several studies have recently tried to determine whether trend productivity accelerated in the United States in the mid-1990s – see, for instance, Filardo and Cooper (1999) for the use of various methods to correct for cyclical influences in the United States, and Maury and Pluyaud (2004) for the application of the Bai-Perron method for several industrial countries.

The approach retained in this paper is described in Box 1 and is relatively simple. Basically, one tries to draw one or several lines through the actual productivity series, thereby producing a stylised representation of the trend, which is taken to be structural productivity (see Bodier et al. (2001), as well as Doisy (2001) for an application for measuring potential GDP growth in France). The main advantage is that countries' estimates are produced in a transparent and homogeneous way, allowing a direct comparison of countries' patterns in trend-productivity gains. A second advantage is the ability to measure the influence played by the state of the business cycle (see below).

14 *Or, alternatively, if the considered sector enjoys higher productivity gains worldwide (relative to other sectors, in all countries) and, because of measurement issues, a relative size that is growing more rapidly in a specific country (compared to the situation in other countries). Another but less likely possibility of statistical distortion would be if, because of statistical measurement, productivity is* accelerating *more rapidly in a sector of a given economy than in the same sector in the other countries. But this last case would signify that productivity gains would be overestimated, that this overestimation would rise over time, and that the impact of that would be sufficient to affect the economic performance of the country considered, which is doubtful. Indeed, looking at US sectoral data, Comin (2003) found that output price measurement "is not a key element (…) in the time series evolution of productivity growth".*

15 *For instance, it has been estimated that the labour productivity gap between the Canada and US economies has widened over the last two decades mainly because of "real" differences, i.e. differences in industrial structure (IMF (2004)).*

16 *For instance, the estimated bias by F W Vijselaar (2003) for euro area GDP growth was roughly the same in the period 1982–1990 and in 1991–2000, implying that this had no effect on changes in productivity gains. The ECB (2004b) stated that the continued decline in euro area productivity growth over the past decades, compared to the United States, is "a feature that results independently of the measure of labour input used (…) and of the economic aggregate chosen." Wadhwani (2000), in contrast, estimates that the understatement of actual output growth in the United Kingdom (compared to the United States) has been rising over time, from 0.10 percentage point (pp) in 1979–89 to almost 0.4 pp in 1994–98, i.e. an acceleration of around 0.3 pp.*

Box 1 – A general framework for testing for time breaks in productivity gains

Measuring trend productivity

The formal approach is relatively straightforward. Let's note P the productivity in levels, GDP total output in real terms and L the labour input (being either total hours worked or the number of persons employed). We thus have:

$$P = GDP / L \quad (1)$$

It is possible to regress $log(P)$ on a time trend:

$$\log P = a\,T + b + u \quad (2)$$

where T is the time trend (equal to 1 at the beginning of the sample period, T is growing by a constant unit of 1 every year, for annual data; hence $T = N$ at the end of the sample of N years); a and b are the parameters; and u is the residual.

We can define trend productivity P^* by:

$$\log P^* = a^*\,T + b^* \quad (3)$$

where a^* and b^* are the estimated parameters of (2).

The differentiation of (3) yields directly p^*, the yearly change in trend (or structural) productivity defined as:

$$p^* = (P^*_{t+1} - P^*_t) / P^*_t$$

since $\Delta P^* / P^*_t$ is small, $p^* = \Delta P^* / P^*_t \approx \log(1 + \Delta P^* / P^*_t) = \log(P^*_{t+1} / P^*_t) = \log P^*_{t+1} - \log P^*_t = \Delta(\log P^*)$

Hence $p^* \approx a^*\,\Delta T$ and, since $\Delta T = 1$ by definition,

$$p^* \approx a^* \quad (4)$$

We thus obtain directly a^*, the yearly rate of increase in trend-productivity over the sample period.

Allowing for temporal breaks

One can reestimate (2) by allowing the possibility to retain different time trends over the sample period. We thus have:

$$\log P = \sum_{i=1}^{i=I} a^i T^i + b + u \quad (5)$$

where I is the number of time trends in the estimation and (I−1) the number of breaks; T^i is the i[th] time trend (equal to 0 before the year y_i, 1 for y_i, T^i is growing by 1 each year following y_i); by construction y_1 is the first year of the sample; $(a^i)_{i=1}^{i=I}$ and b are the parameters; and u is the residual.

This estimation leads to a new way of calculating trend-productivity compared to (4):

$$P^* \approx \sum_{i=1}^{i=I} a^{*i} I_{y \geq y_i} \quad (6)$$

where $I_{y \geq y_i} = 1$ if $y \geq y_i$ and $I_{y \geq y_i} = 0$ if $y < y_i$;

and $(a^{*i})_{i=1}^{i=I}$ and b^* are the estimated parameters of (5).

If, for instance, I=1, than the equation (6) is the same than (2): there is only one time trend and no breaks, $I_{y \geq y_1} = 1$ for each year y and $p^* = a^{*1}$.

If I=2, than there is one time break and $I_{y \geq y_2} = 1$ only for the years following the break-year y_2. Thus $p^* = a^{*1}$ for all the years preceding y_2 and $p^* = a^{*1} + a^{*2}$ for the year y_2 and all the following years. In this case, trend productivity gains would have changed by a^{*2} after the year y_2, i.e. from the period $[y_1\, y_{1+1} \ldots y_{2-1}]$ to the period $[y_2\, y_{2+1} \ldots y_{N-1}\, y_N]$.

The break-years y_i are estimated imposing three conditions:

(i) all the parameters of the equation (5) have to be statistically significant;
(ii) for a given value of the number I of time trends, the quality of the estimation (as summarised by the F^I statistic of equation (5)) is the highest among all the possible combinations of any other break years.
(iii) A minimum period of 6 years must separate two different break-years, the implicit idea being to keep trend productivity gains constant during a sufficient period of time (roughly comparable to the length of the business cycle in industrial countries[17]).

In practice, a step-by-step approach has been adopted. Equation (5) was first estimated for $I=1$ (only one time trend is in the equation); for $I>1$ and an existing combination of $(I-1)$ break years ($y_2 \ldots y_I$), (5) was reestimated for I+1, leading to I break years ($y'_2 \ldots y'_i \ldots y'_{I+1}$) and this new estimation was kept if:

(i) all the new parameters are significant;
(ii) and if $F^{I+1} > F^I$.

Correcting for cyclicality

In order to correct for the cyclical component of productivity growth, a first solution is to improve the regression (5) by adding a variable that can capture the cycle.

We thus have the following equation:

$$\log P = \sum_{i=1}^{i=I} a^i T^i + b + c\ CY + u \qquad (7)$$

where CY is the indicator of the cycle and c another parameter.

CY has to display no temporal trend (in practice its mean was equal to zero over the sample period as it was normalised), so that developments in trend-productivity continue to be given by the expression (6). The way (7) is specified implies that it is the change in CY, and not its level, that has an influence on productivity gains – i.e. on $\Delta(\log P)$.

17 *The basic aim is to try to obtain a sequence of waves in labour productivity, echoing recent views on the US situation (see, for instance, Meyer (2001)). Interestingly, we were not able to find a single break over the past forty years or so in some countries. Estimates with a shorter minimum period between break years (for instance two years) gave somewhat more breaks in the estimation for some countries but produced relatively similar results in terms of trend-productivity gains' patterns.*

4. Productivity gains are influenced by the state of the business cycle

Cyclical influence may distort cross-country comparisons

The approach detailed above might not be sufficient for determining trend-productivity gains. Because of the lags with which labour adjusts to changes in output, it is indeed a well known feature that labour productivity moves procyclically and this might distort international comparisons.

Certainly, productivity gains are not strictly speaking synchronised with the business cycle, since they tend to be the highest during periods when output is accelerating – i.e. when the change in the output gap is the largest, not when the economy is peaking. For instance, firms that have hoarded labour during a recession can raise output without much increase in measured employment once demand picks up: productivity growth therefore surges in the upswing phase before slowing as the labour market begins to recover. These cyclical influences might be particularly important in certain years – implying that they have to be taken into consideration when searching for temporal breaks in productivity gains. For instance, capacity utilisation fell sharply in 2001 in the United States (by around 1½ points of standard deviation), contributing to a decline in labour productivity gains of around 1 percentage point (and more than 3 points for capital productivity gains). But the underlying trend remained strong and measured productivity actually recovered sharply in 2002 and 2003.[18]

18 *See Oliner and Sichel (2002) for an assessment of whether the 2001 IT-led economic downturn changed the US underlying productivity performance; and Greenspan (2002) for the implications of the latest US cyclical downturn.*

The extent to which changes in productivity are cyclically influenced is of particular importance when doing cross-country comparisons. This is because national business cycles are far from being synchronised. In the early 1990s, for instance, the output decline in the major English-speaking countries preceded that in continental Europe and Japan by two years. This desynchronisation was mainly attributable to country-specific disturbances and events, notably the German reunification and the end of the asset price bubble in Japan. The latter also led to protracted balance sheet problems which weighed on activity in Japan throughout the 1990s. The resulting decoupling of Japan from the global business cycle was reinforced by the Asian crisis in 1997–98. The latest downturn in 2001–2002 was somewhat more synchronised, but not entirely. In particular, the US economy started to recover earlier than both the euro area and Japan.

Nevertheless, international comparisons are often made without taking proper account of the influence of business cycles. For instance, observers have focused on comparing national developments since the mid-1990s – in reference to the reported improvement in US productivity during this period. However, cyclical developments might have significantly biased these comparisons: according to the OECD, the output gap in both the euro area and Japan was positive only during one or two years from 1995 to 2000, compared to four years in the United States (and it was even almost always positive during this period in countries such as Australia, the Netherlands, and some Nordic countries). Another example is the current recovery observed since 2002. The United States has experienced a faster growth in demand and this may have raised the "cyclical component" of US productivity gains relative to other countries. From 2002 to 2004, indeed, the US negative output gap has been reduced by 1.7 percentage points while it actually increased slightly for the rest of the OECD area.

Several ways to deal with cyclical fluctuations

We have retained three different ways to disentangle cyclical fluctuations from trend, though a wider range of methods exist (see, for instance, Gordon (2003)).[19] The first and preferred approach we used was to directly correct productivity levels by using an indicator of the cycle (cf Box 1). There are several variables than can be used for this purpose and they don't necessarily move in tandem (see Steindel (2004) for a recent discussion on the divergence between manufacturing production and goods output in the United States). Duval (2000) for instance looked at the degree of slack in the labour market, retaining the job vacancies ratio for characterising the cyclical position of the US economy. The present study has favoured using capacity utilisation in manufacturing (or in the industry when not available), which measures how much productive capacity is in use. These data appear to be relatively homogeneous among OECD countries and are available over a significant period of time.[20] In addition, the shape of the business cycle in manufacturing seems to reflect correctly developments in the whole economy – despite the relative limited size of manufacturing, its contribution to the variance of total output is quite important.[21] Finally, there seems to be no other cyclical indicators easily available and with a sufficient degree of homogeneity across industrial countries.

Needless to say, this approach has some drawbacks. First, and as already noted, capacity utilisation is only measured in the industry, which represents only a small part of today's economies. Furthermore, the way these data are elaborated (type of question, period under review) can differ, depending on each country. In addition, the series are not available for all industrial economies during the same period of time and are even missing in some (rare) countries. For instance, we could estimate the cyclical component of productivity for the

19 *An alternative and simple method would be to compare productivity during whole cycles. Trend productivity gains would be measured as the rate of actual productivity growth between comparable points in the cycle (e.g. peaks or troughs). A similar approach is to compare productivity growth around business cycle peaks (Council of Economic Advisers (2002)). However, the choice of 'extreme' points of the cycle is arbitrary and estimations for the current cycle are by nature difficult. These methods are not well suited for international comparisons, given that business cycles are not the same across countries and that the way they are detected might differ. Finally, and perhaps most importantly, these approaches basically assume that no structural change in productivity can occur during a business cycle, an assumption that seems to be too restrictive. For instance, a general view is that US trend productivity accelerated in the mid-1990s, i.e. about in the middle of the 1991–2001 cycle.*

20 *This is not to say that the degree of capacity utilisation is free of measurement problems. Different surveys can lead to different estimates as observed by Wadhwani (2001) – though these differences are more notable in terms of levels than in terms of changes. In particular, the data used for Australia displayed a significant degree of uncertainty, that may have affected the validity of the results presented in this paper.*

21 *This is the reason why OECD estimates of leading indicators rely on industrial production as the "reference series" for the business cycle.*

whole 1960–2003 period in the United States but only from 1968 onward in Italy. Hence, estimations had to be conducted on different periods and it can be argued that it may affect the comparability of cross-country results.

Hence, we also adopted a second approach and smoothed the original productivity data by running a statistical algorithm through the productivity series (the HP filter proposed by Hodrick and Prescott (1980)).[22] We thus applied the technique described in Box 1 (e.g. equation (5)) and were able to estimate an alternative set of both trend-productivity gains and break years. This second approach has several advantages, such as: the ability to obtain cyclically-adjusted series in a direct way and without having to make any particular economic assumption (especially regarding the reference to the business cycle, in contrast with equation (7) in Box 1); the comparability of country estimates; a larger set of data; and the possibility to incorporate expected developments in productivity in the analysis (since productivity was forecasted by the OECD up to 2005 while the latest capacity utilisation data were only available up to the beginning of 2004[23]).

However, such a statistical method has also well-known limitations (see, regarding precisely the issue of measuring productivity trends, the discussions in O'Mahony and van Ark (2003)): it rests on somewhat arbitrary assumptions especially regarding its degree of smoothness (the choice of λ); the filtered series may retain some undetected pro-cyclicality; one cannot be sure that the correction reflects cyclical developments rather then other unknown factors; and the direct impact of the cycle on productivity gains cannot be directly estimated. But the major problem is the "end-of-sample bias", i.e. the fact that the estimations of the recent HP trends tend to be overly influenced by the latest observations of the sample. This is a major handicap here since our scope is precisely to look at recent developments in trend-productivity gains.

Finally, we also estimated a third set of trend-productivity gains by running the HP filter directly on productivity gains. But the end-of-sample bias, in this case, appeared to be an even more significant problem. In addition, this method does not allow to determine break years (since the trend is equal to the filtered data).

Against this background, the present study has favoured the first approach of directly correcting for cyclical developments. But we also checked the results by comparing them to the two other estimates, i.e. the trends calculated on HP-filtered productivity levels (in log)[24] and the HP-filtered productivity gains themselves. These results, presented in Annex C, give in general a broadly similar picture regarding both relative trends in productivity gains and estimates of break-years. For some countries (generally the smallest ones), however, these alternative methods pointed to relatively different results, arguing for some caution when interpreting them.

5. Recent trends in labour productivity

Influence of the business cycle on labour (and capital) productivity: some empirical evidence

Empirical evidence confirms that productivity displays significant cyclical movement. In particular, capacity utilisation has a positive impact on both labour and capital productivity gains in almost all industrial countries, as shown in Table 5.

The first point to note is that if breaks in productivity trends are not considered, then estimates of the impact of the cycle on productivity gains are often insignificant, or obtain the wrong (negative) sign. However, when allowing for breaks in trend-productivity gains, the impact of the cycle is highly significant and, as expected, positive in almost all countries – the main exceptions being Australia (see footnote 19) and Denmark.

22 *Hence we estimated directly trend productivity by running the HP filter so as to satisfy:*

$$\underset{P_t^*}{Min} \sum (\log P_t - \log P_t^*)^2 + \lambda \sum (\Delta \log P_t^* - \Delta \log P_{t-1}^*)^2$$

Where log P *is the logarithm of productivity and* log P* *the logarithm of trend-productivity. A general feature of this filter is that it takes into account both closeness to actual productivity (the first term of the minimisation) and the variability of the trend (second term). The relative weight of these two criteria is set by the choice of the parameter λ, typically 100 for annual data. For an application of this technique to productivity data in the United States and France, see Gilles and L'Horty (2003).*

23 *No projections of capacity utilisation are available. Hence the correction of the cycle presented in Box 1 is not possible for the two forecasted years (2004 and 2005) – though for 2004 we nonetheless conducted the analysis by taking into consideration developments in capacity utilisation observed in the first part of the year.*

24 *It should be noted, however, that applying the HP filter on productivity levels appears to yield less smoothing as well as a somewhat higher occurrence of break points.*

Table 5 – Impact of the cycle on labour and capital productivity[1]

	Output per person		Output per hour worked		Output per capital	
	Observed[2]	Corrected[3]	Observed[2]	Corrected[3]	Observed[2]	Corrected[3]
Australia	...	−0.7**	...	−0.6**	3.4***	1.1**
Austria	...	0.9***	–	–	2.3***	1.6***
Belgium	−2.4***	1.3***	−3.2***	1.0**	...	2.4***
Canada	1.1***	1.0***	...	0.7***	...	2.5***
Denmark	...	−0.8***	...	−0.9***	−2.0***	−2.0***
Finland	2.3***	1.6***	3.0***	1.1***	...	5.5***
France	...	0.4***	...	0.8***	...	0.8***
Germany	−2.9***	0.8***	−3.5***	0.8***	3.0***	1.9***
Ireland	...	1.5**	...	...	−5.5***	−2.5**
Italy	−2.1***	2.3***	−3.4***	1.7***	0.9**	2.2***
Japan	...	2.0***	...	2.1***	5.2***	4.4***
Netherlands	...	1.4***	...	0.8***	1.0**	1.8***
New Zealand	1.3***	1.3***	0.9*	1.8***	4.8***	2.1***
Norway	...	0.9**	−1.8**	0.8**	2.9*	2.3***
Spain	...	0.4**	−4.6***	0.8**	4.2***	1.5***
Sweden	2.2***	2.0***	...	1.4***	4.6***	2.2***
Switzerland	...	1.2***	...	1.2***	3.1***	2.8***
United Kingdom	...	0.7***	...	−0.8 **	3.0***	2.1***
United States	1.0*	0.8***	...	0.8***	3.1***	2.5***
Euro area	...	1.2***	...	0.5***	2.2***	1.9***
OECD ex. US	...	1.2***	...	1.0***	2.6***	2.2***
OECD	0.5*	1.1***	...	0.5***	2.6***	2.3***

[1]*See footnotes of Table 4 for the exact composition of country groups and footnote 19 for Australia. The ***, **, * represent the significance level of 1, 5 and 10 percent respectively.*

[2]*Elasticity* c *when regressing productivity on a time trend and* CY *(no time-break is allowed, ie:* log P = a T + b + CY*).* CY *is the capacity utilisation in manufacturing (industry in some cases), normalised in order to allow for cross-country comparisons.*

[3]*Same elasticity, but measured when break trends are allowed in the specification as explained in Box 1 (equation 5).*

Sources: OECD; national data; BIS calculations.

Second, the estimated impact of the state of the business cycle on labour productivity gains depends on the way the input of labour is measured. In general, it is higher when labour productivity is measured as the ratio of GDP to the number of persons employed, and lower when it is measured as the ratio of GDP per hours worked – though there are some exceptions. The probable explanation for that is that in response to cyclical developments firms are better able to adjust the number of hours worked per employee than the number of their employees.

A third point is that for almost all countries the impact of the cycle is much more important for capital productivity than for labour productivity. This is not surprising, given that firms have only limited ways to adjust their stock of capital in the near term. If, for instance, the economy is picking up, they might invest more or slow the rate of capital depreciation, but this will have only a modest impact on the stock of capital. As a result, the influence of the cycle on capital productivity gains is rather large and positive.

Turning to cross-country comparisons, it might be thought that the cyclical impact on labour productivity gains (defined as output per person employed) should be the lower, the more flexible labour markets are. Indeed, the influence is relatively lower in the *United States*, high in *Japan* and *Italy*. In the most of other euro area countries, it is between these two extremes – confirming widely shared views about the functioning of labour markets in various economies. There are, however, notable exceptions. In particular, cyclical effects in some large European labour markets often considered as inflexible (e.g. *France* and *Germany*) do not appear to be particularly large.

Cross-country differences look somewhat less significant when considering the productivity of capital. There are nonetheless some notable exceptions. The impact of the cycle looks particularly low in *France*. By contrast, the cyclical influence appears quite high in *Japan*.

Trend labour productivity gains in the business sector have diverged across OECD countries

A general long-term trend in the industrial world has been that labour productivity gains have on average declined since the 1960s. Trend productivity (with input of labour defined as the number of hours worked) was growing by around 3½% per year in the *OECD area* as a whole in the late 1960s/early 1970s. These gains fell sharply thereafter, by around 1½ percentage point, and have stabilised at rather low levels over the past three decades. These developments have been shared by most countries, though in very different ways (Table 6 and Graph 1).

- A unique case is the *United States*, where output per hour worked has significantly accelerated since the late-1990s and now appears to be growing at the same pace as before the 1970s, i.e. at an annual rate of around 3%.
- The performance of the other main industrial countries as a whole has been weaker. The *OECD area (excluding the United States)* has experienced a steady and sharp deceleration in trend labour productivity. Productivity appears to be growing by almost 2% per year since the mid-1990s, around half of the rate registered in the early 1970s.
- This slowdown has been very significant in the *euro area* as a whole, and particularly in *Italy* and *Spain*. In other euro area countries (e.g. *Germany* and the *Netherlands*) as well as in *Japan*, the slowdown has been more uneven. Productivity accelerated in the course of the 1980s, but only temporarily, apparently reflecting periods of strong demand growth (German unification; sharp inflation in asset prices in Japan in the course of the 1980s). However, trend productivity gains appear to have resumed their decline in the 1990s, both in Germany and Japan.
- In a third group of countries (France, the United Kingdom, Canada and the Nordic countries), trend gains in output per hour dropped sharply after the 1960s but have stabilised or even began to increase somewhat in more recent years. The situation in the *United Kingdom* has surprised some observers, since several reforms have been implemented there over the past few decades in order to improve the functioning of various markets – indeed, the HP-method points to a minor acceleration in productivity in recent years. In *France*, very recent data might suggest that productivity gains have improved a little, but this still looks relatively uncertain. In contrast, a significant acceleration in trend productivity appears to have taken place since the 1980s in *Sweden*, *Denmark* and *Norway* – but, in contrast to the United States, productivity gains are still well below the levels recorded several decades ago.
- Trend productivity gains have barely changed in the few remaining countries (e.g. Australia, Ireland and New Zealand) since the 1970s. Latest developments could suggest that the situation might have deteriorated recently in *Australia* (where productivity has significantly decelerated during the 2000–2003 period, but where some uncertainty surrounds the data used here – see footnote 19) and *Ireland* (productivity gains dropped significantly in 2003); but it seems to be too early to get firm conclusions at this stage.

6. Developments in total factor productivity

The general framework

Another important issue when looking at developments in labour productivity is the influence of capital accumulation and technological progress. The reason is that higher gains in labour productivity can result from an increase in output using the same units of labour and capital inputs (so-called "technological progress"), or from using more capital for a given labour input ("capital deepening"), or from a combination of these effects.

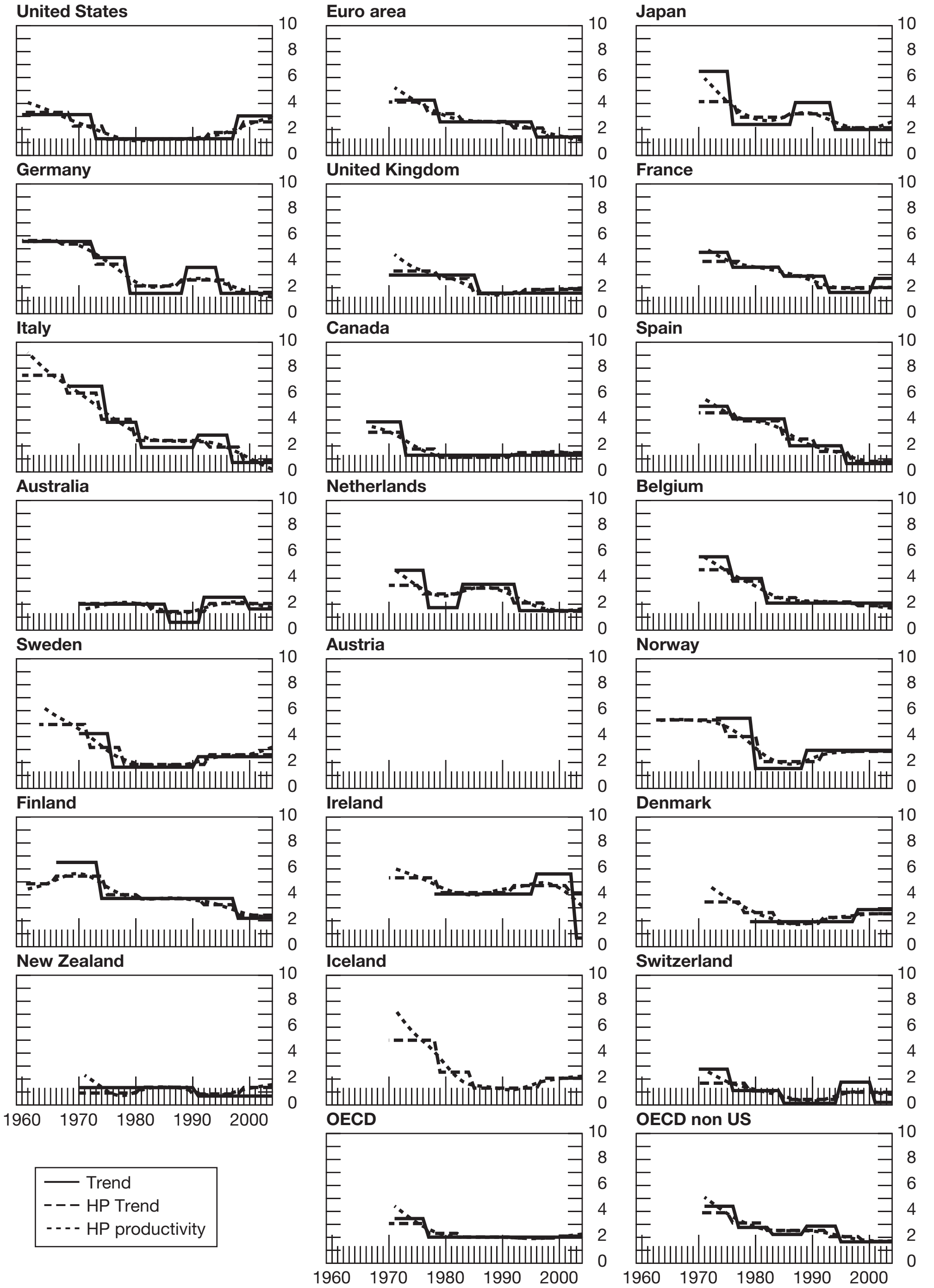
Table 6 – Most recent trends in labour productivity gains[1]

	Previous trend		Current trend		Acceleration	Confirmation by other methods?
	Start year	Average	Start year	Average		
Australia	Early 90s	2½	Early 00s	1½	−1	*Not confirmed*
Austria	Mid 70s	2½	Early 00s	1	−1½	Yes, but smaller decline
Belgium	Mid 70s	4	Early 80s	2	−2	Yes
Canada	Mid 60s	4	Mid 70s	1¼	−2¾	Yes
Denmark	Late 70s	2	Late 90s	3	+1	Yes, but smaller increase
Finland	Mid 70s	3¾	Late 90s	2¼	−1½	Yes
France	Mid 90s	1½	Early 00s	2¾	+1¼	*Not confirmed*
Germany	Late 80s	3½	Mid 90s	1½	−2	Yes , but smaller decline
Iceland	Mid 80s	1¼	Late 90s	2	+¾	–
Ireland[2]	Mid 90s	5½	Early 00s	¾	−4¾	*No, much smaller decline*
Italy	Early 90s	2¾	Late 90s	¾	−2	Yes
Japan	Mid 80s	4	Mid 90s	2	−2	Yes, but smaller decline
Netherlands	Mid 80s	3½	Mid 90s	1½	−2	Yes
New Zealand	Early 70s	1¼	Early 90s	¾	−½	*Not confirmed*
Norway	Early 80s	1½	Late 80s	3	+1½	Yes
Spain	Mid 80s	2	Mid 90s	¾	−1¼	Yes
Sweden	Mid 70s	1½	Early 90s	2½	+1	Yes
Switzerland	Mid 90s	1¾	Early 00s	¼	−1½	*Not confirmed*
United Kingdom	Early 70s	3	Mid 80s	1½	−1½	Yes, but perhaps slight increase recently
United States	Mid 70s	1¼	Late 90s	3	+1¾	Yes
Euro area	Late 70s	2½	Mid 90s	1½	−1	Yes
OECD ex. US	Late 80s	3	Mid 90s	1¾	−1¼	Yes, but smaller decline
OECD	Early 70s	3½	Late 70s	2	−1½	Yes, but perhaps slight increase recently

[1]*Calculated using input of labour, expressed as hours worked (number of employees for Austria, HP filtered data for Iceland); business sector; annual rates in per cent. See footnotes of Table 4 for the exact composition of country groups, footnote 19 for Australia, and table 5 for the methodology applied.*

[2]*Results of the econometric estimation are not significant for Ireland.*

Graph 1 – Labour (hours) productivity



Note: Results of the econometric estimation are not significant for Ireland.
Sources: OECD; BIS calculations.

Box 2 – Total factor productivity – the general framework

Estimating TFP

The starting point remains the model introduced by Solow (1956) that highlights the main sources of economic growth: GDP is obtained by a simple production function that allows capturing the relationship between output and the two substitutable production factors that are capital and labour.

This production function can be written as:

$$GDP = F(K, L, TFP) \tag{8}$$

where *GDP*, *K*, *L* and *TFP* respectively stand for output, the capital stock, labour input, and a residual. *L* can be measured in the number of persons employed or in the number of hours of work (defined as L^h). *K* is the capital stock that is in place. *TFP* is called total factor productivity which captures the contribution of all the factors not incorporated in the measure of labour. Hence, *TFP* reflects the influence of various factors such as technical and organisational progress, effort, and so on; here, it is referred to as the "technological progress".

While *Y*, *K* and *L* (or L^h) are statistically observed, *TFP* has to be estimated as a residual. One common way is to retain for (8) a Cobb-Douglas production function, which has the advantage to be easily manipulated and is usually considered as consistent with stylised facts.[25] This leads to:

$$GDP = TFP\ L^{\alpha}\ K^{1-\alpha} \tag{9}$$

where α is the share of labour in value added and $(1-\alpha)$ the share of capital. On the basic assumption that production factors are remunerated at their marginal productivity and that α is stable over the sample period,[26] one can estimate *TFP* by:

$$\log(TFP) = \log(GDP) - \alpha \log(L) - (1-\alpha)\log(K) \tag{10}$$

Decomposing labour productivity growth

The relation between labour productivity and total factor productivity is also straightforward. Indeed, combining (1) and (9) gives:

$$P = GDP/L = TFP\ L^{\alpha}\ K^{1-\alpha}/L = TFP\ (K/L)^{1-\alpha}$$

$$P = TFP\ R^{1-\alpha} \tag{11}$$

where *R* is the ratio of capital per unit of labour (called "capital depth").

The differentiation of (11) yields directly the relationship between yearly changes in labour productivity (*p*) and in total factor productivity (*tfp*):

$$p = tfp + (1-\alpha)\ r = g - l; \tag{12}$$

where:

$p\ = \Delta P / P \approx \Delta(\log P)$

$tfp \approx \Delta(\log TFP)$

$g\ \approx \Delta(\log GDP)$

$l\ \approx \Delta(\log L);$

$r\ = \Delta(\log R)$ is the yearly variation in the ratio of capital per unit of labour, called "capital deepening" (the impact of which on labour productivity gains is obtained by multiplying by $(1-\alpha)$);

g and *l* are the growth rate of respectively the real GDP and the labour factor.

25 *Needless to say, there are several approaches for estimating production functions. See for instance Gradzewicz and Kolasa (2004) for a recent estimation of a cointegration relationship between production and inputs of factors in the case of Poland.*

26 *In reality, α fluctuates over the cycle. We therefore retained the average value of α over the sample period when conducting the estimations.*

It should be noted that the decomposition (11) for labour productivity (P) defined as *GDP* per employee can also be made for labour productivity defined as *GDP* per hours worked (P_h). Needless to say, the estimation of both TFP_h and R_h in this case has to consider the total of hours worked (not solely the number of employees).

Finally, one can also decompose *TFP* gains in changes in both labour and capital productivity (resp P and P_k). As already noted, we have:

$P = TFP\, R^{1-\alpha}$ thus $p = tfp + (1-\alpha)\, r$

$P_k = TFP\, R^{-\alpha}$ thus $p_k = tfp - \alpha\, r$

Hence:

$$tfp = \alpha\, p + (1-\alpha)\, p_k \quad (13)$$

In other terms, *TFP* gains are the weighted average (the weights being respectively α and $(1-\alpha)$, the respective shares of the product factors in value added) of labour productivity gains (p) and capital productivity gains (p_k).

The long-run view

When using a Cobb-Douglas production function as in (9), the marginal productivity of capital *mpc* can be defined by:

$$mpc = \partial GDP/\partial K = (1-\alpha)\, TFP\, L^{\alpha}\, K^{-\alpha} = (1-\alpha)\, GDP/K$$

If production factors are remunerated at their marginal productivity (a valid assumption in case of perfect competition), then the income of capital can be easily obtained by:

$$income\ of\ capital = mpc\ K = (1-\alpha)\, GDP \quad (14)$$

Hence the share $(1-\alpha)$ of total income going to capital is constant, reflecting one "stylised fact" of advanced economies that the distribution of value added between labour and capital is stable in the long run; hence, capital per worker grows over time and its rate of return is broadly constant. Equation (14) shows that this is compatible with a steady capital-to-output ratio (K/GDP) and a steady rate of return of capital.[27]

In this steady state case, the relation between *TFP* and labour productivity can be simplified:

According to (12), we have:

$$p = tfp + (1-\alpha)\, r = g - l$$

since $R = K/L$ and $\Delta K / K = \Delta GDP / GDP = g$ (stability of the ratio K/GDP in the long run), then:

$$r = \Delta R/R = k - l = g - l = p \quad (15)$$

hence we have:

$$tfp = \alpha\, (g - l) = \alpha\, p \quad (16)$$

In sum, along the steady-state growth path, TFP gains are equal to labour productivity gains multiplied by the income share of labour.

One major implication of the decomposition presented above is that in the longer-run labour productivity gains should be driven by the rate of technological progress (i.e. *TFP* gains). The impact of capital deepening on labour productivity gains would be constant.

This has important implications when looking at changes in labour productivity gains and their long-run sustainability. As noted in (12), we can disentangle these changes in the sum of changes in *TFP* gains and changes in the contribution of capital deepening:

$$\Delta p = \Delta tfp + \Delta((1-\alpha)\, r) \quad (17)$$

27 *These stylised facts have been reported by Kaldor (1961). However, it should be noted that this applies to a closed economy and if both the rate of depreciation of capital and the saving- (or investment-) to-GDP ratio are stable. Indeed, the capital/output ratio may change over the long run, in response to changes in relative prices (witness the ongoing fall in IT prices) or in labour resources and technology.*

with α constant and the sustainability conditions (15) and (16) we get:

$$\Delta tfp = \alpha \, \Delta p \text{ and} \quad (18)$$

$$\Delta \, ((1-\alpha) \, r) = (1-\alpha) \, \Delta r = (1-\alpha) \, \Delta p = (1-\alpha)\Delta tfp/\alpha \quad (19)$$

If, for instance, α=2/3 (roughly the case among most industrial economies), than an acceleration in labour productivity of say one percentage point will be considered as "sustainable" if it is matched by an increase in *TFP* gains by 2/3 pp and by an increase in the contribution of capital deepening by 1/3 pp – the ratio between these two figures being equal to 2, i.e. (α /(1−α)). If, in contrast, *TFP* gains have not changed, then the contribution of capital deepening would amount to one percentage point and would be largely excessive. This would raise the risk of a downward correction of the capital/output ratio latter on, and therefore of a smaller or even negative contribution of capital deepening to future labour productivity gains.

From this perspective, to what extent did the developments noted in Section V in terms of labour productivity reflect differences in accumulating capital, especially IT equipment, or in rates of technological progress? As detailed in Box 2, growth theory is the obvious framework to deal with these issues, as it puts emphasis on supply-side factors such as the labour force and technological progress when looking at long trends in economic growth. TFP can be estimated using a simple production function approach on OECD annual data on GDP, employment, average hours worked and capital stock in the business sector, since 1960.[28] The data are however not available for all industrial countries and over the whole sample (see annex A and results in Table 7).

Disentangling these factors is essential when comparing developments in productivity across countries and, in particular, when judging the sustainability of any acceleration in labour productivity. For instance, if the increase in labour productivity growth has solely been the result the accumulation of equipment, then maintaining this trend would require that capital spending would have to continue to grow at a high rate. This might however not be sustainable in the longer run since a "stylised economic fact", at least in advanced economies, is that capital-output ratios as well as the shares of labour and capital in national income are expected to be constant over long periods.

However, it looks sensible to assume that the difficulties described in Section II (in terms of measuring both output and the inputs of labour) are even worse when measuring the input of capital across countries. For instance, determining what part of technological change should be seen as part of the capital stock is difficult (see van Ark (2001)). Disentangling the contribution of IT and non-IT capital might be important, as argued by Khan and Santos (2002) using a growth-accounting exercise similar to the one presented in Box 2 on Canadian data. Or different ways in aggregating capital stock items or estimating the depreciation profile can yield opposite pictures for the growth contribution of capital, as observed by Wadhani (2001b) and estimated by Oulton and Srinivasan (2003). It should be reminded, from this perspective, that the valuation of existing fixed assets is quite uncertain. In principle, it should be carried out at replacement or market prices. In practice, however, firms record assets at historic acquisition costs in their balance sheet and replacement costs are often derived from bankruptcy procedures and therefore likely to be biased downwards. Hence, statisticians have to use an indirect method (called the "perpetual inventory method") to estimate the market value of assets, based on the application of adequate price index numbers to cumulated flows of gross fixed capital formation (net of depreciation).

Furthermore, uncertainty in measuring the contribution of capital to economic growth affects estimates of TFP growth, a point noted by Schreyer (2001) when dealing with different set of ICT deflators. Moreover, the share of physical capital in real output (hence the choice of α) can have crucial implication for measuring the contribution of TFP to output growth, as argued by Senhadji (2000). This reinforces the view that one should focus on the evolution of TFP gains over time rather then on simply comparing their levels across countries.

28 *This is certainly a very simple way of measuring TFP. Dean and Harper (1998) provide an overview of the questions surrounding the measurement of multifactor productivity, benefiting from the leading expertise of the US Bureau of Labour Statistics. For the specific issues surrounding the measurement of multi-factor productivity and quality adjusted measures of factor inputs (in particular of capital services or human capital), see Groth et al. (2004) and Schreyer (2003). Alternative conceptual measures of the capital stock can also yield different results as argued by Wadhwani (2001). In addition, the productivity of public capital might be an important factor that is not taken into account in these calculations limited to the business sector (see Kamps (2004)).*

Table 7 – Total factor productivity in the business sector

Average annual percentage changes[1]

	Total factor productivity				Total factor productivity (hours worked)			
	1966–1975	1976–1985	1986–1995	1996–2004	1966–1975	1976–1985	1986–1995	1996–2004
Australia	1.0	0.7	0.6	1.2	−0.1	0.8	0.6	1.4
Austria	2.2	0.8	1.1	0.4	–	–	–	–
Belgium	2.2	1.4	0.9	0.5	2.8	1.7	1.2	0.9
Canada	1.2	0.4	0.7	1.3	1.7	0.7	0.7	1.2
Denmark	0.2	0.6	0.6	0.8	1.0	1.5	0.8	1.0
Finland	–	1.7	1.8	2.6	–	2.0	1.9	2.7
France	2.4	0.4	0.6	0.5	0.9	1.0	0.9	1.0
Germany	1.7	1.0	1.3	0.6	2.5	1.4	1.7	0.9
Iceland	2.1	1.6	0.8	1.6	3.1	2.1	0.9	1.6
Ireland	4.6	2.2	3.5	3.4	4.7	2.7	3.7	4.0
Italy	2.6	1.3	0.9	−0.4	2.9	1.6	1.0	−0.3
Japan	2.3	0.5	0.1	0.1	0.4	0.6	0.7	0.3
Netherlands	2.6	1.2	1.1	0.4	2.5	1.7	2.2	0.6
New Zealand	−0.0	−0.0	0.3	0.7	0.1	0.1	0.3	0.8
Norway	2.7	1.1	1.4	1.9	3.5	1.8	1.6	2.2
Spain	2.7	0.6	0.3	0.2	1.9	1.1	0.4	0.2
Sweden	1.6	0.4	1.6	1.5	2.6	0.6	1.2	1.8
Switzerland	0.5	0.2	−0.6	0.1	−0.3	0.7	−0.4	0.2
United Kingdom	2.0	1.7	1.0	0.7	1.9	2.1	1.1	0.8
United States	0.7	0.9	1.1	1.9	1.2	0.9	1.0	2.0
Euro area	1.1	1.0	0.9	0.3	1.7	1.4	1.2	0.6
OECD ex. US	1.0	1.0	0.8	0.4	1.7	1.3	1.1	0.6
OECD	1.1	1.0	0.9	1.1	1.6	1.2	1.1	1.2

[1]*Cross-country comparisons for 1966–1975 might be misleading given that data for some countries were not available for the entire period. See footnotes of Table 4 for the exact composition of country groups.*

Sources: OECD; national data; BIS calculations.

Finally, developments in TFP gains are likely to be positively influenced by cyclical developments, even more than in the case of labour productivity. As already noted, the reason why is that the adjustment of the capital stock to the state of the business cycle is less rapid than in the case of labour.

Given all these elements, the methods described in Box 1 have been applied to TFP estimates, either by directly correcting for cyclical developments, or by using the HP filter for smoothing purposes. The impact of the cycle on TFP gains appears significant for almost all industrial countries and – when changes in trend-TFP gains are taken into consideration – positive (Table 8). This is not so surprising given estimates presented above for both labour and capital productivity and the fact that TFP gains are an average of labour and capital productivity gains. At first sight, cross-country differences appear relatively significant. The influence of the state of the business cycle is relatively smaller in the *United States* and the *United Kingdom* but also in large euro area countries such as *France* and *Germany*. It is quite important in some other European countries, such as *Italy*, and *Japan*.

Cross-country developments in TFP

Leaving aside cyclical influences, long-term developments in total factor productivity reflect the trends in labour productivity described above and also more specific elements. For the OECD area as a whole, TFP gains appear to have been relatively stable over the past few decades (Table 9 and Graph 2). They declined slightly from the 1976/85 period to the 1986/1995 period but have improved marginally in the past decade.

Table 8 – Impact of the cycle on total factor productivity

	Total factor productivity		Total factor productivity (hours worked)	
	Observed	Corrected	Observed	Corrected
Australia	1.7***	−0.7**	1.7***	1.7***
Austria	...	1.3***	–	–
Belgium	−1.3***	1.7***	−1.7***	1.5***
Canada	2.0***	1.6***	1.4***	1.5***
Denmark	...	−0.9***	−1.1***	−1.1***
Finland	...	3.4***	1.6***	3.0***
France	...	0.7***	0.6***	0.7***
Germany	...	1.3***	...	1.2***
Ireland	–2.1*	…	–3.0**	–3.0**
Italy	...	2.2***	...	2.2***
Japan	...	2.7***	1.8***	2.7***
Netherlands	...	1.7***	...	1.1***
New Zealand	2.9***	2.9***	2.8***	2.8***
Norway	...	1.4***	...	1.5***
Spain	...	0.7***	...	1.0***
Sweden	3.0***	2.1***	1.7***	2.1***
Switzerland	0.9**	1.7***	1.3***	1.4***
United Kingdom	...	1.4***	1.3**	1.0***
United States	1.8***	1.4***	1.2**	1.1***
Euro area	...	1.5***	...	1.3***
OECD ex. US	1.1***	1.7***	0.8*	1.4***
OECD	1.4***	1.6***	0.9***	1.3***

*Note: Total factor productivity estimated by considering the labour input as the number of people employed; total factor productivity (hours worked) is estimated by considering the labour input as the number of hours worked. The ***, **, * represent the significance level of 1, 5, and 10 percent respectively. See footnotes of Table 4 for the exact composition of country groups, footnote 19 for Australia, and table 5 for the way the impact of the cycle is estimated.*

Sources: OECD; national data; BIS calculations.

- The first point to note is that the *US* performance looks impressive: trend TFP (per hour) has been steadily accelerating since the late 1970s. It accelerated in the 1980s and again in the 1990s and is now estimated to be growing by more than 2% per year, i.e. almost as rapidly than in the 1960s. From this perspective, the US improvement appears more deeply rooted than when just looking at labour productivity (which accelerated in the course of the 1990s): TFP has accelerated earlier, and in a more continuous way.
- A second, and related, result is that the performance of the United States is even more obvious in relative terms, since trend TFP gains have significantly declined over the past few decades in other main industrial countries as a whole, to as low as ½% per year currently. As a result, the discrepancy between the United States and the *rest of the OECD* has steadily widened (Graph 3).
- Reflecting developments in labour productivity, the slowdown in TFP has been marked in the *euro area* as a whole, particularly in *Italy* and *Belgium*. The situation has also deteriorated but less markedly in *Germany* and *Spain*. In *Japan* as well, the weakening in trend TFP gains over the past few decades has been somewhat more moderate and less continuous than the decline in labour productivity gains; nevertheless, TFP appears to be flat or even decreasing, as output growth is mainly attributable to higher inputs of labour and capital deepening. The *UK* performance has declined since the 1970s, presenting a picture similar to the developments in labour productivity analysed above.
- Compared to their performance in trend labour productivity, the outlook looks somewhat more favourable in terms of TFP gains in Canada and France. In *Canada*, trend TFP accelerated in the late 1990s though this improvement has still to be confirmed (recent data have not been that strong). TFP gains have been almost stable in *France* since the 1970s, with apparently a very slight improvement after the mid-1990s. The acceleration observed in several *Nordic countries* in terms of labour productivity is confirmed by developments in TFP gains. Though substantial uncertainty remains, the situation of *Australia* and *New Zealand* also appears to have been stable or even improving recently.

Table 9 – Most recent trends in total factor productivity gains[1]

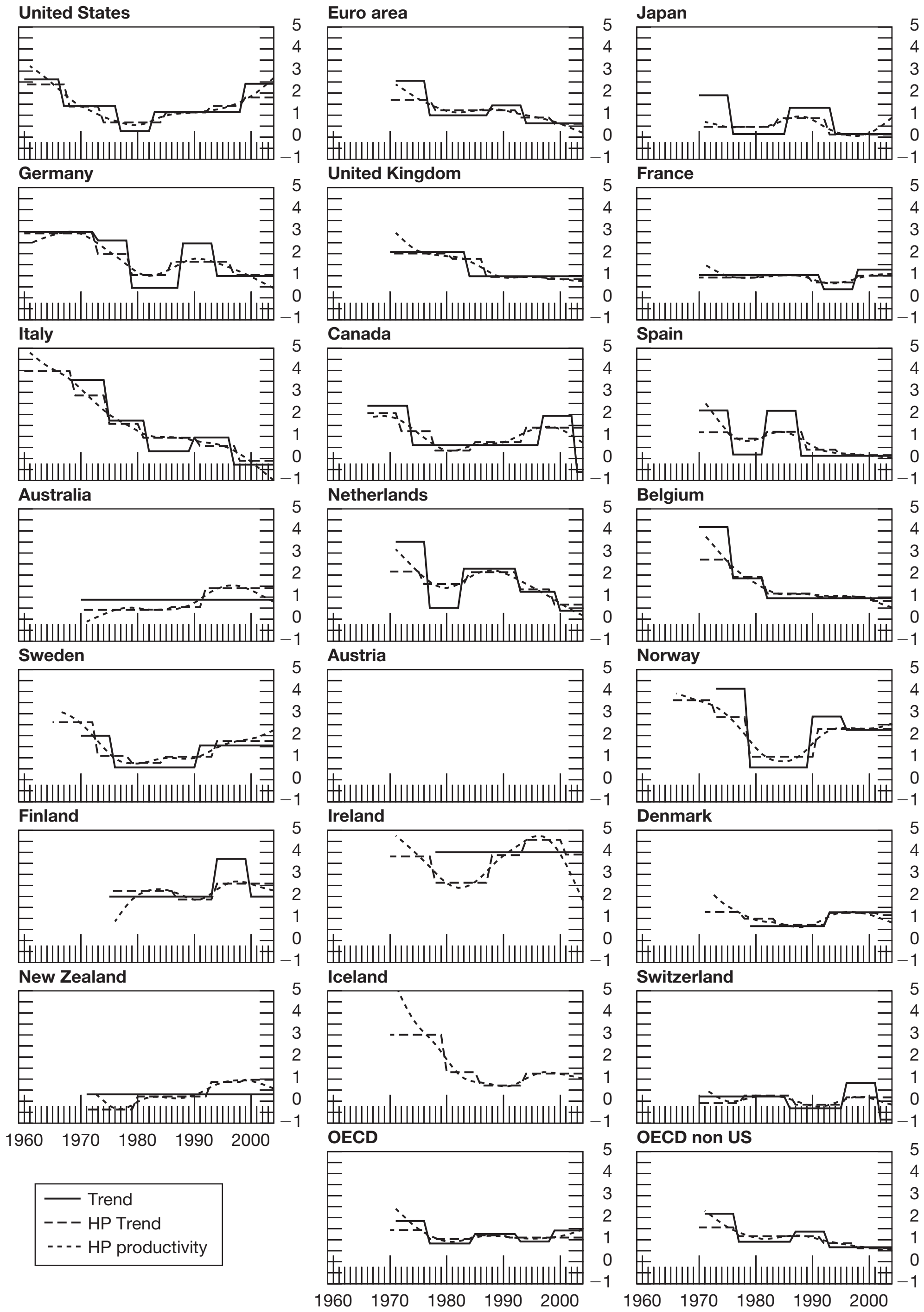
	Previous trend		Current trend		Acceleration	Confirmation by other methods?
	Start year	Average	Start year	Average		
Australia	–	–	Early 70s	1	0	*No, TFP might have accelerated in the 1990s*
Austria	Late 80s	1¼	Early 00s	−¾	−2	Yes, but smaller decline
Belgium	Mid 70s	2	Early 80s	1	−1	Yes
Canada	Late 90s	2	Mid 00s	−½	−2½	*No, the 1990s improvement remains*
Denmark	Late 70s	¾	Mid 90s	1¼	+½	Yes
Finland	Mid 90s	3¾	Early 00s	2	−1¾	*No, the 1990s improvement remains*
France	Early 90s	½	Late 90s	1¼	+¾	Yes, but smaller increase
Germany	Late 80s	2½	Mid 90s	1	−1½	Yes, but smaller decline
Iceland	Mid 80s	¾	Mid 90s	1¼	+½	–
Ireland	–	–	Late 70s	4	0	*No, TFP decelerated perhaps in the 2000s*
Italy	Early 90s	1	Late 90s	−¼	−1¼	Yes
Japan	Mid 80s	1¼	Mid 90s	0	−1¼	Yes
Netherlands	Mid 90s	1¼	Early 00s	½	−¾	Yes
New Zealand	–	–	Early 70s	¼	0	*No, TFP accelerated perhaps in the 1990s*
Norway	Early 90s	3	Mid 90s	2¼	−¾	*No, the 1990s improvement remains*
Spain	Early 80s	2¼	Late 80s	0	−2¼	Yes, but smaller decline
Sweden	Mid 70s	½	Early 90s	1½	+1	Yes
Switzerland	Mid 90s	¾	Early 00s	−¾	−1½	*Not confirmed*
United Kingdom	Early 70s	2	Mid 80s	1	−1	Yes
United States	Mid 80s	1¼	Late 90s	2½	+1¼	Yes
Euro area	Late 80s	1½	Mid 90s	½	−1	Yes
OECD ex. US	Late 80s	1½	Mid 90s	¾	−¾	Yes
OECD	Mid 90s	1	Late 90s	1½	+½	Stability instead of a slight improvement

[1]*Calculated using input of labour, expressed as hours worked (number of employees for Austria); business sector; annual rates in per cent. See footnotes of Table 4 for the exact composition of country groups, footnote 19 for Australia, and Table 5 for the methodology applied.*

Different patterns in capital accumulation

These cross-country developments might reflect different patterns in capital accumulation ("capital deepening") as well as varying rates of technological progress (measured by the growth rate of total factor productivity – TFP), as indicated in Table 10.

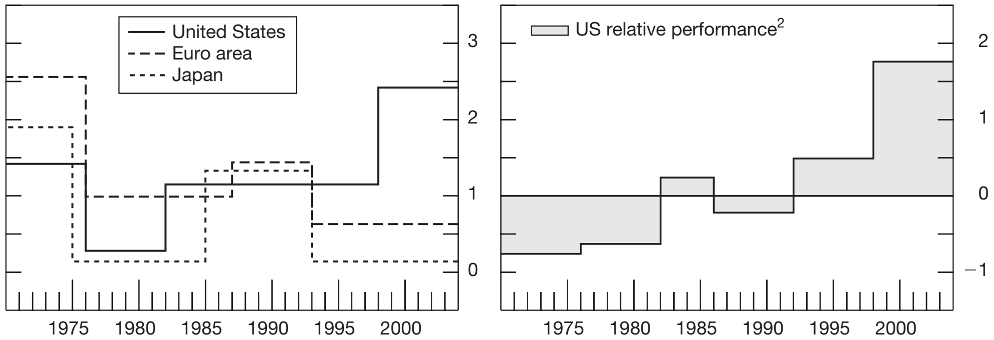
Graph 2 – Total factor productivity (hours)



Sources: OECD; BIS calculations.

- In the *United States*, the main feature behind the recent increase in labour productivity gains has been the acceleration in technological progress. The contribution of capital deepening has been also positive, especially since the mid-1990s, but not excessive: trend TFP increased by 2% annually in the 1996–2004 period while the contribution of capital deepening to annual labour productivity gains was of 0.7 percentage points. Indeed, a major factor behind the US acceleration in TFP over the past few decades has been the sharp improvement in the productivity of capital (see Annex B2).
- In the main *OECD countries excluding the United States*, by contrast, the deceleration in TFP over the past few decades has been accompanied by a relatively strong contribution of capital deepening to labour productivity gains (which, indeed, has been higher than the contribution from TFP gains while it should be significantly lower in the longer run as argued in Box 2).

Graph 3 – Trends in total factor productivity (TFP)[1]
Annual changes, in percent



[1]TFP is calculated using input of labour, expressed as hours worked; business sector; for an explanation, see box 2.
[2]Trend TFP gains in the United States minus those in the group composed of Australia, Canada, euro area countries (see footnote 2 of Table 4), Japan, Sweden and the United Kingdom.

Sources: OECD; BIS calculations.

Table 10 – Trend-productivity gains and changes over the past three decades

	Trend gains, 1996–2004, annual rates in per cent					Changes in trend gains, from 1986–95 to the 1996–2004 period		Changes in trend gains, from 1976–85 to the 1986–1995 period	
	Output per person	Contribution of average hours worked[1]	Output per hour worked	Contribution of capital deepening[2]	TFP (hour worked)	Output per hour	TFP	Output per hour	TFP
Australia	1.8	−0.2	2.0	1.2	0.9	0.7	−0.0	−0.6	0.0
Austria	1.8	–	–	1.4	0.4	−0.7	−0.8	0.0	0.6
Belgium	1.2	−0.9	2.1	1.1	1.0	0.0	0.0	−1.1	−0.5
Canada	1.4	0.1	1.3	0.1	1.2	0.0	0.6	0.0	0.0
Denmark	2.3	−0.4	2.7	1.4	1.3	0.7	0.4	0.0	0.2
Finland	2.3	−0.2	2.5	−0.2	2.7	−1.2	0.4	0.0	0.3
France	1.1	−1.0	2.1	1.0	1.1	−0.4	0.3	−1.0	−0.3
Germany	1.0	−0.6	1.6	0.6	1.0	−1.2	−0.8	0.4	0.7
Iceland	2.0	0.0	2.0	0.7	1.3	0.7	0.4	−1.8	−1.2
Ireland	3.7	−0.8	4.5	0.5	4.0	0.5	0.0	0.0	0.0
Italy	0.7	−0.3	1.0	1.1	−0.1	−1.4	−0.8	−0.5	−0.5
Japan	1.5	−0.5	2.0	1.9	0.1	−1.5	−1.0	1.1	1.0
Netherlands	1.3	−0.2	1.5	0.7	0.8	−1.4	−1.2	0.4	0.6
New Zealand	1.0	0.3	0.7	0.4	0.3	−0.3	0.0	−0.3	0.0
Norway	2.5	−0.4	2.9	0.7	2.3	0.4	0.3	−0.6	0.3
Spain	0.7	0.1	0.6	0.5	0.1	−1.4	−0.4	−2.1	−0.4
Sweden	2.0	−0.5	2.5	0.9	1.6	0.4	0.5	0.4	0.5
Switzerland	0.7	−0.4	1.1	0.8	0.3	0.8	0.6	−0.7	−0.5
United Kingdom	1.6	0.0	1.6	0.6	1.0	0.0	0.0	−1.4	−0.9
United States	2.5	−0.1	2.7	0.7	2.0	1.4	0.8	−0.0	0.5
Euro area	0.9	−0.5	1.4	0.8	0.6	−1.2	−0.6	−0.5	0.0
OECD ex. US	1.3	−0.4	1.7	1.0	0.7	−0.9	−0.5	−0.2	0.1
OECD	1.9	−0.1	2.0	0.8	1.2	0.0	0.1	−0.1	0.2

See footnotes of Tables 4 and 5 for the exact composition of country groups and the methodology applied and footnote 19 for Australia. Number of employees instead of hours worked for Austria; HP filtered data for Iceland; results of the econometric estimations are not significant for Ireland.

[1]Contribution to annual gains in trend output per worker.
[2]Contribution to annual gains in trend output per hour.

This has reflected relatively unfavourable patterns in the productivity of capital. It also raised the risk that some correction in the capital/output might occur at some point, which would dampen labour productivity gains looking ahead. This risk looks significant in the *euro area* as a whole, particularly in *Spain* and even more so in *Italy*. For instance, Italian trend-TFP gains have turned negative in the 1996–2004 period but labour productivity gains have remained relatively resilient, supported by further capital deepening. The situation looks even less favourable in *Japan*: TFP gains have declined over the past few decades, to almost zero, while the impact of capital deepening is still high (it contributed to almost 2 percentage points in trend labour productivity gains in the 1996–2004 period).

- By contrast, the outlook might look more favourable for some countries. As noted above, TFP gains have stopped declining or even started to increase in *Canada* and several *Nordic countries*. But the contribution of capital deepening has been very limited, suggesting further support for higher labour productivity gains looking ahead.

7. Conclusions and policy implications

The US performance stands out

To sum up, the level of US labour productivity is the highest among the major industrial countries and has been rising the fastest in the recent past. Rather than just reflecting stronger capital accumulation, this performance has been associated with a higher rate of technological progress (as measured by TFP) that was maintained during the latest recession. More importantly, perhaps, the US performance has improved in relative terms, as TFP growth has accelerated in the United States but decelerated in most other economies.

Are there any lessons for other countries?

Whether the growing gap between TFP in the United States and that in other countries will diminish is difficult to judge (regarding the elements of this debate, see Oesterreichische Nationalbank (2004)). Certainly, information and communication technologies have positively contributed to economic growth in many industrial countries over the past few decades and this effect has been particularly felt in the United States.[29] But observing *what* happened in the United States is different from explaining *why* it happened, as pointed by Stiroh (2001).

Nevertheless, what is clear is that the steady improvement in US productivity is not just a matter of greater use of IT equipment. The bulk of the accumulation in IT equipment occurred in the 1990s, i.e. well after US TFP started to accelerate. Moreover, IT use has also expanded in other countries over the past decade without preventing a sharp deceleration in TFP. And IT investment might be a necessary but not sufficient factor in order to observe meaningful developments in aggregate productivity (see Greenan et al., 2002).[30] Instead, much of the acceleration in trend TFP in the United States could well have come from the earlier deregulation of markets for goods and services. The resulting increase in competition might have spurred innovation by creating strong incentives to reduce production and distribution costs.[31] Indeed, the United States has seen the emergence of large producers in the IT sector, a sector characterised by both fierce competition and a very high rate of technological progress. Sizeable productivity improvements have also been recorded in retail trade, where competition has again been intense. In addition, the US labour market has been helpful; witness the long-term fall in structural unemployment since the 1970s and the ease with which workers move from declining to growing sectors. In short, the implementation of structural reforms might be

29 *For an overview of the impact of IT investment on industrial countries see Colecchia and Schreyer (2002). For the specific case of the United States see, among others, the seminal paper from Oliner and Sichel (2000).*

30 *This is not to say that this factor did not play a role. But other elements were also in play. Indeed, lower IT equipment in non-US countries was often associated with insufficient investment into the more general determinants of long-run growth and technical progress, such as research, education and the diffusion of new technologies (Aiginger (2004)). Moreover, Ferguson and Wascher (2004) argue that factors pertaining to private-sector initiative play a key role in shaping periods of strong productivity growth that are characterised by technology innovations. Furthermore, Basu et al. (2003) emphasise the role played by investments in intangible organisational capital.*

31 *For a review of the interactions between institutional reforms in markets and macroeconomic performance see Pichelmann and Roeger (2004). Gust and Marquez (2002) found that regulatory environments played a role in explaining differences in IT adoption and thus productivity developments in industrial countries in the 1990s. Looking at the impact of trade, Hung et al. (2003) estimate that competition effects stemming from import prices have been particularly powerful in fostering US productivity.*

one key prerequisite for creating stronger market incentives and duplicating the US innovation process.[32]

Undoubtedly, and as argued by Turner (2003), the above developments do not necessarily imply that those structural reforms have actually to be implemented. Indeed, individual and social choices might well point to opposite preferences (e.g. appetite for leisure; planning restrictions for retailers and the use of land; etc).

Will potential growth rise in the longer run?

Nonetheless, some uncertainty remains as to whether recent patterns in productivity growth will be maintained and whether industrial countries will be able to achieve stronger potential growth rates in the future. On the one hand, the recent improvement in US productivity gains may not last indefinitely. In particular, companies' willingness to cut costs, as well as the lagged impact of past large investments in IT equipment, may have raised the level of productivity, and thus its measured growth for a time (see Gordon (2003) for a discussion on developments in US productivity in the most recent years). Eventually, however, such effects could well fade away (Dudley (2004)).

On the other hand, the tendency for structural reforms, implemented in the past two decades, to increase the demand for less-skilled labour could also have held down measured overall productivity gains in a significant way in the United States and some European countries (e.g. Ireland, the Netherlands and the United Kingdom).[33] A related and positive consequence is that structural unemployment has now come down to lower levels. At some point, such transitory effects might begin to dissipate, possibly revealing higher underlying productivity growth in the coming years.

All in all, the discussion above suggests that trend GDP growth rates have diverged significantly among major industrial economies, though the picture is somewhat more complicated because of substantial differences in labour force growth.[34] In terms of *growth rates*, and according to the OECD, potential output might be currently growing by around 3% per year in the United States (or even 3½% according to some other estimates), compared to around 2½% in the United Kingdom, 2% in the euro area and 1½% in Japan (OECD (2004b)). In terms of *changes in growth rates*, the United States has seen a clear improvement in its relative position: potential growth is still running at roughly the same pace than in the 1980s, while it has sharply decelerated in the euro area and even more so in Japan.

References

Ahmad, N, F Lequiller, P Marianna, D Pilat, P Schreyer and A Wölfl (2003): "Comparing labour productivity growth in the OECD area: the role of measurement", *OECD Statistics Working Paper* 2003/5.

Aiginger, K (2004): "The Economic Agenda: a view from Europe", *Review of International Economics*, 12(2).

Arnold, R W (2003): "Modelling long-run economic growth", *US Congressional Budget Office Technical Paper Series*, 2003–04.

Bailey, A, S Millard and S Wells (2001): "Capital flows and exchange rates", *Bank of England Quarterly Bulletin*, Autumn.

Bank for International Settlements (2004): "Developments in the advanced industrial economies", *Annual Report*, Chapter II.

Basu, S, J G Fernald, N Oulton and S Srinivasan (2003): "The case of the missing productivity growth: or, does information technology explain why productivity accelerated in the United States but not the United Kingdom?", *NBER Working Paper* 10010.

32 *For a recent overview of the factors that can influence long-run growth in TFP, see Arnold (2003) as well as de Serres (2003). For an estimation of long-run effects of labour-productivity determinants in Europe, see Denis et al. (2004) as well as Belorgey et al. (2004). For the policy implications, see OECD (2001b).*

33 *These specific countries have introduced incentives to hire young or unskilled workers. For example, some have cut social security taxes or income taxes for low-skilled workers, while others have encouraged their recruitment through increased labour market flexibility. Such measures can adversely affect measured aggregate labour productivity for a time even though they increase potential growth in the longer run.*

34 *For a recent discussion surrounding these interactions, see Daly (2004). For a long-term view of potential growth rates and estimations see OECD (2000, 2004b).*

Belorgey, N, R Lecat and T P Maury (2004): "Determinants of Productivity per Employee: an Empirical Estimation Using Panel Data", *Bank of France Working paper*, no 110.

Blades, D and D Roberts (2002): "Measuring the non-observed economy", *Statistics Brief*, no 5, OECD.

Blanchard, O (2004): "The economic future of Europe", *NBER Working Paper* 10310.

Bodier, M, Capet S and R Duval (2001): "Estimation des croissances potentielles et des écarts d'activité pour les huit principaux partenaires de la France", Journées 2001 de l'Association Française de Science Economique (AFSE).

Burda, M and Wyplosz C (2001): "Macroeconomics: A European Text", third edition, Oxford University Press.

Colecchia, A and P Schreyer (2002): "The contribution of information and communication technologies to economic growth in nine OECD countries", *OECD Economic Studies*, no 34.

Comin, D (2003): "Using Investment Data to Assess the importance of Price Measurement", Department of Economics New York University.

Congressional Budget Office (2002): "The role of computer technology in the growth of productivity", *US CBO Paper*.

Council of Economic Advisers (2002), Economic Report of the President.

Daly, K (2004): "Euroland's Secret Success Story", *Goldman Sachs Global Economics Paper*, no 102.

de Serres, A (2003): "Structural policies and growth: a non-technical overview", *OECD Economics Department Working Papers*, no 355.

Dean, E R and M J Harper (1998): "The BLS Productivity Measurement Program", US Bureau of Labor Statistics.

Denis, C, K McMorrow and W Röger (2004): "An analysis of EU and US productivity developments (a total economy and industry level perspective)", *European Commission Economic Paper*, no 208.

Deutsche Bundesbank (2001): "Problems of international comparisons of growth – a supplementary analysis", Appendix to the *Monthly Report*, May.

Doisy, S (2001): "La croissance potentielle de l'économie française: une évaluation", *Document de travail de la Direction de la Prévision*, Ministère de l'Economie, des Finances et de l'Industrie, Novembre.

Drew, C, C Richardson and P Vaze (2001): "International Comparisons of Productivity – An update on developments", UK Office for National Statistics.

Dudley, B (2004): "Has the Technology Contribution to Productivity Growth Peaked?", *Goldman Sachs Daily Comment*, 21/07/2004.

Duval, R (2000): "What is America 'New Economy' Worth?", *Economie et Statistique*, no 339–340–9/10.

ECB (2004a): "The natural real interest rate in the euro area", *Monthly Bulletin*, May.

ECB (2004b): "Labour productivity developments in the euro area: aggregate trends and sectoral patterns", *Monthly Bulletin*, July.

Eldridge, L P, M E Manser and P F Otto (2004): "Alternative measures of supervisory employee hours and productivity growth", US Bureau of Labor Statistics, *Monthly Labor Review*, April.

European Commission, Eurostat (2004a): "The statistical guide to Europe", *Eurostat Yearbook*, 2004.

European Commission, Eurostat (2004b): "GDP per capita in Purchasing Power Standards", *Structural indicators – general economic background*, http://epp.eurostat.cec.eu.int/portal/page?_pageid=1090%2C1137397&_dad=portal&_schema=PORTAL

Ferguson, R W and W L Wascher (2004): "Distinguished Lecture on economics in Government: Lessons from Past Productivity Booms", *Journal of Economic Perspectives*, vol 18, no 2, Spring.

Filardo, A J and P N Cooper (1999): "Cyclically-Adjusted Measures of Structural Trend breaks: An Application to Productivity Trends in the 1990s".

Gilles, F and Y L'Horty (2003): "The Productivity paradox in France and the United States: A Reevaluation", *Document de Recherche* EPEE.

Gordon, R J (2003): "Exploding Productivity Growth: Context, Causes, and Implications", *Brookings Papers on Economic Activity*, 2:2003.

Gradzewicz, M and M Kolasa (2004): "Estimating the output gap in the polish economy: the VECN approach", National Bank of Poland.

Greenan, N, Y L'Horty and J Mairesse (ed) (2002): "Productivity, Inequality, and the Digital Economy: A Transatlantic Perspective", MIT Press.

Greenspan A (2002): "Productivity", Remarks at the US department of Labor and American Enterprise Institute Conference, October 23.

Griliches, Z (1994): "Productivity, R&D and the Data Constraint", *American Economic Review*, vol 84, no 1.

Groth, C, M Gutierrez-Domenech and S Srinivasan (2004): "Measuring total factor productivity for the United Kingdom", *Bank of England Quarterly Bulletin*, Spring.

Gust, C. and Marquez J. (2000): "Productivity Developments Abroad", *Federal Reserve Bulletin*, October, Board of Governors of the Federal Reserve System.

Gust, C and J Marquez (2002): "International comparisons of productivity growth: the role of information technology and regulatory practices", *International Finance Discussion Papers*, no 727, May, Board of Governors of the Federal Reserve System.

Hatzius, J (2004): "Has Outsourcing Inflated the GDP Numbers?", *Goldman Sachs US Economics Analyst*, March 26.

Hodrick, R and E Prescott (1980): "Postwar U.S. Business Cycles: An Empirical Investigation", *Discussion Paper*, no 451, Carnegie Mellon University, reprinted in *Journal of Money, Credit and Banking*, vol 29, no 1, 1997.

Hung, J, M Salomon and S Sowerby (2003): "International trade and U.S. productivity", *Technical Paper Series*, US Congressional budget Office.

IMF (2004): "Canada: Selected issues", *IMF Country Report*.

Kaldor, N (1961) Capital Accumulation and Economic Growth, *in* Lutz F., Hague D. (eds), *The Theory of Capital*, New York: St Martin's Press.

Kamps, C (2004): "New Estimates of Government Net Capital Stocks for 22 OECD Countries 1960–2001", *IMF Working Paper* WP/04/67.

Khan, H and M Santos (2002): "Contribution of ICT Use to Output and Labour-Productivity Growth in Canada", *Bank of Canada Working Paper* 2002–7.

Lequiller, F (2001): "The new economy and the measurement of GDP growth", *INSEE Working Paper* G 2001/01.

Levy, M D (2003): "Why does the U.S. Grow Faster than the EU? A View From America", Oesterreichiche Nationalbank 31st Economics Conference.

Maddison, A and B van Ark (1994): "The international Comparison of Real Product and Productivity", *Research Memorandum* 567, Groningen Growth and Development Centre.

Magnien, F, J L Tavernier and D Thesmar (2002) : "Le recul du PIB par habitant de la France traduit surtout l'imperfection des comparaisons internationales", *Economie et Statistique*, no 354.

Maury, T P and B Pluyaud B (2004): "The Breaks in per Capita Productivity Trends in a Number of Industrial Countries", *Bank of France Working paper*, no 111.

Meyer, L H (2001): "What happened to the New Economy?", Speech, June 6.

Nordhaus, W D (2002): "Productivity Growth and the New Economy", *Brooking Papers on Economic Activity* 2002:2.

O'Mahony, M and B van Ark (ed) (2003): "EU productivity and Competitiveness: An industry perspective – Can Europe resume the catching-up process?".

OECD (2000): "Recent growth trends in OECD countries", *Economic Outlook* 67.

OECD (2001a): "Measuring Productivity: measurement of aggregate and industry-level productivity growth", *OECD Manual*.

OECD (2001b): "The Growth Project", Economic Policy Committee Ministerial Report on Growth.

OECD (2004a): "OECD measures of total hours worked", The OECD productivity database, March.

OECD (2004b): Statistical Annex, OECD *Economic Outlook* 75.

Oesterreichische Nationalbank (2004): "Current Issues of Economic Growth", Proceedings of OeNB Workshops, no 2.

Oliner, S D and D E Sichel (2000): "The Resurgence of Growth in the Late 1990's: Is Information Technology the Story?", *Journal of Economic Perspectives*, vol 14, no 4.

Oliner, S D and D E Sichel (2002): "Information Technology and Productivity: Where Are We Now and Where Are We Going?", *Federal Reserve Bank of Atlanta Economic Review*, vol 87 (Fall 2002).

Oulton, N and S Srinivasan (2003): "Capital stocks, capital services, and depreciation: an integrated framework", *Bank of England Working Paper*, no 192.

Pichelmann, K and W Roeger (2004): "The EU Growth Strategy and the Impact of Aging", *Review of International Economics*, 12(2).

Pilat, D (1996): "Labour Productivity Levels in OECD countries: estimates for manufacturing and selected service sectors", *Economics Department Working Papers*, no 169.

Pilat, D and A Wölfl (2004): "ICT production and ICT use: what role in aggregate productivity growth?", OECD.

Richardson, C (2001): "International Comparisons of Productivity", UK Office for National Statistics.

Sakellaris, P and F Vijselaar (2004): "Capital Quality Improvement and the Sources of Growth in the Euro Area", *ECB Working Paper*, no 368.

Scheuer, M (2001): "Measurement and statistical issues related to the 'new economy' with IT equipment and software in Germany and the United States as a case in point", *Deutsche Bundesbank Paper*.

Schreyer, P (2001): "Computer price indices and international growth and productivity comparisons", *OECD Statistics Directorate Working Paper* STD/DOC(2001)1.

Schreyer, P (2003): "Capital stocks, capital services and multifactor productivity measures", *OECD Draft paper Economic Studies*.

Schreyer, P and F Koechlin (2002): "Purchasing power parities – measurement and uses", *Statistics Brief*, no 3, OECD.

Senhadji, A (2000): "Sources of Economic Growth: An Extensive Growth Accounting Exercise", *IMF Staff Papers*, vol 47, no 1.

Solow, R (1956): "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, vol 70, no 1.

Sorense, A and B Schjerning (2003): "Is It Possible to Measure Sectoral Productivity Levels? The Case of Manufacturing", Centre for Economic and Business Research.

Steindel, C (2004): "The relationship between Manufacturing Production and Goods Output", Current *Issues in Economics and Finance*, Federal Reserve Bank of New York, vol 10, no 9, August.

Stiroh, K J (2001): "What Drives Productivity Growth?", *Federal Reserve Bank of New York Policy Review*, March.

Turner, A (2003): "What's wrong with Europe's economy?", Speech, February 5.

van Ark, B (1993): "International comparisons of output and productivity", *Groningen Growth and Development Centre Monograph Series*, no 1.

van Ark, B (2001): "Issues in measurement and international comparison issues of productivity – an overview", OECD.

Vijselaar, F W (2003): "Capital Quality Improvement and the Sources of Growth in the Euro Area".

Wadhwani, S B (2000): "Monetary Challenges in a New Economy", Speech, October 12.

Wadhani, S B (2001): "Do we have a New Economy", Speech, September 14.

Wölfl, A (2003): "Productivity growth in service industries: an assessment of recent patterns and the role of measurement", *OECD STI Working Paper* 2003/7.

Abstract

There has been a growing interest in comparing productivity developments across industrial countries in recent years. The scope of the present paper, in this context, is relatively simple: it is to distinguish advanced economies where trend-productivity gains have improved from those where they have declined over the past years.

The reasons why productivity is a key issue are briefly summarised in Section I. The various difficulties encountered in conducting international comparisons are then described in Section II, which notes that comparing levels as well as changes in productivity can be rather misleading. By contrast, it is argued that focussing on changes in productivity changes (i.e. the second derivative) might be more informative when judging relative economic performances across countries. A general framework for determining trend productivity gains and searching for time breaks is therefore presented in Section III. Section IV also shows that productivity developments are significantly influenced by the state of the business cycle and Section V presents a general overview of trend-labour productivity gains in the main industrial economies.

Section VI deals with the impact of capital accumulation on labour productivity and provides a cross-country analysis of long-term developments in total factor productivity (TFP). A main feature is that the recent performance of the United States clearly stands out. In particular, it is shown that the level of US labour productivity is the highest among the major industrial countries and has been rising the fastest in the recent past. Rather than just reflecting stronger capital accumulation, this performance has been associated with a higher rate of technological progress that was maintained during the latest recession. More importantly, perhaps, the US performance has improved in relative terms, as total factor productivity growth has accelerated in the United States but decelerated in most other economies. These observations might have several policy implications, as argued in Section VII.

Les Skoczylas and Bruno Tissot (Bank for International Settlements)

Annex A

Data availability

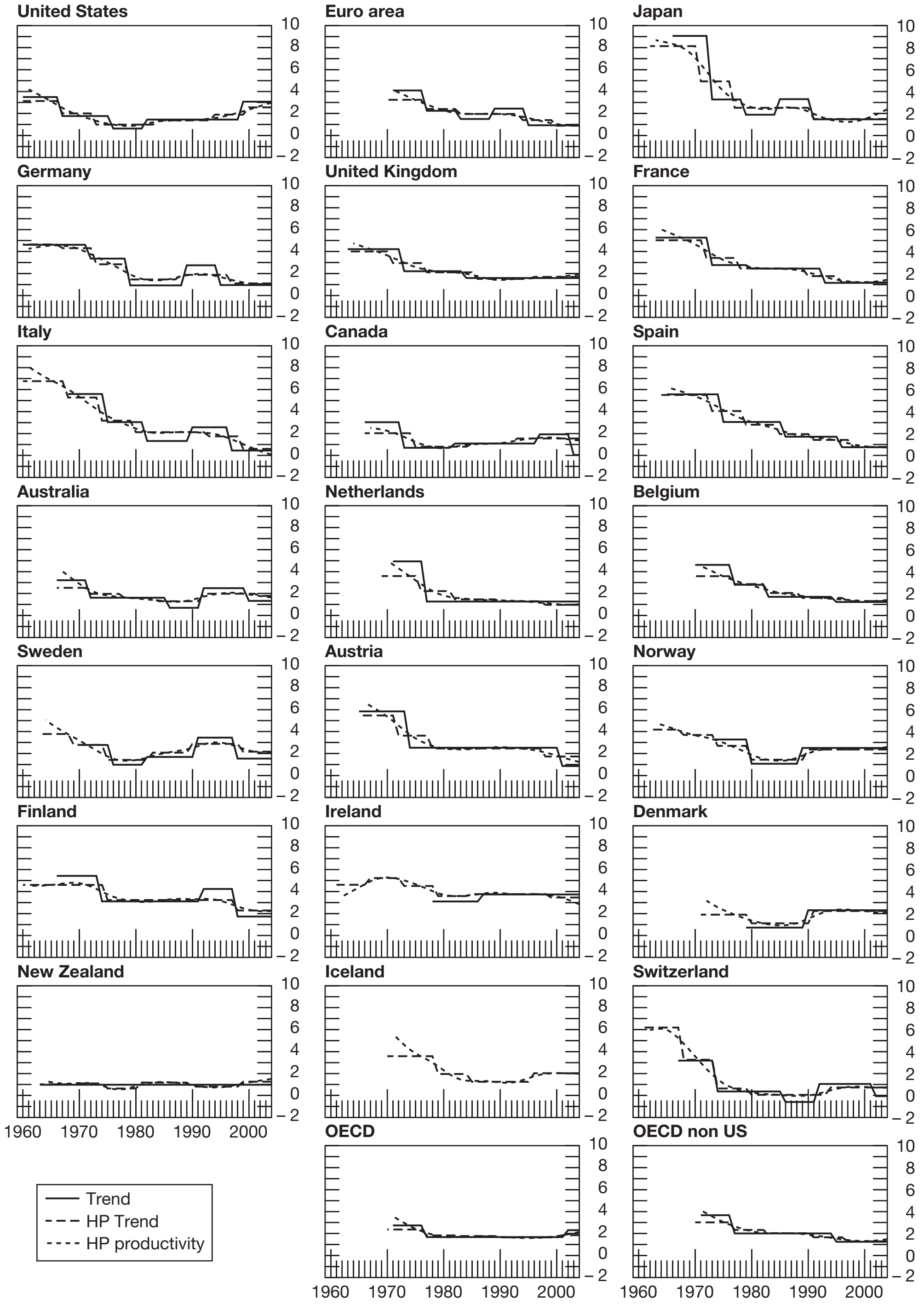
	Average labour share, in %, 1970–2002	Business sector			Manufacturing
		Output per person	Output per hours worked	Output per unit of capital	Capacity utilisation
Australia	56	1966	1970	1966	1965
Austria	62	1965	–	1960	1963
Belgium	60	1970	1970	1970	1962
Canada	62	1966	1966	1966	1961
Denmark	64	1971	1971	1971	1979
Finland	60	1960	1960	1975	1966
France	62	1963	1970	1963	1962
Germany	63	1960	1960	1960	1960
Iceland	62	1970	1970	1970	–
Ireland	55	1961	1970	1961	1978
Italy	51	1960	1960	1960	1968
Japan	56	1962	1970	1965	1966
Netherlands	59	1969	1970	1969	1971
New Zealand	54	1963	1970	1971	1961
Norway	56	1962	1962	1965	1973
Spain	55	1964	1970	1964	1965
Sweden	66	1963	1963	1965	1970
Switzerland	63	1961	1970	1961	1967
United Kingdom	65	1963	1970	1963	1960
United States	63	1960	1960	1960	1948
Euro area	59	1970	1970	1970	1971
OECD ex. US	59	1970	1970	1970	1971
OECD	61	1970	1970	1970	1971

See footnotes of Table 4 for the exact composition of country groups.

Sources: OECD (2004); national data; BIS calculations.

Annex B1

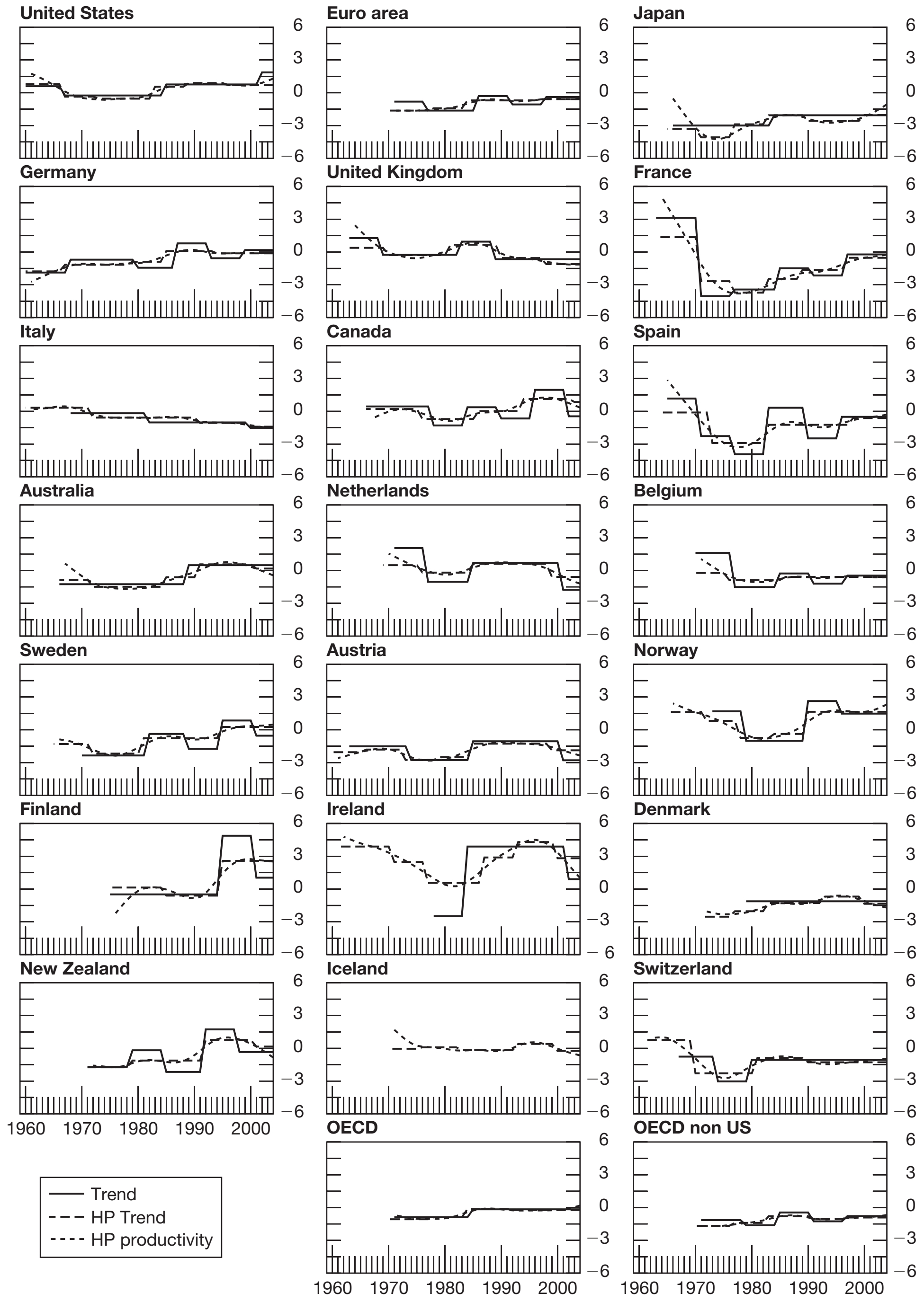
Labour productivity



Sources: OECD; BIS calculations.

Annex B2

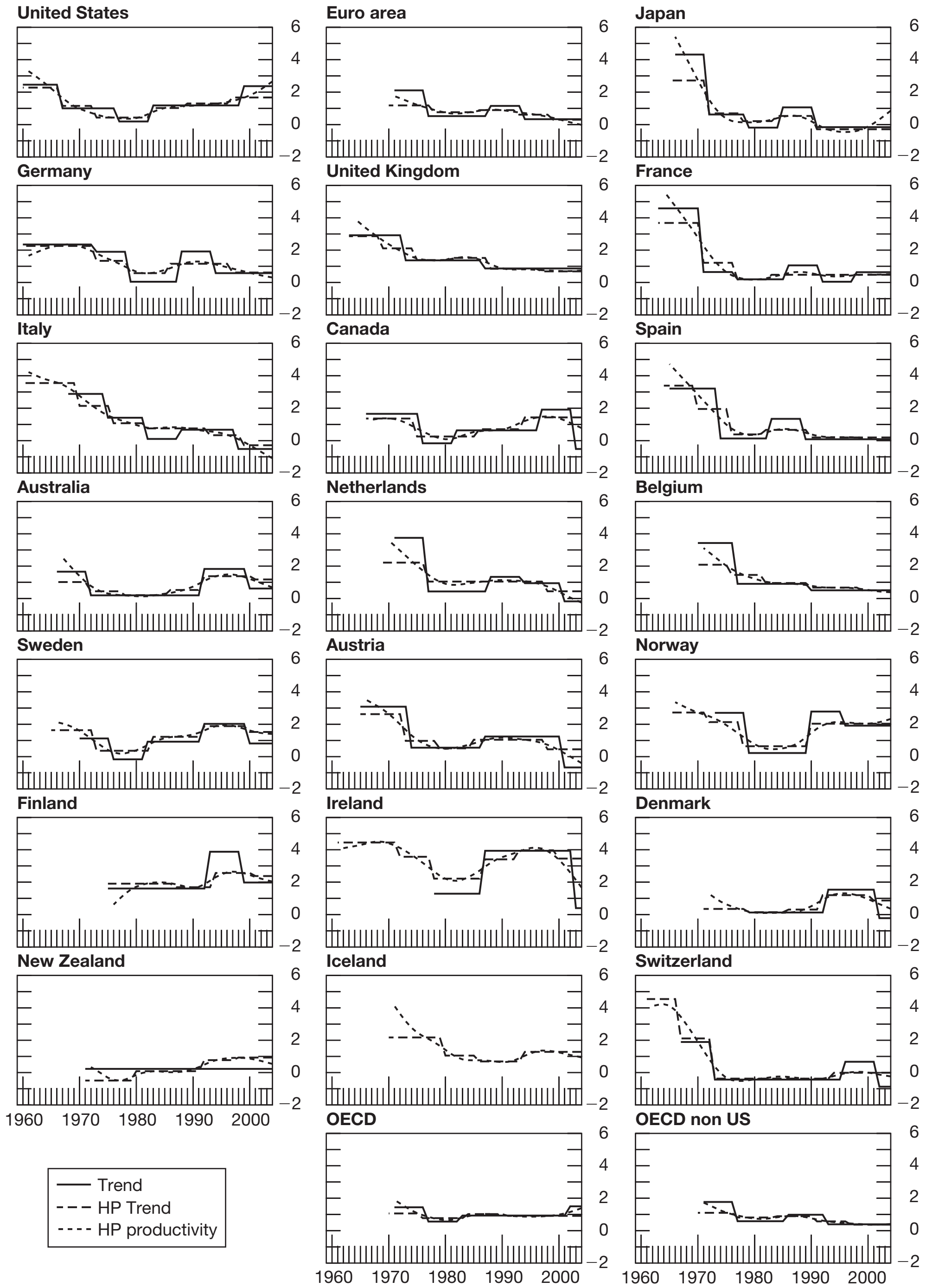
Capital productivity



Sources: OECD; BIS calculations.

Annex B3

Total factor productivity



Note: Results of the econometric estimation are not significant for Ireland.
Sources: OECD; BIS calculations.

Annex C: *Estimation results for the main industrial countries.*

C1 – Capital productivity, HP filtered data

T stat in brackets	Constant	Cycle	Time trend	Time breaks: additional change in trend productivity gains and respective break-years					F-stat
Australia	−0.1497 (−40.9)	−0.0085 (−24.5)	−0.0064 (−15.4) 1972	0.0088 (29.5) 1985	0.0118 (35) 1991	−0.0039 (−8.4) 2001			15973
Austria	−0.3444 (−184.9)	−0.0205 (−48.4)	0.0024 (3.9) 1966	−0.0095 (−19.4) 1973	0.0026 (5.3) 1979	0.0123 (36.9) 1985	−0.0061 (−18.9) 1999		111734
Belgium	−0.9974 (−307.6)	−0.0022 (−9.9)	−0.0064 (−22.9) 1976	0.0027 (25.1) 1986					39381
Canada	−0.5766 (−252.3)	0.0021 (11.8)	−0.0092 (−32.1) 1976	0.0072 (25.5) 1986	0.0112 (35.4) 1994	−0.0028 (−3.4) 2003			2597
Denmark	−0.6806 (−144)	−0.0253 (−80.7)	0.005 (10.7) 1977	0.0075 (25.7) 1983	0.0059 (23) 1993	−0.0066 (−17.6) 2000			57586
Finland	−1.4018 (−169.3)	0.0014 (3.7)	−0.0074 (−12.2) 1985	0.0318 (60.6) 1995					3240
France	−0.2779 (−43.6)	0.0136 (17.8)	−0.0403 (−28.9) 1971	−0.0105 (−7.4) 1977	0.0126 (9) 1983	0.0082 (6.9) 1989	0.0111 (11.5) 1997		14553
Germany	−0.6004 (−596.5)	−0.0177 (−93.9)	0.006 (24.7) 1967	0.0028 (13.5) 1979	0.0098 (42.6) 1985	−0.0021 (−12.8) 1994			47756
Iceland	−0.7345 (−277.2)	−0.0004 (−2.2)	0.0013 (4.8) 1976	−0.0026 (−16.2) 1983	0.0057 (36.9) 1993	−0.0064 (−28) 2000			335
Ireland	−1.8636 (−599.1)	0.0389 (95.9)	−0.0141 (−17) 1971	−0.0192 (−25.5) 1977	0.0233 (30.4) 1987	0.0141 (15.2) 1993	−0.0148 (−14.2) 2000		43257
Italy	−0.9222 (−1264.3)	0.0032 (38.8)	−0.009 (−82) 1972	−0.0049 (−40.6) 1991	−0.0034 (−13) 2000				49963
Japan	0.2901 (79.9)	−0.0332 (−83.5)	−0.0076 (−12.4) 1971	0.0119 (24) 1977	0.0083 (19.2) 1983	−0.0052 (−16.1) 1990	0.0054 (13.5) 2000		188189
Netherlands	−1.023 (−220.9)	0.0049 (14.8)	−0.0068 (−14.8) 1976	0.008 (32.5) 1985	−0.012 (−36.4) 1999				1672
New Zealand	−0.2103 (−42.9)	−0.0174 (−59.1)	0.0062 (16.1) 1979	0.0189 (65) 1992	−0.0061 (−10) 2001				7415
Norway	−0.8035 (−153.6)	0.0163 (30.1)	−0.0081 (−9) 1972	−0.0156 (−18.6) 1978	0.0036 (4.6) 1984	0.0203 (40.3) 1990			5054
Spain	−0.3267 (−62.5)	−0.0011 (−2.1)	−0.0279 (−37.4) 1973	0.0166 (35.4) 1983	0.0062 (9.9) 1998				16479
Sweden	−0.6052 (−183.5)	−0.013 (−39.1)	−0.0088 (−19.5) 1972	0.0141 (59.1) 1981	0.0103 (47.4) 1995				33244
Switzerland	−0.8051 (−217.7)	0.0077 (15.5)	−0.0306 (−42.9) 1970	0.0141 (24.8) 1981	−0.004 (−8.3) 1990				14243
UK	−0.7579 (−323.8)	0.0038 (13.3)	−0.0066 (−17.7) 1970	0.0095 (38.4) 1981	−0.0122 (−41.5) 1990	−0.0056 (−15.9) 1998			1934
United States	−0.2563 (−174.1)	0.0078 (26.4)	−0.0114 (−23.6) 1967	−0.0018 (−5.3) 1973	0.0109 (32.7) 1983	0.0034 (8.7) 1989	−0.002 (−6.4) 1996		4404
Euro area	0.4739 (241)	−0.0163 (−126.2)	0.0021 (9.4) 1978	0.0072 (46.3) 1984	0.0011 (8.7) 1997				94444
OECD ex.US	0.5184 (289)	−0.0167 (−136.4)	0.0027 (13.2) 1977	0.0061 (38.2) 1983	−0.0025 (−16.1) 1991	0.0013 (8.5) 1997			155999
OECD	0.2704 (151.2)	−0.0107 (−91.3)	0.0018 (8.4) 1978	0.0076 (43.3) 1984	−0.001 (−9) 1992				27242

C2 – Labour productivity (persons employed), HP filtered data

T stat in brackets	Constant	Cycle	Time trend	Time breaks: additional change in trend productivity gains and respective break-years					F-stat
Australia	10.3018 (4535.3)	0.0252 (111.9)	−0.0055 (−15.9) 1972	−0.0038 (−13.6) 1978	−0.0029 (−13) 1984	0.0068 (40.2) 1992	−0.0019 (−5.7) 2002		215637
Austria	9.2303 (1921.5)	0.0547 (110.8)	−0.0184 (−23.7) 1972	−0.0117 (−27.9) 1978	−0.0073 (−20.7) 1998				119847
Belgium	9.8448 (2818.3)	0.0358 (149.7)	−0.0072 (−18) 1977	−0.0081 (−22.1) 1983	−0.0045 (−14.9) 1989	−0.0029 (−11.8) 1997			159339
Canada	10.4438 (3909.7)	0.0203 (92.6)	−0.0123 (−37.3) 1975	0.0029 (10.1) 1985	0.0046 (17.4) 1993				46871
Denmark	12.1749 (3226.3)	0.019 (86.5)	−0.0079 (−25.4) 1980	0.0112 (56) 1991					76278
Finland	9.3882 (6938.8)	0.0461 (376.2)	−0.0138 (−81.9) 1975	−0.0095 (−32.9) 1998					454145
France	9.5869 (3674.1)	0.0503 (172.7)	−0.0161 (−29.3) 1972	−0.0096 (−22.5) 1978	−0.007 (−16.8) 1990	−0.0057 (−12) 1996			136712
Germany	9.65 (5034.6)	0.0465 (120.1)	−0.0034 (−5.3) 1967	−0.0147 (−25.1) 1973	−0.0138 (−30.1) 1979	0.0042 (12.3) 1988	−0.0078 (−21.4) 1997		133301
Iceland	13.7081 (2807.6)	0.0358 (115.6)	−0.0162 (−27.7) 1979	−0.0071 (−15.3) 1985	0.0078 (22.1) 1996				47050
Ireland	8.9673 (4528.6)	0.0462 (122.9)	0.0061 (10.5) 1967	−0.0072 (−15.1) 1973	−0.0092 (−20.9) 1979	0.0018 (6.4) 1985	−0.0031 (−10.9) 1999		614259
Italy	9.1302 (3666.7)	0.0677 (150.9)	−0.0149 (−18.3) 1968	−0.0211 (−26.6) 1974	−0.0105 (−19.3) 1980	−0.0038 (−7.1) 1993	−0.0114 (−14.7) 1999		111210
Japan	14.1662 (3070.1)	0.0814 (141.4)	−0.0321 (−30.1) 1971	−0.0241 (−31.6) 1977	−0.0102 (−24.9) 1991				74295
Netherlands	9.8378 (2166.1)	0.0359 (107.9)	−0.0137 (−25) 1976	−0.0075 (−17.4) 1982	−0.0017 (−4.8) 1990	−0.0033 (−8.9) 1998			58550
New Zealand	10.4531 (9198.3)	0.0102 (64.4)	0.0008 (3.1) 1969	−0.0046 (−23.6) 1975	0.0054 (36) 1981	−0.0037 (−32.2) 1990	0.0045 (30.9) 1999		174402
Norway	11.991 (4772.4)	0.0418 (103.3)	−0.0049 (−7.8) 1968	−0.0099 (−20.2) 1974	−0.0126 (−35.5) 1980	0.0093 (47.5) 1990			205642
Spain	9.0047 (3005.1)	0.0551 (181.7)	−0.0146 (−24.4) 1973	−0.0125 (−18.7) 1979	−0.0085 (−12.8) 1985	−0.0053 (−8.7) 1991	−0.0064 (−11.8) 1998		108078
Sweden	12.0315 (3302.9)	0.0378 (74.6)	−0.0099 (−12.8) 1969	−0.0137 (−25.8) 1975	0.0066 (14.4) 1983	0.008 (18.7) 1990	−0.0071 (−15.1) 1999		93573
Switzerland	10.6038 (2474.5)	0.0619 (84.8)	−0.0293 (−24) 1968	−0.0262 (−24.4) 1974	−0.0058 (−8.4) 1980	0.0067 (15.8) 1994			12785
UK	9.3786 (4715.2)	0.0401 (168.5)	−0.0106 (−25.2) 1971	−0.0084 (−24.9) 1977	−0.006 (−25.2) 1986	0.0018 (7.5) 1996			195725
United States	10.3633 (6622.3)	0.0314 (99.7)	−0.0113 (−22.1) 1967	−0.0102 (−30) 1973	0.0039 (16.7) 1984	0.0052 (14.6) 1994	0.0066 (12.8) 2000		88234
Euro area	−1.0108 (−252.2)	0.0325 (118.7)	−0.0084 (−18.8) 1977	−0.0045 (−15) 1983	−0.0056 (−18.5) 1994	−0.0051 (−11.5) 2000			107078
OECD ex.US	−0.974 (−338.2)	0.0302 (153.6)	−0.0069 (−21.1) 1977	−0.003 (−11.6) 1983	−0.0037 (−14.7) 1991	−0.0037 (−15.3) 1997			208875
OECD	−0.8216 (−437.6)	0.0237 (184.8)	−0.0054 (−25.2) 1977	−0.0006 (−3.6) 1983	−0.0012 (−9.7) 1990	0.002 (10.7) 2001			402023

C3 – Labour productivity (hours worked), HP filtered data

T stat in brackets	Constant	Cycle	Time trend	Time breaks: additional change in trend productivity gains and respective break-years					F-stat
Australia	2.7904 (1634.8)	0.0206 (222.6)	−0.0062 (−37.1) 1983	0.0062 (38.7) 1993					193148
Belgium	2.1809 (388.5)	0.0466 (116.7)	−0.0088 (−14.4) 1976	−0.0127 (−26.2) 1982	−0.0033 (−9.2) 1988	−0.0029 (−9.2) 1998			145891
Canada	2.7687 (774.8)	0.0306 (91.6)	−0.0129 (−24) 1973	−0.0064 (−19.5) 1979	0.0036 (19.8) 1992				69441
Denmark	4.4576 (1113.6)	0.0345 (134.6)	−0.0082 (−19.4) 1978	−0.0083 (−24.7) 1984	0.0046 (14) 1992	0.0028 (8.2) 1998			156679
Finland	1.747 (1007.1)	0.0485 (126.7)	0.0059 (11.5) 1966	−0.0143 (−37.9) 1975	−0.0031 (−9.1) 1981	−0.0046 (−13.3) 1992	−0.0081 (−19) 1998		539979
France	2.1497 (636.1)	0.0403 (182)	−0.0044 (−11.7) 1978	−0.0066 (−18.4) 1985	−0.0092 (−35.4) 1991				283807
Germany	1.9995 (870)	0.0564 (121.5)	−0.0027 (−3.4) 1967	−0.0155 (−22) 1973	−0.0168 (−30.5) 1979	0.0046 (11.3) 1988	−0.0095 (−21.7) 1997		166807
Iceland	5.8841 (813.9)	0.0499 (108.8)	−0.0246 (−28.6) 1979	−0.0123 (−18.4) 1985	0.0076 (14.2) 1997				30674
Ireland	1.2656 (350.6)	0.0531 (239.8)	−0.0114 (−37.8) 1979	0.0054 (22.5) 1992	−0.0056 (−8.7) 2002				411633
Italy	1.5272 (579)	0.0745 (156.8)	−0.0137 (−15.8) 1968	−0.0202 (−24) 1974	−0.0163 (−28.2) 1980	−0.0051 (−8.9) 1993	−0.0106 (−12.9) 1999		134923
Japan	6.9637 (1263.9)	0.0415 (111.9)	−0.012 (−23.1) 1977	0.0026 (7.1) 1986	−0.0107 (−32.5) 1994				142047
Netherlands	2.3822 (455.2)	0.0346 (93.6)	−0.0066 (−12.5) 1976	0.0044 (13.2) 1983	−0.0114 (−31.7) 1992	−0.0064 (−16) 1998			161750
New Zealand	2.9243 (1146.8)	0.0092 (61.9)	0.0045 (17.2) 1981	−0.0052 (−19.8) 1990	0.005 (14.5) 1999				34954
Norway	4.3675 (2143.9)	0.0528 (268.6)	−0.0128 (−26.3) 1975	−0.0193 (−40.8) 1981	0.0079 (26.6) 1992				222159
Spain	1.4961 (252.3)	0.0456 (113.6)	−0.0062 (−10.4) 1977	−0.014 (−25.7) 1985	−0.0096 (−15.5) 1991	−0.0076 (−14.3) 1997			84351
Sweden	4.4564 (1450.1)	0.0493 (143.9)	−0.0177 (−27.9) 1972	−0.0131 (−29.4) 1978	0.0075 (28.7) 1993				119652
Switzerland	3.5426 (1246.6)	0.0168 (93.1)	−0.006 (−17.5) 1979	−0.0065 (−23) 1985	0.0055 (27) 1995				30140
UK	1.9097 (642.6)	0.0329 (174.7)	−0.0058 (−16.2) 1979	−0.0116 (−36.9) 1985	0.0032 (14.7) 1994				161930
United States	2.7579 (1758.3)	0.0332 (131.1)	−0.0105 (−22.6) 1969	−0.0102 (−32.8) 1975	0.0052 (16.5) 1993	0.0082 (16) 1999			100460
Euro area	−5.9088 (−1355.2)	0.0412 (138.3)	−0.009 (−18.3) 1977	−0.006 (−16.5) 1983	−0.005 (−13.9) 1992	−0.0068 (−17) 1998			158592
OECD ex.US	−5.85 (−1349.7)	0.0389 (126.3)	−0.0079 (−17.2) 1976	−0.0057 (−20.9) 1982	−0.0048 (−17.9) 1993	−0.0037 (−10.3) 1999			216660
OECD	−5.6207 (−2194.9)	0.0306 (174.6)	−0.0076 (−26.4) 1977	−0.0028 (−13.2) 1983	−0.0008 (−4.8) 1992	0.0015 (6.2) 2000			308183

C4 – Total factor productivity (persons employed), HP filtered data

T stat in brackets	Constant	Cycle	Time trend	Time breaks: additional change in trend productivity gains and respective break-years					F-stat
Australia	5.6963 (2326.2)	0.0102 (42.3)	−0.0057 (−15.5) 1972	−0.0024 (−9.4) 1978	0.0032 (13.1) 1986	0.0085 (36.4) 1992	−0.002 (−5.3) 2001		29322
Austria	5.5498 (1640.3)	0.0262 (79.6)	−0.0165 (−27.6) 1973	−0.0046 (−7.9) 1979	0.0053 (13.4) 1985	−0.0059 (−15.2) 1998			18696
Belgium	5.4962 (1844.7)	0.0209 (98.6)	−0.0063 (−19.7) 1976	−0.005 (−23) 1982	−0.0028 (−16.5) 1990	−0.0017 (−7.8) 1999			67619
Canada	6.2047 (2509)	0.0136 (67.4)	−0.0111 (−37.9) 1975	0.0045 (17.3) 1986	0.0073 (26.5) 1994				20158
Denmark	7.4954 (3013.2)	0.0035 (23)	−0.002 (−7.9) 1979	0.0016 (6.3) 1986	0.0089 (38.3) 1992	−0.0033 (−9) 2001			17578
Finland	5.167 (3471.7)	0.0191 (287.8)	−0.0022 (−14.7) 1987	0.0088 (47.6) 1994	−0.0019 (−6.2) 2001				289650
France	5.7855 (1405.5)	0.0368 (74.8)	−0.0246 (−27.5) 1971	−0.0103 (−12.1) 1977	0.0029 (5.9) 1983				9525
Germany	5.842 (6308.6)	0.0226 (208.4)	−0.0092 (−32.6) 1973	−0.0076 (−22.3) 1979	0.0058 (23.1) 1986	−0.0054 (−21.8) 1997			89561
Iceland	8.2342 (2996.8)	0.0217 (128.6)	−0.0112 (−32) 1980	−0.0036 (−10.6) 1986	0.0059 (22.9) 1994				42865
Ireland	4.0741 (2336)	0.0445 (210.7)	−0.0088 (−18.7) 1972	−0.0135 (−28.6) 1978	0.0119 (24.5) 1987	0.0056 (10) 1993	−0.0051 (−7.1) 2000		210440
Italy	4.2135 (2861.4)	0.0355 (162.6)	−0.0141 (−30.5) 1970	−0.0107 (−20.2) 1976	−0.0032 (−8) 1982	−0.0042 (−10.5) 1993	−0.0062 (−10.5) 1999		46342
Japan	8.1279 (1885.5)	0.0272 (61.2)	−0.0203 (−27.3) 1972	−0.0048 (−7) 1978	0.0032 (5.3) 1984	−0.0082 (−21.1) 1991			3932
Netherlands	5.415 (1539.6)	0.0222 (95.2)	−0.0117 (−43.3) 1977	−0.006 (−22.1) 1998					37746
New Zealand	5.5577 (3102.2)	−0.0049 (−46.8)	0.0058 (39.1) 1980	0.0068 (40.2) 1992	0.0016 (6.6) 1998				10090
Norway	6.3779 (1531.2)	0.0272 (63.5)	−0.0059 (−8.6) 1972	−0.0149 (−34.1) 1978	0.0139 (58.5) 1990				43389
Spain	4.7966 (1018.6)	0.0339 (58.6)	−0.0144 (−16.2) 1970	−0.0157 (−22) 1976	0.0028 (4.9) 1982	−0.0047 (−13.6) 1990			10746
Sweden	7.7821 (2997.8)	0.0163 (66.6)	−0.0127 (−36.5) 1973	0.0086 (32.4) 1983	0.0066 (21.2) 1992	−0.0036 (−7) 2000			37462
Switzerland	6.392 (1308.5)	0.0454 (49.2)	−0.0243 (−18.2) 1967	−0.0248 (−39.8) 1973	0.0036 (9.6) 1994				2908
UK	5.8325 (2643)	0.0286 (92.9)	−0.0075 (−15.9) 1969	−0.0072 (−18.7) 1975	0.0013 (4) 1981	−0.0067 (−25.6) 1988	−0.0016 (−6.3) 1997		88624
United States	6.4664 (4844.4)	0.0228 (84.6)	−0.0113 (−25.8) 1967	−0.0072 (−23.1) 1973	0.0059 (19.8) 1983	0.0028 (8.3) 1989	0.0037 (12.1) 1997		51228
Euro area	−0.3908 (−240.2)	0.0118 (111.7)	−0.0042 (−25.4) 1978	0.0013 (8.3) 1986	−0.0028 (−14.9) 1993	−0.0031 (−13.9) 1999			64792
OECD ex.US	−0.3612 (−246.3)	0.011 (109.9)	−0.0029 (−18.6) 1977	0.0008 (6) 1984	−0.0032 (−23.1) 1991	−0.002 (−14.8) 1997			100399
OECD	−0.3989 (−303.1)	0.0106 (114.6)	−0.0029 (−23.1) 1976	0.0025 (29.2) 1984	−0.0011 (−16.7) 1991				331652

C5 – Total factor productivity (hours worked), HP filtered data

T stat in brackets	Constant	Cycle	Time trend	Time breaks: additional change in trend productivity gains and respective break-years					F-stat
Australia	1.5419 (1024.5)	0.0042 (55)	0.0013 (5.7) 1986	0.0085 (35.7) 1992					32497
Belgium	0.9075 (241.3)	0.027 (101.1)	−0.0079 (−19.5) 1976	−0.0075 (−29.1) 1982	−0.0015 (−8) 1991	−0.0018 (−4.6) 2001			76846
Canada	1.4749 (356.6)	0.0206 (50.3)	−0.0082 (−13.2) 1972	−0.0088 (−19.3) 1978	0.0038 (11) 1985	0.0066 (24.5) 1994			21501
Denmark	2.5892 (1302.8)	0.0129 (101.3)	−0.003 (−14.3) 1978	−0.0028 (−18.5) 1984	0.0054 (45) 1993	−0.001 (−3.2) 2002			103702
Finland	0.6187 (379.5)	0.0225 (310.4)	−0.0039 (−25.5) 1987	0.0072 (46.5) 1994					378932
France	1.4671 (2193.1)	0.0092 (236.9)	0.001 (15) 1981	−0.0032 (−43.2) 1990	0.003 (31.5) 1998				313554
Germany	1.0326 (978)	0.0292 (235.3)	−0.0093 (−29.1) 1973	−0.0096 (−24.8) 1979	0.0061 (21.1) 1986	−0.006 (−21.3) 1997			137132
Iceland	3.3831 (828.6)	0.0301 (120.2)	−0.017 (−32.7) 1980	−0.006 (−11.8) 1986	0.0054 (14.1) 1994				27990
Ireland	−0.042 (−9.3)	0.0381 (130.1)	−0.0119 (−28.2) 1978	0.0125 (29.2) 1988	0.007 (13.2) 1994	−0.0067 (−8.3) 2001			163580
Italy	0.3286 (206.9)	0.0396 (153.2)	−0.011 (−21.5) 1969	−0.0129 (−23.6) 1975	−0.0063 (−14.9) 1981	−0.0037 (−9.9) 1991	−0.0067 (−13.6) 1998		63574
Japan	4.0891 (3223.3)	0.0047 (69.8)	0.0039 (28.7) 1984	−0.0074 (−51.9) 1993					23834
Netherlands	1.005 (252.4)	0.0216 (77.2)	−0.0057 (−15) 1976	0.0054 (23.2) 1984	−0.0078 (−27.6) 1993	−0.0069 (−18.1) 1999			101231
New Zealand	1.4869 (835.6)	−0.0038 (−36.4)	0.0059 (41.3) 1980	0.0066 (39.7) 1993	0.001 (3.7) 1999				12797
Norway	2.0948 (526.5)	0.0361 (93.4)	−0.0077 (−11.4) 1973	−0.0179 (−37.9) 1979	0.0126 (46.2) 1991				63505
Spain	0.846 (237)	0.0119 (46.8)	−0.0029 (−7.5) 1976	0.0031 (10.6) 1982	−0.0081 (−29.2) 1989	−0.0025 (−10.5) 1995			26158
Sweden	2.7597 (869.1)	0.0261 (84.5)	−0.0152 (−27.1) 1973	−0.0032 (−5.7) 1979	0.0028 (6.3) 1985	0.0071 (22.7) 1994			33638
Switzerland	2.1025 (1421.4)	−0.0009 (−9.6)	0.0034 (24) 1978	−0.0041 (−38.1) 1987	0.0033 (28.6) 1996				653
UK	1.0062 (644.8)	0.0201 (217.1)	−0.0024 (−11.7) 1981	−0.0082 (−44.4) 1987	−0.0011 (−5.8) 1998				143487
United States	1.6517 (1332.7)	0.0239 (107.4)	−0.01 (−25.6) 1968	−0.0072 (−25.1) 1974	0.0045 (19.1) 1985	0.003 (8.8) 1993	0.0038 (8.5) 1999		63887
Euro area	−3.263 (−1842.4)	0.0169 (150.5)	−0.0047 (−33.1) 1978	−0.0033 (−22.4) 1993	−0.0028 (−11.5) 1999				132514
OECD ex.US	−3.2301 (−1894.2)	0.0156 (138.9)	−0.004 (−29.7) 1977	−0.0032 (−26.4) 1992	−0.0023 (−12.7) 1998				163858
OECD	−3.3131 (−2290.2)	0.0144 (146.7)	−0.0041 (−28.4) 1977	0.0017 (13.4) 1985	−0.001 (−9.9) 1991				291274

C6 – Capital productivity, adjusted for the cycle

T stat in brackets	Constant	Cycle	Time trend	Time breaks: additional change in trend productivity gains and respective break-years					F-stat
Australia	−0.1224 (−7.7)	0.0106 (2.2)	−0.0125 (−16.5)	0.0174 (12.3) 1989					204
Austria	−0.4037 (−38.8)	0.0161 (6.1)	−0.0151 (−16.5)	−0.0126 (−8.9) 1974	0.0172 (16.9) 1985	−0.0174 (−5.9) 2001			3348
Belgium	−1.2653 (−46.9)	0.0238 (10.2)	0.0162 (8.9)	−0.0313 (−13.4) 1977	0.0124 (7.4) 1985	−0.0093 (−4.9) 1991	0.0074 (4.2) 1997		467
Canada	−0.6093 (−54.5)	0.0248 (12)	0.0044 (5.4)	−0.0174 (−8.3) 1978	0.0167 (5.9) 1984	−0.0103 (−3.7) 1990	0.0262 (9) 1996	−0.0242 (−4.6) 2002	99
Denmark	−0.9748 (−49.4)	−0.0202 (−4.5)	−0.0112 (−18.9)						180
Finland	−1.2743 (−53.6)	0.0555 (8.6)	−0.0049 (−5.5)	0.0537 (13.5) 1995	−0.0384 (−4.3) 2001				89
France	−0.4121 (−40.2)	0.0083 (5.3)	0.0312 (24.8)	−0.0717 (−31) 1971	0.0062 (3.2) 1977	0.0194 (10.2) 1985	−0.0066 (−3) 1991	0.0193 (9.4) 1997	4957
Germany	−0.608 (−97.2)	0.0193 (12.8)	−0.0188 (−18.4)	0.0118 (8.5) 1968	−0.0074 (−5.6) 1980	0.0223 (12.1) 1987	−0.0136 (−6.5) 1993	0.0075 (3.3) 1999	826
Ireland	−0.7678 (−4.5)	−0.025 (−2.1)	−0.0248 (−3.4)	0.0637 (8.4) 1984	−0.0299 (−2.3) 2002				276
Italy	−0.8725 (−99.6)	0.0216 (12.3)	−0.0018 (−3.5)	−0.0084 (−10.4) 1982	−0.0053 (−3.4) 2000				1137
Japan	0.1979 (16.7)	0.0439 (13.6)	−0.03 (−49.8)	0.0094 (11.4) 1984					6517
Netherlands	−1.2439 (−51.3)	0.0185 (11.3)	0.0206 (12.6)	−0.0309 (−14.3) 1977	0.0171 (22.8) 1985	−0.0244 (−15.5) 2001			267
New Zealand	−0.2258 (−4.8)	0.021 (4.9)	−0.017 (−5.9)	0.0152 (3.1) 1979	−0.0198 (−4.5) 1985	0.0389 (8.6) 1992	−0.0207 (−4.7) 1998		114
Norway	−0.8698 (−13.2)	0.0227 (4.6)	0.0169 (4.7)	−0.027 (−6.7) 1979	0.0364 (15.1) 1990	−0.0115 (−4) 1996			189
Spain	−0.4358 (−22.1)	0.0152 (6)	0.0116 (5.3)	−0.0343 (−10.1) 1971	−0.0166 (−6.1) 1977	0.0425 (17.2) 1983	−0.028 (−11.8) 1990	0.0197 (8.7) 1996	1940
Sweden	−0.4506 (−27.2)	0.0217 (6.2)	−0.0235 (−22.8)	0.0197 (7.3) 1982	−0.0136 (−4.2) 1989	0.0259 (7.7) 1995	−0.0139 (−3) 2001		445
Switzerland	−0.6802 (−21.9)	0.0277 (10)	−0.0076 (−2.9)	−0.0227 (−5.7) 1974	0.0197 (9.7) 1980				1015
UK	−0.8237 (−53.1)	0.0211 (8.8)	0.0128 (6.3)	−0.0154 (−6.4) 1969	0.0121 (5.8) 1983	−0.0161 (−7.8) 1989			94
United States	−0.2618 (−39.3)	0.0253 (15.1)	0.006 (4.8)	−0.0085 (−5.7) 1967	0.0101 (17.5) 1985	0.011 (3.5) 2002			286
Euro area	0.3542 (20.7)	0.0194 (14.7)	−0.0081 (−7.2)	−0.0082 (−5.7) 1977	0.0133 (13.2) 1986	−0.0077 (−6.2) 1992	0.0068 (5.6) 1998		1990.039
OECD ex.US	0.4413 (48.1)	0.0219 (18.5)	−0.0115 (−20.1)	−0.0047 (−4.8) 1979	0.0117 (11.9) 1985	−0.0081 (−7.9) 1991	0.0048 (5.4) 1997		4228.469
OECD	0.2438 (39.5)	0.0233 (15.8)	−0.0088 (−29.1)	0.0072 (16.7) 1985					807.5005

C7 – Labour productivity (persons employed), adjusted for the cycle

T stat in brackets	Constant	Cycle	Time trend	Time breaks: additional change in trend productivity gains and respective break-years					F-stat
Australia	10.2395 (463.5)	−0.0065 (−2.3)	0.0321 (14.8)	−0.0159 (−6) 1972	−0.0092 (−5.2) 1986	0.0179 (7.7) 1992	−0.0116 (−3.8) 2000		1857
Austria	9.1788 (793.1)	0.0085 (3.8)	0.0584 (63.4)	−0.0331 (−32.4) 1974	−0.0167 (−7.5) 2001				11213
Belgium	9.6918 (365.1)	0.0133 (5.9)	0.0461 (25.5)	−0.0181 (−7.3) 1977	−0.011 (−8.5) 1983	−0.0047 (−5.1) 1995			5574
Canada	10.3434 (695.1)	0.01 (6.2)	0.0303 (22.3)	−0.0233 (−12.4) 1973	0.0039 (3.8) 1982	0.0084 (5.7) 1997	−0.0188 (−3.1) 2003		1587
Denmark	12.4317 (490.7)	−0.0077 (−3)	0.0072 (7.5)	0.0158 (11.4) 1990					1005
Finland	9.2905 (578)	0.0157 (5.6)	0.0543 (40.8)	−0.0232 (−14.5) 1974	0.0113 (6.4) 1992	−0.0251 (−9.3) 1998			8067
France	9.5631 (2123.3)	0.0043 (5.2)	0.0527 (111.2)	−0.025 (−28) 1973	−0.0032 (−4.7) 1980	−0.013 (−26.7) 1993			36608
Germany	9.6417 (2302.2)	0.0079 (6)	0.0462 (90.9)	−0.0126 (−10.9) 1972	−0.0244 (−18.8) 1979	0.0183 (14.2) 1989	−0.018 (−14) 1995		15238
Ireland	9.3229 (128.8)	0.0152 (2.1)	0.031 (10.9)	0.0064 (2.1) 1987					1909
Italy	9.206 (617.9)	0.0227 (14.1)	0.056 (49.1)	−0.0256 (−17.3) 1975	−0.0172 (−13.8) 1982	0.0125 (10.5) 1990	−0.0212 (−18.6) 1997		11488
Japan	14.0378 (637.2)	0.0199 (6.3)	0.0907 (48.2)	−0.0579 (−19.4) 1973	−0.0139 (−4.6) 1979	0.0142 (5.4) 1985	−0.0184 (−10.3) 1991		5548
Netherlands	9.6498 (309)	0.0142 (6.2)	0.0493 (24.6)	−0.0366 (−16.7) 1977					2839
New Zealand	10.4547 (1172.8)	0.0129 (3.3)	0.0098 (30)						451
Norway	12.0704 (235.2)	0.009 (2.1)	0.0328 (12)	−0.0221 (−6.6) 1980	0.0144 (9.8) 1989				1575
Spain	8.9917 (907.4)	0.0045 (2.1)	0.0557 (66.4)	−0.0252 (−18.8) 1975	−0.0134 (−12.1) 1986	−0.0096 (−7) 1996			9064
Sweden	12.1451 (490.2)	0.0201 (9.5)	0.0278 (15.6)	−0.018 (−6.9) 1976	0.0071 (4.1) 1982	0.0176 (11.7) 1991	−0.0191 (−10.6) 1998		4192
Switzerland	10.8579 (726.8)	0.0122 (8.3)	0.0318 (25.8)	−0.0281 (−18.1) 1974	−0.0096 (−7.7) 1986	0.0164 (11.5) 1992	−0.0111 (−3.6) 2002		459
UK	9.3567 (746.4)	0.0071 (3.2)	0.0422 (33.4)	−0.0201 (−10.9) 1973	−0.0062 (−5.9) 1984				3911
United States	10.3434 (1246.1)	0.0083 (3.6)	0.035 (19.7)	−0.0173 (−6.7) 1967	−0.0114 (−6.2) 1976	0.0082 (5.7) 1982	0.0163 (8.5) 1999		3070
Euro area	−1.1333 (−73.9)	0.0124 (9.7)	0.041 (40.2)	−0.0185 (−12.8) 1977	−0.0075 (−6.2) 1983	0.0095 (7.8) 1989	−0.0152 (−16.9) 1995		10874
OECD ex.US	−1.0645 (−71.2)	0.0116 (8)	0.0367 (38.1)	−0.0166 (−15.4) 1977	−0.0076 (−15.5) 1995				13231
OECD	−0.8707 (−57.5)	0.0111 (7.3)	0.0274 (28.4)	−0.0106 (−10.1) 1977	0.0063 (3.5) 2002				8717

C8 – Labour productivity (hours worked), adjusted for the cycle

T stat in brackets	Constant	Cycle	Time trend	Time breaks: additional change in trend productivity gains and respective break-years					F-stat
Australia	2.8027 (194.1)	−0.0059 (−2)	0.0201 (28.8)	−0.014 (−8.3) 1986	0.0193 (8.6) 1992	−0.009 (−3) 2000			1940
Belgium	2.034 (38.3)	0.0098 (2.5)	0.0566 (15.2)	−0.0167 (−3.6) 1976	−0.019 (−10.2) 1982				3633
Canada	2.6858 (147.5)	0.0071 (3.6)	0.0386 (24.7)	−0.0257 (−15.3) 1973					2657
Denmark	4.768 (268.9)	−0.0088 (−2.9)	0.0193 (32.8)	0.0093 (4.4) 1998					1059
Finland	1.5894 (105.8)	0.0108 (4.3)	0.065 (52.6)	−0.0278 (−19.6) 1974	−0.0153 (−10.9) 1998				14180
France	2.0483 (98)	0.0075 (5.9)	0.0473 (32.6)	−0.0115 (−6.2) 1976	−0.0069 (−6.5) 1985	−0.0125 (−10.8) 1993	0.0108 (4.9) 2001		11719
Germany	1.9934 (608)	0.0079 (7.4)	0.0557 (149.2)	−0.0125 (−12.7) 1973	−0.0275 (−24.1) 1979	0.02 (19.1) 1989	−0.02 (−19.2) 1995		41999
Ireland	1.6476 (23.1)	0.0075 (1)	0.035 (12.8)	0.0125 (4.1) 1988					2228
Italy	1.5645 (99.2)	0.0172 (10.7)	0.0661 (54.4)	−0.0278 (−16.6) 1975	−0.0194 (−15.5) 1981	0.0095 (8) 1991	−0.0211 (−16.1) 1997		14545
Japan	6.635 (193.4)	0.0214 (6.9)	0.0648 (27.5)	−0.0409 (−14.3) 1976	0.0169 (10.7) 1987	−0.0208 (−14.1) 1994			7040
Netherlands	2.2214 (56.8)	0.0077 (2.9)	0.0462 (17.6)	−0.0288 (−7.9) 1977	0.018 (10.1) 1983	−0.0203 (−15.7) 1993			5165
New Zealand	2.8521 (136.1)	0.0183 (3.4)	0.0134 (14.1)	−0.0065 (−2.8) 1991					222
Norway	4.3216 (105.6)	0.0081 (2.4)	0.0541 (24.8)	−0.0387 (−14.5) 1980	0.014 (11.9) 1989				4287
Spain	1.4162 (38.6)	0.0078 (2.4)	0.0506 (20.3)	−0.0097 (−3.2) 1976	−0.0207 (−12.6) 1986	−0.0137 (−7.2) 1996			3289
Sweden	4.5441 (142.1)	0.014 (5)	0.0423 (18.7)	−0.0259 (−9.6) 1976	0.0081 (8.6) 1991				3353
Switzerland	3.3917 (106.1)	0.0122 (5.9)	0.0276 (12.2)	−0.0165 (−5.6) 1976	−0.0098 (−7) 1985	0.0163 (8.9) 1995	−0.0156 (−4.5) 2001		540
UK	1.9508 (98.8)	−0.0082 (−2.2)	0.0298 (31.5)	−0.0139 (−9.5) 1986					1287
United States	2.7587 (466)	0.0077 (4.1)	0.0315 (53)	−0.0186 (−24.8) 1973	0.0176 (13.2) 1998				5252
Euro area	−5.9302 (−447.5)	0.005 (3.3)	0.0426 (53.3)	−0.0167 (−17.1) 1979	−0.0117 (−17) 1996				14483
OECD ex.US	−5.9212 (−316.7)	0.0099 (5.2)	0.0439 (35.4)	−0.0163 (−9.2) 1977	−0.0054 (−3.6) 1983	0.0064 (4.1) 1989	−0.0121 (−10.8) 1995		11672
OECD	−5.6718 (−328.8)	0.0048 (2.8)	0.0344 (31.6)	−0.0143 (−12.3) 1977					12128

C9 – Total factor productivity (persons employed), adjusted for the cycle

T stat in brackets	Constant	Cycle	Time trend	Time breaks: additional change in trend productivity gains and respective break-years					F-stat
Australia	5.6252 (216.5)	−0.0074 (−2.1)	0.0183 (7.1)	−0.0176 (−5.4) 1972	0.0034 (2.2) 1984	0.0181 (6.8) 1994	−0.0176 (−4.2) 2000		243
Austria	5.4882 (477.2)	0.0134 (6.3)	0.0308 (32)	−0.0253 (−20) 1974	0.0069 (8.7) 1987	−0.0191 (−7.8) 2001			1347
Belgium	5.2997 (211.2)	0.0167 (7.7)	0.0343 (20.3)	−0.0252 (−12.7) 1977	−0.004 (−6.4) 1990				1734
Canada	6.1687 (641.7)	0.016 (11)	0.0165 (21.3)	−0.0182 (−11.6) 1976	0.008 (6.4) 1982	0.0128 (9.5) 1997	−0.0243 (−4.5) 2003		841
Denmark	7.544 (463)	−0.0094 (−4.4)	0.0013 (2.2)	0.0141 (10.8) 1993	−0.0176 (−4.2) 2002				225
Finland	5.2277 (425.7)	0.0341 (10.5)	0.0161 (33.4)	0.0227 (11.5) 1993	−0.019 (−6.2) 1999				1856
France	5.7144 (855.4)	0.0072 (7.2)	0.0458 (56)	−0.0393 (−26.2) 1971	−0.0045 (−3.7) 1977	0.0086 (7.4) 1986	−0.0101 (−7) 1992	0.0059 (4.1) 1998	2273
Germany	5.828 (1602.6)	0.0127 (10.9)	0.0235 (56.7)	−0.0045 (−4) 1973	−0.0185 (−14) 1979	0.0187 (15.7) 1988	−0.0134 (−12.2) 1994		5210
Ireland	4.6486 (62.5)	0.0116 (1.6)	0.0129 (4.4)	0.0259 (8.2) 1987					1320
Italy	4.2616 (285.9)	0.0221 (14)	0.0288 (25.2)	−0.0147 (−9.7) 1975	−0.0131 (−9.1) 1982	0.0057 (5) 1988	−0.0119 (−11.4) 1998		1458
Japan	7.9403 (353.9)	0.0274 (9.4)	0.0432 (21)	−0.037 (−12.8) 1972	−0.0081 (−3.1) 1979	0.0125 (5.2) 1985	−0.0122 (−7.4) 1991		200
Netherlands	5.2018 (220.9)	0.0171 (10.3)	0.0375 (23.6)	−0.0331 (−16.2) 1977	0.009 (8.5) 1988	−0.0039 (−3.2) 1994	−0.0112 (−5.8) 2001		2102
New Zealand	5.4305 (508.3)	0.0295 (8.9)	0.0023 (6.6)						66
Norway	6.3471 (124)	0.0137 (3.6)	0.027 (9.6)	−0.0248 (−8) 1979	0.0256 (13.7) 1990	−0.0087 (−3.9) 1996			930
Spain	4.7993 (514.3)	0.0068 (3.7)	0.0321 (36.9)	−0.0308 (−20.8) 1974	0.0121 (8.4) 1983	−0.0126 (−11) 1989			1141
Sweden	7.8633 (331.5)	0.0212 (10.8)	0.0112 (6.6)	−0.0129 (−5.2) 1976	0.0109 (7.1) 1982	0.011 (9) 1992	−0.0121 (−5.8) 2000		1375
Switzerland	6.5959 (349)	0.0171 (10.6)	0.0189 (11.8)	−0.0232 (−13.4) 1973	0.011 (8.5) 1996	−0.0154 (−3.8) 2002			93
UK	5.8173 (609.3)	0.0141 (8.1)	0.0292 (30.7)	−0.0155 (−11.9) 1973	−0.005 (−6.6) 1987				2806
United States	6.4511 (908)	0.0138 (7)	0.0246 (16.5)	−0.0146 (−7) 1967	−0.0081 (−5.5) 1977	0.0099 (7.9) 1983	0.0119 (7) 1999		1926
Euro area	−0.5235 (−43.8)	0.0145 (15.7)	0.0211 (26.7)	−0.0159 (−16.2) 1977	0.0063 (10.4) 1988	−0.0082 (−13.4) 1994			3205
OECD ex.US	−0.4574 (−36.8)	0.0171 (15.3)	0.0177 (21.5)	−0.0119 (−11.2) 1977	0.004 (6.4) 1986	−0.0059 (−10.6) 1993			2583
OECD	−0.4515 (−33.3)	0.0159 (13.2)	0.0144 (16.1)	−0.0088 (−6.8) 1977	0.0037 (6) 1983	0.0057 (3.8) 2002			2958

C10 – Total factor productivity (hours worked), adjusted for the cycle

T stat in brackets	Constant	Cycle	Time trend	Time breaks: additional change in trend productivity gains and respective break-years					F-stat
Australia	1.4449 (94.7)	0.0167 (3.1)	0.0088 (16.4)						170
Belgium	0.6955 (17.1)	0.0148 (5)	0.0418 (14.6)	−0.0233 (−6.5) 1976	−0.009 (−6.3) 1982				1565
Canada	1.4332 (135.9)	0.0147 (11.1)	0.0239 (27.6)	−0.0178 (−18.1) 1974	0.0132 (11.5) 1997	−0.0254 (−4.9) 2003			1494
Denmark	2.7257 (157.3)	−0.0112 (−5)	0.0065 (10.4)	0.0063 (5.5) 1993					398
Finland	0.6729 (57.9)	0.0301 (9.7)	0.0199 (44.7)	0.0171 (9) 1994	−0.0171 (−4.9) 2000				2151
France	1.4465 (277.2)	0.0066 (5.6)	0.0103 (45.8)	−0.0064 (−6.8) 1992	0.0089 (5.6) 1998				1512
Germany	1.0193 (298.1)	0.0125 (11.4)	0.0299 (76.6)	−0.0039 (−3.8) 1973	−0.0215 (−17.4) 1979	0.0202 (18) 1988	−0.0148 (−14.3) 1994		11674
Ireland	−0.183 (−3.5)	−0.0304 (−2.4)	0.04 (24.8)						584
Italy	0.3383 (25.5)	0.022 (15.4)	0.0356 (35)	−0.0184 (−13.9) 1975	−0.0139 (−12.5) 1982	0.0062 (5.9) 1990	−0.0123 (−12.1) 1997		3058
Japan	3.8794 (120.6)	0.0269 (9.3)	0.019 (8.6)	−0.0176 (−6.4) 1976	0.0119 (8.5) 1986	−0.0119 (−9.5) 1994			281
Netherlands	0.8161 (35)	0.0105 (6.6)	0.0351 (22.4)	−0.03 (−13.8) 1977	0.0178 (16.5) 1983	−0.0105 (−10.5) 1993	−0.0086 (−5.4) 2000		4870
New Zealand	1.3667 (130.3)	0.0277 (8.5)	0.0031 (8.9)						82
Norway	1.9742 (43.6)	0.0154 (4.6)	0.0413 (16.7)	−0.0357 (−13) 1979	0.0231 (14) 1990	−0.006 (−3.1) 1996			1820
Spain	0.7026 (30.6)	0.0102 (5.5)	0.0218 (13.7)	−0.02 (−8.3) 1976	0.0198 (10.6) 1982	−0.0204 (−17.4) 1988			718
Sweden	2.8496 (112.6)	0.0213 (9.7)	0.02 (11.2)	−0.0144 (−6.8) 1976	0.01 (13.4) 1991				1538
Switzerland	2.0581 (201.6)	0.0144 (7.6)	0.0021 (4)	−0.0054 (−4.8) 1986	0.0116 (5.6) 1996	−0.0167 (−3.3) 2002			17
UK	0.9983 (76.3)	0.0098 (4.5)	0.0208 (31.5)	−0.011 (−11.8) 1984					1429
United States	1.6369 (257.2)	0.0108 (6.2)	0.0262 (19.6)	−0.012 (−6.5) 1967	−0.0114 (−8.6) 1977	0.0087 (7.8) 1983	0.0127 (8.3) 1999		2997
Euro area	−3.3865 (−239.4)	0.0134 (12.3)	0.0256 (27.3)	−0.0157 (−13.5) 1977	0.0045 (6.3) 1988	−0.0081 (−11.1) 1994			5010
OECD ex.US	−3.3184 (−230)	0.0143 (10.9)	0.0218 (22.9)	−0.0127 (−10.5) 1977	0.0046 (5.9) 1987	−0.0071 (−9.9) 1993			3955
OECD	−3.3695 (−281.9)	0.0129 (11.6)	0.0185 (23.5)	−0.0102 (−9.7) 1977	0.0043 (6.3) 1985	−0.0034 (−4.2) 1993	0.005 (4.8) 1999		4641

Labour productivity vs. total factor productivity[1]

Ulrich Kohli (Swiss National Bank)

1. Introduction

Most headline productivity measures refer to the average product of labour, with productivity growth being typically explained by capital deepening and technological progress. Many economists, however, are more interested in total factor productivity (TFP). Although this is a less intuitive concept, total factor productivity, as indicated by its name, is more general in that it encompasses all factors of production, rather than just one of them. It turns out that TFP is an essential component of the productivity of labour. A contribution of this paper is to document this relationship. A multiplicative decomposition of Swiss labour productivity, 1980–2002, is provided as an illustration.

A second contribution of the paper is to move beyond the usual and rather restrictive two-input, one-output production function setting. Thus, we expand the model by adopting the GDP-function framework that allows for many inputs and outputs, including imports and exports. This makes it possible to show that labour productivity is influenced by additional forces, namely changes in the terms of trade and in the real exchange rate. A complete decomposition of Swiss productivity growth is provided for this case as well.

2. Labour productivity in the production function context

Let the aggregate technology be represented by the following two-input, one-output production function:

$$y_t = f(v_{L,t}, v_{K,t}, t), \tag{1}$$

where y_t measures the quantity of output, $v_{L,t}$ denotes the input of labour services, and $v_{K,t}$ is the input of capital services, all three quantities being measured at time t. The production function itself is allowed to shift over time to account for technological change. We assume that the production function is linearly homogeneous, increasing, and concave with respect to the two input quantities. In what follows, we will also assume competitive behaviour and profit maximization.

The average product of labour ($a_{L,t}$), is defined as:

$$a_{L,t} \equiv \frac{y_t}{v_{L,t}}. \tag{2}$$

In terms of production function (1) we can also write:

$$a_{L,t} = a_L(v_{L,t}, v_{K,t}, t) \equiv \frac{f(v_{L,t}, v_{K,t}, t)}{v_{L,t}}. \tag{3}$$

1 *Paper prepared for the annual Irving Fisher Committee Conference, Bank for International Settlements, Basel, 9–10 September 2004*

2 *Throughout this paper we will use the term labour productivity to designate the* average *productivity of labour. For a discussion of the relationship between* average *and* marginal *productivity, see Kohli (2004c).*

The labour productivity index ($A_{t,t-1}$) can be expressed as one plus the rate of increase in the average product of labour between period $t-1$ and period t:[2]

$$A_{t,t-1} \equiv \frac{a_L(v_{L,t}, v_{K,t}, t)}{a_L(v_{L,t-1}, v_{K,t-1}, t-1)}. \tag{4}$$

Note that it follows from the linear homogeneity of the production function that $a_L(.)$ is homogeneous of degree zero in $v_{L,t}$ and $v_{K,t}$. The same is therefore true for $A_{t,t-1}$, which thus depends on changes in *relative* factor endowments and on the passage of time only.

3. Accounting for labour productivity

We next turn to the task of accounting for the changes over time in labour productivity. Using (4) as a starting point, we can define the following index that isolates the effect of changes in factor endowments over consecutive periods of time:

$$A^L_{V,t,t-1} \equiv \frac{a_L\left(v_{L,t}, v_{K,t}, t-1\right)}{a_L\left(v_{L,t-1}, v_{K,t-1}, t-1\right)}. \tag{5}$$

When defining $A^L_{V,t,t-1}$ we have held the technology constant at its initial (period $t-1$) state. $A^L_{V,t,t-1}$ has thus the Laspeyres form, so to speak. Alternatively, we can adopt the technology of period t as a reference. We then get the following Paasche-like index:

$$A^P_{V,t,t-1} \equiv \frac{a_L\left(v_{L,t}, v_{K,t}, t\right)}{a_L\left(v_{L,t-1}, v_{K,t-1}, t\right)}. \tag{6}$$

Since there is no reason *a priori* to prefer one measure over the other, we can follow Diewert and Morrison's (1986) lead and take the geometric mean of the two indexes just defined. We thus get:

$$A_{V,t,t-1} \equiv \sqrt{\frac{a_L\left(v_{L,t}, v_{K,t}, t-1\right)}{a_L\left(v_{L,t-1}, v_{K,t-1}, t-1\right)} \cdot \frac{a_L\left(v_{L,t}, v_{K,t}, t\right)}{a_L\left(v_{L,t-1}, v_{K,t-1}, t\right)}}\ . \tag{7}$$

Note that if capital deepening takes place, both $A^L_{V,t,t-1}$ and $A^P_{V,t,t-1}$ are greater than one, in which case $A_{V,t,t-1}$ must exceed one as well.[3]

In the same vein, we can define an index that isolates the impact of technological change. That is, we compute the index of labour productivity, allowing for the passage of time, but holding factor endowments fixed, first at their level of period $t-1$, and then at their level of period t:

$$A^L_{T,t,t-1} \equiv \frac{a_L\left(v_{L,t-1}, v_{K,t-1}, t\right)}{a_L\left(v_{L,t-1}, v_{K,t-1}, t-1\right)} \tag{8}$$

$$A^P_{T,t,t-1} \equiv \frac{a_L\left(v_{L,t}, v_{K,t}, t\right)}{a_L\left(v_{L,t}, v_{K,t}, t-1\right)}. \tag{9}$$

Taking the geometric mean of these two indexes, we get:

$$A_{T,t,t-1} \equiv \sqrt{\frac{a_L\left(v_{L,t-1}, v_{K,t-1}, t\right)}{a_L\left(v_{L,t-1}, v_{K,t-1}, t-1\right)} \cdot \frac{a_L\left(v_{L,t}, v_{K,t}, t\right)}{a_L\left(v_{L,t}, v_{K,t}, t-1\right)}}\ . \tag{10}$$

Comparing (7) and (10) with (4), it can then easily be seen that $A_{V,t,t-1}$ and $A_{T,t,t-1}$ together yield a complete decomposition of the labour productivity index:

$$A_{t,t-1} = A_{V,t,t-1} \cdot A_{T,t,t-1}. \tag{11}$$

3 *This follows directly from the slope and linear homogeneity properties of the production function.*

4. Total factor productivity

While labour productivity remains the concept of choice when it comes to the public debate, most economists prefer to think in terms of TFP. The measure of TFP treats all inputs symmetrically. In the production function context, it can be defined as the increase in output that is not explained by increases in input quantities. Put differently, it is the increase in output made possible by technological change, holding all inputs constant. One state-of-the art definition of TFP, $Y_{T,t,t-1}$, is drawn from the work of Diewert and Morrison (1986). It too can be thought of as the geometric average of Laspeyres-like and Paasche-like measures:

$$Y_{T,t,t-1} \equiv \sqrt{\frac{f(v_{L,t-1}, v_{K,t-1}, t)}{f(v_{L,t-1}, v_{K,t-1}, t-1)} \cdot \frac{f(v_{L,t}, v_{K,t}, t)}{f(v_{L,t}, v_{K,t}, t-1)}} \;. \tag{12}$$

In view of (3), it is immediately clear that $Y_{T,t,t-1}$ as given by (12) is in fact identical to $A_{T,t,t-1}$ as defined by (10). That is, TFP in this model is equal to the contribution of technological change when explaining the productivity of labour. Labour productivity will exceed TFP to the extent that capital deepening occurs ($A_{V,t,t-1} > 1$).

5. Measurement

To make the decomposition (11) operational one needs to specify a functional form for the production function (1). One functional form well suited for this purpose is the Translog. In the production function context, and under linear homogeneity, it is as follows:[4]

$$\begin{aligned} \ln y_t = \alpha_0 + \beta_K \ln v_{K,t} + (1-\beta_K) \ln v_{L,t} + \frac{1}{2}\phi_{KK} (\ln v_{K,t} - \ln v_{L,t})^2 + \\ \phi_{KT} (\ln v_{K,t} - \ln v_{L,t})\, t + \beta_T\, t + \frac{1}{2}\phi_{TT}\, t^2 \end{aligned} \tag{13}$$

One option, at this stage, would be to estimate function (13) econometrically, and then to use the resulting parameter estimates to calculate (7) and (10) in order to get the full decomposition of labour productivity.[5] It turns out, however, that as long as the true production function is indeed given by (13), it is not necessary to have estimates of its parameters to be able to proceed. Thus, Diewert and Morrison (1986) have shown that in this case TFP ($Y_{T,t,t-1}$) as defined by (12) can be calculated from knowledge of the data alone in the following way:

$$Y_{T,t,t-1} \left(= A_{T,t,t-1} \right) = \frac{Y_{t,t-1}}{V_{t,t-1}} \;, \tag{14}$$

where $Y_{t,t-1}$is the index of real GDP:

$$Y_{t,t-1} \equiv \frac{y_t}{y_{t-1}} \;, \tag{15}$$

and $V_{t,t-1}$ is a Törnqvist index of input quantities:[6]

$$V_{t,t-1} \equiv \exp\left[\sum_{j \in \{L,K\}} \frac{1}{2} (s_{j,t} + s_{j,t-1}) \ln \frac{v_{j,t}}{v_{j,t-1}} \right] ; \tag{16}$$

4 *See Christensen, Jorgenson, and Lau (1973).*
5 *See Kohli (1990, 1991, 2004c) for such an econometric approach.*
6 *The Törnqvist index is a superlative index in the sense of Diewert (1976).*

$s_{j,t}$ in (16) is the income share of factor j:

$$s_{j,t} \equiv \frac{w_{j,t} v_{j,t}}{p_t y_t}, \qquad j \in \{L, K\}, \tag{17}$$

where $W_{j,t}$ is the rental prices of factor j, and P_t is the price of output. Note that the GDP identity implies that $S_{L,t} + S_{K,t} = 1$.

It follows from the definition of $A_{t,t-1}$ that:

$$A_{t,t-1} = \frac{Y_{t,t-1}}{v_{L,t} / v_{L,t-1}} . \tag{18}$$

Making use of (11), (14), (16), and (18), we thus find that:

$$A_{V,t,t-1} \equiv \exp\left[\frac{1}{2}(s_{K,t} + s_{K,t-1})\left(\ln \frac{v_{K,t}}{v_{L,t}} - \ln \frac{v_{K,t-1}}{v_{L,t-1}}\right)\right]. \tag{19}$$

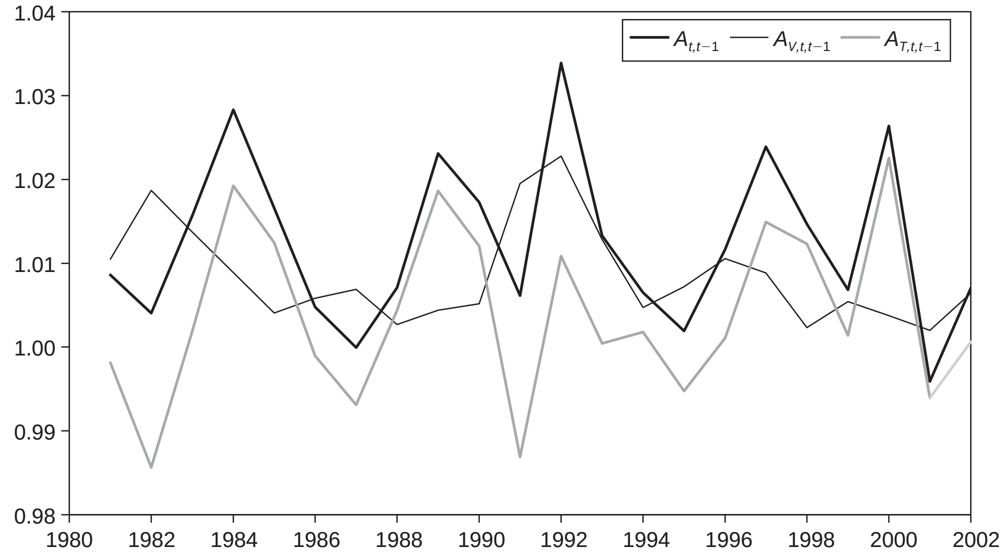
Table 1 reports estimates of the decomposition (11) for Switzerland over the period 1981 to 2002.[7] Cumulated effects and geometric averages for the entire period are shown at the bottom of the table. One can see that labour productivity has increased by about 31% over the entire period; this amounts to about 1.2% *per annum* on average. TFP, on the other hand, has increased by a much more modest 8.8% (about 0.4% per year). The bulk of the increase in labour productivity is due to capital deepening, which added about 20% (0.8% on average annually) to labour productivity over the sample period. One also observes some fairly large annual variations. Thus, labour productivity increased by as much as 3.4% (in 1992), and it actually fell on a couple of

Table 1 – Decomposition of the productivity of labour:
2-input Translog production function

	$A_{t,t-1}$	$A_{V,t,t-1}$	$A_{T,t,t-1}$
1981	1.00860	1.01047	0.99816
1982	1.00405	1.01868	0.98564
1983	1.01565	1.01373	1.00189
1984	1.02829	1.00888	1.01923
1985	1.01658	1.00407	1.01245
1986	1.00478	1.00583	0.99896
1987	0.99994	1.00689	0.99310
1988	1.00708	1.00268	1.00438
1989	1.02309	1.00440	1.01862
1990	1.01728	1.00516	1.01206
1991	1.00614	1.01950	0.98690
1992	1.03386	1.02276	1.01085
1993	1.01328	1.01283	1.00044
1994	1.00650	1.00472	1.00178
1995	1.00192	1.00720	0.99476
1996	1.01165	1.01055	1.00109
1997	1.02389	1.00884	1.01492
1998	1.01468	1.00234	1.01232
1999	1.00682	1.00541	1.00140
2000	1.02636	1.00375	1.02253
2001	0.99591	1.00199	0.99394
2002	1.00704	1.00643	1.00061
1981–2002	**1.31091**	**1.20436**	**1.08847**
Mean	1.01238	1.00849	1.00386

7 *See the Appendix for a description of the data.*

Figure 1 – Labour productivity, capital deepening, and total factor productivity, 1981–2002



occasions (in 1987 and 2001). The contribution of capital deepening, although larger than that of TFP, is much steadier. This is illustrated in Figure 1 that shows the contributions of both components. It is visible that the fluctuations of labour productivity largely reflect those of TFP.

The fact that TFP is found to be rather small, and perhaps less than one would have expected, might be thought to reflect poorly on Switzerland and its capacity to innovate. One must remember, however, that TFP is essentially measured as a Solow *residual*; see expression (14). TFP is the growth in output that cannot be explained by the model, given the input data on hand. TFP, too some degree, is a measure of our ignorance. The more precisely inputs and outputs are measured, the smaller one should probably expect TFP to be. In this study, we have used a superlative index to aggregate outputs and inputs. Furthermore, we have used refined measures of hours worked and of the capital stock. This might explain why the residual is found to be rather modest.[8]

6. Domestic real value added

The model of the production function is rather limiting since it imposes the number of outputs to be one.[9] Moreover, the production function approach does not make it possible to take into account imports and exports. In what follows, we therefore opt for the description of the aggregate technology by a real value added (or real income) function, such as the one proposed by Kohli (2004a). It is based on the GDP function approach to modelling the production sector of an open economy.[10] We assume that the technology counts two outputs, domestic (nontraded) goods (D) and exports (X), as well as three inputs, labour (L), capital (K), and imports (M); imports are treated as a variable input, i.e. as a negative output. This treatment recognises the fact that most foreign trade is in middle products, and that even most so-called finished goods that are imported must still transit through the production sector where they are combined with domestic value added before meeting final demand. We denote output (including import) quantities by y_i and their prices by p_i, $i \in \{D, X, M\}$. Furthermore, we denote the inverse of the terms of trade by q ($q \equiv p_M / p_X$) and the relative price of tradables vs. nontradables by e ($e \equiv p_X / p_D$). Note that for given terms of trade, a change in e can be interpreted as a change in the real exchange rate, an increase in e being equivalent to a real depreciation of the home currency. Let π_t be nominal GDP:

$$\pi_t \equiv p_{D,t} y_{D,t} + p_{X,t} y_{X,t} - p_{M,t} y_{M,t} = p_t y_t \quad . \tag{20}$$

8 *Further progress could probably be made by weighting work hours by their marginal productivity, rather than simply adding them up; see Greenwood and Kohli (2003) for an analysis along these lines.*
9 *Alternatively, one must assume that outputs are globally separable from domestic inputs.*
10 *See Kohli (1978), Woodland (1982).*

Domestic real value added (z_t) – or real domestic income – is defined as nominal GDP deflated by the price of domestic output:

$$z_t \equiv \frac{\pi_t}{p_{D,t}} = y_{D,t} + e_t y_{X,t} - e_t q_t y_{M,t} \,. \tag{21}$$

The difference between real GDP (y_t) and real value added (z_t) lies in the price index that is being used to deflate nominal GDP. In the case of real GDP, an index of the prices of all output components (i.e. including imports and exports) is used (p_t), whereas in the case of real value added (or real income), only the prices of the domestic components are retained ($P_{D,t}$).[11] The effect of this difference in treatment becomes apparent if one considers a change in the terms of trade or in the real exchange rate. An improvement in the terms of trade, for instance, mathematically has little or no impact on real GDP, but it results in an increase in real value added; see Kohli (2004a). For unchanged factor endowments, the productivity of labour is thereby enhanced. Over the past quarter century, it turns out that Switzerland has experienced a significant improvement in its terms of trade and a real appreciation in its currency. This is documented by Figure 2 that shows the terms of trade (measured by $p_{X,t}/p_{M,t}$) and the real exchange rate (measured by $p_{X,t}/p_{D,t}$). The trends in the two series suggest differing paths for real GDP and real value added. This impression is confirmed by Figure 3, which shows the normalized path of the two indices (y_t and z_t). Clearly, real value added has increased more rapidly than real GDP on average.

Let T_t be the production possibilities set at time t. We assume that T_t is a convex cone. The aggregate technology can be described by a real valued added function defined as follows:[12]

$$z_t = g(q_t, e_t, v_{K,t}, v_{L,t}, t) \equiv \max_{y_D, y_X, y_M} \left\{ \begin{array}{l} y_{D,t} + e_t y_{X,t} - e_t\, q_t y_{M,t} : \\ \qquad \left(y_{D,t}, y_{X,t}, y_{M,t}, v_{K,t}, v_{L,t} \right) \in T_t \end{array} \right\} . \tag{22}$$

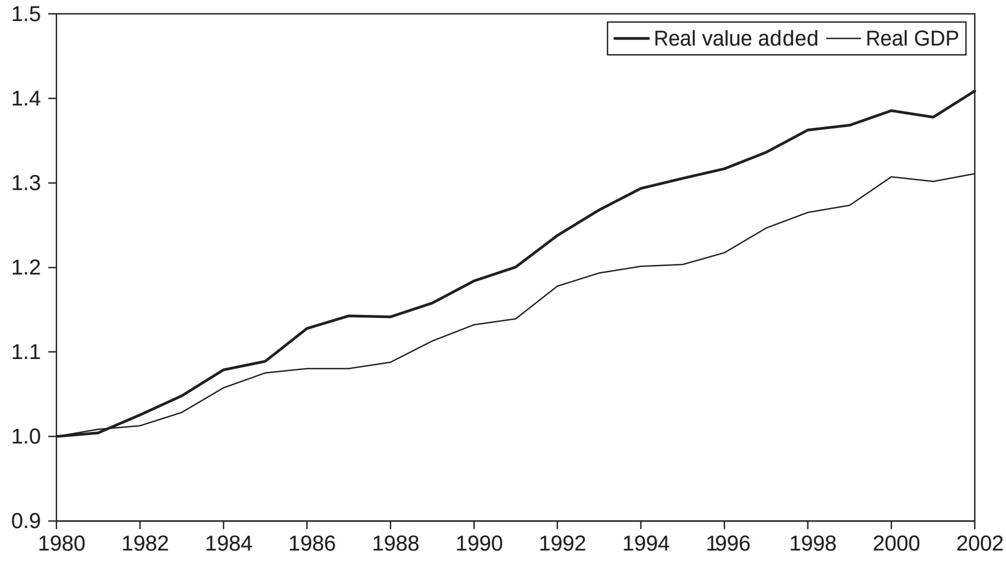
In this context, the real value added per unit of labour ($b_{L,t}$) can be defined as follows:

$$b_{L,t} \equiv \frac{z_t}{v_{L,t}} . \tag{23}$$

In terms of the real value-added function, we get:

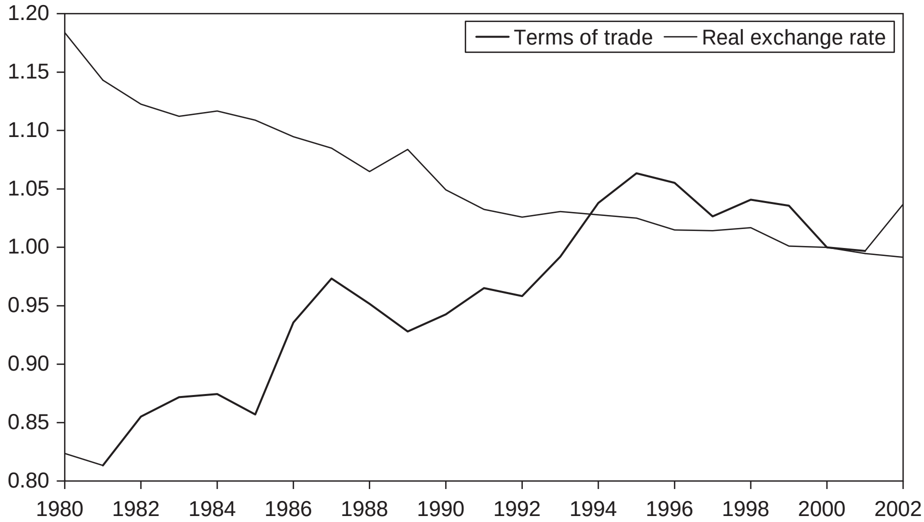
$$b_{L,t} = b_L(q_t, e_t, v_{K,t}, v_{L,t}, t) \equiv \frac{g(q_t, e_t, v_{K,t}, v_{L,t}, t)}{v_{L,t}} . \tag{24}$$

Figure 2 – Real value-added and real GDP, 1981–2002



11 *Törnqvist indices are used in both cases.*
12 *See Kohli (2004a).*

Figure 3 – Terms of trade and real exchange rate, 1980–2002



The labour productivity index is now expressed as:

$$B_{t,t-1} \equiv \frac{b_L(q_t, e_t, v_{K,t}, v_{L,t}, t)}{b_L(q_{t-1}, e_{t-1}, v_{K,t-1}, v_{L,t-1}, t-1)}. \tag{25}$$

The Translog functional form is once again well suited for our purposes. In the context of the real value added function it can be written as:[13]

$$\begin{aligned}\ln z_t = {} & \alpha_0 + \alpha_Q \ln q_t + \alpha_E \ln e_t + \beta_K \ln v_{K,t} + (1-\beta_K) \ln v_{L,t} \\ & + \frac{1}{2}\gamma_{QQ}(\ln q_t)^2 + \gamma_{QE} \ln q_t \ln e_t + \frac{1}{2}\gamma_{EE}(\ln e_t)^2 \\ & + \frac{1}{2}\phi_{KK}(\ln v_{K,t} - \ln v_{L,t})^2 + (\delta_{QK} \ln q_t + \delta_{EK} \ln e_t)(\ln v_{K,t} - \ln v_{L,t}) \\ & + (\delta_{QT} \ln q_t + \delta_{ET} \ln e_t)\, t + \phi_{KT}(\ln v_{K,t} - \ln v_{L,t})\, t + \beta_T\, t + \frac{1}{2}\phi_{TT}\, t^2 \end{aligned} \tag{26}$$

Proceeding along the same lines as in Section 4, we can define the following index to capture the contribution of changes in the terms of trade to the productivity of labour:

$$B_{Q,t,t-1} \equiv \sqrt{\frac{b_L(q_t, e_{t-1}, v_{K,t-1}, v_{L,t-1}, t-1)}{b_L(q_{t-1}, e_{t-1}, v_{K,t-1}, v_{L,t-1}, t-1)} \cdot \frac{b_L(q_t, e_t, v_{K,t}, v_{L,t}, t)}{b_L(q_{t-1}, e_t, v_{K,t}, v_{L,t}, t)}}. \tag{27}$$

Similarly, we can identify the contribution of changes in the real exchange rate as:

$$B_{E,t,t-1} \equiv \sqrt{\frac{b_L(q_{t-1}, e_t, v_{K,t-1}, v_{L,t-1}, t-1)}{b_L(q_{t-1}, e_{t-1}, v_{K,t-1}, v_{L,t-1}, t-1)} \cdot \frac{b_L(q_t, e_t, v_{K,t}, v_{L,t}, t)}{b_L(q_t, e_{t-1}, v_{K,t}, v_{L,t}, t)}}, \tag{28}$$

the contribution of changes in domestic factor endowments:

$$B_{V,t,t-1} \equiv \sqrt{\frac{b_L(q_{t-1}, e_{t-1}, v_{K,t}, v_{L,t}, t-1)}{b_L(q_{t-1}, e_{t-1}, v_{K,t-1}, v_{L,t-1}, t-1)} \cdot \frac{b_L(q_t, e_t, v_{K,t}, v_{L,t}, t)}{b_L(q_t, e_t, v_{K,t-1}, v_{L,t-1}, t)}}, \tag{29}$$

13 *See Diewert (1974).*

and, finally, the contribution of technological progress:

$$B_{T,t,t-1} \equiv \sqrt{\frac{b_L(q_{t-1}, e_{t-1}, v_{K,t-1}, v_{L,t-1}, t)}{b_L(q_{t-1}, e_{t-1}, v_{K,t-1}, v_{L,t-1}, t-1)} \cdot \frac{b_L(q_t, e_t, v_{K,t}, v_{L,t}, t)}{b_L(q_t, e_t, v_{K,t}, v_{L,t}, t-1)}} . \tag{30}$$

Assuming that the real value added function is given by (26), it can be shown that these four effects together give a complete decomposition of the productivity of labour as defined by (25):[14]

$$B_{t,t-1} = B_{Q,t,t-1} \cdot B_{E,t,t-1} \cdot B_{V,t,t-1} \cdot B_{T,t,t-1} . \tag{31}$$

The left-hand side of (31) can readily be computed in the following way:

$$B_{t,t-1} \equiv \frac{z_t / v_{L,t}}{z_{t-1} / v_{L,t-1}} . \tag{32}$$

Under the hypothesis that the true real value added function is indeed Translog, the components on the right-hand side of (31) can be calculated on the basis of the data alone, that is without knowledge of the parameters of (26). One can thus show that:[15]

$$B_{Q,t,t-1} = \exp\left[\frac{1}{2}(-s_{M,t} - s_{M,t-1}) \ln \frac{q_t}{q_{t-1}}\right] \tag{33}$$

$$B_{E,t,t-1} = \exp\left[\frac{1}{2}(s_{B,t} + s_{B,t-1}) \ln \frac{e_t}{e_{t-1}}\right] \tag{34}$$

$$B_{V,t,t-1} \equiv \exp\left[\frac{1}{2}(s_{K,t} + s_{K,t-1})\left(\ln \frac{v_{K,t}}{v_{L,t}} - \ln \frac{v_{K,t-1}}{v_{L,t-1}}\right)\right] \tag{35}$$

$$B_{T,t,t-1} \equiv \frac{Y_{t,t-1}}{V_{t,t-1}} , \tag{36}$$

where s_M is the GDP share of imports ($s_M \equiv p_M y_M / \pi$), s_B is the trade balance relative to GDP ($s_B \equiv (p_X y_X - p_M y_M)/\pi$), and s_K is, as before, the GDP share of capital. It is noteworthy that $B_{V,t,t-1} = A_{V,t,t-1}$ and $B_{V,t,t-1} = A_{V,t,t-1}$. This implies:

$$B_{t,t-1} = B_{Q,t,t-1} \cdot B_{E,t,t-1} \cdot A_{t,t-1} . \tag{37}$$

That is, the difference between the growth in real value added per unit of labour and that of real GDP per unit of labour is due to the terms-of-trade and the real exchange-rate effects. These are precisely the elements that account for the distinction between the two price indices (p_t vs. $p_{D,t}$) used to deflate nominal GDP to get either real GDP or real value added. Indeed, $p_{t,t-1}$, the change in domestic prices over consecutive periods and which is obtained as a Törnqvist index of the prices of domestic sales, exports and imports, can be written as follows:

$$\begin{aligned} p_{t,t-1} &= \exp\left[\sum_{i \in \{D,X,M\}} \pm \frac{1}{2}(s_{i,t} + s_{i,t-1}) \ln \frac{p_{i,t}}{p_{i,t-1}}\right] \\ &= \frac{p_{D,t}}{p_{D,t-1}} + \exp\left[\frac{1}{2}(-s_{M,t} - s_{M,t-1}) \ln \frac{q_t}{q_{t-1}}\right] + \exp\left[\frac{1}{2}(s_{B,t} + s_{B,t-1}) \ln \frac{e_t}{e_{t-1}}\right] \end{aligned} \tag{38}$$

where we have taken into account the fact that $s_D + s_X - s_M = s_D + s_B = 1$. This expression demonstrates that p_t, the commonly used GDP price deflator, encompasses two components,

14 *The demonstration is the same as in Kohli (2004a).*
15 *See Kohli (2004a).*

the terms-of-trade effect and the real exchange-rate effect, that should best be viewed as real – rather than price – elements.

A decomposition of the productivity of labour according to (31) is reported in Table 2. One finds that labour productivity, in terms of real value added, has increased by close to 41% over the sample period. The impact of the real exchange rate was negligible,[16] but favourable movements in the terms of trade have lifted labour productivity by 7.8% (about 0.3% per year). This is nearly as much as the contribution of technological change, as suggested by our estimate of TFP. Our results are further illustrated by Figure 4, which is based on (37). It shows the annual contributions of changes in the terms of trade and in the real exchange rate, together with the real-value-added and GDP labour productivity indices. It appears that at times the movement in the terms of trade has very much dominated the dynamics of real value added. In 1986, for instance, labour productivity increased by 3.6% as indicated by $B_{t,t-1}$, whereas, judging from the estimate of $A_{t,t-1}$, real GDP per unit of labour only increased by 0.6%. Except for a negligible real exchange-rate effect, the difference is due to the massive terms-of-trade effect (3.1%) that Switzerland experienced that year.

Table 2 – Decomposition of the productivity of labour: 2-input, 3-output Translog real domestic value added function

	$B_{t,t-1}$	$B_{Q,t,t-1}$	$B_{E,t,t-1}$	$B_{V,t,t-1}$	$B_{T,t,t-1}$
1981	1.00430	0.99545	1.00028	1.01047	0.99816
1982	1.02121	1.01728	0.99981	1.01868	0.98564
1983	1.02201	1.00639	0.99987	1.01373	1.00189
1984	1.02938	1.00102	1.00004	1.00888	1.01923
1985	1.00917	0.99281	0.99990	1.00407	1.01245
1986	1.03584	1.03118	0.99973	1.00583	0.99896
1987	1.01304	1.01329	0.99980	1.00689	0.99310
1988	0.99909	0.99242	0.99965	1.00268	1.00438
1989	1.01426	0.99112	1.00024	1.00440	1.01862
1990	1.02260	1.00561	0.99962	1.00516	1.01206
1991	1.01377	1.00788	0.99971	1.01950	0.98690
1992	1.03132	0.99776	0.99979	1.02276	1.01085
1993	1.02429	1.01062	1.00024	1.01283	1.00044
1994	1.02019	1.01376	0.99984	1.00472	1.00178
1995	1.00910	1.00731	0.99986	1.00720	0.99476
1996	1.00875	0.99765	0.99949	1.01055	1.00109
1997	1.01464	0.99100	0.99997	1.00884	1.01492
1998	1.01973	1.00484	1.00014	1.00234	1.01232
1999	1.00422	0.99823	0.99918	1.00541	1.00140
2000	1.01272	0.98677	0.99994	1.00375	1.02253
2001	0.99440	0.99874	0.99974	1.00199	0.99394
2002	1.02246	1.01551	0.99982	1.00643	1.00061
1981–2002	**1.40871**	**1.07823**	**0.99665**	**1.20436**	**1.08847**
Mean	1.01570	1.00343	0.99985	1.00849	1.00386

Figure 5 provides a final illustration of our results. It shows the cumulated effects of productivity gains ($A_t \equiv \prod_{h=1}^{t} A_{h,h-1}$, and so on). Thus, it indicates the paths of labour productivity (in terms of real value added and in terms of real GDP) and of TFP over the past quarter century. The vertical distance between $A_{T,t}$ and A_t results from capital deepening, whereas the vertical distance between A_t and B_t is accounted for by the foreign-trade (terms-of-trade and real exchange-rate) effects. One again, we see that capital deepening has dominated in the Swiss case, and that the role of trade effects has been about as important as that of TFP.

16 *This is due to the fact that, in the eighties, when the real exchange rate was dropping rapidly, net exports were close to zero. Later, as a larger current account surplus developed, the real appreciation slowed down markedly. Nonetheless, the real exchange-rate effect must be considered for things to add up, or more precisely, for the multiplicative decomposition to be complete.*

Figure 4 – Labour productivity and the contribution of terms-of-trade and real exchange-rate changes, 1981–2002

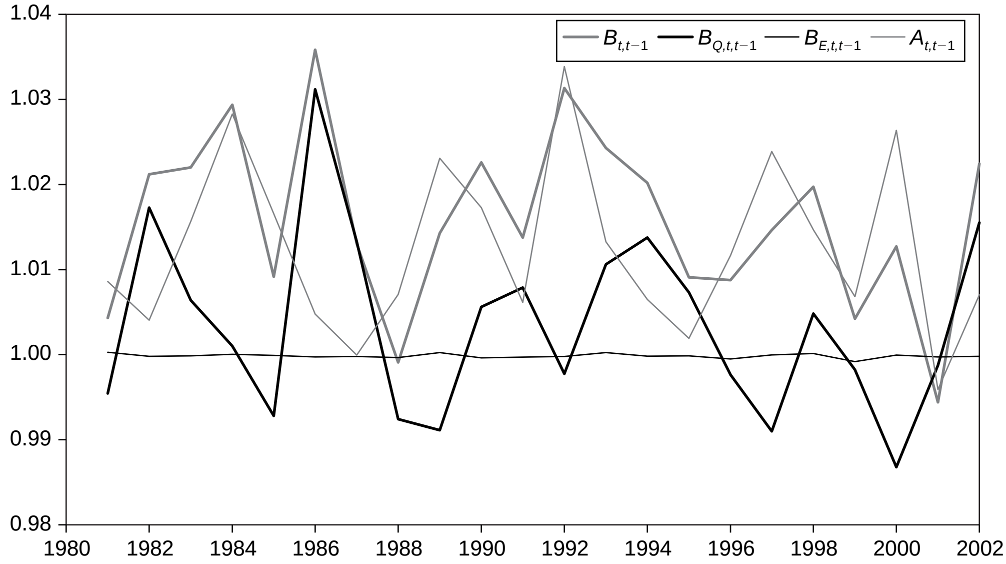
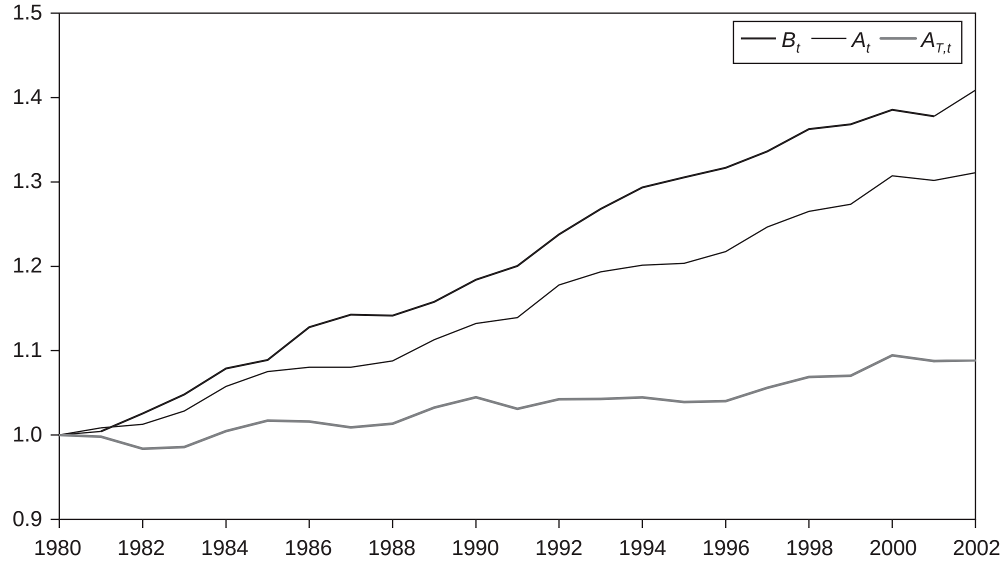


Figure 5 – Labour productivity and totals factor productivity: Cumulated effects, 1980–2002



7. Conclusions

Productivity is an important, yet elusive concept. In this paper we have focused on the two measures of productivity most prevalent in the literature, labour productivity and TFP. Furthermore, we have identified and quantified the main forces at work: technological progress, capital deepening, terms-of-trade changes, and changes in the real exchange rate. We have shown that TFP is an essential component of labour productivity. However, we have found that in the Swiss case, capital deepening has played an even larger role in explaining the growth of labour productivity. Improvements in the terms of trade have played a very substantial role as well.

Appendix: Description of the data

All data are annual for the period 1980 to 2002. We require the prices and quantities of all inputs and outputs. The data for GDP and its components, in nominal and in real terms, are taken from the *Office fédéral de la statistique (OFS)* website. Prices are then obtained by deflation. Data on labour compensation and on the operating surplus are also retrieved from the *OFS* website. The quantity of capital services is assumed to be proportional to the stock. The necessary figures, together with data on labour input (measured in hours worked) are *Swiss National Bank* estimates; see Fox and Zurlinden (2004). The rental price of labour and capital are then obtained by dividing labour and capital income by the corresponding quantity series. For the purpose of Section 4 output is expressed as an implicit Törnqvist index of real GDP; see Kohli (2004b) for details. In Section 5, the price of nontraded goods is computed as a Törnqvist price index of the deflators of consumption, investment and government purchases.

References

Christensen, Laurits R., Dale W. Jorgenson, and Lawrence J. Lau (1973) "Transcendental Logarithmic Production Frontiers", *Review of Economics and Statistics* 55, 28–45.

Diewert, W. Erwin (1974) "Applications of Duality Theory", in Michael D. Intriligator and David A. Kendrick (eds.) *Frontiers of Quantitative Economics*, Vol. 2 (Amsterdam: North-Holland).

Diewert, W. Erwin (1976) "Exact and Superlative Index Numbers", *Journal of Econometrics* 4, 115–145.

Diewert, W. Erwin and Catherine J. Morrison (1986) "Adjusting Output and Productivity Indexes for Changes in the Terms of Trade", *Economic Journal* 96, 659–679.

Fox, Kevin J. and Mathias Zurlinden (2004) "On Understanding Sources of Output Gaps and Growth for Switzerland", Swiss National Bank, unpublished.

Greenwood, Michael J. and Ulrich Kohli (2003) "Exact Unemployment Rate Indices", *Contemporary Economic Policy* 21, 218–226.

Kohli, Ulrich (1978) "A Gross National Product Function and the Derived Demand for Imports and Supply of Exports", *Canadian Journal of Economics* 11, 167–182.

Kohli, Ulrich (1990) "Growth Accounting in the Open Economy: Parametric and Nonparametric Estimates", *Journal of Economic and Social Measurement* 16, 125–136.

Kohli, Ulrich (1991) *Technology, Duality, and Foreign Trade: The GNP Function Approach to Modeling Imports and Exports* (Ann Arbor, MI: University of Michigan Press).

Kohli, Ulrich (2004a) "Real GDP, Real Domestic Income, and Terms-of-Trade Changes", *Journal of International Economics* 62, 83–106.

Kohli, Ulrich (2004b) "An Implicit Törnqvist Index of Real GDP", *Journal of Productivity Analysis* 21, 337–353.

Kohli, Ulrich (2004c) "Labour Productivity: Average vs. Marginal", paper presented at the *SSHRC International Conference on Index Number Theory and the Measurement of Prices and Productivity*, Vancouver, July 2004.

Office fédéral de la statistique, *http://www.statistik.admin.ch*.

Woodland, Alan D. (1982) *International Trade and Resource Allocation* (Amsterdam: North-Holland).

Abstract

In this paper, we examine the relationship between two commonly used measures of productivity, namely labour productivity and total factor productivity. We show that total factor productivity is an essential component of labour productivity. Labour productivity is further influenced by capital intensity, and, in the open economy context, by the terms of trade and the real exchange rate. Complete multiplicative decompositions of productivity are given for Switzerland for the period 1980 to 2002. Our analysis rests on a tight theoretical framework being based on the GDP function approach to modelling the production sector of an open economy.

Keywords: labour productivity, total factor productivity, index numbers, technological change, capital deepening, terms of trade, real exchange rate

JEL classification: D24, O47, E25, F43.

Ulrich Kohli,
Chief economist, Swiss National Bank, and honorary professor, University of Geneva.

Postal address:
Swiss National Bank,
Börsenstrasse 15, P.O. Box 2800,
CH-8022 Zurich, Switzerland.
Phone: +41-44-631-3233; fax: +41-44-631-3188;
e-mail: Ulrich.Kohli@snb.ch;
home page: http//www.unige.ch/ses/ecopo/kohli/kohli.html

Revised November 2004

Measuring total factor productivity for the United Kingdom[1]

Charlotta Groth, Maria Gutierrez-Domenech and Sylaja Srinivasan
(Bank of England)[2]

Summary

A good understanding of productivity growth is important for understanding aggregate supply capacity, and so for the conduct of monetary policy. To understand the sources of supply capacity well, it is important to measure output and factor inputs correctly. This article summarises recent and ongoing research at the Bank of England on improved measures of factor inputs. This work explicitly accounts for changes in the quality of these inputs and for the flow of services available from them, as well as for the costs of adjusting the level and utilisation of the inputs over time. This research was presented at a workshop on "measuring factor inputs" held at the Bank of England in December 2003.

Introduction

The aim of monetary policy is to keep inflation low and stable, in accordance with the target set by the Chancellor. A key influence on inflationary pressure is the balance between the demand for and the economy's capacity to supply goods and services. This capacity depends both on the quantities and qualities of the primary inputs into the production process – capital and labour – and on the efficiency with which they are combined. The latter concept is often referred to as total factor productivity (TFP). A good understanding of past and current productivity growth is thus important for understanding aggregate supply capacity, and so it is relevant for the conduct of monetary policy.

To understand the sources of supply capacity well, it is important to measure output and factor inputs, and therefore productivity, correctly. It is also crucial to recognise and adjust for the changing composition of the aggregate inputs, which may vary over time. This article discusses recent work at the Bank of England on improved measures of factor inputs, which accounts explicitly for changes in their quality and for the flow of services available from them, and for the costs of adjusting the level and utilisation of the inputs over time. These improved factor input estimates can then be used to obtain better measures of total factor productivity growth for the United Kingdom.

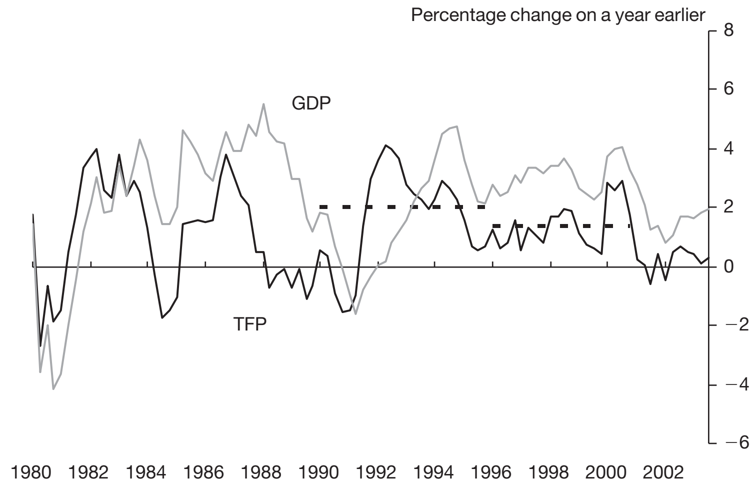
The Solow residual

The standard measure of total factor productivity growth is the Solow residual[3]: that part of output growth that cannot be accounted for by the growth of the primary factors of production, capital and labour.[4] The Solow residual (z) is calculated by subtracting the growth of the primary inputs (weighted by their respective shares in nominal output) from the growth of output:[5]

$$z = y - s_k k - s_l l \qquad (1)$$

1 *This article has been reproduced from the Spring 2004 issue of the* Bank of England Quarterly Bulletin*.*
2 *We would like to thank John Fernald, Steve Nickell, Soledad Nunez and Nick Oulton for valuable comments. We would also like to thank Pablo-Burriel Llombart and Jerry Jones for supplying us with the quality adjusted labour series.*
3 *Total factor productivity as defined in this article is also referred to in the literature as multi-factor productivity. (See the November 2003* Bank of England Inflation Report *for some standard multi-factor productivity estimates for the United Kingdom.)*
4 *For a "biography" of total factor productivity, see Hulten (2001).*
5 *This is a simplified version of the formula actually used in the empirical calculations. The full formula is given in the appendix in equation* ***(A.1)****.*

Chart 1 – Growth of Total Factor Productivity and GDP for the United Kingdom: 1980–2003



where y is the growth rate of output, k is the growth rate of capital input, l is the growth rate of labour input and s_k and s_l are the shares of capital and labour in nominal output respectively.

Chart 1 shows a standard measure of the Solow residual for the United Kingdom.[6] The growth rate of TFP is calculated here using aggregate data, where the capital input is a capital stock measure and the labour input is total hours worked.[7] The growth rate appears to be procyclical – it is positively correlated with GDP growth.[8] But over and above that, a slowing in the growth rate is noticeable in the second half of the 1990s (relative to the first half), in contrast to the United States, which experienced an increase in TFP growth in the late 1990s. Basu, Fernald, Oulton and Srinivasan (2003) discuss possible reasons for the differing productivity growth patterns in the United States and the United Kingdom.[9]

The Solow residual shown above provides us with just one estimate of total factor productivity growth in the United Kingdom. There are, however, a number of well-known measurement issues that need to be considered. First, capital and labour inputs need to be estimated correctly. For example, the capital measure should reflect the productive services available from the capital stock and needs to reflect factors such as the increased use of ICT capital; and the labour measure should reflect the changing composition and skills of the UK labour force. Second, because the movement of resources between industries also affects aggregate productivity, it is preferable to aggregate industry-level data rather than to use aggregated data directly.[10] Third, the basic Solow residual calculation in equation 1 assumes that all factors of production are flexible and fully employed. This may not be the case if there are costs involved in e.g. hiring and firing or in installing new machines and equipment (usually referred to collectively as adjustment costs). Also, if it is costly to adjust inputs, firms may respond to short-run fluctuations in demand by varying the rates at which their existing capital and labour are utilised. The remainder of this article summarises ongoing Bank of England research on each of these measurement issues and considers their impact on UK TFP growth.[11]

Measuring factor inputs

This section discusses measurement issues relating to the factor inputs used in the TFP calculations.

6 *The sources for the data underlying the calculations are given in the appendix.*
7 *A similar TFP growth measure, but using the number of people in employment as the labour input, is summarised in Table 3.A of the November 2003* Bank of England Inflation Report*.*
8 *This is similar to the United States. See Burnside, Eichenbaum and Rebelo (1995) and Basu and Fernald (2000).*
9 *They argue that unmeasured investments in intangible organisational capital – associated with the role of ICT as a "general-purpose technology" – can explain the divergent US and UK productivity performance after 1995.*
10 *See Stiroh (2002) and Bosworth and Triplett (2003) for an explanation of these effects.*
11 *The focus of this article is on total factor productivity. Clearly, a corresponding labour productivity measure can be calculated.*

Capital services

The standard Solow residual is calculated as that part of output growth that cannot be accounted for by growth in capital and labour inputs. The measure of capital that is traditionally used is the stock of capital, which is a measure of economic wealth. As shown in the seminal work by Jorgensen and Griliches (1967), Jorgensen et al. (1987) and Jorgensen and Stiroh (2000), what is in fact needed to measure productivity accurately is a measure of the flow of services that the capital stock generates. This issue was discussed in an earlier *Quarterly Bulletin* article (Oulton (2001)).

The main difference between a capital stock measure and a capital services measure is the way in which different assets are aggregated together. To create the aggregate stock of capital, different stocks of assets are weighted together by their asset (market) price weights.[12] In the capital services measure, on the other hand, different assets are weighted together by their rental price weights.[13] The rental price is the price that a user of the asset would have to pay to rent the asset for a period of time and, in a competitive market, it will reflect the value of the services which can be derived from the asset. The rental price is related to the price of the asset, but it also takes into account the opportunity cost of holding the asset, the cost of depreciation, and any capital gains or losses (including obsolescence) that are expected to be made by holding the asset over a period of time.

An important implication of using a services rather than a stock measure of capital input is that the services measure will give more weight to assets for which the rental price is high in relation to the asset price. If the stocks of such assets are also growing more rapidly than those of other types of assets, the services measure of aggregate capital will grow more rapidly than the stock measure of aggregate capital. In recent years ICT assets have precisely had these characteristics: the growth rates of ICT assets have been high compared with those had non-ICT assets and their rental prices are also high in relation to their asset prices.[14] Altogether, this means that the flow of services from capital has recently been growing faster than the stock of capital. [15]

Chart 2 – Growth of Capital in the United Kingdom: 1980–2003

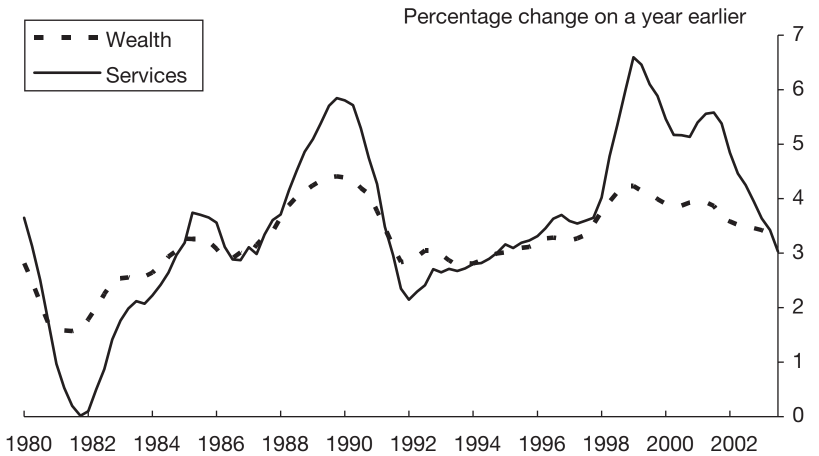


Chart 2 plots the growth rates of a services measure of capital (that accounts separately for ICT assets), against a stock measure of capital (based only on traditional ONS asset classifications: other buildings and structures, transport equipment, other machinery and equipment, intangible fixed assets) for the United Kingdom.[16] The growth of the capital services measure

12 *The asset price weight for each asset is calculated by multiplying the asset price by the asset stock and expressing it as a proportion of aggregate nominal wealth.*

13 *The rental price weight for each asset is calculated by multiplying the Hall-Jorgenson user cost of capital for the asset by the asset stock and expressing it as a proportion of aggregate nominal profits.*

14 *The reason for this is that ICT assets depreciate rapidly. The prices of most ICT assets have also been falling due to rapid technological change. This means that the rental price is high relative to the asset price, since the owner has to be compensated for both depreciation and capital losses.*

15 *For details on the calculations of the stock and services measure of aggregate capital for the United Kingdom and the sensitivity of the calculations to various assumptions on the depreciation rate, and investment prices of individual assets see Oulton and Srinivasan (2003a).*

16 *Chart 3.9 in the February 2004* Bank of England Inflation Report *presents the same data for 1993-2003.*

has been much higher than that of the capital stock measure over much of the past five years. This suggests that the Solow residual estimate in Chart 1 (which is based on a capital stock measure) may overestimate underlying total factor productivity growth over that period.

Quality adjustment of labour input[17]

In order to generate more accurate measures of TFP and aggregate supply, it is also necessary to derive a more accurate measure of aggregate labour input: one that takes into account the quality of labour and allows for changes in its composition over time.

The reason why it is important to adjust for labour quality is that a simple measure of labour input (total hours) disregards the fact that hours of work are not homogeneous: the output they can produce depends on the characteristics of individuals and of jobs. The standard measure of labour input does not capture potential changes in the quality of labour that are linked to changes in, for example, the educational composition of the workforce. For example, even if the amount of labour input (number of people or hours) remained fixed, a shift towards more skilled workers would increase supply capacity.

Determining the quality of labour inputs is not straightforward, since skills are difficult to measure directly. But if we assume that the labour market is competitive, "quality" ought to be reflected in workers' wages since workers would be paid their marginal product. The disadvantage with this approach, however, is that wages might not be a good proxy for skills if there are significant imperfections in the labour market.

Deriving a better measure of labour inputs which reflects these factors requires dividing the working population into groups, according to characteristics linked to different levels of productivity (e.g. age, education and gender),[18] and weighting each group's total hours by its productive quality (i.e. by wages). In practice, the adjusted measure we use is an index (equation 2), aggregating the growth rates of the number of hours of each group and weighting them by the group's contribution to total output,

$$\Delta \ln L_t = \sum_i \left(\frac{s_{i,t} + s_{i,t-1}}{2} \right) \ln \left(\frac{h_{i,t}}{h_{i,t-1}} \right) \quad (2)$$

where $\Delta \ln L_t$ is the growth in the quality-adjusted labour input, $h_{i,t}$ is the number of hours of group i at time t, $s_{i,t}$ is the share in the wage bill of group i, and the weights in the index are given by the average shares in periods t and $t-1$.

This formulation assumes that firms behave competitively in the labour market, so that the contribution of each group of workers to total output is equal to its share of the wage bill: a group is given a higher weight if its members have a higher wage (higher marginal product reflecting higher quality) or work more hours or both. This implies that the quality-adjusted measure will increase by more than the unadjusted measure if the most productive groups of workers (as reflected in their relative wages) experience greater growth in the number of hours (holding the wage bill shares fixed) and/or if the groups with the highest wages experience an increase in their relative wages (holding growth in the number of hours fixed).

This approach parallels the capital services calculations, where each asset is weighted by its rental price weight: in the adjusted labour input measure, each type of labour is weighted by its share in the wage bill.

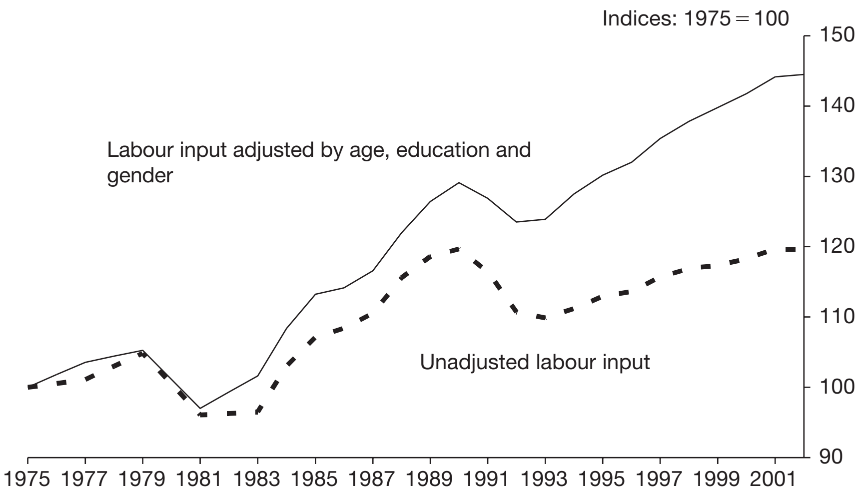
Chart 3 compares indices of unadjusted and adjusted measures of labour input where the adjusted labour input corrects for differences in age, education and gender. It is clear that the measure of labour input is biased downwards if there is no quality adjustment, especially from 1981 onwards.

The difference between the two indices reflects important changes in labour composition (or quality of hours worked). In particular, changes in the educational composition of the workforce have contributed most to the increase in labour quality. This effect has been driven mainly by the fact that highly educated people have experienced the greatest rise in the number of hours worked over these two decades. Changes in the age distribution have had a small positive impact since young people, who are the least productive in terms of hourly wages, have accounted for a declining share of the workforce. Finally, changes in the gender distribution of the workforce

17 *This section is based on ongoing research undertaken at the Bank of England.*

18 *The different groups are constructed by gender, age (16–24, 25–34, 35–44, 45–54, 55–64 (–59 for females)) and education (other qualifications, O level or equivalent, A level or equivalent, degree or equivalent).*

Chart 3 – Labour Input: Unadjusted and Adjusted for Quality



have slightly reduced our measure of labour quality. The latter reflects the fact that more women have joined the workforce, but their wage bill has increased less, partly due to their relative preference for part-time jobs, which have tended to be less well paid per hour than equivalent full-time positions.

Because the adjusted measure of labour input shown in Chart 3 has risen faster than the unadjusted one, a large proportion of what would be considered as TFP growth using raw total hours (i.e. unadjusted labour input) can actually be attributed to labour input. That is, TFP growth is significantly lower once changes in labour quality are allowed for.

There is another dimension of the data that also needs to be considered – namely, that of using disaggregated industry-level data to calculate aggregate productivity growth instead of using aggregate data directly. The following section discusses this issue.

The Bank of England industry data set

This dataset contains data for 34 industries spanning the whole UK economy, for 1970 to 2000.

For each industry, there are data on gross output and inputs of capital services, labour and intermediates, in both nominal and real terms. Capital services cover four types of non-ICT assets (structures, plant and machinery, vehicles, and intangibles), and three types of ICT assets (computers, software, and telecommunications equipment). The real intermediate index is a weighted average of domestic purchases from all other industries and from imports. Labour services are measured as hours worked, both including and excluding labour quality adjustment, based on the work discussed above.

The data set is consistent with the official UK National Accounts (as given in the 2002 *Blue Book*, Office for National Statistics (2002)) in both real and nominal terms before the following adjustments were made. To derive series for real ICT investment (and thus ICT capital), US price indices were employed for computers and software, converted to sterling terms, to deflate investment in current prices. The main reason for this is that US price indices are believed to control better for quality, whereas the UK indices do not do so fully. Since technological progress is high for ICT goods, the quality rapidly improves, and US ICT price indices therefore fall at a faster rate than the official UK ones. Also, a large upward adjustment has been made to the official level of software investment.[19]

The approach to ICT has implications for the other variables in the dataset. Changing the prices used for measuring real investment in computers and software means that the prices used to measure UK output of these products must also be adjusted. The upward adjustment to nominal software investment raises nominal GDP as measured from the expenditure side. To maintain consistency a corresponding adjustment is made to the income side of the accounts.

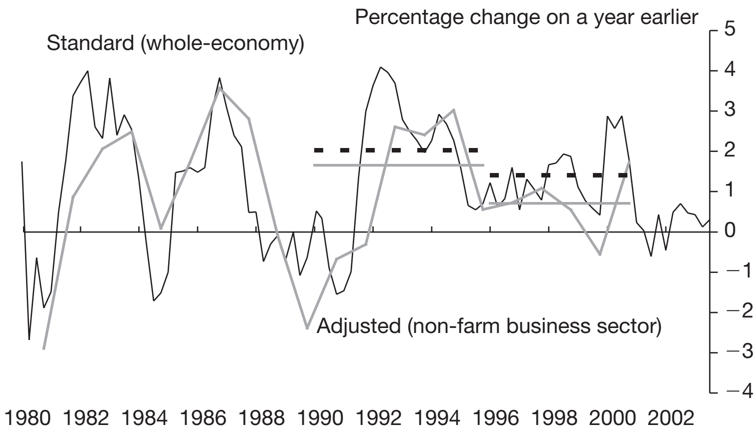
19 *This adjustment is discussed in Oulton (2002).*

Aggregate TFP growth calculated from industry data

The TFP growth rate shown in Chart 1 is calculated from aggregate data. An alternative aggregate TFP growth rate can be constructed by weighting industry-level TFP growth rates appropriately. As pointed out in Basu, Fernald and Shapiro (2001) and Bosworth and Triplett (2003) the two aggregate measures may not be identical if there are differing returns to scale across industries or heterogeneity across industries in the marginal products of identical factor inputs. It is thus preferable to calculate an aggregate TFP growth measure using industry data, since TFP growth calculated using aggregate data includes the above mentioned scale and heterogeneity effects. The Bank of England industry data set was developed to address this and other issues. It contains data for 34 industries spanning the whole UK economy, for 1970 to 2000.[20]

Using this dataset, the growth rate of aggregate TFP can be calculated by weighting industry-level TFP growth rates, which in turn are calculated using industry-specific gross output, capital services, labour and intermediate inputs measures.

Chart 4 – Growth of total factor productivity (TFP) in the United Kingdom (using disaggregated data, capital services and quality-adjusted labour input): 1980–2003



Source for Adjusted (non-farm business sector): Basu, Fernald, Oulton and Srinivasan (2003).

Chart 4 presents an aggregate TFP growth estimate for the non-farm business sector in the United Kingdom. Since the aggregate (for the non-farm business sector) is calculated using a "bottom-up" approach, the hard to measure government sector and agriculture are easy to exclude. Compared with the "top-down" aggregate TFP growth measure in Chart 1, the non-farm measure shown in Chart 4 gives quite different point estimates for some years over the 20-year time period.[21] This indicates that there could be some heterogeneity of inputs across industries. However, the overall picture remains broadly similar. The growth rate is still procyclical and there is a slowdown in UK TFP growth in the 1990s, even after moving to a capital services measure, adjusting for labour quality and aggregating from industry-level data.

Adjustment costs and variable rates of utilisation[22]

So far, we have assumed that the factors of production can be adjusted costlessly in response to changes in economic conditions. The framework can, however, be extended to take into account costs of adjustment and variable rates of utilisation.

20 *Oulton and Srinivasan (2003b) which is available on request describe the Bank of England industry data set.*

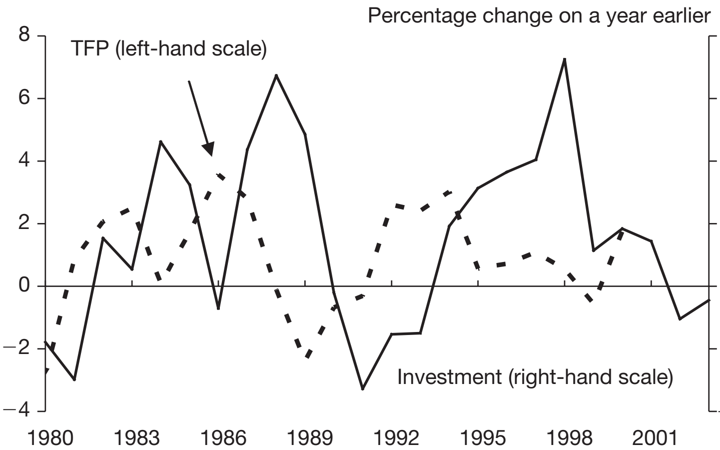
21 *The two lines in Chart 4 must be compared with caution: the standard measure is calculated using data consistent with the 2003* Blue Book *(Office for National Statistics (2003)) whereas the adjusted measure (using data from the Bank of England industry data set) is calculated using data consistent with the 2002* Blue Book *(Office for National Statistics (2002)).*

22 *This section is based on ongoing research undertaken at the Bank of England.*

Capital adjustment costs

The motivation for considering capital adjustment costs is that it may be costly for a firm to increase the amount of capital used for the production of output. One reason for this is that, when firms are investing in new capital, they may need to divert productive resources to installing the new capital rather than producing marketable output. This means that firms are essentially producing two types of products: the final product sold in the market, and the services used within the firm to install new capital. Marketable output may therefore be low during periods of high investment growth, and this would cause a downward bias in estimates of measured productivity growth.

Chart 5 – Growth of business investment and Total Factor Productivity (TFP) in the United Kingdom: 1980–2003



Source for TFP: Basu, Fernald, Oulton and Srinivasan (2003).

Chart 5 shows the growth rates of business investment (measured in chained volume terms) and total factor productivity.[23] It suggests that there may be a relation between investment and productivity growth: productivity growth slowed during the late 1980s and during the second half of the 1990s, when investment grew rapidly.

The measure of productivity growth can be extended to reflect these effects, by defining it as the fraction of output growth that cannot be accounted for by growth in the inputs, where output is defined as the joint product of observed market output and unobserved installation services. Let i be the growth rate of investment and let ϕ denote the (negative) elasticity of output with respect to investment. This measures the percentage change in marketable output that would occur following a percentage increase in investment. The Solow residual calculations can then be amended for adjustment costs – in equation 1, the growth rate of total output (including services to install capital) now equals $y - \phi i$.

The effect on output of installing new capital is not directly observable. But we can estimate it indirectly, by relating the adjustment costs to observable variables. If a firm can adjust capital without incurring any costs, it will always make sure that its productive capital is at its long-run (or normal) level, at which the cost of using one extra unit of capital (given by the rental price of capital) equals the return to one more unit of capital in the production of output. When firms face adjustment costs, the optimal level of capital will still be one at which the cost of installing one more unit of capital equals capital's expected return. But the cost of installing capital now consists of both the rental price and an adjustment cost. And the marginal return to capital consists both of the return in the production of market output and of the contribution to lower adjustment costs in the future. So the optimal level of capital is determined by a dynamic condition, which links current capital to expected future levels of capital. This relation can be used to obtain an estimate of the marginal cost of adjusting capital, from which an estimate of the elasticity of output with respect to investment can be derived.

23 *The TFP measure is the one shown in Chart 4, adjusted for capital services and labour quality, aggregated from industry data.*

Ongoing work at the Bank of England focuses on estimating capital adjustment costs for the United Kingdom, for both non-ICT and ICT assets, using the Bank of England industry data set. The results suggest that capital adjustment costs are quantitatively important, and similar in magnitude to those obtained for the United States.[24] We find that, for every 1% increase in investment in aggregate capital, output falls by between 0.02% and 0.04%. If firms invest in traditional non-ICT capital, such as buildings and plant and machinery, output falls by more, while the opposite holds for investment in ICT capital.[25] The net impact on TFP growth, however, also depends on the growth rates of the different types of investment.

These results thus suggest that the standard measure of productivity growth underestimates actual productivity growth in periods of high investment growth. In particular, the slowdown in UK total factor productivity growth in the late 1990s is less pronounced after taking into account capital adjustment costs, compared with the estimate of TFP growth that only adjusts for capital services and quality-adjusted labour services (as shown in Chart 4).

Variable rates of utilisation

If firms face adjustment costs in undertaking new investment and in hiring and firing workers, they may respond to short-run fluctuations in demand by adjusting the intensity with which labour and capital are used. For example, capital can be utilised more intensively by increasing the number of shifts, and labour can be used more intensively by increasing the effort of workers. The Solow residual would in this case overestimate productivity growth in periods when utilisation is growing rapidly, and vice versa. This would cause measured productivity to vary positively with the economic cycle, as Chart 1 suggests is in fact the case.

A measure of productivity growth that allows for these effects can be defined as the fraction of output growth that cannot be accounted for by growth in inputs or by growth in the utilisation of these inputs. Define s and e as the growth rates of the utilisation of capital and labour, respectively. Equation 1 can now be adjusted to take into account varying rates of utilisation by defining the growth of capital services as $k + s$ and the growth of labour services as $l + e$.[26]

It is not possible to observe the level of utilisation of capital and labour directly; the challenge is again to relate these unobserved variables to something that we can observe. An earlier *Quarterly Bulletin* article by Felices (2003)[27] discussed different approaches to measuring utilisation rates for labour inputs. Here we use an approach that derives links between observed variables and changes in the utilisation rates by using the optimality conditions faced by the firm.[28]

Consider a firm that would like to use more labour. The amount of labour can be thought of as a combination of the number of workers, the number of hours that each worker works, and the effort of each worker. If it is costly to hire more workers, the firm could alternatively consider increasing the number of hours worked, or worker effort. Since the alternative ways of increasing labour tend to come at a cost, it is optimal for the firm to consider all three margins at the same time. This means that the firm makes sure that the cost of a marginal increase in labour is the same irrespective of whether the firm hires more workers, increases the number of hours, or raises effort; when the number of hours is increasing, effort should therefore also be increasing. It should therefore be possible to use observed hours as a proxy for unobserved effort.

Similarly, the utilisation of capital is not observable. But to use capital more intensively, the firm has to use more labour, for example by increasing the number of hours or effort. Moreover, if capital wears out more quickly when utilisation is high, replacement investment should be high when capital utilisation is high. Also, when capital utilisation is rising, the use of intermediate inputs, such as energy inputs, should be increasing. Thus the growth of the number of hours, investment and intermediate inputs could be used as proxies for capital utilisation.

These relationships can be used to obtain an indirect estimate of utilisation. Ongoing work at the Bank of England focuses on this, by relating the growth rates of effort and capital utilisation to the growth rates of the number of hours, investment and intermediate inputs, again using the Bank of England industry data set. Because effort is unobservable, obtaining an appropriate

24 *See for example Shapiro (1986) and Basu, Fernald and Shapiro (2001).*
25 *These estimates are based on average elasticities for the sample period (1979 to 2000).*
26 *This is a simplified formula since we also need to correct the measure of productivity growth for costs of adjusting the capital stock and costs of changing the number of workers. For the exact formula, see Basu, Fernald and Shapiro (2001).*
27 *An alternative approach to modelling and estimating utilization rates for the United Kingdom is also discussed in Larsen, Neiss and Shortall (2002).*
28 *This approach is discussed in Basu and Kimball (1997) and Basu, Fernald and Shapiro (2001).*

proxy requires careful analysis of the data. For example, as discussed by Felices (2003), there has been a strong downward trend in hours in the United Kingdom, driven by mainly structural factors. So hours worked appear to respond not only to cyclical factors, but also change for structural reasons, and taking this into account properly is important when measuring unobserved utilisation.

Initial results suggest that variations in utilisation of both capital and labour may be important and that, by adjusting for variable utilisation rates, the cyclical pattern in total factor productivity growth can be reduced. This is consistent with findings for the United States, as discussed in Basu, Fernald and Shapiro (2001).

Conclusions

The Solow residual is defined as that part of output growth that cannot be explained by the growth in the primary inputs. A standard estimate of total factor productivity growth for the United Kingdom appears to be procyclical and shows a lower growth rate in the late 1990s than in the first half of the decade.

There are, however, a number of well known issues related to the measurement of the factor inputs which need to be corrected for. This article shows that these improvements in measurement could have a material impact on the estimates of total factor productivity growth. For example, using a services measure of capital input instead of a stock measure reduces estimated TFP growth for the United Kingdom in the late 1990s, since the services measure has grown faster than the stock measure. This difference is mainly due to the contribution of services from ICT capital. Using a quality-adjusted measure of labour input instead of an unadjusted measure also reduces TFP growth, since the quality-adjusted measure of labour input has been growing faster than the unadjusted one. This difference is mainly due to changes in the educational composition of the labour force. In contrast, correcting output growth to take into account costs of adjustment to changes in the level of capital input appears to increase TFP growth in periods of high investment growth, such as the late 1990s.

The net effect of these measurement improvements is complex and varies over time. While the overall picture before and after these corrections remains broadly similar, the point estimates are different. It appears that, when all these improvements are made, the decline in the growth rate of aggregate total factor productivity in the late 1990s relative to the first half of that decade is reduced but not eliminated. In addition, if both capital and labour inputs are adjusted for differing degrees of utilisation over time, the correlation of total factor productivity growth with GDP growth is reduced.

This richer treatment of input measurement is also helpful in projecting future supply capacity. This is because it enables a higher proportion of capacity growth to be identified with measurable (and so forecastable) inputs rather than with the unidentified sources of growth represented by TFP. But even after taking into account this “concealed increase in resource expansion” (Abramowitz (1956)), a significant part of output growth remains unexplained by the growth in inputs. Understanding this is the subject of future research.

Appendix

Sources and formula for the data in the charts

The formula used to calculate TFP growth is as follows:

$$z_t = y_t - 0.5^*(s_{k,t} + s_{k,t-1})k_t - 0.5^*(s_{l,t} + s_{l,t-1})l \qquad t = 1980, \ldots, 2000 \qquad \text{(A.1)}$$

TFP growth is calculated as the residual obtained from subtracting a Törnqvist index of the primary inputs (capital and labour) from the growth rate of output (value added).

When using industry-level data, the formula is modified so that the output measure is gross output and an extra term ($0.5^*(s_{m,t} + s_{m,t-1})m$) allowing for intermediate inputs (m), is subtracted from the right-hand side of Equation A.1.

Chart 1: The variables used in the TFP calculations are defined as follows:

Output	GDP at factor cost: ONS code YBHH
Capital	Wealth measure: Variant labelled ONS1 in Oulton and Srinivasan (2003a)
Labour	Total hours: ONS code YBUS
Share of capital	1 – share of labour
Share of labour	Assumed to be 0.7

GDP: GDP at market prices: ONS code ABMM

Chart 2: See Chart 3.9 of February 2004 *Bank of England Inflation Report*.

Chart 3: Bank of England estimates.

Chart 4: The growth rate of total factor productivity for the non-farm business sector is calculated by weighting industry-level TFP growth rates where the weights are the so-called "Domar weights" – the share of each industry's gross output in aggregate value added. For each industry, the output measure is gross output, the capital measure is capital services, the labour input measure is total hours (adjusted by aggregate labour quality growth), intermediate inputs are taken into account and the share of each input (capital, labour, intermediate) is calculated as a proportion of nominal gross output.

The industry level-data are from the Bank of England industry data set and are described in Oulton and Srinivasan (2003b). The UK aggregate TFP measure (for non-farm business sector) is summarized in Table 1 of and described more fully in Basu, Fernald, Oulton and Srinivasan (2003).

Chart 5: Chained-volume measure of Business Investment: ONS code NPEL.

References

Abramowitz, M (1956), "Resource and Output Trends in the United States since 1870", *American Economic Review* Vol 46, No. 2, pages 5–23.

Basu, S and Fernald J G (2000), "Why is productivity procyclical? Why do we care?", *NBER Working Paper no. 7940*.

Basu, S, Fernald, J G, Oulton, N and Srinivasan, S (2003), "The case of missing productivity growth: or, does information technology explain why productivity accelerated in the United States but not in the United Kingdom", forthcoming in *NBER Macroeconomics Annual*, Vol 18, 2003. (Also available as *NBER Working Paper no. 10010*).

Basu, S, Fernald, J G and Shapiro, M (2001), "Productivity growth in the 1990s: technology, utilization or adjustment", *Carnegie-Rochester Conference Series on Public Policy*, Vol 55 pages 117–65.

Basu, S and Kimball, M S (1997), "Cyclical productivity with unobserved input variation", *NBER Working* Paper no. 5915.

Bosworth, B P and Triplett, J E (2003), "Service Sector Productivity in the United States: Griliches' services volume revisited", manuscript, The Brookings Institution, prepared for CRIW Conference in memory of Zvi Griliches. Available at http://www.brookings.edu/dybdocroot/views/papers/bosworth/20030919.pdf

Burnside, C, Eichenbaum, M S and Rebelo, S T (1995), "Capital utilization and returns to scale", in Bernanke B S and Rotemberg J J (eds) *NBER Macroeconomics Annual*, Cambridge MA, MIT Press, pages 67–110.

Felices, G (2003), "Assessing the extent of labour hoarding", *Bank of England Quarterly Bulletin*, Summer, pages 198–206.

Hulten, C R (2001), "Total factor productivity: a short biography", in Hulten, C R, Dean, E R, and Harper, MJ, (eds) *New Developments in Productivity Analysis*, Chicago, Chicago University Press.

Jorgenson, D W, Gollop, F M , and Fraumeni, B M (1987), *Productivity and US economic growth*, Cambridge MA, Harvard University Press.

Jorgenson, D W and Griliches, Z (1967), "The explanation of productivity change", *Review of Economic Studies*, Vol 34, pages 249–83.

Jorgenson, D W and Stiroh, K J (2000), "Raising the speed limit: US economic growth in the information age", *Brookings Papers on Economic Activity*, Vol 1, pages 125–211.

Larsen, J, Neiss, K and Shortall, F (2002), "Factor utilisation and productivity estimates for the United Kingdom", *Bank of England Working Paper no. 162*.

Office for National Statistics (2002), *United Kingdom National Accounts: The Blue Book 2002* London: The Stationery Office.

Office for National Statistics (2003), *United Kingdom National Accounts: The Blue Book 2003* London: The Stationery Office.

Oulton, N (2001), "Measuring capital services in the United Kingdom" *Bank of England Quarterly Bulletin*, Autumn, pages 295–309.

Oulton, N (2002), "ICT and productivity growth in the UK", *Oxford Review of Economic Policy*, Vol 18, pages 363–79.

Oulton, N and Srinivasan, S (2003a), "Capital stocks, capital services, and depreciation: an integrated framework", *Bank of England Working Paper no. 192*.

Oulton, N and Srinivasan, S (2003b), "The Bank of England Industry Dataset", Bank of England mimeo. Available from authors on request. Also available at: www.nber.org/books/macro18/basu-et-al9–22–03_u.k.data-appendix.pdf

Shapiro, M D (1986), "The dynamic demand for capital and labour", *The Quarterly Journal of Economics*, Vol 101, pages 513–42.

Stiroh, K (2002), "Information Technology and the U.S. Productivity Revival: What Do the Industry Data Say?", *American Economic Review*, Vol 92, pages 1559–76.

Charlotta Groth, Maria Gutierrez-Domenech and Sylaja Srinivasan
Bank of England
Structural Economic Analysis Division

Information technology and productivity changes in the banking industry

Luca Casolaro and Giorgio Gobbi (Bank of Italy)

1. Introduction[1]

The rapidly increasing use of computers in producing and delivering goods and services has spurred a large literature on the effects of information technologies (IT) on productivity growth. Recent research has documented an important contribution of IT capital boosting total factor productivity in the US economy in the late 1990s. Even larger effects are expected in the future as long as learning processes gain momentum and innovations are paralleled by appropriate complementary investment. More skeptical views maintain that the observed increase in productivity is largely due to cyclical factors and that it is not necessarily a consequence of IT capital deepening.

The debate, initially based on industry-level data, has rapidly extended to empirical work using detailed information at the firm level. Industry-level studies, in fact, implying an aggregation over a number of different products and services, are potentially biased by the mismeasurement of quality changes in output, thus underscoring the magnitude of productivity changes. Studies using micro-data have found a strong positive correlation between IT capital accumulation and firms' productivity growth. Also, the impact of IT on firm performance appears to be higher when complemented by innovations in workplace organization and a large staff of highly skilled workers.

This paper contributes to the microeconomic strand of research by studying the effects of IT capital deepening in the banking sector. Several reasons make the banking system a particularly interesting industry for testing the impact of new technologies on productivity growth. First, from a macroeconomic point of view financial services account for between 5 and 10 percent of GDP in most industrialized countries. Banks contribute to a large proportion of financial-related output playing an important role in key activities for the functioning of the economy, such as the allocation of capital and the provision of liquidity, payment and safekeeping services. Substantial improvements in the level of efficiency of these services can bring widespread welfare effects (Summers, 2000). Second, the acquisition and the treatment of information is a central activity in banking and the developments of productivity growth in this sector are likely to be highly sensitive to the extensive use of information technology (Berger, 2003). Finally, the stability of the banking system is a primary policy objective. Innovations affecting production costs and profit opportunities can have major consequences on the industry structure and need to be monitored by regulators.

Productivity developments in the banking industry have been extensively studied in relation to deregulation (Humphrey and Pulley, 1997), mergers and acquisitions (Akhavein et al, 1997) and changes in competition (Berger and Mester, 2001). However, limited attention has been paid to the effects of the diffusion of IT, mainly because of the paucity of appropriate quantitative information. Our database, which refers to a panel of more than 600 Italian banks over the period 1989–2000, includes information on both investments and expenses related to computers and software, which allows us to compute estimates of IT capital stock and its user cost at the firm level. Our sample period starts in 1989, when IT technologies had already been adopted by the vast majority of banks, and extends up to year 2000, covering a further wave of IT capital deepening.

1 *The authors thank Allen Berger, Luigi Guiso, Stefano Siviero and all the audience participants at the Bank of Italy and the FED Board seminars for helpful suggestions. We are also indebted with Tim Coelli who kindly made available the econometric software FRONTIER 4.1, used in part of the estimations. Ginette Eramo and Roberto Felici provided outstanding research assistance. All the remaining errors are ours. The views expressed in this paper do not necessarily represent those of the Bank of Italy.*

Because of the heterogeneity of the shocks that affected the banking industry over the sample period (e.g. macroeconomic fluctuations, changes in regulation, demand shifts), any standard measure of total factor productivity growth (e.g. Solow residuals) is likely to reflect the effects of a large number of factors (Kumbhakar, 2000). The availability of data on investment in hardware and software allows us to test whether changes in total factor productivity (TFP) are correlated with the diffusion of the new technologies. We estimate both cost and profit functions using a methodology developed by Battese and Coelli (1993). A stochastic frontier is jointly estimated with an equation modeling the banks' inefficiency vector as a displacement from the best practice. In the frontier equation TFP shifts are also accounted for by the interaction of a time polynomial with output and input prices.

Using this methodology we aim to capture two different effects on productivity. The first is the shift of the frontier traced by the firms using the best practices within the industry. The second is the dynamics of banks' inefficiency, which are estimated as the distance from the most efficient banks on the frontier; this measure reflects the pattern of diffusion of the new technologies, which is likely to differ across firms. A well-documented, stylized fact, is that the usage of new technologies over time typically follows an S-curve (Geroski, 2000). As long as more and more firms introduce the usage of the new, TFP improving technology on a large scale, we should observe a catching-up effect on the average productivity growth of the industry.

We find that IT capital accumulation is strongly correlated with banks' cost and profit efficiency, implying a strong catching-up effect over the decade. We also find a positive correlation between cost and profit frontier shifts and several measures of IT capital intensity. Moreover, our results confirm, albeit indirectly, that organization does matter. In particular, we find that productivity gains are larger for banks combining in-firm IT capital-deepening with outsourcing of a substantial part of IT expenses, especially those implying a higher number of IT dedicated employees.

The paper is organized as follow. Section 2 reviews prior analysis of the effect of IT capital on productivity growth and on the estimation of inefficiency and productivity in banking. Section 3 lays out the methodology for the empirical analysis and the models to be estimated. In section 4 we present the data and the methodology to compute our measure of IT capital and in section 5 we display our empirical results. Section 6 concludes.

2. Related studies and motivation

Over the last 20 years firms have responded to the fall in the (quality-adjusted) price of computers and related technologies by substituting others production factors for IT capital. Computers share with other landmark technological innovations the feature of being a multipurpose technology, which can be adopted into virtually every economic activity. Historically, this kind of innovation has determined major shifts in the productivity of the economy, even if the available empirical evidence varies greatly as to the impact of IT on productivity according to the time period, country and analytical framework.

Studies of TFP growth rates in the US during the '80s and the early '90s generally found that the use of computers had a negligible impact on productivity (Roach, 1987; Berndt and Morrison, 1991).[2] More recent papers (e.g. Jorgenson and Stiroh 2000; Oliner and Sichel, 2000; Jorgenson, 2001) find evidence of a non-negligible contribution of IT capital to the recent productivity growth in the US. Daveri (2001) examines the diffusion of IT in Europe during the '90s finding a strong increase of IT capital in UE countries, albeit with a substantial delay with respect to US experience. He also finds that cross-country differences in growth are correlated with the adoption of IT, even if a causal link has not been yet established. More skeptical scholars argue that productivity gains are largely to be ascribed to cyclical factors, claiming that there is clear evidence of a productivity increase only in the computer-producing industry (e.g. Gordon, 1999).

Most of these studies are based on aggregate industry data, which may underestimate the effect of IT on productivity for two reasons. First, the use of computers is often associated with large changes in the quality of output that are difficult to measure accurately (Boskin et al., 1997), especially aggregating broad ranges of products and services. Second, the use of digital technologies is likely to require time to adjust workplace organization and employees skills, and industry data can average out large cross-firm variability due to different stages in IT adoption.

2 *The negative results of this vintage of papers inspired the much quoted remark by Robert Solow "we see computers everywhere but in the productivity statistics" (Solow, 1987).*

Consequently, a number of papers have addressed the issue using micro-data. Brynjolfsson and Hitt (2000) estimate a production function for a panel of 600 large US firms. They find that over a period of three to seven years the contribution of IT investments to output growth substantially exceeds its factor share, implying a positive effect of computers on lung-run productivity growth. The results also suggest that IT capital deepening is part of a larger system of technological and organizational changes within firms. Bresnahan et al. (2002) also find a positive correlation between the accumulation of IT capital, innovation in workplace organization and the use of more highly skilled workers. They also estimate a positive correlation between firms' IT capital stock and their value added. Using Italian micro data, Bugamelli and Pagano (2001) investigate the diffusion of IT in Italian manufacturing firms. They find evidence of a positive correlation between IT investments, reorganizations of the production process and the level of human capital in the labor force.

To our knowledge, the only two academic papers specifically investigating the effects of computers in the banking industry are Parsons et al. (1993) and Autor et al. (2000). Both of them use data from case studies. Parsons et al. (1993) estimate a cost function using data from a large Canadian bank over the period 1974–1987 and find a weak but significant correlation between productivity growth and the adoption of computers. Autor et al. (2000) examine the introduction of automatic image processing of checks in one of the 20 largest banks in the United States. They argue that computers are a substitute for low skilled works in standardized tasks, but complementary to computer-skilled labor. Furthermore they emphasize the interdependence between technological and organizational change.

A number of studies have tried to infer the effects of technical change on banking from total factor productivity changes, estimated by fitting cost or profit functions to micro-data. The specification of dual functions in place of the standard primal production function is due to the multiple nature of the banks' output and to the measurement problems associated with services provided through the activities of lending and fund-raising (Hanckock, 1991). For instance, Hunter and Timme (1986) estimate a cost function using data on a sample of large US bank holding companies over the period 1972–1982. Technical change is modeled introducing a time polynomial as an argument of the cost function. Their major finding is that, over the sample, technical change contributed substantially to lowering average costs, especially for larger banks.

Average industry TFP changes, however, can be a poor proxy of technical change whenever there is a wide dispersion of efficiency levels across firms: an increase in average TFP over a given time lapse can be ascribed either to technical progress or to a rise in average technical or allocative efficiency (Bauer, 1990). A large body of applied literature has documented that this is in fact the case in the banking industry (Ferrier and Lovell, 1990, Berger, 1993, Berger, Hunter and Timme, 1993, Berger and Mester, 1997). Using both parametric and non-parametric frontier techniques, in the literature TFP growth has been decomposed into changes in average efficiency and frontier shifts. The latter measures the productivity growth of the banks adopting the best practices in the industry, which is likely to be correlated with technical change. Approximating technical change with frontier shifts, both Bauer et al. (1993) and Humphrey (1992) found that it had a negative effect on the productivity of US banks for most of the '80s. More recently Kumbhakar et al. (2001) estimate a stochastic profit frontier for Spanish saving banks during the period 1986–1997, finding a positive effect of technical change on TFP. Berger and Mester (2001) also find a correlation between technical change and productivity growth in US banking in the second half of the '90s. Alam (2001), applying a non parametric approach to US banks, finds that during the '80s, productivity growth in banking was driven mainly by technological changes rather than convergence to the efficient frontier.

The general pattern of negligible or negative effects of technical change on productivity growth in the '80s, followed by positive effects in the '90s, seems to fit the banking industry data as well. The existing literature, however, has two major shortcomings.

First, the identification of technical change with frontier shifts in not straightforward, especially when other shocks from different sources affect the path of productivity growth of banks adopting the best practices in the industry. For example, Berg et al. (1992), in a study on Norwegian banking in the '80s, explain the upward shift in the production frontier as an effect of the absorption of the post-deregulation excess capacity. Grifell-Tatje and Lovell (1996) find that deregulation induced a sharp decline in the productivity of the top performer Spanish saving banks in the late '80s.

Second, the dispersion of bank efficiency levels also contains useful information about the effect of technical change. Several studies have tried to measure the average response of the industry to different types of shocks, such as changes in regulation (Humphrey and Pulley, 1997) or mergers and acquisitions (Akhavein et al., 1997). Others have tried to characterize the efficiency levels in terms of market and bank characteristics (Berger and Mester, 1997), or to focus on particular bank features, such as the ownership structure (Altunbas et al., 2001) or the level of specialization (Vennet, 2002).[3] Lack of data, however, has prevented research to directly explore whether displacements from the best practice are correlated with technological innovation. The pace at which different banks have invested in IT capital may differ substantially. On the one hand, the speed of adoption of the new technologies can be hampered by lack of information on how to implement them or by the need for workforce skills that are obtainable only through learning-by-doing processes (following the "epidemic" models of technology diffusion). On the other hand, different firms may want to adopt the new technology at different times because of idiosyncratic adjustment costs (following the "probit" model of diffusion). A difference in the adoption rate of new technologies across banks is likely to be reflected in the distribution of efficiency levels. The methodology used in this paper allows us to tackle both these drawbacks, thanks to the availability of data on IT capital.

3. Methodology

3.1. Stochastic frontiers and inefficiency

To estimate the pattern of inefficiency of the banking sector we apply the stochastic frontier methodology developed by Battese and Coelli (1992), which allows the estimation of both the frontier shifts and the time patterns of the inefficiency vector. The joint estimation of the frontier and the equation relating the inefficiency levels to a set of covariates overcomes the problem caused by the orthogonality assumption usually imposed between residuals and regressors in the standard inefficiency regressions.[4]

Considering a generic (cost or profit) frontier function, we estimate the following system:

$$\mathrm{F}_{it} = f\left[Y_{nit}, P_{mit}, t\right] + u_{it} + \nu_{it} \tag{1}$$

$$u_{it} = \delta z_{it} + w_{it} \tag{2}$$

where F_{it} is the volume of costs or profits, y_{it} are the values of the n output or price variables and the p_{it} are the m input prices. The residual of equation (1) is a composite term formed by a standard white noise residual v_{it} plus an asymmetric distribution term u_{it}, accounting for inefficiency. In equation (2) the u_{it} term is specified as a stochastic process with mean dependent from a vector of covariates z_{it}, and a random component w_{it} defined as the truncation of an independent normal distribution with mean zero and variance σ_U^2, such that u_{it} is a non-negative truncation of the normal distribution $N(\delta z_{it}, \sigma_U^2)$ (see Battese and Coelli, 1995).[5] This specification partially mitigates the usual problems the stochastic frontier methodology related with the *ad hoc* assumptions imposed to the inefficiency distribution.[6] The model allows also inefficiency levels to vary over time without imposing the same ordering of firms in terms of efficiency for each period, as in Battese and Coelli (1992).[7] In our setting a bank can be more efficient than another one year and less efficient the following.

The u_i term can be identified, following Jondrow et al. (1982), conditional on the estimated residual $e_i = v_i + u_i$. Battese and Coelli (1988) pointed out that the best predictor of cost inefficiency for firm i is the expression e^{ui}, for which the following expression holds (Battese and Coelli, 1993):

$$E\left[e^{u_i} \mid e_i\right] = \left[\frac{\dfrac{\phi}{\sigma_a}\left[(1-\gamma)z_{it}\delta - \gamma e_i\right] - \sigma_a}{\dfrac{\phi}{\sigma_a}\left[(1-\gamma)z_{it}\delta - \gamma e_i\right]}\right] e^{[\gamma e_i + \sigma_a^2/2]} \tag{3}$$

3 *By contrast, the issue is widely debated within the profession. McKinsey (2002) claims that many of the post-1995 IT investments were aimed at rising revenue-increase activities that have yet to be repaid, leading to a decline in banking productivity due to excess capacity.*

4 *The typical procedure in the past literature was first to estimate the stochastic frontier and predict inefficiency under the assumption of identical distributions, then to estimate an inefficiency equation using the predicted* u_{it} *term, in contradiction with the prior distributional assumption (see Khumbhakar, Ghosh and McGukin, 1991).*

5 *The assumption that the* u_{it} *are independent distributed* $\forall$ i *and* t*, made to simplify the model, is a restriction because it does not account for heteroskedasticity or for possible correlation structures among the residuals.*

6 *In the results of our estimates, however, we control for their robustness to different inefficiency estimation methods.*

7 *The relation* $z_{it}\delta + w_{it} > z_{i't}\delta + w_{i't}$ *for* $i \neq i'$ *does not imply* $z_{it'}\delta + w_{it'} > z_{i't'}\delta + w_{i't'}$ *for* $t \neq t'$.

where σ_a is equal to the square root of the expression $\gamma(1-\gamma)(\sigma_U^2 + \sigma_V^2)$, ϕ is the density function of a standard normal random variable $\gamma = \sigma_U^2/(\sigma_U^2 + \sigma_V^2)$, σ_U^2 and σ_V^2 being respectively the estimated variance of the inefficiency term u_{it} and of the white noise v_{it}.

The parameter γ estimated in the regression represents the part of total variance around the frontier that is explained by inefficiency. A value of γ equal to zero means that the deviations from the frontier are entirely due to noise, while a value equal to one indicates that all the deviations are due to cost inefficiency. The expectation of inefficiency u_{it} conditional to the entire residual e_{it} measures the ratio between the costs (profits) of a generic firm i and those of the most efficient firm i^*, which is on the frontier with an inefficiency level equal to zero ($\sigma_{ui}^* = 1$).

Taking the time derivative of the frontier (1), which is assumed to be multiplicative in its arguments, we obtain the value of the cost and profit frontier shift, defined as:

$$\frac{\partial F_{it}}{\partial t} = \sum_{l=1}^{L} \frac{\partial f(t,.)}{\partial t} + \sum_{n=1}^{N} \frac{\partial f(Y,.)}{\partial t} + \sum_{m=1}^{M} \frac{\partial f(P,.)}{\partial t} \quad (4)$$

where L is the power of the time polynomial included in the frontier to account for non-neutral technological progress, N is the number of outputs and M is the number of inputs. For each bank i a time derivative is computed, representing the projection on the frontier of the bank's time shift in cost or profit, after controlling for its input-output mix and for its level of inefficiency.

3.2. The empirical model

We first estimate a stochastic translog cost function with a time polynomial included as a production factor to allow for non-neutral technological progress. The estimated cost frontier is the following:

$$C = \beta_0 + \sum_{n=1}^{N} \beta_{yn} y_n + \sum_{m=1}^{M} \beta_{pm} p_m + \frac{1}{2}\left(\sum_{n=1}^{N}\sum_{l=1}^{N} \beta_{y_n y_l} y_n y_l + \sum_{m=1}^{M}\sum_{k=1}^{M} \beta_{p_m p_k} p_m p_k \right)$$
$$+ \sum_{n=1}^{N}\sum_{m=1}^{M} \beta_{y_n p_m} y_n p_m + \sum_{n=1}^{N} \beta_{ty_n} y_n t + \sum_{m=1}^{M} \beta_{tp_m} p_m t + \beta_{t1} t + \beta_{t2} t^2 + \phi_1 GNP_t + \phi_2 r_t + v + u \quad (5)$$

where the N outputs are denoted by y_n, the M input prices by p_m, all expressed in logs and in differences from the sample mean and the standard symmetry and linear homogeneity conditions are imposed.[8] We include in the model the growth rate of the real *GNP* and the real official discount rate in order to control for the business cycle. The cost inefficiency term u_{it} measures how close the costs of bank i at time t are to those of a bank on the efficient frontier, producing the same output vector under the same technical constraints.

Berger and Mester (1997) argue that in banking analysis the estimation of a profit frontier is more appropriate because the use of a profit function accounts for both output and input side errors in the bank's choice. In the standard profit analysis output prices are taken as exogenous, considering profit inefficiency as a sub-optimal choice with respect to input-output relative prices. This model, however, is not the best suited for the banking industry, where a large share of banks' revenue originates from bilateral contracting in which banks have some bargaining power. We therefore adopt an alternative concept of profit function, proposed by Humphrey and Pulley (1997), where output prices are allowed to vary while output is assumed to be constant, as in the cost function. Using this approach, errors in the choice of outputs do not affect efficiency, while errors in output prices do. The alternative profit approach can be useful when some of the standard assumptions of the profit model cannot be satisfied, as it usually happens in the banking sector. First, in the case of unmeasurable differences in the output qualities between banks, the alternative profit approach can be used because it accounts for the additional output coming from higher quality in services. Also, output prices are usually constructed from quantity data and cannot adequately account for the variance of profits.

The efficiency estimation obtained by this approach is a measure of how close a bank is to its maximum profit, given an input price-output quantity mix, i.e. the proportion of profits that

8 *The subscripts* i *and* t *are omitted for simplicity.*

bank i is earning with respect to maximum profits, due to excessive costs and/or inadequate revenues.[9] We therefore estimate the following translog alternative profit frontier:

$$\pi = \beta_0 + \sum_{n=1}^{N} \beta_{yn} y_n + \sum_{m=1}^{M} \beta_{pm} p_m + \frac{1}{2}\left(\sum_{n=1}^{N}\sum_{l=1}^{N} \beta_{y_n y_l} y_n y_l + \sum_{m=1}^{M}\sum_{k=1}^{M} \beta_{p_m p_k} p_m p_k \right)$$
$$+ \sum_{n=1}^{N}\sum_{m=1}^{M} \beta_{y_n p_m} y_n p_m + \sum_{n=1}^{N} \beta_{ty_n} y_n t + \sum_{m=1}^{M} \beta_{tp_m} p_m t + \beta_{t1} t + \beta_{t2} t^2 + \phi_1 GNP_t + \phi_2 r_t + v + u \qquad (6)$$

where the variable u measures inefficiency in terms of forfeited profits. In both cost and profit analysis the inefficiency term is modeled as follows:

$$u_{it} = \delta_0 + \sum_{j=1}^{J} \delta_j z_{itj} + \delta_t t + w_{it} \qquad (7)$$

where δ_0 is a constant term, z_{it} are individual time-varying covariates and t is a linear time trend to account for time patterns of inefficiency. We jointly estimate equation (5) (or equation (6)) and equation (7) using maximum likelihood techniques[10] that produce asymptotically consistent and efficient estimates.[11]

As a last step we analyze the frontier shift determinants in order to investigate the existence of a correlation between IT capital and the cost and profit shifts of banks over time. The idea behind this exercise is that frontier shifts are driven by the introduction of new best practices along with the increase of IT capital.[12]

We first compute the time derivative of the frontier function for each bank in each year, which is given by:

$$\mathrm{FS}_{it} = \frac{\partial f}{\partial t} = \beta_{t1} + 2\beta_{t2} t + \sum_{i=1}^{N} \beta_{yt_i} y_i + \sum_{j=1}^{M} \beta_{pt_j} p_j \qquad (8)$$

where the f stands for the cost or the profit function. We estimate the panel of banks' frontier shifts for each year with a fixed effects regression of the kind:

$$\mathrm{FS}_{it} = \beta_2 TECH_{it} + \sum_{j=1}^{J} \beta_j w_{jit} + \lambda_i + \eta_t + v_{it} \qquad (9)$$

where $TECH_{it}$ represents a vector of variables that measure the degree of IT adoption for bank i, the w_{jit}s are time-varying bank specific variables, the λ_is are the individual fixed effects and the η_ts are time dummies.

4. Data

4.1. Banks

In the first part of the '90s the IT-related expenses of Italian banks accounted for about 9 percent of total operating costs, peaking above 13 percent in 1999 with the introduction of the single currency and the Y2K (Table 1). Annual investment in hardware, adjusted for hedonic price changes, increased by six times and that on software by three times (Table 2). In 2001 the total real value of hardware, software and EDP specific plants per employee was almost six times higher than in 1989.

9 *In the last years a number of empirical studies on banking efficiency adopted the alternative profit approach. See, between the others (Altunbas et al., 2001, Clark and Siems, 2002, Vennet, 2002).*
10 *The likelihood function of the system is shown in Battese and Coelli (1993).*
11 *We do not include the factor share equations, which are derived from the Shepard's Lemma and Roy's Identity restrictions: the inclusion of factor demand equations would have increased the efficiency of the estimates but would have forced the banking system not to display allocative inefficiency.*
12 *As pointed out by Salter in his seminal empirical investigation on productivity and technical change: ". . . gross investment is the vehicle of new techniques, and the rate of such investment determines how rapidly new techniques are brought into general use and are effective in raising productivity." (Salter, 1966; pag. 17).*

Moreover, during the '90s the banking industry in Italy, as well as in many other developed countries, underwent far-reaching transformations: a new banking law was passed in 1993; virtually all banks under public control have been privatized; a wave of mergers and acquisitions has swept through the industry; large demand shifts have caused prolonged periods of excess capacity. In the first five years of the decade net return on equity fell from 9.91 to 1.21, while the cost-income ratio increased by almost seven percentage points. Performance subsequently recovered along with a decrease in the burden of staff costs and bad loans and with a strong increase in revenues from services.

In our empirical analysis we employ data referring to an unbalanced panel of 618 Italian banks over the period 1989–2000.[13] We use information from the Supervisory Returns Database of the Bank of Italy, which includes detailed information on IT expenses and investment relating to software, hardware, outsourcing and the number of employees working in the computing centers.

We adjust for banks disappearing in the sample period owing to mergers and acquisitions by computing *pro-forma* balance sheets and profit and loss accounts. Possible effects of M&A on productivity are accounted for by introducing the share of the target banks into the joint assets of the target and the acquiring banks among the explicative variables of both the inefficiencies and the frontier-shifts equations. Acquisitions that maintain the charter of the target bank within a banking group are controlled for by means of a dummy variable.

We exclude from the sample branches of foreign banks because their balance sheets and profit and loss accounts are strongly affected by the relationships with their parent banks. We also drop 170 banks presenting data problems. In particular, for a large number of small-sized banks, information on IT is rather noisy due to the small scale of activity. This is the case of banks that outsource a large part of data processing activities for which it has not been possible to rely on investment flow to compute IT capital stock (see below). However, the number of banks for which this information becomes available increases over time.

In order to assess cross-sectional differences we divide the sample banks into three groups according to their average total assets during the sample period. The first group includes the smallest two thirds of the sample and contains almost exclusively cooperative banks: we call it the *small banks group*. In the second group, our *medium banks group*, we include banks between the top 33 and the top 10 percent of total asset distribution, while the *large banks* are those in the top 10 percent of the distribution.

4.2. Outputs and inputs

The appropriate representation of the bank production process and, in particular, the definition of bank output is a long-standing and controversial issue. In this paper we follow what has come to be known as the *value added approach*, in which all the activities having substantial value added are considered outputs of the bank. We classify services produced and supplied by banks into two broad areas.

The first group includes services for which banks charge direct fees such as payment execution, safekeeping, brokerage and asset management.[14] These outputs are likely to be standardized and in principle it would be possible to have appropriate physical measures for them, such as the number of transactions. However, actual data are collected in a rather aggregate form, and owing to quality heterogeneity they are at best only proxies for output levels. Moreover, banks supply a large number of different services which can hardly be accommodated in an econometric model. We have therefore combined the information on three main types of proxies in a single composite output using their shares in fee income as weights. Specifically, we have the turnover of current accounts as a proxy for payment services, the total amount of deposited securities as a proxy for both safekeeping and asset management, and loan guarantees as a proxy for off-balance intermediation activity.

The second group of services stems from typical intermediation activity: raising and lending funds. According to the theoretical micro-foundations of financial intermediation, banks "produce" financial contracts exploiting economies of scale in transaction technologies as well as in screening and monitoring activities. The remuneration for the provision of intermediation services is given by the spread between the interest paid on financial liabilities and the interest charged on assets. Since this kind of activity tends to be customer-specific, there is no clear-cut distinction between

13 *This number refers to the sample obtained after the exclusion mentioned above.*

14 *Clark and Siems (2002) show that banking efficiency is affected by the off-balance-sheet activities of the banks. They find that the inclusion of variables related with non interest income can adequately proxy the effect of the off-balance-sheet activities, contributing to explain interbank differences in profit and cost efficiency.*

prices and quantities. Intermediation services are to a large extent not standardized and it is hard to find compelling physical measures. In the applied literature they are usually assumed to be proportional to the face value of assets and liabilities, meaning that a €100 loan corresponds to two times the amount of services produced with a €50 loan.[15] In this study the outstanding amounts of loans to and deposits from the non-bank sector are assumed to be the main carriers of these kinds of services. The exclusion of other items, such as interbank accounts and securities, can be motivated on the ground that they contribute little to the value added once their user costs are properly taken into account (Hancock, 1991 and Fixler, 1993). We consider separately short-term loans (those with original maturity up to 18 months) and medium- and long-term loans on the grounds that they are likely to differ substantially in terms of the amount of resources they require for their origination, screening and monitoring. In Italy, most of the short-term loans consist of non-collateralized commercial and industrial lending. By contrast medium and long-term loans mainly include mortgaged lending (Pozzolo, 2001). On the input side we have considered labor and two types of physical capital services: IT capital and other premises and equipment. For operating expenses not related to staff and capital expenditure there is no information enabling us to distinguish between prices and quantities. We have therefore assumed that they enter in the production in fixed proportion with other inputs. The IT price is calculated as a usage cost, given by the ratio between IT expenses and the value of the capital stock computed before (see below). As a proxy for the price of labor we use the average wage, while the price of the other type of physical capital is constructed as the ratio of the capital expenses divided by its book value.[16]

Total operating expenses are the left-hand side variable in the frontier estimation. Price homogeneity of the cost function has been imposed by normalizing both costs and input prices by the price of labor. Outputs and inputs have been deflated by the consumer price index in order to avoid the estimated coefficients of the time variable being affected by inflation. In the profit function the dependent variable is the value of total revenue before taxes.

4.3. IT capital

The IT capital stock has been computed by applying the permanent inventory method to the real investment in hardware, software and premises for computing equipment. The value of banks' investment has been deflated using hedonic price indexes of software and hardware developed by the Bureau of Economic Analysis (BEA)[17] and adjusted for the variation in the ITL/USD exchange rate.[18] The US deflator is the only one available for each group of products; its application to Italian data may be justified by the high share of imported IT equipment and the existence of a global market for these products. Capital stock is obtained as the sum of past investment flows, weighted by the relative efficiency in production of different vintages. We assume the depreciation rate is constant in time but different across goods. Following Seskin (1999) and Jorgenson and Stiroh (2000) software is assumed to depreciate at a yearly rate of 44 percent, hardware and dedicated premises by 32 percent yearly.

4.4. Explanatory variables of inefficiency levels and frontier shifts

On the right-hand side of the inefficiency and the frontier-shift regressions we include several explanatory variables to control for banks' individual and environmental characteristics. To investigate the impact of IT capital accumulation on productivity we introduce the variable IT_CAP, representing the amount of IT capital stock per employees, split in a more detailed analysis into software and hardware capital stock (respectively SOFT_CAP and HARD_CAP). The variable IT_STAFF represents staff costs for IT personnel divided by total staff costs; the variable OUTSOURCE is given by the fraction of IT expenses coming from outsourcing of computing services.

The development of alternative distributive channels for bank products is measured by the variable REMOTE, defined as the number of remote customers standardized by the total number of current accounts, and the variable ATM_BR, given by the number of automatic teller machines (ATM) divided by the number of branches. We also try to capture features linked to banks' different specialization through the variables SMALL, which stands for the fraction of

15 *This is known as the intermediation approach and is usually contrasted with the production approach in which output is measured by the number of accounts.*
16 *We also included estimates of the value of leases consistently with the definition of expenses.*
17 *Data have been downloaded from the web site www.bea.org.*
18 *We implicitly assume that exchange-rate movements are entirely translated onto consumers.*

loans to small and medium-sized firms, and SERVICE, the ratio of income from services to gross income. We include three variables to account for the composition and the degree of utilization of the staff. The variable MANAGERS is defined as the ratio of managers' costs to total staff costs, REDUNDANCE represents the fraction of staff made redundant by the bank, STAFF_BR is given by the average number of employees per branch. The variable CAP, the ratio of capital and reserves to total assets, controls for the degree of capitalization of the bank, the variable M&A measures the increase in assets obtained with mergers. We also include the variable AGE, given by the expression: $AGE = 1 - 1/\sqrt{year - foundyear}$ where *foundyear* is the year of foundation of the bank and year is the current date. In this way, differences in the age of the banks are more important for relatively younger banks. The dummies NW, NE, SOUTH, ISL account for geographical differences, while the dummies GROUP1 and GROUP2 have value one if the bank is part of a large group (the top 25 percent of the group's total assets distribution) or of a medium or small group (the last 75 percent) respectively. Summary statistics of the variables used in the empirical analysis are reported in Table 3.

5. Results

5.1. Frontier estimates and "catching-up" effects

Figure 1 contains a plot of the average inefficiencies estimated from the cost frontier:[19] the overall sample inefficiency averages 9.2 percent, with substantial differences across size classes. The result is consistent with previous studies that found X-inefficiency effects to dominate scale inefficiency (Berger and Humphrey, 1991; Berger, 1993). The results of the profit inefficiency analysis are similar to those of the cost one, even if the displacements from the profit frontier present a smaller estimated average value (8.1 percent) because of the inclusion of quality improvements in the profit specification. Coherently with previous research (Berger and Mester, 2001), the patterns of inefficiency computed with the profit and cost specifications present substantial differences. Small banks always exhibit the lowest inefficiency levels, while medium and large banks display higher distance from the efficient frontier. The time pattern is the same for all bank classes, displaying an increasing inefficiency over the decade. This finding can be attributed to the different speed at which banks respond to the structural changes and the changes in demand experienced by the industry in the '90s.

The parameters of the inefficiency equations, jointly estimated with the translog cost and profit functions, shed some light on the determination of the inefficiency pattern in the Italian banking industry over the decade (Table 4). The effect of IT capital stock on inefficiency, controlling for geographical characteristics and changes in the structure of the market, is negative and statistically significant in both equations. This result confirms that the most IT-capitalized banks are those with the strongest catching-up effect and those that are more likely to be closer to the efficient cost and profit frontier. In particular, for the median bank, an increase of one standard deviation in the level of IT capital per employee will generate a 3.7 percent cut in costs and a 7.5 percent profit boost due to decreased inefficiency.

The variable SERVICE, representing the fraction of earnings not coming from loans, has a negative impact on inefficiency, meaning that the banks which have responded more quickly to shifts on the demand side have also benefited in terms both of lower costs and higher profits.

The coefficient of the variable REMOTE is negative, albeit significative only in the profit regression, indicating a positive correlation between efficiency and innovation in delivery channels of banking services, even if we cannot exclude a reverse direction of causality. On the contrary, banks with a larger number of ATM per branch display greater inefficiency. One interpretation is that, in the majority of cases, the downsize in staff required by large-scale automation of some basic activities takes time or it is not fully adjusted, leading to some excess capacity. The existence of frictions in the adjustment of labor, which is supported by the estimated negative correlation between inefficiency and the fraction of staff made redundant, is a consequence of major restructuring plans (REDUNDANCE).

Inefficiency is also positively correlated with the ratio of capital to total assets (CAP_ASS). Again the causality link is ambiguous. Large capital ratios may follow from excess capacity or may be due to agency problems: more efficient banks can signal their quality through a high leverage as in Berger and Bonaccorsi di Patti (2002).

19 *The estimated coefficients of the traslog cost and profit functions, not reported in the paper, are available on request.*

Mergers and acquisitions are negatively correlated with inefficiency. Since we use pro-forma data, this result simply means that banks which buy other banks have above-average efficiency scores, a well-established result in the M&A literature (Focarelli et al., 2002). The dummy variables identifying the partnership in a large (GROUP1) or small (GROUP2) banking group have estimated coefficients with positive signs in the cost function and negative signs in the profit regression. Most of these banks have been acquired by a bank holding company during the sample period. Target banks were usually poorly performing, and restructuring within the banking group has initially focused on the revenue side and only later gradually extended to the cost side (Focarelli et al., 2002).

The AGE variable is negatively correlated with inefficiency: this may be due to the time needed by the *de novo* banks to reach their optimal level of capacity utilization. Finally, the geographical dummies show that banks with headquarters in the richest regions of the country (NE and NW) are closer to the efficient frontier than banks located elsewhere.

5.2. Frontier shift regressions

Tables 5 and 6 report the results of the panel data estimation of equation 9 relatively to cost and profit shifts. A negative sign of a coefficient in Table 5 implies that an increase in the value of the corresponding variable is associated with a reduction in industry best practice costs. In a similar way a positive sign of a coefficient in Table 6 implies that an increase in that variable is associated with an increase in best practice profits. The first three columns report the results of different specifications of the IT variables for the entire sample. In the last three columns we check for differences owing to bank size.

In columns A, D, E and F, IT capital deepening is measured simply by total IT capital per employee. The results are consistent with the hypothesis that computers and related technologies lead to substantial increases in total factor productivity, here indirectly measured by frontier shifts. The coefficient of the IT capital variable is different from zero at high significance levels, and the sign is negative for the cost frontier regression and positive for the profit frontier one. The result holds true across all bank size classes (columns D, E, and F) and does not depend on the type of frontier we consider. In particular, as expected the IT capital effect is greater for smaller banks, which are less likely to be closer to the optimal level of IT capital stock and so to suffer excess capacity. For the median bank, an increase of one standard deviation in the level of IT capital per employee will generate a favorable frontier cost shift of the order of the 8 percent both on costs and profits.

In columns B we have split IT capital into two broad categories: hardware and software. The coefficient of the HARD_CAP variable is negative in the cost function and positive in the profit function and statistically significant, indicating a favorable effect of computers on productivity. The coefficients of SOFT_CAP are estimated less precisely and have an opposite pattern of signs. A possible explanation is that software capital stock is poorly measured because hedonic prices do not account properly for quality improvements (Jorgenson and Stiroh, 2000).

In columns C we introduce a variable aiming to capture the impact of IT on labor organization: the fraction of salary expenses devoted to IT staff. This variable is available only for a sub-sample of banks (about half of the total number). The sign of the coefficient of this variable is positive in the cost frontier and negative in the profit frontier, meaning that the within-firm production of IT services is detrimental to TFP growth. Accordingly, the variable OUTSOURCE, measuring the amount of IT-related expenses for outsourced activities, is positively correlated with favorable frontier shifts in all the columns in Tables 5 and 6. These results are consistent with the findings about the importance of within firm organization for achieving productivity growth through IT capital accumulation.

Once we rule out the effect of excess capacity, considering only fully efficient banks (i.e. the projection of the banks on the efficient frontier) we find that the variables SERVICE, CAP_ASS and ATM_BR are positively correlated with productivity. The number of staff per branch (STAFF_BR), the share of managers' salaries in total salaries (MANAGERS) and that of small business loans in total loans (SMALL) are negatively correlated with TFP growth. All these three variables capture the specialization of retail banks for which the route to innovation has been less straightforward.

6. Conclusions

In this paper we have investigated the effect of information technologies on the Italian banking industry over the period 1989–2000. We have estimated stochastic cost and profit functions allowing for individual banks' displacements from the efficient frontier and non-neutral technological change. Data on IT capital stock for individual banks enabled us to distinguish between movements along the efficient frontier and shifts of the frontier owing to the adoption of new technologies. Several studies have emphasized the role of technological innovation as a primary source of structural changes within the financial industry. However, empirical evidence on the effects of technical progress in increasing TFP and abating unit costs is still scarce.

We found that both cost and profit frontier shifts are strongly correlated with IT capital accumulation. Banks adopting IT capital intensive techniques are also closer in average to the best practice of the banking industry, implying ceteris paribus a higher level of efficiency. We interpret this last result as evidence of a catching-up effect consistent with the usual pattern of diffusion of the new technologies.

Table 1 – Banks' IT expenses

Year	IT expenses as a percentage of		Composition of IT expenses					
	Gross income	Operating costs	Hardware	Software	Staff	Outsourcing	Others	Total
1989	5.6	9.0	26.1	9.9	32.2	13.9	17.9	100.0
1990	5.8	9.4	27.2	10.8	30.3	14.4	17.3	100.0
1991	6.1	9.5	27.1	10.7	29.8	14.8	17.7	100.0
1992	6.0	9.2	24.4	12.4	28.9	14.9	19.5	100.0
1993	5.5	9.1	23.4	13.2	26.5	16.4	20.5	100.0
1994	6.6	9.8	22.0	13.4	26.8	17.7	20.0	100.0
1995	6.7	9.9	19.2	13.6	25.1	22.3	19.8	100.0
1996	6.4	9.6	16.6	15.2	21.7	25.3	21.2	100.0
1997	6.4	9.4	16.9	16.7	21.2	25.6	19.7	100.0
1998	7.4	12.2	21.3	20.7	16.5	17.8	23.7	100.0
1999	7.6	12.8	19.7	19.7	15.0	26.6	19.0	100.0
2000	7.6	13.5	14.9	18.3	12.4	38.7	15.8	100.0
2001	8.0	14.5	15.6	20.8	10.4	34.1	19.2	100.0

Sources: Bank Supervisory Reports.

Table 2 – Banks' IT investment and capital stock*

(1989=100)

Year	Nominal IT investments		Real IT investments		IT Capital	
	Hardware	Software	Hardware	Software	Total	Per employee
1989	100	100	100	100	100	100
1990	101	126	135	154	129	125
1991	114	149	154	166	155	148
1992	112	179	158	188	176	167
1993	98	191	133	164	179	168
1994	101	222	164	205	193	182
1995	96	295	185	270	215	202
1996	97	268	258	264	254	243
1997	99	337	302	302	300	292
1998	123	430	523	405	421	414
1999	91	370	430	300	458	455
2000	105	393	535	289	522	514
2001	107	453	656	312	612	599

* *Adjusted for inflation.*

Sources: Bank Supervisory Reports, ISTAT and BEA.

Table 3 – Descriptive statistics: Variables used in the estimation

All financial variables are measured in ITL million and are adjusted for inflation. Statistics refer to the panel of 618 banks for the period 1989–2000 with the exception of internet banking accounts, available from 1995.

Symbol	Description	Obs.	Mean	Min	Max	St.dev
COST	Log of total cost (divided by the price of labor)	5231	5.077	1.676	10.696	2.006
PROF	Log of gross profit (divided by the price of labor)	5231	8.2889	2.480	10.279	0.316
Y1	Log of short-term loans	5231	4.341	−5.276	10.902	1.932
Y2	Log of long-term loans	5231	4.000	–1.440	10.952	1.744
Y3	Log of services weighted commission income	5231	5.167	–2.624	10.786	1.611
Y4	Log of deposits	5231	−0.181	−5.291	7.151	2.107
P1	Log of IT capital price (divided by the labor price)	5231	2.069	0.253	3.744	0.620
P2	Log of branches price (divided by the labor price)	5231	−0.137	−1.266	1.833	0.530
CAP_ASS	Capital divided by total assets	5231	0.093	0.017	0.396	0.034
AGE	$1-1/\sqrt{\text{bank age}}$	5231	0.842	0.000	0.875	0.082
SERVICE	Net services value revenue divided by gross income	5231	0.060	−0.983	0.556	0.060
SMALL	Value of small loans (under 5 billion lire) over total loans	5231	0.675	0.023	1.000	0.197
GROUP1	Part of a big group (top 25% of the groups' total asset distribution)	5231	0.059	0.000	1.000	0.236
GROUP2	Part of a medium or small group (last 75% of the groups' total asset distribution)	5231	0.088	0.000	1.000	0.283
MERACQ	Relative increase in total assets of the bank due to M&A	5231	0.011	0.000	1.536	0.066
STAFF_BR	Number of employees divided by number of branches	5231	11.200	2.000	242.00	10.057
MANAGERS	Wages paid to managers divided by total wages	5213	36.189	0.520	331.46	26.645
REDUNDANCE	Number of redundant employees divided by total staff	5231	0.001	0.000	0.519	0.012
IT_CAP	Stock of real IT capital per employee (billion lire)	5231	0.056	0.001	0.482	0.044
HARD_CAP	Stock of real hardware capital per employee (billion lire)	5231	0.051	0.000	0.467	0.042
SOFT_CAP	Stock of real software capital per employee (billion lire)	5231	0.004	0.000	0.060	0.004
OUTSOURCE	Share of IT expenses due to outsourcing of services	5100	0.3689	0.000	1.000	0.2996
IT_STAFF	Wages paid to IT personnel divided by total wages	3031	0.055	0.000	1.000	0.070
REMOTE	Number of phone and electronic banking accounts of households divided by total number of current accounts	5231	0.018	0.000	1.000	0.091
ATM_BR	Number of ATM divided by number of branches	5231	0.526	0.000	20.00	0.750
NW	Headquarters in the North-West	5231	0.201	0.000	1.000	0.401
NE	Headquarters in the North-East	5231	0.408	0.000	1.000	0.492
CE	Headquarters in the Centre	5231	0.211	0.000	1.000	0.408
SOUTH	Headquarters in the South or islands	5231	0.180	0.000	1.000	0.384

Sources: Bank Supervisory Reports.

Table 4 – Cost and profit inefficiency correlates

The dependent variables are respectively the cost and the profit inefficiencies jointly estimated from the translog specification of Table 4. A complete description of the independent variables is given in Table 3.

Variables	Cost inefficiency	Profit inefficiency
CONSTANT	0.2126***	0.8587***
	0.0568	0.0651
TREND	0.0117***	0.0352***
	0.0038	0.0068
IT_CAP	−0.9100***	−1.7889***
	0.1939	0.3683
SERVICE	−1.5195***	−1.2455***
	0.2018	0.0482
REMOTE	−0.0294	−0.0592*
	0.0551	0.0347
ATM_BR	0.0192***	0.0417***
	0.0078	0.0077
REDUNDANCE	−0.0148	−0.8690***
	0.2841	0.2184
CAP_ASS	0.0139	3.9501***
	0.1510	0.4436
MERACQ	−0.0293	−0.2367***
	0.0721	0.0422
GROUP1	0.0661***	−0.7449***
	0.0210	0.0961
GROUP2	0.1665***	−0.6888***
	0.0288	0.0877
AGE	−0.0095	−0.6347***
	0.0569	0.0266
NW	−0.3572***	−0.3838***
	0.0609	0.0212
NE	−0.4983***	−0.0438*
	0.0943	0.0228
CE	−0.2293***	−0.0156
	0.0384	0.0266
σ^2	0.0521***	0.0669***
	0.0039	0.0086
γ	0.4419***	0.8320***
	0.0511	0.0282
Observations = 5231		
Cross-sections = 618		
Time periods = 12		

*Note: Statistically different from zero, respectively, at: *** 99%, **95% and *90% significance level.*

Table 5 – Cost function shifts

The dependent variable is the time derivative of the cost function estimated in Table 4. A complete description of the independent variables is given in Table 3. The full sample regression refers to the sample employed for estimations in Tables 4 and 5, reduced by the missing values in the OUTSOURCING and the IT_STAFF variables. The Large banks *sample refers to the top 10 per cent of banks in total asset distribution, the* Medium banks *sample include banks from the 33 to the 10 percent and the* Small banks *the last 66 percent of banks. For the estimated coefficients, a positive sign implies an increasing effect on costs and a correspondent decreasing effect on productivity.*

Variables	(A) Full sample	(B) Full sample	(C) Full sample	(D) Large banks	(E) Medium banks	(F) Small banks
CONSTANT	0.1296***	0.1297***	0.0826***	0.0399	0.0746***	0.2700***
	0.0119	0.0119	0.0129	0.0281	0.0163	0.0211
IT_CAP	−0.1597***		−0.1412***	−0.1022**	−0.0747***	−0.1953***
	0.0131		0.0198	0.0501	0.0216	0.0163
HARDW_CAP		−0.1753***				
		0.0142				
SOFTW_CAP		0.2759*				
		0.1486				
IT_STAFF			0.0851***			
			0.0173			
OUTSOURCING	−0.0109***	−0.0098***	−0.0220***	−0.0345***	−0.0252***	−0.0018
	0.0024	0.0024	0.0041	0.0105	0.0042	0.0029
ATM_BR	−0.0362***	−0.0364***	−0.0368***	−0.0355***	−0.0385***	−0.0326***
	0.0009	0.0009	0.0012	0.0030	0.0016	0.0011
REMOTE	−0.0007	0.0002	−0.0014	0.0092	−0.0172***	0.0265***
	0.0058	0.0058	0.0080	0.0332	0.0062	0.0099
CAP_ASS	−0.7754***	−0.7739***	−0.7348***	−0.2278**	−0.6437***	−0.7691***
	0.0250	0.0250	0.0300	0.0973	0.0440	0.0318
SMALL	0.0761***	0.0744***	0.0949***	0.0789***	0.0672***	0.0706***
	0.0063	0.0063	0.0089	0.0259	0.0124	0.0074
SERVICE	−0.2267***	−0.2229***	−0.1980***	−0.1207***	−0.2288***	−0.2819***
	0.0117	0.0118	0.0149	0.0303	0.0176	0.0185
STAFF_DEN	0.0023***	0.0023***	0.0024***	0.0022***	0.0028***	0.0025***
	0.0001	0.0001	0.0002	0.0002	0.0003	0.0002
REDUNDANCE	0.0838*	0.0777*	0.0592	−0.0934	−0.1494	0.1383**
	0.0446	0.0446	0.0619	0.1041	0.1109	0.0552
MANAGERS	0.0002***	0.0002***	0.0001***	0.0000	0.0000	0.0004***
	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000
MERACQ	0.0133**	0.0134**	0.0159*	−0.0289	0.0096	0.0216***
	0.0066	0.0066	0.0084	0.0429	0.0129	0.0074
GROUP2	−0.0231***	−0.0234***	−0.0251***	−0.0324***	−0.0241***	−0.0263
	0.0026	0.0026	0.0027	0.0114	0.0025	0.0244
GROUP1	−0.0162***	−0.0167***	−0.0174***	−0.0151***	−0.0205***	–
	0.0026	0.0026	0.0027	0.0053	0.0029	
AGE	−0.0957***	−0.0966***	−0.0607***	−0.0363	−0.0140	−0.2764***
	0.0125	0.0125	0.0124	0.0272	0.0156	0.0239
Observations	5082	5082	2921	435	1318	3329
Banks	612	612	396	49	146	417
R-squared (within)	0.756	0.757	0.787	0.698	0.836	0.760
R-squared (between)	0.001	0.001	0.014	0.049	0.329	0.016
R-squared (overall)	0.440	0.439	0.451	0.405	0.688	0.375

*Note: Statistically different from zero, respectively, at: *** 99%, **95% and *90% significance level.*

Table 6 – Profit function shifts

The dependent variable is the time derivative of the profit function estimated in Table 4. A complete description of the independent variables is given in Table 3. The full sample regression refers to the sample employed for estimations in Tables 4 and 5, reduced by the missing values in the OUTSOURCING and the IT_STAFF variables. The Large banks *sample refers to the top 10 percent of banks in total asset distribution, the* Medium banks *sample include banks from the 33 to the 10 percent and the* Small banks *the last 66 percent of banks. For the estimated coefficients, a positive sign implies an increasing effect on profits and a correspondent increasing effect on productivity.*

Variables	(A) Full sample	(B) Full sample	(C) Full sample	(D) Large banks	(E) Medium banks	(F) Small banks
CONSTANT	−0.1376***	−0.1379***	−0.0940***	−0.0507**	−0.0825***	−0.2575***
	0.0086	0.0086	0.0092	0.0203	0.0119	0.0150
IT_CAP	0.1351***		0.1362***	0.1271***	0.0699***	0.1545***
	0.0094		0.0141	0.0363	0.0158	0.0116
HARDW_CAP		0.1419***				
		0.0102				
SOFTW_CAP		−0.0706				
		0.1071				
IT_STAFF			−0.0705***			
			0.0124			
OUTSOURCING	0.0124***	0.0118***	0.0205***	0.0258***	0.0222***	0.0060***
	0.0017	0.0017	0.0029	0.0076	0.0031	0.0021
ATM_BR	0.0258***	0.0259***	0.0262***	0.0278***	0.0287***	0.0221***
	0.0006	0.0006	0.0008	0.0022	0.0011	0.0008
REMOTE	0.0001	−0.0003	0.0067	0.0255	0.0134***	−0.0238***
	0.0042	0.0042	0.0057	0.0240	0.0046	0.0070
CAP_ASS	0.5722***	0.5721***	0.5461***	0.2003***	0.4855***	0.5625***
	0.0180	0.0180	0.0214	0.0705	0.0322	0.0225
SMALL	−0.0518***	−0.0509***	−0.0634***	−0.0480**	−0.0440***	−0.0473***
	0.0045	0.0045	0.0064	0.0188	0.0091	0.0052
SERVICE	0.2333***	0.2317***	0.2019***	0.1234***	0.2293***	0.2899***
	0.0084	0.0085	0.0107	0.0220	0.0129	0.0131
STAFF_DEN	−0.0020***	−0.0020***	−0.0021***	−0.0018***	−0.0022***	−0.0023***
	0.0001	0.0001	0.0001	0.0002	0.0002	0.0002
REDUNDANCE	−0.0579*	−0.0549*	−0.0222	0.0888	0.1299*	−0.1162***
	0.0321	0.0321	0.0442	0.0754	0.0812	0.0391
MANAGERS	−0.0003***	−0.0003***	−0.0002***	−0.0002***	−0.0001***	−0.0005***
	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000
MERACQ	−0.0026	−0.0026	−0.0018	0.0264	−0.0025	−0.0081
	0.0048	0.0048	0.0060	0.0311	0.0094	0.0052
GROUP2	0.0190***	0.0192***	0.0195***	0.0259***	0.0194***	0.0409***
	0.0019	0.0019	0.0019	0.0083	0.0018	0.0173
GROUP1	0.0152***	0.0155***	0.0160***	0.0157***	0.0178***	–
	0.0019	0.0019	0.0019	0.0038	0.0021	
AGE	0.0694***	0.0700***	0.0393***	0.0197	0.0074	0.2134***
	0.0090	0.0090	0.0089	0.0197	0.0114	0.0169
Observations	5082	5082	2921	435	1318	3329
Banks	612	612	396	49	146	417
R-squared (within)	0.8016	0.8015	0.833	0.7852	0.871	0.8002
R-squared (between)	0.0462	0.0441	0.152	0.0677	0.307	0.0331
R-squared (overall)	0.4935	0.490	0.551	0.484	0.706	0.4334

*Note: Statistically different from zero, respectively, at: *** 99%, **95% and *90% significance level.*

Figure 1 – Inefficiency Estimates for the cost function model

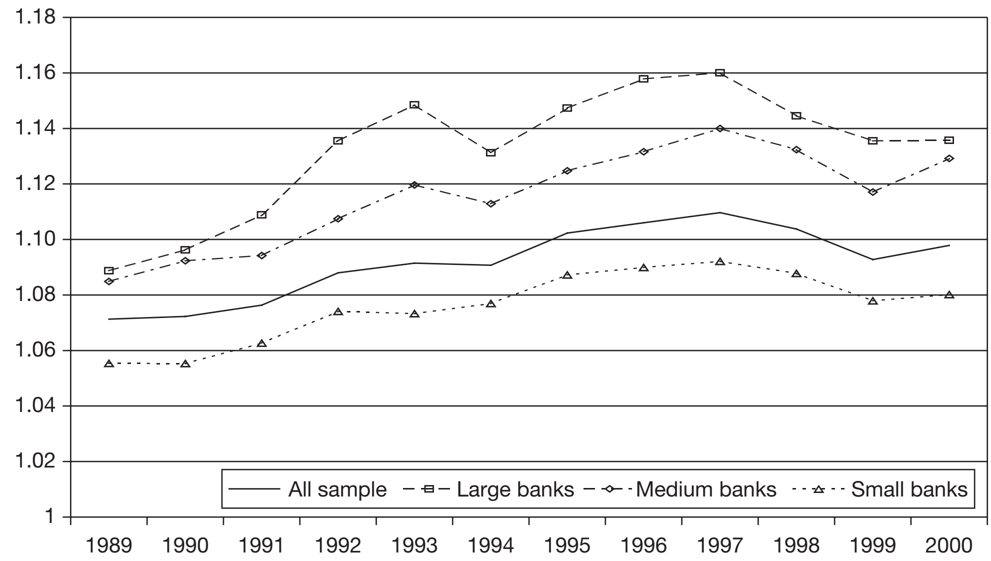
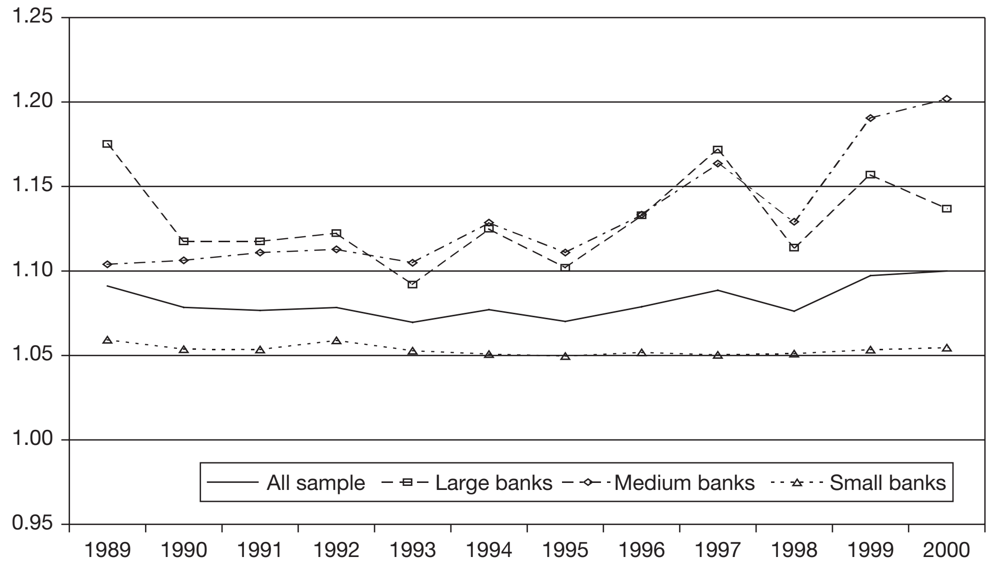


Figure 2 – Inefficiency Estimates for the profit function model



References

Akhavein, J.D., A.N. Berger and D.B. Humphrey, (1997), "The Effects of Megamergers on Efficiency and Prices: Evidence from a Bank Profit Function", *Review of Industrial Organization*, 12: 95–139.

Alan, I.M.S. (2001), "A Nonparametric Approach for Assessing Productivity Dynamics of Large US Banks", *Journal of Money, Credit and Banking*, 33: 121–39.

Altunbas, Y., L. Evans and P. Molineux (2001), "Bank Ownership and Efficiency", *Journal of Money, Credit and Banking*, 33: 926–54.

Autor, D.H., F. Levy and R.J. Murnane (2000), "Upstairs, Downstairs: Computer-Skill Complementarity and Computer-Labor Substitution on Two Floors of a Large Bank", *NBER Working Paper* 7890.

Battese, G.E. and T.J. Coelli (1988), "Predictions of Firm Level Technical Efficiencies with a Generalised Frontier Production Function and Panel Data", *Journal of Econometrics*, 38: 387–400.

Battese, G.E. and T.J. Coelli (1993), "A Stochastic Frontier Production Function Incorporating a Model for Technical Inefficiency Effects", *Working Papers in Econometrics and Applied Statistics*, 69, Department of Econometrics, University of New England, Armindale.

Battese, G.E. and T.J. Coelli (1995), "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function with Panel Data", *Empirical Economics*, 20: 325–32.

Bauer, P.W. (1990), "Recent Developments in the Econometric Estimation of Frontiers", *Journal of Econometrics*, 46: 39–56.

Bauer et al. (1993), "Efficiency and Productivity Growth in US Banking" in Fried, H.O. et al. (ed.) "*The Measurement of Productive Efficiency: Techniques and Applications*", Oxford University Press: Oxford.

Berg, S.A., F. Finn and E.S. Jansen (1992), "Malmquist Indices of Productivity Growth During the Deregulation of Norwegian Banking 1980–89", *Scandinavian Journal of Economics*, 94 Supplement: 211–28.

Berger, A.N. (1993), "Distribution Free Estimates of Efficiency in the US Banking Industry and Tests of the Standard Distributional Assumption", *Journal of Productivity Analysis*, 4: 261–92.

Berger, A.N. and E. Bonaccorsi di Patti (2002), "Capital Structure and Firm Performance: A New Approach to Testing Agency Theory and an Application to the Banking Industry", Federal Reserve Board, *Finance and Economics Discussion Series*, 54.

Berger, A.N. and D.B. Humphrey (1991), "The Dominance of Inefficiency over Scale and Product Mix Economies in Banking", *Journal of Monetary Economics*, 28: 117–48.

Berger, A.N., W.C. Hunter and S.G. Timme (1993), "The Efficiency of Financial Institution: a Review and Preview of Research Past, Present and Future", *Journal of Banking and Finance*, 17: 221–49.

Berger, A.N. and L. Mester (1997), "Beyond the Black Box: What Explains Differences in the Efficiency of Financial Institutions?", *Journal of Banking and Finance*, 21: 895–947.

Berger, A.N. and L. Mester (2001), "Explaining the Dramatic Changes in the Performances of US Banks: Technological Change, Deregulation and Dynamic Changes in Competition", *Working Paper 01-6*, Federal Reserve Bank of Philadelphia.

Berger, A.N. (2003), "The Economic Effect of Technological Progress: Evidence from the Banking Industry", *Journal of Money, Credit and Banking*, 35: 141–76.

Berndt, E.R. and C.J. Morrison (1991), "Assessing the Productivity of Information Technology Equipment in US Manufacturing Industries", *NBER working paper* 3582.

Boskin, M.J. and D.W. Jorgenson (1997), "Implications of Overstating Inflation for Indexing Government Programs and Understanding Economic Progress", *American Economic Review*, 87: 89–93.

Bresnaham, T., E. Brynjolfsson and L. Hitt (2002), "Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm Level Evidence", *Quarterly Journal of Economics*, 17: 339–76.

Brynjolfsson, E. and L.M. Hitt (2000), "Computing Productivity: Firm-Level Evidence", *mimeo, MIT*.

Bugamelli, M. and P. Pagano (2001), "Barriers to Investments in ICT" Banca d'Italia, *Temi di Discussione*, 420.

Clark, J.A. and T.F. Siemes (2002), "X-efficiency in Banking: Looking Beyond the Balance Sheet", *Journal of Money, Credit and Banking*, 34: 987–1013.

Daveri, F. (2001), "Information Technology and Growth in Europe", mimeo*, University of Parma*.

Ferrier, G.D. and C.A.K. Lovell (1990), "Measuring Cost Efficiency in Banking: Econometric and Linear Programming Evidence", *Journal of Econometrics*, 46: 229–45.

Fixler, D.F. (1993), "Measuring Financial Service Output and Prices in Commercial Banking", *Applied Economics*, 25: 983–93.

Focarelli, D., F. Panetta and C. Salleo (2002), "Why Do Banks Merge?", *Journal of Money, Credit and Banking*, 34: 1047–66.

Geroski, P.A. (2000), "Models of Technology Diffusion", *Research Policy*, 29: 603–25.

Grifell-Tatje, E. and C.A.K. Lovell (1996), "Deregulation and Productivity Decline: The Case of Spanish Savings Banks", *European Economic Review*, 40: 1281–303.

Gordon, R.J. (1999), Has the "New Economy Rendered the Productivity Slowdown Obsolete?", *Working Paper*, Northwestern University, June.

Hancock, D. (1991), "*A Theory of Production for the Financial Firm*", Boston, Kluwer Academy Publisher.

Hunter, W.C. and S.G. Timme (1986), "Technical Change, Organizational Form, and the Structure of Bank Production", *Journal of Money, Credit and Banking*, 18: 151–66.

Humphrey, D.B. (1992), "Flow Versus Stock Indicators of Banking Output: Effects on Productivity and Scale Economies Measurement", *Journal of Financial Service Research*, 6: 117:30.

Humphrey, D.B. and L.B. Pulley (1997), "Banks' Responses to Deregulation: Profits, Technology and Efficiency", *Journal of Money, Credit and Banking*, 29: 73–93.

Jondrow, J., C.A.K. Lovell, I.S. Materov and P. Schmidt (1982), "On the Estimation of the Technical Inefficiency in the Stochastic Frontier Production Model", *Journal of Econometrics*, 19: 233–38.

Jorgenson, D.W. (2001), "Information Technology and the US Economy", *American Economic Review*, 91: 1–32.

Jorgenson, D.W. and K.J. Stiroh (2000), "Raising the Speed Limit: US Economic Growth in the Information Age", *Brooking Papers on Economic Activity*, forthcoming.

Kumbhakar, S.C. (2000), "Estimation and Decomposition of Productivity Change When Production is not Efficient: a Panel Data Approach", *Econometric Reviews*, 19: 425–460.

Khumbhakar, S.C., S. Ghosh and J.T. McGukin (1991), "A Generalized Production Function Approach for Estimating Determinants of Inefficiency in US Dairy Firm", *Journal of Business and Economics Statistics*, 9: 279–86.

Kumbhakar, S.C., A. Lozano-Vivas, C.A.K. Lovell and I. Hasan (2001), "The Effects of Deregulation on the Performance of Financial Institutions: The Case of Spanish Savings Banks", *Journal of Money, Credit and Banking*, 33: 101–20.

McKinsey Quarterly (2002), number 1 "*Banking: The IT Paradox*".

Oliner, S.D. and D.E. Sichel (2000), "The Resurgence of Growth in the Late 1990s: Is Information Technology the Story?", *Journal of Economic Perspectives*, 14: 3–22.

Parsons, D., C.C. Gotlieb and M. Denny (1993), "Productivity and Computers in Canadian Banking", *Journal of Productivity Analysis*, 4: 95–114.

Pozzolo, A.F. (2001), "An Empirical Investigation of Bank Secured Lending", Banca d'Italia, Lavoro Preparatorio per la Relazione Annuale del 2000.

Roach, S.S. (1987), "America's Technology Dilemma: A Profile of the Information Technology", *Morgan Stanley Special Economic Study*.

Salter, W.E.G. (1966), *Productivity and Technical Change,* 2nd edition*,* Cambridge, Cambridge University Press.

Schmidt, P. and C. Lovell (1979), "Estimating Technical and Allocative Inefficiency Relative to Stochastic Production and Cost Function", *Journal of Econometrics*, 9: 343–66.

Schmidt, P. and R.C. Sickles (1984), "Production Frontiers and Panel Data", *Journal of Business and Economics Statistics*, 2: 367–74.

Seskin, E.P. (1999), "Improved Estimates of the National Income and Product Accounts for 1959–98: Results of the Comprehensive Revision", *Survey of Current Business*, BEA.

Solow, R.M. (1987), "We'd Better Watch Out" *New York Review of Books*, July 12.

Summers, L.H. (2000), "International Financial Crises: Causes, Prevention and Cures", *American Economic Review, Papers and Proceedings*, 90: 1–16.

Vennet, R.V. (2002), "Cost and Profit Efficiency of Financial Conglomerates and Universal Banks in Europe", *Journal of Money, Credit and Banking*, 34: 254–82.

Abstract

This paper analyzes the effect of information technologies (IT) in the financial sector using micro-data on a panel of over 600 Italian banks over the period 1989–2000. We estimate stochastic cost and profit functions allowing for individual banks' displacements from the efficient frontier and for non-neutral technological change. Data on IT capital stock for individual banks enable us to distinguish between movements along the efficient frontier and shifts of the frontier owing to the adoption of new technologies. We find that both cost and profit frontier shifts are strongly correlated with IT capital accumulation. Banks adopting IT capital intensive techniques are also more efficient. We interpret this last result as evidence of a catching-up effect consistent with the usual pattern of diffusion of the new technologies.

JEL classification: D24, G21, O33.

Keywords: banking, productivity, efficiency, information technology.

Luca Casolaro and Giorgio Gobbi
Bank of Italy
Economic Research Department

DISCUSSANTS' COMMENTS

Discussant's summary on Workshop C: Productivity Statistics

Hlabi Morudu (South African Reserve Bank)

For Monetary Policy purposes, for inflation to remain low, an increase in aggregate demand has to be adequately matched by an increase in output growth.

From the papers that discuss statistics in the measurement of productivity, it is evident that there remain several measurement problems both in defining the numerator and the denominator in the usual calculation of productivity. For instance,

(a) how aggregated should the numerator (usually output) be? Is it defined identically throughout countries?

(b) should the denominator include both labour and capital? How should labour be treated? How should capital be defined?

The papers make useful suggestions on correcting some of the major problems in measuring productivity. Note for instance, the constructive proposals on improvements that may be made for cross-sectional country productivity comparisons in Bruno Tissot, et al., Ivan Odonnat, et al.

Also note some of the useful suggestions for improvements in the measurement of productivity that may be made within a country's data as in Sylaja Srinivasan, et al. Further also note a potentially useful approach of focusing on productivity at a more disaggregated level – i.e. at the "micro" level – as in Luca Casolaro, et al.

For policy purposes, however, there remains one particular significant question. The papers all acknowledge that there is a strong correlation between productivity growth and business cycles. Are changes in productivity a result of business cycles as employment increases/decreases, or capital deepening?

To conclude, improvements in the measurement of productivity remain highly essential in conducting monetary and other policies.

References

List of papers presented on "Productivity Statistics" at the Irving Fisher Conference on Central Bank Issues Regarding National and Financial Accounts, 9–10 September 2004, BIS, Switzerland

1. Les Skoczylas, Bruno Tissot, "Revisiting recent productivity developments across OECD countries", BIS.
2. Charlotta Groth, Maria Gutierrez-Domenech and Sylaja Srinivasan, "Measuring total factor productivity for the United Kingdom", Bank of England.
3. Ivan Odonnat, "Synthetic review of the methods and results considered and obtained by the Banque de France on productivity", Banque de France.
4. Luca Casolaro and Giorgio Gobbi, "Information technology and productivity changes in the Italian banking industry", Banca d'Italia.

Hlabi Morudu (South African Reserve Bank)

WORKSHOP D

Output gaps and capacity

Chair: Radha-Binod Barman (Reserve Bank of India)

Papers:

Estimating output gap for the Turkish economy: A semi-structural non-linear time series approach
Çağrı Sarıkaya, Fethi Öğünç, Dilara Ece, Ayşe Tatar Curl, Arzu Çetinkaya (Central Bank of the Republic of Turkey)

Using additional information in estimating the output gap in Peru: a multivariate unobserved component approach
Gonzalo Llosa and Shirley Miller (Central Reserve Bank of Peru)

Capital stock in Indonesia: measurement and validity test
Noor Yudanto, Gunawan Wicaksono, Eko Ariantoro and A. Reina Sari (Bank Indonesia)

Comments on the workshop about output gap and capacity
Ahmed Al Kholifey (Saudi Arabian Monetary Agency)

Estimating output gap for the Turkish economy: A semi-structural non-linear time series approach[1]

Çağrı Sarıkaya, Fethi Öğünç, Dilara Ece, Ayşe Tatar Curl, Arzu Çetinkaya (Central Bank of the Republic of Turkey)

1. Introduction

This study seeks to develop an output gap measure for the Turkish economy. The emergence of price stability as the overriding goal of the monetary policy in the last decade has led central banks to utilize all available information in the economy to foresee the future course of price dynamics. In this respect, the output gap, which can be defined as the difference between the actual output and its "potential" level, is closely monitored by the central banks for the implementation of the monetary policy. Output gap is a key indicator of inflationary pressures among various measures of resource utilization. As far as inflation dynamics are concerned, potential output is often defined as the level of output that can be sustained without putting pressure on production costs and thus on inflation. Therefore, a level of actual output above potential may signal inflationary pressures in the near future.

Output gap has gained popularity in the "monetary policy rules" literature, which has been subject to considerable interest in the recent years. The progress following the pioneer work of Taylor (1993) and especially the success of inflation-targeting regimes, which require effective use of short-term interest rates as the policy instrument, has attached a significant role to the output gap as a response variable in feedback rules. A positive output gap is often perceived as a signal of "excess demand", which may require an increase in the interest rate to prevent the economy from overheating.

The primary issue to be tackled, while conducting an output-gap estimation exercise, would pertain to the underlying technique in question. After all, the information content of the conventional output gap measures (especially for emerging market economies) may be limited.[2] First, these series cannot be directly observed. It is not surprising to see a wide range of detrending methodologies utilizing univariate models as well as multivariate filtering models to come up with alternative estimates of the output gap. The studies about developing proper alternative techniques have further intensified after the shortcomings of the Hodrick-Prescott (HP) filter – the most commonly used methodology – have been realized.[3] Specifically, the values suggested by Hodrick and Prescott are criticized as they are appropriate for the US data and may lead to misspecification of the underlying economic structure of other economies. Also, since the HP filter is a two-sided optimization procedure, which uses both lead and lagged information, the accuracy of the filter diminishes at the end of sample due to missing lead information. In fact, this is unfortunate given that the last observations are often the most vital ones for monetary policy analysis in the sense that these are the ones used for forecasting inflation. In addition, HP filter is subject to criticisms as being purely statistical and having no economic content since it does not exploit any additional information other

1 *This study is an ongoing joint project conducted by the Research and Statistics Departments of the Central Bank of the Republic of Turkey under the supervision of Hakan Kara and Ümit Özlale. We would like to thank Levent Özbek for his valuable assistance and suggestions.*

2 *See Billmeier (2004) and Orphanides (1999) for an account of the information content of the output gap estimates widely used in the literature.*

3 *The weighting parameter in the objective function, which in fact represents the relative magnitude of aggregate demand and supply shocks, is subjectively determined by the user. As stated by Butler (1996), rather than imposing the weighting parameter, it is also possible to estimate this parameter in a multivariate setting. Boone et al. (2001) also shows that the optimization procedure utilized by the HP filter can be specified as a univariate unobserved components model where the smoothing parameter represents the relative variances of a shock to trend component and a shock to temporary component, which can be estimated by prediction error decomposition via the Kalman filter.*

than the series to be detrended. Moreover, the cyclical component always sums up to zero in HP filter; ruling out the case that actual output is below the potential for a longer time than it is above it. Finally, HP filter cannot capture the high volatility of the macroeconomic environment and rapidly changing dynamics, which are two well-known factors peculiar to emerging markets. These criticisms make it necessary to handle the issue of measuring output gap within the context of more structural models, which also take country-specific factors into account.

In this context, the production function approach can be seen as more appropriate to estimate the potential output. However, these kinds of applications need accurate and healthy data for at least capital, labor and productivity, which is fairly not so in our case. In that sense, statistical methods remain as a natural candidate to estimate the output gap for the emerging markets, including the Turkish economy. However, one can still enrich the structural content of the output gap estimates by using macroeconomic relations that are supposed to be relevant for inflation and output gap dynamics. At this point, multivariate filtering techniques utilizing information from theoretical macroeconomic relationships, such as the Phillips curve, enter into the picture.[4] As stated by Kuttner (1994), the problem is to estimate the parameters to obtain the unobserved variable, the output gap, which is most consistent with the observed inflation. In this respect, one appropriate estimation technique is the Kalman filter, which is a recursive algorithm for optimally forecasting the unobserved component, given the observed variables and the imposed economic structure.

However, these techniques often impose strong restrictions in defining the relationships between key macroeconomic variables. For example, it is assumed that the relationships that govern the dynamics of the economy stay intact over the sample period. Although these models address – and hence improve upon – the criticisms made for both HP filter and other statistical procedures, there is still room for improvement, especially for economies experiencing frequent structural changes. There is no doubt that this argument applies strongly for emerging markets: Adopting different monetary policy regimes and experiencing frequent fiscal and financial restructuring periods affect the behavior of economic agents over time. For example, the effects of the transition to an inflation-targeting framework after the collapse of the exchange rate based stabilization program and the intensified attempts to ensure fiscal discipline cannot be adequately captured in a model where the system parameters are assumed to be constant over time. Therefore, relaxing the typical restriction that system parameters stay constant over time may provide insightful results if the output gap in question is pertaining to an emerging market economy.

On the other hand, it is not a trivial task to derive output gap in an unobserved components model, where parameters are also to be estimated in a time-varying fashion. For one thing, when the state variables (including output gap and/or potential output) and the system parameters are to be estimated simultaneously, the model takes a non-linear characteristic and the standard Kalman filter (SKF henceforth) needs to be modified. In this case, extended Kalman filter (EKF henceforth) emerges as an appropriate estimation procedure to be employed.

Estimation of the output gap by the extended Kalman filter methodology allows us to observe the changing dynamics in the economy in question. Accordingly, this study presents a multivariate unobserved components model to estimate both the output gap and the time-varying system parameters for the Turkish economy. Needless to say, the main motivation is the need for an output gap estimate in the construction of near term economic forecasts as a future indicator of inflationary pressures. Such pressures become even more important when the monetary authority is committed to maintain price stability in a forward-looking fashion.

We believe that the findings in this paper will serve as a reference in two distinct ways. First, we propose a new methodology that incorporates time-varying parameter framework into output gap estimation (an unobserved components model). Second, the resulting output gap series and its time-varying relationship with other variables will reveal important information for the changing transmission mechanisms in Turkey.

Needless to say, one should refrain from putting too much emphasis on a single output gap measure. While it is a functional tool in aiding the understanding and forecasting of inflation developments, it has some weaknesses as well. After all, it is an unobserved variable, highly model-specific, and also its link with inflation is not always stable. Moreover, it is not the main variable that drives inflation. For example, many recent studies that are based on micro-founded models point out that inflation is mostly driven by marginal costs, which is not necessarily

4 *As an example Öğünç and Ece (2004) find in their study that incorporating the supply side into the output gap system reduces the parameter uncertainty and the total standard error, hence improves output gap estimate.*

correlated with *contemporaneous* output gap measures.[5] Also, there may be some factors, such as the exchange rate dynamics, which affect both inflation and the output gap at the same time. In this case, the seemingly positive relationship between the two variables can be derived by a third factor. Finally, and importantly, some recent studies cast doubt on the positive relationship between inflation and the output gap implied by the Phillips curve notion.[6] Therefore, the output gap measures derived in this paper should by no means perceived as the sole determinants of future inflation in Turkey. It should rather be evaluated together with a range of other indicators of inflationary pressures.

The rest of the paper is organized as follows. In the next section, we present and discuss the model along with a state-space representation. The estimation methodology is also introduced in this section. In section 3, we present both the estimated output gap series and the estimated time varying parameters. A sensitivity analysis is performed in section 4, where we analyze whether the results remain robust to different specifications about inflation and the output gap. The comparison of the estimated gap series with the ones obtained from the HP filter and the SKF procedure is also displayed in this section. Finally, section 5 concludes.

2. The model

Empirical model

In this section, we present the general form of the model that is employed throughout the paper. Given the fact that univariate methodologies only make use of the information on the series to be detrended, thus lack economic content, we present a multivariate specification including a Phillips curve equation and a system of equations representing the output gap dynamics. Therefore, the parameters and the gap series are estimated to obtain consistent figures with the inflation rate and the underlying output gap dynamics. The general form of the model is as follows:

(1) *Inflation-Output Gap Dynamics:*

$$\pi_t = \alpha_{1,t}\pi_{t-1} + \alpha_{2,t}\pi_{t-2} + \alpha_{3,t}gap_{t-1} + \alpha_{4,t}reer_t + \nu_t$$

(2) *Actual Output Decomposition:*

$$y_t = y_t^* + gap_t$$

(3) *Potential Output Equation:*

$$y_t^* = y_{t-1}^* + \mu_{t-1} + \eta_t$$

(4) *Potential Output Growth Rate Equation:*

$$\mu_t = (1-\rho_t)\mu_0 + \rho_t\mu_{t-1} + \varepsilon_t$$

(5) *Output Gap Dynamics:*

$$gap_t = \gamma_{1,t}gap_{t-1} + \gamma_{2,t}r_t + \gamma_{3,t}DI_t + \gamma_{4,t}reer_t + \zeta_t$$

where π_t is the inflation rate defined as the logarithmic difference of quarterly seasonally adjusted consumer price index (CPI), gap_t is the unobserved output gap, $reer_t$ is the logarithmic difference of the real effective exchange rate, y_t is the logarithmic seasonally adjusted real gross domestic product, y_t^* is the unobserved potential output, μ_t is the potential output growth rate, r_t is the ex-post real interest rate based on 3-month Treasury auction rates, and DI_t is the demand

5 *See, for example, Clarida, Gali and Gertler (2002), in which this argument is forcefully demonstrated with special reference to European inflation dynamics.*

6 *As an example, Özlale and Özcan (2003) find evidence about the validity of a time-inconsistency problem for the Turkish economy in the last decade, implying a negative relation between inflation and the output gap. Also, Özbek and Özlale (2004) show that there is not strong evidence regarding a positive relationship between these two variables.*

index, which is constructed from the Business Tendency Survey of the Central Bank of the Republic of Turkey. The derivation of the demand index along with other data descriptions are presented in Appendix 3. Finally, ν_t, η_t, ε_t, and ξ_t represent shocks to the system, which are assumed to be i.i.d. with zero mean and constant variances.

It is important to remind that the parameters of the system are time varying. Therefore, one has to make a time-series specification for the evolution of these parameters. It is assumed that each time-varying parameter follows a random walk. Such a specification can be defended on theoretical grounds: Since any structural change on the dynamics of the model – thus the system parameters – cannot be known a priori, it is intuitive to specify a random walk process for each parameter. As a result, the system includes nine more equations, where each time-varying parameter follows a random walk process.

Equation (1) is a fairly standard reduced form Phillips curve specification including lagged inflation terms, lagged output gap, and the change in the real effective exchange rate. Accordingly, persistence in inflation is reflected in inertial terms up to two lags. Output gap is assumed to affect inflation with a lag since it takes time for the pressure on production costs to be revealed and for prices to be adjusted in response to a demand shock. On the other hand, changes in the real effective exchange rate capture the effects of the exchange rate dynamics on the inflation both through the "cost of production channel" and through the prices of imported final goods.[7]

Equation (2) is the identity defining output as the sum of the potential output (trend component) and the output gap (transitory component). Equation (3) defines potential output as a random walk with a drift model, implying that shocks to trend output are permanent. Moreover, the drift term, trend growth, is allowed to vary over time and the persistence can be shaped with respect to different values of ρ. Needless to say, trend growth may change over time along with labor force, productivity or technology developments. Moreover, in a recent study, Aguiar and Gopinath (2004) state that emerging markets are subject to extremely volatile shocks to the stochastic trend and provide evidence that emerging market business cycles are driven by shocks to the trend growth rate which may result from extreme and relatively frequent changes in economic policies. Taking these factors into account, potential growth rate is modeled as time-variant. In this respect, equation (4) defines potential growth as a first-order autoregressive process with long-run average growth rate of μ_0 and autoregressive coefficient $0 < \rho < 1$ representing the persistence in trend growth. The magnitude of ρ shows the persistence of the deviations from the long run growth rate μ_0.[8] The system allows for a time-varying estimation of the persistence parameter. An estimated ρ close to one means the potential output can deviate from the steady state for substantially long periods. In this respect, the setting also provides a framework to test the "cycle is the trend" hypothesis in emerging market economies, which is discussed in Aguiar and Gopinath (2004).

Equation (5) specifies the output gap dynamics. Rather than modeling the output gap by a purely stochastic process – as most models do – we include variables that can provide extra information on the evolution of the output gap. We choose the variables employed in the equation so as to characterize a set of broad macroeconomic variables that may affect the actual output, but not the potential output. The main variables included are the lagged output gap, real interest rate, the expectations of the business sector participants in the economy and the changes in the real effective exchange rate. Although these factors are vital in explaining the output performance in the short-run, they are viewed to be more effective on the demand side, and thus, neutral for the behavior of potential output in the short run. Therefore, these variables appear in the output gap equation to account for the deviation of actual output from its potential level.

In this respect, the real interest rate undertakes its traditional role as affecting the consumption and the investment behavior in the economy. The inclusion of DI_t captures the information content that is embedded in the private sector expectations. In other words, expectations index is supposed to capture the firms' prospects about demand conditions, which may be a major determinant of the output gap along with the real interest rate.

Although the role of expectations and the real interest rates in the whole system are clear, the role of the exchange rate dynamics on the output gap is less certain and thus should be discussed in details. As mentioned in the first section, the seemingly positive relationship between inflation and the output gap can partly be driven by another factor, which could affect both of the variables contemporaneously. In this context, since the real exchange rate can play such role, it is included both in the first and the fifth equations. Moreover, the real exchange rate affects the output gap

7 *See Leigh and Rossi (2002) for more on exchange rate pass-through in Turkey.*

8 *The sustainable steady-state real growth rate for the economy is assumed to be 4.5 percent on annual basis, which corresponds to 1.106 percent per quarter. As shown in the sensitivity part, our output gap computations are fairly robust to the underlying assumption for the steady state growth rate.*

through two distinct and opposite channels. First, an appreciation leads to a decline in the relative prices of imported goods and leads to a temporary hike in actual demand, raising the gap between the actual and the potential. Second, in a country where capital goods are mostly imported, *ceteris paribus*, an exchange rate appreciation induces firms to substitute capital for labor, leading to an increase in labor productivity, and thereby increasing the potential output as well. Also, such an appreciation will lead to a decrease in the cost of imported intermediate goods, which would increase the supply in the economy. The question of which of these factors dominates may depend on the specific state of the economy. Therefore, the net effect of the changes in the exchange rate on the output gap is not clear, at best. Various factors such as the exchange rate elasticity of net exports, the magnitude of the exchange rate pass-through, the importance of imported capital goods in the overall production should be taken into account in order to reach a robust conclusion.

It can also be argued that several other exchange rate measures could be used both in the inflation specification and the output gap equation. For example, import price inflation could be a better candidate in explaining the exchange rate pass-through to prices while deviations of the real exchange rate from its long-run trend could be used in the output gap specification.[9] However, we believe that a common variable, which could be effective on both the inflation and the output gap dynamics, could be more appropriate to identify the inter-linkages among the dynamics of output, inflation and the exchange rates. Therefore, the changes in the real effective exchange rate have been used in both the inflation and the output gap specification.

State-space representation of the model

State-space modeling has been extensively used in the estimation of potential output in recent years. It does not only provide the opportunity for building encompassing models, but also simplifies the formulation of rather complicated problems. Besides, once the model is written in a state-space form, it becomes straightforward to obtain the required estimates by utilizing the Kalman filter algorithm. The general form of the state-space formulation can be represented as:

$$x(t) = Fx(t-1) + Gu(t) + e_1(t)$$
$$z(t) = Hx(t) + e_2(t)$$

where e_1 and e_2 denote vectors of normally distributed i.i.d. shocks which are assumed to be uncorrelated and have covariance matrices R_1 and R_2, respectively. Furthermore, $u(t)$ is the vector of exogenous variables. In this respect, our measurement equation, where the evolution of the observed variables (inflation and output) is described as a function of the unobserved state variables, is as follows:

$$\begin{bmatrix} \pi_t \\ y_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \pi_t \\ \pi_{t-1} \\ y_t^* \\ \mu_t \\ gap_t \end{bmatrix}$$

The second measurement equation is the identity specified in equation (2) of the model, which states that the actual output is equal to the sum of the potential output and the output gap. The unobserved variables (potential output, potential output growth rate and the output gap) and inflation rate evolve according to the following transition equation:

$$\begin{bmatrix} \pi_t \\ \pi_{t-1} \\ y_t^* \\ \mu_t \\ gap_t \end{bmatrix} = \begin{bmatrix} \alpha_{1,t} & \alpha_{2,t} & 0 & 0 & \alpha_{3,t} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & \rho_t & 0 \\ 0 & 0 & 0 & 0 & \gamma_{1,t} \end{bmatrix} \begin{bmatrix} \pi_{t-1} \\ \pi_{t-2} \\ y_{t-1}^* \\ \mu_{t-1} \\ gap_{t-1} \end{bmatrix} + \begin{bmatrix} \alpha_{4,t} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & (1-\rho_t) & 0 & 0 \\ \gamma_{4,t} & 0 & \gamma_{2,t} & \gamma_{3,t} \end{bmatrix} \begin{bmatrix} reer_t \\ \mu_0 \\ r_t \\ DI_t \end{bmatrix} + \begin{bmatrix} \nu_t \\ 0 \\ \eta_t \\ \varepsilon_t \\ \zeta_t \end{bmatrix}$$

9 *Employing Prior Consistent filter, Benes and N'Diaye (2003) uses such a measure. However, following Meese and Rogoff (1983) and Edwards and Savastano (1999), there is not a commonly agreed methodology in determining the equilibrium real exchange rate, and thus the deviations from such equilibrium.*

where η_t, ε_t, ξ_t and ν_t are assumed to be independent and identically distributed normal white-noise processes.

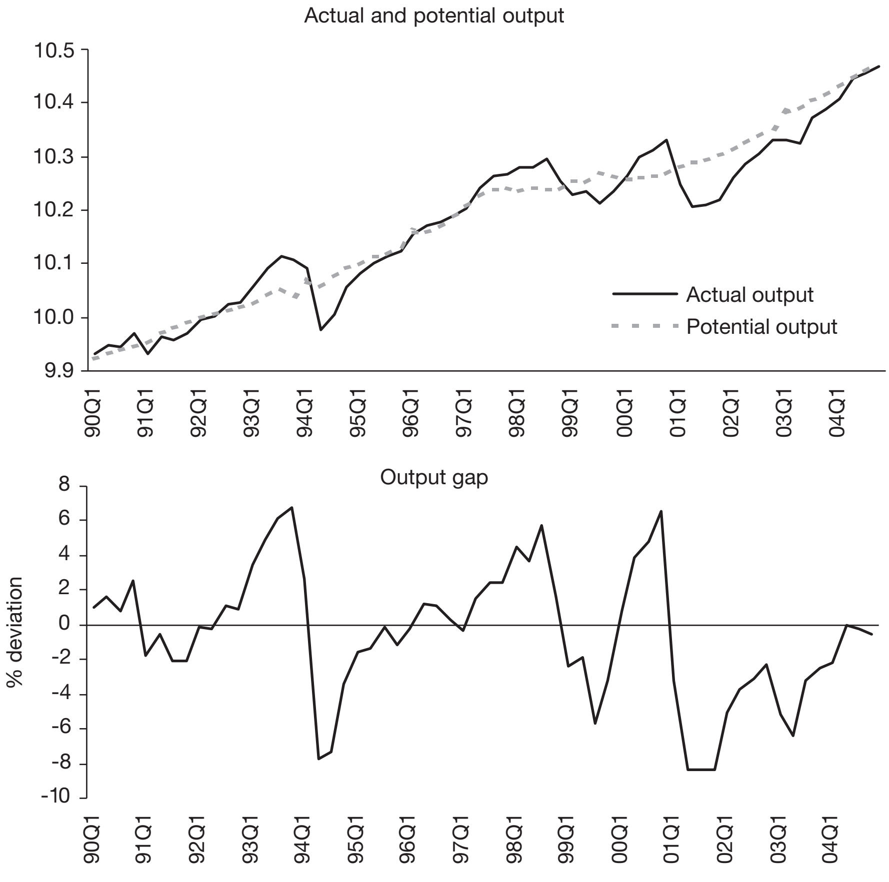
Non-linearity in state space models and EKF

When the transition equation is analyzed, it can be seen that the transition matrix consists of time-varying parameters to be estimated. Moreover, both these parameters and the state variables, which are to be estimated simultaneously, are presented in multiplicative form. Such a representation causes the state space model to take a non-linear form, where SKF becomes inappropriate to employ. In this context, EKF emerges as the estimation methodology, which consists of using the SKF equations to the first-order approximation of the non-linear model about the last estimate. In this case, the time-varying parameters, which were all assumed to follow random walk, are treated as a new state vector and added to the initial transition equation. Then, EKF procedure can be applied to estimate the new state vector, which contains the parameter vector as one of its components. The EKF procedure, its application and the smoothing algorithm are described more comprehensively in Appendix 1.A.

3. Empirical findings

Based on the discussion above, first, the parameter vector, which contains nine equations, is formed. Next, the state vector and the parameter vector are combined to produce the new enhanced transition equation. Finally, EKF is applied to obtain the estimates for both the state variables and the time-varying parameters in the model. This section displays and interprets these estimation results. Initially, we start with the estimated potential output and the output gap, which are presented below in Figure 1.

Figure 1 – Actual output, potential output and output gap



Interpretation

The output gap estimate displays the fact that the Turkish economy has been subject to severe shocks – economic crises that created an unstable environment, ending with rapid contractions in the economic activity – for several times in the sample period. As a consequence, not only the actual output but also the output gap estimates exhibit sharp movements, which could be interpreted as a common characteristic of emerging market economies. Indeed, the estimated potential output is far less smooth than a potential output that could be obtained using a standard filtering technique. These findings are also in line with Aguiar and Gopinath (2004), which argue that shocks to potential output are vital in explaining the business cycles in emerging markets.

According to the output gap figure, expansion periods were generally interrupted by economic crises so that the last decade seems to have witnessed three separate periods of recession. Such an observation validates both the excessive boom-bust cycle peculiar to emerging markets and the "unsustainable growth path" that the Turkish economy followed in the last decade. At the second quarter of 1994, due to the deep financial crisis, output declined drastically. The devastating effects of the crisis brought about a negative output gap of almost 8%. Following the trough at 1994Q2, the economy started to recover and actual output exceeded its potential level approximately after seven quarters. Between 1996 and 1998, the actual output remained above its potential level where 1998Q3 was a peak for the economy. This was a period, where high interest rates attracted short-term capital, fiscal policy was expansionary and the real exchange rate was targeted in the pursuit of financial stability with an accommodative monetary policy. However, it was obvious that such an expansionary period increased the fragility of the economy, and thus destined to be rather short-lived. The period starting from 1998Q3 has witnessed several internal and external shocks hitting the Turkish economy. The Russian crisis at 1998Q3 and the devastating earthquake at 1999Q3 were the major shocks driving the economy into a recession, which manifested itself as a trough at 1999Q3.

In order to permanently solve the prolonged problem of high inflation and unsustainable growth, Turkey announced an IMF-supported exchange rate-based stabilization program in the context of a crawling peg regime to be adopted by the beginning of 2000. It did not take a long time for the economy to recover. Indeed, the revival was quick, characterized by a rapid expansion in the output and thus the output gap. The early stages of the program witnessed a sizeable drop in the interest rates, rapid credit growth, and appreciation of the Turkish lira in real terms, eventually boosting the domestic demand. Consistent with the underlying story, our measures point out a positive output gap by the beginning of 2000Q1.

In the year 2000, the high-rated rises in imports, due to the expanding domestic demand and production, caused concerns about the sustainability of the current account. Moreover, due to the fragile structure of the banking sector, having potential problems such as maturity mismatches and open foreign exchange positions, the economy was even more vulnerable to speculative attacks. In addition to the anxiety associated with economic dynamics, the political problems such as the reluctance of the government in delivering structural reforms became evident at that time. As perceptions of the vulnerability of the economy became more evident, the economic agents were already questioning the prospects of the macroeconomic program. The first signal of the failure of the program came in November 2000. Succeeding collapse of the crawling peg regime at February 2001 dragged the economy into the deepest recession that Turkey has experienced ever. The crisis had been very detrimental to the economic activity, which led to a trough at 2001Q4. In the post-crisis period, the implementation of sound macroeconomic policies induced significant achievements on the way to economic stability. In this respect, along with the declining inflation, high growth rates were attained after 2002. As a consequence, output gap has closed significantly, approaching to zero as of the second quarter of 2004.

The figures reveal that, although the output gap seems to have closed as of the first quarter of 2004, its slow convergence has contributed dramatically to the disinflation process since the February 2001 economic crisis. On the other hand, one should be careful in interpreting the output gap measures. The measure we present does not capture explicitly the labor market conditions in the economy. In that sense, it may be misleading to make judgments on future inflation just by looking at the output gap. Output gap is only one of the many indicators of overall resource utilization in the information set of the policy maker. One should also carefully evaluate the labor market conditions, unemployment rate, and the other cost factors such as exchange rates, wages and energy prices to have a more reliable assessment of future inflationary pressures.[10]

10 *As it will be shown later, the output gap estimate remained robust to alternative specifications.*

Time-varying parameter estimates

Time-varying parameter estimates of the system are consistent with the observations of the last decade and thus easy to interpret. The evolutions of these parameters, which can be analyzed due to the recursive nature of the EKF algorithm, are displayed in Appendix 2. The common emphasis of the recursive estimates is that financial crises had a significant influence on the parameters. Especially, dramatic changes in the parameter estimates can be observed after the 1994 crisis.

The autoregressive coefficient of potential growth is estimated between 0.75 and 0.85 through the whole sample, and 0.77 at the end, implying that potential growth is fairly persistent. This means, in the absence of shocks, output growth would converge within 1 percent of the steady-state rate in just about 5 years. Therefore, we can interpret shocks to the trend growth as "near-permanent" in Turkey. On the other hand, the sum of the coefficients regarding π_{t-1} and π_{t-2} is high, pointing to a significant degree of persistence implying a strong inflationary inertia (Table 1).

Table 1 – End-sample estimates for the model parameters

Model parameters	Baseline model	Alternative model 1	Alternative model 2
α_1	0.37	0.26	0.61
α_2	0.59	0.65	0.31
α_3	0.43	0.12	0.27
α_4	−0.03	−0.07	–
rho	0.77	0.71	0.78
γ_1	0.21	0.52	0.25
γ_2	−0.08	−0.07	−0.02
γ_3	0.19	–	0.21
γ_4	0.02	–	0.02

When the parameter estimates in the inflation specification are examined, it is seen that the depreciation of the currency and the positive output gap exert pressure on inflation, as expected. The depreciation of the domestic currency increases inflation both by an increase in the prices of imported final goods and by an increase in the cost of production through imported capital goods. On the other hand, the positive relationship between inflation and the output gap is consistent with the Phillips curve notion.[11]

The role of the real exchange rate on the output gap measures, which is shown with the parameter γ_4, should also be discussed. Keeping in mind that an increase in the series point out an appreciation of the Turkish Lira, we find a positive relationship between the real exchange rate changes and the output gap. Such a finding implies that the cost channel of the exchange rate on the output gap dominates the demand channel. In other words, appreciation of the domestic currency leads to a significant decrease in the cost of the imported capital goods, thereby results in an increase in the production. Such an increase outweighs possible decrease in the net exports that may be caused by the appreciation of the Turkish Lira.

As a result, the findings suggest that, due to the heavy importance of the imported goods on both the production process and the consumer basket, the periods of low inflation and positive output gap coincide with the periods of appreciation. In addition, even after controlling for the role of the exchange rates, we still find that a positive output gap causes an inflationary pressure in the economy.

Real interest rate is of expected sign in explaining the output gap dynamics, except for the beginning of the sample period. Moreover, the effect of real interest rate on output gap as of 2004Q2 is estimated as −0.08. Until 1996, we observe that real interest rates and output gap have a positive relationship. It is interesting to observe that, after 1996, the relationship between real interest rate and output gap turns out to be negative, in line with economic theory. The unstable empirical coefficient of the effect of the interest rate on output gap in Turkey can be

11 *In a previous study, Öğünç and Ece (2004) also find a positive relation between inflation and output gap with a parameter estimate of 0.24.*

attributed to several factors. First of all, interest rates had never been used as a policy instrument to attain the inflation target, prior to the floating regime. Instead, most of the sample period is dominated by crawling pegs/fixed exchange rate regimes, where the exchange rate developments have been the primary determinant of inflation and output dynamics in Turkey. In this respect, during 80's and 90's, interest rate sensitivity of aggregate demand was limited and conventional monetary transmission mechanism was not evident in Turkey. Second, in an unstable economy with high and persistent inflation and uncertainty, real interest rates generally remained at so high levels. This, in turn, rendered the economic agents' decisions as insensitive to intertemporal shifts in the real interest rates, weakening the link between the interest rates and spending decisions.

Time varying parameter estimates suggest that, transition to a financial restructuring period and a new policy regime – implicit inflation targeting along with the floating exchange rate regime – in February 2001 have had significant effects on underlying economic dynamics in Turkey. Since then, short-term interest rates have been actively used as the policy instrument to attain announced targets for the inflation rate. Moreover, there is evidence on weakened exchange rate pass-through on prices. As a consequence, developments after 2001 crisis indicate that, the relative weight of the interest rates, as a determinant of inflation and output dynamics, has significantly risen. Therefore, our findings on time varying parameter of the real interest rate may be a sign of increased effectiveness of interest rate as a policy instrument in expectations management and output dynamics.

As a result, all of the parameter estimates in the model are reasonable when the characteristics of the Turkish economy in the last decade are taken into account. Moreover, the changing impact of the real interest rate on the output gap signals that interest rates are more of a policy instrument in the recent years than they used to have been.

4. Sensitivity analysis

In order to analyze whether the results presented in the previous section are sensitive to the specification of the model (and to the magnitude of the shocks), we estimate both the output gap and the system parameters along with the other two specifications about the output gap dynamics. Also, we analyze whether our output gap estimates are sensitive to the underlying steady state growth rate of potential output. It is seen that, our output gap computations are fairly robust to the underlying assumption for the steady state growth rate. Finally, in order to test our results against alternative methods, we estimate the output gap with both the HP filter and the SKF to make a sound comparison.

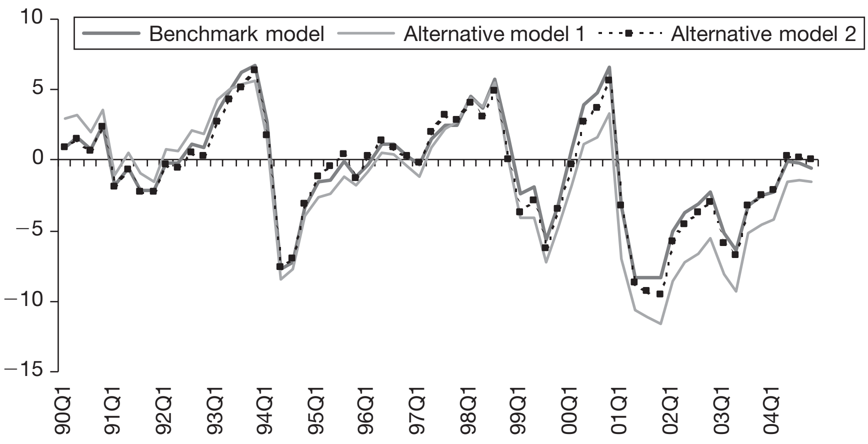
Alternative models

As mentioned above, in view of the inherent uncertainty regarding the main characteristics of Turkish economy, we have developed two alternative models in addition to the one described above. One of these models implies a distinct output gap dynamics while the other one excludes real exchange rate from the inflation equation. Monitoring the results of different specifications will provide flexibility for better evaluation and interpretation of the output gap estimates.

In the first alternative model, we have only used the lagged value of the output gap and the real interest rate in the output gap specification. Such a restriction imposes the assumption that the output gap is affected solely by the monetary policy actions. Such an imposition can also be defended on the grounds that expectations may improve as well when there is an increase in the potential level of the economy. Therefore, both temporary and permanent developments can be reflected in forming the expectations in the economy.

Finally, in our second alternative, we specify inflation only as a function of its lagged values and the output gap while remaining equations appear as in the benchmark model. Assuming that changes in the exchange rate affect prices with a lag, it can be argued that the effects of the exchange rate are already inherent in the lagged values of the inflation. Figure 2 displays the output gap estimates obtained under three specifications, our benchmark model and two alternative models. As mentioned above, in the first alternative, only the monetary policy actions are assumed to affect the output gap equation. Finally, in the second alternative setup, assuming that the impact of the changes in the real exchange rate on the current inflation are also reflected in the lagged values of inflation, real exchange rate variable is omitted in the inflation specification.

Figure 2 – Output gap estimates under different models



The end-sample coefficient estimate of the output gap in the inflation specification for the first alternative model is 0.12. On the other hand, in the second alternative model, when the coefficient of the real exchange rate is restricted to zero, output gap still remains to be a significant determinant of inflation with an estimated coefficient of 0.27 at the end of the sample (Table 1). Therefore, the role of the output gap on the inflation dynamics is found to be robust regardless of the presence of the exchange rate dynamics in the inflation specification. The exercise of estimating the output gap under different model specifications implies that, in line with economic theory, a positive (negative) output gap is inflationary (disinflationary) in Turkey.

The figure reveals that, the three models exhibit a similar pattern qualitatively, whereas there are only minor differences in quantitative terms. Moreover, these discrepancies mostly occur in the peak or the trough of the cycle. All of the recessionary and the expansionary periods are evident. All in all, it can be conveniently claimed that all of the models point out a similar output gap path and the results are fairly robust to different specifications about output gap and inflation dynamics.

Alternative steady state growth rate assumptions

The method employed in this paper primarily focuses on estimating the output gap considering its assumed own dynamics as well as its relationship with inflation. In this framework, potential output is defined as following a purely stochastic process, rather than being a structural estimate. On the other hand, production function approaches derive potential output from structural models integrated with main determinants of growth such as physical and human capital, labor, and technology. Hence, given the absence of such a measure for potential output, one have to consider the uncertainty on the long-run potential growth rate parameter, μ_0, which is exogenously imposed to the model. The figure points to the robustness of the output gap estimate to alternative steady state trend growth assumptions (Figure 3).

Sensitivity to data revisions (comparison with HP filter)

One major drawback of the typical output gap estimates used in the literature is the sensitivity to the data revisions or new data. Moreover, because the filters use both past and future data, there is a problem at the end of a sample due to absent future data. This problem is generally entitled as “end point” problem. In order to assess whether our results suffer from the same syndrome and to see how sensitive they are to the inclusion of new data or revisions, our estimates will be compared with the HP filter results on the basis of uncertainty for the last estimates. It is expected that, the output gap estimates should not be revised too much and exhibit large swings when new observations are added into the analysis.

Figure 4 compares the results from two filtering methods, when the sample is cut at two alternative points – 2002:Q1 and 2003:Q1 – in addition to full sample. The resulting HP filter estimates exhibit major differences. Just to give a striking example, when the sample is cut at 2003:Q1, output gap is positive with a magnitude of 3.2 percent, where as the full sample

Figure 3 – Output gap estimates under different steady state growth rate assumptions

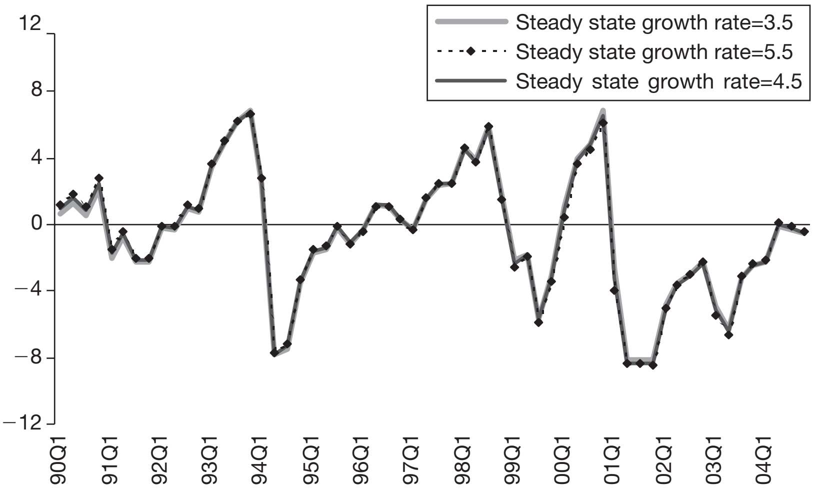
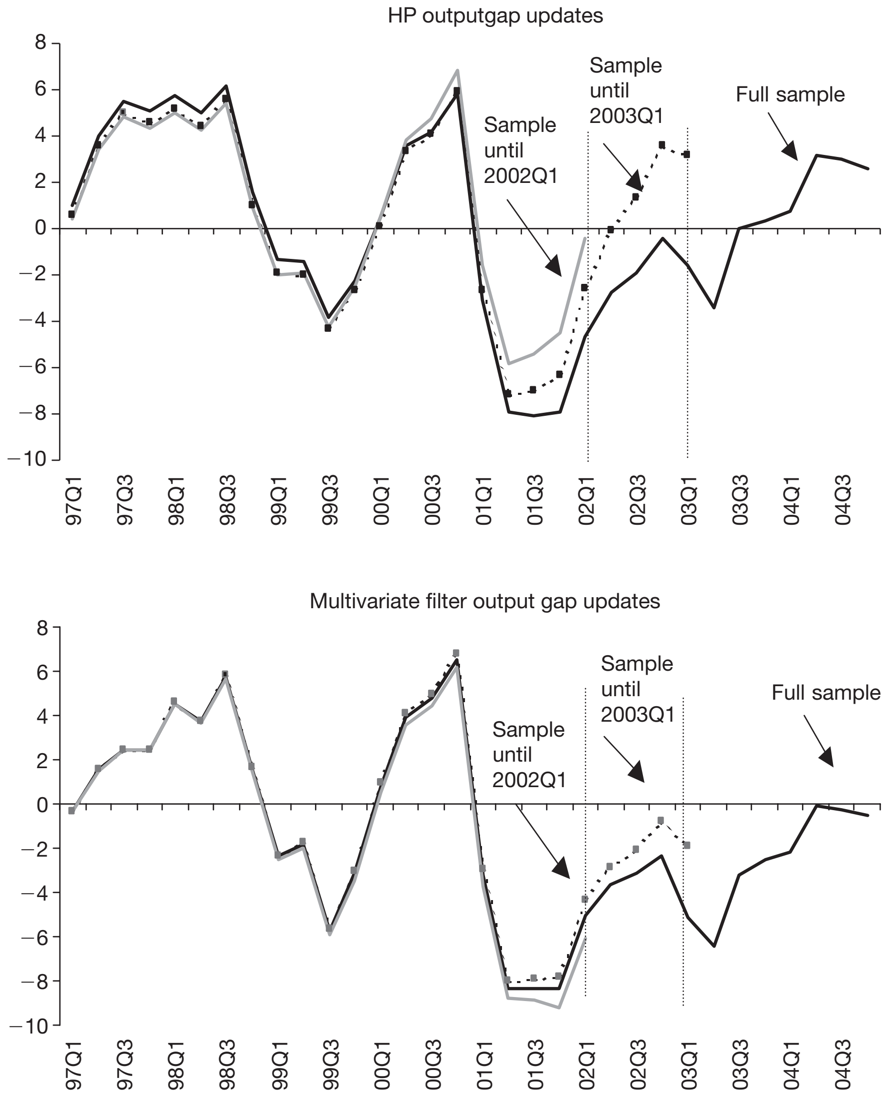


Figure 4 – HP versus MV filter revisions

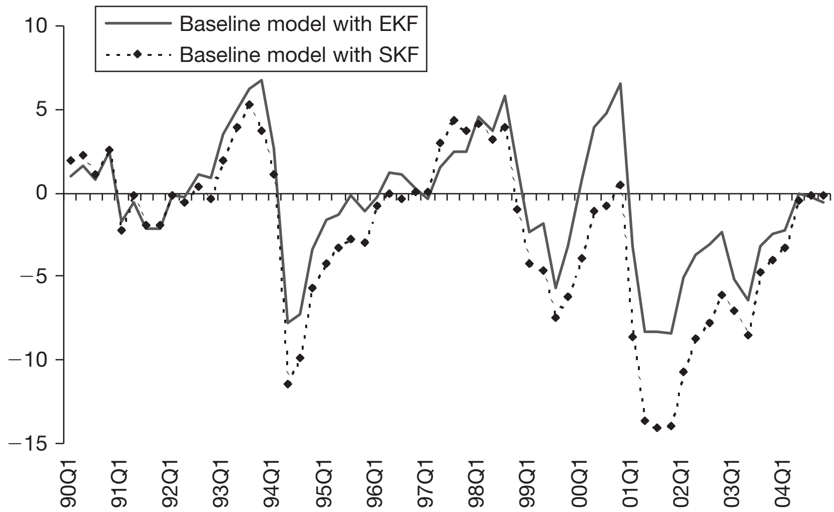


estimates point to a negative gap with a magnitude of 1.6. However, the same exercise for the multivariate filter does not produce diverging results. The gap estimates are close to each other and fairly insensitive to the inclusion of new data.

Extended vs. Standard Kalman Filter

The major contribution of this study to the existing literature on the estimation of output gap is the time-varying parameter framework, which is believed to be more appropriate regarding the economic dynamics specific to the Turkish economy. Such a methodology imposes a non-linearity to the model and thus requires the use of EKF. In this section, it will be investigated that whether the assumption of time-varying parameters significantly differs from that of time-invariant parameters, in terms of the resulting estimates for the output gap. In order to do so, we estimate our baseline model with the parameters assumed to be constant over time by utilizing SKF and compare the results with those obtained from the EKF (Figure 5).

Figure 5 – Comparison of EKF and SKF results for the output gap



The alternative output gap figures exhibit a resemblance in terms of their turning points whereas significant level differences are observed especially for the two crisis periods – 1994 and 2001 – and the expansion period of 2000. Contrary to EKF estimates; even high economic activity observed throughout the year 2000 could not push the output gap to positive levels according to SKF results. On the other hand, end-sample information revealed by two estimates is similar, with actual output approaching to its potential level.

5. Summary and concluding remarks

This paper has presented a time-varying parameter methodology for constructing a reliable estimate of potential output and thus output gap by exploiting the EKF technique in a multivariate setting.

Estimating the potential output and output gap in a multivariate setting has several advantages over univariate techniques, such as the HP filter, which have a number of well-documented shortcomings. First, our multivariate model incorporates more economic content by the inclusion of inflation and output gap dynamics. Second, considering the volatile economic structure peculiar to emerging market economies including Turkey, it is more appropriate to employ a time-varying parameter model to account for policy changes, structural shifts and thus varying economic relationships.

Our output gap estimate points out that business cycle of Turkey displays sharp turning points rather than exhibiting a smooth pattern. Sharp transitions between expansion and recession periods reflect the volatile nature of the Turkish economy, which has been subject to several internal and external disturbances causing severe economic crises in the last ten years. The uncertainties in such an environment and the dominance of the previous regimes where inflation and output dynamics were primarily determined by exchange rate movements, real interest rate appears as impotent in explaining output gap dynamics in Turkey. However, we think that the extent of the relation has been subject to change with the transition to the new policy regime after February 2001. In this respect, the behavior of the time-varying coefficient of the real interest rate observed towards the end of the sample may be perceived as a signal of this consideration.

Being an unobserved variable, output gap can only be estimated, thus it turns out to be model-specific. Therefore, establishing the best model representing the true structure of the economy is of great importance. Moreover, data revisions give rise to criticisms about real-time accuracy of the output gap estimates, since the information available to the policy maker at the time that policy decisions are taken, can be subject to significant changes. Furthermore, due to the lags in the announcement of the data, the projections for the observed variables gain importance especially for the reliability of the end-sample estimates. Though having a number of shortcomings and practical problems, output gap, as an aggregate measure, is still being used extensively by central banks as an indicator of inflationary pressures. However, it would also be better to recognize that output gap is just one of many indicators of resource utilization and inflationary pressures. Hence, simultaneous interpretation of the output gap measures along with the information provided by other indicators would give rise to more reliable evaluations. In this way, any relevant information that is not captured by the model dynamics can be taken into account as well. For instance, labor market dynamics is not represented explicitly in our model. When the non-inflationary, albeit high-rated, growth performance of the Turkish economy in the last three years is considered, one should not undervalue the role of labor market developments characterized with high unemployment rates and declining real wages. In other words, the rigidities in the labor market, which has been contributing to recent disinflationary process in Turkey, should also be taken into account in policy analysis. Therefore, the joint estimation of output gap and unemployment gap may be the one remaining issue for future research in order to account for labor market dynamics as well.

References

Aguiar, M. and Gopinath, G. (2004) "Emerging Market Business Cycles: The Cycle is the Trend", mimeo, University of Chicago.

Anderson, B.D.O and Moore, J.B. (1979) Optimal Filtering. Prentice Hall.

Benes, J. and N'Diaye, P. (2003) "A Multivariate Filter for Measuring Potential Output and the NAIRU", The Czech National Bank's Forecasting and Policy Analysis System, Chapter 5, 99–118.

Billmeier, A. (2004) "Ghostbusting: Which Output Gap Measure Really Matters?", IMF Working Paper No: 04/146.

Boone, L., Juillard, M., Laxton, D. and N'Diaye, P. (2001) "How Well Do Alternative Time-Varying Parameter Models of the NAIRU Help Policy Makers Forecast Unemployment and Inflation in the OECD Countries?", IMF Working Paper, Draft Version.

Butler, L. (1996) "A Semi-Structural Method to Estimate Potential Output: Combining Economic Theory with a Time-Series Filter", The Bank of Canada's New Quarterly Projection Model (QPM), Part 4.

Clarida, R., Gali, J. and Gertler, M. (2002) "A Simple Framework for International Monetary Policy Analysis", NBER Working Paper No: 8870.

Edwards, S. and Savastano, M.A. (1999) "Exchange Rates in Emerging Economies: What Do We Know? What Do We Need To Know", NBER Working Paper No: 7228.

Gomez, V. and Maraval, A. (1998) "Seasonal Adjustment and Signal Extraction in Economic Time Series", Banco de Espana, No. 9809.

Harvey, A.C. (1990) Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press.

Kuttner, K.N. (1994) "Estimating Potential Output as a Latent Variable", Journal of Business and Economic Statistics, Vol.12 (3), 361–68.

Kuttner, K.N. (1999) "A Time-Series Approach to Potential Output", Prepared for the meeting of the CBO Panel of Economic Advisers, Federal Reserve Bank of New York.

Laxton, D., Isard, P., Faruqee, H., Prasad, E. and Turtelboom, B. (1998) "MULTIMOD Mark III: The Core Dynamic and Steady-State Models", IMF Occasional Paper, No.164.

Leigh, D. and Rossi, M. (2002) "Exchange Rate Pass-Through in Turkey", IMF Working Paper No: 02/204.

Ljung, L. and Soderstrom, T. (1983) Theory and Practice of Recursive Identification. The MIT Press.

Meese, R. and Rogoff, O. (1983) "The Out-of Sample Failure Of Empirical Exchange Rate Models", in Exchange Rate and International Macroeconomics (Ed.) J. Frenkel, University of Chicago Press, Chicago.

Orphanides, A. (1999) "The Reliability of Output Gap Estimates in Real Time", Finance and Economics Discussion Series 1999-38, August 1999.

Öğünç, F. and Ece, D. (2004) "Estimating the Output Gap for Turkey: An Unobserved Components Approach", Applied Economics Letters, Vol.11 (3), 177–182.

Özbek, L. and Özlale, Ü. (2004) "Employing The Extended Kalman Filter In Measuring The Output Gap", forthcoming in Journal of Economic Dynamics and Control, 2004.

Özlale, Ü. and Metin-Özcan, K. (2003) "Does Time Inconsistency Problem Apply For Turkish Monetary Policy", Manuscript, Bilkent University, Department of Economics, Ankara, Turkey.

Taylor, J. (1993) "Discretion versus Policy Rules in Practice", Carnegie-Rochester Conference Series on Public Policy, Vol.39, pp. 195–214.

Çağrı Sarıkaya, Fethi Öğünç, Dilara Ece, Ayşe Tatar Curl, and Arzu Çetinkaya
(Central Bank of the Republic of Turkey)

November 2004

APPENDIX

1.A Extended Kalman filter algorithm[±]

Following algorithm presents the results when the extended Kalman filter is applied to the estimation of the parameter vector of $\theta = (\alpha_1,\alpha_3,\rho,\gamma_1,\alpha_2,\alpha_4,\gamma_2,\gamma_3)$. In addition to the usual Kalman filter algorithm, in here we have also random parameters, which are assumed to be evolving according to the random walk process. Simply, in the EKF case, because of the non-linear relationship, we linearize the process and measurement functions at the current state estimate by using the partial derivatives and then apply the usual Kalman filter algorithm.

The unobserved state vector X can be seen as partitioned into two parts: one is the usual unobserved state variables and other is the unknown parameter vector:

$$X(t) = \begin{pmatrix} x(t) \\ \theta(t-1) \end{pmatrix},\ \bar{K}(t) = \begin{pmatrix} K(t) \\ L(t) \end{pmatrix},\ \bar{P}(t) = \begin{pmatrix} P_1(t) & P_2(t) \\ P_2^T(t) & P_3(t) \end{pmatrix} \quad \textbf{1.A.1}$$

where $\bar{K}$ and $\bar{P}$ are the Kalman gain and covariance matrix for the extended state. Then the general algorithm:

$$S_t = H_t P_1(t) H_t^T + H_t P_2(t) D_t^T + D_t P_2^T(t) H_t^T + D_t P_3(t) D_t^T + R_2 \quad \textbf{1.A.2}$$

$$L(t) = \left[P_2^T(t) H_{t-1}^T + P_3(t) D_t^T \right] S_t^{-1} \quad \textbf{1.A.3}$$

$$K(t) = \left[F_t P_1(t) H_t^T + M_t P_2^T(t) H_t^T + F_t P_2(t) D_t^T + M_t P_2(t) D_t^T + R_{12} \right] S_t^{-1} \quad \textbf{1.A.4}$$

$$\begin{aligned} P_1(t+1) = {} & F_t P_1(t) F_t^T + F_t P_2(t) M_t^T + M_t P_2^T(t) F_t^T + M_t P_3(t) M_t^T \\ & - K(t) S_t K^T(t) + R_1 \end{aligned} \quad \textbf{1.A.5}$$

$$P_2(t+1) = F_t P_2(t) + M_t P_3(t) - K(t) S_t L^T(t) \quad \textbf{1.A.6}$$

$$P_3(t+1) = P_3(t) - L(t) S_t L^T(t) + Q \quad \textbf{1.A.7}$$

$$\hat{x}(t+1) = F_t \hat{x}(t) + G_t u(t) + K(t)\left[z(t) - H_t \hat{x}(t) \right] \quad \textbf{1.A.8}$$

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)\left[z(t) - H_{t-1} \hat{x}(t) \right] \quad \textbf{1.A.9}$$

Here, $F_t = F(\hat{\theta}(t)),\ G_t = G(\hat{\theta}(t)),\ H_t = H(\hat{\theta}(t))$ **1.A.10**

$$M_t = M(\hat{\theta}(t), \hat{x}(t), u(t)), \text{ with } M(\hat{\theta}, x, u) = \frac{\partial}{\partial \theta}\left[F(\theta) x + G(\theta) u \right]\Big|_{\theta = \hat{\theta}} \quad \textbf{1.A.11}$$

and $D_t = D(\hat{\theta}(t-1), \hat{x}(t))$ with $D(\hat{\theta}, x) = \frac{\partial}{\partial \theta}\left[H(\theta) x \right]\Big|_{\theta=\hat{\theta}}$ **1.A.12**

Finally, regarding the initial values:

$$P_1(0) = P_{1,0},\ P_2(0) = 0,\ P_3(0) = P_{3,0},\ \hat{x}(0) = x_0,\ \hat{\theta}(0) = \theta_0 \quad \textbf{1.A.13}$$

[±] *The alogrithm is discussed in details in "Theory and practice of recursive identification" by Ljung and Soderstorm (1983).*

1.B The smoothing algorithm

The potential output can be estimated in two different ways depending on what information is used. The filtered estimate at time t is one-sided and it uses information up to time t ($y^*_{t|t}$). Therefore, Kalman filter, used as a real-time or online algorithm, estimates the state vectors exploiting the current and past information. On the other hand, a smoothed value is two-sided and uses information from the whole sample, up to time T ($y^*_{t|T}$ where $0 \le t \le T$). In this way, the smoothing algorithm allows for considering future information as well, in the estimation of potential output. Unless there is some immediate real-time constraint, state estimates can be improved by using the smoothing algorithms. Referring to the fixed interval-smoothing (Rauch-Tung-Striebel Two-Pass Smoother) algorithm, the smoothed estimator can be defined as:

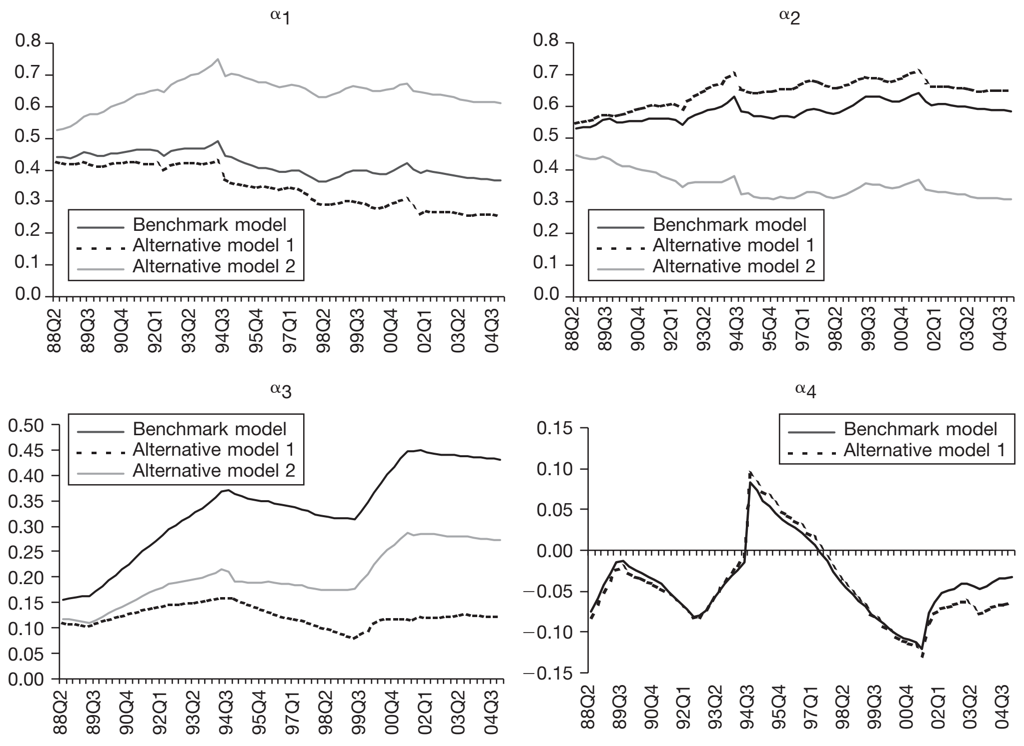
$$X(t-1|T) = X(t-1|t-1) + B(t)[X(t|T) - X(t|t-1)] \qquad \textbf{1.B.1}$$

with its corresponding covariance matrix,

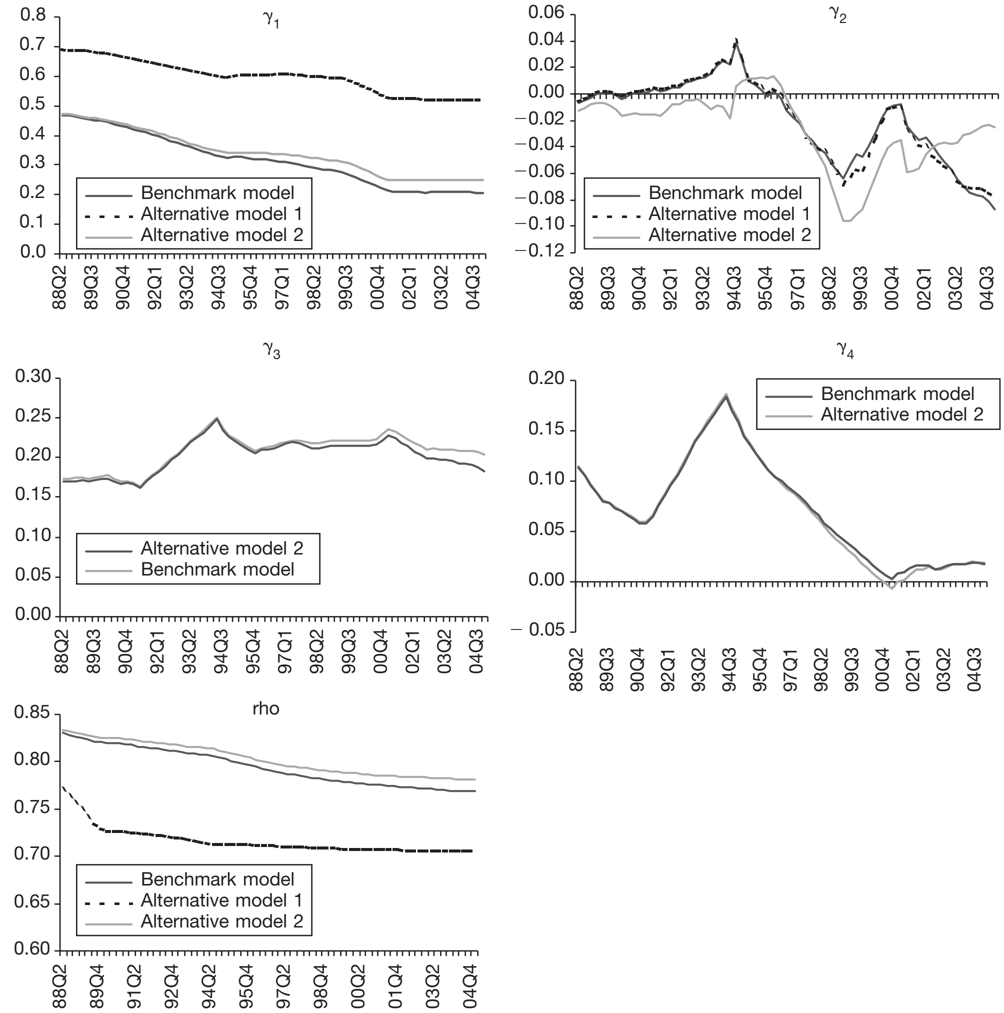
$$\bar{P}(t-1|T) = \bar{P}(t-1|t-1) + B(t)[\bar{P}(t|T) - \bar{P}(t|t-1)]B(t)' \qquad \textbf{1.B.2}$$

where $B(t) = \bar{P}(t-1|t-1)f(t)'\bar{P}^{-1}(t|t-1)$ for $t = T \ldots 1$ and $f(t) = \begin{bmatrix} F_t & 0 \\ 0 & I_{k\times k} \end{bmatrix}$ with k representing the number of the time-varying parameters. Since the smoother is based on more information than the filtered estimator, it will have an MSE that is smaller than that of the filtered estimator.

2. Time-varying parameter estimates under alternative models



2. Time-varying parameter estimates under alternative models (continued)

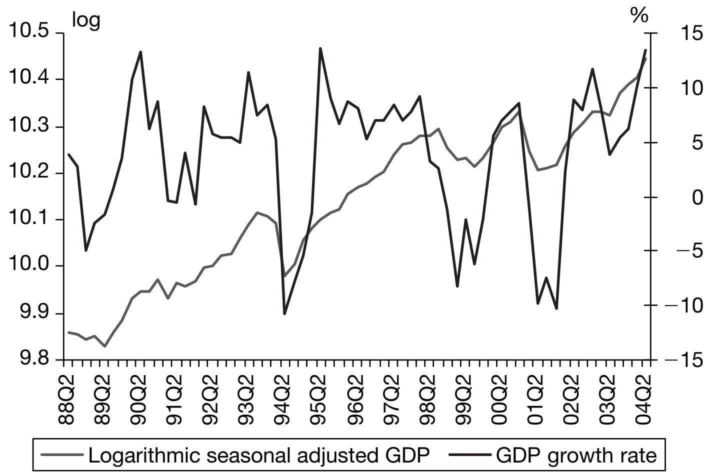


3. Data description

Sample period covers quarterly data between 1988:Q2 and 2004:Q2. The issue of seasonality is handled with the commonly used program named TRAMO/SEATS (Gomez and Maravall, 1998).

y_t : Logarithmic seasonally adjusted gross domestic product at 1987 constant prices
π_t : Log difference of quarterly seasonally adjusted consumer price index (1994=100)
$reer_t$: Log difference of CPI-based real effective exchange rate, (1995=100)
r_t : Ex-post real interest rate on 3-month average discounted Treasury action rates
DI_t : Demand index constructed from the Business Tendency Survey (BTS) of CBRT. Components of the index: BTS question 9 (total amounts of orders received this month), question 16 ((trend of next three months) volume of goods sold in domestic market), and question 18 ((Trend of next three months) volume of raw-material stocks).

Figure – Actual GDP growth and seasonal adjusted GDP



Using additional information in estimating the output gap in Peru: a multivariate unobserved component approach[1]

Gonzalo Llosa and Shirley Miller (Central Reserve Bank of Peru)

1. Motivation

One of the key elements for the implementation of Inflation Targeting regime is the right identification of inflationary or disinflationary pressures. It is important to have a reliable indicator of these pressures because the central bank will use it for guiding its monetary policy to achieve its inflation target. The central bank will engage on tight (expansive) policy whenever the indicator signs inflationary (disinflationary) pressures that risks achieving its target.

In general, the indicator used is the output gap. This variable tries to measure the short run pressures of marginal costs over inflation generated by a demand expansion and an inaccurate distribution of the productive factors of the economy. Unfortunately, the output gap is an unobservable variable and its value must be inferred from the information contained in other economic variables. To this respect, the estimation of the output gap has been the focus of considerable research effort of many central banks.[2]

The most common techniques are based on univariate filters, which only use gross domestic product (GDP) information.[3] These methodologies assume that output is an isolate process from the rest of macroeconomic time series. In most of the cases, this simplicity implies a high degree of uncertainty in the output gap measure, specially at the end of the sample.[4] Moreover, in the cases that other relevant variables have affected output gap, these univariate approaches do not allow to identify them,[5] thus disturbing the decisions of monetary policy.[6]

As an alternative, different multivariate methods have been developed, each one is based on a particular theory and implementation technique. One of the most common multivariate methods is the Production Function approach, which consists on a neoclassical production function with different inputs, generally capital stock, labor force and total factor productivity (Solow residual). Often, researchers attempting to apply this technique use an univariate method to estimate the trend of productivity.[7] As a consequence, uncertainty remains on this component affecting the output gap reliability.

1 *This paper was published on Money Affairs Journal, Volume XVII, N° 1, January-June, 2004, Center for Latin American Monetary Studies. The authors are thankful to José Dorich, Hugo Perea, Vicente Tuesta, Marco Vega and Diego Winkelried for useful comments to this paper. We are particularly grateful to Jaromir Benes and David Vavra (Czech Republic National Bank) for advice in Kalman filter technique. We also benefited from comments by participants in seminars at Central Reserve Bank of Peru, VII Meeting of the Network of American Central Bank Researchers (Venezuela, Nov. 2003), XXI Meeting of Economist (Peru, Feb. 2004) and IFC Conference on Central Bank Issues Regarding National and Financial Accounts (Switzerland, Sep. 2004). The views expressed in this paper are those of the authors and do not reflect those of the Central Reserve Bank of Peru.*

2 *See for example, Benes and N'Diaye (2002) and Butler (1996).*

3 *See for example, Hodrick and Prescott (1997), Beveridge and Nelson (1981), Baxter and King (1995) and Harvey and Jaeger (1993).*

4 *Several studies have addressed this problem in univariate methods. For example, Orphanides and van Norden (1999) studied uncertainty in US output gap estimation process and Gruen et al. (2002) do the same for Australian GDP. Their results confirm that end of problem is the principal source of uncetainty affecting output gap estimation.*

5 *For example, Haltmaier (2001) uses cyclical indicators to adjust Japanese output gap estimates derived from the Hodrick and Prescott filter over the most recent period.*

6 *Smets (2002) and Gaudich and Hunt (2000) found that the bigger the uncertainty surrounding output gap estimates, the smaller the reaction of monetary policy to it.*

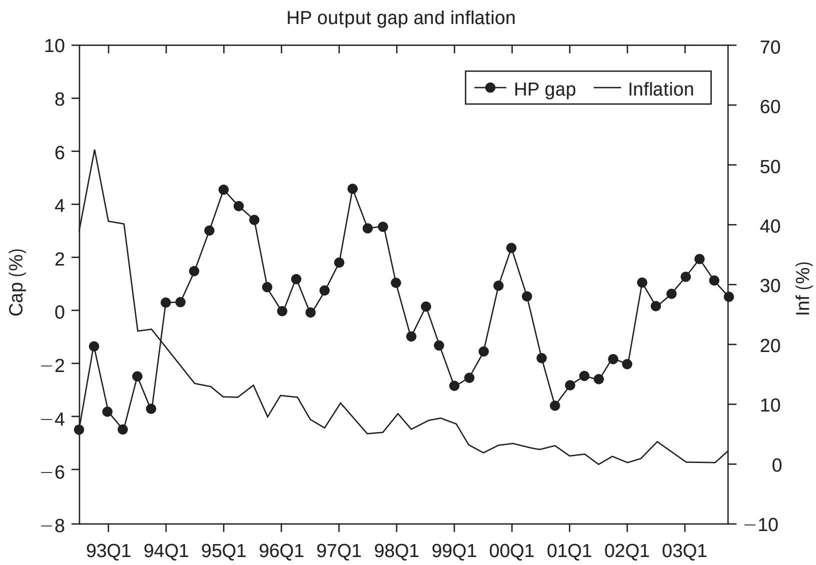
7 *See for example, Miller (2004a) and Texeira (2002).*

Another way to impose structural restrictions is using the SVAR identification of Blanchard and Quah (1989). The SVAR output gap is the component not affected by permanent shocks and related to the employment rate or inflation in a transitory way. This method has several limitations, it is not accurate to identify permanent and transitory shocks and its performance could be undermined by omitted variable problems.[8]

More recently, a new group of multivariate methods use unobserved component models, which combine structural relationships with properties of statistical filters. Their main characteristic is that they include an explicit relation between output gap and inflation (Phillips Curve), and/or between the output gap and the unemployment rate (Okun's law). Several authors have used multivariate techniques based on unobserved component models, whose estimation is carried out via the Kalman filter algorithm.[9] This approach benefits from correlation in the data and model structure, mixing this information according to the lowest prediction error. This technique has been successfully applied, increasing the accuracy and reliability of output gap estimations.[10]

In order to show the limitations of univariate methods in figure (1) we plot the annual variation of the core Consumer Price Index and the output gap, estimated with the univariate Hodrick-Prescott (HP) filter for the quarterly period 1992–2003. The inflation process in Peru presents two episodes. The first one (1992–1994) is characterized by a continuous disinflation process from high (more than 80 percent during 1992) to moderate inflation rates (around 20 percent in 1994). In the second episode (1995–2003), the inflation rate continues decreasing, but at a lower pace, moving from moderate (around 11 percent in 1995) to low inflation rates (one-digit inflation in 1997 and lower than 5 percent since 1999).

Figure 1 – Output gap corresponds to Hodrick and Prescott estimate with smoothing parameter of 1600. The inflation rate is calculated on quarterly base (annualized)



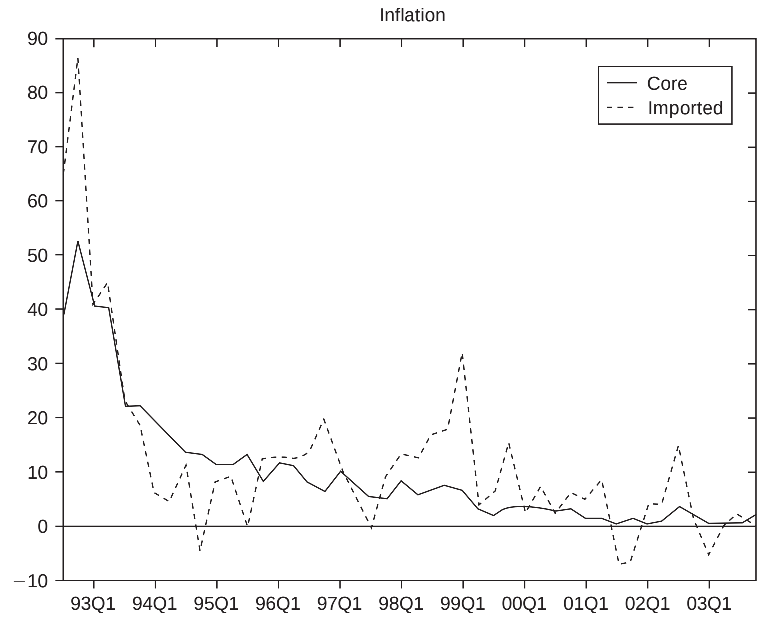
On the other hand, the HP output gap during the first episode (specifically in 1994–1996) indicates high excess demand conditions, which implies the presence of strong inflationary pressures. Nevertheless, this result does not seem to be in line with the persistent decline in inflation along the nineties. Another similar episode to highlight is observed at the end of the sample, where inflation is relatively stable but the HP output gap is positive, indicating inflationary pressures. In this context, the results obtained with the HP filter do not permit to analyze and explain correctly the evolution of the inflation, particularly during periods where output was growing significantly and inflation was falling or stable. This univariate technique only captures

8 *See for a technical details van Norden (1995), Blanchard and Quah (1989) and Cerra and Chaman (2000).*
9 *For example, see De Brouwer (1998), Scott (2000b), Camba-Mendez and Rodriguez-Palenzuela and Benes (2001) and N'Diaye (2002). Alternatively, Laxton and Tetlow (1992) and Hirose and Kamada (2001) propose a multivariate Hodrick and Prescott filter.*
10 *See for example, Rünstler (2002). Although, it worths to say that the improvement depends on the structure imposed, and calibration or parameters estimation, see Butler (1996).*

the output process, without taking into account any structure or the dynamics process of other important macroeconomic variables.

Given that Peruvian economy is a small open economy, many other variables (for example, imported inflation) are critical in understanding inflation dynamics. Figure (2), plots the Core CPI inflation and imported inflation rates.[11] Core inflation has been evolving together with imported inflation except in two remarkable cases: 1994–1996 and 1998–1999. During the former period, inflation is higher than imported inflation suggesting that some inflationary pressures might have restrained the total pass-through. The opposite happens in the second case: imported inflation is higher than core inflation, and this coincides with a weak output phase. This analysis suggests that imported inflation is a key variable that have to be considered in the determination of the output gap.

Figure 2 – Quarterly core inflation and imported inflation. Imported inflation is computed by PPP condition



For the Peruvian case, most of the studies had based on univariate filters.[12] Given the advantages of multivariate unobserved component models, the aim of this paper is to provide an estimation of the output gap using this technique. The model employed relies on an explicit short run relation between the output gap and inflation rate (Phillips Curve) and structural restrictions over output dynamics. We estimate the model via Kalman Filter for the period 1992–2003.

The results show that the multivariate unobserved component output gap (MUC) is less sensible to end of sample problems and presents a better relation with the Peruvian inflation process than other estimates, calculated with the Hodrick-Prescott filter and the production function approach. In particular, in periods of high output growth together with disinflationary or stable inflation environments, MUC output gap is lower than the ones obtained with the alternative methods mentioned. Besides, MUC identification is quite related to pass-through effect from imported prices to consumer prices. In particular, whenever imported inflation was higher (lower) than domestic inflation, the system found a negative (positive) output gap. Furthermore, the diagnostic statistics report that MUC estimate is more reliable than other alternatives and increases out of sample predictive power for inflation.

The document is organized in the following form. In the second section, the structure of the model used, as well as its implementation and the data, are explained and analyzed. In next section, we present the most important features of MUC estimate, and some of its properties: updating properties and inflation forecasts power. Finally, in the fourth section, we conclude.

11 *Imported inflation is calculated from US inflation and nominal exchange rate depreciation (appreciation).*

12 *For example, Cabredo and Valdivia (1999), Caballero and Gallegos (2001) and Miller (2004a) compare different output gap estimates using those techniques. Their results indicate that production function output gap is the best indicator of inflation pressures in Peru.*

2. The model

We use a semi-structural model for a small open economy. The system is based on three behavioral equations:

1. Uncovered interest parity.
2. Phillips Curve.
3. Aggregate demand.

The uncovered interest parity allows us to estimate the permanent and transitory components of real interest rate and real exchange rate. Combining the gaps of real interest rate and real exchange rate, we construct a real monetary condition index.[13] Taking this index as an exogenous variable, we use the aggregate demand equation and Phillips curve to calculate the output gap related to the evolution of real activity and inflation. The model takes the following form,

$$y_t = \bar{y}_t + \hat{y}_t \qquad (1)$$

$$B(L)\,\hat{y}_t = \kappa RMCI_t + \eta_t^y \qquad \eta_t^y \sim N(0, \sigma^2_{\eta y}) \qquad (2)$$

$$\pi_t = \tilde{\pi}_t + \varepsilon_t^\pi \qquad \varepsilon_t^\pi \sim N(0, \sigma^2_{\varepsilon\pi}) \qquad (3)$$

$$\tilde{\pi}_t = \alpha_1 \tilde{\pi}_{t-1} + \alpha_2 \tilde{\pi}_t^m + (1 - \alpha_1 - \alpha_2)\tilde{\pi}_{t,t+1}^e + \gamma \hat{y}_t + \eta^\pi \qquad (4)$$

$$\eta_t^\pi \sim N(0, \sigma^2_{\eta\pi})$$

From equation (1), output y_t (in logarithms) is decomposed into potential output $\bar{y}_t$ and the output gap $\hat{y}_t$. The second equation describes the output gap dynamics influenced by the real monetary condition index, $RMCI_t$. The lag polynomial is defined by $B(L) = 1 - \beta$, which represent an AR(1) stationary process. Equation (3) decomposes the CPI core inflation, π_t, into its forecastable component,[14] $\hat{\pi}_t$, and an stochastic shock η_t^π. The underlying inflation is modeled using a Phillips curve for a small open economy, equation (4). According to this equation, this measure is influenced by its own inertia, imported inflation $\tilde{\pi}_t^m$, inflation expectations $\tilde{\pi}_{t,t+1}^e$ and the output gap $\hat{y}_t$.

Potential output $\bar{y}_t$ follows a random walk process with a stochastic slope μ_t. The slope is modeled as an stationary autoregressive process with constant, $\bar{\mu}$, reflecting the growth rate of potential output in steady state.[15]

$$\bar{y}_t = \bar{y}_{t-1} + \mu_t \qquad (5)$$

$$\mu_t = \phi \mu_{t-1} + (1 - \phi)\bar{\mu} + \eta_t^u \qquad \eta_t^u \sim N(0, \sigma^2_{\eta\mu}) \qquad (6)$$

The model is completed by the assumption that stochastic shocks $\varepsilon_t^\pi, \eta_t^\pi, \eta_t^y$, and η_t^u are normally and independently distributed and mutually uncorrelated.

2.1. Inflation expectations

One important issue is the measurement of inflation expectations. Typically, the New Keynesian Phillips curve stresses on forward looking behavior in the price setting process.[16]

13 *This index captures the general orientation of monetary policy affecting the aggregate demand with the aim to control the inflation rate, see Dennis (1997) for technical discussion. Given the trends in the data, the index is constructed using gaps instead of levels. The details are shown in appendix A.*

14 *Forescastable inflation may be interpreted as a measure of underlying or trend inflation, which is filtered from high frequency fluctuations. Arguably, a central bank should be responsible primarily for development in underlying inflation and not for high frequency inflation, which is unable to control.*

15 *A local linear trend model for potential output was proved. The results indicate that a steady state growth rate of potential output reduces end of sample revisions. For a technical discussion of local level and local linear trend models see Harvey (1993).*

16 *See Calvo (1983) and Clarida et al. (2002)*

Particularly, the forward-looking component on inflation is quite important during disinflation episodes.[17]

On the other hand, empirical work usually assumes totally backward looking expectations,[18] implying that the output gap has permanent effects on the level of inflation rate.[19]
Confronting this trade off between theoretical and empirical grounds, we consider a simple error correction mechanism for inflation expectations which allows us to incorporate the deceleration on Peruvian inflation without assuming totally backward looking expectations,[20]

$$\tilde{\pi}^e_{t,t+1} = \overline{\pi}_{t+1} + (\tilde{\pi}_t - \overline{\pi}_t) = \tilde{\pi}_t + \Delta\overline{\pi}_{t+1} \tag{7}$$

where $\tilde{\pi}_t$ is underlying inflation, $\tilde{\pi}^e_{t,t+1}$ represent inflation expectations over next quarter, and $\overline{\pi}_t$ is interpreted as inflation target rate.[21] Given (7), if underlying inflation is higher (lower) than the target, inflation expectations raises (decreases). If inflation is aligned to the target, expectations do not change. Considering this structure, replace (4) with,[22]

$$\tilde{\pi}_t = \frac{\alpha_1}{\alpha_1+\alpha_2}\tilde{\pi}_{t-1} + \frac{\alpha_2}{\alpha_1+\alpha_2}\tilde{\pi}^m_t + \frac{(1-\alpha_1-\alpha_2)}{\alpha_1+\alpha_2}\Delta\overline{\pi}_{t+1} + \frac{\gamma}{\alpha_1+\alpha_2}\hat{y}_t + \eta^\pi_t \tag{8}$$

We assume that in the long run, real exchange rate depreciation is zero and inflation target is constant. Thus, there is not a relationship between output gap and the inflation in the steady state.[23]

2.2. The state space form

For estimation, the model must be put in its state space form, which comprises two equations.[24] Measurement equation (9) relates observations x_t at time t, $t = 1, \ldots, T$, to the unobserved state vector α_t.[25] Transition equation (10) denotes the stochastic dynamic behavior governing the state vector.

$$x_t = Z\alpha_t + \varepsilon_t \tag{9}$$

$$A_0\alpha_t = c + A_1\alpha_{t-1} + B\varphi_t + R_0\eta_t \tag{10}$$

where:

$x_t = \begin{bmatrix}\pi_t & \Delta y_t\end{bmatrix}'$ is the observable vector,

$\alpha_t = \begin{bmatrix}\tilde{\pi}_t & \hat{y}_t & \hat{y}_{t-1} & \mu_t\end{bmatrix}'$ is the state vector,

$\varphi_t = \begin{bmatrix}\pi^m_{t-1} & \Delta\overline{\pi}_{t+1} & RMCI_t\end{bmatrix}'$ is the exogenous vector,

$\varepsilon_t = \begin{bmatrix}\varepsilon^\pi_t & 0\end{bmatrix}'$ and $\eta_t = \begin{bmatrix}\eta^\pi_t & \eta^y_t & 0 & \eta^\mu_t\end{bmatrix}'$ are innovation vectors

Matrices are given by:

$$Z = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & -1 & 1\end{bmatrix},$$

17 *See Mankiw and Reis (2001).*
18 *See for example, Rünstler (2002).*
19 *Also, this structure implies that the level of inflation rate is unforecastable, which is a clear contradiction to inflation targeting policy framework.*
20 *A more suitable technique could be related to adaptative learning expectations, see Evans and Honkapohja (2001).*
21 *Formally, inflation targeting was adopted as a monetary policy framework in Peru in 2002. Nevertheless, the Central Reserve Bank of Peru has been announcing inflation targets since 1994. See Rossini (2001) for details.*
22 *Although the error term has changed, we maintain its nomenclature.*
23 *This also implies that in the long run inflation rate is constant and output gap is zero.*
24 *Because output has a unit root, the model was stationarized differencing output equation. Differentiation permits direct calculation of Kalman Filter initial conditions from the model structure and data, without applying diffuse priors for initial conditions, which modifies severely the results at the beginning of the sample.*
25 *We present an structural version of state equation, which incorporates contemporaneous effects between underlying inflation and output gap. To get the autoregressive form, invert the left matrix of the state system.*

$$A_0 = \begin{bmatrix} 1 & \frac{-\gamma}{\alpha_1+\alpha_2} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$c = \begin{bmatrix} 0 \\ 0 \\ 0 \\ (1-\phi)\bar{\mu} \end{bmatrix},$$

$$A_1 = \begin{bmatrix} \frac{\alpha_1}{\alpha_1+\alpha_2} & 0 & 0 & 0 \\ 0 & \beta & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \phi \end{bmatrix},$$

$$B = \begin{bmatrix} \frac{\alpha_2}{\alpha_1+\alpha_2} & \frac{1-\alpha_1-\alpha_2}{\alpha_1+\alpha_2} & 0 \\ 0 & 0 & \kappa \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$R_0 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Innovations η_t and ε_t are mutually uncorrelated and have diagonal covariance matrices. Both are modeled as multivariate gaussian distributions. Matrices A_0, A_1, B, R_0 and vector c depend on unknown hyperparameters.[26] After fixing hyperparameters, prediction, updating and smoothing algorithms are applied.

To get the usual state space representation, take into account the following equalities,

$T = A_0^{-1}A_1$, is the transition matrix,
$d = c + A_0^{-1}B\varphi_t$, summarizes exogenous variables,
$R = A_0^{-1}R_0$.

2.3. Calibration

The model (1)–(6) incorporates several hyperparameters, coefficients $\{\beta, \kappa, \alpha_1, \alpha_2, \gamma, \phi, \bar{\mu}\}$ and variances $\{\sigma^2_{\varepsilon\pi}, \sigma^2_{\eta\pi}, \sigma^2_{\eta y}, \sigma^2_{\eta\mu}\}$. This hyperparameters can be estimated using maximum likelihood procedure. However there are several issues with this approximation. First, for the sample selected, the inflation rate shows a persistent dynamics can be explained by a non-stationary homogenous component in the stochastic dynamic equation of inflation.[27] Second, the quarterly sample used is too short to permit a reliable econometric estimation. Third, we suspect that structural breaks, due to institutional changes and structural reforms in Peru, could prevent a suitable econometric identification.[28]

26 *The next section focus on the criteria utilized for hyperparameters calibration.*
27 *This phenomenon may invalidate any econometric estimation. For a technical discussion, see Enders (1995) chapter 1.*
28 *There exist some evidence about structural breaks in Peruvian data, see Quispe (1999). In general, structural breaks could distort inflation – output relationship, see Clark and McCracken (2003).*

As an alternative, we choose to calibrate the model using external information. In order to get priors, a Phillips curve and a aggregate demand function, similar to equations (2) and (8), were estimated econometrically using Hodrick and Prescott output gap.[29] The table 1 reports the selected values.

Table 1 – Model parameters

Parameter	Calibrated value
Inflation persistence (α_1)	0.70
Pass through (α_2)	0.15
Output gap to inflation (γ)	0.70
Output gap persistence (β)	0.70
RMCI to output gap (κ)	0.10
Potential slope persistence (ϕ)	0.80
Potential quarterly growth ($\bar{\mu}$)	0.01

In the Phillips curve, we calibrate the parameter α_1 in 0.7. The inflation elasticity to output gap (γ) is calibrated in 0.7. This value is higher than the ones found for other countries.[30] However, it reflects the low sacrifice ratio during the disinflation process in the last ten years.[31] Additionally, we set the pass-through effect from imported inflation over CPI core inflation captured by α_2 in 0.15, according to those found by Miller (2004b) and Winkelried (2004).

For the output gap equation we use the econometric estimation to set the inertia parameter β in 0.7, the effect of the real monetary condition index κ in 0.1 and the value of ϕ in 0.8. The steady state growth rate of potential output was fixed in 4 percent (annualized), according to the mean growth rate of potential output calculated using the production function approach.[32]

All variances, except that for the growth rate of potential output, were normalized. For filtering process, we have to identify the permanent and transitory components of output. The signal extraction problem is basically related to the variance ratio between growth rate of potential output and output gap,[33] $\sigma^2_{\eta\mu}/\sigma^2_{\eta y}$. We set this value to 1/64. This smoothes potential output and increases the relation between the cyclical component of output and inflation.

2.4. The data

We use quarterly data form the Central Reserve Bank of Peru. The sample spans from 1992 to 2003. We utilized the real GDP calculated using 1994 prices. Inflation is represented by core CPI inflation and nominal exchange rate by soles/US$ parity. As an international interest rate we use monthly LIBOR rate. External inflation is approximated by United States CPI inflation rate. Imported inflation is constructed using PPP condition: $\pi_t^m = \pi_t^{US} + \Delta e_t$, where π_t^{US} is US CPI inflation and Δe_t is the exchange rate depreciation (appreciation).

The real exchange rate is measured by the imported prices index deflated by core consumer prices index. On the other hand, the ex-post real interest rate is measured as: $r_t = i_t - \pi_t^{core}$, where i_t is the annualized interbank interest rate and π_t^{core} is year-to-year core inflation rate. Real monetary condition index is constructed with real interest rate and real exchange rate gaps. The risk premium is calculated as the uncovered interest parity condition residual.

Finally, the inflation target rate is the HP filtered of core inflation, restricted to the last announced target (2.5 percent) as a final level prior since 2002.[34]

29 *We took these results with caution because the presence of persistent dynamics on inflation can invalidate statistical inference and also the use of an incorrect output gap measure distort coefficient values. Nevertheless, the estimations give useful information about the parameter values and their uncertainty. In particular, we found that {γ,κ} are blurred by tremendous uncertainty.*

30 *For Germany 0.40 and United States 0.44, see Ball (1994), and Czech Republic 0.22, see Benes et al. (2002).*

31 *See Zegarra (2000).*

32 *See Miller (2004a).*

33 *Signal extraction problem is practically intractable without imposing some ad-hoc restrictions, see Quah (1992) for a technical discussion. For example, the direct estimation of variance ratio between transitory and permanent component of a time series tend to differ to those recommended by Hodrick and Prescott, see for example, Blith et al. (2001).*

34 *For a discussion of prior's inclusion on Hodrick and Prescott filter, see St. Amant y van Norden (1997).*

3. Results

In this section, first we describe the MUC output gap, comparing it to HP filtered and the production function estimates. Then, properties of revisions in the output gap estimates and inflation forecast performance are discussed. Finally, we evaluate which output gap estimate improves inflation predictability. The results indicate that the MUC estimate shows more relation with the Peruvian inflation process, reduces end of sample uncertainty and improves inflation forecast.

3.1. Output gap estimates

Panel (a) of figure (3) plots the MUC output gap estimate using the multivariate unobserved component approach, based on the model defined by (1)–(6) and (8).

According to the results, Peruvian output gap has fluctuated inside the range of −7 to 2 percent. Four periods of inflationary pressures can be identified: 1994Q2–1995Q4, 1997Q1–1997Q4, 1999Q4–2000Q2, and more recently 2002Q2–2002Q4. The first two periods have been the most outstanding and the longest, reaching levels near 2 percent. With regard to the disinflationary pressures' episodes, they have been longer and have presented a higher average magnitude than inflationary ones. Four periods have been also identified: 1992Q3–1994Q2, 1996Q1–1996Q4, 1998Q2–1999Q4, and 2000Q4–2002Q1, being the first one the most significant, reaching values near to −7 percent.

Panel (b) of figure (3) plots quarterly underlying inflation and imported inflation. Both series show a high correlation during the ninety's. However, this relation breaks in two remarkable periods: 1994–1995 and 1998–1999. In the first one, underlying inflation is higher than imported inflation. At the same time, a positive output gap is identified, explaining the incomplete pass-through. The opposite happens in the second period: underlying inflation is lower than imported inflation, phenomenon accompanied with a negative output gap.[35]

Figure 3 – Output gap correspond to MUC smoother estimate. Underlying inflation is computed using semi-structural model and imported inflation is computed by PPP condition

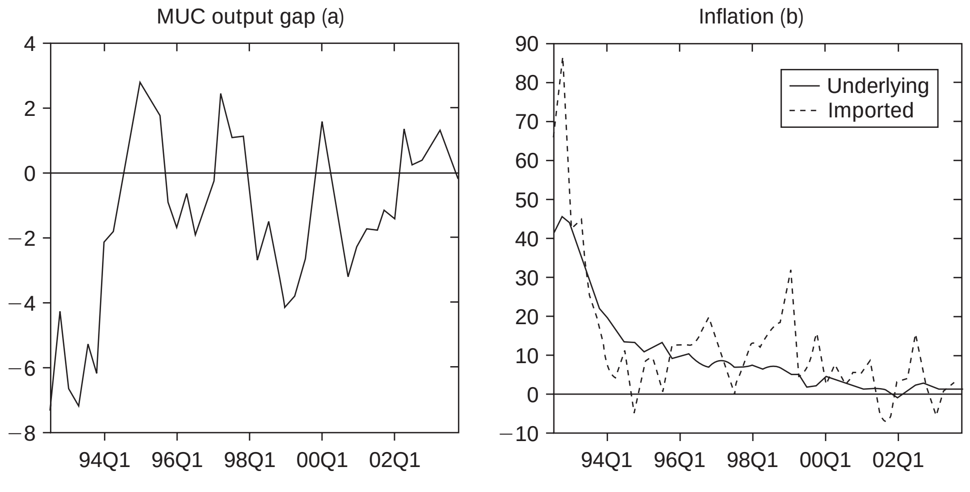
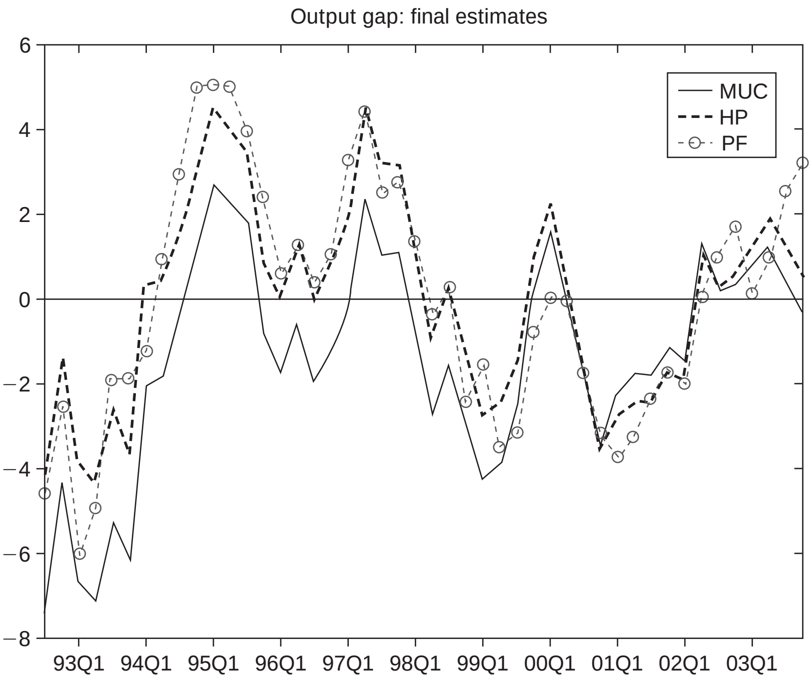


Figure (4) displays the MUC output gap estimate alongside to the Hodrick-Prescott filtered (HP) and the production function (PF) estimates.

MUC, HP and PF output gap estimates are very similar for the entire sample. However, our estimate is lower than the alternatives in two periods: 1994–1997 and at the end of the sample (2003). The most remarkable feature about those episodes is that they combine high output growth rates with a disinflationary process (1994–1997) or stable inflation environment (2003). On one hand, HP and PF methods tends to link the output gap evolution with the economic cycle, even when this cycle had not affected the inflation rate. On the other hand, MUC estimate is influenced not only by output behavior but also by domestic and imported inflation.

35 *This kind of non linear pass-through has been discussed recently in Winkelried (2004).*

Figure 4 – Smoother output gap estimate



3.2. Properties of revisions: end of sample problem

In this section, we analyze the updating properties of MUC, HP and PF estimates. Measuring output gap revisions due to additional observations is one way of evaluating the uncertainty surrounding different methods.[36] In fact, the higher uncertainty around output gap estimate, the lower is the sensitivity of monetary authority reactions to it.[37]

In order to compare the reliability of each method we calculated the real-time estimates and the final estimates of the output gap.[38] The uncertainty that each method introduce in the output gap estimation is determined by the comparison between the final estimate and the real time estimate.

The results of this exercise are shown in the figure (5). The left panel plots the real-time and final estimates calculated with MUC, HP and PF methods. It is evident that revisions of the MUC estimate of the real-time output gap in response to new data are much smaller than those of HP and PF. In the right panel, the scatter graphs between real-time and final estimates are presented. Those graphs allow to have a clearer picture of the uncertainty degree surrounding each method. The graphs are divided in 4 areas: Areas I and III show the points where the final and real-time estimates provide contradictory signals, while areas II and IV present those occasions in which both estimations give similar signals. The results indicate that MUC estimates are grouped around the 45 line (areas II and IV), while HP and PF provide contradictory signals (areas I and II).

With the aim of quantifying the uncertainty degree, we calculated the correlation coefficients and concordance indices.[39] Additionally, we test the reliability of output gap estimates using the Pesaran and Timmermann (1992) test.[40] The results are summarized in table 2. The correlation coefficients indicate that the output gaps calculated (final and real-time) with MUC

36 *See Gruen etal. (2002) and Orphanides and van Norden (1999).*

37 *See Gaudich and Hunt (2000) and Smets (2002).*

38 *The real time estimates correspond to updated state estimate in Kalman Filter recursion, conditional to past and current observation of the published data available. This means that we are not considering the ex-post revisions of published data, see for example Orphanides and van Norden (1999). On the other hand, final estimates are equivalent to smoother estimates in Kalman Filter, reflecting all available information to forecast sequentially observable variables. To get updated estimates of Hodrick and Prescott filtered we use its state space representation, see Scott (2000b).*

39 *The concordance index is simply a non parametric statistic method that measures the time proportion in which two time series are in the same state. Thus, the degree of concordance will be 1 if both output gap measures have the same sign for a determined period. By contrast, it will take a zero value if the sign of both measures (final and real-time) are always opposite. For more details of this indicator see McDermott and Scott (1999).*

40 *Quoted in Camba-Mendez and Rodriguez-Palenzuela (2001).*

present higher co-movements (0.65) than those obtained with HP (0.26) and PF (0.16). In the same way, the concordance statistic indicates that real-time and final estimates with MUC provides similar signals (0.73), better than HP (0.63) and PF (0.52) do. Moreover, the application of the Pesaran and Timmermann test shows that the acceptance probability of similar signals in the case of MUC is 70.89 percent, in contrast with the 0.01 percent and 0.00 percent of HP and FP, respectively. Those results suggest that the multivariate approach provides more reliable estimates.[41]

Figure 5 – Final estimates correspond to smoother estimates.
Real-Time estimates correspond to updated estimates.

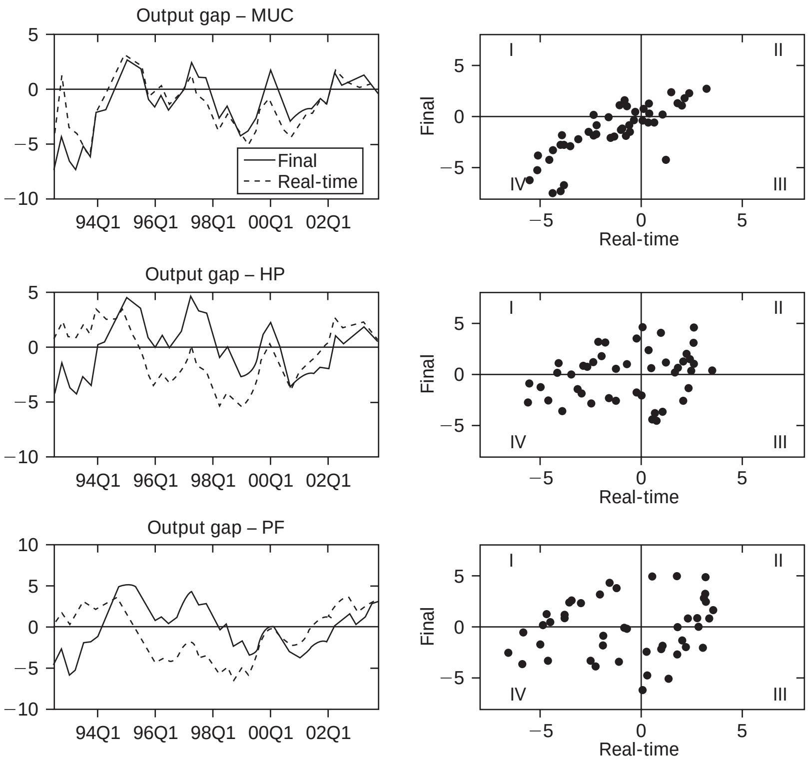


Table 2 – Evaluation statistics

	MUC	HP	PF
Correlation coefficient	0.65	0.26	0.16
Concordance index	0.73	0.63	0.52
Pesaran and Timmerman test*	70.89%	0.01%	0.00%

**Acceptance probability of the null hypothesis that the sign of real-time and final estimates is the same.*

What explains these results? Because future data always contains relevant information to the current decomposition of transitory and permanent shocks, the most recent estimates of output gap will invariably change as the persistence characteristics of past shocks become more apparent. With structural restrictions, the MUC approach exploits the correlation in the data, guiding the output gap estimation at the end of the sample.

41 *A variety of studies evaluates the reliability of different univariate and multivariate methods, comparing the real-time estimates relative to final ones. Butler (1996), Conway and Hunt (1997), Camba-Mendez and Rodriguez-Palenzuela (2001), De Brouwer (1998) and Scott (2000a) compare the updating properties of different output gaps measures for Canada, European Union, United States, Australia and New Zealand. Their results suggest that multivariate output gaps estimates are statistically more reliable. Rünstler (2002) concentrated only on unobservable components methods, univariate and multivariate. As well as in the preceding cases, his study indicates that bigger information levels increase the confidence of the output gap estimate in European Union.*

3.3. Inflation forecast

The predictive power of the output gap for inflation through the short run supply curve (Phillips curve) is an essential precondition for the economic validity of any output gap estimate. This section test the information content of different real-time output gap estimates as a leading indicator for future inflation change. For this purpose, we analyze the following regression,

$$\Delta\tilde{\pi}_t = \theta\hat{y}_{t|t} + \sum_{i=1}^{k}\Delta\tilde{\pi}_{t-i} + \varepsilon_t \qquad (11)$$

where $\hat{y}_{t|t}$ is the real time output gap estimate and $\tilde{\pi}_t$ is the underlying inflation measure calculated from Kalman Filter recursion.[42]

We apply this equation on real-time output gap estimates computed with different methods: MUC, HP and PF. Additionally an ARIMA regression is estimated, which is taken as a benchmark. In all cases, order lags k is found from Akaike criterion minimization.

We evaluate out-sample performance using the following steps. First, the equation (11) and ARIMA equation are estimated for the sample selected. Second, the out of sample forecasts of inflation changes over the next four quarters are computed. Third, another observation to the sample is added and the first two steps are applied. We start this procedure with a sample from 1992Q1 to 1997Q4, expanding it until 2002Q2.

Table 3 reports the mean square error of forecast for underlying inflation using different output gap estimates in relation to the benchmark equation.

Table 3 – Out-sample forecasts performance

MSE-ratio (in percentages)	Forecast horizon			
	+1	+2	+3	+4
MUC	57	55	65	67
HP	66	73	77	84
PF	75	77	78	85

MSE-ratio denotes the mean square error of the inflation forecast relative to MSE of the random walk forecast. MSE's of the out-of-sample random walk forecast are given 1.56, 1.61, 1.76, and 1.29 for 1, 2, 3, 4 quarters ahead, respectively. Initial sample: 1992Q1–1997Q4.

Out-of-sample forecasts using output gap estimates improve substantially on the random walk forecast at all the four quarters ahead. However, the improvement varies across different methodologies considered. MUC output gap increases inflation predictability more than HP and PF gaps estimates do. Additionally, the forecast performance of the MUC estimate is nearly stable as the forecast horizon increases. In that sense, HP and PF perform worse, showing MSE-ratios increments as the forecast horizon is expanded. At best, the four quarter forecast of HP and PF improve only slightly relative to the random walk model.

4. Final remarks

With the objective of improving output gap measurement in Peru, we develop a semi-structural model for a small open economy. The model was estimated as a multivariate unobserved component model using the Kalman Filter technique. The system incorporates explicitly a short run relation between output gap and inflation process through a Phillips Curve and also adds some other structural restrictions over potential output dynamics. The model parameters were calibrated using external information sources. Our results indicate that the MUC output gap estimate outperforms alternatives such as the HP filter or PF estimates.

The results indicate that the MUC output gap is quite similar to alternatives measures. However, in periods of high output growth rate together with a disinflation or stable inflation

42 *The use of inflation changes eliminates the excessive persistence on inflation. Econometrically, this approach is optimal since improves the short run forecast.*

context, our estimate indicate lower demand pressures than other estimates do. In particular, at the end of the sample (characterized by an environment of stable inflation), HP and FP are biased toward excess demand conditions. Besides, MUC output gap identification is quite related to pass-through effect from imported prices to consumer prices. In particular, whenever imported inflation was higher (lower) than domestic inflation, the system found a negative (positive) output gap.

Furthermore, we studied updating properties comparing the smoother estimates and the updated estimates of the three competitive approaches. The diagnostic statistics report that MUC estimate is the most reliable of the group. Finally we explore the out-sample predictive power for inflation of different output gap estimates. The results indicate that the MUC estimates forecast better inflation changes, confirming the essential precondition for the economic validity of any output gap estimate.

The advantages above-mentioned prove the importance of adding structural information on output gap calculation. For monetary policy purposes, this outcome could imply a significant uncertainty reduction and could improve future inflation control. Given that, a future research agenda could be oriented to explore additional cyclical indicators to improve output decomposition, in that sense, we recommend Rünstler (2002). Further, as the model presented here was calibrated, uncertainty involved in this process must be quantified. Regarding to this, Bayesian analysis of posterior densities of hyperparameters as in Harvey et al. (2002) could be implemented.

References

Ball, L. (1994): "What determines the sacrifice ratio". *Monetary Policy*, Ed. Gregory Mankiw, National Bureau of Economic Research, The University of Chicago Press.

Baxter, M. and R. King (1995): "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series". NBER Working Paper No. 7872, Cambridge.

Benati, L. (2001): "Band pass filtering, cointegration, and business cycle analysis". Working Paper, Bank of England.

Benes, J., T. Hledik, D. Vavra and J. Vlcek (2002): *The Czech National Bank's Forecasting and Policy Analysis System: The Quarterly Projection Model and its Properties*. Ed. W. Coats, D. Laxton and D. Rose, Czech National Bank.

Benes, J. and P. N'Diaye (2002): *The Czech National Bank's Forecasting and Policy Analysis System: A Multivariate Filter for Measuring Potential Output and the NAIRU*. Ed. W. Coats, D. Laxton and D. Rose, Czech National Bank.

Beveridge, S. and C.R. Nelson (1981): "A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle". *Journal of Monetary Economics*, No. 7, 151–174.

Blanchard, O. and D. Quah (1989): "The dynamics effects of aggregate demand and supply disturbances". *American Economic Review*, Volume 79, Edition 4, 655–673, 1989.

Blith, C., J. Reeves, J. Small and C. Triggs (2000): "The Hodrick – Prescott filter, a generalization, and a new procedure for extracting an empirical cycle from a series". *Studies in Nonlinear Dynamics and Econometrics*, Volume 4, No. 1, MIT Press.

Butler, L. (1996): *The Bank of Canada's New Quarterly Projection Model (QPM): A semi structural method to estimate potential output: combining economic theory with a time-series filter*. Bank of Canada.

Caballero, M. and J. Gallegos (2001): "La incertidumbre sobre la brecha producto y la función de reacción bajo un esquema de inflation targeting". *Revista de Estudios Económicos*, No. 7, Central Reserve Bank of Peru.

Cabredo, P. and L.Valdivia (1999): "Estimación del PBI potencial: Perú 1950–1997". *Revista de Estudios Económicos*, No. 5, Central Reserve Bank of Peru.

Camba-Mendez, G. and D. Rodriguez-Palenzuela (2001): "Assessment criteria for output gap estimates" Working Paper Series No. 58, European Central Bank.

Calvo, G. (1983): "Staggered prices in a utility-maximizing framework". *Journal of Monetary Economics*, No. 12, 383–398.

Cerra, V. and S. Chaman (2000): "Alternative methods of estimating potential output and the output gap: an application to Sweden". IMF Working Paper, WP/00/59.

Clarida, R., J. Gali and M. Gertler (2002): "A simple framework of international monetary policy analysis". *Journal of Monetary Economics*. No. 49, 879–904.

Clark, T. and M. McCracken (2003): "The predictive content of output gap for inflation: resolving in-sample and out-sample evidence". RWP 03–06, Research Division, Federal Reserve Bank of Kansas City.

Conway, P. and B. Hunt (1997): "Estimating potential output: a semi structural approach". Discussion Paper Series, G97/9, Reserve Bank of New Zealand.

Dennis, R. (1997): "A measure of monetary conditions". Discussion Paper Series, G97/1, Reserve Bank of New Zealand.

De Brouwer, G. (1998): "Estimating output gap". Research Discussion Paper No. 9809, Reserve Bank of Australia.

Enders, W. (1995): *Applied econometrics time series*. John Wiley and Sons, INC.

Evans, G. and S. Honkapohja (2001): *Learning and expectations in macroeconomics*. Princeton University Press.

Gaudich, V. and B. Hunt (2000): "Inflation Targeting under potential output uncertainty". Discussion Paper Series, DP2000/08, Reserve Bank of New Zealand.

Gruen, D., T. Robinson and A. Stone (2002): "Output gaps in real time: are they reliable enough to use for monetary policy?". Research Discussion Paper, Bank of Australia.

Hamilton, J. (1994): *Time Series Analysis*. Princeton, N.J., Princeton University Press.

Haltmaier, J. (2001): "The use of cyclical indicators in estimating the output gap in Japan". International discussion paper, No. 701, Board of Governors of the Federal Reserve System.

Harvey, A. (1993): *Time Series Models*. Harvester Wheatsheaf, London.

Harvey, A. and A. Jaeger (1993): "Detrending, stylized facts, and the business cycle". *Journal of Applied Econometrics*, Vol. 8, 231–247.

Harvey, A., T. Trimbur and H. van Dijk (2002): "Cyclical component in economic time series: a Bayesian approach". DAE working paper No. 0302.

Hirose, K. and Y. Kamada. (2001): "A new technique for simultaneous estimation of output gap and Phillips curve". Working Paper, Bank of Japan.

Hodrick, R.J. and E.C. Prescott (1997): "Postwar US business cycles: an empirical investigation". *Journal of Money, Credit and Banking*, No. 29, 1–16.

Laxton, D. and R. Tetlow. (1992): "A Simple Multivariate Filter for Measurement of Potential Output". Technical Report No. 59, Bank of Canada.

Mankiw, G. and R. Reis (2001): "Sticky information versus sticky prices: a proposal to replace the new Keynesian Phillips Curve". NBER working paper, No. 8290.

McDermott, C. and A. Scott (1999): "Concordance in Business Cycles". Discussion Paper Series G99/7, Reserve Bank of New Zealand.

Miller, S. (2004a): "Métodos alternativos para la estimación del PBI potencial: una aplicación para el caso del Perú". *Revista de Estudios Económicos*, No. 10, Central Reserve Bank of Peru.

Miller, S. (2004b): "Estimación del pass though del tipo de cambio a precios: 1995–2002". *Revista de Estudios Económicos*, No. 10, Central Reserve Bank of Peru.

Orphanides, A. and S. van Norden (1999): "The reliability of output gap estimates in real time". Working Paper, Federal Reserve.

Quah, D. (1992): "The relative importance of permanent and transitory components: identification and some theoretical bounds". *Econometrica*, Volume 60, Edition 1, 107–118.

Quispe, Z. (1998): "Una aproximación a la demanda de los principales agregados monetarios en el Perú: junio 1991 – mayo 1997". *Revista de Estudios económicos*, No. 3, Central Reserve Bank of Peru.

Rossini, R. (2001): "Aspectos de la adopción de un régimen de metas de inflación en el Perú". *Revista de Estudios Económicos*, No. 7, Central Reserve Bank of Peru.

Rünstler, G. (2002): "The information content of real-time output gap estimates: an application to the Euro Area". Working Paper Series, No. 182, European Central Bank.

Scott, A. (2000a): "Stylized facts from output gap measures". Discussion Paper Series. DP2000/07, Reserve Bank of New Zealand.

Scott, A. (2000b): "A multivariate unobserved component model of cyclical activity". Discussion Paper Series. DP2000/04, Reserve Bank of New Zealand.

Smets, F. (2002): "Output gap uncertainty: does it matter for the Taylor rule?". *Empirical Economics*, No. 27, 113–117.

St. Amant, P. and S. van Norden (1997): "Measurement of the output gap: a discussion of recent research at Bank of Canada". Technical Report No. 79, Bank of Canada.

Texeira, T. (2002): "Estimating Brazilian potential output: a production function approach". Working Paper Series 17, Central Bank of Brasil.

van Norden, S. (1995): "Why is so hard to estimate the current output gap?". Working Paper, Bank of Canada.

Winkelried, D. (2004): "No linealidad del pass though en el Perú. Nuevas evidencias, 1993–2002". *Revista de Estudios Económicos*, No. 10, Central Reserve Bank of Peru.

Zegarra, L. (2000): "El ratio de sacrificio y los efectos reales de la política monetaria en Perú". *Revista Moneda* No. 124, Central Reserve Bank of Peru.

Abstract

One of the key elements for inflation targeting regime is the right identification of inflationary or disinflationary pressures through the output gap. In this paper we provide an estimation of the Peruvian output gap using a multivariate unobserved component (MUC) model, relying on an explicit short run relation between the output gap and inflation rate (Phillips Curve) and structural restrictions over output dynamics. The results show that the MUC output gap estimate is less sensible to end of sample problems and exhibits closer dynamics with the inflation process than the standard output gap estimates.

JEL classification: E32, E31, C51, C52.

Keywords: Output gap, Inflation, unobserved component model.

Gonzalo Llosa
e-mail: gllosa@bcrp.gob.pe

Shirley Miller
e-mail: smiller@bcrp.gob.pe

Econometric Models Unit, Central Reserve Bank of Peru

Technical appendix

Appendix A: The real monetary condition index

A real monetary condition index summarizes the main transmission channels of monetary policy: real interest rate and real exchange rate channels.[43] The index is calculated as a linear combination of real interest rate and real exchange rate gaps,

$$RMCI_t = -\theta \hat{r}_t + (1-\theta)\hat{q}_t \tag{A.1}$$

where $\hat{r}_t$ and $\hat{q}_t$ are the gaps of real interest rate and real exchange rate, respectively. The coefficient θ measures the relative importance of the real interest.[44] A positive (negative) real monetary condition index implies an expansionary (contractionary) monetary policy stance.
Real interest rate and real exchange rate gaps were computed using a Kalman filtering technique. The model is based on uncovered interest parity condition.

$$r_t - r_t^* = 4\Delta q_t + \rho_t \tag{A.2}$$

where r_t is the domestic real interest rate; r_t^* is the external real interest rate; q_t is the real exchange rate (in logarithms) and ρ_t represents the risk premium level. We can decompose every variable in the UIP equation into transitory (gap) and trend components.

$$r_t - 4\Delta q_t - \rho_t = \hat{r}_t + \bar{r}_t - 4\Delta \hat{q}_t - 4\Delta \bar{q}_t - \tilde{\rho}_t - \bar{\rho}_t$$
$$r_t - 4\Delta q_t - \rho_t = \left[\hat{r}_t - 4\Delta \hat{q}_t - \tilde{\rho}_t\right] + \left[\bar{r}_t - 4\Delta \bar{q}_t - \bar{\rho}\right]$$

Taking UIP as a cointegration relation implies that the real interest rate, the change in real exchange rate and the risk premium move together around a long run equilibrium.[45]

$$\bar{r}_t - \bar{r}_t^* = 4\Delta \bar{q}_t + \bar{\rho}_t$$

Considering the above-mentioned, we rearrange the UIP equation as,

$$x_t = \left[\hat{r}_t - 4\Delta \hat{q}_t - \hat{\rho}_t\right] + \bar{x}_t$$

where $x_t = r_t - 4\Delta q_t - \rho_t$.

We need to specify the stochastic laws of motion of the real interest rate, the real exchange rate and the risk premium. These three variables follow a local level model ($z_t = r_t, \Delta q_t, \rho_t$) [46] as,

$$z_t = \bar{z}_t - \hat{z}_t$$
$$\bar{z} = \bar{z}_{t-1} + \bar{\eta}_t \qquad \bar{\eta}_t \sim N\left(0, \sigma^2_{\bar{\eta}}\right)$$
$$\hat{z}_t = \hat{\eta}_t \qquad \hat{\eta}_t \sim N\left(0, \sigma^2_{\hat{\eta}}\right)$$

where $\bar{z}_t$ and $\hat{z}_t$ represent the permanent (trend) and transitory (gap) component, respectively. To compute the gaps we calibrate the signal extraction ratio between transitory and permanent shocks.

43 *See Dennis (1997). Typically, monetary condition indexes are calculated in level. However, to be consistent to the semi-structural model employed, we center the monetary condition index detrending its component.*

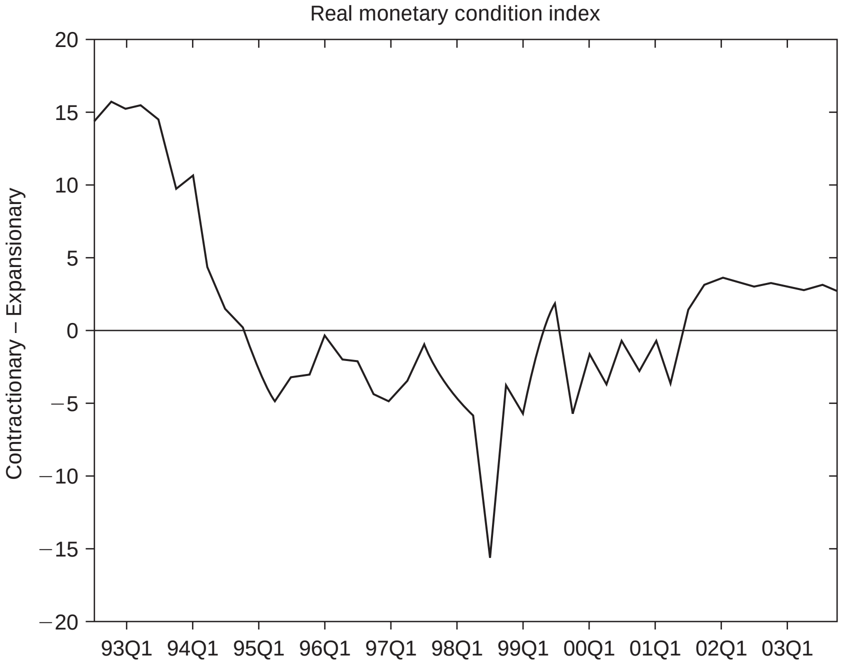
44 *A higher value of* θ *indicates that real interest rate channel is more important than real exchange rate channel. Therefore, a higher real depreciation (appreciation) is required to off set the effects of real interest rate increment (reduction).*

45 *In this long run equilibrium, external real interest rate is taking as an exogenous variable, which do not adjust to any domestic disequilibrium.*

46 *For technical discussion of local linear trend models and local level models, see Harvey (1993).*

Figure (6) plots the real monetary condition index calculated. We set $\theta = 0.9$.[47] The results indicate two stages of monetary policy stance. Expansionary during 1992Q3–1995Q4 and 2001Q1–2003Q4 and contractionary during 1996Q1–2001Q4.

Figure 6 – Real monetary conditions index is computed as a linear combination of the real interest rate and the real exchange rate gaps



47 *To calibrate this parameter we estimate equation (2) for HP output gap with the real interest rate and the change in real exchange rate in gaps separately. The results show that the real interest rate gap is more important in output gap determination than real exchange rate depreciation (appreciation) gap. Furthermore, we explore alternative specifications with primary and non-primary sectors instead of the total GDP. We found that the real interest rate gap has more effect on non-primary output gap, while the real exchange rate depreciation (appreciation) gap has more effect on primary. Given that the contribution of non-primary sector on total GDP is more than 60% (in average), the results obtained with total GDP output gap are reasonable.*

Capital stock in Indonesia: measurement and validity test[1]

Noor Yudanto, Gunawan Wicaksono, Eko Ariantoro and A. Reina Sari (Bank Indonesia)

Introduction

The main purpose of the study is to measure capital stock data for Indonesia with broader coverage by economic sector and type of capital goods. The previous studies (Keuning, 1988 and 1991; BPS-Statistics Indonesia 1995; Timmer, 1999) were generally stressed on manufacturing industrial sector only. Due to its limited coverage data availability, investment data became broadly common proxy for capital stock data necessitated on national output analysis. Bank Indonesia (BI), the central bank of Indonesia, in cooperation with BPS-Statistics Indonesia started the projects to recalculate capital stock data since 1980. As first part of continuing plans to calculate capital stock, this study is directed to present wealth capital stock. While estimating productive capital stock will be the following agenda. BI concerned of the data closely significance to the central bank plan to introduce new monetary policy framework namely inflation targeting framework (ITF). The new framework requires a number various data and information inputs for inflation and other macro indicators projection models. The framework is conditioned central bank to announce the next year inflation rate as government target to the public every end of current year. In this case, capital stock data is needed to estimate national potential output. The difference between actual output and potential output, namely output gap, is theoretically accepted as one of important information to notice the approaching inflation.

The first study of BI and BPS in 2000 was more on searching the suitable measurement method of capital stock according to finance and expertise considerations. The study main suggestion was to adopt indirect method of perpetual inventory (PIM) instead of direct method. On following study (2001), there was more an effort of BPS to define Gross Fixed Capital Formation (GFCF) comprehensively in terms of economic sectors and type of capital goods that was utilized on this study. This made a distinct difference of the study to preceding Indonesia capital stock studies. On implementing PIM to measure capital stock, the availability of qualified GFCF data and reliable assumptions are really required (Meinen, Verbiest, Paul deWolf, 1998).

In the next section of this paper, concept and definitions of capital stock are discussed and followed by detailed on type, classification, and valuation of capital goods. On subsequent section, there is elaboration on perpetual inventory method to measure capital stock. In the same section how gross-up method applying to estimate the earliest period of capital stock also discussed. Capital stock measurement results and its development are outlined on the following section. The last section assesses the validity tests of GFCF and capital stock data on several statistical tests and Cobb-Douglas production model.

Concept and definition

Capital stock

General definition of capital stock is a long term accumulation of capital goods such as buildings, machines, transportation means, livestock, and others at certain period of time. The process of economic activity to make capital goods available and usable commonly termed as investment or capital formation particularly in forms of physical investment. It should be notice that society productive activity to fulfill the needs and desires of immediate consumption is out

1 *Presented on Irving Fisher Committee (IFC) – Conference, Basel, September 9–10, 2004.*

of the definition. Additionally, capital stock could be explained as factor of production that are used repeatedly or continuously in production over several accounting periods, usually more than one year. It also could be upgraded and expanded to increase its productivity. Flows approach could be assessed when we discussed addition and deduction process of physical capital goods. Capital goods meant here are various forms of real or physical capital that can increase the efficacy of productive effort. Out of the definition is capital goods that only have intermediate role (intermediate goods) even it has over a year lifetime, for example calculator, stationary, etc.

Capital formation and capital goods

This study used Gross Fixed Capital Formation (GFCF) term for capital formation or investment as proposed by Statistics of National Account (SNA) 1968. In particular, GFCF covered total amount of physical domestic investment by excluding financial investment. In case of Indonesia, GFCF consists of the acquisition on new capital goods domestically produced, and new or second-hand capital goods imported from abroad. The method applied to estimate Indonesia GFCF is commodity flow approach. This approach utilized the data on supply of goods (fixed assets) intended for capital formation.

GFCF is a one of GDP by final consumption expenditure components. This GDP consists of goods and services used by households, government, and producers to fulfill their individual or collective needs and wants. Therefore the activity of GFCF is restricted to institutional units in their capacity as producers and can be defined as the value of their acquisitions less disposals of fixed assets. Fixed assets are produced assets that can be used repeatedly or continuously in process of production over a year generally and it has significant value as well. Additionally, GFCF data also can be derived from Input-Output (I-O) Table. In purpose to provide more comprehensive coverage of GFCF by economic sectors, this study made use of the GFCF matrices data that has been successfully constructed by BPS based on investment series data on GDP and I-O table for period of 1980–2002.

Capital goods classification according to I-O table is more detail than to GDP. Actually, there is no significant discrepancy between those two data sources. Related to this study we used both sources, so we should follow reclassification process between capital goods of GDP (four types) and capital goods of I-O Table (thirty-four types).

Type of capital goods will be used of this study is as follows:

a. Building (Construction) that covered half or full built, i.e., residential buildings, non-residential buildings (offices, factories, etc.), roads, bridges, telecommunication installation, electricity power installation, ports, gas pipe networks, etc. Any major improvement on this capital goods type is taking into account.
b. Machines, that covered all machines and its equipments used for production process on agriculture, manufacturing industry, mining, roads and bridges project, buildings project, etc.
c. Transportation means: trains, ships, automobiles, ferry, etc.
d. Livestock, which covered all that used for agriculture sector and dairy production.
e. Equipments, that covered all devices of electric, metal product, fabrics, leather, etc.
f. Others, e.g., electronic devices, music instruments, optic devices, household appliances, glass and ceramics product, sports devices, etc.

System of capital goods recording

Flows and stocks

Flows approach is recording system which reflected capital goods changing within certain period, while stocks approach when consider capital goods position on certain period.

Double entry system

The double entry system means the recording system that accommodates two sides, by producer and by user. For example locomotive, by producer side it recorded as product of manufacturing industry while by user side it recorded as capital goods of land transportation industry. These capital goods recording codes do require consistency and balance.

Time of recording

Capital goods recorded when it is legally get hold of users. Capital goods that remain on producer as stock but not sold yet are acknowledged as part of capital formation.

On this study, these three types' capital goods recording were used altogether. On gross fixed capital formation matrices construction, flows concept was used. While on capital stock matrices construction, stock concept was used.

System of valuation

Accrual base

Valuation of capital goods on this study applied accrual base (real transaction flows) as recommended by SNA 1968 instead of cash base (cash transaction flows). Therefore capital goods value is recorded as its economic values appeared rightly by ignoring its ownership.

Transaction value based on purchasing prices

Capital stock valuation could be based on current purchasing prices and book value. In order to reflect its value at the real time, current purchasing price is employed. While to reflect its value by taking into accounts its depreciation, book value is employed. Book value obtained by deducting capital goods purchasing price with its accumulation depreciation. Capital goods book value actually reflected its capacity on current and upcoming production process.

Transaction value based on current and constant prices

Capital stock valuation could be based on current price and constant price. Capital goods current value is a result of multiplication between capital goods volume and its price on current year. While, capital goods constant value is a result of multiplication between capital goods volume and its price on base year. Choosing a certain year as a base usually considered when national economy on normal condition. The common calculation to have constant value is usually by deflating current value with related price indexes.

According to this study, firstly the measurement of capital stock data will be provided for 1980–2002 period in line with the availability of GFCF data period both in gross and net concept. On the next part, capital stock data will be expanded to 1960 by applying gross-up method as we assumed that 1960 is starting period of Indonesia capital stock.

The method: perpetual inventory (PIM)

There are mainly two methods used to measure capital stock, direct and indirect. The first method utilizes primary data mostly collected directly from government or private sector asset reports and also by various special surveys. Efficiency and lack of expertise are main reasons for many countries to avoiding to implement direct method. As alternative for costly considered direct method is indirect method that depends mostly on secondary data and supporting relevant assumptions. Indirect method that used to adopt is perpetual inventory method (PIM).

On applying PIM, there are two main requirements: (1) availability of capital formation data that ideally covered as many as economic sectors and specified by type as well, and (2) assumptions about asset life, asset discard or survival patterns, and asset depreciation method. Therefore, capital stock data resulted tends to be sensitive with those assumptions.

According to the method of perpetual inventory, In PIM capital stock defined as GFCF accumulation for certain periods by considering its service life, its retirement value, and its depreciation. General formula to measure capital stock is:

$$GCS = \Sigma\ GFCF - \Sigma\ RET \tag{1}$$

$$NCS = \Sigma\ GFCF - \Sigma\ DEP \tag{2}$$

$$= (GCS + \Sigma\ RET) - \Sigma\ DEP \tag{3}$$

Where:

GCS : Gross Capital Stock
NCS : Net Capital Stock
RET : Retirement
DEP : Depreciation
GFCF: Gross Fixed Capital Formation, data were provided by BPS-Statistics Indonesia

It should be noted here, that in the calculation of depreciation, it has already included the retirement effect to the asset life time. In the other word, depreciation has been calculated directly from asset life distribution resulted from the survival function.

Survival function calculation is a distribution of asset life calculated from the stream of an asset life time which whenever assigned values of GFCF is then called as Gross Capital Stock. Since depreciation calculation is also based on distribution patterns of survival function and then applied directly to GFCF then it implies that depreciation calculation used in here also takes into account the retirement effect. Extracting the effect of retirement, the above equation of NCS becomes:

$$NCS = GCS - (\Sigma\, DEP - \Sigma\, RET) \quad (4)$$

which explains that depreciation already included the retirement effect.

Assumptions taken were:

(1) *Asset life* describes how long certain assets have its service on average. This study followed such specific asset life that had been practicing in other countries where the similar method had been implemented. Asset life assumption taken by BPS also has been taken as reference as well.

(2) *Discard pattern or survival function*, Discard patterns mean a pattern that describing a declining value of assets as it was used overtime. There were five discard patterns mostly applied in PIM, namely (1) *Standard*, (2) *Linear*, (3) *Delayed Linear*, (4) *Logistic* and (5) *Weibull*. This study conducted a simulation of all these discard patterns and assumed they had similar survival pattern. Previous study by Wicaksono and Ariantoro (2001) concluded that capital stock measured were not sensitive to variation of these discard patterns application but more sensitive on asset service life assumption.

(3) *Depreciation method.* This study assumed the assets depreciation was based on *straight line depreciation method* for its simplicity and validity empirically as best practices in many countries.

Gross-up method

Another crucial issue on capital stock measurement is determining the starting year of capital stock. Earlier starting time tends to produce better capital stock. This statement is relevant due to the definition of capital stock that is an accumulation of investment overtime. Relating to this issue, this study stretched the data series back to 1960 as starting year. The reason to choose the particular year was considering the availability of investment data on pre 1980 period. BI and BPS agreed to use gross-up terms for backcasting GFCF data to 1960 based on 1980–2002 data trend. On early study, trend linear method has been applied. However, it found that the method tended to ignore investment acceleration. This study improved gross-up method by what we called three steps estimation method. First step was estimate GFCF for 1960–1980 periods by making use of investment data on that period. The following step was adjustment of GFCF 1980 before grossing-up with the GFCF 1980 grossing up result. And the last step was running the program to produce capital stock based on GFCF 1960–2002 by taking into account assumptions of asset disposal and depreciation.

The result: capital stock measurement

Assumptions of assets life and distribution patterns were as follows:

Table 1 – Assets service life assumption

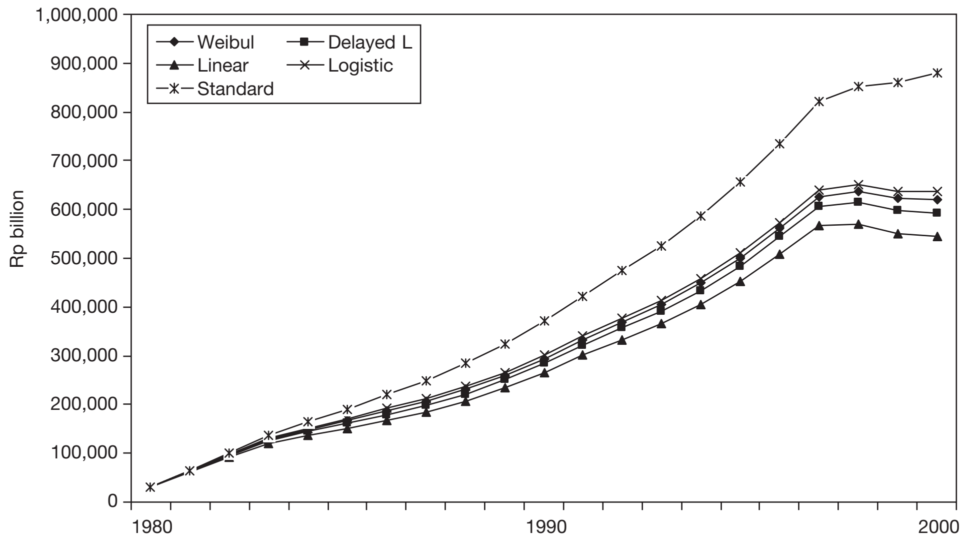
No.	Type of capital goods	Service life (year)
1	Buildings	20
2	Machines I	16
3	Machines II (electricity and its equipment)	18
4	Transportation	10
5	Livestock	3
6	Equipment – Electricity	10
7	Equipment – Metal	5
8	Equipment – Cloth & Leather	5
9	Others	16

Table 2 – Distribution patterns

Asset lifetime (age)	Survival probability
to 25% × age	93.1%
to 50% × age	61.9%
to 75% × age	23.3%
to 90% × age	9.0%
to 95% × age	6.1%

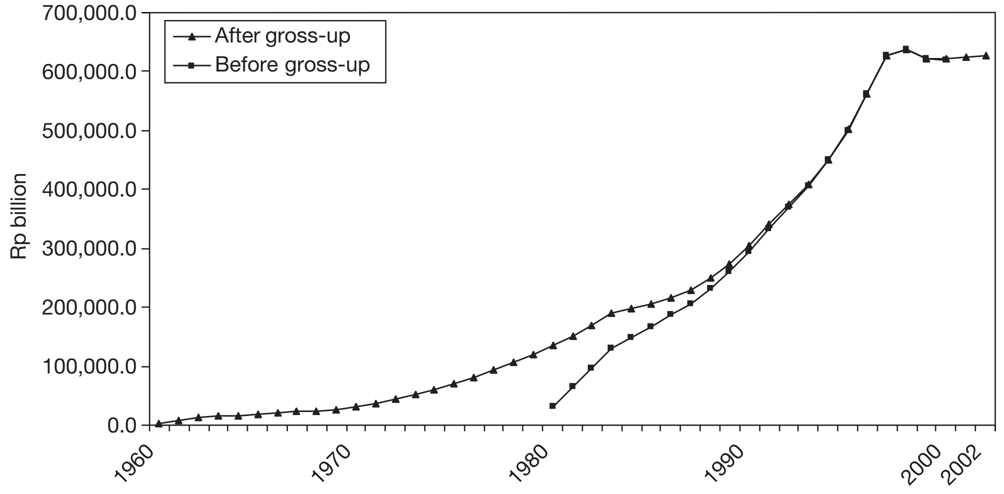
As a result of application of distribution patterns of Standard, Linear, Delayed Linear, Logistic, and Weibull, it seemed that the magnitude of capital stock of each patterns had not such significant differences. But, they apparently showed almost similar trend or patterns. Capital stock produced with the Standard distribution pattern showed the biggest magnitude. Its line also demonstrated a bigger positive slope as compare to the others particularly for the years after economic crisis begun in 1997. While, based on Linear, Delayed Linear, Logistic, and Weibull distribution patterns a deceleration and even contraction of capital stock in 1998 to 2000 were seen more obviously due to national investments weakening. And, among these distributions, Weibull distribution has been recommended as the significant one. Its magnitude and pattern could reflect the real economy condition better.

Figure 1 – NCS of all distributions



In order to have longer series of capital stock, gross-up method has been implemented to have 1960–2002 data series instead of 1980–2002. As recommended, capital stock exercised was according to Weibull distribution pattern. The method consequently lifted capital stock up, for example 1980 NCS before grossing-up was Rp 31,223 billion and after applying gross-up became Rp 135,245.6 billion or multiplied 4.3 times. But, later on the ratio became less and flat on around one times.

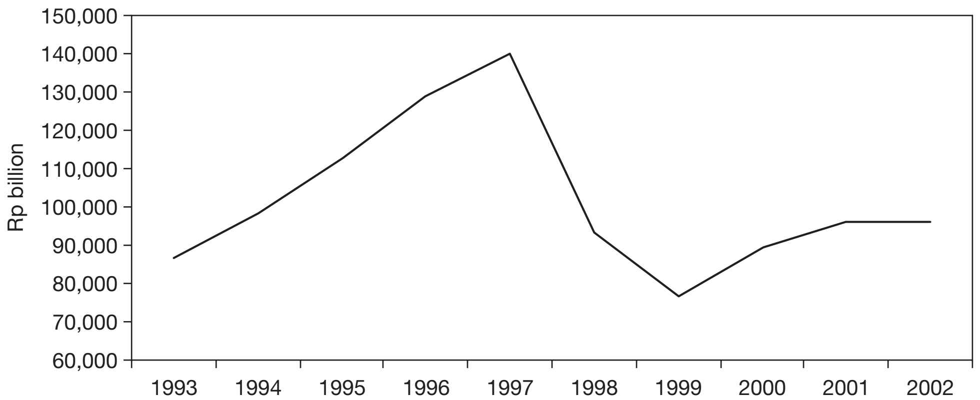
Figure 2 – NCS 1960–2002 (1993 prices), after and before Gross-up



GFCF development

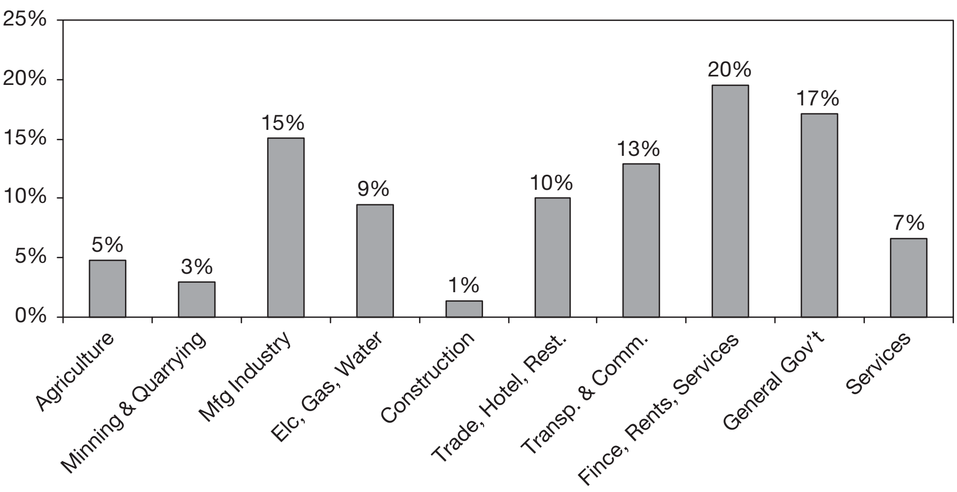
GFCF at constant 1993 prices development during the last ten years observation from 1980 to 2002 showed a sharp up and down GFCF development. Since 1993 to 1997, there was an uprising GFCF development as physical investment increased during that period. Its highest point at Rp 139.73 trillion (approximately USD 15.19 billion). But, that peak point turned to decline almost instantly and reached its lowest point in 1999 (Rp 76.57 trillion). This sharp declining was obviously reflecting national economic condition as economic crises hit the country started mid 1997. On its subsequent phase, there was a turning point to economic recovery in 1999, but at slow pace.

Figure 3 – GFCF development (last 10 years)



In terms of GFCF by economic sector contribution to total GFCF, sector of financial, rents, and business services had 20% or demonstrated its domination for the last ten years. Sector of government followed with contribution of 17%, and then the sector of manufacturing industry had 15% (see graph below).

Figure 4 – GFCF by sector contribution

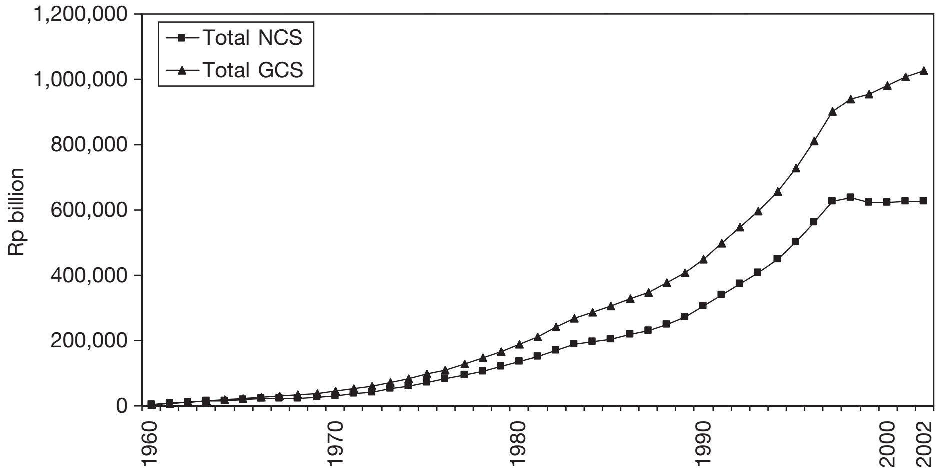


Capital stock development

According to net capital stock (NCS) at constant 1993 price for 1960–2002 periods, it had average annual growth 11.4%. If the period observed was divided into 1960–1980 and 1980–2002, each part period had average annual growth 16.0% and 7.3% respectively. Indonesia ever had such remarkable NCS annual growth on period of 1970 to 1980 as it reached 15.9%. These constant NCS highly annual growth seemed to have relation with the Government consistency on implementation of continuous five years economic plans (PELITA). But, NCS annual growth jumped to 4.2% in 1985 which was predicted to have relationship with Government policy to treat the economy overheating phenomenon. National capital stock development was also being noticed with economic crisis in 1997. Its aftermath was abviously reflecting on the NCS contraction in 1999 and 2000 consecutively.

In 2002, NCS had positive growth already at 0.45% or from Rp 625.40 trillion in 2001 to Rp 628.20 trillion in 2002. In terms of its magnitude, 2002 NCS remained lower than 1998 NCS at Rp 636.58 trillion. This meant that the accumulation of capital stock during 1999–2002 was not as many as the depreciation of existed capital goods during that period. A declining of NCS which was opposed to GCS increasing during that period also convinced that investment of new capital goods was quite weak. During that economy recession, any such capital stock growth mostly came from capital goods re-utilizations process instead of new investments. A slightly positive growth of NCS in 2002 shed a light on expected national economic recovery. On Indonesia case, there was longer and slower economic recovery as compare to other East Asian countries that had similar impact of regional financial crises in 1997. Nevertheless, there was improvement on several macroeconomic indicators performance, e.g., inflation rate, exchange

Figure 5 – GCS and NCS 1960–2002, constant 1993 price

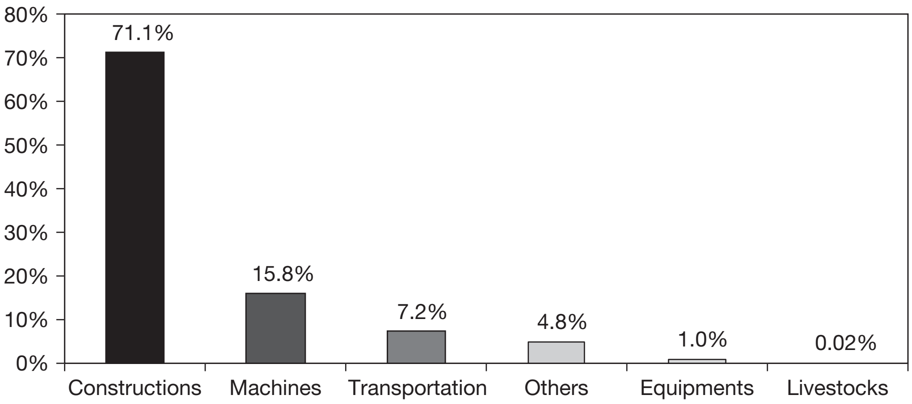


rate, interest rate, in 1999 to present. But, these positive gained was not accompanied with improvement on realization on real investments.

Capital stock by type

In term of capital stock distribution during 1960–2002, building or construction recorded 71.1% or the highest, the second place was machines at 15.8%, and followed by Transportation at 7.2%. A rapid investment on property sector particularly on real estate and shopping centers before economic crises period came was able to explain why construction sector on capital stock distribution more dominated on 1990–2002 than on 1980–1990 periods. On the other hand, contribution of investment on machines to total on 1996–2002 was 15.3%, slightly lower than that of on 1960–1965 at 16.1%. While on transportation sector, there was a consistent capital goods investment declining since 1980 as it was showed by its contribution to the total capital stock from 10.4% (1980–1985) to 4.4% (1996–2002).

Figure 6 – Capital stock by type



Capital stock by sector

Based on capital stock distribution by economic sectors, it was not equally distributed since 1980. In the period of 1980–2002, it tended to concentrate only on two economic sectors, sector of general government and sector of financial, rents, and business services whose on average 38.1% and 19.5% respectively. Capital stock on sector of general government mainly was as result of investments on buildings, roads, bridges, and other public infrastructures. But, as private sectors role to national economy increasing there was a declining portion of general government capital stock as showed 45.8% (1960–1965), 42.7% (1981–1985), and 20.2% (1996–2002).

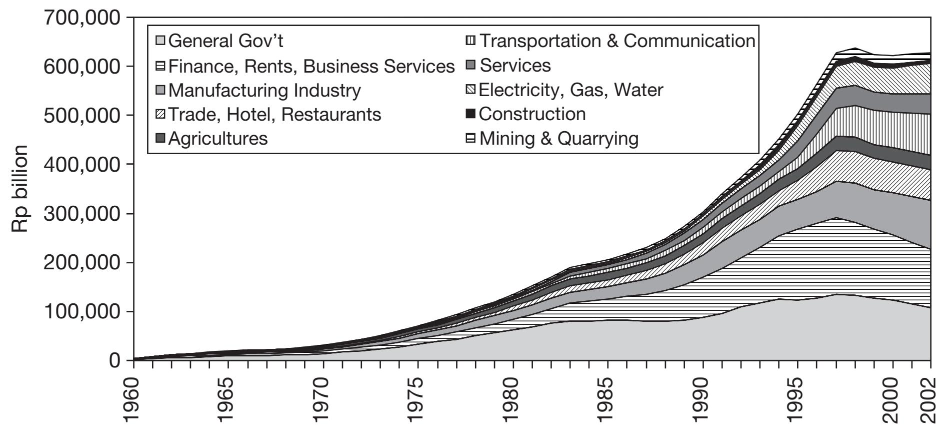
Contrastly, capital stock of sector of financial, rents and services had increasing portion on period of 1986–1990 and 1991–1995. This condition was closely related to a series of financial deregulation policies taken by government since 1983. But, financial sector expansion was halted due to monetary crises in 1997. Its portion dropped to 22.5% averagely on 1996–2002 periods.

Table 3 – Development of capital stock distribution by sector

Years	Gov't	Financial, Rents, Services	Mfg Industry	Trade, Hotel, Restaurant	Agriculture	Transport & Communication	Services	Electricity, Gas, Water	Construction	Mining & Quarrying
1960–1995	45.8%	15.4%	14.2%	6.2%	8.1%	2.8%	1.7%	1.8%	1.5%	2.6%
1966–1970	47.1%	15.3%	13.8%	5.8%	8.0%	2.5%	1.6%	1.7%	1.5%	2.6%
1971–1975	47.0%	15.3%	13.8%	5.9%	8.0%	2.6%	1.6%	1.7%	1.5%	2.6%
1976–1980	46.8%	15.4%	13.7%	6.0%	8.0%	2.7%	1.6%	1.7%	1.5%	2.5%
1981–1985	42.7%	19.0%	12.3%	6.9%	7.3%	3.5%	3.2%	1.5%	1.5%	2.1%
1986–1990	32.8%	25.1%	13.8%	7.7%	6.5%	4.2%	4.8%	1.8%	1.4%	1.8%
1991–1995	27.7%	27.8%	14.0%	7.6%	5.3%	4.0%	5.6%	3.7%	2.0%	2.3%
1996–2002	20.2%	22.5%	13.3%	10.0%	4.6%	10.7%	6.5%	8.2%	1.3%	2.7%

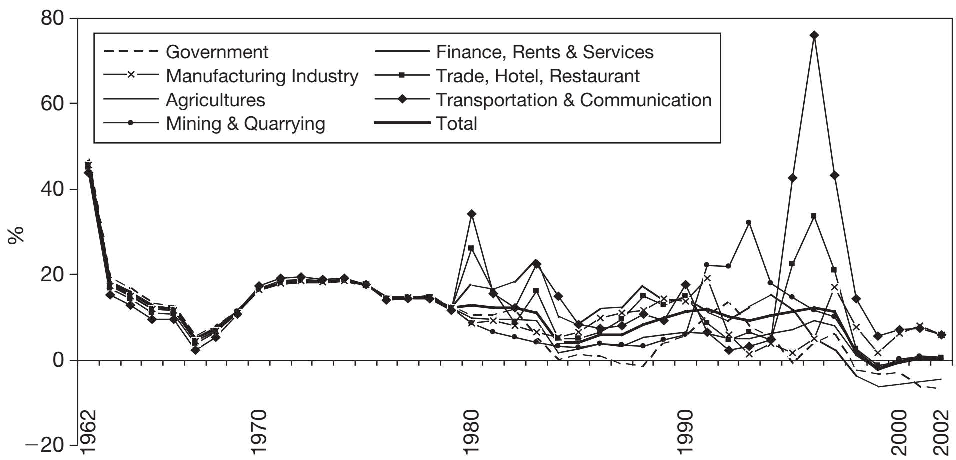
Sector of transportation and communication and sector of electricity, water, and gas had also showed such increasing portion. In 1996–2002, portion of transportation and communication sector went to 10.7% or much higher than of its average at 5%. While portion of electricity, gas, and water sector increased from 1.8% (1960–1965) to 8.2% (1996–2002).

Figure 7 – NCS development by economic sector



In term of growth, capital stock of Indonesia had average annual growth of 11.4% during 1960–2002 periods. Its growth recorded such acceleration in 1990s, specifically on sector transportation and communication, sector of mining, and sector of financial, rents, and business services. As economic crises hit in 1997, capital stock growth dropped to 1.5% in 1998, even contracted in 1999 and 2000 at 2.1% and 0.3% respectively. Positive growth was recorded again in 2001 and 2002 as capital stock grew at 0.69% and 0.45% consecutively. By economic sector, sector of construction and sector of financial, rents, and services were experience their worst contraction at 11.6% and 6.4% respectively due to the drop of their physical investments in 1999.

Figure 8 – NCS growth by selected economic sector



Validity test results

The main purpose of validity test is to ensure whether capital formation and capital stock data could be employed on further economic analysis. First step, the accuracy of capital formation (GFCF) to estimate GDP should be checked by taking simple linear regression on total and sectoral based. The linear regression results proved that GFCF could estimate GDP, both on total and sector based, significantly (see table below). Additionally, it has been taken statistical test to check whether financial investment had any result on physical investment as well. Linear regression results between GFCF as dependent variable and financial investment (INVFINC) as independent variable showed such significant relationship. GFCF variability could be explained significantly by financial investment variable.

Table 4 – Regression of GDP and GFCF sector result

Dependent variable	Independent variable	R	R^2	Adj.R^2	b	t
GDPtot	GFCFtot	0.940	0.883	0.877	2.907	12.300
GDPagric	GFCFagric	0.887	0.786	0.776	6.991	8.576
GDPind	GFCFind	0.905	0.819	0.810	5.726	9.526
GDPegw	GFCFegw	0.968	0.937	0.934	0.502	17.236
GDPconst	GFCFconst	0.712	0.506	0.482	0.536	4.529
GDPthr	GFCFthr	0.933	0.871	0.864	0.432	4.235
GDPtrcom	GFCFtrcom	0.935	0.874	0.868	0.118	11.798
GDPfrs	GFCFfrs	0.750	0.562	0.540	1.136	5.067
GDPserv	GFCFserv	0.757	0.573	0.552	0.348	5.179

Table 5 – Regression of GFCF and financial investment result

Dependent variable: PMTB

Variable	Coefficient	Std. error	r-Statistic	Prob
C	84030.2	23363.98000	3.59657	0.00290
INVFINC	0.05368	0.03001	1.78899	0.09530
AR(1)	0.80247	0.14530	5.52274	0.00010
R-squared	0.822237	Mean dependent var		81802.75
Adjusted R-squared	0.796843	S.D. dependent var		29501.56
SE. of regression	13297.23	Akaike info criterion		21.98728
Sum squared resid	2.48E+09	Schwarz criterion		22.13432
Log likelihood	−183.8919	F-statistic		32.37833
Durbin-Watson stat	1.60775	Prob(F-statistic)		0.000006

Note: Dependent Variable PMTB means GFCF.

Cobb-Douglas function

In purpose of capital stock data test to be variable candidate of potential output estimator, there had been exercising of simple Cobb-Douglas production model which employed capital stock data. In this model, capital stock and employment were treated as estimators of potential output. Input contribution to economic growth also estimated. This estimation result will be benefited for further estimation of output gap which is calculated by making difference between real output (GDP) and potential output.

The model:

$$Y = T K^{\alpha} L^{\beta} \tag{5}$$

Where,

- Y = Total output (GDP)
- K = Capital stock (NCS)
- L = Employment (LPROD)
- T = Technology
- α, β = Elasticity of NCS and LPROD to GDP

Transform the model into log-linear function:

$$\ln Y = \ln T + \alpha \ln K + \beta \ln L \tag{6}$$

Table 6 – Cobb-Douglas production model result

Dependent variable: PDBTOT

Variable	Coefficient	Std. error	r-Statistic	Prob
C	1.90237	0.31787	5.98471	0.00060
LNNCSTOT	0.25555	0.02602	9.82151	0.00000
LNLPROD	0.90128	0.06394	14.09686	0.00000
R-squared	0.99575	Mean dependent var		12.76231
Adjusted R-squared	0.99454	S.D. dependent var		0.16160
SE. of regression	0.01194	Akaike info criterion		−5.77445
Sum squared resid	0.00100	Schwarz criterion		−5.68367
Log likelihood	31.87224	F-statistic		820.69840
Durbin-Watson stat	2.054646	Prob(F-statistic)		0.00000

Note: Dependent variable PDBTOT means GDPTOT.

The regression resulted parameters each estimator as $\alpha = 0{,}256$; $\beta = 0{,}901$; and constant lnT = 1,902, therefore the function is $Y = 1{,}902\ K^{0{,}256}\ L^{0{,}902}$.

In order to calculate input factors of K and L contribution to economic growth (GDP), the model (5) above has been derived into:

$$(dY/dt)(1/Y) = (dT/dt)(1/T) + \alpha(dK/dt)(1/K) + \beta(dL/dt)(1/L) \quad (7)$$

$$(dy/dt) = (dT/dt) + \alpha(dK/dt) + \beta(dL/dt) \quad (8)$$

or,

$$r_{(Y)} = r_{(T)} + \alpha\ r_{(K)} + \beta\ r_{(L)} \quad (9)$$

where,

$r_{(Y)}$ = annual rate of output changes

$\alpha\ r_{(K)}$, $\beta\ r_{(L)}$ = K and L growth contribution

$r_{(T)}$ = technology growth contribution or *total factor productivity* (TFP)

Estimation conducted was employed data for 1990–1999 periods to suit employment data availability. Simulation was being focused on input factor contribution growth on the two difference periods, 1991–1996 (pre economic crises) and 1997–1999 (after economic crises). The simulation results were:

Table 7 – Input factors (K,L) contribution to growth

Period	$r_{(Y)}$	$r_{(T)}$	$\alpha\ r_{(K)}$	$\beta\ r_{(L)}$
Pre economic crisis	0.0754	0.0015	0.0262	0.0477
Pasca economic crisis	−0.0443	−0.0966	−0.0005	0.0529

It can be seen that input factor of employment contributed to national economic growth more than capital stock contribution, both on two difference periods. Contraction on economic growth 4.43% mainly was caused by input factor declining of technology (−9.66%) and of capital stock (−0.05 %).

Comparison to other related studies

There were several preceding studies on Indonesia capital stock measurement as reported by Keuning (1988, 1991), BPS-Statistics Indonesia (1995), and Timmer (1999). These studies coverage deliberately focused on manufacturing industry capital stock measurements only. Likewise, capital goods classification provided also limited. However, comparative study between this capital stock studies to previous studies remained relevant to conduct especially on manufacturing industry sector (table 8).

NCS of manufacturing industry on this study had almost similar distribution percentage to Timmer' study results, but the magnitude differs. Meanwhile, this study showed the growth of GCS was bigger than Keuning found for the same period on 1975–1985. In general, capital stock resulted by this study demonstrated bigger value than what Timmer and Keuning studies result. The difference results have been believed due to the difference of measurement method, base year, and *gross-up* method conducted.

Table 8 – NCS of manufacturing industry sector by composition, 1995 as comparison to Timmer' study

Manufacturing Industry	Code	Results of			
		BI-BPS		Marcel P. Timmer*	
		(1993, Rp billion)	%	(1983, Rp billion)	%
Foods, beverages, tobacco	S-31	7,448	12.01	5,998	14.60
Textile and leather	S-32	15,471	25.11	9,629	23.44
Woods, bamboos, etc	S-33	4,905	8.39	4,281	10.42
Papers and printings	S-34	6,636	11.68	4,148	10.10
Chemical	S-35	11,306	18.87	6,141	14.95
Non-metal minerals	S-36	5,483	8.86	3,100	7.55
Basic metals	S-37	1,905	3.07	1,277	3.11
Goods of metal made	S-38	7,005	11.35	6,025	14.67
Others	S-39	384	0.65	472	1.15
Total NCS	S-30	60,544	100	41,071	100

**Indonesia's Ascent on the Technology Ladder: Capital Stock and Total Factor Productivity in Indonesian Manufacturing, 1975–1995 (Bulletin of Indonesian Economic Studies, Vol.35, April 1999, Marcel P. Timmer, Endhoven University of Technology).*

Table 9 – GCS manufacturing industry sector 1975–1985 as comparison to Steven J. Keuning' study

Year	Results of				
	BI-BPS		Steven J. Keuning		
	(1993, Rp billion)	Growth %	(1980, Rp billion)	Eqivalen 1993, Rp billion*	Growth %
1975	9,811.41	17.6	66,996	20,353	---
1976	11,227.48	14.4	73,082	22,210	9.1
1977	12,867.86	14.6	79,360	24,109	8.6
1978	14,758.97	14.7	85,451	25,959	7.7
1979	16,541.87	12.1	94,343	28,660	10.4
1980	17,980.66	8.7	102,484	31,133	8.6
1981	19,646.11	9.3	111,983	34,019	9.3
1982	21,210.47	8.0	122,756	37,292	9.6
1983	22,579.76	6.5	135,274	41,095	10.2
1984	23,781.05	5.3	148,223	45,028	9.6
1985	25,295.36	6.4	159,706	48,517	7.7

**Equivalent to 1993 by GDP deflator approach.*

Conclusion

1. Wealth capital stock data has been successfully calculated by applying perpetual inventory method. Tests on both gross fixed capital formation and capital stock data proved its quality to be used for further economic analysis. The descriptive analysis also proved that the data could were able to reflecting Indonesia capital stock and economic development. As compare to previous studies, this study was more comprehensive in area of data time series produced (1960–2002), coverage of economic sectors (10 sectors), and type of capital goods (6 types).
2. Variation of distribution patterns applied to capital stock measurement, apparently gave an almost similar pattern or trend even with different levels. Standard discard pattern had the biggest level of capital stock produced as it had also been found on other studies. However according to different discard pattern and asset life assumptions simulation, it was found that asset life was more having influence on capital stock. The study result suggested that capital stock calculates on Weibull discard pattern was the most recommended to be utilized for the reasons of its level on the middle among others.
3. Further studies or special surveys to provide better assumptions incorporated on PIM operation are deserved to be middle or long term agenda. While, a preparation to apply direct observation of capital stock method are also initiated. For macroeconomic analysis significance purpose, productive capital stock is actually required.

Appendix

Table A.1 – Net Capital Stock (NCS) by distribution assumptions (Rp billion, constant 1993 price)

Year	Weibul	Delayed linear	Linear	Logistic	Standard
1980	31223.0	31223.0	31223.0	31223.0	31223.0
	63924.2	63656.2	62573.5	64086.9	64900.8
	97077.6	95967.1	92538.9	97756.5	100998.2
	129062.9	126592.6	120630.0	130595.6	137915.0
	148455.4	144325.9	135864.9	151081.8	164191.0
1985	166465.3	160683.5	150232.8	170250.6	190556.7
	186670.9	179519.5	167693.4	191491.8	219997.7
	206299.8	198122.0	184945.3	211996.3	249588.6
	230781.2	221816.1	207179.3	237216.0	284624.8
	259732.2	250073.6	233813.7	266751.9	324689.6
1990	294241.1	283835.0	265584.4	301734.6	371073.3
	332966.2	321671.6	301068.1	340880.6	422896.0
	369284.8	356937.8	333616.3	377605.0	473850.4
	405692.5	392118.3	365964.9	414475.9	526127.9
	448943.2	434032.2	405134.1	458291.3	586145.9
1995	500485.8	484153.0	452267.5	510512.3	655442.2
	561857.9	543881.2	508445.8	572745.4	735732.3
	626578.4	606570.8	566901.8	638607.1	820877.5
	636446.2	613974.9	569529.2	649930.9	853042.9
	622058.3	597217.3	549701.9	636998.7	861212.1
2000	619701.3	593504.6	546057.0	635541.5	879187.1

Table A.2 – Net Capital Stock (NCS) by economic sector (Rp billion, constant 1993 price)

Year (1)	General Gov't (2)	Finance, Rents, Business Services (3)	Manufacturing Industry (4)	Trade, Hotel, Restaurants (5)	Agricultures (6)	Transportation & Communication (7)	Services (8)	Electricity, Gas, Water (9)	Construction (10)	Mining & Quarrying (11)	Total (12)
1960	1595.74	542.73	505.64	223.24	286.46	104.22	59.77	62.56	53.16	93.33	3526.85
	3804.05	1289.91	1198.99	528.38	680.35	245.67	141.80	148.69	126.25	221.35	8385.43
	5600.87	1887.87	1746.93	766.93	994.50	353.59	206.74	217.59	184.48	322.66	12282.16
	6678.57	2231.71	2052.33	895.54	1174.04	407.63	243.01	257.13	217.59	379.35	14536.90
	7794.19	2582.46	2361.88	1023.49	1357.73	459.70	279.60	297.32	251.25	436.92	16844.54
1965	8827.73	2902.42	2642.87	1136.96	1525.98	504.09	312.53	333.72	281.84	489.31	18957.44
	9937.75	3247.40	2947.33	1259.93	1707.74	552.42	348.01	372.67	314.82	546.09	21234.16
	10498.25	3406.07	3077.49	1306.31	1790.75	564.90	362.84	389.76	329.50	570.75	22296.62
	11307.00	3651.90	3288.07	1389.86	1919.06	595.23	387.28	416.52	352.69	610.24	23917.83
	12598.44	4065.75	3654.76	1544.51	2135.29	659.59	430.34	462.36	392.32	678.50	26621.86
1970	14627.76	4732.42	4255.62	1804.10	2484.08	773.34	501.29	537.17	456.69	789.97	30962.43
	17197.35	5581.24	5023.15	2137.71	2928.02	921.21	592.14	632.92	538.82	932.20	36484.77
	20310.42	6609.60	5953.43	2542.21	3465.83	1100.82	702.41	749.40	638.38	1104.54	43177.04
	23954.79	7810.11	7039.11	3012.77	4094.10	1309.07	831.09	885.80	754.62	1305.67	50997.14
	28347.21	9253.89	8346.24	3577.13	4850.79	1558.23	985.84	1050.29	894.47	1547.83	60411.92
1975	33296.56	10873.80	9811.41	4206.38	5700.53	1834.13	1159.08	1235.01	1051.29	1819.38	70897.56
	38157.23	12448.91	11227.48	4808.76	6526.45	2093.47	1326.29	1414.30	1203.38	2082.12	81288.39
	43787.21	14272.30	12867.86	5505.42	7483.27	2393.05	1519.79	1621.79	1379.44	2386.51	93216.63
	50266.44	16372.30	14758.97	6308.57	8585.78	2738.77	1742.69	1860.63	1582.25	2737.44	106953.84
	56471.45	18365.90	16541.87	7060.07	9631.13	3056.48	1952.69	2086.47	1774.22	3068.66	120008.95
1980	62299.72	21536.71	17980.66	8893.97	10572.28	4098.67	2266.11	2229.68	2038.21	3329.62	135245.63
	68776.48	25079.34	19646.11	10338.21	11580.66	4737.82	3318.49	2413.30	2294.07	3549.98	151734.46
	76737.32	29685.75	21210.47	11239.30	12683.11	5312.86	4379.29	2616.11	2612.02	3737.26	170213.47
	80497.54	36517.48	22579.76	13069.72	13835.28	6499.95	6507.14	2798.35	2956.91	3891.56	189153.69
	80373.65	40178.17	23781.05	13732.71	14083.52	7478.36	7507.84	2954.08	3028.01	4017.89	197135.27
1985	81325.96	43436.68	25295.36	14419.56	14441.72	8105.91	7911.79	3148.58	3092.18	4138.66	205316.41
	81910.14	48575.16	27750.04	15360.78	14984.51	8709.51	8908.67	3506.10	3189.83	4292.23	217186.96
	80997.84	54518.71	30834.40	16831.55	15476.09	9401.38	10147.59	3931.85	3282.89	4438.17	229860.47
	79766.89	63767.43	34462.27	19328.88	16300.52	10422.38	12016.86	4509.95	3478.93	4580.15	248634.26
	82904.19	72427.67	39367.45	21837.33	17263.26	11390.44	13911.61	5214.19	3869.41	4797.13	272982.67
1990	87602.49	82875.79	44818.27	25115.52	18375.18	13403.63	16117.78	6172.13	4292.17	5076.90	303849.87
	96421.02	92083.52	53441.48	27295.37	19495.68	14279.71	17727.59	7149.92	5650.00	6203.63	339747.92
	109251.39	100431.44	56613.55	28595.04	20450.37	14625.40	19414.85	9818.81	7148.06	7568.83	373917.74
	117787.76	112786.79	57372.51	30465.57	21458.36	15095.53	21434.89	13589.13	8735.73	10007.65	408733.93
	124491.96	130043.03	59520.63	31931.75	22776.41	15792.02	24079.53	20274.36	10121.12	11808.78	450839.58
1995	123187.79	144704.07	60543.58	39082.05	24404.97	22538.27	34036.43	29700.73	10003.50	13540.80	501742.21
	128030.87	152051.74	63558.13	52167.06	26675.77	39716.44	38683.41	36722.39	9868.08	15102.68	562576.57
	135453.81	154939.85	74371.43	63028.56	28818.66	56943.56	42366.58	44543.77	9854.04	16619.76	626940.02
	132223.46	149183.10	80019.91	64587.57	29171.22	65042.98	41561.01	48858.57	8942.27	16991.37	636581.45
	127659.68	139568.73	81386.92	63648.65	28504.97	68594.92	38081.19	51148.78	7905.92	16668.24	623168.00
2000	123774.37	131431.75	86373.40	63352.02	28329.09	73551.00	36189.27	54230.70	7145.12	16721.64	621098.35
	115722.73	124681.91	93268.13	63536.10	28353.02	79048.57	39441.80	57966.37	6542.14	16844.16	625404.93
2002	107827.40	119068.67	98844.18	63778.15	28380.86	83674.09	42439.86	61319.51	6006.62	16862.48	628201.82

References

Ariantoro, Eko (2001), *Penetapan Asumsi dalam Perhitungan Stok Kapital*, Bagian Statistik Sektor Riil dan Keuangan Pemerintah, Bank Indonesia.

Australia Bureau of Statistics (1997), *Australia's Methodology for Compiling Estimates of Capital Stock and Consumption of Fixed Capital*, Background Paper of Capital Stock Conference in Canberra.

Badan Pusat Statistik (1997), *Estimation of Capital Stock and Investment Matrix in Indonesia*, Background Paper of Capital Stock Conference in Canberra.

Badan Pusat Statistik (1995), *Study of Estimation of Capital Stock in Indonesia 1979–1994*, Jakarta.

Badan Pusat Statistik, *Tabel Input Output*, berbagai seri, Jakarta.

Bagian Statistik Sektor Riil dan Keuangan Pemerintah (2000), *Kajian Kemungkinan Pengumpulan Data Stok Kapital Sektor Industri Pengolahan*, Direktorat Statistik Ekonomi dan Moneter, Bank Indonesia.

Bagian Statistik Sektor Riil dan Keuangan Pemerintah (2001), *Draft Stok Kapital*, Hasil Kerjasama Direktorat Statistik Ekonomi dan Moneter, Bank Indonesia dengan Badan Pusat Statistik.

Bagian Statistik Sektor Riil dan Keuangan Pemerintah (2001), *Proses Pengumpulan Matriks Pembentukan Modal Tetap Bruto (PMTB)*, Hasil Kerjasama Direktorat Statistik Ekonomi dan Moneter, Bank Indonesia dengan Badan Pusat Statistik.

Divisi Pengembangan Statistika dan Komputasi (2001), *Materi Workshop: Probabilitas dan Distribusi*, Fakultas MIPA, Institut Pertanian Bogor.

Frenken, Jim (1992), *How to Measure Tangible Capital Stock?: The Choice between Two Methods*, CBS Netherlands.

Keuning, Steven J. (1988), *An Estimate Of Fixed Capital Stock By Industry and Type Of Capital Good in Indonesia*, Statistical Analysis Capability Programmed, Working Paper Series No. 4, Jakarta.

Saleh, Kusmadi (1997), *The Measurement of Gross Domestic Fixed Capital Formation in Indonesia*, Paper of Capital Stock Conference in Canberra.

Meinen, Gerhard (1999), *Measuring Capital Stock: Explanatory Notes for the Manual on Capital Stock Statistics*, Statistics Netherlands.

Meinen, Gerhard and Bert Verlinden (1997), *Statistics on Tangible Capital Stock: Direct Observation at Statistics Netherlands*, Paper of Conference on Capital Stock (Preliminary Version) in Canberra.

Meinen, Gerhard, Piet Verbiest and Peter-Paul de Wolf (1998), *Perpetual Inventory Method: Service Lives, Discard Patterns and Depreciation Methods*, Department of National Accounts, Statistics Netherlands.

Timmer, Marcel, *The Dynamics of Asian Manufacturing: A Comparative Perspective*, 1963–1993, Eindhoven Centre for Innovation Studies, Netherlands.

Timmer, Marcel (1999), *Indonesia's Ascent on The Technology Ladder: Capital Stock and Total Factor Productivity in Indonesian Manufacturing, 1975–95*, Bulletin of Indonesian Economic Studies, Eindhoven University of Technology, Netherlands.

Wicaksono, Gunawan (2001), *Perhitungan Stok Kapital: Beberapa Pengamatan terhadap Model Perhitungan Stok Kapital dengan Perpetual Inventory Method Metode Australia dan Metode Belanda*, Bagian Statistik Sektor Riil dan Keuangan Pemerintah, Bank Indonesia.

Abstract

The paper presents a study of Indonesia capital stock measurement by Bank Indonesia in cooperate with BPS-Statistics Indonesia based on indirect method of perpetual inventory. Comparing to previous studies, the data estimates were more comprehensive in terms of economic sectors and its type. Assumption discard patterns of Standard, Linear, Delayed linear, Logistic, and Weibull were employed. The wealth capital stock data could successfully be provided for 1960–2002 period. Capital stock data of each pattern resulted apparently had similar pattern or trend with insignificance discrepancy levels. However capital stock estimates with Weibull discard pattern was most likely recommended one. The data estimates were provided both on net and gross concept. Capital stock development showed annual growth 11.4% on average for 1960–2002 periods. Graph analysis demonstrated that capital stock growth start to decelerate as economic crises hit the country in 1997 (it grew 1.5% in 1998 and even contracted 2.1% in 1999). A following step to test data quality by incorporating fixed gross capital formation and capital stock data on several statistical tests and also on simple Cobb-Douglas production function proved its quality for further macroeconomic analysis.

Noor Yudanto, Gunawan Wicaksono, Eko Ariantoro and A. Reina Sari

Researchers of Real Sector and Government Finance Statistics Division
Bank Indonesia
E-mail: srkp@bi.go.id

DISCUSSANTS' COMMENTS

Comments on the workshop about output gap and capacity

Ahmed Al Kholifey (Saudi Arabian Monetary Agency)

Thank you Mr. Chairman. Let me at the outset thank the organizers of this workshop for giving me an opportunity to comment on the four papers just delivered by the distinguished speakers. I admire the authors for the straightforwardness of the papers given the complexity of the topic, and agree with them in their argument regarding other ad-hoc methodologies of finding the Output Gap.

I found the issue of output gap and the methods of estimating it using the multivariate approach, very interesting. There seems to be a general understanding in the literature of advantages of multivariate approaches over univariate ones. The latter are accused of being mechanical for they leave economic aspects of the variables aside. Comparison of results of the estimation of output gap in Peru lends some support to that argument. I view a move towards multivariate approaches as a step closer to reality as it means moving away from being pure statistician and to add more economic flavor to the system of equations that is needed to estimate the unobservable output gap.

The Kalman technique is a good method for predicting variables and updating the parameter estimates in the predicting equations. Applying the techniques of Kalman filtering to the economies of Peru and Colombia, and Extended Kalman to that of Turkey is a good contribution to the literature on output gap estimation. The challenge of finding ways to estimate output gap stems from the fact that we are dealing with an unobserved variable whose size and direction are of great importance to policy makers since it is widely known that the wider the output gap today, the higher will be inflation tomorrow. Added to this challenge is the difficulty of deciding whether a given output change is cyclical or permanent.

Before I comment on papers, I would like to congratulate the speakers who tackled the issue of output estimation and contributed to the literature in the form of case studies about their own countries. I have a couple of general observations. First, the reader of the papers related to specific economies would find brief background about the structure and potentials of these economies very helpful in having a good picture of the economies the paper tended to address. Second, one way of testing the robustness of the estimation of the model is to conduct an out of sample forecasts for the output gap. This has been done by Mr. Cobo, and I think, it would be beneficial to have similar forecasts in other papers.

I would like now to consider some specific comments about each paper. Nevertheless, I will not go into technical details because surely I am less cognizant of these techniques than the authors of the papers.

Estimating output gap in Peru

Let me begin first with the paper by *Ms. Miller* and others about estimating output gap *in Peru*. In this paper, the authors applied a multivariate unobserved component (MUC) model to estimate the Peruvian's output gap, relying on an explicit short-term relation between the output gap and inflation rate (Phillips Curve) and structural restrictions over output dynamics. Using Kalman Filter Method, the authors estimated their model for the period 1992–2003. The authors found that MUC output gap was less sensible to end of sample problems and presented a better relation with the Peruvian inflation process than other estimates that had applied Hodrick-Prescott filter and the production function method.

I think the paper is well organized and nicely written. It explains the system of behavioral equations in an easy to understand way. However, I would prefer having a discussion of the results with the findings of other authors that applied multivariate techniques and specifically Kalman Filter.

I believe a brief discussion of the structural breaks in the data, resulting from institutional changes in Peru may be needed. This would give the reader a better understanding of the issue.

Estimating output gap in Turkey

The paper by Mr. Sarikaya and others about estimating output gap in Turkey is written with beautiful English. It provides nice and clear empirical model with interesting results supported by empirical findings. However, I would like to point out the following remarks:

In the introduction, the paper emphasized the importance of exchange rate movements to the Turkish economy, as is the case in other emerging markets. Therefore, the variable named TRQ, a proxy for short term ER, was incorporated in output gap dynamics equation. The authors stated that "in Turkey, imported intermediate and capital goods constitute a significant share of the total inputs, hence the ER emerges as a primary input cost affecting production decisions". In addition, the paper states that exchange rate developments have been the primary determinant of inflation. Yet, when it came to the choice of appropriate model the authors chose, I think for the sake of simplicity, model 2, which excludes the proxy of ER variable. My own view is that if just for the sake of simplicity an important variable would be dropped then this might be at the expense of the robustness of the model.

The paper emphasized the point that "given the high and persistent inflationary environment, real interest rates remained at so high levels that its link to consumption and investment decisions disappeared". I believe this opinion is partially supported by results obtained by authors in which interest rate coefficient produced conflicting signs, depending on the period we deal with. For instance, after 1996, the relationship between interest rates and output gap turns out to be negative, in line with economic theory. But as described by authors, in 1998 surprisingly there was positive output gap along with high interest rates! So it seems there is still unclear relationship between interest rates and output gap.

When talking about output gap as only one of the indicators enriching the information set of policy makers, the authors mentioned "capacity utilization rate" as one of the other variables that could be considered. The definition of "capacity utilization rate" is the actual to potential output expressed in ratio terms. I wonder what additional information will capacity utilization rate offer to policy makers that the output gap estimate will itself not offer?

A word of caution about the recursive process. In the presence of volatility as the case in the Turkish economy for the period tested, non-linearity nature my not be consistent overtime. The fear is that recursive process might become an error generating process.

Output gap in Colombia

I come to the paper by Mr. Cobo on Output Gap in Colombia. It's a comprehensive and informative paper with good explanation of the various methods of estimation. However, in the section about multivariate methods (section 3.3 and 3.4), it is not clear whether the author himself developed the system of equations and generated the empirical findings presented, or he is just presenting the findings by others.

The paper itself layed out the shortcomings of the model developed by Perez (2004), one of which is that the equations except Philips and IS curves are ad-hoc equations. I think a provision of reasoning of that statement will be more convincing. Besides, the author states that the variable of demand situation indicator (ddgap) can't be compared to other estimates due to its different nature (i.e. qualitative). The question is what is the purpose of incorporating this variable if it's not comparable. Also, this variable (ddgap) is narrowly focused since it is the demand by the manufacturing sector that is used as a proxy for gap in demand.

Furthermore, for clarity purposes a brief explanation should be given about the non-availability of coefficients of GDP trend corresponding to the last three variables of table 1 (ICUGAP, NAIRUGAP and DDGAP), which were given n.a.

Capital stock in Indonesia

The paper by Mr. Yudanto and others about measurement of capital stock in Indonesia, unlike the other papers, attempts to recalculate Capital Stock Data in Indonesia for earlier period

(1980) as a means of estimating national potential output. It, however, did not go further to estimate the output gap.

This paper gave brief explanations of the concepts and definitions of most of economic terms used by authors, which in my opinion, is a very useful approach.

I noticed that the Gross Fixed Capital Formation (GFCF) definition was drawn from the SNA of 1968. I wonder why not more recent edition (i.e. 1993) was used? Using recent edition would provide more inclusive definition of GFCF as it applies to the evolving nature of assets especially intangible assets (e.g. Computer Softwares).

When talking about fast growth in Net Capital Stock (NCS) in the 1970s and the role of the government in that, I became curious about the influence of oil market developments during that time. Did it have a strong push towards more oil facilities build up?

The paper gave assumptions about assets service life. But it does not spell out the basis of these assumptions: Are they arbitrarily chosen? or borrowed from the literature about other countries, or based on some sample survey. Besides, as it appears, the same assumptions have been used across all sectors. I wonder whether this treatment would give the right weight to different sectors of the economy.

The authors checked the accuracy of GFCF which is basically the published investment data. In my view there is hardly any need to check their accuracy to estimate GDP. However, if there is such a need, then it be used as a regressor for change in GDP, rather GDP itself.

Regression results show unstable parameters in Cobb-Douglas Function where $\alpha+\beta$ exceeds 1. This is probably what caused the oddness in the input factors contribution to growth (Table 7), where contribution of technology has changed from virtually zero before the crisis to negative 10% after the crisis and of labor from 4.8% to 5.3%. Estimating a constrained production function will actually test the validity of the capital stock data complied by the indirect method.

Finally, Mr. Chairman, I would like again to congratulate the authors for their work, which I believe will inspire further research in that field. Thank you to all of you.

Ahmed Al Kholifey (Saudi Arabian Monetary Agency)

WORKSHOP E

Output and output indicators

Chair: Peter Charleton (Central Bank and Financial Services Authority of Ireland)

Papers:

Measurement of gross domestic product (GDP) and challenges for the new base year 2003 series: the Chilean case
José Venegas Morales (Banco Central de Chile)

Constructing a composite leading indicator for the Turkish economic activity and forecasting its turning points
Aslıhan Atabek, Evren Erdoğan Coşar and Saygin Şahinöz
(Central Bank of Turkey)

Real-time data and business cycle analysis in Germany
Jörg Döpke (Deutsche Bundesbank)

GDP flash monthly estimates in DRC
Gerard Mutombo Mule Mule and Marie-Jose Ndaya Ilunga
(Central Bank of Congo)

Quality measures for quarterly national accounts and related external statistics
Jorge Diz Dias and Wim Haine (European Central Bank);
Ingo Kuhnert (European Commission)

Measurement of gross domestic product (GDP) and challenges for the new base year 2003 series: the Chilean case[1]

José Venegas Morales (Banco Central de Chile)

Introduction

The calculation of GDP, the traditional national accounts indicator, has not undergone any great transformation as regards methods and sources. However, the present day situation has changed the setting in which data are defined and methods are applied and thus poses several challenges to the task of measuring GDP. In this paper we discuss four aspects:

- GDP is a variable involved in a complex set of accounts.
- More frequent and timely measurements of GDP must be made.
- The present situation requires substantial changes in the definition of activities, goods and services.
- Analytical perspectives have widened remarkably and therefore require a far more professional approach to relational database design and management.

In Section 1 we describe Chilean GDP measurement methodology. In the following three sections we discuss the first three aspects on the above list and the final Section 5 draws attention to the need for a framework for data modeling and information systems to resolve the problem of how to harmonise the various sources, different analytical perspectives and systems of classification.

1. Background to GDP calculation in Chile

In 1940, Chile was one of the first countries in Latin America to begin with official measurement of GDP.[2] Two factors were influential in this beginning:

- the theoretical development of accounting schemes based on J.M Keynes' effective demand approach which guided the progress of the *Conferences on Research in National Income and Wealth* during 1937 to 1943.
- the economic thrust in America and Europe following the Depression of the early '30s and World War II.

These stimulated the creation of the Chilean Economic Development Agency (CORFO), dependent on the Ministry of Economy. For the corporation to formulate and carry out its plans, it was considered that technical studies of the nation's economic reality must be made and these resulted in the first official publication of the national accounts covering the period 1940 to 1945 (CORFO, 1946). Later series published were 1940–1954 (CORFO, 1957), 1950–1960 (CORFO, 1960), 1940–1962 (CORFO, 1962) and 1958–1963 (CORFO, 1964).

In 1965, within CORFO, the National Planning Office (ODEPLAN) was created and later became independent in 1967 as an advisory office to the Presidency. As part of ODEPLAN, the National Accounting Department was created by transferring professionals from the Economic Research Section of the CORFO who were responsible for compiling accounts.

1 *Bank for International Settlements- Basel- Switzerland, Conference of the Irving Fisher Comittee on Central Bank Statistics, 9–10 September 2004.*

2 *Mamalakis, 1978, provides a thorough and well-documented history of national accounts in Chile covering the whole period of CORFO and ODEPLAN up to 1976 aproximately.*

The main aim of this recently created National Accounting Department was to compile the input-output tables for the year 1962 (ODEPLAN, 1968). During this task, CORFO's 1959–1961 series (ODEPLAN, 1965) was updated and the 1960–1966 series (ODEPLAN, 1967) was published.

The input-output study began in August 1962 and ended in May 1967. These input-output tables were updated in 1965 and that year became the first base year for constant and current prices measurement of the national accounts.

In ODEPLAN, the methodology of national accounts underwent changes. The United Nations' *A System of National Accounts* (SNA) first published in 1953 had just been revised (United Nations, 1965) and the new version considered Education and Public Health as separate activities from the central Public Administration. It also made a distinction between corporations and establishments and it introduced methods of deflation for the measurement of GDP at constant prices. The SNA provided the guidelines for the Chilean National Accounts for the period 1960–1976.

Parallel to the compilation of input-output tables 1962, the first study was carried out of sources and uses of funds (ODEPLAN, Central Bank of Chile, 1967) for the period 1962–1964. The aim of this study was to know the country's capital market structure, the magnitude, origen and destination of financial resources and the nature of capital market instruments.

National accounts measurements with base year 1965 continued to be published periodically by ODEPLAN: the series 1960–1970 (ODEPLAN, 1970), 1960–1971 (ODEPLAN, 1972), 1965–1972 (ODEPLAN, 1972), 1960–1971 (ODEPLAN, 1975), 1965–1973 (ODEPLAN, 1976), 1960–1975 (ODEPLAN, 1977), 1972–1977 (ODEPLAN, 1978).

ODEPLAN also initiated regional measurements of GDP, producing the first official publication of this nature in 1978.

The last national accounts publications in ODEPLAN referred to the input-output tables for 1977 (ODEPLAN, 1981) which defined the change of base year 1965 of the earlier series of national accounts. These new tables partially adopted the recommendations of the SNA Rev.3 (United Nations, 1968).

The 1974–1980 series with 1977 as base year were published by ODEPLAN in 1981. However, economic authorities and users of national accounts were anxious to have a longer series of the main aggregates and pressed for the compilation of the 1960–1973 series with base year 1977. This was not published in ODEPLAN because of the changed circumstances of national accounts which towards the end of 1981, in order to concentrate the whole macroeconomic information function in one institution, were transferred to the Central Bank of Chile. The corresponding team of professionals was also transferred to the Bank and only the regional accounts unit remained in ODEPLAN until 1986 when it also moved to the Central Bank.

In its new residence, national accounts published the series with base year 1977 corresponding to 1960–1980 (Central Bank of Chile, 1982), 1960–1982 (Central Bank of Chile, 1983), 1960–1983 (Central Bank of Chile, 1984), 1974–1985 (Central Bank of Chile, 1985).

At the same time, new demands for information for economic analysis and decision-making led to the development of short-term economic measurements and thence began the publication of the quarterly GDP (Central Bank of Chile, 1982) and the Monthly Indicator of Economic Activity IMACEC (Venegas, 1985).

In the light of the structural changes in the economy and the need to improve economic measurements, the Central Bank established a new base year 1986, compiling the third input-output matrix for the Chilean economy (Central Bank of Chile, 1992). This was a milestone in the development of the methodology used to compile Chile's national accounts because the new guidelines of the SNA 1993 were followed even though they were still only available in draft form. A new input-output matrix 1986 and a set of institutional accounts proposed by the SNA were drawn up and they formed the most complete and up to date implementation of the SNA 1993 recommendations in Latin America, from the point of view of both methodology and variety and detail of information. Only the households sector was not published separately due to lack of information and no table of cross classification of activities and institutional sectors (CCAS) was included.

The 1986 input-output study, albeit more complete than in the previous base years 1965 and 1977, still limited national accounts to the sphere of activity production accounts and supply and use of goods and services accounts. Institutional sector accounts had not advanced since the pioneering study of flow-of-funds accounts published by ODEPLAN and the Central Bank in 1967. To fill in this gap a study of sources and uses of funds 1986–1989 was made (Central Bank, 1994). The scheme of integrated institutional accounts could not be completed because at that time income and outlay accounts of less relative development were not yet available.

After these studies, national accounts with base year 1986 were published covering the series 1985–1992 (Central Bank, 1994). That series did not include the functional distribution of income which was only incorporated into the tables published in the Central Bank's monthly report as from March 1994.

In these years of intense publishing activity, the first regional GDP 1986–1992 was published in 1994, once the regional unit that had been left behind in ODEPLAN was at last transferred to the Central Bank. A GDP series with base year 1986 retropolated for the period 1974–1989 was also compiled (Central Bank, 1993).

As from 1998 an annual publication of national accounts concentrates all the existing series of production accounts, institutional accounts, regional and quarterly GDP and the series of IMACEC. The first national accounts year books base 1986 corresponded to the 1985–1995 series (Central Bank, 1997), 1985–1996 (Central Bank, 1998) and 1985–1998 (Central Bank, 1999).

The input-output tables year 1996 were compiled between 1997 and 2000 (Central Bank, 2001). In this new base year, the SNA 1993 recommendations were fully applied and the gaps in information on households and the cross-classification of activities and institutional sectors (CCAS) that had been missing in 1986 were completed. All the progress made in information was given over to producing a publication that covered both the input-output matrix and the integrated institutional accounts.

Meanwhile, the development of the system of short-term economic indicators went ahead, as did the greater integration of national products in the decision-making process and diagnosis on the part of monetary authorities and institutional economic agents of the Chilean economy. The publication of the year book, suspended in 2000 due to the change of base year, began again in 2001 with the inclusion of the 1996–2000 series (Central Bank, 2002). This annual publication has continued with Year Book 2002, 1996–2001 series (Central Bank, 2003) and Year Book 2003, 1996–2003 series (Central Bank, 2004).The regional GDP base year 1996 has continued to be published in the 1996–2000 series and 1996–2001 series (Central Bank, 2004).

Progress in measurement and awareness of the increasingly frequent structural changes besetting our small economy, open to international trade, make it necessary to reduce the period between base years. Thus, in 2002, it was decided to embark on a new base year 2003. This program shows marked differences in the organisation and management of national accounts, not only on the part of the Central Bank but also national statistics institutions.

Among other management changes, a strategy of sectoral and intersectoral projects is being developed which can be forward of, backward from or contemporary with the year 2003 which is the reference year. This strategy allows for a more efficient distribution of professionals and shortens the time for publication of the input-output matrix 2003 (2006) and of the integrated institutional accounts and series of annual accounts base year 2003 (2007).

These projects began in May 2002 and now, after two years' work, the projects of construction and stock of capital are finished, the sectoral compilation of services and trade are nearly completed and the projects of foreign trade, agriculture, livestock and forestry and fishing are well under way. Towards the middle of 2004, the projects of mining, manufacturing industries, energy, transport, communications, hotels and catering, government, health and education were all begun. The pre-balanced sectoral and intersectoral input-output information should be ready by the second quarter of 2005. There will thus be six months to balance supply and uses of goods and services in order to meet the proposed publication date of March 2006.

Two points should be made about the development of national accounts in Chile.

- The progress made in methodology of national accounts has not been accompanied by similar development of the national statistics system which provides the basic information. There were advances in the statistics system during the '70s but in the following decades there has been a weakening of the coverage and technical performance in the work of basic statistics on production, consumption and investment. This has affected the work of national accounts which has had to turn to other complementary sources of information. Consequently, these other sources such as administrative tax registers and foreign trade registers have been strengthened. Nevertheless, the deficient institutionality and lack of investment in national statistics is evident.
- The progress made in developing national accounts methodology and short-term economic indicators has been a national task. Only on certain occasions has foreign technical expertise been required and only with regard to specific aspects of measurements. There have thus been more benefits than costs. Not least of the benefits is the creation of independent autonomous professional teams, thereby avoiding the cost of depending on foreign know-how and benefitting from the challenge of finding our own solutions in applying methodology and in creating operational platforms for measurement. The downside of this way of developing is the lack of an

international institutional counterpart to validate our national methodology and to examine and assess methodological options to confront new areas. Efforts are being made to overcome this lack by establishing permanent relations with international institutions such as the OECD or Eurostat and national accounts departments of countries in the forefront of methodology.

2. GDP calculation in the context of national accounts

GDP measurement is subject to a series of rules of consistency. We can conceive the rules of consistency or soundness of the figures using central trend measures of GDP data in a dimensional space *n*. Deviations within a normally accepted range can be considered except for justifiable reasons of economic behaviour of types of activity or supply and use of goods and services.

GDP consistency in the balance between supply and use of goods and services

In the early stages of GDP measurement, the rules of consistency were naturally more general but as GDP is calculated in greater detail and complexity with regard to valuation, frequency and geographical space rules become more specific. In the case of Chile before the 1980s the rules considered a basic dimension of currency for measurement: current prices and constant prices. This consistency supposes that there must be a sound measure in the implicit prices *p* resulting from the comparison between the two. If we symbolise the GDP consistency rules as r_{GDP}, for this particular rule we would have an function of the type:

$$r_{GDP} = f(p)$$

Also as a primary definition of consistency we must consider the coverage of products *i* and activities *j*, the greater the coverage of products and activities, the greater the difficulty of achieving consistency. This was evident when in the base year 1977 the number of types of activity was increased from the 54 of the 1965 base year to 66. These 12 additional activities posed problems of consistency. The effect of greater breakdown of activities and products can be taken to an extreme when *j* is extended to the level of establishments and *i* to the level of specific products. In Chile, the massive introduction of personal computers in the ’80s meant that electronic spreadsheets and production accounts were made using Excel and Access, allowing the aggregation and segregation of corporations’ data to an extent unheard of for base years 1965 and 1977. As from base year 1986 the large corporations’ carried out detailed processing from which production accounts were made and thus added value or GDP obtained for each establishment. We can call these dimensions of balancing agent dimension (*j*) and object dimension (*i*). Thus

$$r_{GDP} = f(p, i, j)$$

Another fundamental rule of consistency that makes measurement more complex arises when input-output tables are made to define a base year (3). The input-output coefficients (a_{ij}) define a structural correlation in the behaviour of the GDP of the various types of activity. If coefficients are defined in an extreme range of 0 to 1, closer to 1 means greater correlation between the types of activity concerned. This implies that the rates of variation of GDP between the activities defined in the coefficient should move in an ever tighter range if the coefficient is closer to 1. The behaviour of the figures cannot alter this rule of consistency without falling into a structural change that clearly distorts intersectoral input-output relations. Thus, this structural dimension that arises from the input-output tables in a base year can be added to the consistency function:

$$r_{GDP} = f(p, i, j, a_{ij})$$

It should also be pointed out that input-output tables may be of different structural qualities. If we consider them as an interweaving of relations between supply and use of goods and services, the cloth produced may be thick, weak and loosely woven or fine, strong and densely woven. In the Chilean experience, we see that the input-output matrix for 1977 was of the first type compared to the later 1986 and 1996 input-output matrixes which were of the second type.

Strictly speaking, the difference in quality is due to the treatment of valuation. Considering valuation in its most abstract form it is nothing more than the breakdown of production value into its costs once out of the factory or field (basic price). If we think of the production value at basic price like the heart of an onion, valuation is the layers of onion added to the product once it leaves the production site. Valuation is thus no more than another production or transfer (in the case of taxes) value and therefore another product *i* in the interweaving of input-output.

Let's take for example the value chain of tobacco. The following could be considered the layers of valuation:

  Tobacco in the factory at basic price (pb)
\+ special tax on tobacco (iit)
\+ value-added tax (vat)
\+ margin for transport (mgt)
\+ insurance (seg)
\+ storage (alm)
\+ advertising (pbl)
\+ trade margin (mgc)

---
tobacco at market price

As we see from this example, a layer of valuation is added for any cost not included in the producer's inputs. For example purpose, advertising, insurance and storage are explicit distribution cost instead of production cost. A product's layers of valuation are not explicitly stated in the consolidated GDP at market price because they are part of the production value of some activity.

Similarly, taxes are an "added value" but one that is imposed by the government for it to fulfil its protectionist or redistributive aims. However, aims have no bearing from the point of view of valuation, taxes are like any other layer of valuation that constitutes the production value.

National accounts valuation requirements define new rules of consistency of the type a_{ij} widened to an i^* that is a layer of valuation. In simple terms, the behaviour or results of the GDP should keep a certain coherence in terms of the stability of the coefficients of distribution margins and taxes and other layers of valuation defined in the input-output tables. We can therefore widen the function as follows:

$$r_{GDP} = f\,(p,\ i,\ j,\ a_{ij},\ a_{i^*j})$$

In general, when the behaviour of GDP between activities (a_{ij}) or in relation to the layers of valuation (a_{i^*j}) begins to go beyond a normal range of variation, there is an evident deterioration in the structural relation in the economy and this signals the need to redefine a base year. There are various measures of this structural discrepancy (Fontela, Lopez, Pulido, 2000) but their correct application and interpretation are not simple, faced with the reality of the data compiled in the input-output tables of each country.

The rules of consistency must also extend to other dimensions, one of which is the geographical dimension defined with the calculation of regional GDP. In the case of Chile, regional GDP measurements go back to the '60s and have continued annually to date. This rule of consistency (*e*) assumes that there should be a certain correlation between the GDP of an activity at a regional level and at the level of the whole country. We can add this rule to the function:

$$r_{GDP} = f(p,\ i,\ j,\ a_{ij},\ a_{i^*j},\ e)$$

The number of rules of consistency continue to increase as new analytical dimensions are added. In the case of Chile, among these dimensions are the following.

The increase in the frequency of direct or indirect measurements of GDP – from yearly to quarterly to monthly. This is the time dimension (*t*) which imposes another rule of consistency similar to the geographical dimension. That is, there should be a certain correlation between the annual GDP and the quarterly or monthly measurements of GDP. In Chile, official calculation of quarterly GDP began in 1984 and the calculation of IMACEC began in 1987. Therefore:

$$r_{GDP} = f(p,\ i,\ j,\ a_{ij},\ a_{i^*j},\ e,\ t)$$

But the increase in the number of rules of consistency has not ended. We are in the initial stages of developing numerous other interactions among different sectors of the economy that are widening GDP calculation, we can call it the satellite dimension (*s*). In Chile measurements of varying degrees of elaboration have already been made for tourism, health and culture. Thus we now have

$$r_{GDP} = f(p, i, j, a_{ij}, a_{i*j}, e, t, s)$$

Among all this quantity of rules of consistency that potentially should be considered in the additive quantification of GDP a fundamental distinction must be made. Once we open up the field of GDP measurement, all these analytical dimensions no doubt have a bearing. But, since it is impossible to measure GDP from "atomic" or singular observations that compound an additive effect on GDP, we establish a convention that in practice is implicit:

- There are dimensions that have an incidence in the additive measurement of GDP.
- There are dimensions that require an adjustment to the overall GDP.

The first are determinant in the whole country GDP measurement. These are currency, coverage of activities and products defined in the input-output matrix of the base year, coefficients of input-output and valuation.

In the case of the second ones that imply an adjustment, in Chile the ones to be considered are the time (quarterly GDP, IMACEC), regional and satellite dimensions. In other words, any of these dimensions could be "determinant" and not "adjustable" if the production accounts were compiled from short-term, regional or satellite measurements.

The consistency of GDP in integrated national accounts[3]

The complexity of the context in which GDP is recorded is clearly exposed in the base year measurements. The difficulty in expressing the system of national accounts simply and in the aggregate form can be seen by examining the tables in the 1993 SNA. The integrated economic table (Table 2.1), the set of "T" accounts of Annex V and the matrix format given in chapter XXI are all pretty dense and hardly didactic according to economists, users in general and even some national accounts specialists. However, the new practical national accounts manual (United Nations, 2003) is an improvement on this situation. In Chile for the 1996 base year publication a panoramic vision of matrix-focused integrated accounts was chosen. Strictly speaking, in the form of an aggregated Social Accounting Matrix (SAM). This matrix was called Matrix of Integrated Economic Accounts (MIEA) (Aceituno, Venegas, 2001) and it serves as an instrument of cross consistency for the various national accounts. It is also possible to appreciate all the interrelations present in GDP measurement. The MIEA integrates current flows and accumulation flows and, according to the specific analytical objective, it is possible to incorporate physical or value data from satellite or complementary measurements. Table 1 shows the aggregate version of the MIEA for the year 1996.

Table 1 – Aggregated MIEA of the Chilean economy 1996

(billions of 1996 pesos)

	Uses ↓		Current account				Accumulation account					RoW	Total
	Sources →		**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	
Current account	Income and Outlay	**A**		20.619		23.129						648	44.396
	Institutionals sectors	**B**	42.925		3.463	523							46.911
	Goods and services	**C**		23.211		26.354	313	8.241				8.521	66.639
	Activities	**D**			54.129								54.129
Accumulation account	Capital	**E**		3.081		4.122					7.468	2.701	17.373
	Activities	**F**					8.241						8.241
	Financial instruments	**G**									11.679	3.134	14.813
	Adjustments and discrepancies	**H**									264	−264	0
	Institutionals sectors	**I**					6.118		13.294				19.411
Rest of the world		**J**	1.471		9.048		2.701		1.519				14.740
Total			44.396	46.911	66.639	54.129	17.373	8.241	14.813	0	19.411	14.740	286.652

3 *My acknowledgements to Gerardo Aceituno P. for his contribution to this and the next section (Aceituno, Venegas, 2001).*

Table 2 – Description of the MIEA modules

A,B	Outlay of institutional sectors	E,D	Consumption of fixed capital
A,D	Payments to production factors	E,I	Investment and net lending
A,J	Factors income from abroad	E,J	Net lending of the rest the world
B,A	Income of institutional sectors	F,E	Gross fixed capital formation. Institutional sect.
B,C	VAT, import duties and other taxes on products	G,I	Net increase of financial assets
B,D	Net taxes on production	G,J	Net increase of financial assets. Rest of the world
C,B	Final consumption expenditure		
C,D	Intermediate consumption of goods and services	H,I	Adjustment and discrepancies
		H,J	Adjustment and discrepancies
C,E	Changes in inventories	I,E	Financial resources for capital formation
C,F	Gross fixed capital formation	I,G	Net increase of liabilities
C,J	Exports of goods and services	J,A	Factor payments to the rest of the world
D,C	Output	J,C	Imports cif
E,B	Saving	J,E	Current external balance
		J,G	Net increase of liabilites. Rest of the world

The MIEA has been divided into 10 rows and columns A to J that present information in 27 modules. Each module is distinguished by its coordinates *(i,j)*. The content of the modules is described as follows:

The detailed MIEA resulting from the publication of the base year (Central Bank 2001) presents a breakdown into 7 institutional sectors (rows and columns B and I), 12 activities (rows and columns D and F), 13 national and imported goods and services (row and column C), 15 income instruments (row and column A), 10 financial instruments (row and column G) and another 13 capital and current account flows (including one of adjustments and discrepancies). This matrix brings together all the economic information on Chilean national accounts for the new base year 1996, comprising approximately 800 data that are the combination of analytically significant crosses of sectors, activities, goods and services and other flows already mentioned.

In table 3, three modifications have been made to the MIEA shown in the previous table:

Table 3 – Rearrangement and breakdown of the input-output table in the MIEA 1996
(billions of 1996 pesos)

		Income	Goods and services				Activities			Final consumption			Accumulation					Total
		A	C				D			B			E	F	G	I	J	
			Goods	Trade	Services	Others	Goods	Trade	Services	Private Sector	Government	Financial S.	Capital Form.	Activities	Financial Inst	Institut. Sectors	Rest o. World	
A							10.949	2.662	9.518	12.691	2.269	5.660					648	44.396
B	PS	34.641																34.641
	GO	2.743	688			2.775	91	125	307									6.729
	FS	5.542																5.542
C	1						11.147	720	2.814	6.822			272	7.534			6.475	35.784
	2						147	220	268	232							220	1.086
	3						2.842	1.846	4.339	7.939	3.426						1.504	21.895
	Oth.						1.082	62	867	4.792			41	707			322	7.874
D	1		27.136	9	207	465												27.817
	2		84	945	225	4.596												5.850
	3		18	5	20.401	38												20.462
E							1.559	214	2.349	2.165	1.034	−118				7.468	2.701	17.373
F													8.241					8.241
G																11.679	3.134	14.813
H																264	−264	0
I													6.118		13.294			19.411
J		1.471	7.857	128	1.063								2.701		1.519			14.740
Total		44.396	35.784	1.086	21.895	7.874	27.817	5.850	20.462	34.641	6.729	5.542	17.373	8.241	14.813	19.411	14.740	286.652

- the order of the modules has been changed to show the relations of the input-output tables.
- Products and activities have been opened up into goods, trade and services.
- Institutional sector has been opened up to private sector (SP), government (GO) and financial corporations (SOC).

Row D shows the cost structure of economic activities. The (C,D) module corresponds to intermediate goods and services consumption. The (A,D) module registers primary production income. Taxes on production are registered in the (B,D) module while fixed capital consumption is registered in (E,D). By excluding the production registered twice (intermediate consumption B,D) production net of duplications (GDP) can be obtained as follows:

Table 4 – GDP recording in the MIEA

Income approach			Production approach			Expenditure approach		
	Module (A,D)	23.129		Total columna (D)	54.129		Module (C,B)	23.211
+	Module (B,D)	523	−	Module (C,B)	26.354	+	Module (C,E)	313
+	Module (E,D)	4.122	+	Module (B,C)	3.463	+	Module (C,F)	8.241
+	Module (B,C)	3.463				+	Module (C,J)	8.521
						−	Module (J,C)	9.048
=	GDP	31.237	=	GDP *	31.237	=	GDP*	31.237

** Total GDP differences are caused by rounding errors.*

This structure allows us to review the rules of validation of GDP measurement in relation to income distribution and capital formation. As in the case of the rules of validation shown in the previous point, here we can review how GDP behaves in relation to other macroeconomic aggregates included in institutional accounts. These rules of validation can only be established for GDP at current prices due to the difficulty in breaking down the values of institutional account balances into price and quantity.

The MIEA in table 3 provides a register of the figures for variables commonly used in economic analysis. For example, it is possible to consider an economy with a goods and services market; the income redistribution process; the financial services market (money and other financial assets) and four agents (private, public, financial and foreign sectors).

One possible approach is to express the identity of GDP of table 4 in terms of the income and outlay flows of institutional sectors (Aceituno, Venegas, 2001) which would be familiar to national accountants.

Given that Y_g is government income, Y_p private sector, Y_f financial sector, Y_e factor payments abroad and, as counterpart, G_g is government expenditure, G_p is private expenditure and X_n goods and services exports. Economic policy hypotheses and macroeconomic models put forward some given relationships between the variables of financial system flows which are identifiable with the financial deficits and surpluses of institutional sectors. If the income and outlay components of each institutional sector are isolated, the income-outlay gaps can be calculated and how they relate to institutional sectors' net borrowing or loans. Thus the following expressions register these sectors' income and outlay gaps or budgetary restrictions.

$$D_g = Y_g - G_g \tag{1}$$

MIEA modules $(E, B_{GO}) = (B_{GO}, A) + (B_{GO}, C) + (B_{GO}, D) - (A, B_{GO}) - (C, B_{GO})$

Table 3 data $1.034 = 2.743 + 3.463 + 523 - 2.269 - 3.426$

Represents government budgetary restrictions

$$S_p = Y_p - G_p \tag{2}$$

MIEA modules $(E, B_{SP}) = (B_{SP}, A) - (A, B_{SP}) - (C, B_{SP})$

Table 3 data $2.165 = 34.641 - 12.691 - 19.785$

Corresponds to private sector budgetary restrictions

	S_f	$= Y_f - G_f$	(3)
MIEA modules	(E, B_{FS})	$= (B_{FS}, A) - (A, B_{FS})$	
Table 3 data	-118	$= 5.542 - 5.660$	

Refers to financial sector budgetary restrictions

	S_e	$= Y_{de} - X_n$	(4)
MIEA modules		$= (J, A) + (J, C) - (A, J) - (C, J)$	
Table 3 data	1.351	$= 1.471 + 9.048 - 648 - 8.521$	

Finally, corresponds to foreign sector surplus.

Income-outlay gaps accounted in this way close ex-post with fixed-capital investment in the savings-investment identity. Thus:

	$(D_g + S_p + S_f) + I_r + S_e = I_g + I_p + I_f$	(5)
MIEA modules	$(E, B_{GO}) + (E, B_{SP}) + (E, B_{SF}) + (E, D) + (E, J) = (C, E) + (C, F)$	
Table 3 data	$3.081 + 4.122 + 1.351 = 773 + 7.626 + 155$	
	$3.081 + 4.122 + 1.351 = 8.554$	

Where I_g, I_p and I_f stand for gross capital formation or investment of the government, private sector and financial sector respectively (see detailed in (E, I) table 5). Net investment is obtained by discounting fixed capital consumption (I_r). (i.e. 8,554 − 4,122 = 4,432).

Thus (5) segregates investment and its financing considering both domestic saving (3.081) and foreign saving (1.351, the nation's current account deficit).

This savings-investment process has its counterpart in the variations or flows of financial assets and liabilities registered in the institutional sector financial account. In terms of the MIEA, it is registered in the detail of the rows and columns of the modules G, H, I, J as shown in table 5.

Institutional sectors either finance or require funding for investment with the net variation of their financial assets and liabilities. For example, the government is a net lender according to (1) and (5) since its investment was less than its saving (net lending 564). This net positive flow is expressed in the net increase of money and credit assets in the government portfolio.

Table 5 – Breakdown of flow of funds in the MIEA

(billions of 1996 pesos)

			A a E	G					I			J	Total
				M	T	P	A	P-S	GO	FS	PS	RoW	
		A a C										9.168	
E	Net lending/net borrowing								564	−203	−1.447	1.351	
	Capital formation								773	155	7.626		
G	Currency and deposits	M							128	57	1.904	−25	2.064
	Securities	T							−1	1.924	−33	569	2.459
	Loans	P							230	1.873	1.816	456	4.376
	Shares and other equities	A							3	589	1.516	2.134	4.241
	Pensions funds, insurances	P-S								94	1.579		1.674
H	Adjustments and discrepancies								45	−42	261	−264	0
I	Government	GO	2.076		−2	−147		36					1.962
	Financial Sector	FS	−168	2.086	1.394	−873	380	1.638					4.458
	Private Sector	PS	4.210		452	5.098	3.232						12.992
J	Rest of the World	RoW	13.220	−23	615	298	629						14.740
Total				2.064	2.459	4.376	4.241	1.674	1.962	4.791	12.659	14.740	

In this detail we can appreciate the importance of the financial sector as an intermediate in the saving-investment process, unlike its marginal role in the production and distribution of income as registered in table 3.

Finally, for the economy as a whole the variations of consolidated assets and liabilities abroad reflect the net balance of resources that contribute to financing investment in the period.

$$[\Delta A_M + \Delta A_T + \Delta A_P + \Delta A_A + \Delta A_j] - [\Delta P_M + \Delta P_T + \Delta P_P + \Delta P_A] = S_e \qquad (6)$$

MIEA modules $[(J, G_M) + (J, G_T) + (J, G_P) + (J, G_A) + (H, J)] - [(G_M, J) + (G_T, J) + (G_P, J) + (G_A, J)] = (E, J)$

Table 5 data $[-23 + 615 + 298 + 629 + 264] - [-25 + 569 + 456 + 2.134] = -1351$

$1.783 - 3134 = -1351$

Where ΔA are asset variations, ΔP liability variations abroad in the form of money (M), securities (T), borrowing and loans (P) shares and capital participation (A) abroad.[4]

As we see, the MIEA integrates GDP measurements with the balance of sources and uses of the economy's financial funds. Detailed data can establish the role of the Central Bank, or commercial banks in the financial sector, as source of financing for other institutional sectors, thereby connecting the "real" flows and "financial" flows.

This integration presents a set of GDP consistencies at current prices with the income and financial flows that must be faced each year and which, among other relations, partly explains the differences between national income and GDP and the losses and gains for goods and services terms of trade.

3. Frequency and timeliness of GDP measurement

In Chile, before the National Accounts Department was transferred to the Central Bank in 1982, GDP was only measured once a year. Economic policy in the past two decades has made short-term analysis a fundamental requirement.

The formulation and execution of economic policy has reached a level of specialisation and complexity far different from that of the '60s and '70s. It was essential to take a close look at the mechanisms of transmission of economic policy decisions and the time taken for their effects to pass through.

In this context, macroeconomic management requires hypothesis on the evolution and its impact on activity, consumption and investment. To study these effects of economic policy relevant data must be gathered and interpreted. The quality of the hypothesis depends on data supply, especially if we consider the scope and depth of the econometric techniques applied which require a great variety of long-term statistical series.

In 1991, Chile – like other countries of the region (Peru in 1994, Mexico, Brazil and Colombia in 1999) – adopted the inflation target as the nominal anchor to ensure price stability. Inflation targeting establishes specific and measurable objectives which the Central Bank must be sure to meet. Thus a virtuous circle of transparency and responsibility is set up which lends credibility to monetary authorities' aims. Public aims become private sector expectations and the ample, open information available ensures that no economic agent has privileged information regarding the current economic situation and its outlook or the suppositions on which inflation targets and growth forecasts are grounded. Inflationary targeting has been strengthened since September 1999, favouring and improving short-term data collection and forecasting techniques within the Central Bank.

To further this task, the Central Bank of Chile has developed a series of instruments associated with GDP to measure economic activity: quarterly GDP (Central Bank, 1983), IMACEC (Venegas, 1985), Economic Expectations Survey (began in May 2004), Quarterly Stocks Survey (in trial period, to be first published in October 2004), Quarterly Consumer Goods Sales Survey (final results evaluation in trial period), Quarterly Accounts (currently being prepared to be first published in March 2005).

These measurements are compiled using data from the Balance of Payments, Finance Ministry (Public Finances), National Statistics Office and other public and private institutions and organisations that contribute with their own particular expertise to the national information system.

In short, macroeconomic policy uses, on one hand, all the available statistical indicators and, on the other hand, it considers the evolution of some macroeconomic aggregates that national accounts estimates each month or quarter.

The following table shows when national accounts and their outlook are published.

4 *There is an entry for adjustments and discrepancies* (A_j) *that measures the difference between net borrowing calculated below and above the line which was made explicit in national accounts base year 1996.*

Table 6 – GDP and national accounts. Frequency and timeliness of publishing
(Year n, quarter t)

	Version	Year/ quarter	Lag	Current prices Total	Current prices Breakdown	Constant prices Total	Constant prices Breakdown
Yearly accounts							
Production by activity	Preliminary	n-1	3 months			✓	✓
	Provisional	n-2	15 months	✓	✓	✓	✓
	Revised	n-3	27 months	✓	✓	✓	✓
Supply and use	Preliminary	n-1	3 months	Start:March 05		✓	
by product	Provisional	n-2	15 months	✓	✓	✓	✓
	Revised	n-3	27 months	✓	✓	✓	✓
Income and	Provisional	n-2	15 months	✓			
outlay account	Revised	n-3	27 months	✓	✓		
Capital account	Provisional	n-2	15 months	✓			
Capital and financial account	Revised	n-3	27 months	✓	✓		
Quarterly accounts							
Production by activity	Preliminary	t-1	2-2,5 months	Start:March 05	Start:March 05	✓	✓
	Provisional	t-1	15 months	Start:March 05	Start:March 05	✓	✓
	Revised	t-1	27 months	Start:March 05	Start:March 05	✓	✓
Supply and use	Preliminary	t-1	2-2,5 months	Start:March 05		✓	
by product	Provisional	t-1	15 months	Start:March 05		✓	
	Revised	t-1	27 months	Start:March 05		✓	

Except for the financial accounts, the main accounts at current and constant prices base year 1996 are published quarterly. They are aggregated except for production which is breakdown by activity at constant prices. At the moment, we are working on improving the aggregated measurements by segregating the main products and activities.

Later, it is planned to have quarterly series for the period 1996–2004, breakdown production by activity at current prices and the tables of supply and use by products at current and constant prices of 1996. Anyway these breakdown will be only use for balancing purpose.

This process of breakdown within the framework of considering the base year in quarters will make future quarterly results more robust and consistent since they will be the result of balancing supply-use, similar to what is currently done with annual accounts, but applied to a smaller number of products. We are improving the source of the figures by making better use of registers of exports and imports, VAT and income. We also intend to strengthen the measurement of quarterly home consumption by incorporating information provided by surveys of stocks and goods and services at point of sale which, moreover, will contribute to improve the measurement of household consumption.

Balancing data in the short term

Short-term measurements are biased to a greater or lesser extent according to two factors:

- The proximity of the quarter being calculated. For example, figures produced two months after the quarter in question will be less reliable than after four months when government and financial sector figures and the corrected quarterly balance of payments figures are available. And so on. In time, quarterly financial statements of private and state-owned corporations might even be incorporated.
- The proximity of the definitive annual national accounts. Integrated accounts for the quarter can only be consolidated with the publication of the annual accounts when distribution coefficients can be determined, real (for goods and services) and nominal (for income and financial instruments). These corrections, that have to be incorporated to the previous four quarters or more, make it possible to improve measurements for the following quarters and their projection.

Thus to balance data is to find the consistency between the sources and uses at the level of each institution and by activity. As far as possible, any differences should be resolved by distributing them between the agents, weighing up the main data or complementary data from different sources or by simple criteria of choice according to previous experience. Finally, any differences that cannot be resolved in this way have to be assigned to "adjustments and discrepancies" (module, row and column H in the MIEA).

Adjustments and discrepancies are normal in short-term measurements, it would be more worrying if there were none. They are first detected in the annual accounts. For example, in base year 1996 there is a discrepancy of some 2.2% of GDP in net household borrowing (see table 5, adjustments and discrepancies) which, given the relatively sure information on financial flows, can only be attributed to a deficit in non-detected income. The difference may be due to "underground or illegal income, underestimation of small enterprises not constituted as corporations, under-declaration of income on income tax declarations, income from abroad that has not been registered, or other such causes" (Central Bank, 1991, p. 84). These unavoidable discrepancies in annual calculations are accentuated in short-term calculations by seasonal and irregular factors that affect the different markets, the impact of terms of trade variations and all the capital gains and losses that affect the figures over and under the line of current and capital measurements.

Different versions of data

As we see, short-term accounts are subject to correction as new data becomes available and that is why the concept of a version of a series associated to a date of issue is so important.

Ideally the information platform where short-term calculations are made and stored should be sufficiently flexible to allow new data and recalculations to be incorporated. However, apart from the work that is going on permanently, specific versions must be published, as table 6 shows these are:

- First version. Preliminary, corresponds to the most timely date of issue for macroeconomic analysis. Recommendable at 45 days from the generation of the main data on goods and services, government and foreign trade. Later, probably about 90 days after the end of the quarter, this version can be corrected once more or less complete short-term balance of payments, public finances and financial system data becomes available.
- Second version. Provisional, adjusted to the preliminary annual accounts, issued by the end of the first quarter of the following year.
- Third version. Revised, adjusted to the revised annual accounts, issued two years after.

Each version thus loses on timeliness but improves on the previous version.

4. Redefinition of activities and products in the new GDP measurements

Changes in the world economy are felt not only in production technology but also in the definition of products and activities.

These changes have really upset traditional national accounts definitions and methods with their "recommended classifications" of activities (ISIC Rev.3) and products (CPC Rev. 1.1). In a way, the national accounts perspective is still "Taylorian" in that it continues with a sequencial or analytical approach instead of an integral or systemic approach.

In the specific case of classifications, the traditional view is of "flat" self-sufficient classifications whereas they should be seen as interrelated networks of classification that respond to different analytical perspectives. The concept of traditional activity implies a typification that today is far too rigid. How can we speak of the "agroindustry" or "fishing" when they clearly cover primary activity and manufacturing and where even the "raw material" loses all reference to the market (having almost no price value) and becomes a mere manufactured product?

Products are also ever less standardised (or like assembly line mass production of Taylor's industrial system) and become differentiated according to the customer's individual tastes. Products that are differentiated by the added value of service incorporated into the market price (De Bono, 1994).

Both elements, the technological complexity of production and the products with post-production value added, are part of a reciprocal cause phenomenon.

By way of illustration, in the Chilean experience, changes can be made in production functions that imply, on one hand, vertical integration and "loss of market signals" of primary products and, on the other hand, a segregation into layers of service in an industry that originally was integrated.

In all cases, it is essential to establish the difference between "products" and "activities". Software is not the same as the activity of producing software. The activity of producing software has to do with industrial organisation whereas the product software has its own dynamics of modernisation and producer and consumer markets.

Case of vertical integration

Primary activity and manufacturing activity are becoming increasingly integrated, making it difficult to put a value on production and to determine the added value of the primary products. This integration is being brought about by heavy investment in large-scale manufacturing industries which are driving out the small-scale suppliers of raw materials.

An example in Chile is the poultry and pig farming industry in which a few large corporations now cover the entire productive cycle from breeding and raising the animals, (which is the traditional livestock activity) production of animal feed, slaughter (which corresponds to the traditional maunfacturing industry) right through to distribution to retailers, including the transport and marketing. This leaves no room for the small producer and consequently there is no market price reference for the primary production of live animals. The industry's added value and thus its production surplus is integrated from start to finish and today the main surplus is generated in the manufacturing phase (slaughter and meat preparation), the commercial phase (packing and supplying supermarkets and exportation) and the transport. In this way, the value of the live animals, even as a cost item, is very low, because of economy of scope and scale, and impossible for an independent producer to compete. As we see, it is not possible to state the price of these primary products independently of the final consumer product of which the main added value is that of the industrial services that have gone into its commercialisation. Even if there were a producer selling live animals, it would not reflect the real market price because of the absence of competition.

The Chilean wine industry, currently enjoying an export boom, is experiencing a similar situation. It is very difficult to register the value of grapes for wine, especially in the case of high quality wine types. Wines may be considered in different categories: ordinary, premium, super-premium and varietal, so, to estimate the production and added value, we have to determine what amount of each kind of grape is destined to each class of bottled wine and what is its reference price. These prices are very difficult to determine since there are no homogeneous categories of grape destined to each type of industrial product and thus no market information for this production.

Similar cases are fishing and mining except that here, these industries buy part of their production from independent producers so there is a market valuation. Nevertheless, the referential value is arguable because of the quality or "mixtures" or "impurities" of the bought raw material. In the case of copper mining in Chile, national accounts have preferred to integrate the primary activity (extraction) with the manufacturing activity (refining) because of analytic advantages of keeping a traditionally key sector of the economy integrated and at the same time eliminating the problem of the valuation of the primary product. The challenge remains of progressing towards treating the two activities separately in national accounts, even if only as a complement to the traditional presentation.

Attempts are to be made to solve these problems, in particular those referring to livestock and grapes for wine, by giving a value to the primary production. Proxy prices will be used, based on international market prices if they exist, or on prices estimated by establishing activity cost that will divide the production surplus pro rata among the products integrated into the industrial chain.

Case of electric power

This is an almost homogeneous product, but coming from a diversity of sources with varying industrial organisation, functions and production cycle.

The first distinction corresponds to the cycle of supply and use of electricity. Generation, strictly speaking, corresponds to electricity production. Transmission refers to the transformation of high power electricity to the low power network for residential and non-residential use. Distribution is the transmission of the transformed electricity to the specific users.

The second distinction corresponds to the energy sources currently used and those of incipient or potential development. With regard to this point, two aspects should be reviewed: the origin and cost of the resource. Chile has abundant water resources and hydroelectricity has the lowest production cost and represents the largest share of supply. On the contrary, the resources used to generate thermoelectricity, natural gas, coal and oil, are today mainly imported, although there is some domestic availability in each of the three cases. The order – gas, coal, oil – reflects the rising production cost of electricity obtained from each of these sources with oil now reaching the point of being considered an "inefficient" source.

This industry has undergone considerable organisational change over the past twenty years. During the '80s it was concentrated in a few great generation and distribution corporations whereas today there are dozens of corporations specialised in one source or another or thermoelectric corporations diversified between several sources.

The main problem when calculating GDP lies in separating the production structure of each power plant of each of these diversified corporations. Their accounting information does not distinguish activity costs, except for their main input.

This separation would be irrelevant if the production structure by type of source or input were constant but, on the contrary, the source varies greatly depending on rainfall. Thus, calculating GDP faces a second problem. In dry years, the inputs used are more inefficient or expensive, producing a strong impact on GDP, even though gross electricity production remains the same. This change can be reflected at current prices with the available information but, to measure GDP at constant prices, the cost structure for each input should be known and therefore we have to separate the costs for each of the power plants of each of the diversified corporations. This is a challenge now facing us with the new base year.

In the future, the energy matrix is expected to be wider and incorporate renewable resources. Chile is expected to be less self sufficient with regard to electricity supplies and to seek greater international integration. Moreover, industrial organisation is likely to be affected by changes in ownership and productive specialisation. All these factors present more challenges to the calculation of GDP in the coming years.

Case of information technology (IT)

Contrary to the case of electric power, IT products are tremendously varied, they change year by year and have a great diversity of prices.

There are two main types of information products: hardware and software. In software there is a generally-used subclassification that situates products according to their increasing cost and decreasing demand. Office software (spreadsheets, word processors) is the most massively used and the cheapest. At the other extreme, management software, such as Enterprise Resource Management ERM or Customer Relationship Management CRM, is expensive and of more specialised use in large corporations or designed for specific customer needs. Prices range from US$ 500 to US$ 60,000 or even more.

On the other hand, we have the value chain of the IT product as regards the most significant aspects for national accounts. First, we must consider that there is an intangible capital production which, apart from the physical product (disk, tape or other media, including virtual media), generates a flow of ownership rights implicit in the user's purchase price. The value chain includes reproduction or cloning of the original, its commercialisation, maintenance or technical assistance that includes upgrading, its distribution as a consumer article or capital formation, exports and imports, stock variation, consumption and loss of capital. All these items have varying degrees of importance in each type of software and generate diverse types of problems on the national accounts register. We will consider these problems of the value chain, distinguishing general problems and specific problems.

General problems

a) Constant loss of value of software as it rapidly becomes technologically obsolescent.
b) Prices going down, even for improved products, due to greater competition and the faster amortisation of investment involved in launching a new product, mostly paid off in the first year of production.
c) There are no references between periods that allow the evolution of market prices of a homogeneous product defined by class, model and uses.
d) The value chain is irregular and its individual components are hard to separate out to make a cost structure.

Specific problems

e) Intellectual property. Software has intellectual value so the original or master is irrelevant. In the first year of sale the value of the software is incorporated into gross capital formation. In the following years it generates a right (income and outlay account) which for mass software is totally recovered but not in the case of specialised software.

f) Reproduction and commercialisation. Reproduction belongs to the printing and publishing industry but the problem in Chile is that apart from the formal industry there is an illegal activity which is very difficult to measure because current technology makes copying so easy.
g) Technical assistance is mainly given for design and management software but, like franchising, the problem is that the cost of support is normally mixed with copyright payments mentioned in e).
h) Use of software. Office and educational or entertainment software in corporations is mainly classed as current expenses and other types as capital formation. It is difficult to integrate both types of software going on accounting information.
i) Foreign trade. Most software in Chile is imported. Unfortunately, the import tariff code is not enough to make a distinction between the disk itself and its intellectual content and even a distinction between other kinds of recordings.
j) Chile exports specialised software on mining and electric power which, like all exported services, do not necessarily cover the total amount exported, even though there are tax benefits for this kind of exports.
k) Variations in the capital in software. If we take the initial value of software stocks and compare it with the final value we will find a great variation, beyond the year's investment.

The points mentioned pose a challenge to national accounts even at the level of corporations. There are no specific methods or data to allow this increasingly important capital component to be measured.

5. New analytical perspectives and data management challenges

GDP measurement is facing an ever more complex situation requiring new specialised studies that, on one hand, open up the traditional national accounts classifications in new directions and, on the other hand, require yet more detailed calculations of costs, production, income and flow-of-funds.

There is no adequate consistent response to this challenge as regards national accounts data systems. Each new field of analysis that opens implies particular solutions in terms of databases, processing and analytical tables that are out of the scope of the core of the system. Existing systems offer "corporative" solutions mounted on top of relational databases or local solutions based on commercial software such as Microsoft Office using Access or Excel.

These solutions lack integration and articulation between the various parts. That is, there is no flow of solutions to processing problems, for example, between accounts data from the rest of the world, economic surveys and production accounts, administrative registers. In general, what is lacking is a transversal focus of input and output information. Moreover, these solutions are at the core of national accounts and leave little room for the introduction of satellite accounts or regional or short-term accounts. In short, there is a gap between a highly developed meticulous conceptual accounts system framework and the practical work of data processing. This gap reflects the lack of international guidelines on database modelling that ought to accompany the various manuals or account and statistics plans.

The following function, used in a paper on satellite accounts modelling (Venegas, 2004) serves to develop this point. It can be applied by analogy to various GDP and national accounts items in different analytical dimensions. According to that, analytical or statistical charts or account plans can be defined as products of an information system, not as its organising axes. If these products are defined as a demand or requirement d to a system, then we have

$$d = c(X, I, J, P, T, E, V, M, G, R, F)$$

Where:
d = any requirement (account, chart, indicator) to the information system
c = combination of transaction registering elements
X = economic transaction
I = data network of the transaction object
J = data network of the transaction agent
P = patrimonial position of the transaction object
T = dates, frequency and other time references of the agent, transaction or object
E = country, region, commune and other geographical references of the agent, transaction or object
V = Valuation
M = units of measurement of currency and volume

G = aim or purpose of the transaction
R = type of register or data: price, quantity, value, primary datum, index, parameter
F = date and data source

To clarify the role of the informative outputs, it is worth centring on the concept of accounts plan. This concept can be extended to any systematic body of statistical or accounting information.

A panorama of these "account plans" that result from current economic information systems is presented in C. Carson (2001). Inspired by that focus, the integration base of the manuals most directly related to national accounts and GDP measurement could be illustrated according to the table 7.

The first part of the table registers the manuals that constitute "account plans". The columns of agents, objects and transactions establish the coordinates for compiling information.

Table 7 – Some accounts plans and economic statistics

Manuals	Agents		Objects	SNA position
	Institutionals	Activities		
System of National Accounts 1993	All	All	All	Central
Households	Households	Non	All	Complementary
Non Observed Economy	All	All	All	Complementary
Monetary and Financial	Financial corporations	Non	Financial instruments	Related
Non Profit Institutions	Non profit institutions	All	All	Satellite
Government Finance	Government	Non	All	Related
Balance of Payments	Rest of the world	Non	All	Related
System of Health Accounts	All	Health	All	Satellite
Capital Stocks	All	All	Assets (capital goods)	Complementary
Tourism	Tourism characteristics	Tourism characteristics	Tourism characteristics	Satellite
Environmental System	Environmental character.	Environmental character.	Environmental character.	Satellite
Classifications				
ISIC Rev.3		All activities		Central
CPC Rev. 1.1			All goods and services	Central
SITC Rev.3			All goods	CCP related
BEC			All goods	CCP related
HCO2			All goods	CCP related
COICOP	Households			Complementary
COFOG (functions)	Goverment			Complementary
COPNI (functions)	Non profit institutions			Complementary
COPP	All	All activities		Complementary

If the economic data generated in a country refer to transactions *xij* that relate information from various analytical domains it is clear that in the database of said plans there are several relations that imply: redundancy, belonging, breakdown and cross classification. We can define a sort of x_{ij} database where account plans are combinations of elements $d = 1,2,..., n$ that presents the following characteristics: they are not exclusive, they can relate just as if they were sets of elements in terms of intersections, unions, belonging to and the elements themselves can relate one to one, one to several, or several to several, among said sets.

Each account plan *d* is defined in function of a need, be it of analysis, control or programming. If it really has these characteristics, then what happened in the past that a specific $d = 1,2,..., n$ need for programming, analysis or control defined an "information system", this is the negation of the real concept of system. It is impossible to imagine *n* systems in the field of public information management without there being redundancies, interruptions, discontinuities, lack of standardisation and, in the end, inefficient use of professional, technological and informational resources.

Conclusions

GDP measurement today faces a series of distinct challenges:

Validation rules

There is a need for internal consistency as regards criteria for valuation, sectoral, temporal and geographic dimension of data and the new satellite analytical dimensions that break with the traditional classification into activities and goods and services.

There is a need for external consistency where GDP measurements should form part of a set of integrated accounts related not only with income distribution and use but also with the stock and formation of capital.

Technology

Marked changes in ways of production and in the products themselves, far removed from the old form of homogeneous, sequential production, generate important changes in technical coefficients of production, division of labour, distribution of goods and services, all of which must be registered in GDP.

Data handling

Considering the statistical requirements that must be registered, there has to be a fundamental change in the focus of data processing, incorporating relational database solutions based on integrated data modelling. The relation database-enquiries (analytical tables and accounts) will allow us to be increasingly more timely and flexible in meeting data requirements.

References

Aceituno, G. y J.Venegas (2001). “Serie de Cuentas Económicas Integradas de Corto Plazo: Una aplicación para Chile”, Documento de Trabajo, Seminario Regional sobre los Usos de las Cuentas Nacionales, United Nations Statistical Division, Comisión Económica para America Latina y el Caribe, Banco Central de Chile, Santiago 12–15 de noviembre 2001.

Banco Central de Chile (2004). “Anuario de Cuentas Nacionales 2003”, Santiago.

Banco Central de Chile (2003). “Anuario de Cuentas Nacionales 2002”, Santiago.

Banco Central de Chile (2002). “Anuario de Cuentas Nacionales 2001”, Santiago.

Banco Central de Chile (2001). “Matriz de Insumo-Producto de la Economía Chilena 1996”, Santiago.

Banco Central de Chile (2000). “Anuario de Cuentas Nacionales 1999”, Santiago.

Banco Central de Chile (1999). “Anuario de Cuentas Nacionales 1998”, Santiago.

Banco Central de Chile (1998). “Anuario de Cuentas Nacionales 1997”, Santiago.

Banco Central de Chile (1995). “Cuentas Financieras de la Economía Chilena 1986–1990”, Santiago.

Banco Central de Chile (1994). “Cuentas Nacionales de Chile 1985–1992”, Santiago.

Banco Central de Chile (1992). “Matriz de Insumo-Producto para la Economía Chilena 1986”, Santiago.

Banco Central de Chile (1985). “Cuentas Nacionales de Chile 1974–1985”, Santiago.

Banco Central de Chile (1984). “Cuentas Nacionales de Chile 1960–1983”, Santiago.

Banco Central de Chile (1983). “Cuentas Nacionales de Chile 1960–1982”, Santiago.

Banco Central de Chile (1983). “Cuentas Nacionales Trimestralizadas, 1980–1983”, Santiago.

Banco Central de Chile (1982) “Cuentas Nacionales de Chile 1960–1980”, Santiago.

Banco Central de Chile (2004) “Producto Interno Bruto Regional 1996–2000”, *Boletín Mensual*, Santiago, enero 2004.

Banco Central de Chile (2004). “Producto Interno Bruto Regional 1996–2001”, *Boletín Mensual*, Santiago, julio 2004.

Carson, C., L. Laliberté (2001). “Manuals on Macroeconomic Statistics. A Stocktakingto Guide Future Work”, *IMF Working Paper, WP/01/183*, New York.

CORFO. Corporación de Fomento de la Producción (1964). “Cuentas Nacionales de Chile 1958–1963”, Santiago.

CORFO. Corporación de Fomento de la Producción (1962). “Cuentas Nacionales de Chile 1940–1962”, Santiago.

CORFO. Corporación de Fomento de la Producción (1960). “Cuentas Nacionales de Chile 1950–1960”, Santiago.

CORFO. Corporación de Fomento de la Producción (1957). “Cuentas Nacionales de Chile 1940–1954”, Editorial del Pacífico, Santiago.

CORFO. Corporación de Fomento de la Producción (1946). “Renta Nacional 1940–1945”, Imprenta Universitaria, Santiago.

De Bono, E. (1994). “Más allá de la competencia”, Paidos, Buenos Aires.

EUROSTAT, Comisión Europea (1996). “Sistema Europeo de Cuentas, SEC 1995” Luxemburgo.

Mamalakis, M.J. (1978). “Historical Statistics of Chile. National Accounts”, Volume 1. Greenwood Press, Westport, CT; London, England.

Fontela, E., A. Lopez, A. Pulido (2000). “Structural Comparison of Input-Output Tables”, XIII International Conference on Input-Output Techniques de la International Input-Output Association, University of Macerata, Italia.

ODEPLAN. Oficina Nacional de Planificación Nacional (1986). “Metodología y Series Regionalizadas del PGB”. Mimeo. Santiago.

ODEPLAN. Oficina Nacional de Planificación Nacional (1981). “Metodología y Serie de Cuentas Nacionales 1974–1980”. Mimeo. Santiago.

ODEPLAN. Oficina Nacional de Planificación Nacional (1981). “Matriz de Insumo-Producto de la Economía Chilena 1977”, Santiago.

ODEPLAN. Oficina Nacional de Planificación Nacional (1978). “Regionalización Gasto del Producto Geográfico Bruto a Precios Constantes del año 1965. Serie 1960–1976”, Mimeo. Santiago.

ODEPLAN. Oficina Nacional de Planificación Nacional (1978). “Cuentas Nacionales de Chile 1972–1977”. Mimeo. Santiago.

ODEPLAN. Oficina Nacional de Planificación Nacional (1977). “Cuentas Nacionales de Chile 1960–1975”. Mimeo. Santiago.

ODEPLAN. Oficina Nacional de Planificación Nacional (1976). “Cuentas Nacionales de Chile 1965–1973”. Mimeo. Santiago.

ODEPLAN. Oficina Nacional de Planificación Nacional (1975). “Cuentas Nacionales de Chile 1960–1971”. Mimeo. Santiago.

ODEPLAN. Oficina Nacional de Planificación Nacional (1972). “Cuentas Nacionales de Chile 1965–1972”. Mimeo. Santiago.

ODEPLAN. Oficina Nacional de Planificación Nacional (1972). “Distribución del Ingreso y Cuentas de Producción 1960–1971”. Mimeo. Santiago.

ODEPLAN. Oficina Nacional de Planificación Nacional (1970). “Distribución del Ingreso y Cuentas de Producción 1960–1970”. Mimeo. Santiago.

ODEPLAN. Oficina Nacional de Planificación Nacional (1968). “Cuadro de Transacciones Intersectoriales para la Economía Chilena 1962”, Santiago.

ODEPLAN. Oficina Nacional de Planificación Nacional Banco Central de Chile (1967). “Estudio de Fuentes y Usos de Fondos Chile 1960–1964”, Editorial del Pacífico, Santiago.

ODEPLAN. Oficina Nacional de Planificación Nacional (1967). “Cuentas Nacionales de Chile 1960–1966”. Mimeo. Santiago.

ODEPLAN. Oficina Nacional de Planificación Nacional (1965). “Cuentas Nacionales de Chile 1959–1961”. Mimeo. Santiago.

Naciones Unidas (2000), “Manual sobre la compilación y análisis de los cuadros de insumo-producto”, División de Estadística, *Estudios de Métodos*, Serie F, Nº 74, Nueva York.

United Nations, "National Accounts: A Practical Introduction" (2003). Handbook of National Accounting, Department for Economic and Social Affairs, Statistics Division, *Studies in Methods*, Series F, N° 85, ST/ESA/STAT/SER.F/85, New York.

United Nations, "A System Approach to National Account Compilation. A technical report" (1998). Handbook of National Accounting, Department for Economic and Social Affairs, Statistics Division, *Studies in Methods*, Series F, N° 77, ST/ESA/STAT/SER.F/77, New York.

United Nations, Commision of the European Communities, International Monetary Fund, Organisation for Economic Cooperation and Development, World Bank (1993). "System of National Accounts 1993", Brussels, Luxembourg, New York, París, Washingnton D.C.

Venegas J. (1998). "Introducción al Programa de Nuevo Año Base de las Cuentas Nacionales de Chile", Documento de Trabajo Programa Nuevo Año Base de Cuentas Nacionales, Santiago.

Venegas J. (1995). "Matriz de Cuentas Sociales 1986: Una SAM para Chile", Serie *de Estudios Económicos* N° 39, Banco Central de Chile, Santiago.

Venegas, J. (1995). "The use of main frame and micro computer software for national accounts compilations (Issues for implementing the 1993 SNA)", Interregional Seminar on the 1993 SNA in Concept and Practice, United Nations Statistical Division/DESIPA, ESA/STA/AC.49/19/A, enero de 1995.

Venegas, J. (1994). "Elementos para un Sistema de Información Macroeconómica Industrial", Instituto Nacional de Estadísticas, *Revista Estadística & Economía*, N° 8, Santiago, junio de 1994.

Venegas J. (1986). "Indicador Mensual de Actividad Económica (IMACEC). Metodología y Serie 1982–1986", *Serie de Estudios Económicos*, N° 28, Banco Central de Chile, Santiago.

Abstract

This paper describes the Chilean experience of calculating GDP as an indicator of national accounts and examines the main methodological challenges encountered in the complex framework of integrated accounts with a growing set of rules of validation and with the redefinition of activities and products that has evolved over the past decade. We discuss operational problems such as restrictions on the timeliness and frequency of publication of GDP measurements and the need to incorporate technological advances in relational database design and data management. These points are analysed in the context of the compilation of the new series of national accounts using 2003 as the base year.

José Venegas Morales
Departamento de Cuentas Nacionales
Gerencia de Información e Investigación Estadística
Banco Central de Chile

Agosto 2004

Constructing a composite leading indicator for the Turkish economic activity and forecasting its turning points

Aslıhan Atabek, Evren Erdoğan Coşar and Saygin Şahinöz
(Central Bank of Turkey)

1. Introduction

Economic growth is the main determinant of the robustness and prosperity in the economy. For this reason, the issue of economic growth has long been a central concern of nations. There are many theories and studies concerning characteristics, sources and the pattern of the economic growth. In most of the studies, it is observed that economic activity exhibits two types of fluctuations. One of them is the upward trend, indicating long-term changes and the other short-term oscillations, representing temporary changes. There has always been an interest in measuring long-term trends in economic growth itself partly because analysts of business cycles are interested in measuring deviations from long-term trends. Business cycles can be defined as the fluctuations of economic time series around the long-run trend value, after the seasonal fluctuations have been removed. Early detection of business cycle turning points has always been a major concern to policy makers, businessmen and investors. Clearly, early recognition would allow them to trigger countercyclical policy measures. There exists an extensive literature, which attempts to find reliable forecasting tools for business cycle turning points, from the early landmark study by Burns and Mitchell (1946) to the more sophisticated study of Stock and Watson (1989).

An efficient way to predict business cycle turning points is to use leading indicators. Leading indicators are data series that tend to lead business activity. However, experience in many countries have shown that, it is not reliable to use just one economic indicator for short term forecasting because some leading series may produce false signals of future changes. In order to provide a more comprehensive measure of economic activity, composite leading indicators have been developed in many countries. The composite leading indicators are based on a basket of economic indicators, which show a leading relationship with the economic activity. The composite leading indicator (CLI) enables government and businesses to track the economy's performance and forecast this performance over the near term.

There are several analyses about the cyclical movements in the Turkish economy, and several studies have constructed a composite leading indicator index, some of which are Ozatay (1986), Altay et al. (1991), Neftci and Ozmucur (1991), Canakcı (1992), Selcuk (1994), Ucer et al. (1998), Kucukciftci and Senesen (1998), Murutoglu (1999) and Alper (2000).

This paper presents the results of the joint work of the CBRT and OECD on the construction of a new composite leading indicator for the Turkish economic activity. Following the OECD Composite Leading Indicators System, growth cycle approach is used that means the trend and the seasonal components are removed from the series. As the seasonal adjustment methodology, TRAMO/SEATS technique is used. And the trend components of the series are eliminated using the Phase-Average-Trend (PAT) method developed by the United States NBER. The composite leading indicator (CLI) consists of seven leading indicators covering the demand, supply, and policy variables. As a further study the turning points of the constructed CLI are estimated using the Neftci's (1982) sequential probability algorithm. The paper is organized as follows. Section 2 discusses and proposes an economic activity indicator to be used as a reference series and describes the candidate leading indicators that are to be brought in relation with it. Section 3 sets out the cyclical characterization of the reference series and of the candidate leading indicators considered and tackles the construction of the composite leading indicator and its use as an economic activity indicator. Section 4 gives a brief description of the Neftci's sequential probability algorithm and Section 5 presents the results of the estimated turning point probabilities. Finally, the main conclusions of this study are drawn in Section 6.

2. The reference series and the candidate leading indicators

The preliminary step in the composite leading indicator approach is to choose a proxy for the economic activity, which is called the reference series. Generally, Gross Domestic Product or Industrial Production Index is used as a measure of economic activity. In the composite leading indicator approach, a series that is available at high frequency and published with less delay is preferred as the reference series. Gross Domestic Product (GDP) is available on a quarterly basis and it is published about one quarter after the quarter to which it refers. For this reason, in this study, Industrial Production Index (IPI) is chosen as the reference series, which has the advantage of being a monthly reported variable with 5 weeks delay and its turning points being in line with those of the Gross Domestic Product. Since the turning points of IPI are not too different from those of GDP, the cyclical component of IPI is considered as a good proxy for the fluctuation of the overall economic activity. Besides, in the OECD CLI system the Industrial Production Index is used as the reference series for most of the countries.

In some of the empirical works like Stock and Watson (1989), a coincident economic indicators index is constructed and used as the reference series. The main reason behind this approach is the idea that the reference cycle is best measured by looking at co-movements across several aggregate time series. In this approach, the series that cover other sectors of economic activity rather than manufacturing (like agriculture or service sector) and other macroeconomic variables like sales and employment are aggregated in one index. However in Turkey, no regularly published data on a monthly basis is available related to sales, consumption or labor statistics (like wage or employment). Therefore, in line with the OECD system, the IPI is chosen as the reference series.

A useful leading indicator of economic activity should have the following properties: First, the series must be easily and quickly available and not subject to major revisions that would change earlier conclusions based on them. Second, the cyclical movements in the indicator should precede the reference series with a predictable relationship. Third, the lead time of the indicator must be long enough to give policy makers time to react. A detailed information on the selection criteria can be found in Nilsson (2000). Besides these statistical properties, the indicators should also have economic significance. As given in Nilsson (2000), in the OECD CLI system the candidate leading indicators are classified according to five types of economic reasoning. The indicators are classified into the following categories:

- early stage indicators; indicators which measure an early stage of production like new orders, order books, new company formation, etc.
- rapidly responsive indicators; indicators which respond rapidly to changes in economic activity such as average hours worked, stocks, etc.
- expectation-sensitive indicators; indicators which measure expectations or are sensitive to expectations such as stock prices, expectations based on business survey data about production, etc.
- prime movers; indicators which are relevant to monetary and fiscal policies and foreign economic developments such as money supply, terms of trade, indicators for foreign countries, etc.
- other indicators; indicators such as interest rates, motor vehicle registration, price indices, capacity utilization, employment, etc.

In accordance with this classification a data base covering all sectors of economy is constructed and the leading performance of the series in this data base are examined.

For the comparison between the reference series and the candidate leading indicators, the cyclical components of the series are used. The cycles of the series are obtained in the OECD CLI system framework.

3. The cyclical properties of the reference series and the candidate leading indicators

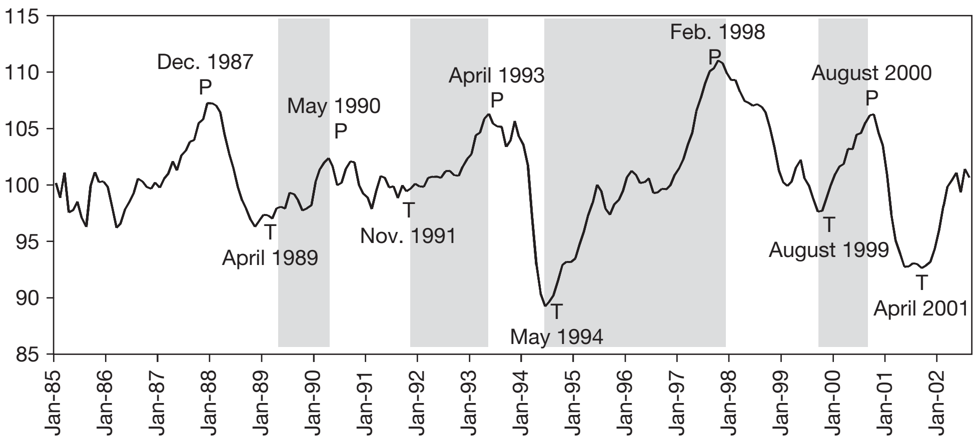
All the series in the analyses are seasonally adjusted and detrended following the TRAMO/SEATS and PAT techniques, respectively.

The first step of the analysis is the determination of the cyclical pattern and the turning points of the reference series which is called the reference chronology. The turning points of the IPI are given in Table 1. Figure 1 plots the reference series along with the cycles and the turning points identified. The shaded areas represent the decreasing periods, and the unshaded areas represent the increasing periods of the reference series.

Table 1 – Reference chronology of industrial production index

Turning points		Duration in months	
Trough	Peak	Deceleration	Acceleration
–	December 1987	–	–
April 1989	May 1990	16	13
November 1991	April 1993	18	17
May 1994	February 1998	13	45
August 1999	August 2000	18	12
April 2001	–	8	–
Mean		14.6	21.8
Median		16.0	15.0

Figure 1 – Cycles of industrial production index



In the OECD CLI system the performance of the series are evaluated according to the criteria given in Section 2 as well as some statistical properties such as: mean and the median average lead of the indicators over the reference cycle at turning points; the standard deviation of the median lead times at turning points; missing or extra cycles in the indicator series with respect to the reference series; smoothness of the candidate leading indicators (small MCD value) and cross-correlation of the candidate leading indicators at different lead lengths. From selected potential leading indicators several CLIs are constructed and out of the constructed composite leading indicators, the series that shows the best performance is chosen as the final CLI. The details of the criteria used for choosing the best composite leading indicator can be found in Nilsson (2000).

The final CLI chosen according to these criteria consists of the following series:

- Production Amount of Electricity
- Discounted Treasury Auctions Interest Rate Weighted by the Amount Sold
- Import of Intermediate Goods
- CBRT Business Tendency Survey (BTS) – Stocks of Finished Goods
- CBRT BTS – Amount of New Orders Received From Domestic Market
- CBRT BTS – Export Possibilities
- CBRT BTS – Employment.

Import of intermediate goods, electricity production and expectations about employment are all supply side indicators. Import of intermediate goods is an important indicator for IPI, since intermediate goods are critical inputs for production. Electricity production is another plausible supply indicator, since approximately 50 percent of the electricity produced is used by industry. The importance of employment in economic activity is obvious; if more is to be produced, more workers have to be used in production. Hence, expectations about the changes in employment reflect the expectations in the output. In Turkey, there is no reliable quantitative demand variable, so expectations are used for this reason. Expectations about export possibilities, stocks and new orders represent demand variables. The discounted treasury auctions interest rate is a policy indicator. This variable represents the cost of production, and has a countercyclical relationship with IPI. The discounted treasury auctions interest rate is preferred to the interest rates of three, six and twelve-month time deposits since it has a longer lead period.

Figure 2 gives the cyclical pattern of the constructed composite leading indicator and IPI. As it can be seen from the figure, CLI is anticipating IPI with similar cyclical pattern. Cyclical profile of CLI reveals that turning points at troughs are much more sharper than the turning points at peaks.

The turning points and the statistical properties of the CLI are given in Tables 2 and 3, respectively. An extra cycle, which is not observed in the reference series, exists in the CLI between July 1995 and November 1996. Since this cycle is not observed in the GDP and IPI and it has minor amplitude, it is considered as an extra cycle.

Figure 2 – Cycles of the CLI

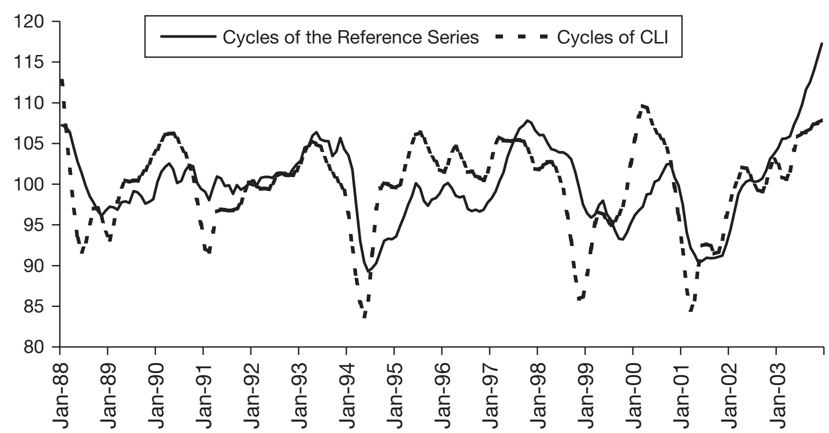


Table 2 – Turning points of the CLI

Turning points		Lead in months	
Trough	Peak	Trough	Peak
June 1988	April 1990	10	1
February 1991	April 1993	9	0
May 1994	July 1995	0	–
November 1996	August 1997	–	6
November 1998	April 2000	9	4
March 2001	–	1	–

Table 3 – Statistical properties of the CLI

	Extra or missing cycles	MCD	Mean lead (+) at turning points (TP)			Median lead (+) at turning points (TP)			Standard deviation	Cross correlation	
			Peak	Trough	All TP	Peak	Trough	All TP		Lead (+)	Coef.
CLI	1X	1	4	6	5	3	9	4	4.4	3	0.70

The turning points of the IPI were forecasted correctly by CLI hence there are no missing cycles. In addition to visual inspection, leading performance is investigated with some statistical criteria. The average number of lag lengths for peaks and troughs do not differ from each other considerably. The standard deviation of the lead time at peaks and troughs are 4.4. Cross-correlation analysis indicates high correlation between CLI and IPI. The maximum correlation occurs at the third lag with a value of 0.70.

4. Predicting the turning points of the CLI

4.1. Neftci's Methodology

The aim of the Neftci's method is to detect the cyclical turning points which determinate the beginning or end of a cyclical downturn. In this approach, sequential analysis is used to calculate the probability of a cyclical turning point and the composite index (X_t) is assumed to have a stochastic behavior. This stochastic process, $\{X_t\}$, pass by two different unobservable states s1 and s2, which are expansion and recession states, respectively. As the number of observations increases, number of states that CLI pass by also increases.

Let Z (Z') be an integer valued unobservable random variable denoting the date following a peak (trough). $Z = i$ ($Z' = i$), for $i = 2, \ldots, t$ with $T \geq t \geq 2$, means that a turning point has appeared between dates $i-1$ and i. It is assumed that the probability distribution of X_t in two states is different and independent from each other and the observations of the stochastic process between and within two states are independent.

If a peak has occurred between dates i and $i-1$, that is $Z = i$ ($T \geq t > i \geq 2$), then:

$$P(X_1 \leq x_1, ..., X_i \leq x_i, ..., X_t \leq x_t) = F^1(x_1, ..., x_{i-1})F^2(x_i, ..., x_t)$$

where $F^1(.)$ and $F^2(.)$ are the cumulative distribution functions for the expansion and recession periods, respectively.

Generally probability distribution functions for the two different states ($f^1(x_t)$ for state 1 and $f^2(x_t)$ for state 2) are assumed to be normal but there are also studies using different distributions (Jorrat and Cerro (2000)).

The value of Z is not directly observable and the observations of X_t up to time t is used to see whether a turning point has started ($Z \leq t$) or not ($Z > t$).

In the sequential algorithm, the prior probabilities of the cyclical turning points are assumed to be known. Let T_t be the a priori transition probability of the change from upward to downward regime,

$$T_t = P(Z = t \mid Z > t-1),$$

and T_t' be the a priori transition probability of the change from downward to upward regime,

$$T_t' = P(Z' = t \mid Z' > t-1).$$

With these information set up to time t, the aim is to maximize the posterior probability of a change in the economic activity. The optimal estimators (posterior probabilities) calculated by using Bayes' rule for peaks (P_t) and troughs (P_t') are given as follows respectively:

$$P_t = \frac{[P_{t-1} + (1-P_{t-1})T_t]f^2(x_t)}{[P_{t-1} + (1-P_{t-1})T_t]f^2(x_t) + [(1-P_{t-1})(1-T_t)]f^1(x_t)}$$

$$P_t' = \frac{[P_{t-1}' + (1-P_{t-1}')T_t']f^1(x_t)}{[P_{t-1}' + (1-P_{t-1}')T_t']f^1(x_t) + [(1-P_{t-1}')(1-T_t')]f^2(x_t)},$$

where the posterior probabilities P_t and P_t' are initialized to 0 for the first observation. While calculating the posterior probabilities, the likelihood that the latest observation in the CLI is from the recession or recovery sample and the likelihood of a recession (recovery) given the current length of the expansion (recession) relative to its historical average are combined with previous months probability estimates.

To estimate the parameters of the probability distribution functions ($f^1(x_t)$ and $f^2(x_t)$) and the a priori probabilities (T_t' and T_t), firstly two separate samples are obtained from observations that belong to upward and downward regimes. Then the probability density functions are estimated by fitting a density function to observations of X_t in each regime. Following Neftci (1982), a priori transition probabilities are assumed to be duration-dependent and they are estimated by using the average duration of upward and downward regimes.

4.2. Evaluation of Probability Forecasts

We used the techniques given in Diebold and Rudebusch (1989) to evaluate the probability forecasts. The turning point forecasts are evaluated by using accuracy and calibration as attributes.

Accuracy refers to the closeness, on average, of predicted probabilities (P_t) and observed realizations (R_t). Rt equals one if a turning point occurs over the horizon H and equals zero otherwise. Accuracy is measured by using Brier's (1950) Quadratic Probability Score, which is the probability forecast analog of mean squared error:

$$QPS = 1/T \sum_{i=1}^{T} 2(P_t - R_t)^2$$

The *QPS* ranges from 0 to 2 and a score of 0 corresponds to perfect accuracy. *QPS* is the unique proper scoring rule that is a function only of the discrepancy between realizations and assessed probabilities as shown by Winkler (1969).

Accuracy is also measured by the Log Probability Score (LPS), which is another strictly proper scoring rule:

$$LPS = -1/T \sum_{t=1}^{T} \left[\left(1 - R_t\right) \ln\left(1 - P_t\right) + R_t \ln\left(P_t\right) \right]$$

The *LPS* ranges from 0 to ∞, and a score of 0 corresponds to perfect accuracy. The difference between *QPS* and *LPS* is that large mistakes are penalized more heavily under *LPS*. Calibration refers to the closeness of forecast probabilities and observed relative frequencies. Overall forecast calibration is measured by Global Squared Bias:

$$GSB = 2(\bar{P} - \bar{R})^2$$

where $\bar{P} = 1/T \sum_{t=1}^{T} P_t$ and $\bar{R} = 1/T \sum_{t=!}^{T} R_t$. The *GSB* ranges from 0 to 2, and a score of 0 corresponds to perfect global calibration.

5. Empirical results

In this section, the results of the empirical studies to forecast the turning point probabilities of the Turkish economic activity are presented. Although normal distribution is used in most of the empirical studies on turning point forecasting, the use of a distribution that adjusts better to the observed data may give superior forecasts. From this point of view, the Q-Q plots of the deceleration and acceleration periods of the CLI are used to give an idea about the proper distribution. According to the Q-Q plots and the goodness-of-fit tests, lognormal and Weibull distributions are used in the calculations in addition to the normal distribution.

The results of the calculated peak and trough probabilities for the CLI show that the peak probabilities obtained from Weibull and normal distributions are very similar but the peak probabilities obtained from lognormal distribution are better since they are smooth and increase continuously. For the trough probabilities, Weibull distribution gives better results than normal distribution. But the calculated probabilities obtained using lognormal distribution are superior to Weibull and normal distributions.

Since the probabilities obtained by using Weibull and normal distributions are similar and lognormal distribution gives better results, the graphs of the calculated peak and trough probabilities obtained from Weibull distribution are not given in this paper. The graphs of the

Figure 3 – Peak probabilities of the CLI; lognormal distribution

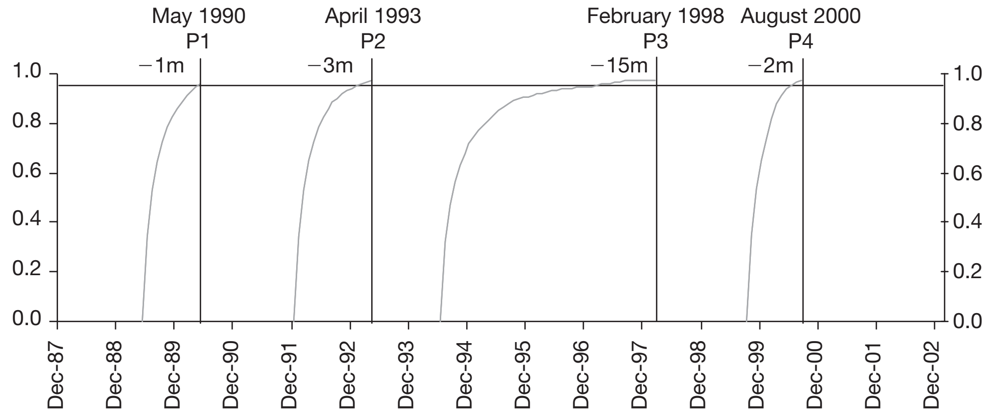


Figure 4 – Trough probabilities of the CLI; lognormal distribution

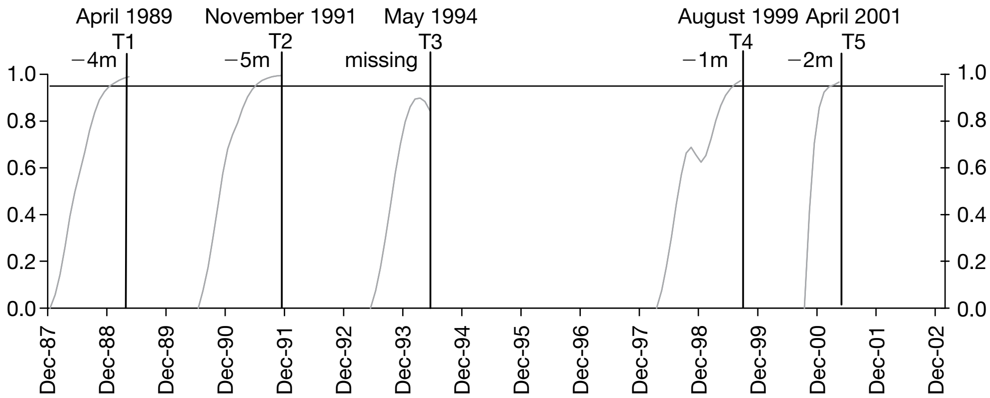
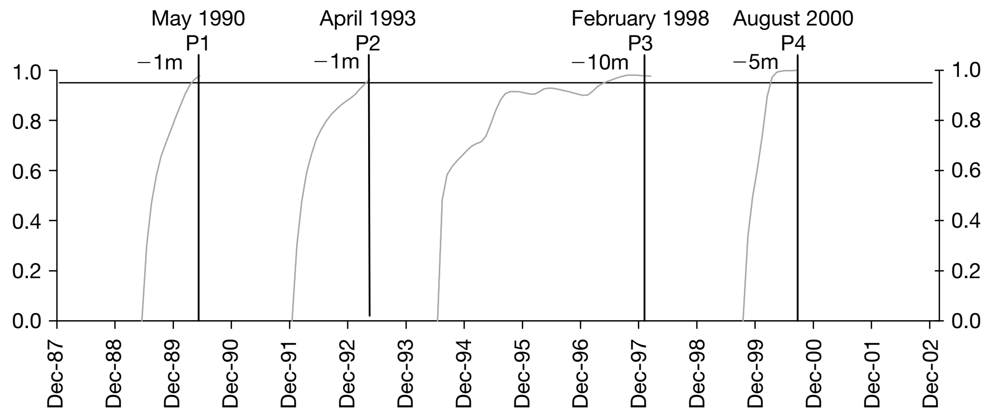


Figure 5 – Peak probabilities of the CLI; normal distribution



calculated peak and trough probabilities of the CLI for lognormal and normal distributions are given in Figures 3 to 6. The threshold value of signaling a peak or trough is chosen as 0.95.

The detailed investigation of the CLI graphs show that the lead times of the peak signals calculated using normal and lognormal distributions are very similar. But for the troughs, the probabilities obtained by lognormal distribution give better signals than normal distribution since the probabilities do not decrease before the occurrence of a trough.

Table 4 gives the peaks and troughs observed in the reference series and CLI together with the leading performance of the CLI and the calculated posterior probabilities over the reference series turning points. The predictive power of the CLI and the calculated probabilities are evaluated by comparison of the leading time, signal leading time and recognition lag values. As given in Zhang and Zhuang (2002), the leading time is calculated as the difference between the

Figure 6 – Trough probabilities of the CLI; normal distribution

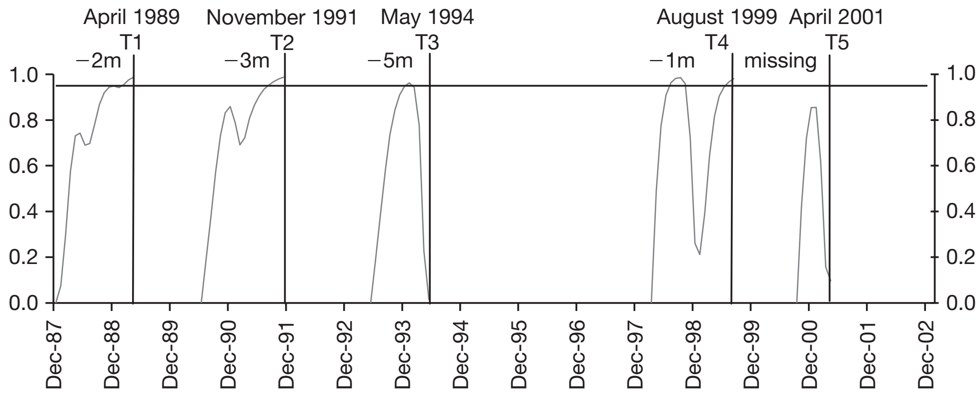


Table 4 – Lead times of the CLI and the posterior probabilities

				Lognormal distribution		Normal distribution	
	CLI	Reference series	Leading time	Signal leading time	Recognition lag	Signal leading time	Recognition lag
Peak 1	–	Dec 1987	–	–	–	–	–
Peak 2	Apr 1990	May 1990	−1	−1	0	−1	0
Peak 3	Apr 1993	Apr 1993	0	−3	−3	−1	−1
Peak 4	Aug 1997	Feb 1998	−6	−15	−9	−10	−4
Peak 5	Apr 2000	Aug 2000	−4	−2	2	−5	−1
Average			**−2.8**	**−5.3**	**−2.5**	**−4.3**	**−1.5**
Trough 1	Jun 1988	Apr 1989	−10	−4	6	−2	8
Trough 2	Feb 1991	Nov 1991	−9	−5	4	−3	6
Trough 3	May 1994	May 1994	0	missing	n.a.	−5	−5
Trough 4	Nov 1998	Aug 1999	−9	−1	8	−1	8
Trough 5	Mar 2001	Apr 2001	−1	−2	−1	missing	n.a.
	Average		**−5.8**	**−3.0**	**4.25**	**−2.8**	**4.25**

**Negative sign denotes lead and positive sign denotes lag.*

time when a turning point in the CLI appears and the time when the corresponding turning point in the reference series that the CLI attempts to predict arrives. The signal leading time is the difference between the time when the calculated probabilities signal a turning point and the time when the turning point in the reference series arrives. The recognition lag is the time required to recognize that a turning point in the CLI signals a turning point in the reference series and it is the difference between the leading time and signal leading time.

According to Table 4, the average leading time for the troughs is longer than the average leading time for the peaks whereas the reverse is true for the signal leading time. The calculated probabilities (both for normal and lognormal distributions) signal the fourth peak before CLI. By looking at the table, it may be said that the use of the calculated probabilities does not bring extra gain to predict the future turning points since one of the troughs could not be predicted, leading time and signal leading times do not differ from each other considerably and no exact rule could be developed due to the different signal leading times at each turning point.

Like Jorrat and Cerro (2000) and Zhang and Zhuang (2002), the successes of the probability forecasts are evaluated by looking at the Quadratic Probability Score (QPS), Log Probability Score (LPS) and Global Squared Bias (GSB) statistics. The QPS, LPS and GSB are calculated for probability forecasts of the CLI under different distributions and exposed on Table 5.

Table 5 – Evaluation measures of probability forecasts

	Horizon	Peaks						Troughs					
		0	1	2	3	4	5	0	1	2	3	4	5
Normal distribution	QPS	1.31	1.24	1.17	1.11	1.05	0.98	1.20	1.06	0.90	0.75	0.63	0.53
	LPS	1.99	1.85	1.73	1.63	1.53	1.43	2.35	2.08	1.72	1.36	1.05	0.82
	GSB	1.16	1.05	0.94	0.84	0.74	0.65	0.79	0.58	0.41	0.26	0.15	0.07
Weibull distribution	QPS	1.26	1.19	1.13	1.07	1.01	0.96	1.14	0.99	0.84	0.69	0.58	0.50
	LPS	1.92	1.74	1.62	1.53	1.44	1.36	2.47	2.12	1.74	1.37	1.06	0.95
	GSB	1.12	1.00	0.90	0.80	0.70	0.62	0.81	0.60	0.42	0.27	0.16	0.07
Lognormal distribution	QPS	0.86	0.81	0.77	0.73	0.69	0.65	1.28	1.11	0.94	0.77	0.63	0.51
	LPS	1.08	1.03	0.98	0.94	0.90	0.86	2.93	2.40	1.90	1.46	1.08	0.79
	GSB	0.74	0.65	0.56	0.48	0.41	0.34	1.00	0.76	0.56	0.38	0.24	0.14

We used these measures to assess and compare the predictive power of the probability forecasts under different distributions.

It can be seen from Table 5 that for the lognormal distribution, the peak predictions have smaller and the trough predictions have higher QPS, LPS and GSB than the other distributions. For the trough predictions, QPS is the smallest in Weibull distribution whereas LPS and GSB are the lowest in normal distribution.

By looking at the QPS, LPS and GSB statistics for the peak predictions, it can be said that the lognormal distribution gives better peak forecasts than the other distributions. And the QPS, LPS and GSB statistics for the trough predictions show that the Weibull distribution gives better trough forecasts than the other distributions.

According to Table 5, for the lognormal distribution, at shorter forecast horizons, the peak predictions have lower QPS and GSB than the trough forecasts and this may show that using lognormal distribution, the peaks are predicted more accurately than the troughs. The results of the probability evaluation statistics are in harmony with the graphs of the peak and trough probabilities given in Section 5. Looking at the graphs it can be said that, the peaks are better forecasted than the troughs and lognormal distribution improves clearly the forecasts.

6. Conclusions

This paper presents the results of the joint work conducted with OECD on the construction of a composite leading indicator to predict the cyclical pattern and the turning points of the Turkish economic activity. The Industrial Production Index (IPI) is used as a proxy for the economic activity, and an analysis is carried out with a broad set of demand, supply and policy variables. From the broad set of series, the seven variables with the most desirable features are selected as the leading indicators. The selected final components constitute a balanced subset of demand, supply and policy variables. After selecting the leading indicators, they are combined into a composite leading indicator in order to increase efficiency.

According to the results of this study, there is an asymmetry between the acceleration and deceleration periods of the series in the sense that the duration of the acceleration periods are longer than the deceleration periods. In addition to this, the probabilities obtained for the longest acceleration period give the signal of a peak nearly one year beforehand in line with the Neftci methodology, which assumes that the longer the economy remained in one state the more likely it was to change to the other. But it is observed that as the duration of a phase increases, the reliability of the estimated probabilities decrease since they give very early (sometimes 10–15 months in advance) signals of a turning point. And this makes harder to forecast the date of the turning point accurately.

As given in Niemira (1991), due to its dynamic characteristic, Neftci's method provides additional information about the strength of a signal, hence increases the possibility of screening out false signals. Therefore the use of the CLI together with the Neftci's sequential probability algorithm may be more efficient in calling the future turning points.

References

Alper, E. (2000) "Business Cycles, Excess Volatility and Capital Flows: Evidence from Mexico and Turkey", Bogazici University, Department of Economics Discussion Papers No.11.

Altay, S., Arıkan, A., Bakır, H. and A. Tatar (1991), "Leading Indicators: The Turkish Experience", Paper presented at the 20th CIRET Conference, Budapest, October 1991.

Anas, J. and L. Ferrara (2002) "Detecting Cyclical Turning Points: the ABCD Approach and Two Probabilistic Indicators", 26th CIRET Conference, Taiwan.

Anas, J. and M. Nguiffo-Boyom (2001) "A New Indicator Based on Neftci's Approach for Predicting Turning Points of the Euro-Zone Growth Cycle", Quarterly Journal of Economic Research, Vol. 3, pp. 364–376.

Artis, M.J., Bladen-Hovell, R.C. and W. Zhang (1995) "Turning Points in the International Business Cycle: An Analysis of the OECD Leading Indicators for the G-7 Countries", OECD Economic Studies, June, pp.125–65.

Atabek, A., Cosar, E.E. and S. Cevik (2002) "Composite Leading Indicator for Turkish Economic Activity", Paper presented at METU/ERC International Conference in Economics VI, Ankara, 11–14 September 2002.

Brier, G.W. (1950) "Verification of Forecasts Expressed in terms of Probability", *Monthly Weather Review*, 75 (January), 1–3.

Burns A. F. and W. C. Mitchell (1946) "Measuring Business Cycles", NBER, Columbia University Press.

Canakcı, İ. (1992) "Kısa Vadeli Tahmin Yöntemleri ve Türkiye icin bir Deneme (Oncu Gostergeler Yaklasimi)", DPT Uzmanlik Tezi, Ankara.

Diebold, F.X. and G.D. Rudebush (1989) "Scoring the Leading Indicators", Journal of Business, 62:3.

Jorrat, J.M. and A.M. Cerro (2000) "Computing Turning Point Monthly Probability of the Argentine Economy According to the Leading Index. 1973–2000", Estudios de Economia, Vol.27, 2:179–95.

Kucukciftci, S. and U. Senesen (1998) "A Composite Leading Indicator Index for Turkey", Istanbul Technical University, Discussion Paper in Management Engineering, 98/3.

Murutoglu, A. (1999) "Leading Indicators Approach for Business Cycle Forecasting and a Study on Developing a Leading Economic Indicators Index for the Turkish Economy", ISE Review, Vol. 3, No. 9, January–March.

Neftci, S.N. (1982) "Optimal Prediction in Cyclical Downturn", *Journal of Economic Dynamics and Control*, 4:225–41.

Neftci, S. and S. Ozmucur (1991) "Turkiye Ekonomisi icin TUSIAD Oncu Gostergeler Endeksi", TUSIAD, June 1991, İstanbul.

Niemira, M.P. (1991) "An International Application of Neftçi's Probability Approach for Signalling Growth Recessions and Recoveries Using Turning Point Indicators", In K. Lahiri and G.D. Moore, eds., Leading Economic Indicators: New Approaches and Forecasting Records, Cambridge University Press, pp. 91–108.

Nilsson, R. (2000) "OECD System of Leading Indicators", Workshop on Key Economic Indicators, 22–25 May 2000, Bangkok.

OECD (1987) "OECD Leading Indicators and Business Cycles in Member Countries", Department of Economics and Statistics, OECD.

OECD (2001) "OECD Cyclical Analysis and Composite Indicators System User Guide", Paris 2001.

Ozatay, F. (1986) "Turkiye Ekonomisinde Devresel Hareketler", Unpublished Ph.D. Thesis, Ankara University.

Selcuk, F. (1994) "TUSIAD Oncu Gostergeler Indeksi", Ekonomiyi İzleme Sempozyumu, 5–6 May 1994, Hacettepe University.

Stock, J.H. and M.W. Watson (1989) "New Indexes of Coincident and Leading Economic Indicators", Macroeconomics Annual, Vol. 4, MIT Press.

Ucer, M., Van Rijckeghem, C. and R. Yolalan (1998) "Leading Indicator of Currency Crises", Yapi Kredi Economic Review, 9:2, 3–23.

Winkler, R.L. (1969) "Scoring Rules and the Evaluation of Probability Assessors", *Journal of the American Statistical Association*, 64, 1073–78.

Zhang, W. and J. Zhuang (2002) "Leading Indicators of Business Cycles in Malaysia and the Philippines", Asian Development Bank ERD Working Paper Series No.32.

Aslıhan Atabek, Evren Erdoğan Coşar, Saygin Şahinöz (Central Bank of Turkey)

Real-time data and business cycle analysis in Germany[1]

Jörg Döpke (Deutsche Bundesbank)

1. Introduction

A correct diagnosis of the stance of the business cycle is crucial both for academic purposes, i.e. estimating or testing business cycle theories, and policy issues, such as for the decisions on stabilisation policy measures. It has been argued, however, that the diagnosis of business cycles might be distorted in real time, i.e. based on the data set available at a certain point in time (Orphanides and van Norden (1999, 2002)). Though the discussion of this topic has a long tradition in Germany (see e.g. Rinne 1969 and, more recently, Braakmann 2003), this problem recently has been addressed for US data only. The present study tries to figure out how important the problems with real-time data and estimates are for business cycle analysis in Germany.

First, the problem of calculating output gaps is addressed. An output gap is defined as the difference between actual real output and the potential or trend output. The importance of this figure for both macroeconomic theory and practical business cycle analysis can hardly be overstated. Unfortunately, it is not directly observable and hence has to be estimated. A wide range of methods have been suggested for that purpose.[2] Moreover, several criteria for the empirical evaluation of output gaps have been suggested in the literature.[3] Among these, the behaviour of output gaps in real time has gained considerably more attention recently. In this paper, we focus on rather simple methods of estimating the output gap. Nevertheless, these methods are both popular and important for practical business cycle analysis.

Second, the paper deals with the problem of inflation forecasts based on real-time output gaps. Simple forecasting equations are considered. While these equations are surely oversimplified, they can still be justified as a crude form of a Phillips-curve relation linking the level of the output gap to future inflation. Additionally, a vectorautoregressive (VAR) model is used to discover the dynamic interaction between the stance of the business cycle as it appears on the base of different data sets, short-term interest rates, and the inflation rate.

Third, the robustness of (at least one) famous stylised fact of the business cycle is sketched. To this end, the contemporaneous correlation of prices and output is addressed. The question arises as to whether such a stylised fact remains unaffected by the use of real-time data.

The paper is organised as follows. Section 2 gives a short summary of the literature on real-time data. Section 3 describes the data and the methods used to estimate the output gap. The following section evaluates the real-time estimations. In particular, the following problems are addressed: what are the statistical properties of the real-time estimates and of the revisions, i.e. the difference of real-time and final output gaps? Are real-time output gaps rational forecasts of the final outcome of this series? Do real-time output gaps provide information regarding future inflation? Do real-time output gaps match business cycle turning points? Does the identification of economic shocks driving the business cycle depend on the use of real-time data? The last section offers some conclusions.

The results of this study strongly support scepticism regarding the reliability of output gap estimates in real time. In particular, it is shown that the first estimates of output gaps are not rational expectations of gaps calculated on the basis of the last available data set. Additionally, the revisions of the first estimates do not behave purely randomly. In contrast, the information content regarding future inflation does not seem to be strongly affected by the use of real-time data.

1 *The author wishes to thank Michael Dear, Christine Gerberding, Uli Fritsche, Heinz Herrmann, Christian Schumacher, Karl-Heinz Tödter and Franz Seitz as well as seminar participants at the German Institute for Economic Research, at the Annual Meeting of the German Economic Association in Zurich, and at the Bundesbank for helpful comments on earlier drafts of this paper. All remaining errors are, of course, mine. The views presented in this paper are the author's and do not necessarily reflect those of the Deutsche Bundesbank.*

2 *For comprehensive surveys see Schumacher (2002) and the European Central Bank (2000).*

3 *See e.g. the discussion in Gamba-Mendez and Rodriguez-Palenzuela (2001) and Rünstler (2002).*

A comparison of output gap measures to results for simple growth rates reveals that the main source of the poor performance of the output gap estimates is not the revision of the underlying data. Rather, the end-of-sample problem inherent in any trend/cycle decomposition seems to be the crucial problem. The modelling of the dynamics of the interaction of inflation and output gaps as well as the correlation of prices and output also appear to be affected by the use of real-time data, though to a much lesser extent.

2. Selected related literature

Economists have recognised the importance of data revisions for a long time. Initial estimates of aggregate output are based on an incomplete set of information. But this is not the only source of data revisions. Following Rinne (2002), one might distinguish three sources of revisions: i) *Statistical revisions*. This kind of revision occurs because the underlying data themselves are still incomplete or estimated when the first figure is published. As regards this point, the statistical office faces a trade-off between an exact and a most timely publication of certain figure. ii) *Revisions due to changing definitions*. From time to time, definitions of the national account statistics change. The last recent example of a revision of this kind is the introduction of the European System of National Accounts (ESA 95). iii) *Methodological revisions*. This category includes revisions due to different methods of price deflation, seasonal adjustment and other transformations of the primary statistics. Beside these categories of revisions, the more recent literature has an even wider understanding of the term revision. For example, Orphanides and van Norden (1999, 2002) call it a revision when an *estimated* figure changes due to additional data, even if the underlying data do not change.[4]

The recent scientific discussion on revisions in that broad sense and their impact on economic research may be summarised as follows (Stark 2002, but see also Rinne 2002). One broad strand of research analyses the magnitude and statistical properties of revisions. For example, this line of research focuses on the question of how large revisions are, both by historical and international standards. Faust, Rogers and Wright (2000) analyse the revisions of the preliminary announcements of output growth rates for the G-7 countries. They conclude that the magnitude of the revisions is quite large, albeit with considerable differences between the countries under investigation. The authors also contribute to a second theme often stressed in this branch of the literature: they present evidence that the revisions are not just white noise but to a surprisingly large extent predictable. This finding is of particular interest, since it suggests that the inclusion of information on the revision process might help to improve the predictability of the latest data. However, though the authors find such evidence for a number of countries, they also point out that the degree of predictability is rather modest.

A second line of argumentation looks at the question of whether the use of preliminary data has consequences for the quality of economic forecasts (Stark and Croushore 2002). For example, statistics used to evaluate forecasts differ considerably depending on whether they are calculated based on preliminary or finally revised data. Additionally, the choice of the appropriate model to generate forecasts might be influenced by data revisions.

The most recognised area of research might be seen in the discussion regarding a possible influence of data revisions and output gap mismeasurement on political decisions. In particular, Orphanides (2002) has argued that the course of monetary policy conducted by the Fed can be understood by the means of errors in gauging the true level of the output gap. Nelson and Nikolov (2001) have confirmed this result using data for the UK and the Bank of England. A large part of the papers on policy analysis is devoted to the question of how the decision on short-run interest rates can be understood. In particular, it is argued that the course of monetary policy in the early and mid-seventies is not, in the first place, due to a less inflation-averse central bank, but due to the fact that the real-time data suggest a deep negative output gap for this time (Orphanides 2002). As a result, that cannot be confirmed based on the final data set.

Real-time data are also suitable to analyse the robustness of empirical findings on macroeconomic relationships. For example, as Croushore and Stark (2000) point out, the estimated response of a certain macroeconomic variable to a shock may well depend on the data set from which this response is estimated (Croushore and Evans 2002). Last but not least, the real-time discussion is important for research on financial markets, since financial markets normally respond to news concerning economic fundamentals (see Stark 2000 and the literature cited therein).

4 *In other words, the revision of an Output gap estimates might be decomposed in two parts: the revision of the underlying data on the one hand and the revision of the estimate du to additional available information as time goes by.*

3. Data and business cycle measures

3.1. Data

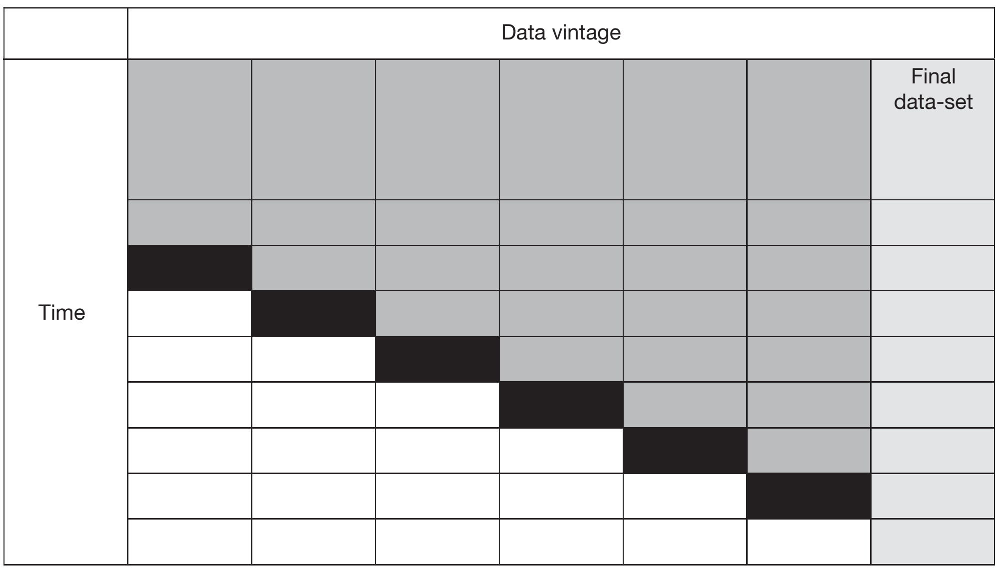
The estimation of output gaps in this paper relies on data on real gross domestic product (GDP). The underlying data are taken from the German Federal Statistical Office (*Statistisches Bundesamt*), which has regularly published quarterly national accounts statistics from 1968 on. As a general rule, the data are published in March, June, September and December. However, additionally a first rough estimate of the annual growth rate is published. Additional information is provided in the Federal Statistical Office's monthly periodical "*Wirtschaft und Statistik*". Thus, to take into account all possible data revisions, data have been collected for each month of a year. As a consequence, a "real-time" series for real GDP is available for each month.

For Germany, however, there are additional problems compared to the US situation, most of which are related to German unification. To begin with, for the latest data release West German data end in 1991, and unified German data start at the same time. As regards the real-time data, the first estimates of data for Germany as a whole were not available before September 1995. Thus, to make the data comparable and to approximate as closely as possible the situation policymakers faced in the early nineties, we shall refer to western Germany up to 1998. Beginning with 1999 the estimates will be based on unified German data (hereafter German data). To be able to calculate real-time data matching this convention, it is necessary, however, to refer not to the latest possible data release, but the release of 1999. These data provide real GDP for western Germany up to 1998. Unfortunately, these data rely on the "old" system of national accounts instead of the "new" ESA data. Hence, the data available in 1999 have been used as the "final" data set. With the exception of figures giving the change over previous year, all data have been seasonally adjusted using the Census X-11 procedure.

Following Orpahnides and van Norden (2002) two types of real-time estimates of the output gap are calculated in this paper. First, data based on the data sets given at a certain point of time, i.e. real-time data, and estimates based on the last available data set with the estimation period restricted, i.e. "quasi-real-time data". Figure 1 illustrates the different concepts. For each point in time, i.e. for each column in figure 1 the output gap is estimated based on the available information at that time, i.e. based on the data in this column. The result of this task is the real-time data set. The quasi-real time data set is based on the final data set, i.e. on the last column only, but proceeds recursively, i.e. row by row to make the data comparable to the Real-time data set. Finally, the final data set refers to the last column only and uses all available data (ie all rows).

As regards real-time data for output gaps, they are constructed as follows (Orphanides and van Norden 2002: 541). In a first step, each and every data vintage is detrended, i.e. in every

Figure1 – Scheme of Different data sets



quarter the output gap is estimated using the data available in this quarter.[5] In a second step, a series is constructed containing the latest available output gap estimate for each quarter. This series is the real-time output gap and represents the timeliest estimate that is possible for each quarter. Note, however, that, possibly in contrast to the US experience, the time span between the data release and the latest available quarter for which GDP data have been released may differ sharply over time. Mostly due to German unification, but also due to other changes in data definitions, there have been some periods without regular published GDP data.

3.2. Business cycle measures

The selection of business cycles measures is mainly motivated by practical purposes, i.e. by the relevance of the respective measure for day-by-day business cycle analysis. Hence, the list of methods comes close to the one considered by the German Council of Economic Experts (2003), though it is restricted by the availability of real-time data. As a consequence, promising multivariate approaches as surveyed by Chagny, Lemoine and Pelgrin (2003) and used in Schumacher (2002a,b) are beyond the scope of this paper.

3.2.1. Changes over previous periods

For the sake of comparison and since these numbers are still by far the most popular measures of the business cycle the changes over previous years and previous quarters are being investigated. As regards the former, the data are not seasonally adjusted. The latter have been seasonally adjusted as described above. They are calculated at an annual rate.

3.2.2. Linear detrending

Linear detrending is widely used for estimating trend and cycle. Moreover, this method is of interest because it is an important input to the production function-based methods of estimating potential GDP. Within these approaches "technical progress", total factor productivity or the potential level of the capital/output ratio are frequently estimated on the basis of a linear trend.[6] Hence, the real-time properties of these methods are strongly affected by the real-time properties of the linear trend model. This can be described as follows: If we let y_t denote the log of real GDP at time t, then the estimation of potential GDP is based on the simple OLS regression

$$y_t = \beta_0 + \beta_1 t + u_t \tag{1}$$

The fit of this equation gives an estimate of potential GDP and the residual u_t is the estimated output gap. Since we have a logarithmic specification, the estimate $\hat{\beta}_1$ gives the average trend growth over the period under investigation. The estimation implies some normalisation since the residuals have a mean of zero.

3.2.3. Quadratic detrending

Recently, some authors have argued that quadratic detrending might give a good approximation to the output gap (see e.g. Gali, Gertler and Lopez-Salido (2001)). The method is implemented by simply regressing the log of output on a trend and its quadrate.

$$y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + u_t \tag{2}$$

Again, the residual u_t provides the estimate of the output gap.

5 *Both real-time and quasi-real-time data make use of the maximum of available data, i.e.* no *"rolling window" approach is applied.*

6 *See e.g. Deutsche Bundesbank (1995).*

3.2.4. Hodrick-Prescott filter

The Hodrick-Prescott (HP) filter (Hodrick and Prescott 1997) has probably become the most popular way of detrending economic time series in the last recent years. This is mainly due to the fact that it can be very easily calculated and implemented in virtually any econometric software package. If *y* denotes real GDP, the filter is defined as

$$min\sum_{t=1}^{T}(y_t - y_t^*)^2 + \lambda\sum_{t=2}^{T-1}\left[(y_{t+1}^* - y_t^*) - (y_t^* - y_{t-1}^*)\right]^2 \tag{3}$$

y_t^* being the smooth component which gives the estimate of potential GDP in this context. An HP filter is more or less a "moving average for snobs" (Kuttner 1994). Broadly speaking, the procedure described in [3] contains two commands: (i) minimise the distance between the actual and the trend value of the time series and (ii) minimise the change in the trend value. Obviously, the commands contradict each other. Therefore, a weight has to be given to both aims. This is done by choosing the factor λ. For quarterly data, a smoothing factor of 1,600 has become something of an "industrial standard". Though this assumption can be justified,[7] the arbitrary choice of the smoothing parameter is one of the major criticisms of the filter. However, in this paper we follow the most frequently used practice.

It is well known that the Hodrick/Prescott filter has an end-of-sample problem, i.e. at the end of the sample the estimates are particular unreliable. To take this fact into account, an approach often adopted by practitioners is also considered here: to make the most recent output gap estimates more reliable forecasted values are added to the filtered series. To calculate the forecasts a simple AR(4) process of the rate of change of real GDP is used.

3.2.5. Band-pass filter

The logic of the band-pass filter suggested by Baxter and King (1999) rests on the grounds of spectral analysis. It assumes that the phenomenon "business cycle" can be described as a fluctuation of a certain frequency. For example, the authors argue that the tradition of Burns and Mitchell suggests that a typical business cycle lasts between 6 and 32 quarters. Fluctuations of a shorter length belong to irregular components of the time series, whereas fluctuations of a lower frequency should be attributed to the trend of time series. Once the upper and lower bound of frequencies which shall define the cycle are given, it is still not possible – at least in a finite sample – to calculate the ideal filter which will remove all fluctuations of that length. Instead, it is only possible to approximate this ideal filter by a moving average. The longer the moving average is, the closer the calculated filter comes to the ideal one. Thus, in implementing the band/pass filter, the three parameters need to be set: the upper bound of frequencies, defining the trend of the time series (32 in this paper), the lower bound defining the irregular part of the series (6 in this paper) and the length of the centred moving average (30 in this paper).

A variant of the Baxter and King filter is suggested by Chistiano and Fitzgerald (2003). The main difference between the two procedures is, that the optimal filter is approximated by a *two-sided* filter in case of the Baxter and King (1999) method and a *one-sided* filter in case of the Chistiano and Fitzgerald (2003) approach. Hence, to calculate recent output gaps based on the former method some values have to be forecasted. In this paper, we follow a common approach and make use of a simple AR(3) process to extrapolate the growth rate of real GDP. In contrast, when applying the Chistiano and Fitzgerald (2003) method the cyclical component of the series can be calculated using a one-sided moving average. The authors argue that their filter has better real-time properties as compared to the Baxter and King (1999) variant.

7 *In their original paper Hodrick and Prescott argue that "a five percent cyclical component is moderately large as is a one-eighth of one percent change in the rate of growth in a quarter" (Hodrick and Prescott 1997: 4). This leads to* $[5/(1/8)]^2 = 1600$*. Some studies discuss the appropriate setting of the smoothing parameter. For a full discussion see of this topic see Mohr (2001) and Tödter (2002).*

[5/(1/8)]2=1600

3.2.6. Unobserved component model

As regards the rather broad class of unobserved component approaches to estimate the output gap, this paper refers to the most simple model of Watson (1986). According to his approach, potential GDP is models by a simple random walk with drift (v):

$$y_t^* = v + y_{t-1}^* + \varepsilon_{1,t} \tag{4}$$

with $\varepsilon_{1,t}$ as a white noise error. The output gap, in turn, is assumed to follow an AR(2) process:

$$gap = \alpha_1 gap_{t-1} + \alpha_2 gap_{t-2} + \varepsilon_{2,t} \tag{5}$$

Again, $\varepsilon_{2,t}$ is a white noise error term. Furthermore, it is assumed that the error terms are uncorrelated.

4. Evaluating the measures

Calculations like in the previous sections may be undertaken for several reasons. One might be interested in the long-run trend of economic activity or the current stance of the business cycle, or in a long time series for analytical purposes, e.g. to estimate an equation. Thus, the criteria to evaluate these data depend on the purpose of the investigation and might hence be quite different. The list of criteria applied in the following section will reveal that the main perspective taken by the present paper is on current business cycle analysis.[8] As a consequence, the main questions are whether real-time business cycle measure are reliable and whether they provide information on possible inflationary pressure.

4.1. Summary statistics

Some summary statistics on the output gap time series are given in table 1. The table compares final, real-time and quasi-real-time estimates as described above. To begin with, the differences in the means of the series are striking in some cases, though the differences between the alternative data sets of the underlying data seem to be of a small magnitude only. For simple detrending methods, the differences are about one percentage point on average. Given a standard deviation of the same order magnitude, this would imply on average a serious misinterpretation of the true stance of the business cycle. It is noteworthy that the filter techniques perform somewhat better according to this criterion. To make things even worse, not even the sign of correction is the same for all methods. Whereas the linear detrending method and, though to a much lesser degree, the band/pass filter lead on average to an upward revision, i.e. imply an underestimation of the output gap based on real-time data, quite the opposite is true for the estimation based on a quadratic trend.

As regards the standard deviation of the time series, the differences between real-time, quasi-real-time and final estimations are – compared to the means – much less a matter of concern. Whereas the competing methods tell alternative stories of how strong economic fluctuations around a trend are, within a given method these differences seem to be minor. This does not rule out the possibility that for single observations the magnitude of revision might be important. However, as regards the minimum and maximum of the output gaps this does not seem to be the case here. Another interesting bit of descriptive information is the correlation of the real-time and quasi-real-time output gaps with their final counterparts. As a general rule, this correlation is only small compared to the respective correlation of the growth rates.

Table 2 shows the summary statistics on the revisions for each time series. The revision is defined as the actual measure minus the first estimate of the measure. Ideally, the revision should have a zero mean, indicating no systematic difference in the output gaps based on different data inventories. Unfortunately, this is not the case for both deterministic trend extraction methods. The filter techniques perform much better in this respect. The standard deviation of the revisions is of roughly the same order of magnitude as the respective standard deviation of the output gap measures itself. This fact is also illustrated by the noise-to-signal ratio of the preliminary estimates. As is done in Orphanides and van Norden (2002), this measure is calculated as the ratio of the standard deviation of the revision to the standard deviation of the final estimate.

8 *For other list of criteria compare e.g. Gamba-Mendez and Rodriguez-Palenzuela (2001).*

Table 1 – Summary statistics of output gap measures in Germany, 1980 I to 2001 IV

Method	Mean	Standard deviation	Minimum	Maximum	Correlation with final estimate
Change from previous year					
Real-time	1.82	1.75	−3.27	6.07	0.95
Quasi real-time	NA	NA	NA	NA	NA
Final estimate	1.92	1.98	−3.80	6.99	1.00
Change from previous quarter					
Real-time	1.57	3.80	−9.86	10.74	0.89
Quasi real-time	NA	NA	NA	NA	NA
Final estimate	1.84	4.23	−10.63	13.53	1.00
Linear trend					
Real-time	−1.39	2.37	−6.36	3.49	0.77
Quasi real-time	−1.30	2.87	−6.89	6.05	0.82
Final estimate	−0.21	2.56	−4.34	6.54	1.00
Quadratic trend					
Real-time	1.29	2.26	−2.37	6.59	0.42
Quasi real-time	1.43	2.65	−2.36	8.43	0.44
Final estimate	−0.53	2.77	−5.26	6.46	1.00
Hodrick-Prescott filter					
Real-time	−0.07	1.33	−3.63	2.87	0.40
Quasi real-time	−0.02	1.50	−4.38	2.71	0.44
Final estimate	0.02	1.35	−2.42	3.89	1.00
Hodrick-Prescott filter, incl. forecasts					
Real-time	−0.10	0.81	−2.11	3.38	0.51
Quasi real-time	−0.03	1.62	−5.00	2.93	0.42
Final estimate	−0.04	1.40	−2.42	3.89	1.00
Band-pass (6,32) filter					
Real-time	−0.12	0.70	−2.07	1.97	0.64
Quasi real-time	1.08	0.88	−0.81	3.33	0.68
Final estimate	0.00	1.16	−2.26	2.98	1.00
Unobserved component model					
Real-time	0.05	1.45	−3.53	3.53	0.49
Quasi real-time	−0.66	1.46	−3.59	3.39	0.57
Final estimate	−0.75	2.61	−5.12	5.95	1.00

The table comprises summary statistics on output gap measures for different concepts. See text for additional information.

If this measure exceeds one, the information provided by the initial estimate appears to be rather useless. Though this is the case for the quadratic trend only, some numbers come very close to one. All in all, the results point to a great importance of revisions for gauging the stance of the business cycle. In some particular cases the revision becomes extremely large, as is indicated by the minimum and maximum observations in the series. An upward or downward revision of seven or eight percentage points will lead the business cycle researcher to a completely different judgement of the cyclical situation. Last but not least, it is noteworthy that most of the revision series show a strong degree of autocorrelation, suggesting that revisions are not just white noise but show some systematic behaviour.

Table 2 – Summary statistics on revisions: final versus real-time estimates, 1980 I to 2001 IV

Method	Mean	Standard deviation	Noise-to-signal ratio	Minimum	Maximum	Auto-correlation
Change from previous year	0.13	0.64	0.33	−1.42	2.01	0.74
Change from previous quarter	0.32	1.97	0.47	−5.37	6.97	−0.16
Linear trend	1.18	1.67	0.65	−1.30	4.66	0.95
Quadratic trend	−1.95	2.77	1.00	−7.27	1.66	0.98
Hodrick/Prescott filter	0.10	1.48	1.09	−2.68	3.67	0.93
Hodrick/Prescott filter, incl. forecasts	0.12	1.21	0.86	−2.52	3.22	0.84
Band-pass (6,32) filter	0.12	0.89	0.77	−1.75	2.68	0.90
Unobserved component model	−0.79	2.28	0.87	−5.17	5.22	0.78

The table comprises summary statistics on revisions of output and output gap measures for different concepts. Revision is final minus real-time estimation. See text for additional information.

4.2. Are real-time output gaps a rational forecast of the final data?

Following Mankiw and Shapiro (1987) and Mankiw, Runkle and Shapiro (1984), we will consider whether the real-time output gaps and the quasi-real-time output gaps can be seen as rational forecasts of the results. To this end, the following estimation is used:

$$gap_t^{final} = \alpha + \beta gap_t^{real\text{-}time} + u_t \qquad (6)$$

In the case of a rational forecast, the null hypothesis $H_0 = \begin{pmatrix} \alpha = 0 \\ \beta = 1 \end{pmatrix}$ cannot be rejected. Table 3 gives the result of such tests for the data under investigation. At the 5% level the rational

Table 3 – Are preliminary output gaps rational forecasts of the final estimate?

Method	$\hat{\alpha}$	$\hat{\beta}$	R^2	Test (F-value)
Change from previous year	−0.41 (−0.73)	1.26 (5.63)***	0.27	0.71
Change from previous quarter	0.34 (1.50)	0.98 (17.6)***	0.78	1.17
Linear trend	0.96 (4.67)***	0.84 (11.2)***	0.60	24.52***
Quadratic trend	−1.24 (−3.67)***	0.50 (4.03)***	0.17	31.29***
Hodrick-Prescott filter	0.06 (0.43)	0.40 (3.98)***	0.16	17.70***
Hodrick-Prescott filter, incl. forecasts	0.12 (1.00)	0.89 (5.83)***	0.28	0.84
Band-pass filter	0.12 (1.27)**	1.05 (7.64)***	0.40	0.82
Unobserved component model	−0.79 (−3.22)***	0.89 (5.26)***	0.24	5.58***

*"Test" indicates an F-test on the rationality of the forecast. t-values in brackets. See text for details. *** (**, *) denotes rejection of the null hypothesis at the 1 (5, 10) percent level. Estimation period is 1968:1 to 2001:4.*

expectation hypothesis cannot be rejected for the simple rates of change. In the case of the changes over the previous year, the null has to be rejected at the 10% level. These findings are in sharp contrast to the results for the output gap figures. In all cases the null has to be rejected even at the 1% level. This result confirms that the real-time estimate has limited informative value. Moreover, it becomes apparent that the revision of the underlying data is not the main problem.

4.3. Business cycle turning points

An important feature of real-time data is that they might help to better understand business cycle forecast errors. A full discussion of this topic[9] would require a full real-time data set, which is not available for Germany yet. However, an important part of the problem is the behaviour around business cycle turning points (Dynan and Elemedorf 2001). The business cycle forecaster might miss the "true" turning point if he relies on real-time data. Therefore, he might misdiagnose the current situation and, as a consequence, be more likely to make the wrong prediction.

Figure 2 depicts the output growth rates and the output gap measures based on different data sets and compares the implied business cycle turning points. It becomes apparent that the dating of a turning point depend crucially on the use of final, real-time and quasi-real-time data. At least this holds for the detrending methods. It is noteworthy that the growth appears to be much more robust to the choice of the data set. This is also apparent in figure 3. The exhibit shows the revisions of the respective business cycle measure, i.e. the difference between the first estimate and the final data set. The magnitude of output gap revision is remarkably high and has the same order of magnitude than the output gap measure itself.

Even if the measures based on different data sets differ regarding the level of the output gap, it might still be the case that they tell at least the same business cycle story qualitatively. Thus, the series may show at least the same sign for a certain period. To highlight whether this is indeed the case, we will make use of a test on information content. This test rests on the classification given in table 4.

From this classification it is possible to evaluate the information content of the real-time estimate using the measure

$$I = \frac{O_{II}}{O_{II} + O_{JI}} + \frac{O_{JJ}}{O_{JJ} + O_{IJ}}.$$

In a "coin flip", we have $O_{II} \approx O_{JI}$ and $O_{JJ} \approx O_{IJ}$ and therefore $I \to 1$. If the real-time estimates fits the final-data set results perfectly than $O_{JI} = O_{IJ} = 0$ and $I = 2$. Therefore, any value of $1 < I \leq 2$ indicates a positive information content. Furthermore, the statistical significance of the information content can be tested (cf. Diebold and Lopez, p. 257): We estimate the expected cell counts under the null of no information content. The consistent estimator for the cell counts is given by $\hat{E} = O_{i.}O_{j.} / O$. Finally, one constructs the following test statistic $C = \sum_{i,j=1}^{2} (O_{ij} - \hat{E}_{ij})^2 / \hat{E}_{ij} \sim \chi^2(1)$. If the empirical value exceeds the critical one, the null hypothesis of no information content has to be rejected. The results of this task are given in table 5.

The results indicate that there are numerous cases in which the sign of the output gap based on real-time estimates does not match the respective number based on the final data set. However, the null of no information content has still to be rejected for all methods under investigation.

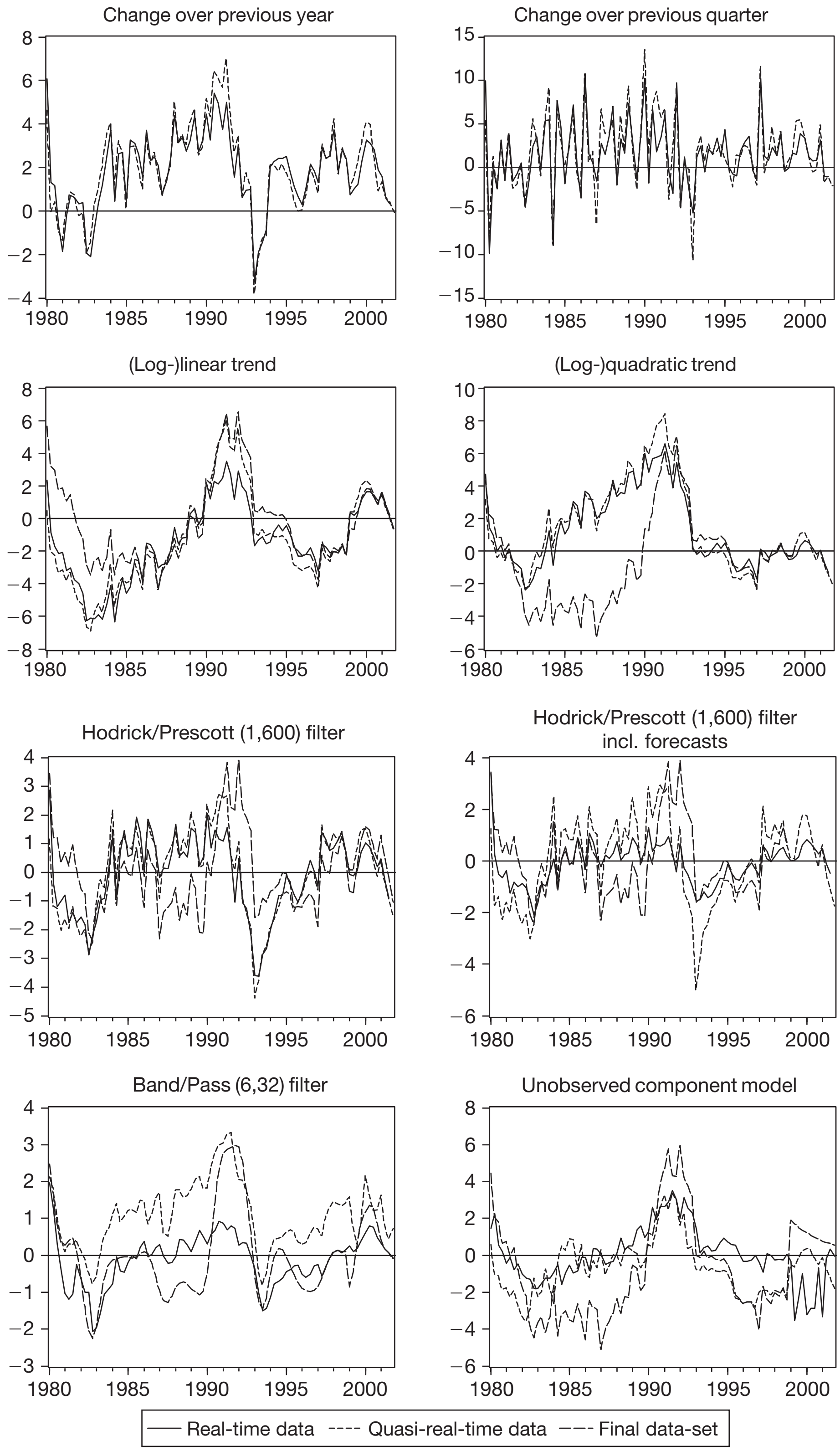
Table 4 – Classification of revisions

		Final data-set		
		Positive output gap	Negative output gap	Sum
Real-time data set	Positive output gap	O_{ii}	O_{ij}	$O_{i.}$
	Negative output gap	O_{ji}	O_{jj}	$O_{j.}$
	Sum	$O_{.i}$	$O_{.j}$	O

Source: The classification follows the classification of forecast errors, see e.g. Diebold and Lopez (1996, S. 257).

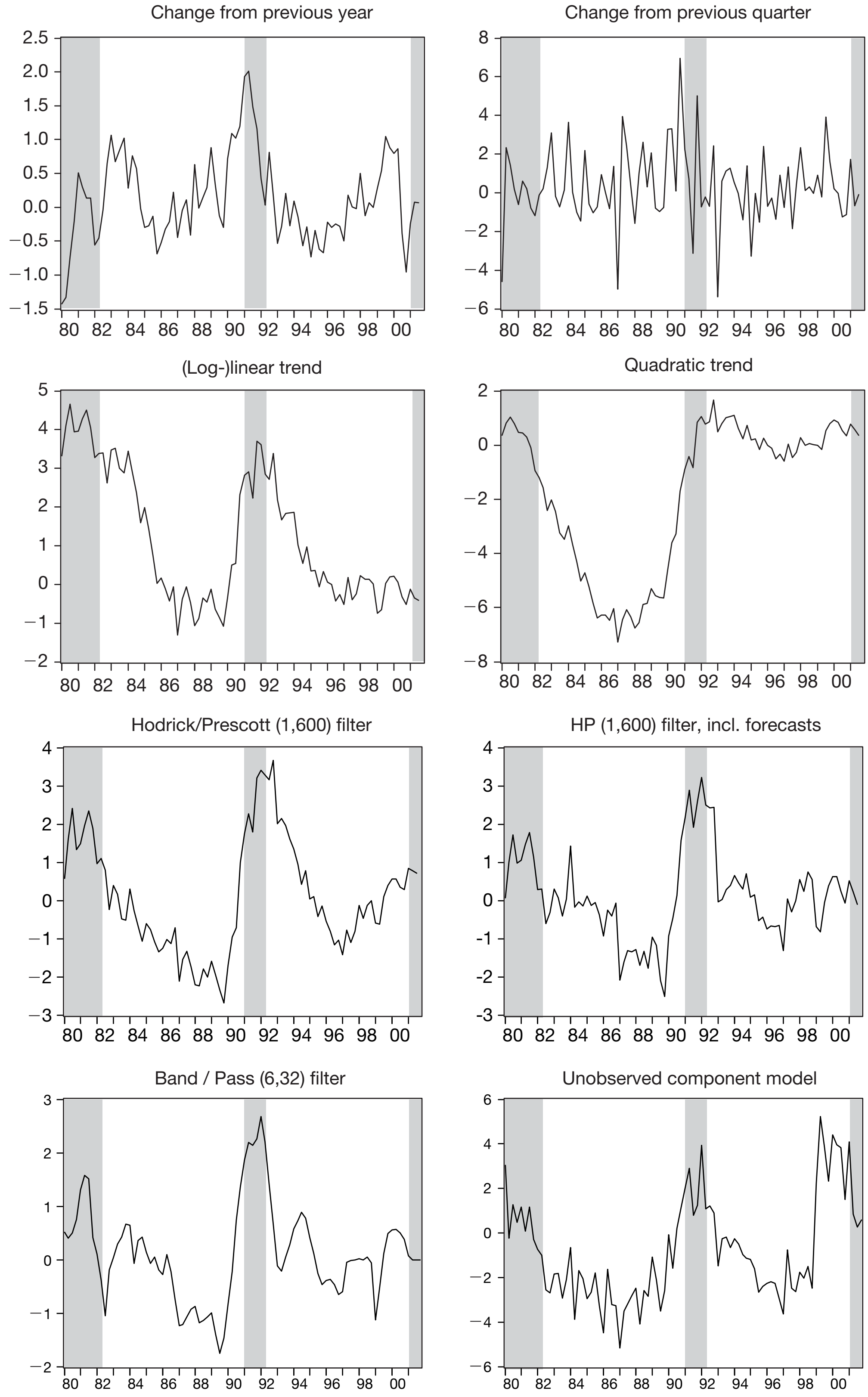
9 *See, for example, Stark and Croushore (2002).*

Figure 2 – Output and output gaps in Germany based on different data sets



Notes: Output gaps are expressed as a percentage of the trend level. See text for additional information.

Figure 3 – Revisions of output gap estimates in Germany



Notes: Revisions are expressed as a percentage of the trend level. The shaded areas are recession according to the definition of the Economic Cycle Research Institute (2003).

Table 5 – Test of information content regarding the sign of the final output gap estimate

Sign of real time estimate:	+	+	−	−	Information content	Test value	p-value
Sign of final estimate:	+	−	+	−			
	%				χ^2		
Change from previous year	84	2	5	9	1.64	64.41	0.00
Change from previous quarter	61	9	3	26	1.75	47.05	0.00
Linear trend	25	17	6	52	1.49	32.50	0.00
Quadratic trend	34	10	28	28	1.27	9.39	0.00
HP-filter	33	14	22	31	1.29	8.64	0.00
HP, incl. forecast	34	13	16	37	1.43	17.36	0.00
Band-pass filter	32	10	15	43	1.50	30.45	0.00
UC model	20	15	27	38	1.16	24.46	0.00

See text for additional information.

From a practitioner's viewpoint it matters whether the data revisions are random or contain a systematic component. If the latter is the case, the forecast of the final outcome might be improved using information on the revisions. It has already been noted in table 2 that the revisions show a large degree of autocorrelation and, thus, are predictable. A related topic is the question of whether macroeconomic variables may help to forecast revisions. If this were the case, the real-time properties of the output and output gap measures might be improved by looking at these variables. To test whether this is the case, tests for Granger non-causality have been performed. As macroeconomic variables linked to the stance of the business cycle, the short-term interest rate (as a prominent leading indicator) and survey data on capacity utilisation are used. Both series have the advantage of not being subject to revision.

The results of tests for Granger-non-causality are given in table 6. Unfortunately, neither series helps to predict revisions. In some cases, however, at least a feedback relationship can be established. This points to the possibility that additional data may help to interpret the current stance of the cycle.

Furthermore, for the purposes of business cycle forecasting it particularly matters whether revisions make it difficult to detect a business cycle turning point. A good deal of forecasting rests on the stylised facts, i.e. on the assumption that, once a turning point has been reached, the forecaster can rely on a "typical" pattern of, say, an upswing. Thus, table 7 compares the order of magnitude of revisions around major business turning points according to the NBER-style definition of turning points.[10] To this end, the revisions' means are calculated for the quarters in which a turnaround has emerged plus and minus one quarter. The results are not clear-cut, however. Some numbers suggest that revisions might be larger in these periods, but no systematic evidence can be found. From this it follows that the problems in detecting turning points are not related to the turning point itself. If the method of determining turning points is independent of the output gap measure, no systematic bias occurs. Again, this points to the conclusion that the problems arise from detrending the series, rather from the underlying stance of the cycle.

4.4. Information content for future inflation

Another empirical criterion for evaluating estimates of the output gap is whether they contain information about future inflation (Claus 2000). The underlying argument is that the output gap is an indicator of excess demand or supply in the aggregated goods market. Thus, if excess demand increases, inflationary pressures should also increase. To analyse this aspect, a simple VAR containing inflation and the respective output or output gap measure equation is estimated (see also Orphanides and van Norden 2003).

10 *See Economic Cycle Research Institute (2003) for a discussion of different concepts of turning points. This implies that the turning point do not themselves depend on real-time data.*

Table 6 – Tests for Granger non-causality of output gap revisions and macroeconomic variables

	H0: Variable does not Granger-cause revision	H0: Revision does not Granger-cause variable	Test decision
Variable		Change from previous year	
Short-term interest rate	0.82	1.43	No causality
Survey data on capacity utilisation	0.54	1.36	No causality
Variable		Changes from previous quarter	
Short-term interest rate	0.06	2.04*	Revisions Granger-cause variable
Survey data on capacity utilisation	1.41	1.24	No causality
Variable		Linear trend	
Short-term interest rate	2.78*	1.13	Variable Granger-causes revisions
Survey data on capacity utilisation	2.58*	4.90***	Feedback
Variable		Quadratic trend	
Short-term interest rate	0.40	1.34	No causality
Survey data on capacity utilisation	6.21***	1.51	Variable Granger-causes revisions
Variable		Hodrick-Prescott filter	
Short-term interest rate	2.16*	1.38	Variable Granger-causes revisions
Survey data on capacity utilisation	3.99***	4.08***	Feedback
Variable		Hodrick-Prescott filter; incl. forecasts	
Short-term interest rate	1.80	0.78	No causality
Survey data on capacity utilisation	1.49	6.47***	Revisions Granger-cause variable
Variable		Band-pass filter	
Short-term interest rate	3.38**	2.23*	Feedback
Survey data on capacity utilisation	1.87	7.16***	Revisions Granger-cause variable
Variable		Unobserved component model	
Short-term interest rate	1.75	2.06*	Revisions Granger-cause variable
Survey data on capacity utilisation	1.41	2.81**	Revisions Granger-cause variable

**** (**, *) denotes rejection of the null hypothesis at the 1 (5, 10) percent level.*

$$X_t = \begin{bmatrix} \pi_t \\ gap_t \end{bmatrix}; \; X_t = \Theta(L) X_{t-1} + \varepsilon_t \qquad (7)$$

If there is information content stemming from the respective gap series, the system in [7] should produce significantly better inflation forecasts than a simple autoregressive process. To test this implication, ex ante forecasts have been computed. To this end, I refer to the end-year data available from 1977 to 1997. These data include the third quarter of the respective year. Thus, for each vintage, data running up to the second quarter of the previous year are available. Based on these data sets, both the VAR and a simple autoregressive process are used to compute forecasts for the period until the end of the next year, i.e. for the coming five quarters. For these forecasts

Table 7 – Revisions around business cycle turning points, 1980 I to 2001 IV

Mean of revision series		
Method	"Normal" periods	Around turning points
Change from previous year	0.12	0.17
Change from previous quarter	0.23	0.81
Linear trend	1.00	1.96
Quadratic trend	−1.99	−0.30
Hodrick-Prescott filter	−0.06	0.92
Hodrick-Prescott filter, incl. forecasts	0.01	0.66
Band-pass (2,32) filter	0.04	0.55
Unobserved component model	−1.03	0.46

Business cycle turning points are: 1973, 3rd quarter (peak), 1975, 2nd quarter (trough), 1980, 1st quarter (peak), 1982, 2nd quarter (trough), 1991, 1st quarter (peak), 1992, 2nd quarter (trough) and 2001, 1st quarter (peak). These data are from: Economic Cycle Research Institute (2003).

the mean squared error (MSE) is computed. With these numbers at hand, it is possible to obtain the loss differential

$$\bar{d} = MSE^{autoregressive} - MSE^{gap} \tag{8}$$

If the inclusion of the gap variable improves the forecasts, the loss differential should be lower than zero. Diebold and Mariano (1995) have developed a test of whether this improvement is significant. In practice, the loss differential is regressed on a constant. If it is significantly higher than zero, the VAR forecast is significantly better, and the output gap measure provides information with regard to future inflation. To ensure white-noise residuals an autoregressive moving average (ARMA) process is added to the test equation. This is likewise recommended by Diebold (1998).

Table 8 – Inflation content for future inflation, 1980 to 1998

Forecasted year	Changes over previous year	Changes over previous quarter	Linear trend	Quadratic trend	Hodrick-Prescott filter	Hodrick-Prescott filter, incl. forecasts	Band-pass filter	UC model	Survey data
1980	0.05	−0.03	0.16	0.09	0.03	0.01	0.11	−0.14	0.13
1981	−0.01	−0.02	−0.17	0.25	0.10	0.13	0.24	0.01	−0.12
1982	0.18	0.01	0.29	0.17	0.14	0.07	0.29	−0.02	0.29
1983	0.10	0.04	0.22	−0.03	0.03	0.04	0.02	0.02	0.18
1984	0.21	0.05	0.25	0.04	0.07	0.10	0.19	0.02	0.20
1985	0.04	0.01	0.09	−0.13	−0.06	−0.01	−0.05	0.01	−0.04
1986	0.40	0.11	0.32	−0.67	−0.24	−0.13	0.55	0.03	−0.06
1987	−0.10	−0.05	0.02	0.04	0.03	0.03	−0.15	−0.06	0.01
1988	0.19	0.13	0.15	0.02	0.12	0.13	0.13	0.07	0.17
1989	0.01	0.00	0.00	0.03	0.01	0.01	0.02	0.01	0.02
1990	0.01	0.02	0.03	−0.21	0.03	0.04	0.04	0.01	−0.12
1991	0.06	0.06	0.07	−0.05	0.05	0.05	0.04	0.06	0.06
1992	−0.01	0.01	−0.15	−0.47	−0.04	−0.08	−0.17	−0.14	−0.13
1993	0.00	0.00	0.06	0.03	0.05	0.06	−0.07	0.03	0.02
1994	0.19	0.16	0.10	0.06	0.13	0.12	0.14	0.07	0.13
1995	0.00	−0.01	−0.04	−0.10	−0.01	−0.01	−0.01	−0.04	0.05
1996	0.01	0.01	0.00	−0.02	0.00	0.00	0.00	0.01	−0.01
1997	−0.03	−0.01	−0.01	0.00	0.00	0.00	−0.02	0.00	−0.01
1998	0.23	0.16	0.25	0.12	0.10	0.14	0.17	0.10	0.15
DM statistic	0.07***	0.04**	0.07***	−0.02	0.03*	0.03***	0.07**	0.01	0.04**
Non-param-test	5.21***	4.16***	4.17***	0.26	3.35***	3.54***	3.35***	2.01**	2.04*

*The table shows the mean loss differential of a forecast based on the lagged inflation rate only (benchmark model) and a VAR forecast based on the inflation rate and the respective output or output gap measure. A positive number indicates that the latter forecast is better. *** (**, *) denotes rejection of the null hypothesis of equal forecast accuracy at the 1 (5, 10) percent level according to the modified Diebold and Mariano (1995) test.*

Table 8 presents the results of the analysis. In general, the methods perform quite well according to this criterion. Except the one based on a quadratic trend output gap measures seems to be significantly helpful in forecasting German inflation. Thus, the results presented here are in some contrast to the findings of Orphanides and van Norden (2003), who argue that virtually no output gap measure is useful to predict inflation. It is noteworthy, however, that survey data are also useful in predicting inflation. Since these variables are not revised at all, it seems to be reasonable at least to double-check an inflation forecast using this variable.

4.5. Identification of macroeconomic shocks

To gain further insight into the importance of the use of real-time data for output gaps for macroeconomic research, a vectorautoregressive (VAR) model is considered (Croushore and Stark 2000). The model analysed in this paper can be justified on the grounds of the Taylor consensus (Taylor 2000). This model can be seen as the workhorse of modern macroeconomics. In a benchmark system this model is built on three equations: an IS function relating the output gap to a real interest rate, a simplified Phillips curve equation linking the development of inflation to the output gap, and a monetary reaction function which stipulates that the authorities react to both the inflation rate and the level of the output gap. Thus, the dynamic interaction of three variables is of particular interest: the inflation rate (π_t), the short-term interest rate (i_t), and the output gap (gap_t). Consequently, the vector of endogenous variables is given by $X_t = \begin{bmatrix} \pi_t \\ gap_t \\ i_t \end{bmatrix}$. The reduced-form VAR model takes the form:

$$X_t = \Theta(L)X_{t-1} + \varepsilon_t \tag{9}$$

In [9], the matrix polynomial $\Theta(L)$ contains the coefficients to be estimated, and the residuals (ε_t) have the variance-covariance matrix $Var(\varepsilon_t) = \Sigma_\varepsilon$. To begin with, the VAR is estimated for a given sample and data vintage, namely for the period 1968–1998 and based on data available in 1998. The length of the lag polynomial was set equal to 2. This choice was based on information criteria and on lag exclusion tests, which are not reported here but are available upon request from the author. Figure 4 depicts the impulse response functions of the VAR obtained in the following order: output gap, inflation, and interest rate (see also Rudebusch and Svensson 1998 and Giordani 2001). In a nutshell, the response functions match the results of previous studies. Hence, it is useful to check, whether these impulse response functions depend on the data set used.

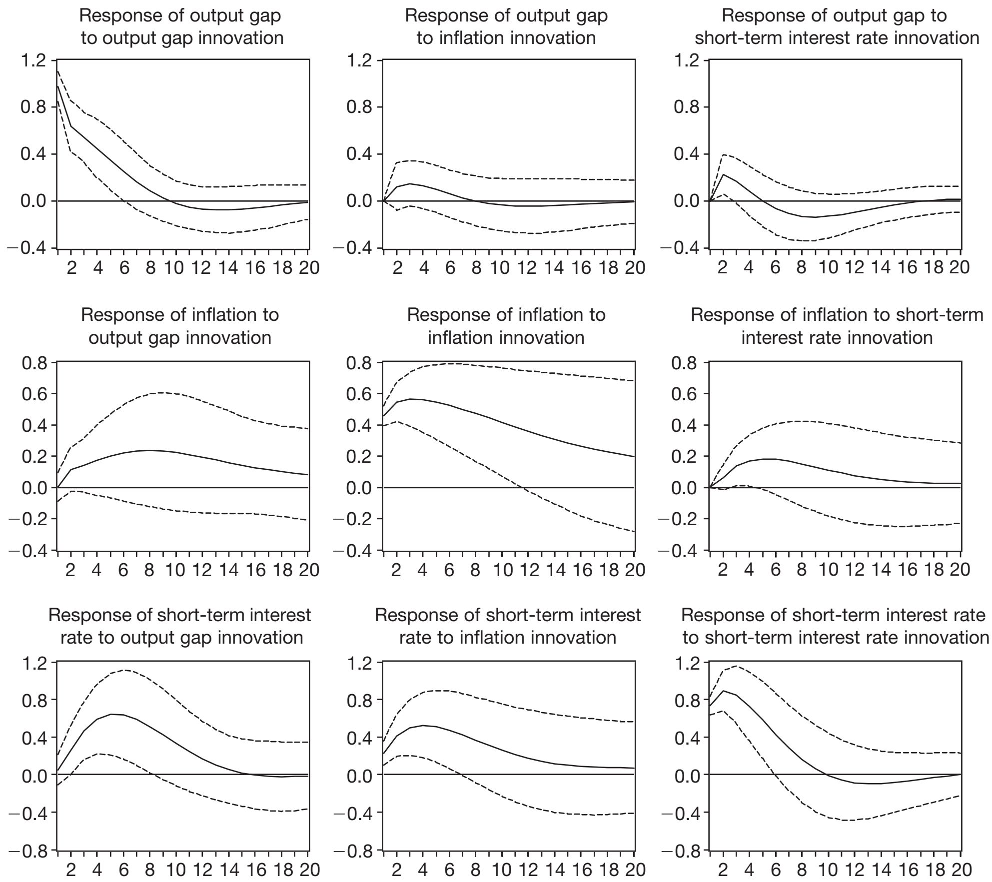
The output gap responds positively to innovations in the output equation, slightly and insignificantly positively to shocks in the inflation equation and negatively to shocks in the interest rate equation. This impulse response function is, however, positive in the first quarters after the shock, which is at odds with an interpretation of this innovation as a monetary shock. Despite the difficult-to-understand short-run behaviour, the medium-term response meets economic prejudice: a higher interest rate lowers the output gap for a while, but not permanently, since the output gap is a stationary variable. As the negative impact is relatively small, the model might still serve as a benchmark.

As regards the inflation rate, two impulse response functions are in line with common expectations: the inflation rate responds positively to its own innovations, which reflects the well-documented fact of inflation persistence, and it also responds positively to output gap innovations, i.e. booms tend to increase the inflation rate. The last impulse response function confirms a puzzle frequently documented in the relevant literature. An increase in the short-run interest rate tends to raise rather than lower the inflation rate. Several explanations have been offered for this puzzle (see Giordani 2001 for a survey). For the purpose of this analysis, it is not necessary to "solve" this puzzle. Rather, this analysis will check whether the use of real-time data helps us to understand this puzzle – or makes it even less comprehensible.

The responses of the short-term interest rate to innovations also unveil a prominent puzzle, the "liquidity puzzle". While an increase in the interest rate in the light of a positive output gap and inflation innovations is plausible, the strong persistence of short-term interest rates raises the question as to why monetary authorities respond to their own decisions on short-term rates with yet another interest rate move in the same direction. Again, we will leave the puzzle for now and discuss instead whether real-time data are helpful in explaining it.

The first question is whether the dynamic interaction between the three variables at hand has changed over time. To illustrate this point, figure 5 compares the response of the output gap to

Figure 4 – Impulse-response functions of three variable VAR (quarters after shock)



Note: The figure depicts the dynamic response of the inflation rate to a one standard deviation innovation.

an innovation in the short-term interest rate. The results suggest that, first, the choice of the output gap measure matters for the judgement on the dynamics of the VAR. Even for a given data set, the impulse-response function differs considerably. Second, the estimations show clearly that the impulse-response function changes over time. The responses for a given method differ for alternative time periods.

This fact, however, is not a clear-cut indication that the real-time problem has any influence on the impulse-response functions. Rather, the possibility of a structural break cannot be ruled out, i.e. the dynamics captured in the VAR itself may have changed. To gain further insights into this problem, a VAR for a given sample (the 1970s) has been estimated based on different data vintages. For this purpose, the impulse-response functions that represent the two puzzles mentioned above have been chosen.

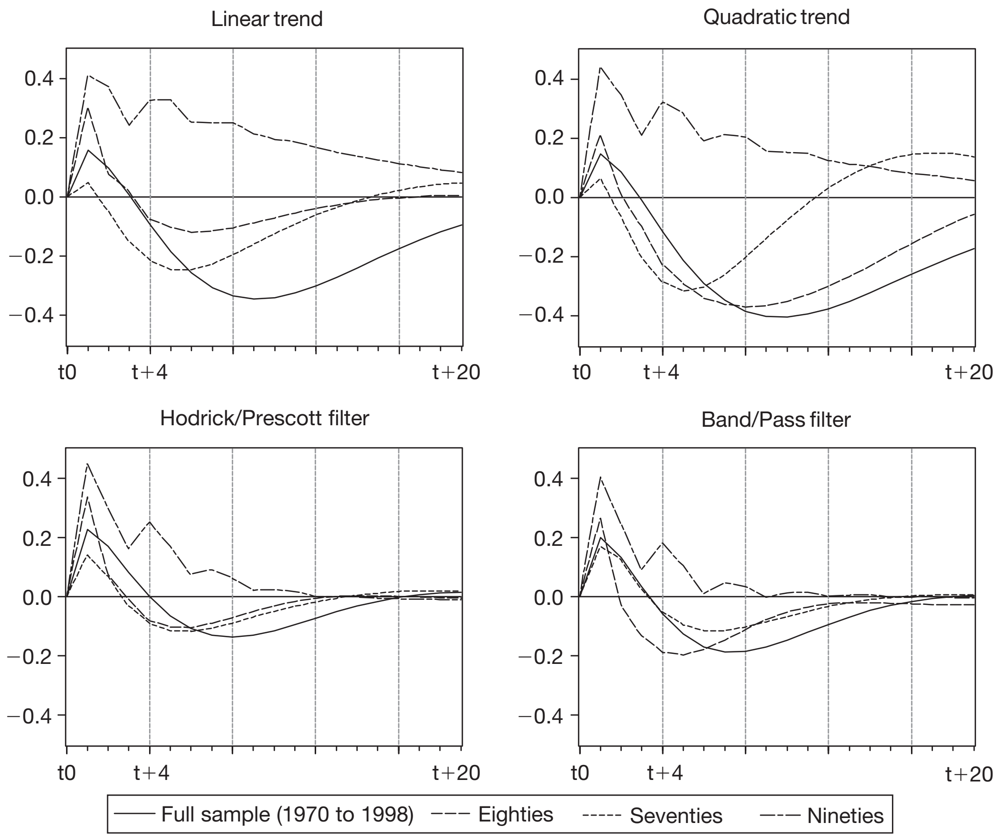
Figure 6 shows the amount of the price puzzle in a VAR estimated for the 1970s. Qualitatively, the results do not differ at all. Quantitatively, however, the response of inflation to interest rate shocks differs depending on the data set used. The price puzzle is largest for the most recent data set and becomes smaller the more data have been used to estimate the gap.

As can be seen from figure 7, the persistence of short-run interest rates is almost completely independent of the underlying data set. Thus, taking the evidence together, the impact of different data sets on the qualitative behaviour of impulse responses seems to be quite small, though sometimes the impact might be quantitatively important. All in all, the results are broadly in line with the findings of Croushore and Evans (2003), who conclude that "(. . .) the use of revised data in VAR analyses of monetary shocks may not be a serious limitation".

4.6. The robustness of stylised facts

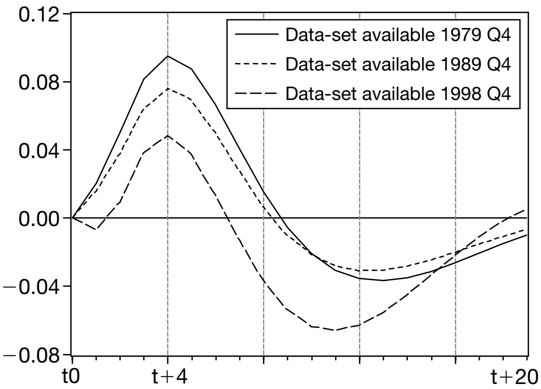
Given that real-time and final estimates differ considerably, the question arises as to whether well-established stylised facts of the business cycle are robust against the choice of a data vintage. Following Stark (2000), the owing discussion will focus on the contemporaneous

Figure 5 – Impulse responses of the output gap to innovations in the interest rate equation based on different data sets and alternative methods of estimating the output gap



Note: The figure depicts the dynamic response of the output gap to a one standard deviation innovation in the interest rate equation.

Figure 6 – Data vintage and impulse-response functions: the "price puzzle"

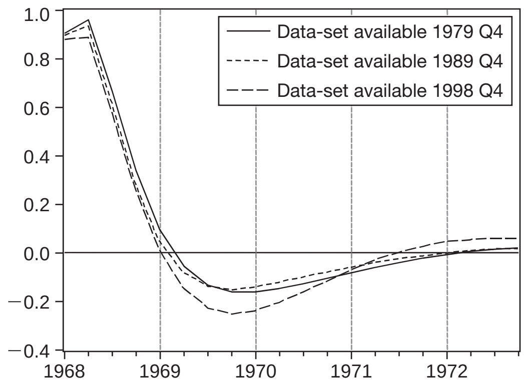


Note: The figure depicts the dynamic response of the inflation rate to a one standard deviation innovation in the interest rate equation.

correlation of the output gap and the detrended price level and the inflation rate. Authors who argue that supply-side shocks might be important for business cycle fluctuations have stressed this correlation. Table 9 shows the correlation of real-time and final output gaps with the respective price figure.

As regards the correlation with the detrended price level, nothing changes qualitatively when real-time data are used in place of the latest data. The picture is less clear for the correlation with the inflation rate. In this case, the differences are generally large. Sometimes even the sign of the

Figure 7 – Data vintage and impulse-response functions: interest rate persistence



Note: The figure depicts the dynamic response of the interest rate to a one standard deviation innovation in the interest rate equation.

Table 9 – Real-time data and the output-price correlation, 1980 I to 2001 IV

Method	Contemporaneous correlation based on	
	Real-time data	Final data
	With detrended price level	
Change from previous year	−0.43*	−0.42*
Change from previous quarter	−0.15	−0.20*
Linear trend	−0.54*	−0.27*
Quadratic trend	−0.51*	−0.26*
Hodrick-Prescott filter	−0.46*	−0.16
Hodrick-Prescott filter, incl. forecasts	−0.57*	−0.16
Band-pass (6,32) filter	−0.62*	−0.20*
Unobserved component model	−0.24*	−0.29*
	With inflation rate	
Change from previous year	−0.34*	−0.29*
Change from previous quarter	−0.20*	−0.20*
Linear trend	−0.05	0.50*
Quadratic trend	−0.07	0.37*
Hodrick-Prescott filter	−0.47*	0.24*
Hodrick-Prescott filter, incl. forecasts	−0.32*	0.25*
Band-pass (6,32) filter	−0.20*	0.26*
Unobserved component model	0.22*	0.39*

*Notes: The table gives contemporaneous correlation coefficients. *a correlation significantly different from zero according to the rule of thumb* $2/\sqrt{n} = 0.18$.

correlation changes. Hence, the results suggest that major stylised facts of the business cycle might not be robust against the use of real-time data.

5. Conclusions

As usual, the most obvious conclusion of this paper is, that it points to the need of further research. The first and most urgent item on the agenda is the inclusion of additional methods of estimating the output gap in the analysis. However, the results of Orphanides and van Norden (2002) suggest that the problems with real-time output gaps cannot be resolved by using more sophisticated methods. Moreover, the simple trend extraction methods used in this paper are of some practical relevance for the analysis of the German business cycles, since production function approaches (which are the dominant method of estimating Germany's output gap) depend heavily on the assumed trend model. Second, more real-time data are necessary. The availability of such data would make it possible to use multivariate methods to estimate potential GDP. Furthermore, the discussion of the stability of prominent empirical results and stylised facts of the business cycle would rest on more solid footing if all involved data were real-time data. Third, on the methodological side, possibilities of reducing the measurement error of output gaps should be discussed. For example, it is well known that the end-of-sample properties of filters may be improved by using forecast data (Mohr 2001). Fourth, the consequences of real-time data for both policy decisions and forecasting must be addressed more carefully. Given the preliminary nature of this paper, all these topics are left for further research.

Given the limitations mentioned above, it would be premature to draw too far-ranging conclusions from the results. However, some conclusions can and should be drawn. First, the notion that the quality of real-time estimates of the output gap is rather poor is strongly confirmed by the German data. Hence the results strongly support the scepticism on the usefulness of output gaps estimates in real time raised by Orphanides and van Norden (2002), among others. Of course, the methods differ in respect to the alternative criteria, which are used to evaluate the real-time estimates as it is summarised in table 10. However, it is not possible to find a single method that dominates that others according to all criteria. From this it might be concluded, that the problem lies in output gap estimates itself, or, more precisely, in the information available in real-time, when estimating a gap, rather than in the limitation of one particular method applied.

Table 10 – Overview over selected results

Method	Revisions[1]	Rational expectation?	Information content for final estimate?	Information content for inflation?[2]
Change from previous year	Small	Yes	Yes	High
Change from previous quarter	Small	Yes	Yes	Medium
Linear trend	Medium	No	Yes	High
Quadratic trend	Large	No	Yes	None
Hodrick/Prescott filter	Large	No	Yes	Low
Hodrick/Prescott filter, incl. forecasts	Large	Yes	Yes	High
Band-pass (6,32) filter	Large	Yes	Yes	Medium
Unobserved component model	Large	No	Yes	None

[1]*Small: Noise-to-signal ratio* <0.5, *Medium:* 0.5< *Noise-to-signal ratio* 0.75, *Large: Noise-to-signal ratio* >0.75 *(see table 1).*

[2]*High: DM-test significant at the 1% level, Medium: DM-test significant at the 5% level, Low: DM test significant at the 10% level, None: DM-Test insignificant (see table 8).*

Second, it should be noted that the main source of the revisions of the output gap measures is *not* the revision of the underlying data set but the end-of-sample problem of the estimators used. For example, the results regarding simple growth rates appear to be much more robust to changes in the data vintage than output gap estimates. Third, the information content for future inflation, the dynamic interaction between inflation and the output gap and some stylised business cycle facts are apparently affected by the use of real-time data. However, the impact on these techniques seems to be rather limited and less systematic.

To sum up, the degree of uncertainty regarding the level of the current output gap is enormous. This is, of course, a challenge for stabilisation policy. If the current business cycle position is not clear, stabilisation policy is hard to justify. It would be premature, however, to argue that policy authorities should ignore the output gap. There are several ways we can try to improve our knowledge. For example, one can try to find methods of estimating the output gap with better real-time properties. But even if one has to admit that estimating the current level of the output gap is likely to remain difficult, this does not imply that this series should be ignored completely. A broad strand of the literature addresses the question of how to deal with output gap uncertainty. For example, it is possible to refer to a so-called "speed limit" policy, i.e. a policy relying on the change of the output gap rather than on its level (see Orphanides and Williams 2002).

References

Braakmann, A (2003), Qualität und Genauigkeit der Volkswirtschaftlichen Gesamtrechnungen. Allgemeines Statistisches Archiv 87(2): 183–199.

Baxter, M and R King (1999), Measuring business cycles: Approximate band-pass filters for economic time series. *The Review of Economics and Statistics* 81, 575–593.

Claus, I (2000), Is the Output Gap a Useful Indicator of Inflation? *Reserve Bank of New Zealand Discussion Paper* No 2000/05. Wellington.

Clausen, JR (2003), Real-Time Output Gap Estimates for Germany – Implications for Monetary Policy, mimeo, University of Cologne.

Chagny, O, M Lemoine and F Pelgrin (2003), Assessment of multivariate of output gap in the Euro area: An aggregate versus disaggregate approach. Paper presented at the 4th Eurostat and DG Ecfin Colloqium on Modern Tolls for Business Cycle Analysis, Brussels.

Christiano, LJ and TJ Fitzgerald (2003), The Band Pass Filter. *International Economic Review* 44: 435–465.

Croushore, D and CL Evans (2003), Data Revisions and the Identification of Monetary Shocks. Working Paper No 03-1. *Federal Reserve Bank of Philadelphia*.

Croushore, D and T Stark (2000), A Funny Thing Happened On the Way to the Data Bank: A Real-Time Data Set for Macroeconomists. *Federal Reserve Bank of Philadelphia Business Review*, September/October: 15–27.

Croushore, D and T Stark (2003), A Real-Time Data Set for Macroeconomists: Does the Data Vintage Matter? *Review of Economics and Statistics* 85(3): 605–617.

Chauvet, M and JM Piger (2003), Identifying Business Cycle Turning Points in Real Time. *Federal Reserve Bank of St. Louis Economic Review* March/April: 47–62.

Deutsche Bundesbank (1995), Production potential in Germany and its determinants, *Monthly Report*, August 1995: 39–53.

Dynan, KE and DW Elemendorf (2001), Do Provisional Estimates of Output Miss Economic Turning Points? Board of Governors of the Federal Reserve System, mimeo. Washington DC.

Economic Cycle Research Institute (2003), International Reference Cycle Dates. http://www.businesscycle.com/research/intlcycledates.php#germany

European Central Bank (ECB) (2000), Potential Output Growth and Output Gaps: Concept, Uses and Estimates. *ECB Monthly Bulletin*, October 2000: 37–47.

Faust, J, JH Rogers and JH Wright (2000), News and Noise in G-7 GDP Announcements. Board of Governors of the Federal Reserve System, International Finance Discussion Paper No 690. Washington DC.

Gali, J, M Gertler and JD Lopez-Salido (2001), European Inflation Dynamics. *European Economic Review* 45: 1237–1270.

Gerberding, C, F Seitz and A Worms (2004), Real-Time Data and the Monetary Policy Reaction Function of the Deutsche Bundesbank. Mimeo, Deutsche Bundesbank. Frankfurt.

German Council of Economic Experts (2003), Annual Report 2003/04: Consolidate public finances – reform tax system. Wiesbaden.

Giordani, P (2001), An Alternative Explanation of the Price Puzzle. SSE/EFI working paper series in economics and finance No 414, Stockholm.

Gamba-Mendez, G and D Rodriguez-Palenzuela (2001), Assessment Criteria for Output Gap Estimates. ECB Working Paper No 54. Frankfurt a M.

Hodrick, RJ and EC Prescott (1997), Postwar US Business Cycles: An Empirical Investigation. *Journal of Money, Credit, and Banking* 29: 1–16.

Jacobs, J and JE Sturm (2003), Do Ifo indicators help explain revisions in German industrial production? Paper prepared for the Victor Zarnowitz Seminar RWI Essen.

Knetsch, T (2003), The Inventory Cycle of the German Economy. Paper presented at the CIRET conference in honour of V. Zarnowitz June 2003, Essen.

Koenig, EF, S Doams and J Piger (2003), The Use and Abuse of Real-Time Data in Economic Forecasting. *Review of Economics and Statistics* 85(3): 618–628.

Mankiw, NG and MD Shapiro (1987), News or Noise? An Analysis of GNP Revisions. NBER Working Paper No 1339. Cambridge MA.

Mankiw, NG, DE Runkle and MD Shapiro (1984), Are Preliminary Announcements of the Money Stock Rational Forecasts? *Journal of Monetary Economics* 14: 15–27.

Mitchell, J (2003), Business cycle uncertainty in real-time. Paper presented at the 4th Eurostat and DG Ecfin Colloqium on Modern Tolls for Business Cycle Analysis, Brussels.

Mohr, M (2001), Ein disaggregierter Ansatz zur Berechnung konjunkturbereinigter Budgetsalden für Deutschland: Methoden und Ergebnisse. Discussion paper 13/01, Economic Research Group of the Deutsche Bundesbank. Frankfurt a M.

Nelson, E and K Nikolov (2001), UK Inflation in the 1970s and 1980s: the role of output gap and mis-measurement. Bank of England Working Paper No 148. London.

Newey WK and KD West (1987), A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.

Orphanides, A (2002), Monetary Policy Rules and the Great Inflation. *American Economic Review* 92: 115–120.

Orphanides, A and S van Norden (1999), The Reliability of Output Gap Estimates in Real Time. Board of Governors of the Federal Reserve System Finance and economics discussion series; 1999-38.

Orphanides, A and S van Norden (2003), The Reliability of Inflation Forecasts Based on Output Gaps Estimates in Real-Time. CIRANO Working Paper No 2003s-01. Montreal.

Orphanides, A and S van Norden (2002), The Unreliability of Output Gap Estimates in Real Time. *The Review of Economics and Statistics* 84: 569–583.

Orphanides, A and JC Williams (2002), Robust Monetary Policy Rules with Unknown Natural Rates. *Brookings Paper on Economic Activity*: 63–146.

Remsperger, H (2003), Das System der Volkswirtschaftlichen Gesamtrechnungen als Informationsquelle für die Geldpolitik der EWU. Allgemeines Statistisches Archiv 87(2): 133–147.

Rinne, H (1969), Die Verlässlichkeit von Sozialproduktsdaten auf der Basis von Revisionen. Meisenheim.

Rinne, H (2002), Revision. In: Brümmerhoff, D and H Lützel (eds) Lexikon der Volkswirtschaftlichen Gesamtrechnungen. Munich: 328–333.

Rudebusch, G and LEO Svensson (1998), Policy Rules for Inflation Targeting. NBER working paper No 65, 12. Cambridge MA.

Runkle, D (1998), Revisionist History: How Data Revisions Distort Economic Policy. *Federal Reserve Bank of Minneapolis Quarterly Review* 22. 3–12.

Rünstler, G (2002), The Information Content of Real-Time Output Gap Estimates: An Application to the Euro Area. *European Central Bank Working Paper* No 182. Frankfurt a M.

Schumacher, C (2002a), Alternative Schätzansätze für das Produktionspotenzial im Euroraum. Baden Baden.

Schumacher, C (2002b), Forecasting trend output in the Euro area. *Journal of Forecasting* 21: 543–558.

Scott, A (2000), Stylised Facts from Output Gap Measures. Reserve Bank of New Zealand Discussion Paper No 2000/07. Wellington.

Stark, T (2002), A summary of the Conference on Real-Time Data Analysis. *Federal Reserve Bank of Philadelphia Business Review* Q.1: 5–11.

Stark, T and D Croushore (2002), Forecasting with a real-time data set for macroeconomists. *Journal of Macroeconomics* 24: 507–531.

Tödter, KH (2002): Exponential smoothing as an alternative to the Hodrick-Prescott filter?, in: Mittnik, S and I Klein (ed) Contributions to modern econometrics – From data analysis to economic policy, Amsterdam: 223–237.

Abstract

This paper examines the consequences of using "real-time" data for business cycle analysis in Germany based on a novel data set covering quarterly real output data from 1968 to 2001. Real-time output gaps are calculated. They differ considerably from their counterparts based on the most recent data. Moreover, they are not rational forecasts of the final series. The consequences of using real-time data for inflation forecasts, the dynamic interaction of output gaps and inflation, and stylised facts of the business cycle are also addressed. The results suggest that revisions of data and estimates can seriously distort research and policy implications.

Non Technical Summary

This paper examines the consequences of using "real-time" data for business cycle analysis in Germany. "Real-time" data are data, which have been available for a certain point of time. Based on a data set covering quarterly real output data real-time output gaps are calculated. They differ considerably from their counterparts based on the most recent data ("final data set"). Moreover, they are not rational forecasts of the final series. The consequences of using real-time data for inflation forecasts, the dynamic interaction of output gaps and inflation, and stylised facts of the business cycle are also addressed. The problematic nature of real-time output gaps is not due to revisions of the underlying data, but due to the end-of-sample problem of estimates of potential GDP. Thus, the results suggest that revisions of estimates can seriously distort research and policy implications.

Keywords: Real-time data, business cycles, output gap, VAR, inflation, Germany.

JEL-Classification: E32, C53.

Jörg Döpke
Deutsche Bundesbank, Economics Department
Wilhelm-Epstein-Strasse 14, 60014 Frankfurt a. M.
Germany
Tel: +69 95 66 3051
Email: joerg.doepke@bundesbank.de

GDP flash monthly estimates in DRC[1]

Gerard Mutombo Mule Mule and Marie-Jose Ndaya Ilunga
(*Central Bank of Congo*)

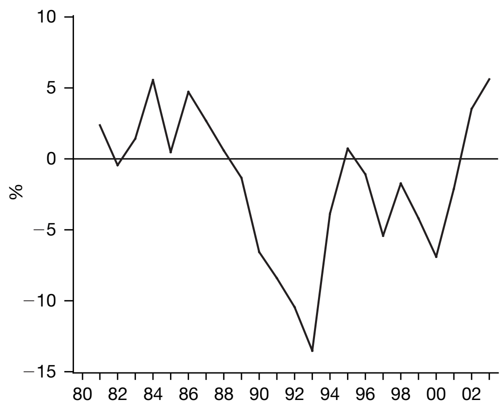
Part 1: The Democratic Republic of Congo background information and framework of national accounts production

1.1. Background

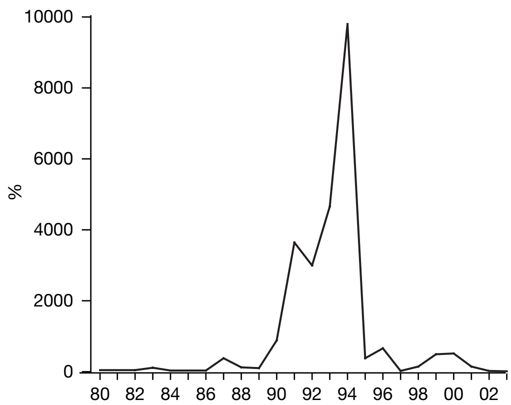
Located in the heart of Central Africa, the Democratic Republic of Congo (DRC) is huge territory extending to 2,245,000 square kilometres. With an average annual demographic rate of 2.7 % and a density of 24.1 inhabitants per square kilometre, its population is currently estimated at 58 million.

Major cities include Kinshasa, the capital, located in the West with a population of 6 million; Kisangani in the Oriental Province with about 5 million; Lubumbashi, the Chief city of Katanga, in the far South with a population of 4 million and Mbuji Mayi in the heart of Eastern Kasaï with more or less 2.5 million.

Graph 1 – Annual real GDP growth (1980–2003)



Graph 2 – Annual inflation (1980–2003)



Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques.

Since cooperation renewal with Bretton Woods Institutions in 2001, frozen so far following 1991–1993 lootings, strong macroeconomic reforms undertaken with the IMF and the World Bank technical assistance have helped the country stabilize the macroeconomic framework and renew with the GDP growth rate reached 3.5 % against an average annual rate of –5.3 % between 1990 and 2001.

This development increased the GDP per capital around USD 113 presently, against 96.1 % in 2001. Similarly, the inflation rate has decreased to around 3.0 % this year, against an average annual rate of 811.7 % in the nineties.

1.2. Production system of real sector statistics in DRC

Regulating the real sector statistical activity under DRC legal arrangements, the Institut National de la Statistique (INS), is the only official body of economic and socio-demographic statistics production and distribution. However, due to the weak output of this institution following the

1 *Paper presented at the Conference of Irving Fisher Committee on Central Bank Statistics in Basel.*

1991–1993 crisis when it lost all the statistical equipments in the lootings, the Central Bank of Congo (BCC) was forced to deal more with real sector statistics production, namely those related to national accounts and to household expenditure price general index number.

To date, the Central Bank of Congo backs basic data collection enquiries and issues the information in an Annual Report and its weekly and monthly publications.

It is also worth mentioning that the UNDP Office in DRC has recently assisted the INS appropriately in the publication of national accounts in a specialised document.

1.3. Institution and analysis framework

1.3.1. Institution framework

At the institutional level, national accounts are dealt with the DRC by the Commission of Statistical Surveys and National Accounts (C.E.S.C.N.).

Created by the CAB/PLAN/90 departmental order of April 14th 1990, this Commission is supervised by the INS and the Central Bank of Congo. Its members include delegates from the following institutions: Presidency, Ministries of Planning, Economy, Finance, Budget, agriculture, Labour and Social Security, the Duty Office, the Institute of Economical and Social Research, Fédération des Entreprises du Congo and Accounting Institution named “Conseil Permanent de la Comptabilité”.

1.3.2. Reference framework: methodology

The drawing up methodology of accounts is mainly based on the United Nations Accounting System of 1998 (SCN 68). It should be pointed out, however, that at the request of the IMF and the National Accounts Sub-Committee of the Southern African Development Community (SADC) of which DRC is a member within the framework since 2001 the move to SCN 1993, by drawing up a crossing grille of the old system accounts to the new one in accordance with accountancy standards and by extending the field of data collection.

Nevertheless, it is worth stating that DRC has not implemented yet the first phase of SCN 1993 to date.

The Gross Domestic Product (GDP) is annually valued mainly by the approach of the product in terms of the sum of added values of all areas of national economy sectors.

Given the changes undergone by the structure of economy for the last decade mainly characterized by a deep economic recession (See Tables 1 and 2) with a GDP average annual growth rate of –5.3 %, the approach by aggregate demand is used to determine the other components, which are Civil Services consumption, public and private investment, stock variation, exports and imports. Due to the lack of recent enquiries about the structure of private consumption, this aggregate is calculated in residual form.

Economic research enquiries monthly made by the Central Bank of Congo help collect information on a sample of products of major sectors of economy. These data are used to value the GDP on a monthly basis. Estimates of goods and services are calculated with constant price lists.

1.3.3. Classification system of economic activities

The United Nations International Standard classification by industry (CITI Updated 3) is the nomenclature used. Key sectors are:

- Agriculture, hunting and forestry;
- Production activities;
- Power, gas and water production and distribution;
- Construction;
- Wholesale and retail trade;
- Hotels and restaurants;
- Transport, warehousing and communications;
- Financial intermediary services;
- Real estate, hiring and other firm activities;
- Civil services;
- Education, health and other collective, social and individual services.

1.3.4. Fields of application and data sources

All resident commercial and industrial production units as well as traditional and unofficial sectors are covered by the national accountancy. Data collected monthly and annually from the modern sector production units of goods and services (statistical and customs declarations of firms) and from Civil Services (implementation of the Central government's budget and that of provinces, customs statistics) through economic enquiries and economic research surveys organized by the Central Bank of Congo are the key information sources.

The balance of payments is another national accounts data source.

Theses sources also include newly collected information from specific enquiries organized by the INS to get production indicators of sectors of activities, not taken into consideration so far in the aggregates formerly calculated, namely craft, the self-consumed production of peasantry or the unofficial trade which both constitute the hidden part of the country's economic iceberg.

Part 2: GDP flash monthly estimates

2.1. Introduction

Valuing the GDP on a monthly basis is recent in DRC as national accounts used to work traditionally on annual accounts describing the development of economic flows for the whole year as well as the GDP quarterly estimates. However, the relatively long deadline between the publications of accounts provisional results of the year n to the term $n+1$ makes the GDP flash monthly estimates useful to help economic policy decision makers get indications on economic policies implemented and adapt them to the situation.

Due to all types of problems met while valuing the monthly GDP, economic research regularly made by the Central Bank of Congo help basic data to undertake this important exercise.

As soon as they are validated by the C.E.S.C.N., these estimates are published by the Central Bank of Congo in such documents as, the Summary of Statistical Information and Recent Economic and Financial Development, to meet the demand of many users from business, academics and diplomacy.

This part describes the content of the GDP monthly estimates, the methodology used and their uses.

2.2. Content of the GDP monthly estimates

The GDP monthly estimates only state the development by activity sector. It is mainly the product approach which is used. Components of total demand are not calculated monthly due to problems met in getting information for such short periods.

These estimates are presented in tabular form showing value added development in value and volume by sector.

2.3. Estimation methodology of monthly GDP

Given the lack of full and reliable statistical data, estimations are made through a sample of products representing the general trend of the sector I the whole. Moreover, the GDP resorts to econometric methods to fill in the gap noticed in some sectors.

The GDP is calculated by added values aggregate of economic activity of different sectors. The added value is extrapolated by applying a production volume index at the base period. This index is obtained according to the nature of available information by using quantitative data or by deflating the production current value by an appropriate price index number.

$$\mathrm{VAGW} = ((\mathrm{VA}_t/\mathrm{VA}_{t-1})^{1/12} - 1)*\ 100 \qquad (1)$$

where

VAGW = added value growth in monthly percentage change;
VA_t = added value at constant price (2000 = 100) current year;
VA_{t-1} = added value at constant price (2000 = 100) past year.

The variation of the relation (1) can be used to calculate the monthly GDP on the base of annual data (2).

$$GDPGW = ((GDP_t/GDP_{t-1})^{1/12} - 1)*100 \qquad (2)$$

where

GDPGW = GDP growth in monthly percentage change;
GDP_t = GDP at constant price (December 2002 = 100) current year;
GDP_{t-1} = GDP at constant price (December 2002 = 100) past year.

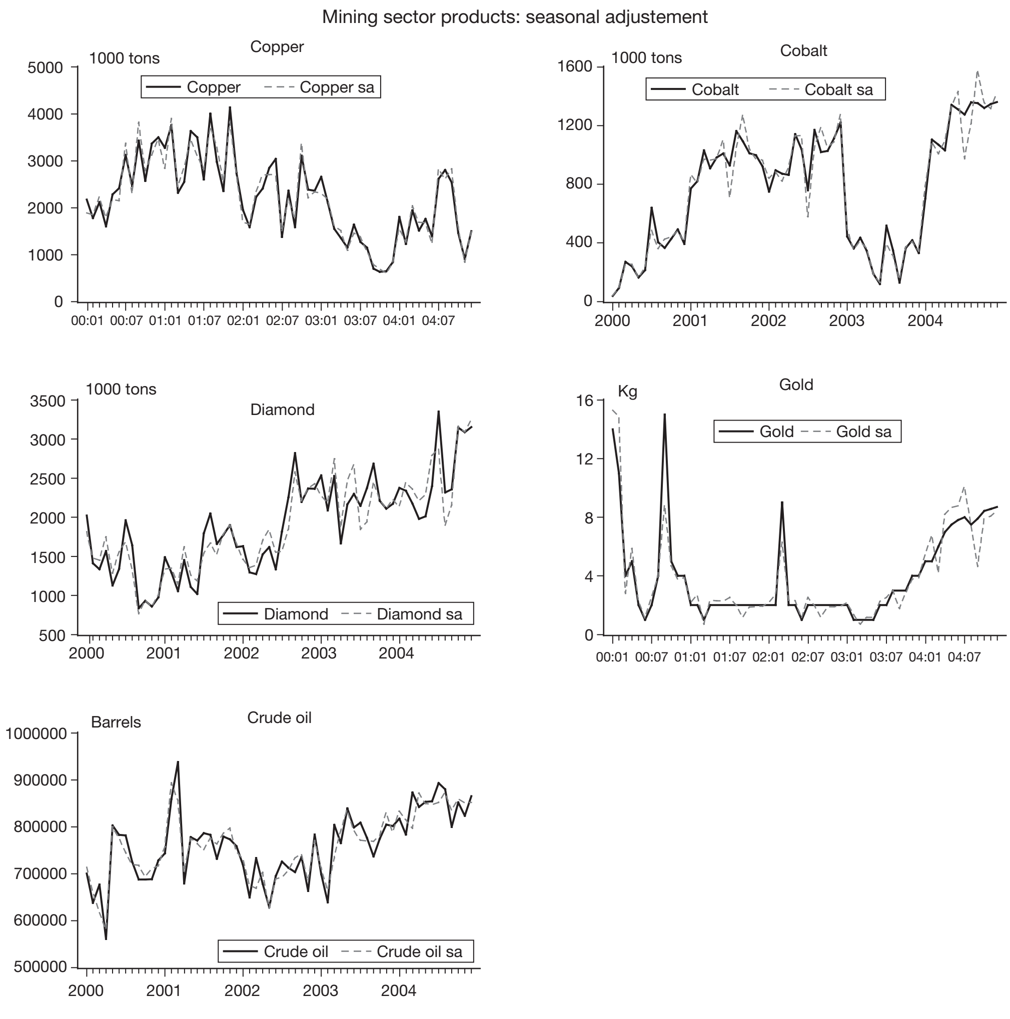
Theses estimates results are presented in current values and in volume in a synoptic table at the year 2000 prices.

Given the fact that the annual base is the key factor for national accounts in constant values, monthly developments by sector are calculated on an annual basis to determine added values of which sector growth rate will be calculated on a monthly basis with the compound average formula (1).

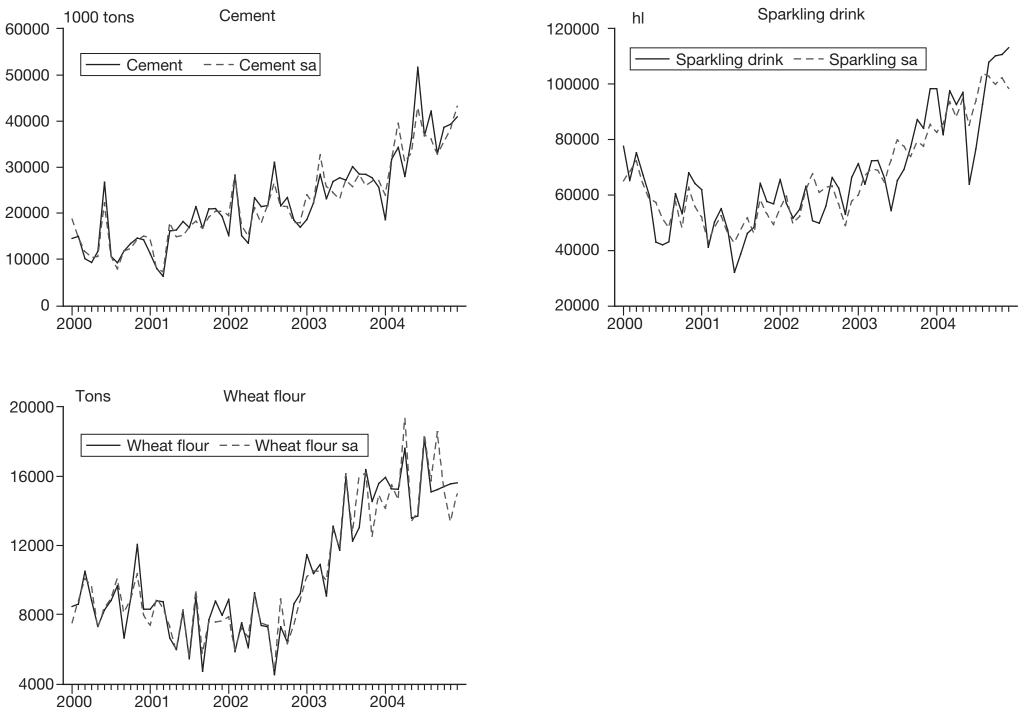
The relation variant (1) can also be used to calculate the monthly GDP from annual data (2).

In principle, monthly added values aggregate by sector should give the annual GDP. However, some gaps between the annual GDP calculated traditionally and the monthly GDP might remain. Added value monthly estimates concern the following sectors:

- Agriculture, fishing, farming and forestry;
- Mining and crude oil extraction;
- Manufacturing industries;
- Power and water production;
- Construction;
- Trade;
- Transportations, warehousing and communications;



Manufacturing products: seasonal adjustment



- Trading services;
- Non trading services;
- Import duties and taxes.

In the sighting of different statistics, the Central Bank of Congo corrects its monthly series of seasonal factors in applying ways for smoothing the data in view of stabilizing the monthly GDP. The above graphics illustrate the unseasonable series for the mining and hydrocarbons sector as well as the manufacturing sector.

2.3.1. Agriculture, breeding, hunting and forestry

The monthly production index of the agricultural sector includes major products of traditional and modern sectors from key production regions of the country. Working jointly with the Food and Agriculture Organisation (FAO), the Service National des Statistiques Agricoles, SNSA,

Table 1A – Average agriculture product monthly growth rate

Products	2003	2004
Agricultural crops products (A01)		
Maize	−0.001	0.002
Rice	−0.015	0.019
Mil	0.051	0.062
Cassava	0.008	0.003
Potato	0.121	0.047
Indian potato	−1.165	−0.045
Bean	0.076	0.073
Arachide	0.097	0.097
Banana	5.548	−0.043

Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques et INS. Calculs faits sur base des données du SNSA.

provides data. It is worth pointing that agriculture seasonal factors that characterized the agriculture sector make monthly data unavailable.

Monthly trends are generated from quarterly and annually data calculated on a monthly basis.

With regard to the agriculture which represents an important part in the African economies and in taking into account some difficulties to get monthly data of this sector, the Central Bank of Congo collects its annual data on a monthly basis in order to let appearing a monthly GDP. However, efforts are being set up in order to intercept the monthly data of this sector directly.

Monthly trends are calculated product by product by applying the annual average growth rate of production of the preceding period. The production level is often disturbed by information on rainfalls and epidemic risks which affect different cultures.

$$Prod_t = Prod_{t-1} * (1+gwr/100) \quad (3)$$

where

$Prod_t$ = annually product at current year;
$Prod_{t-1}$ = last annual product;
gwr = average monthly product growth rate.

Table 1B – Agriculture annually products

Products (10^3 t)	weight	2002	2003	2004	Product index (2004/2003)	P*W
		Product (in millions cdf)				
Agricultural crops products						
Maize	10.10	11665.50	11663.48	11665.80	1.00	10.10
Rice	17.00	5355.00	5345.31	5357.21	1.00	17.04
Mil	19.00	687.80	691.98	697.11	1.01	19.14
Cassava	3.90	58225.44	58283.82	58307.03	1.00	3.90
Potato	12.60	2771.69	2812.19	2828.07	1.01	12.67
Indian potato	33.10	3250.42	2824.09	2808.87	0.99	32.92
Bean	23.10	2480.94	2503.81	2525.75	1.01	23.30
Peanut	28.70	10202.85	10321.67	10442.50	1.01	29.04
Banana	6.90	5488.79	10492.35	10438.80	0.99	6.86
Subtotal of agriculture crops products	154.4					154.98
Percentage change						0.37

Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques et INS.
Calculs faits sur base des données du SNSA.

Table 1C – Agriculture products

Agricultural non crops products	Weight	2002	2003	2004	Product index (2004/2003)	P*W
		Product (t)				
Coffee	90.50	164227.33	130939.31	130939.31	1.00	90.50
Cocoa	30.90	43466.00	41023.86	41023.86	1.00	30.90
Rubber	109.20	3057.60	247484.55	247484.55	1.00	109.20
Palm oil	30.00	123554.84	117628.29	117628.29	1.00	30.00
Subtotal	260.60					260.60
Percentage change						0.00

Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques et INS.
Calculs faits sur base des données du SNSA.

Table 1D – Breeding products

Breeding products	Weight	2002	2003	2004	Product index (2004/2003)	P*W
		Products (10^3 head)				
Bovines	124.80	1566614.40	1510655.29	1510655.29	1.00	124.80
Sheep	139.30	385721.70	379823.40	379823.40	1.00	139.30
Caprine	139.30	2565209.50	2525438.81	2525438.81	1.00	139.30
Porcine	132.40	3139204.00	2992542.89	2992542.89	1.00	132.40
Poultry	73.80	777852.00	777999.60	777999.60	1.00	73.80
Game	152.90	14220410.30	14220403.34	14220403.34	1.00	152.90
Subtotal	762.50					762.50
Percentage change						0.00

Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques, Economiques et INS.
Done by using the SNSA data.

Table 1E – Forest products

Products	Weight	2002	2003	2004	Product index (2004/2003)	P*W
		Product (in millions cdf)				
Rough timber (10^3 m^3)	8.7	146629.8	473436.6	569848.6978	1.203643102	10.4717
Firewood (t)	3	9420.84	9694.04436	9975.171645	1.029	3.087
Coal wood (t)	5.8	13553.904	13946.9672	14351.4293	1.029	5.9682
Subtotal	17.5					19.527
Percentage change						11.58

** En thousand of cube metre.*
Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques et INS.
Done by using SNSA data.

Table 1F – Fishing products

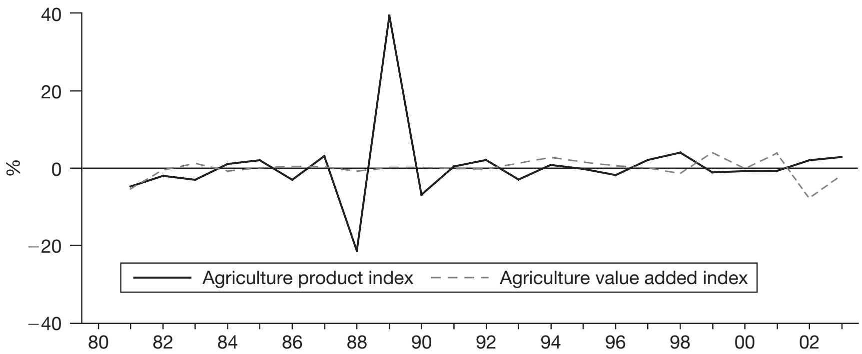
Fishing products (B)	Weight	2002	2003	2004	Product index (2004/2003)	P*W
		Product (10^3 t)				
Fresh fish	113.7	63669.73	63444.60	65157.60	1.03	116.77
Subtotal	113.7					116.77
Percentage change						2.70
Sum agriculture	1308.70	23322720.36	23607393.56	24230331.1	1.026387391	1343.23
Percentage change						2.6387

Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques et INS.
Done by using SNSA data.

The tables 1A to 1F give the way to calculate the index volume of agricultural index in 2004 from first quarter data which are annualized. The different steps are the following:

1) Production index numbers are calculated by the production ratio volume of production of 2004 and 2003;
2) Theses index numbers are multiplied by balance of each product weighting according to their relative importance in the structure of household expenditure. Theses weightings derive from last household budget- expenditure enquiry which dates back to 1986;

Graph 3 – Relationship within agriculture added value index & agriculture product index (1980–2003)



Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques.

3) The growth rate in volume of sub sector and sector is obtained by the variation of point (2) products aggregate and the total of weightings;
4) This rate is then applied to the base period added value, that is VA_{t-1} to determine the of VA_t level.

It is worth that added values thus calculated are expressed in annual values at the year 2000 prices. They are divided by 12. The subdivision on months is done according to the adjusted development of data related to rainfall and the cultivated area.

These steps are taken into consideration while calculating the added value of other sectors of economy activities. Merely monthly available data for others economic sectors. Then it is possible to calculation product monthly indexes.

2.3.2. Mining, metallurgy and crude oil

This production index number of this sector is monthly calculated from production data of mining and crude oil exploitation companies.

Table 2 – Mining and crude oil product

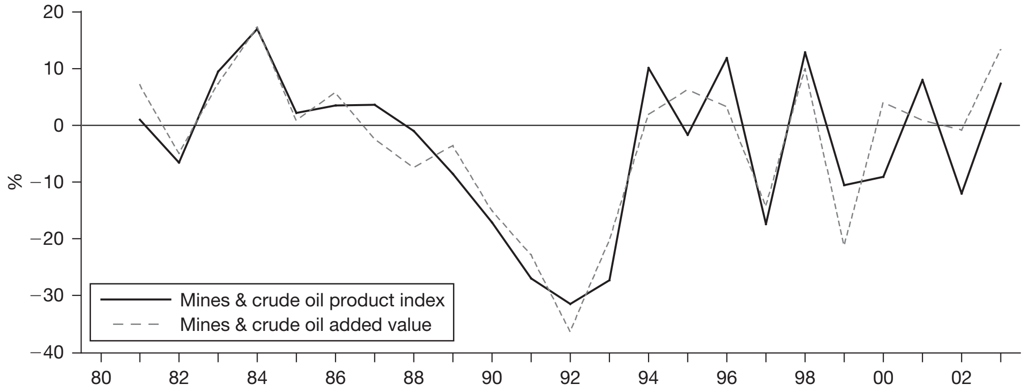
Products	Weight	Dec. 2003	Jan. 2004	Feb. 2004	Product index Jan. 2004	Product index Feb. 2004	P*W Jan. 2004	P*W Feb. 2004
Copper (t)	30	830	1,562	1,271	1.88	0.81	56.49	24.42
Cobalt (t)	54	345	791	1,090	2.30	1.38	124.01	74.42
Diamonds (10^3 carats)	585	2,230	2,134	2,448	0.96	1.15	559.81	671.31
Gold (kg)	28	4	5.5	6.8	1.41	1.24	39.47	34.71
Crude oil (barrel)	303	788,688	833,187	814,585	1.06	0.98	320.10	296.24
Total	1000						1099.868	1101.088
Percentage change							1.10	1.10

Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques et INS.
Done by using data from enquiries.

The monthly value added is calculated according to the monthly production index numbers of the mining sector calculated during the inquiries.

1,004 is the elasticity calculated on annual series of the added value index and the mining production index. The monthly data of the production sector and hydrocarbon production index to determine a close indicator of this sector added value.

Graph 4 – Mines crude oil added value added index & mines product index (1980–2003)



Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques.

2.3.3. Manufacturing industries

The sector key product helps calculate the production index number of this sector.

The analysis of the added value of annual statistical series in volume of manufacturing industries and production index of the same sector shows an elasticity of 1,005. This can be regarded as the basis in building up volume indexes to extrapolate a monthly indicator representing the sector added value.

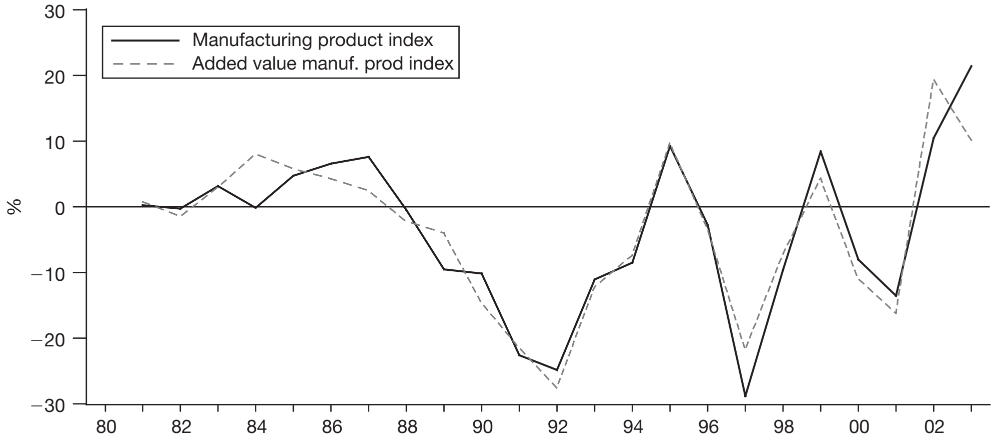
The progression pace of the sector monthly combined index is used to adjust the close indicator of the sector added value.

Table 3 – Manufacturing industries

Products	Weight	Dec. 2003	Jan. 2004	Feb. 2004	Product index Jan. 2004	Product index Feb. 2004	P*W Jan. 2004	P*W Feb. 2004
Wheat flour (t)	156	14,929	14,159	15,514	0.95	1.10	147.95	170.93
Palm oil (t)	86	737	742	738	1.01	0.99	86.58	85.54
Biscuits (t)	4	20	23	23	1.15	0.98	4.60	3.91
Beers (hl)	210	118,758	126,251	130,197	1.06	1.03	223.25	216.56
Sparkling drink	52	85,592	82,552	85,652	0.96	1.04	50.15	53.95
Cigarettes (10^6 unit)	117	233	234	240	1.01	1.03	117.60	120.00
Printed fabric (m^2)	190	378	380	375	1.01	0.99	191.01	187.50
Shoes (10^3 unit)	71	557	560	707	1.01	1.26	71.38	89.64
Paint (t)	6	4,421	4,562	4,538	1.03	0.99	6.19	5.97
Saop (10^3 unit)	48	898	903	902	1.01	1.00	48.27	47.95
Cement (t)	60	27,079	23,932	31,720	0.88	1.33	53.03	79.53
Total	1000						1000.009	1061.478
Percentage change							1.00	1.06

Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques et INS.
Done by using data from enquiries.

Graph 5 – Added value manufacturing product index (1980–2003)



Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques.

2.3.4. Electricity and water production

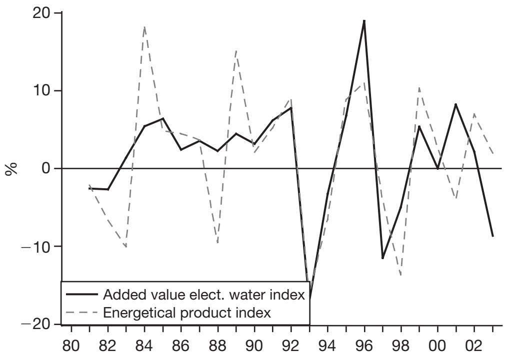
The monthly added value is determined by the adjustment of the progression pace of the energy sector combined monthly index. This adjustment is done by applying the elasticity of 1.01 to the combined index of water and electricity consumption.

Table 4 – Electricity and water production

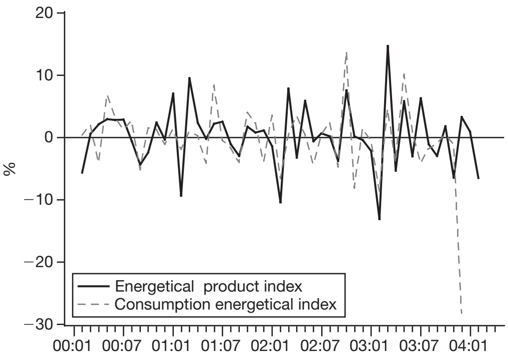
Products	Weight	Dec. 2003	Jan. 2004	Feb. 2004	Product index Jan. 2004	Product index Feb. 2004	P*W Jan. 2004	P*W Feb. 2004
Water (10^3 m^3)	653	17,338	17,139	16,123	0.99	0.94	645.51	614.26
Electricity (Mwh)	347	494,187	518,882	478,894	1.05	0.92	364.34	320.26
Total	1000						1009.851	934.5195
Percentage change							1.009851	0.934519

Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques et INS.
Done by using data from enquiries.

Graph 6 – Electricity, water added value index & energetical product index (1980–2003)



Graph 7 – Energetical product index & energetical consumption index (December 2000–March 2004)



Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques.

2.3.5. Construction

The sector added value monthly estimate is thought to develop according to consumption index of cement, coal timber and paint consumption done by public works, individual construction as well as rural construction. All those constructions are explained by the demographic growth rate.

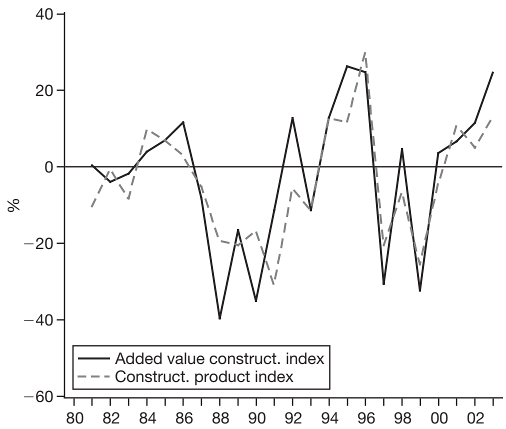
Table 5 – Construction

Products	Weight	Dec. 2003	Jan. 2004	Feb. 2004	Product index Jan. 2004	Product index Feb. 2004	P*W Jan. 2004	P*W Feb. 2004
public works (10^6 cdf)	230	375.00	379.22	380.18	1.01	1.00	232.59	230.58
Buildings								
– Consumpt° cement (t)	497	20236.29	14822.00	25833.00	0.73	1.74	364.03	866.21
– Coal wood (m^3)	47	752.51	3611.00	3652.24	4.80	1.01	225.54	47.54
– Paint (t)	26	4,421	4,562	4,538	1.03	0.99	26.83	25.86
– Rural buildings (10^6 cdf)	200	725.73	727.33	728.93	1.0022	1.0022	200.4400	200.4400
Total	1000						1049.42	1370.638
Percentage change							1.05	1.37

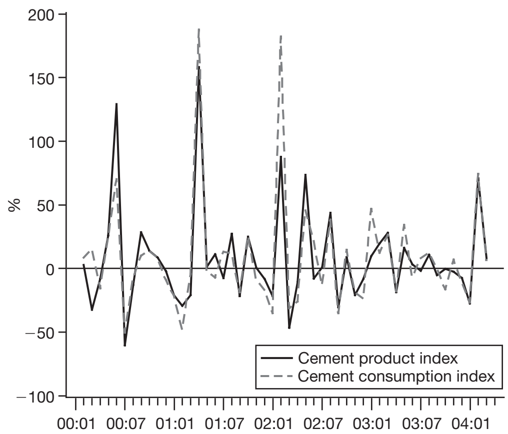
Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques et INS.
Done by using data from enquiries.

Thus historical series of the added value index in construction volume and that of cement consumption show a constant elasticity of 1.03.

Graph 8 – Added value construction index & cement consumption index (1980–2003)



Graph 9 – Cement product index & cement consumption index (December 2000–March 2004)



Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques.

2.3.6. Trade

Although wholesale and retail trade production is measured by the total value of profit margins from goods that wholesalers and retailers buy fro resale, the weighted total of agriculture and manufacturing industries production indexes and that of import index show a rough estimate of the sector activity on a monthly basis.

Table 6 – Trade & commerce

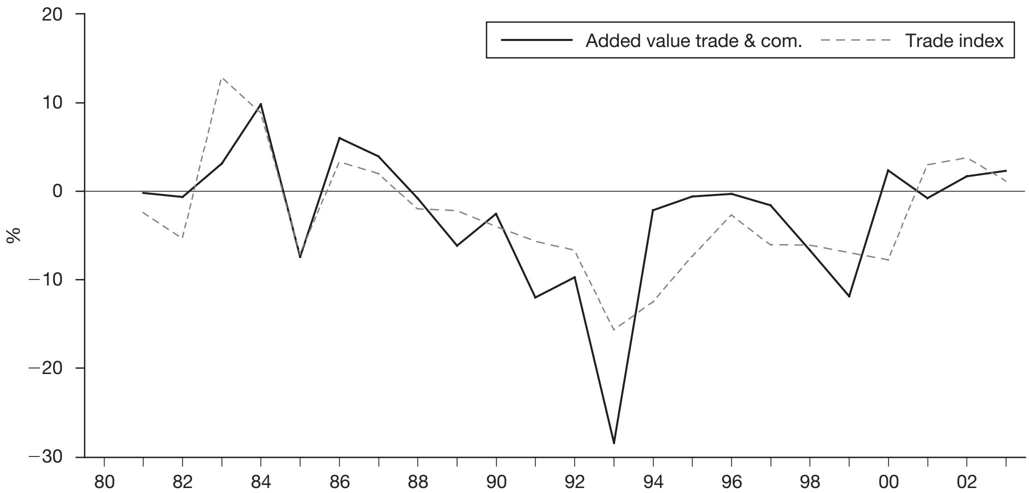
Products	Weight	Dec. 2003	Jan. 2004	Feb. 2004	Product index Jan. 2004	Product index Feb. 2004	P*W Jan. 2004	P*W Feb. 2004
Agriculture	182	1.08	0.95	0.98	0.98	0.97	178.36	176.54
Manufacturing	236	1.04	1.00	1.06	0.92	1.05	218.23	246.97
Imports (10^6 cdf)	582	46564.87	45324.87	46763.98	0.72	1.42	419.79	824.20
Total	1000						816.38	1247.71
Percentage change							−18.36	24.77

Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques et INS.
Done by using data from enquiries.

Obtained on the basis of production development in volume of coffee, cocoa and rubber as well as that of manufacturing and goods and services import, the agriculture composite weighted index number is the added value proxy variable of the month's wholesale and retail trade.

Given the open economy of the DRC, mainly natural products export all over the world, goods and services import volume index is the trade activity measure on a monthly basis. 1.01 is the elasticity value between the trade added value and this indicator.

Graph 10 – Trade & commerce added value & trade index (1980–2003)



Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques.

2.3.7. Transport and telecommunication

The sector index number of transport, warehousing and communication is determined by using the relation between the sector and that of transportable and import goods. In other words, transport and communication activity estimation is obtained by taking as indicators the weights total of travellers and goods traffic, warehousing and communication by mobile telephone. Traveller traffic which increases according to the growth rate of population and that of communication by mobile telephones expressed in terms of calling time.

The composite index number of the travellers and goods traffic, warehousing and telecommunication is the major indicator capable of providing more or less reliable indicators of the added value measure of the sector monthly activities.

Table 7 – Transportations and telecommunications

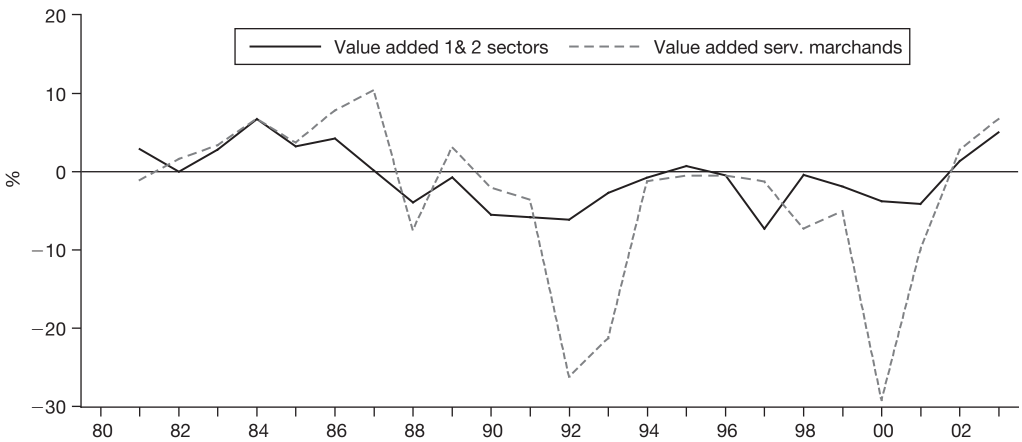
Products	Weight:	Dec. 2003	Jan. 2004	Feb. 2004	Product index Jan. 2004	Product index Feb. 2004	P*W Jan. 2004	P*W Feb. 2004
Traveller's traffic	100	256.87	257.44	258.00	1.06	1.21	100.22	100.22
Goods traffic	198	205.87	212.05	218.41	1.13	1.14	203.94	203.94
Warehousing	239	155.72	142.35	166.93	0.91	1.07	218.48	280.26
Communications	100	1739.81	1662.68	1700.26	0.96	0.98	95.57	102.26
Total	637						618	687
Percentage change							0.97	1.08

Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques et INS.
Done by using data from enquiries.

2.3.8. Trading services

Given problems met in getting monthly information about trading services production, the quantitative survey of the 1980–2002 period historical and chronological series shows that there is a more or less strong correlation between goods sectors added value development, that is primary and secondary sectors, and that of trading services as shown on the table below:

Graph 11 – Primary & secondary value added and services marchands (1980–2003)



Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques.

The bi-logarithmic function estimation of the primary and secondary sectors' added value and that of the years 1980–2003 trading services has given the following results:

Dependent Variable: LVASERMAR
Method: Least Squares
Date: 06/01/04 Time: 12:06
Sample (adjusted): 1981 2003
Included observations: 23 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	−6.757278	1.184068	−5.706834	0.0000
LVASERMAR(−1)	0.563127	0.090395	6.229604	0.0000
LVASECTBIENS	1.406105	0.248256	5.663939	0.0000
R-squared	0.970094	Mean dependent var		3.820543
Adjusted R-squared	0.967103	S.D. dependent var		0.412888
S.E. of regression	0.074887	Akaike info criterion		−2.224557
Sum squared resid	0.112162	Schwarz criterion		−2.076449
Log likelihood	28.58240	F-statistic		324.3796
Durbin-Watson stat	1.707080	Prob(F-statistic)		0.000000

where

LVASERMAR = Trading services added value logarithm;
LVASERMAR = Trading services added value logarithm of the preceding year;
LVASECTBIENS = Primary and Secondary sectors added value aggregate logarithm

The econometric relation shows 1.41 elasticity between the development of trading services added value and that of primary and secondary sectors' added value aggregate.

This helps extrapolate monthly trading services added value as soon as those of agriculture, mining and crude oil exploitation, manufacturing industries, construction and water and power generation.

Table 8 – Services marchands

Products (10^6 cdf)	Dec. 2003	Jan. 2004	Feb. 2004	Product index Jan. 2004	Product index Feb. 2004	P*W Jan. 2004	P*W Feb. 2004
Primary value added sector	112,437	123,568	147540.2	1.099	1.194		
Secondary value added sector	23857.5	22306.76	24247.45	0.935	1.087		
Total	136,294	145,875	171,788	1.070	1.178		
Percentage change						9.848	24.88706

Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques et INS. Done by using data from enquiries.

Advertisement, hostelry and mobile telephones price index numbers are proxy variables for the sector monthly added value estimation.

2.3.9. Non trading services

Non trading services monthly extrapolation is obtained by deflating the public sector wages total amount value by the average index number of consumption prices.

Table 9 – Services non marchands

Products (10^6 cdf)	Dec. 2003	Jan. 2004	Feb. 2004	Product index Jan. 2004	Product index Feb. 2004	P*W Jan. 2004	P*W Feb. 2004
Public sector wage	4,799	4951.73	4954.20	1.03	1.00		
Total							
Percentage change						3.17	0.05

Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques et INS. Done by using data from enquiries.

2.3.10. Import duties and taxes

Data related to duties and taxes given monthly by the Office des Douanes et Accises (OFIDA) provide a rough estimate of this sector.

Import duties and taxes volume index number is calculated by deflating the face value by the import price index number. However, the unavailability of the import price index number on a monthly basis leads us to circumvent this difficulty by using the following relation:

$$P_{imp} = P_{int} * e_t \quad (4)$$

where

P_{imp} = Import price index;
P_{int} = domestic price index (at consumption);
e_t = Nominal exchange rate of national currency to the US$.

By logarithmic transformation, we obtain the relation (5)

$$Lp_{imp} = Lp_{int} + Le_t \qquad (5)$$

where

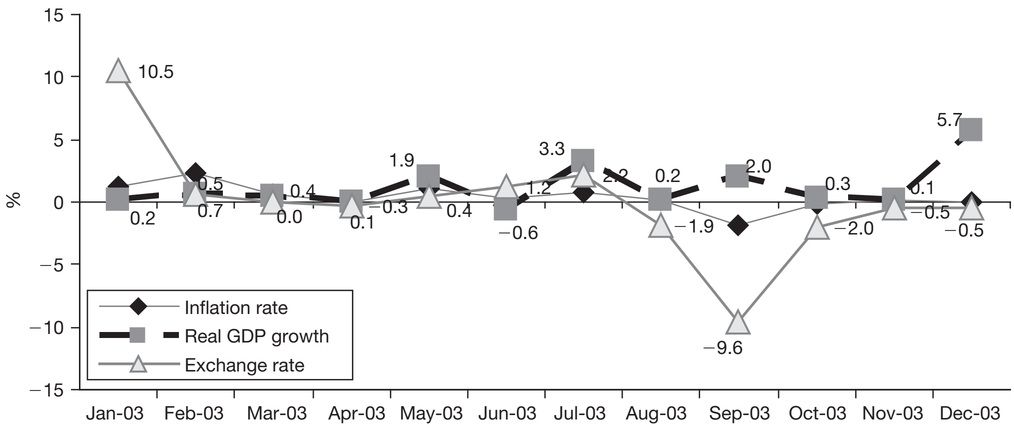
Lp_{imp} = Percentage change of import price index;
LP_{int} = Domestic inflation rate;
Le_t = Percentage change of exchange rate (CDF/1 US$).

Table 10 – Import duties

Products (10^6 cdf)	Dec. 2003	Jan. 2004	Feb. 2004	Product index Jan. 2004	Product index Feb. 2004	P*W Jan. 2004	P*W Feb. 2004
Import duties	4,936	4665	5.12E+03	0.95	1.10		
Total				−5.50	9.70		
Percentage change							

Sources: Banque Centrale du Congo, Direction des Etudes, Statistiques Economiques et INS.
Done by using data from enquiries.

Graph 12 – GDP growth, inflation rate & exchange rate (January–December 2003)



Graph 12 illustrates the monthly evolution in percent of the GDP, the inflation rate and the nominal exchange rate.

According to correlation that exists between the inflation rate and the change in exchange rate, this graph indicates that the pick periods of appreciation of Congolese franc to the American dollar are revealed by the regression of inflation and a tendency to an acceleration of growth.

An econometric study to be done on a long period should be done to confirm this analysis.

2.4. Use of the estimated monthly GDP

The elaboration of estimated monthly GDP is done to respond to triple goals that are:

- the monthly of annual GDP;
- the rapid description of macro economical activity;
- the availability of data in order to do previsions.

2.4.1. GDP monthly estimate

Unlike national accounts estimated annually for mid- and long term planning works, the monthly GDP estimate helps a more or less precise follow up of development in order to identify different steps of development and value reaction deadlines between key macroeconomic variables. Besides, the GDP monthly estimate helps to get long series so as to improve economic studies.

2.4.2. Description

Getting quickly relatively through macroeconomic information is the result expected of the GDP monthly estimate.

Establishing this process has helped to create some development indicators to provide major information on infra-annual fluctuations of the economy.

2.4.3. Estimates making

The GDP monthly estimate helps to make estimates within more or less than two years time. In fact, these accounts help to detect any dates of change and the appreciation of deadline structures. They also profit by long statistical series useful in making up a data bank.

Conclusion

The fast estimate of the monthly GDP is an important exercise for the DRC economy as the Government Economic Program implementation backed financially by the IMF and the World Bank, requires available statistical information for the follow up of quantitative criteria.

The GDP monthly estimate will help to make up macroeconomic indicators which will help decision makers to improve the ways of leading correctly the economy.

Despite gaps between the GDP growth rate estimated monthly and annually, the DRC delegation wishes to capitalize the Basel Conference attainments by adopting the GDP monthly estimate methodology so as to improve how to read and estimate the economic activity.

Gerard Mutombo Mule Mule and Marie-Jose Ndaya Ilunga (*Central Bank of Congo*)

Quality measures for quarterly national accounts and related external statistics[1]

Jorge Diz Dias and Wim Haine (European Central Bank);
Ingo Kuhnert (European Commission) [2]

Executive summary

The paper presents different quality indicators for **euro area**[3] Quarterly National Accounts (QNA) main aggregates and external statistics (balance of payments (BOP) and international investment position (IIP)). These indicators were developed in the context of a joint European Commission (Eurostat)/European Central Bank Task Force (TF) that was established in order to assess the various dimensions of output quality in QNA and external statistics. The IMF's Data Quality Assessment Framework (DQAF), supplemented by the European Statistical System (ESS) quality definitions, was used as a reference structure.

Quality indicators for QNA

Four DQAF elements (scope, source data, revisions and timeliness) were identified for which **quantitative indicators** and recommendations concerning their publication were developed.

Scope describes the QNA main aggregates data availability and is assessed separately within each of the various dimensions of prime interest to users. Press releases should include a statement on the nature of the (seasonal and working-day) adjustment applied to the series. After significant changes in availability, there should also be a note on the length of time series available for the euro area.

Source data indicators describe the quantity of information available to Eurostat for compiling the estimates of quarterly euro area GDP. Press releases should mention the proportion of euro area GDP supported by the available GDP data from Member States. They may also include the proportion of final quarterly information that underlies the euro area estimate, after this summary indicator has been sufficiently tested.

It is useful to include in press releases some information on **revisions** to euro area GDP estimates. A brief statement on the average revisions to estimates, complemented by a frequency distribution of past revisions, is considered most appropriate.

On **timeliness**, press releases should show the date of the next GDP quarterly release (indicating the timeliness of the release) and a link to the published release calendar on the website.

In addition to these quantitative indicators, four DQAF elements were identified for which **qualitative statements** should be drawn up. These relate to **relevance** (detailing the consultation process), **transparency** (e.g. announcing major methodological changes), **metadata accessibility** and **consistency** (detailing the various dimensions and degree of consistency, e.g. time consistency and consistency with other data sets). Apart from major methodological changes and a hyperlink to metadata, information on these DQAF elements should not go into press releases, but rather in a background quality document. In any case, a background quality document should cover all quantitative and qualitative indicators in more detail.

The Committee on Monetary, Financial and Balance of Payments statistics (CMFB) supported the above recommendations and suggested their implementation in a co-ordinated

1 *Prepared for the IFC Conference in Basel, September 2004.*

2 *The views expressed in this paper are those of the authors and do not necessarily reflect the views of the European Central Bank or the European Commission.*

3 *The quality indicators discussed in this paper also apply to European Union (EU) aggregates.*

procedure between Eurostat and the National Statistical Institutes. Eurostat with the support of the ECB intends to publish this type of quality information for the first time end 2004 or early 2005.

Quality indicators for external statistics

For the external statistics three elements are identified for which the calculation of quantitative indicators on a regular basis are envisaged. These comprise indicators about **revision studies**, **consistency** and, additionally, **plausibility**. Regarding qualitative statements, four dimensions/elements of the DQAF, namely **methodological soundness**, **timeliness**, **revision policy and practice** and **accessibility** may be assessed in the quality report.

The external statistics quality report will be prepared to address simple characteristics of each dimension/element. It will also provide in-depth analysis for those quality aspects, which require improvements. To accomplish this, the quantitative indicators will be calculated and analysed twice a year (in May and November). This exercise is used as input for an in-depth annual quality report for the euro area external statistics covering both quantitative indicators and qualitative assessments.

1. Introduction

For monetary policy purposes high quality economic and financial statistics are indispensable. The 2002 Fourth Progress Report of the Economic and Financial Committee (EFC) on the Statistical Requirements in the EMU stated that *more work is needed to operationally assess the various dimensions of quality*. The Statistical Programme Committee (SPC), in collaboration with the CMFB, was invited to make proposals in this respect.

As a result, a joint European Commission (Eurostat)/European Central Bank (Directorate General Statistics) Task Force (TF) on quality was established in order to assess the dimensions of output quality in Quarterly National Accounts (QNA) and in Balance of Payments statistics (BOP), in view of offering guidance to data users. The IMF's Data Quality Assessment Framework (DQAF), supplemented by the European Statistical System (ESS) quality definitions, was used as a reference structure.

Eurostat publishes three releases of **euro area** QNA main aggregates each quarter. The flash estimate covers quarterly GDP growth and is published 45 days after the end of the reference quarter. Two subsequent releases ($t + 65$ and $t + 105$ days) comprise additional data on the output, expenditure and income side. The European Central Bank (ECB) publishes monthly BOP statistics for the **euro area** at $t + 56$ days. More detailed quarterly BOP and annual international investment position (IIP) statistics are publish at $t + 91$ and $t + 330$, respectively. From March 2005 onward the ECB will publish quarterly IIP for the euro area at $t + 91$ days.

The paper discusses quality indicators for **euro area** QNA main aggregates and external statistics. It builds on the TF's findings, which were supported by the CMFB in 2003 and 2004. For each statistical domain, the paper briefly explains the quality framework that was used as a reference. Second, quantitative quality indicators are presented to gauge. Third, quality areas that are not easily quantified but merit special attention are addressed. Moreover, the paper puts forward suggestions for communicating quality indicators to all users. Finally, the paper presents some concluding remarks.

2. Quality indicators for QNA

At present, the information that National Statistical Institutes (NSIs) provide on the quality of their GDP estimates varies widely, and only few examples exist of systematic, comprehensive publications of such information. The US Bureau of Economic Analysis (BEA) has a longstanding tradition of publishing detailed revisions studies, and also supplies substantial background information on the information sources that went into the compilation of National Accounts. The Australian Bureau of Statistics (ABS) is running a comprehensive quality review, but has not published its results yet. Statistics Canada has endorsed mandatory documentation standards in its 2002 quality framework. While the standards are adhered to, the available information does not allow for any easy quality assessment. Although there is a great awareness of quality issues in national accounts amongst EU Member States' NSIs, users may find it difficult to form an opinion on the basis of information given in regular publications.

However, it is worth mentioning the UK Office of National Statistics (ONS) in this context for the background notes attached to its press notices.

Reporting on quality is usually done within a quality framework, which may be constructed by the NSI itself, or adapted from a "standard" set forth by for example the IMF, the OECD or the European Statistical System.

For the purpose at hand, the IMF's Data Quality Assessment Framework (DQAF[4]) was taken as a common language and structure on which to base the work, with the European Statistical System (ESS) quality definitions supplementing it.[5] One particular advantage of the DQAF is that the IMF has published a quality assessment structure specifically for national accounts.

The DQAF follows a cascading structure, from the abstract/general to the more concrete/specific, and is divided into five dimensions, each comprising a number of elements that are in turn associated with a group of potential indicators.

Within the IMF's DQAF framework, elements of quality were divided into three groups:

- The **first group** consists of those four elements for which **quantitative indicators** could be meaningfully constructed for **quarterly euro area GDP and its expenditure components**.
- The **second group** consists of those four elements for which, although meaningful quantitative indicators could not be identified, publishable **qualitative statements** could be made.
- The **third group** consists of those elements that were considered of little relevance in terms of the TF's mandate or did not offer prospects for successful construction of quality indicators.

The following table shows a full list of the IMF quality elements, and how each of these elements was categorised into the three groups:

Table 1 – Overview of IMF Data Quality Assessment Framework

Quality dimensions	Elements	TF recommendations for indicators
0. Prerequisites of quality	0.1 Legal and institutional environment	–
	0.2 Resources	–
	0.3 Relevance	**Qualitative**
	0.4 Other quality management	–
1. Assurances of Integrity	1.1 Professionalism	–
	1.2 Transparency	**Qualitative**
	1.3 Ethical standards	–
2. Methodological soundness	2.1 Concepts and definitions	–
	2.2 Scope	**Quantitative**
	2.3 Classification/sectorisation	–
	2.4 Basis for recording	–
3. Accuracy and reliability	3.1 Source data	**Quantitative**
	3.2 Assessment of source data	–
	3.3 Statistical techniques	–
	3.4 Assessment and validation of intermediate results and statistical outputs	–
	3.5 Revision studies	**Quantitative**
4. Serviceability	4.1 Periodicity and timeliness	**Quantitative**
	4.2 Consistency	**Qualitative**
	4.3 Revision policy and practice	–
5. Accessibility	5.1 Data accessibility	–
	5.2 Metadata accessibility	**Qualitative**
	5.3 Assistance to users	–

4 *Carson, Carol S. "Toward a Framework for Assessing Data Quality," IMF Working Paper, WP/01/25. IMF 2001. For the most recent version of the DQAF, refer to http://www.imf.org/external/np/sta/dsbb/2003/eng/dqaf.htm#ap.*
5 *"Quality in the European statistical system – the way forward", European Commission (Eurostat), 2002.*

2.1. Quantitative indicators

Four DQAF quality elements were identified for which quantitative indicators should be regularly computed. These are discussed below.

2.1.1. Methodological soundness – scope

The DQAF defines scope largely in terms of being in accordance with internationally accepted standards (SNA). It lists the range and breakdown of national accounts aggregates that should be regularly compiled and published. The argument is that the higher the quantity and diversity of data within this area of statistics, the better users' needs are met and the easier it is for users to interpret the data.

In order to put this concept into practice, scope is described in terms of length of time series, transaction breakdowns, availability of seasonally adjusted and unadjusted series, etc. These dimensions are derived from the compulsory transmission programme used between NSIs and Eurostat. The construction of a single indicator that includes all dimensions at once by comparing actual scope to maximum scope, i.e. calculating the percentage share of data cells actually provided is straightforward, but does not give convincing results. In particular the length of time series tends to dominate the other dimensions, and integrating additional dimensions has an unduly negative effect on the scope indicator. Solving these problems by introducing a weighting system on the various dimensions, attributing more importance to the most important aggregates and the current quarters proved to be complex and incomprehensible to users.

Therefore, scope is shown separately within each of the dimensions that are deemed to be of focal interest to most users. These are:

- **Length of time series** – Eurostat's stated aim is to have all quarterly euro area national accounts main aggregates back to 1980 Q1. Currently the full set of GDP (and expenditure components) data for **euro area** only goes back to 1991 Q1 (i.e. 53 rather than 97 observations).
- **Number of transaction variables** – this measures the number of main aggregates currently available for the **euro area**. The number available can be compared with the number of variables required by the transmission programme specified in the Council Regulation on the European System of Accounts[6] (ESA). Currently 71 (about three-quarters) of the 96 main aggregates required in table 1 of the transmission programme are available.
- **Seasonal adjustment** – this states how many types of series are available. Currently two (unadjusted, and seasonally adjusted with partial correction for working day variation) out of the four possible types (unadjusted, seasonally adjusted, not seasonally adjusted but working day adjusted, and seasonally and working day adjusted) are available.

Breakdowns by industry or consumption purposes might be considered as additional dimensions that could be referred to in a quality assessment.

2.1.2. Accuracy and reliability – source data

The DQAF has an extensive definition of this element of quality, covering adequacy and timeliness of source data and its consistency with national accounts definitions. Covering this comprehensively is a difficult task, in particular at the European level, given the multitude of different sources used across all euro area countries. Therefore, this indicator is defined in terms of the quantity of information available to Eurostat for compiling each of the estimates of quarterly **euro area** GDP.[7]

One way of showing the information content in the GDP estimates is to show the proportion of **euro area** GDP covered by the GDP of the countries whose data have been used in their compilation. This source indicator takes no account of the quality of the national GDP estimates transmitted by countries. In particular, no account is taken of the fact that countries use incomplete (and sometimes different) information when compiling their early estimates of GDP. The source indicators can be refined by adjusting each countries' contribution

6 *Council Regulation (EC) No 2223/96 of 25 June 1996, published in the Official Journal L310 of 30/11/1996.*

7 *While only GDP estimates are referred to in the following paragraphs, the same arguments apply to the estimates of expenditure components of euro area GDP.*

to **euro area** GDP to take account of this varying information content in their national GDP estimates.

In a pilot study, National Statistical Institutes (NSIs) were asked to indicate – on the basis of a set of guidelines – what proportion of final quarterly information (i.e. not taking annual sources into account) is available to them when they transmit data to Eurostat to compile the "flash", first and second estimates of **euro area** GDP growth for the latest quarter. It is however clear that a fully harmonised, objective, and accurate assessment of the proportion is not possible. Instead, Member States were asked to provide a figure to the nearest 10%. The aim of the indicator is to give users a good, rather than accurate impression of the quantity of source data supporting the euro area estimates of GDP.

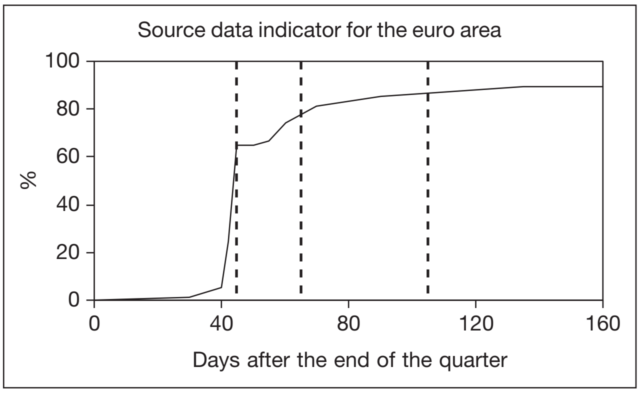
It should be stressed that the approach will require further testing across all euro area countries before a final assessment of its usefulness can be made. The following table exemplifies the indicator values for each of the three quarterly euro area GDP estimates:

Table 2 – Source data coverage for Eurostat's GDP releases

Estimation	Number of countries available for the estimation	% of GDP covered by these countries	Approx. % of total final quarterly information available for the estimation
	(1)	(2)	(3)
Flash (t + 45)	7 out of 12 (58%)	83%	65%
1st regular (t + 65)	8 out of 12 (67%)	93%	75%
2nd regular (t + 105)	10 out of 12 (83%)	98%	85%

While the indicators in columns (1) and (2) can be unambiguously calculated and interpreted, those in column (3) are the result of preliminary calculations and subject to further assessment, but they already offer valuable insight into the quality of the estimation. The GDP flash estimate for instance covers a reassuringly large part of 83% of euro area GDP, but is based on 65% of the available final quarterly information. An advanced source data indicator like the one presented in column 3 may – in combination with a revision analysis – prove a useful tool for assessing the reliability of subsequent estimates.

While the table shows this indicator for three specific points in time only, the graph below shows how it evolves over time, from the first data becoming available after thirty days to close to 90% after 160 days. The indicator does not attain 100% because some countries do not compile quarterly accounts, and others will complete their final quarterly information set only during the following quarters.



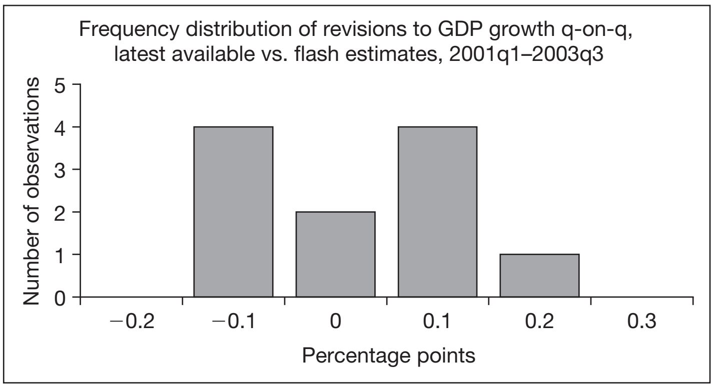
2.1.3. Accuracy and reliability – revisions

The DQAF describes this element in terms of undertaking regular analyses of revisions. Of all quality elements, it is perhaps for this one the easiest to produce large numbers of different, and sometimes quite complicated, quantitative indicators. However, these indicators may not always give a true reflection of quality. They can be misleading in the sense that few or small revisions do not necessarily imply accurate (high quality) data.

Nevertheless, users should be provided with a few simple indicators, and detailed analyses could be undertaken (and published) to provide further background information. Revision indicators should be easy to interpret from a theoretical perspective (understanding the concept) and a practical perspective (grasping the dimension of the numbers attached to these indicators). A wide range of possible indicators was tested and some proved more appropriate for the communication with the public at large than others.

In this respect, the mean (absolute) revisions and the range of revisions are considered most informative when describing bias and variance. These can be complemented with a chart showing the frequency distribution of revisions and some explanatory sentences covering information on *i.a.* the number of upward/downward revisions, the number of changes in the sign of the growth rate and how well the acceleration or deceleration of growth is captured. More complicated indicators like the Root Mean Square Error (RMSE) (combining variance and squared bias) could however be included in a more comprehensive publication.

The revision analysis found an average revision of **0.0 percentage points** of the flash estimate[8] of euro area GDP volume growth for the period 2001q1–2003q3. In addition, a recent analysis showed that in the period under investigation the sign of GDP growth was never revised and that acceleration or deceleration of GDP growth was confirmed in 8 out of 11 cases. The chart below details the distribution of the revisions to the flash estimate.



To date, the significant improvement in timeliness that was achieved with the publication of flash estimates, has not come at the cost of higher revisions to early **euro area** GDP estimates. However, the results should be interpreted with some caution, as the period under investigation is relatively short and characterised by a low level of quarterly **euro area** GDP growth. There is no guarantee that the favourable reliability of recent GDP estimates holds for the future, and it cannot be excluded that future major benchmark revisions of GDP results (as they are planned to be implemented for national accounts in 2005) will somewhat change the current picture.

2.1.4. Serviceability – timeliness

The DQAF defines timeliness in terms of its own dissemination standards. It seeks publication of quarterly GDP within three months after the end of the reference quarter. The EU has its own dissemination standards. The ESA 95 regulation sets out its timeliness requirements for GDP transmissions from Member States. In recent years, timeliness targets for the publication of **euro area** GDP have been specified in the EMU Action Plan, and in the context of Principal European Economic Indicators (PEEIs).

This aspect of quality can be shown both in terms of its past improvement, and potential for improvement towards the PEEI targets. Simply studying the changing timeliness of Eurostat's press notices does not necessarily show how performance has improved (or might improve), because over time additional press notices have been introduced, old ones have been suppressed

8 *Eurostat started publishing flash estimates with 2003q1. The earlier flash estimates used in this study were "hypothetical releases" (i.e. real time results obtained by Eurostat in the test period prior to the publication of the first flash estimate, based on the information available at that time). That information may have been somewhat different from the information Eurostat has available now. The results of these hypothetical releases were given to the ECB's DG-S on a confidential basis.*

or merged, and contents has been shifted from later to earlier press notices.[9] It is thus more appropriate to look at when key statistics such as GDP growth, GDP level etc. have actually been published.

As the table below shows, there has been a considerable improvement in the timeliness of the publication of **euro area** quarterly GDP data over the last three years. Except for GDP growth, there is still considerable room for improvement before PEEI targets are achieved.

Table 3 – Time of first release of GDP (days after end of quarter)

Variable	3 years ago	Over last 4 quarters	PEEI target
GDP growth	71	44	45
GDP level	99	74	60
GDP basic breakdowns, growth	71	65	60
GDP basic breakdowns, levels	99	74	60

2.2. Qualitative statements

As explained above, the feasibility of quantitative quality indicators has been studied for all DQAF elements. In some cases, it was found that indicators could be constructed, but that these would not be very informative to users, giving the wrong impression of exactly quantifiable quality, when in truth a substantial degree of arbitrariness is involved.

Four DQAF elements were identified for which quantitative indicators could not be meaningfully constructed, but for which qualitative statements should be published instead.

2.2.1. Serviceability – relevance

The IMF defines this element in terms of the statistics meeting users' needs, consultation processes, and whether specific actions are taken to address users' concerns. There is little doubt that euro area QNA statistics are relevant to many users: GDP press notices are amongst the most consulted of all Eurostat's releases. Furthermore, the ESA 95 Regulation (i.e. Council Regulation (EC) No 2223/96 of 25 June 1996, published in the Official Journal L310 of 30/11/1996) went through an extensive legislative process that involves consultation with users and others outside the statistical system. This consultation process should be explained to users in all openness.

2.2.2. Integrity – transparency & 2.2.3. Accessibility – metadata accessibility

Other DQAF elements that warrant special consideration are "transparency", which covers the terms and conditions under which statistics are collected, processed and disseminated, as well as "metadata accessibility", which covers the adequacy and accessibility of metadata. Users should be made aware of major changes in methodology with a brief statement on the change, its expected date of introduction, and a reference to a more detailed explanation. Similarly, users should have an easy access to adequate metadata.

2.2.4. Serviceability – consistency

According to the DQAF, "consistency" comprises internal consistency, consistency over time, and consistency with other major data sets. Internal consistency and consistency over time are deemed reasonably satisfactory in euro area national accounts main aggregates. Nevertheless, the euro area national accounts data set is large, multi-dimensional, and complicated. It is becoming increasingly difficult to maintain consistency in all dimensions. QNA for the euro area are not geographically

9 *The content of Eurostat's former second and third releases has recently been merged. Previously, the second release (t + 100 days) contained the goods and services breakdown for exports and imports – in addition to GDP and expenditure components and value added by industry which are now published in the first release. The third release (t + 120 days) additionally provided income data as well as the product breakdown for gross fixed capital formation and the industry breakdown for compensation of employees.*

consistent as not all Member States currently compile QNA.[10] Furthermore, there are inconsistencies at the national level between balance of payments and national accounts (i.e. rest of world transactions) statistics, which lead to respective inconsistencies at the euro area level. These inconsistencies are likely to become more prominent with the development of institutional sector accounts. The various dimensions and degrees of consistency should be explained to users.

2.3. Communicating on quality in QNA press releases and publications

Eurostat press releases already include some quality statements. For example, the "flash" (45 day) release comments on the number of countries that have reported their GDP in time for inclusion in the compilation of the euro area GDP. This release also gives some indication of past revisions to the early estimate of GDP.

This is already a useful start and it is considered useful to provide additional information on quality. At the same time, the dangers of including too much quality information in press releases are recognised, particularly as this could distract from the actual numbers and as most of it is expected to change little from quarter to quarter. Against this background, recommendations on the publication of quality indicators and statements in press releases and in a comprehensive quality report are detailed below.

2.3.1. The quality statement in the press releases

Comments on quality included in press notices should be brief, easy to understand, and specific to the release. What is included in Eurostat's three quarterly GDP press releases will vary somewhat according to the time of the release (45-, 65- and 105-day releases). However, the broad content should be similar for each release. The following quality issues should be included in Eurostat's press releases:

- On *timeliness*: the date of the next GDP quarterly release (indicating the timeliness of the release); and a link to the published release calendar on the website.
- On *revisions*: first, a statement on the revision of the estimate of growth in the latest quarter since the last publication. This would be included only in the 65- and the 105-day releases. Users need to know whether the estimate is unchanged since last time or has been revised. If the latter, a footnote on any major reason for revision should be added. Second, a brief statement on the average revisions to estimates at this stage in the quarterly sequence. This should be in tabular form, particularly in the 65-day and 105-day releases, if information on expenditure components is also given. Third, there should be a frequency distribution of past revisions – either as a chart or spelt out (simply) in words. Finally, there should be a link to more detailed analyses of revisions.
- On *source data*: the proportion of euro area GDP supported by the available GDP data from countries. A summary indicator of the proportion of final quarterly information that underlies the euro area estimate may also be included, after it has been sufficiently tested.
- On *scope*: a statement on the nature of the (seasonal and working-day) adjustment applied to the series. After significant changes in availability, there should be a note on the length of time series available for the euro area.
- On *metadata accessibility*: a link to Eurostat documentation on national accounts compilation and in particular to ESA 95.
- On *transparency*: as appropriate, major in particular forthcoming changes to national accounts methodology (e.g. chain linking) should be trailed in quarterly GDP releases.
- On *quality in general*: there should be a link to the standing quality document on Eurostat's website.

2.3.2. The quality document

With respect to comprehensive quality information that goes beyond the scope of a press release, the following is recommended:

- A quality document should be prepared, and structured on the basis of the IMF's DQAF. It should be published on Eurostat's web-site, and reviewed annually, and updated when necessary.

10 *Thus, they cannot be calculated by summing and must be estimated from the available data. The estimation procedure, however, gives results whose development is consistent with that of the sum of available Member States' data.*

- It should concentrate on the elements identified above as the most relevant, but, for completeness, it may make brief comments on other elements.
- All elements should be discussed in the website document, showing both the poorer aspects as well as the good ones. References should be made to planned improvements and to barriers to improvements.

At its launch, the new document should be heralded with an article in one of Eurostat's paper publications.

2.4. Timetable

The CMFB supported the above recommendations and suggested their implementation in a coordinated procedure between Eurostat and the National Statistical Institutes. Eurostat with the support of the ECB intends to publish this type of quality information for the first time end 2004 or early 2005.

3. Quality indicators for external statistics

The euro area BOP and IIP are based on the aggregation of statistics provided by countries, reflecting the transactions and positions between their residents and non-residents of the euro area. The legal framework for the provision of data to the ECB is established by the Guideline ECB/2004/15 of 16 July 2004. Taking into account, *inter alia*, the variety of methods and sources at the national level, no simple assessment, neither qualitative nor quantitative, can fully reflect the quality of the euro area statistics. However, some dimensions of quality have been assessed qualitatively, and quantitative indicators[11] have been designed to help users in the analysis of the data.

Accordingly to the IMF's DQAF framework, dimensions/elements of quality may be divided into three groups:

- The **first group** consists of those three elements for which **quantitative indicators** could be meaningfully constructed for BOP
- The **second group** consists of those four dimensions/elements for which, although meaningful quantitative indicators could not be identified, publishable **qualitative assessments** could be made.
- The **third group** consists of those dimensions/elements that were outside the mandate of the TF or did not offer prospects for successful construction of quality indicators. Notwithstanding, the TF recommends a periodic *ad hoc* control for some dimensions/elements.

The following table provides a summary of the dimensions/elements contained in each of the groups defined above:

11 *Documentation on the indicators used in this report are available on the CMFB Website (http://www.cmfb.org/) entitled: "Task Force on Quality – Report on the quality assessment of balance of payments and international investment position statistics".*

Table 4 – Overview of IMF Data Quality Assessment Framework for BOP

Quality dimensions	Elements	TF recommendations for indicators
0. Prerequisites of quality	0.1 Legal and institutional environment 0.2 Resources 0.3 Relevance 0.4 Other quality management	Periodic *ad hoc* check
1. Assurances of Integrity	1.1 Professionalism 1.2 Transparency 1.3 Ethical standards	Periodic *ad hoc* check
2. Methodological soundness	2.1 Concepts and definitions 2.2 Scope 2.3 Classification/sectorisation 2.4 Basis for recording	Qualitative assessment
3. Accuracy and reliability	3.1 Source data 3.2 Assessment of source data 3.3 Statistical techniques 3.4 Assessment and validation of intermediate results and statistical outputs **3.5 Revision studies** *3.6 Plausibility*	– – – – Quantitative *Quantitative*
4. Serviceability	**4.1 Periodicity and timeliness** **4.2 Consistency** **4.3 Revision policy and practice**	Qualitative assessment Quantitative Qualitative assessment
5. Accessibility	5.1 Data accessibility 5.2 Metadata accessibility 5.3 Assistance to users	Qualitative assessment

3.1. Quantitative indicators

The TF identified three elements of the DQAF for which the calculation of quantitative indicators on a regular basis is recommended. The quality indicators derived should be understood in relative terms and not interpreted as an absolute measure of quality. The readings from any aggregated quantitative indicator should always be performed with caution. Statistical averages and variances may hide very different realities as a few outliers can influence the results considerably.

3.1.1. Accuracy and reliability – revision studies

Although measuring the size and direction of revisions is relatively easy, a variety of possible indicators and analyses can be performed. For example, an analysis may be limited to simple descriptive statistics or, alternatively, more sophisticated statistical/econometric calculations can be performed. The choice may depend on the characteristics of the data (e.g. gross or net figures) and on the phenomenon the compiler wishes (or is able) to measure (e.g. a simple assessment of the size of revisions; a comparison of revisions across items and/or countries).

The ECB will use the mean absolute percentage error (MAPE) for measuring revisions in the current and capital accounts of the BOP For the financial account, where the data have positive and negative values, the ECB will use of the root mean square relative error (RMSRE).

The indicators should be calculated over series with a minimised deterministic component, i.e. limited trend and cyclical parts. It is important to remove the deterministic component, using, for example, first differences, growth rates or other methods. The non-deterministic component of the series should in principle be stationary, i.e. have a constant (unconditional) mean and variance.

The RMSRE indicator may further distinguish the systematic revisions from unsystematic revisions. This classification is critical because systematic revisions contain regular patterns, while unsystematic revisions are connected to unpredicted changes. The systematic component can in certain cases be addressed by statisticians within the production process. In addition, the RMSRE penalises series with a higher volatility less, as there is more uncertainty associated. Hence, it allows for higher total revisions in more volatile series since their non-deterministic component is more important.

The TF developed other support indicators like the upward revisions, which gives the percentage of positive revisions, and the directional reliability indicator, which expresses the percentage of cases in which earlier and later assessments move in the same direction.

3.1.2. Accuracy and reliability – plausibility

Although not included in the DQAF, the plausibility describes the likelihood of the data. Significant *outliers* or sudden and unexpected changes in the trend need to be investigated, especially when there is virtually no economic and/or methodological explanation for them.[12]

While the plausibility should be regularly assessed, no specific quantitative indicator is recommended. Plausibility may be covered by multiple quantitative indicators, but the indicators used are typically changing over time.

3.1.3. Serviceability – consistency

Concrete indicators measuring overall consistency across statistical series are broken down into the following sub-categories:

- internal consistency, e.g. within the integrated statistics like BOP/IIP;
- consistency over time (for example, in the case of methodological or institutional – i.e. euro area enlargement – changes, historical data are compiled as far back as reasonable);
- external consistency (between different sources of data and/or different statistical frameworks, including mirror statistics). Conceptual consistency, as highlighted by the IMF, fosters the international comparability of statistics, even when compiled by different institutions. In addition, different measurements of the same phenomenon should not result in unreasonably different data.[13]

The ECB will use the root mean square error (RMSE) of net errors and omissions for the study of internal consistency in the BOP statistical framework. This indicator is classified as a key indicator. Other supporting indicators were studied, like the count of positive net errors and omissions.

For external consistency, the ECB will use the RMSRE, like in revision studies, which is adapted straightforward to measure the discrepancy stemming from two different data sources. These two indicators are classified as potential indicators.

3.2. Qualitative assessments

The TF considers relevant to address the following four dimensions/elements in the quality report. These are discussed below.

3.2.1. Methodological soundness

The methodologies observed by Member States when compiling the BOP and IIP are covered in the country chapters of the ECB's yearly publication "European Union BOP and IIP statistical methods".[14] This publication describes the collection system in each Member State and includes details about the reporting population, administrative sources, periodicity of surveys, estimations and legal framework. In addition it also gives an overview of the compilation of the

12 *As an example, the simplest way to formalise a plausibility check is to standardise the variable x under observation. The average* $(\bar{x})$ *and the variance* (σ_x^2) *of the variable can be calculated over a specific historical time range. The plausibility check is therefore performed on the observation* x_i *with the formula* $|s_i| = |(x_i - \bar{x})/\sigma_x| > c$*, where c is the limit over which the observation is considered implausible.*

13 *For instance, consistency between aggregated BOP statistics compiled by different international organisations.*

14 *ECB, November 2003.*

euro area aggregate figures by explaining the compilation procedures and underlying methodological concepts agreed by the EU Member States. The compilation methods for international reserves (flows and outstanding amounts) are described in a separate report.[15] In addition, the ECB website contains a methodological note specific for the euro area BOP and IIP[16] focusing on common methodological issues, as well as on the aggregation procedures at the euro area level. It is updated whenever changes occur.

Changes in this dimension are better reflected through a qualitative assessment, and commented as such in the quality reports.

3.2.2. Serviceability – periodicity and timeliness

Timeliness is well covered and monitored by defining and publishing an advanced release calendar for data dissemination (including a contribution to the euro area/EU aggregate) in line with internationally accepted dissemination standards. In addition, simple indicators concerning deviations from the established timeliness, where relevant, can easily be constructed (for example, the number of delays and average/total days of delay with regard to reporting/dissemination timetables).

3.2.3. Serviceability – revision policy and practice

Like the previous element, the quality report should address changes in the revision policy/practice with a qualitative assessment describing the changes and the impact on the statistics.

3.2.4. Accessibility

Information that is not made accessible to users must be labelled poor quality, regardless of its accuracy. The developments and achievements in this dimension should be noted in the quality reports.

3.3. Communicating on quality in external statistics

The quality indicators will be calculated and analysed by compilers twice a year (in May and November). They form the input for an in-depth quality report of euro area (and possibly the EU) external statistics to be published once a year.[17] Together with it a statement will be published on the progress made and achievements so far, as well as the necessary caveats and information which will help in the interpretation of these indicators. In particular, special attention should be paid to a balanced consideration of both the quantitative as well as the qualitative indicators, noting where appropriate the trade-offs between different dimensions and elements of quality (e.g. timeliness vs. reliability).

4. Conclusion

To facilitate users' understanding of economic and financial statistics, it is important to provide them with information on the quality of statistics. Statistical authorities are best placed to inform users about the various dimensions of data quality (incl. scope, source data, revisions, and timeliness). In the absence of such information, users may infer their own conclusions. A recent article by Goldman Sachs on the extent of and reasons for revisions to euro area quarterly GDP growth may be taken as an example.[18] While the article touches upon a relevant issue, some of the findings are based on incomplete or inaccurate information and presented in an overly negative manner.

15 *Statistical treatment of the Eurosystem's international reserves; ECB, October 2000.*
16 *ECB's Monthly Bulletin – Euro area statistics, Methodological notes: Chapter 7, BOP and IIP of the euro area (including reserves).*
17 *First release is expected to occur in early 2005.*
18 *"Statisticians selling Euroland short?", Goldman Sachs – Euroland Weekly Analyst' of 16 July 2004.*

Based on the findings of the joint European Commission (Eurostat)/ECB (DG Statistics) Task Force, Eurostat and DG Statistics developed different quantitative indicators and qualitative statements that allow for an operational assessment of the output quality of euro area QNA and external statistics. The findings will be implemented in press releases and comprehensive quality reports published by end 2004 or early 2005.

Together with the efforts undertaken by other international organisations (IMF, OECD) and leading NSIs, this ECB (DG-Statistics)/Commission (Eurostat) initiative shows that it is desirable and feasible to foster an effective communication on the quality of statistics vis-à-vis policy makers, advanced users and the public at large. This initiative may be extended to other statistical areas and implemented in a harmonised and standardised way across countries, as it is recommended for EU Member States by the CMFB. This will contribute significantly to continued quality awareness among users and producers.

5. References

Barklem, Adèle, March 2000, *Revisions analysis of initial estimates of key economics indicators and GDP components*, UK Office of National Statistics, *Economic Trends* No. 556.

Carson, Carol S. and Laliberté, Lucie, February 2002, *Assessing accuracy and reliability: a note based on approaches used in national account and balance of payments statistics*, International Monetary Fund working paper – Statistics Department.

Damia, V., Israël, J. and Picón-Aguilar, C., May 2004, *Positive effects of measuring revisions in balance of payments statistics*, European conference of quality and methodology in official statistics, Mainz, Germany, 24–26 May 2004.

Data Quality Assessment Framework for balance of payments statistics, International Monetary Fund – Statistics Department, July 2003.

Diz Dias, J., April 2004, *Measuring revisions of statistical outputs*, Joint UNECE/Eurostat/OECD Meeting on National Accounts, Geneva, 28–30 April 2004.

European Union BOP and IIP statistical methods, November 2003, European Central Bank.

Fixler, Dennis J. and Bruce T. Grimm, January 2002, *Reliability of GDP and related NIPA estimates*, US Bureau of Economic Analysis.

Grimm, Bruce T. and Robert P. Parker, December 1998, *Reliability of the quarterly and annual estimates of GDP and gross domestic income*, US Bureau of Economic Analysis.

Harmonisation of revision practices for the euro area/EU BOP and IIP statistics, European Central Bank, CMFB May 2004.

Zarb, J., *Quality measures for systems of economic accounts*, Analytical Services Branch, Methodology Division, Australian Bureau of Statistics.

Joint ECB-Eurostat Task Force on Quality – Report on the quality assessment of the balance of payments and international investment position, CMFB (Website http://www.cmfb.org/).

Joint ECB-Eurostat Task Force on Quality – The output quality assessment of quarterly national accounts, CMFB.

McLennan, W., February 1996, *Quality of Australian balance of payments statistics*, Australian Bureau of Statistics. (This report was presented at the BOPCOM meeting, Canberra 21–25 October 1996 under the title "Revisions in Australia's balance of payments (BOP) statistics".)

Nesbit, Shirley, October 1996, *Revisions in the New Zealand balance of payments*, Statistics New Zealand, Paper for the BOPCOM meeting, Canberra, 21–25 October 1996.

Stewart, T. R. and Lusk, C. M. *Seven components of judgmental forecasting skill: Implications for research and the improvement of forecasts*, Journal of Forecasting, No. 13, 1994, pp.579–599.

Symons, Peter, March 2001, *Revisions analysis of initial estimates of annual constant price GDP and its components*, UK Office of National Statistics, *Economic Trends* No. 568.

Trade-off between timeliness and accuracy: ECB requirements for general economic statistics, European Central Bank Monthly Bulletin, April 2001.

Verbeek, Jos, December 1998, *The World Bank's unified survey projections: How accurate are they? An ex-post evaluation of US91-US97*, World Bank.

Young, Allan H., August 1987, *Evaluation of the GNP estimates, US Bureau of Economic Analysis*. (Carol S. Carson, Frank de Leeuw and Robert P. Parker also contributed to this article.)

Young, Allan H., October 1993, *Reliability and accuracy of the quarterly estimates of GDP*, US Bureau of Economic Analysis.

Abstract

The paper discusses different quantitative indicators and qualitative statements that allow for an operational assessment of the output quality of euro area quarterly national accounts main aggregates and BOP statistics. Proposals to communicate these quality measures in press releases and comprehensive quality reports are presented. These proposals will be implemented by end 2004 or early 2005.

Jorge Diz Dias and Wim Haine
European Central Bank
Directorate General Statistics
Kaiserstrasse 29, D-60311 Frankfurt am Main
Germany
jorge.diz_dias@ecb.int
wim.haine@ecb.int

Ingo Kuhnert
European Commission
Eurostat
Bât. Jean Monnet, Rue Alcide de Gasperi
L-2920 Luxembourg
Ingo.Kuhnert@cec.eu.int

August 2004