# Heterogeneity and Aggregate Fluctuations

Minsu Chang*
*Georgetown University*

Xiaohong Chen
*Yale University*,
*Cowles Foundation*

Frank Schorfheide
*University of Pennsylvania*,
*CEPR, PIER, NBER*

This Version: October 29, 2019

**Abstract**

We develop a state-space model with a state-transition equation that takes the form of a functional vector autoregression and stacks macroeconomic aggregates and a cross-sectional density. The measurement equation captures the error in estimating log densities from repeated cross-sectional samples. The log densities and the transition kernels in the law of motion of the states are approximated by sieves, which leads to a finite-dimensional representation in terms of macroeconomic aggregates and sieve coefficents. We use this model to study the joint dynamics of technology shocks, per capita GDP, employment rates, and the earnings distribution. We find that the estimated spillovers between aggregate and distributional dynamics are generally small, a positive technology shocks tends to increase the fraction of individuals earning less than the labor share of per capita GDP, and shocks that raise the inequality of earnings have ambiguous effects on per-capita GDP. (JEL C11, C32, C52, E32)

*Key words:* Econometric Model Evaluation, Earnings Distribution, Functional Vector Autoregressions, Heterogeneous Agent Models, Technology Shocks

* Correspondence: M. Chang: Department of Economics, Georgetown University, Washington, D.C., 20057-1036. Email: minsu.chang@georgetown.edu (Chang). F. Schorfheide: Department of Economics, University of Pennsylvania, Philadelphia, PA 19104-6297. Email: schorf@ssc.upenn.edu (Schorfheide). X. Chen: Department of Economics and Cowles Foundation for Research in Economics, Yale University, New Haven, CT 06520-8281. Email: xiaohong.chen@yale.edu. We thank Yongsung Chang and Dirk Krüger as well as seminar participants at various seminars and conferences for helpful suggestions. Schorfheide gratefully acknowledges financial support from the National Science Foundation under Grant 1424843.

# 1 Introduction

Models with heterogeneity on the household side or the firm side have long been used to study distributional effects of macroeconomic policies. Heterogeneity evolves dynamically and in some of these models interacts closely with aggregate fluctuations. This is particularly true in models with financing constraints that try to capture the large downturn during the recent Great Recession. While the macroeconomics literature has demonstrated that dynamics in heterogeneous agent (HA) models can be different from their representative agent (RA) counterparts, it is an open question whether in the data there is strong evidence that the dynamics of macroeconomic aggregates interacts, at business cycle frequencies, with the evolution of the cross-sectional distribution of income and wealth on the household side and the distribution of productivity and capital on the firm side.

The goal of this paper is to develop and apply econometric tools that can provide semi-structural evidence about the interaction of aggregate and distributional dynamics. More specifically, we specify a state-space model with a state-transition equation that takes the form of a functional vector autoregression (VAR) and stacks macroeconomic aggregates and cross-sectional distributions. We motivate the specification of our functional state-space model by a linearization of a reduced form model in which dynamics of aggregate variables and a function of the lagged cross-sectional distribution of individual-level decisions or states, and the units (households or firms) base their decisions on lagged macroeconomic aggregates and lagged cross-sectional distributions. To make the functional analysis tractable, we approximate the log-densities of the cross-sectional distributions as well as the transition kernels in the functional autoregressive law of motion of the states by sieves. This leads to a finite-dimensional approximation that is expressed in terms of the aggregate variables as well as the coefficients of the sieve approximations.

The functional analysis is implemented with spline basis functions. Log-splines are a popular tool in the statistics literature to approximate densities non-parametrically. In a first step, we estimate the coefficients of the log-spline density approximation for each time period based on a finite sample of cross-sectional observations. We treat these coefficients as noisy measures of a finite-dimensional population approximation of the cross-sectional densities. The measurement errors capture the estimation uncertainty associated with the log-spline coefficients. We then construct an approximate state-space model that stacks the latent population spline coefficients and the macroeconomic variables in a vector of state variables that evolves according to a linear vector autoregressive law of motion. The coefficients of this

state-space model are estimated using Bayesian techniques and the subsequent substantive analysis is based on the estimated state-space model.

This paper makes several contributions. First, we develop a Bayesian implementation of the estimation procedure. Our procedure is able to account for estimation errors in the cross-sectional densities. We also allow for seasonal adjustments of the cross-sectional densities if they are combined with seasonally-adjusted aggregate data. Prior distributions are used to regularize a potentially high-dimensional estimation problem. Second, we make an empirical contribution, by documenting the effect of technology shocks on the cross-sectional distribution of income as well as the effect of distributional shocks that raise inequality on macroeconomic aggregates. Third, an expanded future version of this paper will provide a large sample theory for our estimation procedure. We will provide convergence rates for estimators under the assumption that the number of draws from period $t$ cross-sectional distributions, $N$, the number of time periods, $T$, and the number of terms in the sieve approximations of the cross-sectional densities, $K$, tend to infinity.

Using mostly calibrated HA models, the quantitative macro literature has been examining three types of research questions related to distributional dynamics. First, does micro-level heterogeneity affect the propagation of aggregate shocks to aggregate variables? Second, what is the effect of an aggregate shock on cross-sectional distributions? Third, what is the effect of a change in the cross-sectional distribution on macroeconomic aggregates?

A seminal paper assessing the first question is Krusell and Smith (1998), henceforth KS. The paper combines a neoclassical stochastic growth model with a heterogeneous agent economy in which households face uninsurable idiosyncratic income risk. The equilibrium in the KS model has exactly the features described above: households' decisions depend on the aggregate technology shock as well as the cross-sectional distribution of skills and wealth. In turn, the entirety of the household-level decisions determine the cross-sectional distribution. One of the key findings was that in the stationary stochastic equilibrium, the behavior of the macroeconomic aggregates can be almost perfectly described using only the mean of the wealth distribution.

In extensions of the benchmark model, this approximate aggregation result is no longer true and these models exhibit a richer interaction between aggregate and distributional dynamics. For instance, Chang, Kim, and Schorfheide (2013) consider a HA model with indivisible labor supply. According to their findings, it is important to include a labor supply shock in the RA model to approximate the dynamics of the HA economy well. This aggregate

labor supply shock essentially captures time-varying aggregation errors. Their simulation results imply that modeling the dynamics of cross-sections is not of first-order importance if the analysis focuses on macroeconomic aggregates. However, there is an important caveat: preference and technology parameter estimates of the RA model are not invariant to policy changes and the bias in the RA model's policy predictions is large compared to predictive intervals that reflect parameter uncertainty.

A popular tool to study whether micro-level heterogeneity affects the propagation of aggregate shocks at the macro level is a comparison of impulse responses from a models with and without micro-level heterogeneity that share an otherwise identication structure. Recent examples of this work include Ottonello and Winberry (2018) as well as Ahn, Kaplan, Moll, Winberry, and Wolf (2018) and Kaplan and Violante (2018). The second research question, that is, the effect of aggregate shocks on cross-sectional distributions, has been analyzed in a number of recent papers. For instance, the abovementioned paper by Ahn, Kaplan, Moll, Winberry, and Wolf (2018) also studies the effect of factor-specific productivity shocks on inequality dynamics. Coibion, Gorodnichenko, Kueng, and Silvia (2017) and Kaplan and Violante (2018) examine the distributional effects of monetary policy shocks. Mongey and Williams (2017) analyze the effect of macro shocks on the dispersion of sales growth. Finally, an example of research examinig the third question is the paper by Auclert and Rognlie (2018) which studies the effect of an exogenous rise in equality on macroeconomic dynamics.

In our functional state-space framework, after applying suitable identification schemes, we are able to construct impulse responses of cross-sectional distributions to aggregate shocks as well as responses of aggregate variables to shocks that primarily move the cross-sectional distribution. Moreover, we can assess the significance of the coefficients that capture spill-overs from lagged distributional coefficients to current macroeconomic variables and, vice versa, from lagged macroeconomic variables to the current cross-sectional distribution. If the VAR coefficients in the state-transition equation are block-diagonal, then there is no benefit from modeling the the cross-sectional heterogeneity if the goal is to understand aggregate dynamics. Thus, the methods developed in this paper allow researchers to examine the three abovementioned question in a semi-structural framework that does not require the solution and calibration or estimation of a heterogeneous agent model.

There is an extensive literature on the statistical analysis of functional data. General treatments are provided in the books by Bosq (2000), Ramsey and Silverman (2005), and Horvath and Kokoszka (2012). Much of the literature assumes that the functions are observed

without error. Examples of such functions are the reading of a thermometer over the span of twenty-four hours or the price of a stock from opening to closing time of a stock exchange. Each day delivers a new observation of the curve. In our application, the functions are a log probability densities, which are not observed, but can be estimated from cross-sectional observations. A fundamental model in the functional time series literature is the functional autoregressive model of order one. Bosq (2000) provides a detailed analysis of this model as well as more general functional linear processes. Functional autoregressive models can be estimated by functional principal component analysis, which approximates the functions by linear combinations of the eigenfunctions of the sample covariance operator associated with the $K$ largest eigenvalues. Rather than using what the literature considers to be an optimal (in a least squares sense) *empirical* orthonormal basis, we use a spline basis that is chosen independently of the data.

Applications of functional data analysis in macroeconometrics are rare. The state-space model Diebold and Li (2006) could be interpreted as a functional model for yield-curve data, but there is no infinite-dimensional aspect to the analysis in the sense that it is assumed (and empirically justified) that the yield-curves can be represented by three time-varying parameters. Inoue and Rossi (2018) estimate what they call a VAR with functional shocks, which uses a similar representation of the yield curve as in Diebold and Li (2006) and focuses on the identification of a functional monetary policy. The paper imposes restrictions of the evolution of the yield curves and abstracts from the nonparametric aspects of functional modeling. Hu and Park (2017) develop an estimation theory for a functional autoregressive model with unit roots and fit it to yield curve data and Chang, Kim, and Park (2016) use a functional time series process to capture the evolution of earnings densities with a focus on unit-root components. Both papers use functional principal components analysis.

Both our functional VAR analysis as well as the solution of models with heterogeneous agents requires a parsimonious representation of cross-sectional distributions. For instance, Krusell and Smith (1998) represent the wealth distribution by its mean. Reiter (2010) uses a discretization of the wealth distribution and a Markov transition matrix to characterize movements in the wealth distribution. The transition probabilities are functions of the aggregate states. Algan, Allais, and Den Haan (2008) and Winberry (2017) use moments to characterize the distribution of productivity and capital. From these moments, it is then possible to recover a density with a class of distributions that belong to the exponential family.[1] Childers (2015) considers linearization methods for models with function-valued states,

[1]This approach is somewhat related to the density estimator discussed in Efron and Tibshirani (1996).

using a wavelet representation. These approximations can be interpreted as sieves with different types of basis functions. Our analysis uses spline basis functions to approximate log densities, which dates back to Kooperberg and Stone (1990).

The remainder of this paper is organized as follows. In Section 2 we outline a functional state-space model for a group of macroeconomic time series and a sequence of cross-sectional distributions that are estimated based on cross-sectional data. We estimate the functional state-space model using Bayesian techniques. Implementation details such as the choice of basis functions, the choice of prior distributions, and the posterior sampler are discussed in Section 3. The empirical application is presented in Section 4 and Section 5 concludes. Supplemental derivations and information about the empirical analysis are relegated to the Online Appendix.

# 2 A Functional State Space Model – Heuristics

VARs can be viewed as approximations to the equilibrium dynamics arising from linearized RA models and have proved to be useful for the evaluation and development of dynamic stochastic general equilibrium (DSGE) models. Moreover, VARs are widely used in empirical macroeconomics independently of DSGE models, to study business cycle fluctations, the propagation of shocks, and to generate macroeconomic foreasts. In this paper, we develop a functional vector autoregression (fVAR) that is embedded in a state space model and can play a similar role as traditional VARs in environments in which macroeconomic aggregates interact with cross-sectional distributions. While we not will establish a formal link between our functional model and the solution of HA models, it does capture the salient features and provides a natural reference model for the evaluation of HA models. Moreover, just as VARs, our functional model can be used as a stand-alone tool for empirical work in macroeconomics.

In this section we provide an informal presentation of the functional model. A formal treatment is relegated to Section D. The variables in the model comprise an $n_z \times 1$ vector of macroeconomic aggregates $Z_t$ and a cross-sectional density $p_t(x)$. In our application $Z_t$ consists of (log) total factor productivity growth, per-capita GDP growth, and the log employment rate. The cross-sectional variable $x$ is earnings as a fraction of per-capita GDP. Throughout this paper, we will work with log densities defined as $\ell_t(x) = \ln p_t(x)$. We decompose $Z_t$ and $\ell_t$ into a deterministic component $\big(Z_*, \ell_*(x)\big)$ and fluctuations around the

deterministic component. Let

$$Z_t = Z_* + \tilde{Z}_t, \quad \ell_t = \ell_* + \tilde{\ell}_t. \tag{1}$$

For notational convenience we assumed that the deterministic component is time-invariant and could be interpreted as a steady state. This assumption could be easily relaxed by letting $(Z_*, \ell_*)$ depend on $t$. We assume that the deviations from the deterministic component $\big(\tilde{Z}_t, \tilde{\ell}_t(x)\big)$ evolve jointly according to the following linear functional VAR law of motion:

$$\begin{aligned} \tilde{Z}_t &= B_{zz}\tilde{Z}_{t-1} + \int B_{zl}(\tilde{x})\tilde{\ell}_{t-1}(\tilde{x})d\tilde{x} + u_{z,t} \\ \tilde{\ell}_t(x) &= B_{lz}(x)\tilde{Z}_{t-1} + \int B_{ll}(x,\tilde{x})\tilde{\ell}_{t-1}(\tilde{x})d\tilde{x} + u_{l,t}(x). \end{aligned} \tag{2}$$

Here $u_{z,t}$ is mean-zero random vector with covariance $\Sigma_{zz}$ and $u_{l,t}(x)$ is a random element in $L_2$ with covariance function $\Sigma_{ll}(x,\tilde{x})$. We denote the covariance function for $u_{z,t}$ and $u_{l,t}(x)$ by $\Sigma_{zl}(x)$.

To condense the notation, we define integral operators with kernels $B_{zl}(\tilde{x})$ and $B_{ll}(x,\tilde{x})$ as

$$\boldsymbol{B}_{zl}[g] = \int B_{zl}(\tilde{x})g(\tilde{x})d\tilde{x}, \quad \boldsymbol{B}_{ll}[g](x) = \int B_{ll}(x,\tilde{x})g(\tilde{x})d\tilde{x}.$$

Using the operator notation, we can write (2) more compactly as

$$\begin{aligned} \tilde{Z}_t &= B_{zz}\tilde{Z}_{t-1} + \boldsymbol{B}_{zl}[\tilde{\ell}_{t-1}] + u_{z,t} \\ \tilde{\ell}_t(x) &= B_{lz}(x)\tilde{Z}_{t-1} + \boldsymbol{B}_{ll}[\tilde{\ell}_{t-1}](x) + u_{l,t}(x). \end{aligned} \tag{3}$$

For now, (3) should be interpreted as a reduced-form fVAR in which $u_{z,t}$ and $u_{l,t}(x)$ are one-step-ahead forecast errors. In principle we could add further lags, but our empirical analysis will be based on a single lag. The system will subsequently serve as the state-transition equation in a functional state-space model.

## 2.1 Sampling and Measurement Equation

We assume that in every period $t = 1, \ldots, T$ the econometrician observes a measurement of the macroeconomic aggregates $Z_t^o$ (here the $o$ superscript indicates "observed") as well as a sample of $N_t$ draws $x_{it}^o$, $i = 1, \ldots, N_t$ from the cross-sectional density $p_t(x)$. In practice, $N_t$ is likely to vary from period to period, but for the subsequent exposition it will be more convenient to assume that $N_t = N$ for all $t$. We also assume that the draws $x_{it}^o$ are

independently and identically distributed (iid) over the cross-section as well as over time. The measurement equations take the form

$$Z_t^o = Z_* + \tilde{Z}_t \tag{4}$$

$$x_{it}^o \sim iid\ p_t(x) = \frac{\exp\{\ell_*(x) + \tilde{\ell}_t(x)\}}{\int \exp\{\ell_*(x) + \tilde{\ell}_t(x)\}dx}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T. \tag{5}$$

Here we assumed that the macroeconomic aggregates are observed without errors and have been transformed to have time-invariant mean/steady state $Z_*$. The assumption of $x_{it}$ being $iid$ across $i$ and $t$ is consistent with data sets that comprise repeated cross sections.[2] It is also approximately consistent with panel data sets if the unit index $i$ is randomly re-assigned in every period $t$. Thus, to the extent that the cross-sectional densities $p_t(x)$ are estimated from a panel data set, there is some potential loss of information in our functional modeling approach. However, on the positive side, the functional modeling approach does not require the econometrician to make assumptions about the evolution of $x_{it}$ at the level of an individual, a household, or a firm.

## 2.2 Sieve Approximations

Equations (3), (4), and (5) define a state-space model for the observables $\{Z_t^o, x_{1,t}^o, \ldots, x_{2,t}^o\}_{t=1}^T$. The state variables are $(\tilde{Z}_t, \tilde{l}_t)$. To implement the estimation of the functional model we approximate infinite-dimensional elements by sieves. We will formalize the nature of the approximation in Section D. Let

$$\tilde{\ell}_t(x) \approx \tilde{\ell}_t^{(K)}(x) = \sum_{k=0}^{K} \tilde{\alpha}_{k,t}\zeta_k(x) = \big[\zeta_0(x), \zeta_1(x), \ldots, \zeta_K(x)\big] \cdot \begin{bmatrix} \tilde{\alpha}_{0,t} \\ \vdots \\ \tilde{\alpha}_{K,t} \end{bmatrix} = \zeta_t(x)\tilde{\alpha}_t \tag{6}$$

and

$$\ell_*(x) \approx \ell_*^{(K)}(x) = \zeta(x)\alpha_*.$$

Here $\zeta_0(x), \zeta_1(x), \ldots$ is a sequence of basis functions. We let $\alpha_t = \alpha_* + \tilde{\alpha}_t$ such that $\ell^{(K)}(x) = \ell_*^{(K)}(x) + \tilde{\ell}_t(x)$. Moreover, we adopt the convention that $\zeta_0(x) = 1$ such that $\alpha_{0,t}$ becomes the normalization constant for the density $p_t(x)$.

[2] If the data exhibit spatial correlation, then our estimation approach below essentially replaces the likelihood function for $x_{1t}^o, \ldots, x_{Nt}^o$ by a composite likelihood function that ignores the spatial correlation; see Varin, Reid, and Firth (2011).

To approximate the measurement equation of the cross-sectional observations in (5), we let $X_t^o = [x_{1t}^o, \ldots, x_{Nt}^o]$ and define the $K$-dimensional vector of sufficient statistics

$$\bar{\zeta}(X_t^o) = \sum_{i=1}^{N} \zeta(x_{it}^o).$$

This allows us to write a $K$'th order approximation to the density of $X_t^o$:

$$\begin{aligned} p^{(K)}\left(X_t^o|\alpha_t^o\right) &= \exp\left\{\bar{\zeta}(X_t^o)\alpha_t^o\right\} \\ \text{where} \qquad & \alpha_{0,t}^o = -\ln \int \exp\left\{\sum_{k=1}^{K} \alpha_{k,t}^o \zeta_k(x)\right\} dx \\ & \alpha_{k,t}^o = \alpha_{k,*} + \tilde{\alpha}_{k,t}, \quad k = 1, \ldots, K. \end{aligned} \tag{7}$$

We introduced the auxiliary coefficients $\alpha_{k,t}^o$ to capture the re-normalization, which is determined outside of the state-transition equation (see below) that will govern the evolution of $\alpha_t$.

For the state-transition equation (3) we approximate the kernels $B_{zl}(\tilde{x})$ and $B_{ll}(x, \tilde{x})$, the function $B_{lz}(x)$, and the functional innovation $u_{l,t}(x)$ as follows:

$$\begin{aligned} B_{zl}(\tilde{x}) &\approx B_{zl}^{(K)}(x) = \sum_{j=0}^{J} B_{zl,\cdot j}\xi_j(\tilde{x}) = B_{zl}\xi'(\tilde{x}) \\ B_{ll}(x, \tilde{x}) &\approx B_{ll}^{(K)}(x, \tilde{x}) = \sum_{k=0}^{K}\sum_{j=0}^{J} B_{ll,kj}\zeta_k(x)\xi_j(\tilde{x}) = \zeta(x)B_{ll}\xi'(\tilde{x}) \\ B_{lz}(x) &\approx B_{lz}^{(K)}(x) = \sum_{k=0}^{K} B_{lz,k}\zeta_k(x) = \zeta(x)B_{lz} \\ u_{l,t}(x) &\approx u_{l,t}^{(K)}(x) = \sum_{k=0}^{K} u_{\alpha k,t}\zeta_k(x)p_*(x) = \zeta'(x)u_{a,t}. \end{aligned} \tag{8}$$

Here $\xi_0(x), \xi_1(x), \ldots$ is a second sequence of basis functions and, for a given $x$, $\xi'(x) = [\xi_0(x), \xi_1(x), \ldots, \xi_J(x)]$ is a $1 \times (J+1)$ vector. The matrix $B_{zl}$ is of dimension $n_z \times (J+1)$, $B_{ll}$ is of dimension $(K+1) \times (J+1)$ and $B_{lz}$ is of dimension $(K+1) \times n_z$. Let $\boldsymbol{B}_{zl}^{(K)}[\cdot]$ and $\boldsymbol{B}_{ll}^{(K)}[\cdot](x)$ be the operators associated with the kernels $B_{zl}^{(K)}(x)$ and $B_{ll}^{(K)}(x, \tilde{x})$. An approximation of (3) is given by

$$\begin{aligned} \tilde{Z}_t &= B_{zz}\tilde{Z}_{t-1} + \boldsymbol{B}_{zl}^{(K)}[\tilde{\ell}_{t-1}^{(K)}] + u_{z,t} \\ \tilde{\ell}_t^{(K)}(x) &= B_{lz}^{(K)}(x)\tilde{Z}_{t-1} + \boldsymbol{B}_{ll}^{(K)}[\tilde{\ell}_{t-1}^{(K)}](x) + u_{l,t}^{(K)}(x). \end{aligned} \tag{9}$$

Combining (6), (8), and (9) yields the following vector autoregressive system for the macroeconomic aggregates and the sieve coefficients:

$$\begin{bmatrix} \tilde{Z}_t \\ \tilde{\alpha}_t \end{bmatrix} = \begin{bmatrix} B_{zz} & B_{zl}C_\alpha \\ B_{lz} & B_{ll}C_\alpha \end{bmatrix} \begin{bmatrix} \tilde{Z}_{t-1} \\ \tilde{\alpha}_{t-1} \end{bmatrix} + \begin{bmatrix} u_{z,t} \\ u_{\alpha,t} \end{bmatrix}, \tag{10}$$

where

$$C_\alpha = \int \xi'(\tilde{x})\zeta(\tilde{x})d\tilde{x}.$$

Overall, we obtain a state-space representation that comprises the measurement equations (4) and (7) and the state-transition equation (10). Due to the measurement equation for the $x_{it}$s, the state-space representation is nonlinear and the computation of the exact likelihood requires a nonlinear filter. To avoid the use of a nonlinear filter in the empirical application, we consider the following simplification. Using (7), define the log likelihood function and the maximum-likelihood estimator $\hat{\alpha}_t^o$ as

$$\mathcal{L}(\alpha_t|X_t^o) = \ln p^{(K)}(X_t^o|\alpha_t^o), \quad \hat{\alpha}_t^o = \text{argmax}_{\alpha_t^o}\, \mathcal{L}(\alpha_t|X_t^o). \tag{11}$$

Provided that the number of cross-sectional observations $N$ is large relative to the dimension $K$, the maximum likelihood estimator has an asymptotically normal sampling distribution, that is,

$$\hat{\alpha}_t^o|\alpha_t^o \overset{\text{approx}}{\sim} N\big(\alpha_t^o, V_{\alpha_t^o}\big),$$

where $V_{\alpha_t^o}$ is the inverse of the negative Hessian. To simplify the measurement equation (7), we pursue a limited-information approach and condition inference about $\alpha_t^o$ on $\hat{\alpha}_t^o$:

$$\hat{\alpha}_t^o = \alpha_* + \tilde{\alpha}_t + N^{-1/2}\eta_t, \quad \eta_t \sim N(0, V_{\hat{\alpha}_t^o}). \tag{12}$$

Here, the Hessian is evaluated at $V_{\hat{\alpha}_t^o}$. By replacing (7) by (12) we obtain a linear Gaussian state-space model and the likelihood function can be evaluated with the Kalman filter.

## 2.3 Relationship to Structural HA Models

In HA models, individual choices depend on the cross-sectional distribution of unit-specific variables, e.g., wealth and skill, and on some macroeconomic state variables, e.g., total factor productivity. As a consequence, macroeconomic variables that aggregate choices at the micro

level will also depend on cross-sectional distributions in addition to the macroeconomic state variables. We capture these relationships in the following, very stylized model:

$$\tilde{Z}_t = B_{zz}\tilde{Z}_{t-1} + \int B_{zp} x p_{t-1}(x)dx + \sigma\eta_t, \quad \eta_t \sim N(0,1) \tag{13}$$

$$x_{it} = B_{xx}x_{it-1} + \int B_{xp} x p_{t-1}(x)dx + B_{xz}\tilde{Z}_{t-1} + \epsilon_{it}, \quad \epsilon_{it} \sim N(0,1). \tag{14}$$

Here $\tilde{Z}_t$ is scalar, $p_t(x)$ is the cross-sectional distribution of the $x_{it}$s, and both $\tilde{Z}_t$ and $x_{it}$ only depend on the mean of the $x_{it-1}$ distribution. Because this stylized model is backward-looking, it is straightforward to characterize the law of motion for the cross-sectional density $p_t(x)$ by integrating (14) over $i$:

$$p_t(x) = \int \phi_N\left(x - B_{xx}\tilde{x} - \int B_{xp} x p_{t-1}(x)dx - B_{xz}\tilde{Z}_{t-1}\right) p_{t-1}(\tilde{x})d\tilde{x}, \tag{15}$$

where $\phi_N(\cdot)$ is the probability density function of a $N(0,1)$.

To cast the model into the form (2) the law of motion for $p_t(x)$ has to be linearized. The steady state density $p_*(x)$ is obtained by shutting down aggregate uncertainty and setting $\sigma = 0$, which leads to the functional equation:

$$p_*(x) = \int \phi_N\left(x - B_{xx}\tilde{x}\right) p_*(\tilde{x})d\tilde{x}. \tag{16}$$

The solution is given by the density of a $N\big(0, 1/(1-B_{xx})\big)$ random variable. Let $\ell_t(x) = \ln p_t(x)$. Linearization of (15) with respect to $\tilde{\ell}_t$ and $\tilde{Z}_t$ around $\ell_*(x) = \ln p_*(x)$ and 0 yields

$$\begin{aligned}
\tilde{\ell}_t(x) &= -\frac{1}{p_*(x)} B_{xz}\left[\int \phi_N^{(1)}\left(x - B_{xx}\tilde{x}\right) p_*(\tilde{x})d\tilde{x}\right]\tilde{Z}_{t-1} \\
&\quad -\frac{1}{p_*(x)}\left[\int \phi_N^{(1)}\left(x - B_{xx}\tilde{x}\right) p_*(\tilde{x})d\tilde{x}\right]\int B_{xp}\tilde{x} p_*(\tilde{x})\tilde{\ell}_{t-1}(\tilde{x})d\tilde{x} \\
&\quad +\frac{1}{p_*(x)}\int \phi_N\left(x - B_{xx}\tilde{x}\right) p_*(\tilde{x})\tilde{\ell}_{t-1}(\tilde{x})d\tilde{x} \\
&= B_{lz}\tilde{Z}_{t-1} + \boldsymbol{B}_{ll}[\tilde{\ell}_{t-1}].
\end{aligned} \tag{17}$$

Thus, in this stylized example the fVAR state-transition equation in (3) is obtained as a linear approximation.

# 3 Implementation Details

We now provide some of the implementation details. The choice of basis functions is described in Section 3.1 and some preliminary transformations of the estimated basis function

coefficients is discussed in Section 3.2. Section 3.3 provides details on the specification of the measurement equation for the basis function coefficients and the state-transition equation of the empirical state-space model. Piors and the computation of posteriors for the parameters of the state-space model are discussed in Section 3.4 and Section 3.5 explains how forecasts and impulse response functions for the basis function coefficients can be converted back into cross-sectional densities.

## 3.1 Basis Functions

A convenient basis for the log density is a spline of degree $m = 3$. This connects the analysis to log-spline density estimation; see Kooperberg and Stone (1990). A spline is a piecewise polynomial functions with knots $x_s$, $s = 1, \ldots, S$:

$$\begin{aligned}\mathrm{Spl}(m,S) &= \sum_{k=0}^{m} a_k \left(x^k \mathbb{I}\{x \leq x_S\} + x_S^k \mathbb{I}\{x > x_S\}\right) \\ &\quad + \sum_{s=1}^{S-1} b_s \left(\left[\max\{x - x_s,\, 0\}\right]^m \mathbb{I}\{x \leq x_S\} + (x_S - x_s)^m \mathbb{I}\{x > x_S\}\right) \\ &\quad + \sum_{k=1}^{m} c_k \left[\max\{x - x_S,\, 0\}\right]^m.\end{aligned}$$

Kooperberg and Stone (1990) suggest to make the function linear and upward sloping on the interval $(-\infty, x_1)$ and linear and downward sloping on the interval $[x_S, \infty]$. Thus, for $m = 3$ this would lead to $a_2 = a_1 = 0$ and $c_2 = c_3 = 0$. This leads to tails of a Laplace density, which are a bit thicker than Gaussian tails. Translating this specification into our $\zeta_j(x)$, $j = 0, \ldots, K = S + 1$ notation, we obtain:

$$\begin{aligned}\zeta_0(x) &= 1 \\ \zeta_1(x) &= (x\mathbb{I}\{x \leq x_S\} + x_S \mathbb{I}\{x > x_S\}) \\ \zeta_2(x) &= \left(\left[\max\{x - x_1,\, 0\}\right]^3 \mathbb{I}\{x \leq x_S\} + (x_S - x_1)^3 \mathbb{I}\{x > x_S\}\right) \\ &\vdots \\ \zeta_S(x) &= \left(\left[\max\{x - x_{S-1},\, 0\}\right]^3 \mathbb{I}\{x \leq x_S\} + (x_S - x_{S-1})^3 \mathbb{I}\{x > x_S\}\right) \\ \zeta_{S+1}(x) &= \max\{x - x_S,\, 0\}.\end{aligned} \tag{18}$$

## 3.2 Sieve Coefficients

The first step in the estimation of the functional model is the construction of the sequence of estimated log-spline coefficients. For each period $t$, we estimate a log-spline density based on the cross-sectional data $x^o_{1t}, \ldots, x^o_{Nt}$; see (11). We denote the resulting maximum likelihood estimates by $\hat{\alpha}^o_{k,t}$, $k = 0, \ldots, K$. We drop the sequence of coefficients, $\alpha_{0,t}$, that simply normalize the densities.

**Seasonal Adjustment.** In our application $x^o_{it}$ is based on quarterly earnings data from the Current Population Survey (CPS). Unlike the macroeconomic variables stacked in $Z^o_t$, the quarterly earnings data are not seasonally adjusted. To seasonally adjust the cross-sectional densities, we estimate two regressions for $k = 1, \ldots, K$ by OLS:

$$\text{Constant Mean} \quad : \quad \hat{\alpha}^o_{k,t} = \bar{\alpha}_k + \text{residual} \tag{19}$$

$$\text{Quarterly Dummies} \quad : \quad \hat{\alpha}^o_{k,t} = \sum_{q=1}^{4} s_q(t)\bar{\alpha}_{q,k} + \text{residual}, \tag{20}$$

where $s_q(t) = 1$ if period $t$ is associated with quarter $q$ and $s_q(t) = 0$ otherwise. For each $k$ we use the Schwarz information criterion (BIC) to choose between the two specifications and then define $\hat{\alpha}^r_{k,t}$ as the residual from the selected regression.

**Compression.** In our experience, the $\hat{\alpha}^r_{k,t}$ series exhibit collinearity. We remove perfect collinearities by expressing the coefficient series as linear combinations of a $\tilde{K} \times 1$ vector $\hat{\alpha}^c_t$ using principal component analysis. Here $\tilde{K} \leq K$. Let

$$\hat{\alpha}^r_t = \Lambda' \hat{\alpha}^c_t + \text{residual}, \tag{21}$$

where $\Lambda'$ is a $K \times \tilde{K}$ matrix of loadings. Now consider the following eigenvalue decomposition of the sample covariance matrix of the $\hat{\alpha}^r_t$'s:

$$\hat{\mathbb{V}}[\hat{\alpha}^r_t] = \frac{1}{T}\sum_{t=1}^{T} \hat{\alpha}^r_t \hat{\alpha}^{r'}_t = V' \Xi V, \tag{22}$$

where $V$ is a matrix of eigenvectors and $\Xi$ is a diagonal matrix of eigenvalues. We define the $T \times \tilde{K}$ matrix $\hat{\alpha}^c$ with rows $\hat{\alpha}^{c'}_t$ where $\hat{\alpha}^{c'}_t = \hat{\alpha}^{r'}_t V_{\cdot 1}$ and $V_{\cdot 1}$ collects the eigenvectors associated with non-zero (in practice less than $10^{-10}$) eigenvalues. The loadings can be computed as

$$\hat{\Lambda} = (\hat{\alpha}^{c'} \hat{\alpha}^c)^{-1} \hat{\alpha}^{c'} \hat{\alpha}^r, \tag{23}$$

where $\hat{\alpha}^r$ is the $T \times K$ matrix with rows $\hat{\alpha}^{r'}_t$. With a cut-off value for the eigenvalues of $10^{-10}$ the estimated loadings generate essentially a perfect fit in our application. Thus, we can use $\hat{\Lambda}$ to convert the $\hat{\alpha}^c_t$'s back into $\hat{\alpha}^r_t$'s.

## 3.3 State-Space Representation

Because of the seasonal adjustment procedure, the $\hat{\alpha}^c_{k,t}$ time series have mean zero. Thus, we can simplify (12) to obtain the following measurement equation:

$$\hat{\alpha}^c_t = \tilde{\alpha}^c_t + N^{-1/2}\eta^c_t, \quad \eta^c_t \sim N(0, V_{\hat{\alpha}^c_t}). \tag{24}$$

Note that compared to (12) the dimension of the $\hat{\alpha}_t$ vector has changed because of the compression step. The covariance matrix $V^c_{\hat{\alpha}_t}$ of the "measurement" errors is defined as

$$V_{\hat{\alpha}^c_t} = -\big(\hat{\Lambda}\hat{\Lambda}'\big)^{-1}\hat{\Lambda}\big[\mathcal{H}(\hat{\alpha}^o_t)\big]^{-1}\hat{\Lambda}'\big(\hat{\Lambda}\hat{\Lambda}'\big)^{-1}, \tag{25}$$

where $\hat{\Lambda}$ is the estimated loading matrix in (23) and $\mathcal{H}(\hat{\alpha}^o_t)$ is the Hessian associated with the original log likelihood function $\mathcal{L}(\alpha_t|X^o_t)$ defined in (11). The Hessian depends on the observations $X^o_t$ only through $\hat{\alpha}^o_t$.

The state transition is essentially given by (10) but we need to adjust it for the compression of the $\alpha_t$ vector. Moreover, we now absorb the matrix $C_\alpha$ into the matrices of regression coefficients:

$$\begin{bmatrix} \tilde{Z}_t \\ \tilde{\alpha}^c_t \end{bmatrix} = \begin{bmatrix} B_{zz} & B^c_{z\alpha} \\ B^c_{\alpha z} & B^c_{\alpha\alpha} \end{bmatrix} \begin{bmatrix} \tilde{Z}_{t-1} \\ \tilde{\alpha}^c_{t-1} \end{bmatrix} + \begin{bmatrix} u_{z,t} \\ u_{\alpha^c,t} \end{bmatrix}. \tag{26}$$

We assume that the innovations are normally distributed and write the state transition more compactly as

$$w_t = \Phi_1 w_{t-1} + u_t, \quad u_t \sim N(0, \Sigma), \tag{27}$$

where $w_t = [\tilde{Z}_t, \tilde{\alpha}^c_t]$.

## 3.4 Priors and Posteriors

Prior to the estimation of the state-space model, we remove the deterministic components from $Z^o_t$ and $\hat{\alpha}^o_t$. Thus, the unknown coefficients are concentrated in the state-transition equation (27), which takes the form of a multivariate linear Gaussian regression model. The state transition can be expressed in matrix form as

$$W = \Phi X + U,$$

where $W$, $X$, and $U$ have rows $w'_t$, $x'_t$, and $u'_t$, respectively and $\Phi = \Phi'_1$. Defining $\phi = \text{vec}(\Phi)$ we use a prior distribution of the form

$$\Sigma \sim IW(\underline{\nu}, \underline{S}), \quad \phi|\lambda \sim N\big(\underline{\mu}_\phi, \underline{P}^{-1}_\phi(\lambda)\big), \tag{28}$$

where $IW(\cdot)$ is the Inverse-Wishart distribution with $\underline{\nu}$ degrees of freedom and scale matrix $\underline{S}$.

The prior precision matrix $\underline{P}_\phi(\lambda)$ is a function of a vector of hyperparameters $\lambda = [\lambda_1, \lambda_2, \lambda_3]'$ and takes the form

$$\underline{P}_\phi(\lambda) = \lambda_1 \begin{bmatrix} (\underline{\Sigma}^{-1})_{zz} \otimes \begin{bmatrix} I_{n_z} & 0 \\ 0 & \lambda_2 I_{n_\alpha} \end{bmatrix} & (\underline{\Sigma}^{-1})_{z\alpha} \otimes \begin{bmatrix} \sqrt{\lambda_3} I_{n_z} & 0 \\ 0 & \sqrt{\lambda_2} I_{n_\alpha} \end{bmatrix} \\ (\underline{\Sigma}^{-1})_{\alpha z} \otimes \begin{bmatrix} \sqrt{\lambda_3} I_{n_z} & 0 \\ 0 & \sqrt{\lambda_2} I_{n_\alpha} \end{bmatrix} & (\underline{\Sigma}^{-1})_{\alpha\alpha} \otimes \begin{bmatrix} \lambda_3 I_{n_z} & 0 \\ 0 & I_{n_\alpha} \end{bmatrix} \end{bmatrix}. \quad (29)$$

Here the partitions of $\Sigma^{-1}$ conform with the partition $w_t = [\tilde{Z}_t, \tilde{\alpha}_t^c]$. The hyperparameter $\lambda_1$ controls the overall precision of the prior distribution; $\lambda_2$ scales the relative precision of the prior distribution for the coefficients that control the effect of $\tilde{\alpha}_{t-1}^c$ on $\tilde{Z}_t$; likewise, $\lambda_3$ scales the relative precision of the prior distribution for the coefficients that control the effect of $\tilde{Z}_{t-1}$ on $\tilde{\alpha}_t^c$. Unlike the more commonly used matrix-Normal Inverse-Wishart prior that mimicks the Kronecker structure of the likelihood function, the prior in (28) allows us to control the degree of spillovers from distributional dynamics to the aggregate dynamics and vice versa. If the prior mean $\underline{\mu}_\phi$ is zero, then as $\lambda_2, \lambda_3 \longrightarrow \infty$, the posterior distributions of $B_{\alpha z}$ and $B_{z\alpha}$ concentrate around zero, which shuts down spillover effects.

However, The problem with the aforementioned prior is that this prior does not account for the different scales of the $y_t$ elements when determining the hyperparameter. Consider the first equation of a simplified model:

$$Z_{1t} = \phi_{z_1 z_1} Z_{1t-1} + \phi_{z_1 z_2} Z_{2t-1} + u_{1t}.$$

Suppose that the regressors have mean zero and that

$$\phi_i \sim N(0, 1).$$

Then the contribution of the first term conditional on $Z_{t-1}$ is a $N(0, Z_{1t-1}^2)$ whereas the contribution to the second term is a $N(0, Z_{2t-1}^2)$. Thus, a priori, depending on the scale of the two regressors, the relative contribution of the two regressors can be quite different. To avoid this problem, one can use a prior of the form

$$\phi_{z_1 z_1} \sim N\big(0, 1/\sigma_{z_1}^2\big), \quad \phi_{z_1 z_1} \sim N\big(0, 1/\sigma_{z_2}^2\big)$$

In order to implement this idea in the more general specification, we should replace:

$$I_{n_z} \text{ by } \begin{bmatrix} \sigma_{z1}^2 & 0 \\ 0 & \sigma_{z2}^2 \end{bmatrix}, \quad I_{n_\alpha} \text{ by } \begin{bmatrix} \sigma_{\alpha 1}^2 & 0 \\ 0 & \sigma_{\alpha 2}^2 \end{bmatrix}.$$

So the analogous modification is applied to the prior stated in (29).

Conditional on $\lambda$, it is straightforward to sample from the posterior distribution of $(\phi, \Sigma)$ using a Gibbs sampler following the approach in Carter and Kohn (1994) that iterates over the blocks:

$$\phi|(\Sigma, W_{1:T}), \quad \Sigma|(\phi, W_{1:T}) \quad W_{1:T}|(\phi, \Sigma, \tilde{Z}_{1:T}, \hat{\alpha}^c_{1:T}).$$

Here it is important to note that conditional on $W_{1:T}$ the observations $(\tilde{Z}_{1:T}, \hat{\alpha}^c_{1:T})$ do not contain any information about $(\phi, \Sigma)$.

### 3.5 Recovering Cross-Sectional Densities

Based on the estimated state-transition equation (26) we can generate forecasts and impulse response functions for the compressed coefficients $\tilde{\alpha}^c_t$. However, the dynamics of these coefficients in itself is not particularly interesting. Thus, we have to convert them back into densities using the following steps (which can be executed for each draw of $\tilde{\alpha}^c_t$ from the relevant posterior distribution). First, use (24) to turn $\tilde{\alpha}^c_t$ into $\hat{\alpha}^c_t$. If the goal is to generate forecasts of observed cross-sectional densities then the conversion should account for the measurement error $\eta^c_t$. If the goal is to generate impulse response functions for the densities, then we recommend setting $\eta^c_t$ equal to zero. Second, use (21) and (23) to transform $\hat{\alpha}^c_t$ into $\hat{\alpha}^r_t$. Third, use (19) to recover $\hat{\alpha}^o_{k,t}$. If the goal is to compute impulse responses, use as intercept the average of the seasonal dummies $\frac{1}{4}\sum_{q=1}^4 \bar{\alpha}_{q,k}$. Forth, compute the normalization constant $\alpha^o_{0,t}$ as well as the density $p^{(K)}(X^o_t|\alpha^o_t)$ according to the definition in (7).

## 4 Empirical Analysis

The empirical analysis focuses on the joint dynamics on total factor productivity, real per-capita GDP, and employment at the aggregate level and the cross-sectional distribution of earnings. Using our functional state-space model, we examine the following three questions: (i) Do distributional dynamics affect aggregate dynamics? (ii) What is the effect of a technology shock on the cross-sectional distribution of earnings? (iii) What are the dynamic responses to different types of distributional shocks? We begin with a description of the data set and the basic parameter estimates in Section 4.1 and then examine the aforementioned questions in Sections 4.2 to 4.4.

## 4.1 Data and Model Estimation

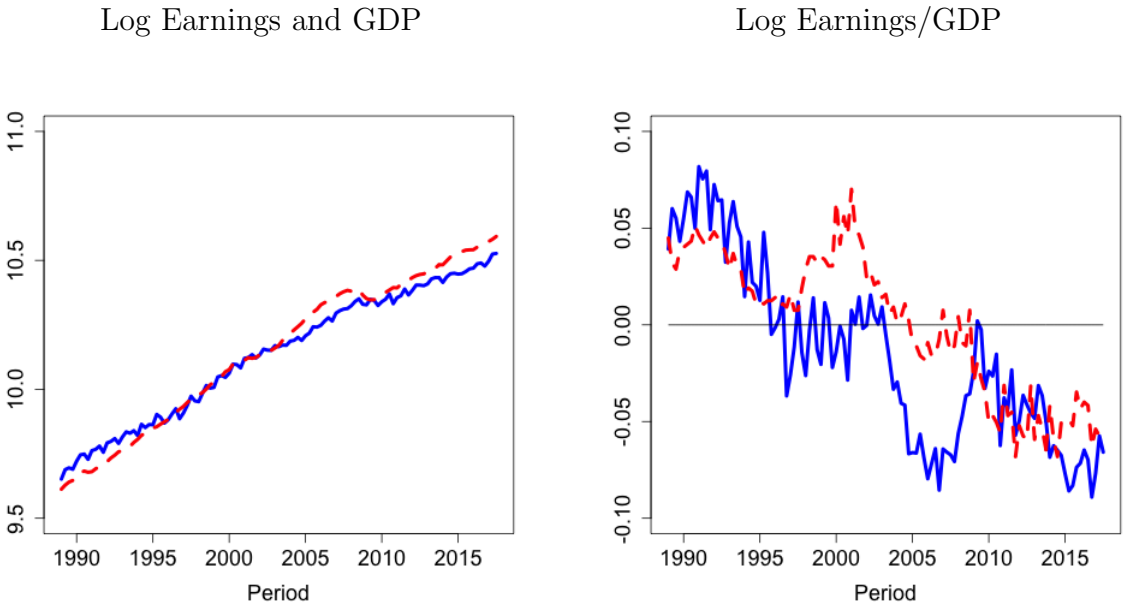
**Data.** We use three macroeconomic aggregates in our empirical analysis: total factor productivity (TFP), real per-capita GDP, and the employment rate. In addition, we use cross-sectional data on earnings. Real per-capita GDP (*A939RX0Q048SBEA*) is provided by the Federal Reserve Bank of St. Louis' FRED database and the TFP series (*dtfp*) is obtained from Fernald (2012). Weekly earnings (*PRERNWA*) are obtained from the monthly Current Population Survey (CPS) through the website of the National Bureau of Economic Research (NBER). Based on the CPS variable *PREXPLF* "Experienced Labor Force Employment" we construct an employment indicator which is one if the individual is employed and zero otherwise. This indicator is used to compute the aggregate employment rate.

We pre-process the cross-sectional data as follows. We drop individuals if (i) the employment indicator is not available; and (ii) if they are coded as "employed" but the weekly earnings are missing. In addition, we re-code individuals with non-zero earnings as employed and set earnings to zero for individuals that are coded as not employed. Weekly earnings are scaled to annual earnings by multiplying with 52. A CPS-based unemployment rate is computed as the fraction of individuals that are coded as not employed. By construction this is one minus the fraction of individuals with non-zero weekly earnings, which is used to normalize the cross-sectional density of earnings. It turns out that the CPS-based unemployment rate tracks the aggregate unemployment rate (*UNRATE* from FRED) very closely. The Online Appendix contains a figure that overlays the two series.

In the left panel of Figure 1 we plot average log nominal earnings computed from the cross-sectional data and log nominal per-capita GDP. We scale per-capita GDP by a factor of 2/3 to account for the labor share.[3] After this re-scaling the mean of log earnings and log per-capita GDP have approximately the same level. However, the mean log earnings grow more slowly than per-capita GDP. In the right panel of the Figure we plot the average log earnings-to-GDP ratio (here per-capita GDP is again scaled by 2/3) and the demeaned log labor share of the nonfarm business sector (obtained from the Bureau of Labor Statistics). The drop in the log earnings-to-GDP ratio is of the same order of magnitude as the fall in the labor share over the sample period.

[3]Nominal per-capita GDP is obtained by multiplying real per-capita GDP by the GDP deflator (*GDPDEF* from FRED). The factor 2/3 is a rule-of-thumb number that happens to align the levels in the left panel. The average labor share of the nonfarm business sector over the sample period is 0.6.

Figure 1: Earnings and GDP




*Notes:* Left panel: average log earnings (blue, solid) and log per capita GDP (red, dashed). Right panel: average log earnings-to-GDP ratio (blue, solid) and demeaned log labor share (red, dashed) of the nonfarm business sector. In both panels per-capita GDP is scaled by 2/3 to account for the labor share.

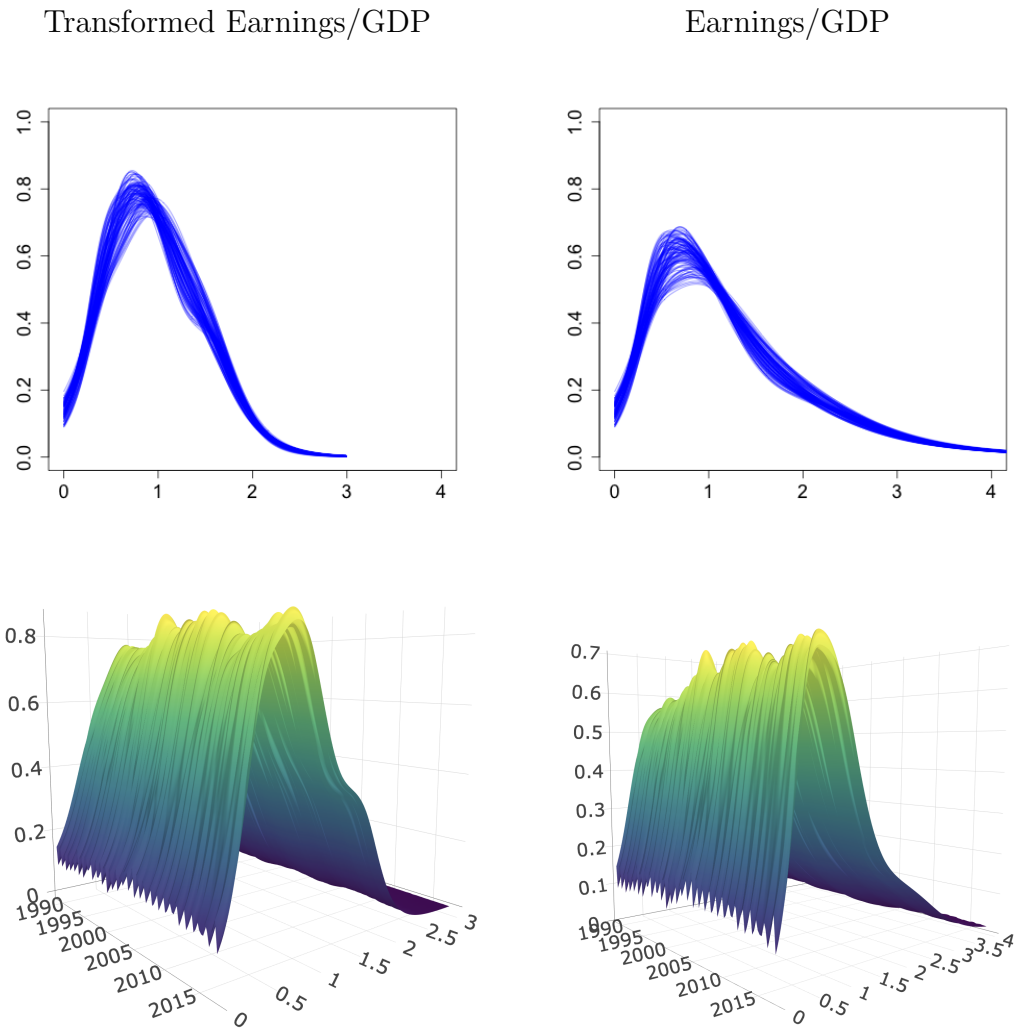
In the remainder of this paper we simply standardize individual-level earnings by (2/3) of nominal per-capita GDP. Rather than taking a logarithmic transformation of the earnings data, we apply the inverse hyperbolic sine transformation, which is given by

$$g(x|\theta) = \frac{\ln(\theta x + (\theta^2 x^2 + 1)^{1/2})}{\theta} = \frac{\sinh^{-1}(\theta x)}{\theta}, \quad x = \frac{\text{Earnings}}{(2/3) \cdot \text{per-capita GDP}}. \tag{30}$$

The function is plotted in the Online Appendix. We set $\theta = 1$. For small values of $x$ the function is approximately equal to $x$ and for large values of $x$ it is equal to $\log(x) + \log(2)$. This transformation avoids the thorny issue of applying a log transformation to earnings that are close to zero.

**Log-Spline Density Estimation.** We take the time period $t$ to be a quarter. For each quarter from 1989:Q2 to 2017:Q3 we estimate a cross-sectional density for the transformed earnings-to-GDP ratio; see (30). Seven knots for the log-spline density estimation are placed at the 0.01, 0.1, 0.25, 0.5, 0.75, 0.9, and 0.98 quantiles of the distribution of the pooled (across-time periods) transformed earnings data. It is important to note that the knot locations are identical for each period $t$, to ensure that the basis functions associated with the polynomial spline are time invariant. All the time variation in the densities is captured by

Figure 2: Estimated Log Earnings Distributions




*Notes:* Each hairline corresponds to the estimated density of earnings for a particular quarter $t$, where $t$ ranges from 1989:Q1 to 2017:Q3. Transformation is inverse hyperbolic sine transformation in (30).

the time coefficients which are estimated separately for each quarter. As mentioned above, in each period, we are normalizing the cross-sectional density by the fraction of individuals who reported to be employed. In Figure 2 we overlay the log-spline estimates of the cross-sectional densities. The left panel shows the density of the transformed earnings whereas the right panel shows the densities of the original earnings-to-GDP ratio which is obtained by a change-of-variables.

**Functional State-Space Model Estimation.** The log-spline density estimation generates the coefficients $\hat{\alpha}^o_{k,t}$, $k = 0, \ldots, K = 8$ and for $t$ ranging from 1989:Q2 to 2017:Q3. After

dropping the coefficients $\hat{\alpha}_{0,t}$ that normalize the densities, we proceed with the seasonal adjustment and the compression described in Section 3.2. This leads to the coefficient vector $\hat{\alpha}_t^c$ in (24) which turns out to be of dimension $6 \times 1$ after the compression step. The vector $Z_t^o$ in (5) is composed of TFP growth, real per-capita GDP growth, and the unemployment rate computed from the CPS data. We set $Z_*$ equal to the mean of these three series. Subsequent results about the dynamics of the cross-sectional densities are obtained by converting $\tilde{\alpha}_t^c$ vectors back into $\hat{\alpha}_t^o$ vectors as described in Section 3.5. We re-normalize the densities so that they integrate to the period $t$ employment rate, and then apply the change-of-variable formula to obtain a density for the earnings-to-GDP ratio.

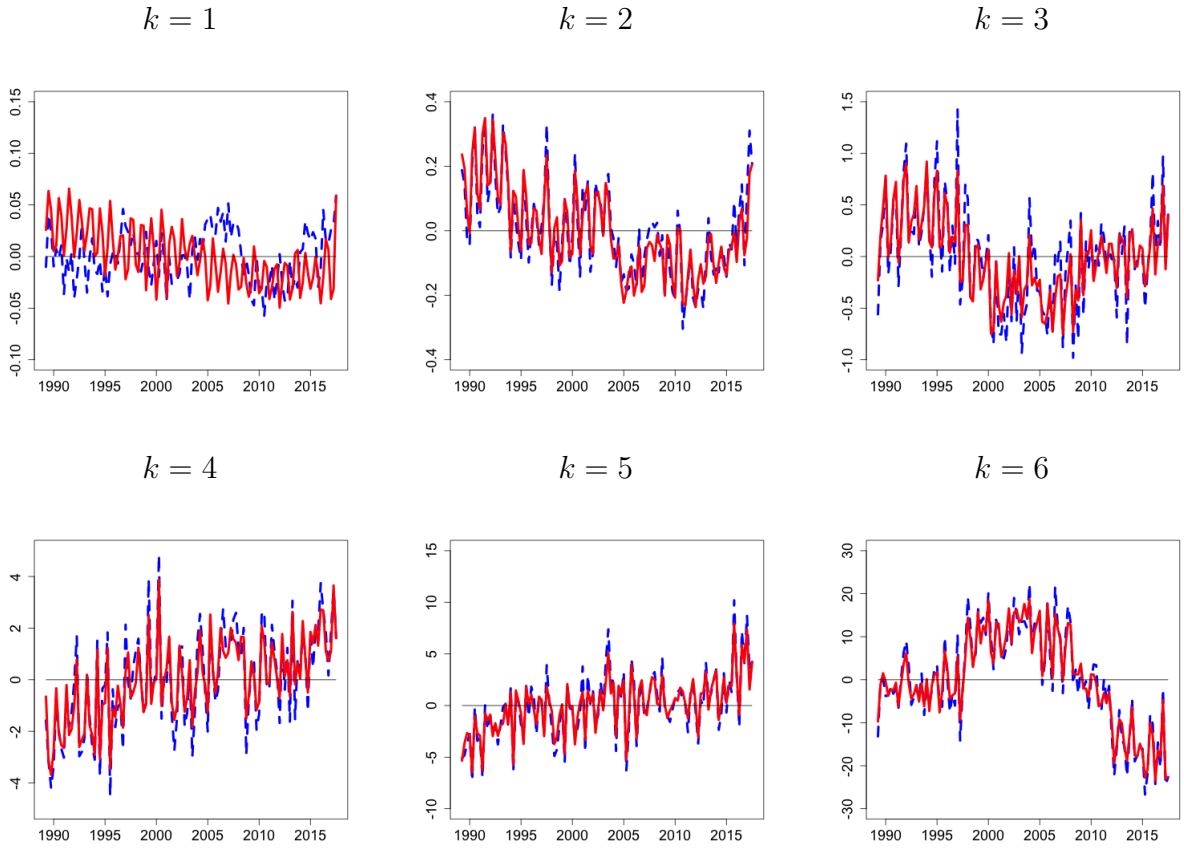
The estimation of the functional state-space model is based on the prior distribution in (28) and (29). We set $\underline{\nu} = n_w + 5$, where $n_w$ is the dimension of the vector $w_t$ in the state-transition equation (27), and let $\underline{\Sigma} = \underline{\nu}\hat{\Sigma}$, where $\hat{\Sigma}$ is the OLS estimator of $\Sigma$ in (27) that obtains when the latent $\tilde{\alpha}_t^c$ in the definition of $w_t$ is replaced by the observable $\hat{\alpha}_t^c$. The prior for $\phi$ is centered at $\underline{\mu}_\phi = 0$. We defer the discussion of the choice of the hyperparameter $\lambda$ to Section 4.2 below. Our estimation sample, after computing growth rates for TFP and GDP growth, ranges from 1989:Q2 to 2017:Q3.

Figure 3 overlays the observed $\hat{\alpha}_{k,t}^c$ versus the smoothed $\tilde{\alpha}_{k,t}^c$ generated as output of the Gibbs sampler. The discrepancy is the measurement error $\eta_{k,t}^c$, which is generally small. The observed and the smoothed series exhibit a strong comovement. For $k \geq 2$ the smoothed series are less volatile than the original series. Recall that the observed $\hat{\alpha}_{k,t}^c$ series are demeaned. All of the series show low frequency movements around zero in combination with some high frequency fluctuations. The $k = 4, 5$ exhibit a upward trend whereas the other series do not. The coefficient series as well as the estimated parameters of the functional state-transition equation are difficult to interpret, which is why we will examine their implications for the dynamics of the aggregate variables and the cross-sectional earnings distribution in the remainder of this section.

## 4.2 Do Distributional Dynamics Affect Aggregate Dynamics?

**Shrinking Toward Block-Diagonality.** The spillovers between distributional and aggregate dynamics are affected by the coefficient matrices $B_{z\alpha}^c$ and $B_{\alpha z}^c$ in (26). In fact, if these matrices are equal to zero then the aggregate variables do not Granger-cause the cross-sectional income distribution and vice versa. The prior covariance matrix $\underline{P}_\phi(\lambda)$ allows us to vary the degree of shrinkage for the off-diagonal elements which leads to a continuum

Figure 3: Observed $\hat{\alpha}^c_{k,t}$ versus Smoothed $\tilde{\alpha}^c_{k,t}$




*Notes:* The red solid lines correspond to the smoothed $\tilde{\alpha}^c_{k,t}$ series, whereas the blue dashed lines represent the observed series $\hat{\alpha}^c_{k,t}$.

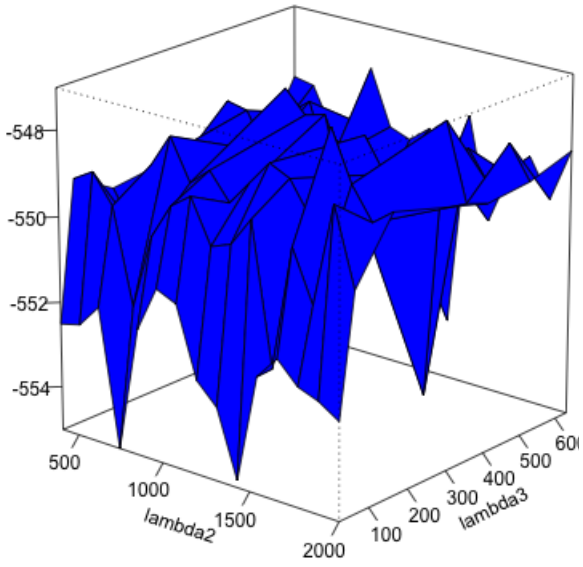
of model specifications of varying degree of spillovers. The fit of these model specifications can be characterized by their respective marginal data densities (MDDs) which include a penalty for model complexity. We plot the log marginal data density as a function of $\lambda_2$ and $\lambda_3$ in Figure 4 at the chosen hyperparameter $\hat{\lambda}_1 = 1.54$. To simplify the computation of the MDD surface, we set the measurement error covariance matrix $V_{\hat{\alpha}^c_t}$ to zero. While this would affect the ranking between the functional state-space model and other model classes, given the small magnitude of the estimated measurement errors in (3), we expect the distortion of inference about $\lambda$ to be small. We define

$$\hat{\lambda} = \text{argmin}_{\lambda \in \Lambda} \ \ln \text{MDD}(\lambda),$$

where $\Lambda$ is the grid that was used to generate Figure 4.

Recall that the hyperparameters scale the precision of the prior. The small value of $\hat{\lambda}_1 = 1.54$ indicates that the overall prior variance for the $\phi$ parameters is large. However,

Figure 4: Log Marginal Data Density as Function of $\lambda_2$ and $\lambda_3$




*Notes:* The log MDD surface is computed after setting the measurement error covariance matrix $V_{\hat{\alpha}_t^c}$ in (24) to zero. The log MDD is maximized at $\hat{\lambda}_1 = 1.54, \hat{\lambda}_2 = 1657.88$, and $\hat{\lambda}_3 = 159.4$ with $\ln \text{MDD}(\hat{\lambda}) = -547.03$.

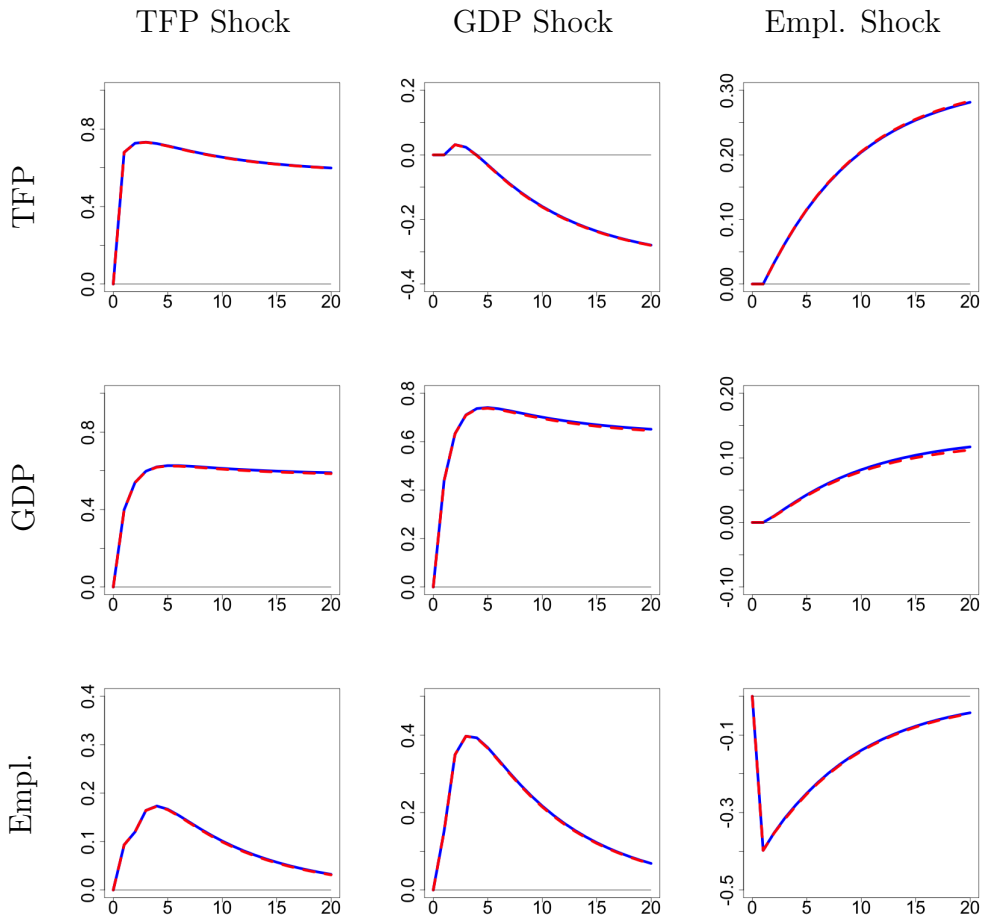
the large value of $\hat{\lambda}_2 = 1657.88$ implies that the fit (accounting for model complexity) improves by shrinking the parameters in the off-diagonal blocks $B_{z\alpha}^c$ to zero. Thus, overall the estimated spillovers are small. The empirical results in the remainder of this section condition on the hyperparameter estimate $\hat{\lambda}$.

**Impulse Response Comparison.** To illustrate the economic significance of the estimated off-diagonal blocks, we compare impulse response functions (IRFs) of the aggregate variables to the aggregate shocks under two parameterizations: (i) the posterior mean estimates of the matrices $(B_{zz}, B_{z\alpha}^c, B_{\alpha z}^c, B_{\alpha\alpha}^c, \Sigma)$; (ii) the posterior mean estimates of the matrices $(B_{zz}, B_{\alpha\alpha}^c, \Sigma)$ and zero off-diagonal blocks $B_{z\alpha}^c = 0$ and $B_{\alpha z}^c = 0$. In the vector $\tilde{Z}_t$ we order TFP growth first, GDP growth second, and the employment rate third. We use a Cholesky factorization $\Sigma = \Sigma_{tr}\Sigma_{tr}'$ to orthogonalize the vector of reduced-form innovations $u_t$:

$$u_t = \Sigma_{tr}\epsilon_t. \tag{31}$$

We interpret the first element $\epsilon_{1t}$ as technology innovation and the elements $\epsilon_{2t}$ and $\epsilon_{3t}$, broadly, as shocks to GDP growth and the employment rate. Results are depicted in Figure 5, which depicts responses of the employment rate and the log levels of TFP and GDP to a one-standard deviation shocks. Overall, the IRFs of the estimated system and the block-diagonal system are small. Despite these small difference, overall, the estimated dynamics

Figure 5: Impulse Responses of Aggregate Variables to Aggregate Shocks



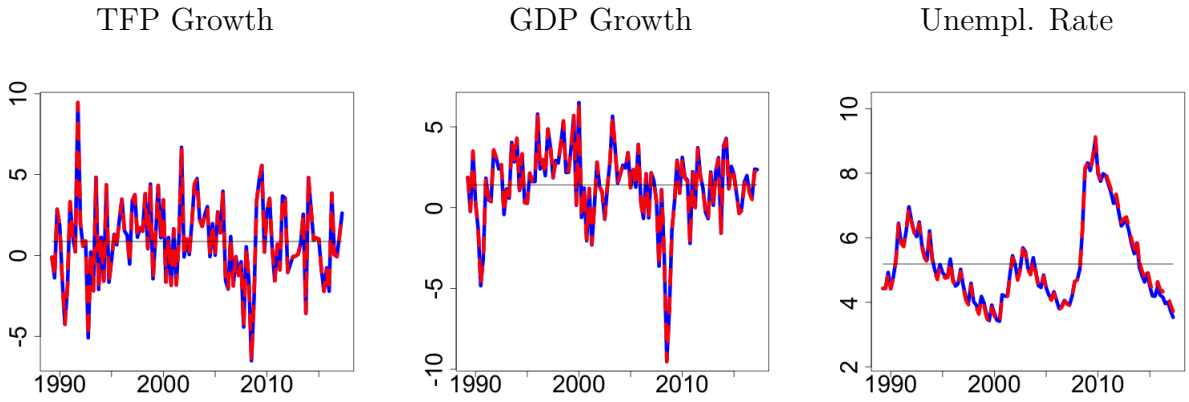

*Notes:* IRFs for one-standard deviation aggregate shocks (orthogonalized via Cholesky factorization; see (31)). Panels depict responses of the log level of TFP and GDP, scaled by 100, and responses of the employment rate in percent. Solid blue responses are based on posterior mean estimates of $B^c_{z\alpha}$ and $B^c_{\alpha z}$; dashed red responses are based on setting $B^c_{z\alpha} = 0$ and $B^c_{\alpha z} = 0$.

are very similar to that of a block-diagonal system, indicating that there is not strong channel that leads from aggregate shocks to distributional responses and feeds back into aggregate dynamics.

**Historical Decompositions.** We now examine the contribution of aggregate shocks to the fluctuations of TFP growth, GDP growth, and the employment rate. We do so by computing conditional on the posterior mean estimates (denoted by "hats" below) of $(\phi, \Sigma)$ and the partially unobserved states $\{w_t\}_{t=1}^T$, the historical sequence of orthogonalized innovations:

$$\hat{\epsilon}_t = \hat{\Sigma}_{tr}^{-1}(\hat{w}_t - \hat{\Phi}_1 \hat{w}_{t-1}). \tag{32}$$

Figure 6: Historical Decomposition of Aggregate Series




*Notes:* Actual paths (blue, solid) and counterfactual paths generated by innovations to macroeconomic aggregates only (red, dashed). Q-o-Q TFP and GDP growth rates are in annualized percentages. The unemployment rate is in percent.

Then, starting from $w_0 = \hat{w}_0$, we iterate the VAR difference equation forward based on a subset of the historical innovations:

$$w_t^* = \hat{\Phi}_1 w_{t-1}^* + \hat{\Sigma}_{tr} M_\epsilon \hat{\epsilon}_t, \tag{33}$$

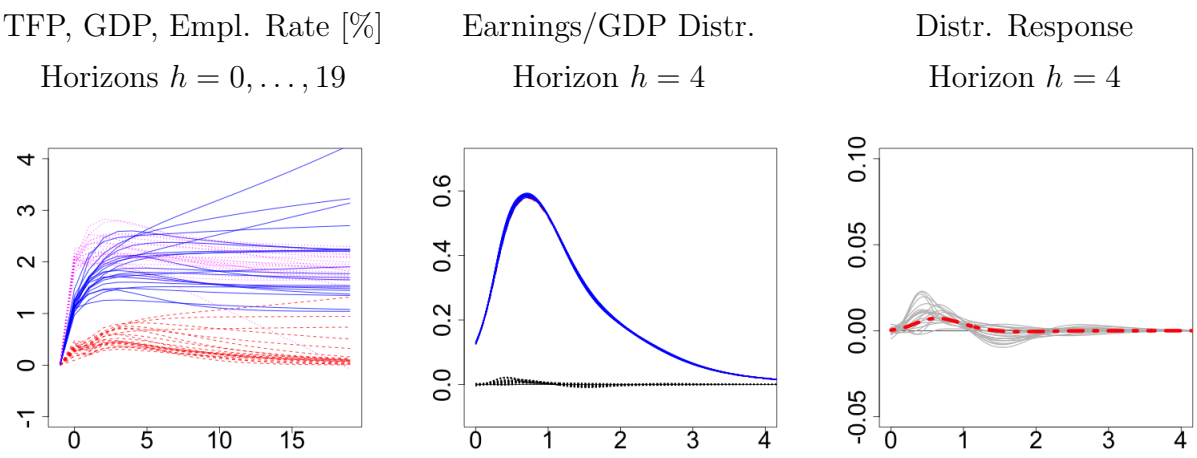
where $M_\epsilon$ is a diagonal selection matrix with zeros and ones on the diagonal that selects a subset of the elements of the $\hat{\epsilon}_t$ vector. The counterfactual path also reflects the effect of the initial condition. Recall that $w_t = [\tilde{Z}_t, \tilde{\alpha}_t]$. For the aggregate series, we simply overlay $\tilde{Z}_t$ and the counterfactual $\tilde{Z}_t^*$.

Results for the historical decompositions are depicted in Figure 6. The actual and counterfactual paths of TFP growth, GDP growth, and the unemployment rate are very similar, meaning that the contribution of the distributional shocks is negligible. This comes from the structure that's close to block-diagonal with the selected hyperparameters.

## 4.3 Effects of Aggregate Shocks

**Impulse Responses to a TFP Shock.** Recall that in the vector $w_t$ TFP growth is ordered first, GDP growth is second, the CPS-based unemployment is third, and the income distribution is last. Under the assumption that shocks to GDP growth and the income distribution do not affect measured TFP contemporaneously, the TFP growth innovation is identified as $\epsilon_{1t}$ based on the Cholesky factorization in (31). Figure 7 shows impulse

Figure 7: Impulse Responses to a TFP Shock

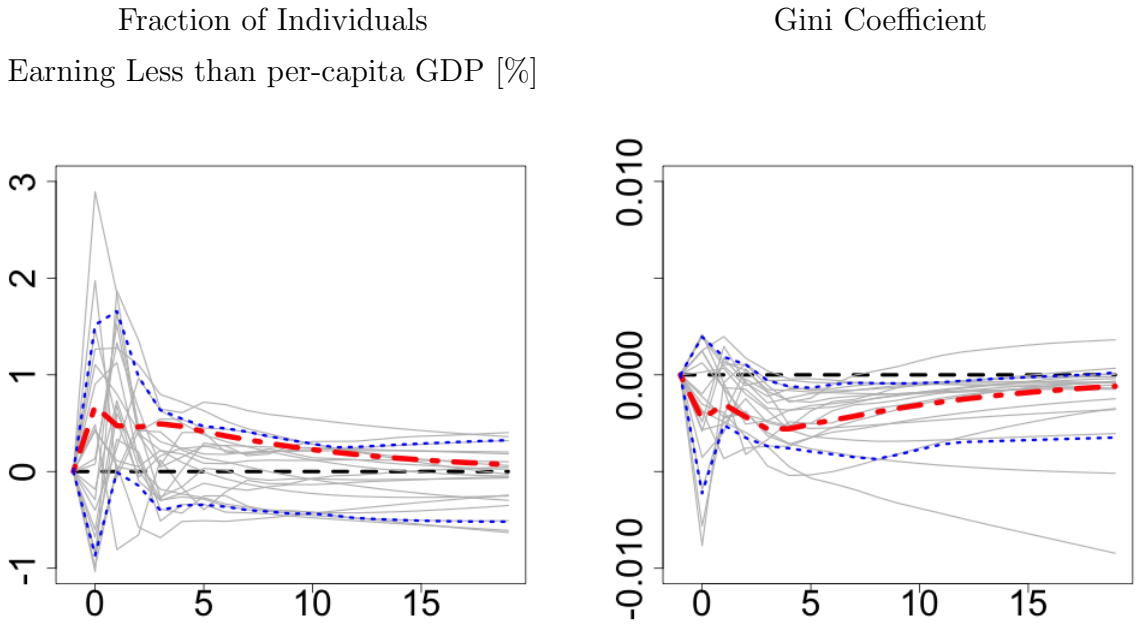



*Notes:* Responses to a 3-standard-deviations shock to TFP. The system is in steady state at $h = -1$ and the shock occurs at $h = 0$. Each hairline corresponds to a draw from the posterior distribution. Left panel: TFP (magenta, dotted), GDP (blue, solid) and Employment Rate (red, dashed). Center panel: steady state earnings/GDP density (red dashed), shocked density (blue solid), and difference (black). Right panel: zoomed-in difference between shocked and steady state density. Red dashed line is the posterior mean response.

responses to a three-standard deviation TFP innovation. We take a subsample of 20 draws from the posterior distribution of the VAR parameters and compute IRFs for each of these parameter draws. According to the left panel, in response to this shock the level of TFP rises in between 180 and 300 basis points (the figure depicts the response of log TFP, scaled by 100) in the log-run, whereas log per-capita GDP rises in between 150 and 300 basis points. Some of the TFP responses exhibit a slight hump-shaped pattern whereas most of the GDP responses are monotonic (recall, the responses are based on a functional VAR(1) state-transition equation). We also show responses of the employment rate. Here a number of one means that the employment rate increases by 100 basis points (or equivalently, the unemployment falls by one percentage point).

The center panel shows the estimated steady state density of the income distribution (obtained from the mean coefficients $\bar{\alpha}$) as well as the response of this density to the TFP shock after $h = 4$ quarters. Visually, the two densities are difficult to distinguish. We also overlay the difference between the steady state and the shocked density and then plot this difference separately in the right panel of Figure 7. For the center and right panels, a one on the $x$-axis refers to an individual whose earnings are equal to GDP per capita (adjusted for a labor share of 2/3). Maybe surprisingly, there seems to be a shift of probability mass

Figure 8: Impulse Responses to a TFP Shock – Continued




*Notes:* Responses to a 3-standard-deviations shock to TFP. The system is in steady state at $h = -1$ and the shock occurs at $h = 0$. Red dashed line is the posterior mean response. Blue lines stand for the 90th-quantile and the 10th-quantile of the responses.

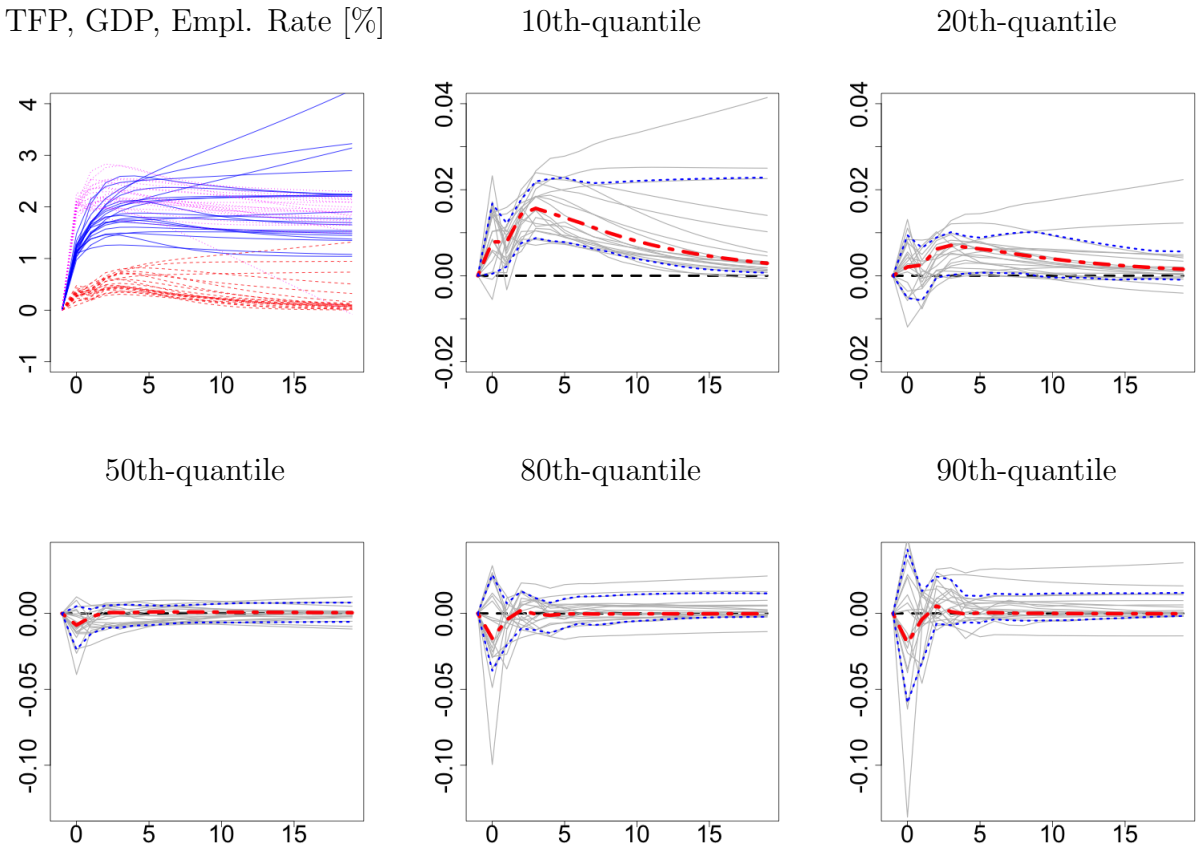
in the distribution of the log earnings-to-GDP ratio to the left. Now consider a hypothetical individual who shifts from 1.25 to 0.75 in this distribution in response to a TFP shock. Prior to the shock, this individual's earnings were 25% above the per-capita GDP (adjusted for a labor share of 2/3) benchmark, whereas after the shock it dropped to being 25% below the benchmark. Meanwhile the level of per-capita GDP rises by about 2.4%, making the individual significantly worse off. However, dynamics of the distribution of log earnings are mean reverting and the effect dies out fairly quickly.

To assess the overall effect of a technology shock on the earnings distribution we plot the response of the probability mass assigned to individuals with an earnings-to-per-capita-GDP ratio less than one in the left panel of Figure 8. This figure includes the band consisting of the 10th and the 90th quantile of the responses. The response is positive but fairly small – less than 2% for most hairlines and around 0.5% on average. The positive response is consistent with a model in which individuals are heterogeneous with respect to their skills and during expansions more low-skilled individuals enter the labor force. The right panel of the figure depicts the response of the Gini coefficient. The uncertainty band indicates a

slight decline of the Gini coefficient (decrease in inequality) after the shock, but the effect is quantitatively small.

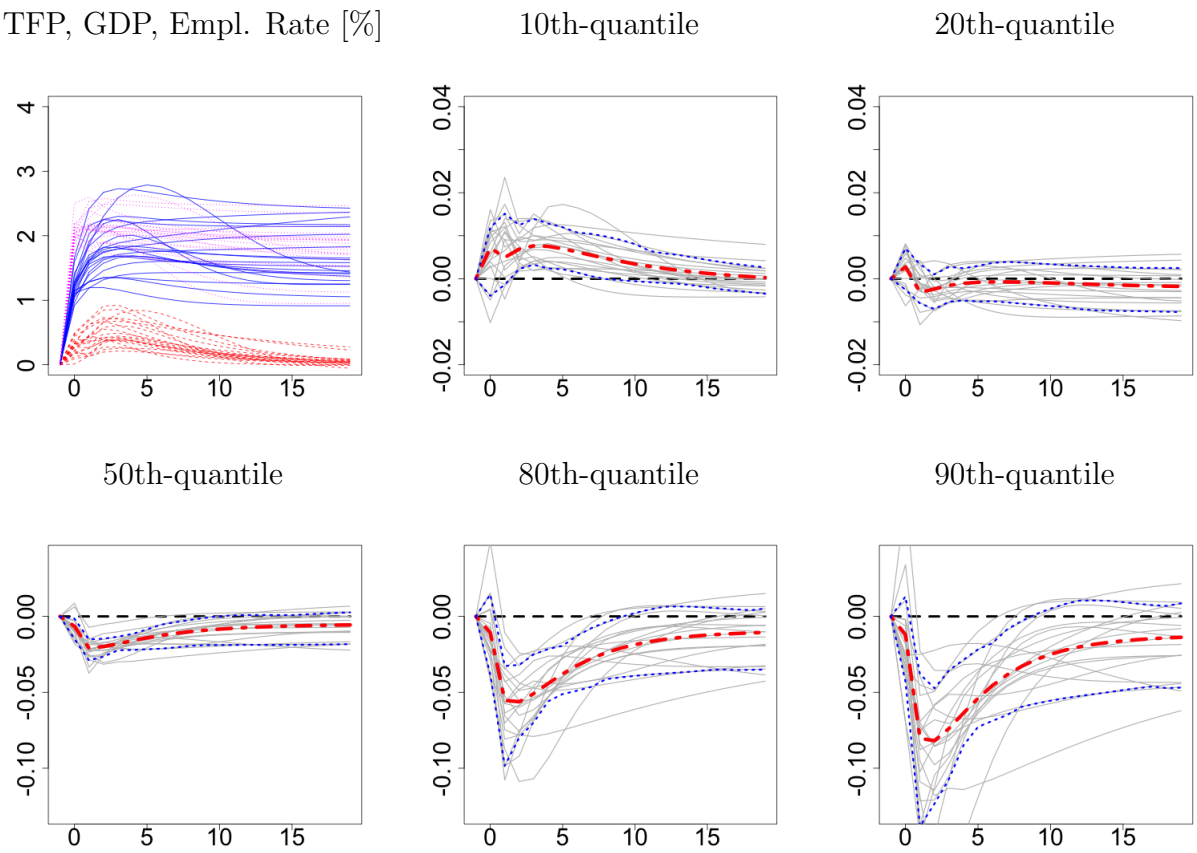
Another way to look at the overall effect of a technology shock on the distribution is to look at how its quantiles respond. Figure 9 show the responses of various quantiles (10, 20, 50, 80, and 90th). With a positive technology shock, there is an increase in 10th-quantile of the distribution as more low-skilled individuals start to earn. The median of the distribution does not respond much to this technology shock. For higher quantiles, there is more uncertainty involved, but the posterior mean response indicates that there is decrease in 80th and 90th-quantiles. Overall, Figure 9 show that the distribution is slightly more compressed, which is consistent with the decrease in Gini coefficient.

Figure 9: Responses to a TFP Shock – Continued




*Notes:* Responses to a 3-standard-deviations shock to TFP. The system is in steady state at $h = -1$ and the shock occurs at $h = 0$. Red dashed line is the response at the posterior mean. Blue lines stand for the 90th-quantile and the 10th-quantile of the responses.

Figure 10: Responses to a TFP Shock




*Notes:* Responses to a 3-standard-deviations shock to TFP. The system is in steady state at $h = -1$ and the shock occurs at $h = 0$. Red dashed line is the response at the posterior mean. Blue lines stand for the 90th-quantile and the 10th-quantile of the responses.

**Comparison to VAR with Quantiles.** One may wonder whether the quantiles are sufficient to span the information as in $\tilde{\alpha}$'s. In order to answer this question, we conduct a canonical correlation analysis to find linear combinations of quantiles and $\tilde{\alpha}$'s that have maximum correlations and are orthogonal to each other. The following is the output of the canonical correlation analysis using 10, 20, 50, 80, and 90th quantiles:

$$0.999,\ 0.997,\ 0.961,\ 0.868,\ 0.415$$

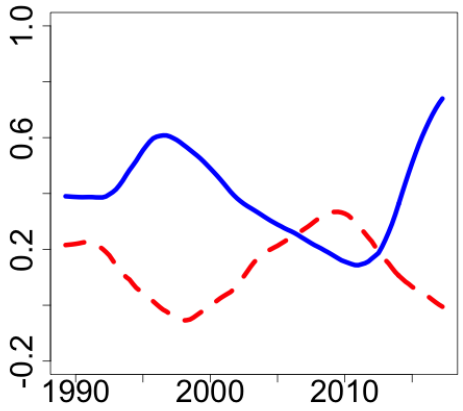
Although some correlations are high (as in first three numbers), the quantiles do not seem to span all the information as in $\tilde{\alpha}$'s.

We also run VAR with the aggregate variables and the 10-20-50-80-90 quantiles instead of adding $\tilde{\alpha}$'s to represent the distsribution. Hyperparameter selection gives $\lambda_1 = 0.22, \lambda_2 =$

1315.76, and $\lambda_3 = 9.83$ for the VAR with the quantiles. Then we obtain the analogous figure including IRFs of the quantiles to compare with Figure 9. Figure 10 shows the IRFs from the VAR with the aggregates and the quantiles. Based on the posterior mean response, the 10th-quantile increases whereas the median, 80th, and 90th-quantile decrease. Qualitatively, Figure 9 and Figure 10 show the compression of the earnings distribution after a positive technology shock. However, the VAR with the quantiles renders the bigger and significant drop in higher quantiles.

**Historical Decomposition.** We now examine what fraction of the fluctuation of the earnings densities is due to aggregate shocks. Starting point are the sequences $\hat{\tilde{\alpha}}_t^c$ and $\tilde{\alpha}_t^{c*}$, where the latter is generated by (33), shutting down the distributional shocks. Using the steps described in Section 3.5 we convert the compressed coefficients into densities $\hat{p}_t(x)$ and $p_t^*(x)$. In addition, we generate a sequence $w_t^0$ by setting $M_\epsilon = 0$ in (33).

Figure 11: Historical Decomposition of Variation in Earnings Densities




*Notes:* Importance measure for aggregate innovations (red, dashed) and distributional innovations (blue, solid); see (34).

Let $p_t^0(x)$ be the sequence of densities that corresponds to $\tilde{\alpha}_t^{c0}$. This sequence only captures the effect of the initial $\hat{w}_0$, but is not perturbed by $\epsilon_t$ innovations. Let $\Delta_{L_1}(p_t, q_t)$ be the $L_1$ distance between the densities $p_t(x)$ and $q_t(x)$. Finally, we define our measure of the importance of a particular group of innovations $\epsilon_{i,t}$. Suppose that $p_t^*(x)$ is generated by shutting down the innovations $\epsilon_{i,t}$. Then

$$IMP_t^*(\epsilon_i) = 1 - \frac{\Delta_{L_1}(p_t^*, p_t^0)}{\Delta_{L_1}(\hat{p}_t, p_t^0)} \tag{34}$$

can be viewed as a measure of the importance of the omitted innovation. If omitting the innovation has essentially no effect on $\tilde{\alpha}_t^*$, then $\hat{p}_t^*(x) \approx \hat{p}_t(x)$ and $IMP_t^*(\epsilon_i) \approx 0$. If the shock explains most of the fluctuations in $\tilde{\alpha}_t$, then $\Delta_{L_1}(\hat{p}_t^*, \hat{p}_t^0)$ is close to zero and $IMP_t^*(\epsilon_i)$ is approximately one. Because the mapping from the $\tilde{\alpha}_t^c$ coefficients into the $L_1$ distances between the three densities is nonlinear, the specific magnitudes (other than the extremes of zero and one) are unfortunately difficult to interpret.

Results are depicted in Figure 11. We show the importance measure for the aggregate and distributional shocks because due to the nonlinearity of the transformation the two measures do not add up to one. The figure shows that while there are some spillovers from the aggregate shocks to the distributional dynamics, most of the fluctuations in the earnings distributions are driven by distributional shocks.

## 4.4 Effects of Distributional Shocks

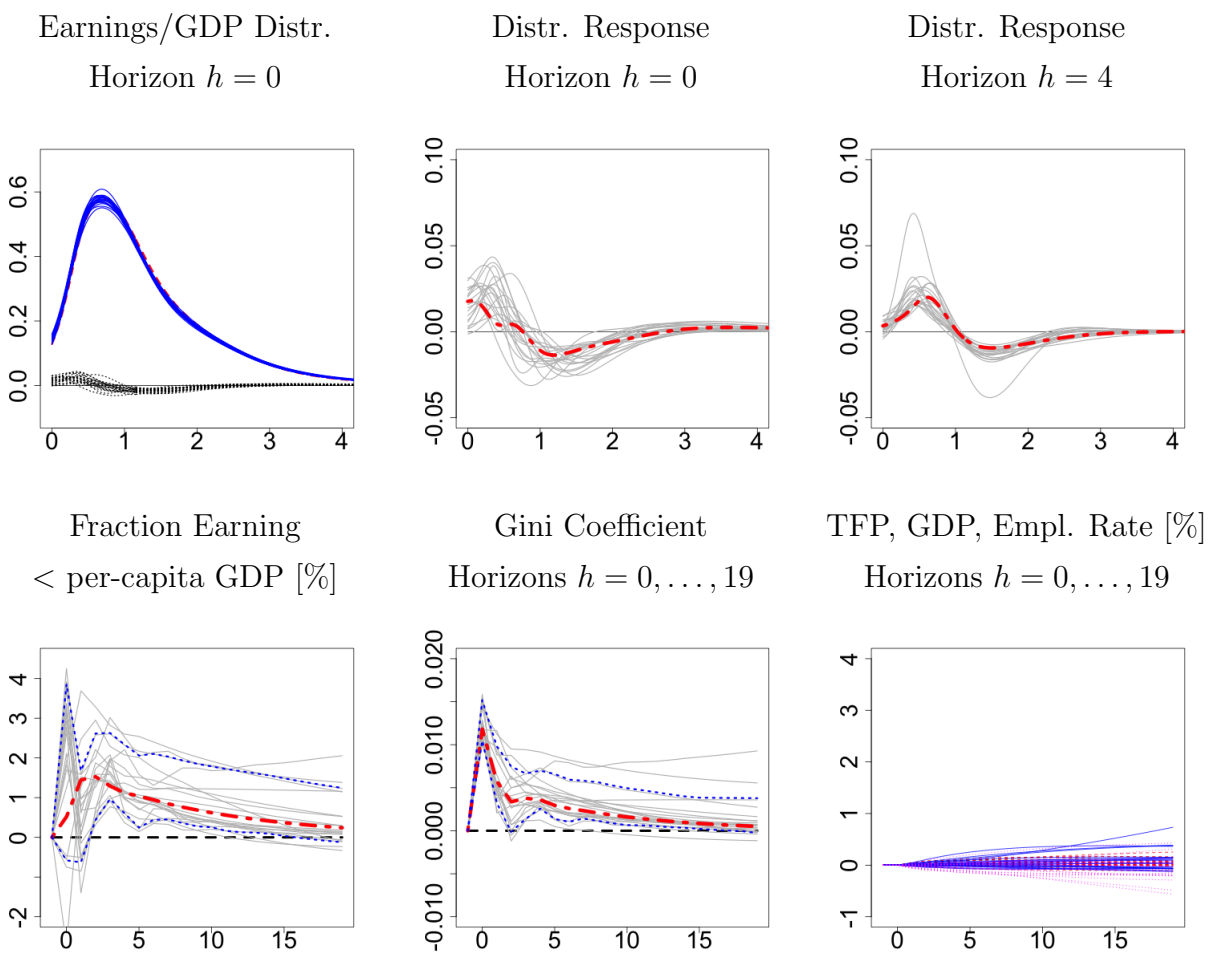
The identification of a distributional shock requires additional assumptions. Recall that $u_t$ has the partitions $u_{z,t}$ and $u_{\alpha^c,t}$. We denote the conforming partitions of $\Sigma_{tr}$ by $\Sigma_{tr,zz}$, $\Sigma_{tr,z\alpha} = 0$, $\Sigma_{tr,\alpha z}$, and $\Sigma_{tr,\alpha,\alpha}$. Now let $q = [0, q_\alpha]$ be a unit-length vector with partitions that conform with the partitions of $\Sigma_{tr}$.

We then define the impact vector of the distributional shock by

$$\begin{bmatrix} \Sigma_{tr,zz} & 0 \\ \Sigma_{tr,\alpha z} & \Sigma_{tr,\alpha\alpha} \end{bmatrix} \begin{bmatrix} 0 \\ q_\alpha \end{bmatrix}.$$

We determine the vector $q_a$, which maximizes the effect on the Gini coefficient. We consider a shock that generates a maximal increase in the Gini coefficient upon impact. The results are depicted in Figure 12. The average response of the Gini coefficient spikes and fades away by construction of the distributional shock. The response in the fraction of earnings that are less than the labor share of per-capita GDP increases and this change stays for a while. The effect of distributional shock on the aggregate variables is ambiguous in the short run and in the long run. Hence, it is hard to say that aggregate variables show significant responses given a shock that maximizes the Gini coefficient based on the earnings distribution.

Figure 12: Responses to a Distributional Innovation $\epsilon_t^*$ (Maximize Gini-Coefficient)




*Notes:* Responses to a 3-standard-deviation distributional innovation $\epsilon_t^*$. The system is in steady state at $h = -1$ and the shock occurs at $h = 0$. Each hairline corresponds to a draw from the posterior distribution. Dashed lines in panels (1,2), (1,3), (2,1) and (2,2) are posterior mean responses. Panel (1,1): steady state log earnings density (red dashed), shocked density (blue solid), and difference (black). Panels (1,2) and (1,3): difference between shocked and steady state density. Panel (2,3): TFP (magenta, dotted), GDP (blue, solid) and Employment Rate (red, dashed).

# 5 Conclusion

We developed a vector autoregressive model that stacks macroeconomic aggregates and and cross-sectional distributions to provide semi-structural evidence about the interaction of aggregate and distributional dynamics. We applied the model to examine the effect of a technology shock on the earnings distribution and the effect of a shock that moves the earnings distribution on macroeconomic aggregates. The technique developed in this paper should be useful more broadly for the evaluation of heterogeneous agent macro models.

# References

Ahn, S., G. Kaplan, B. Moll, T. Winberry, and C. Wolf (2018): "When Inequality Matters for Macro and Macro Matters for Inequality," in *NBER Macroeconomics Annual 2017*, ed. by M. Eichenbaum, and J. Parker, pp. 1 – 75. University of Chicago Press.

Algan, Y., O. Allais, and W. Den Haan (2008): "Solving Heterogeneous-Agent Models with Parameterized Cross-Sectional Distributions," *Journal of Economic Dynamics and Control*, 32(3), 875–908.

Auclert, A., and M. Rognlie (2018): "Inequality and Aggregate Demand," *Manuscript, Stanford University*.

Bosq, D. (2000): *Linear Processes in Function Spaces*. Springer Verlag, New York.

Carter, C., and R. Kohn (1994): "On Gibbs Sampling for State Space Models," *Biometrika*, 81, 541–553.

Chang, Y., C. S. Kim, and J. Park (2016): "Nonstationary in Time Series of State Densities," *Journal of Econometrics*, 192, 152–167.

Chang, Y., S.-B. Kim, and F. Schorfheide (2013): "Labor Market Heterogeneity, Aggregation, and the Policy-(In)variance of DSGE Model Parameters," *Journal of the European Economic Association*, 11(S1), 193–220.

Childers, D. (2015): "On the SoluSolu and Application of Rational Expectations Models with Function-Valued States," *Manuscript, Carnegie Mellon University*.

Coibion, O., Y. Gorodnichenko, L. Kueng, and J. Silvia (2017): "Innocent Bystanders? Monetary Policy and Inequality," *Journal of Monetary Economics*, 88, 70–89.

Diebold, F. X., and C. Li (2006): "Forecasting the Term Structure of Government Bond Yields," *Journal of Econometrics*, 130, 337–364.

Efron, B., and R. Tibshirani (1996): "Using Specially Designed Exponential Families for Density Estimation," *Annals of Statistics*, 24(6), 2431–2461.

Fernald, J. G. (2012): "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity," *FRBSF Working Paper*, 2012-19.

Horvath, L., and P. Kokoszka (2012): *Inference for Functional Data with Applications*. Springer Verlag, New York.

Hu, B., and J. Y. Park (2017): "Econometric Analysis of Functional Dynamics in the Presence of Persistence," *Manuscript, Department of Economics, Indiana University*.

Inoue, A., and B. Rossi (2018): "The Effects of Conventional and Unconventional Monetary Policy: A New Approach," *Manuscript, Vanderbilt University and Pompeu Fabra*.

Kaplan, G., and L. Violante, Giovanni (2018): "Microeconomic Heterogeneity and Macroeconomic Shocks," *Journal of Economic Perspective*, forthcoming.

Kooperberg, C., and C. J. Stone (1990): "A Study of Logspline Density Estimation," *Computational Statistics & Data Analysis*, 12, 327–347.

Krusell, P., and A. A. Smith (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106(5), 867–896.

Kurmann, A., and C. Otrok (2013): "News Shocks and the Slope of the Term Structure of Interest Rates," *American Economic Review*, 103, 2612–2632.

Mongey, S., and J. Williams (2017): "Firm Dispersion and Business Cycles: Estimating Aggregate Shocks Using Panel Data," *Manuscript, New York University*.

Ottonello, P., and T. Winberry (2018): "Financial Heterogeneity and the Investment Channel of Monetary Policy," *NBER Working Paper*, 24221.

Ramsey, J. O., and B. W. Silverman (2005): *Functional Data Analysis*. Springer Verlag, New York, 2nd edn.

Reiter, M. (2010): "Approximate and Almost-Exact Aggregation in Dynamic Stochastic Heterogeneous-Agent Models," *IHS Working Paper, Economics Series*, 258.

Uhlig, H. (2005): "What Are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure," *Journal of Monetary Economics*, 52, 381–419.

Varin, C., N. Reid, and D. Firth (2011): "An Overview of Composite Likelihood Methods," *Statistica Sinica*, 21, 5–42.

Winberry, T. (2017): "A Toolbox for Solving and Estimating Heterogeneous Agent Macro Models," *Working Paper, Chicago Booth*.

# Online Appendix: Heterogeneity and Aggregate Fluctuations

**Minsu Chang, Xiaohong Chen, and Frank Schorfheide**

# A State-Space Model

## A.1 Measurement Error Covariance Matrix

We construct the measurement error covariance matrix from the Hessian of the log-likelihood function for the cross-sectional observations. Let $\zeta_k(x)$ be the basis functions for the log spline approximation of the cross-sectional density $p_t(x)$ and let $X_t^o = \{x_{1t}^o, \ldots, x_{Nt}^o\}$. The density of the observations $X_t^o$ in period $t$ can be expressed as

$$\mathcal{L}(\alpha_t|X_t^o) = p^{(K)}(X_t^o|\alpha_t) = \exp\left\{\bar{\zeta}(X_t^o)\alpha_t\right\}, \tag{A.1}$$

subject to the restriction that the density normalizes to one:

$$\alpha_{0,t} = \int \exp\left\{\sum_{k=1}^{K} \alpha_{k,t}\zeta_k(x)\right\} dx. \tag{A.2}$$

Imposing this restriction, the log likelihood function becomes

$$\mathcal{L}(\alpha_t|X_t^o) = \sum_{k=1}^{K} \alpha_{k,t}\bar{\zeta}_k(X_t^o) - N \ln \int \exp\left\{\sum_{k=1}^{K} \alpha_{k,t}\zeta_k(x)\right\} dx. \tag{A.3}$$

The first-order derivatives with respect to $\alpha_k$ for $k = 1, \ldots, K$ are given by

$$\mathcal{L}_k^{(1)}(\alpha_t|X_t^o) = \bar{\zeta}_k(X_t^o) - N\frac{\int \zeta_k(x) \exp\left\{\sum_{k=1}^{K} \alpha_{k,t}\zeta_k(x)\right\} dx}{\int \exp\left\{\sum_{k=1}^{K} \alpha_{k,t}\zeta_k(x)\right\} dx} \tag{A.4}$$

The second-order derivatives are given by

$$\mathcal{L}_{kl}^{(2)}(\alpha_t|X_t^o) = -N\frac{A}{B}, \tag{A.5}$$

where

$$\begin{aligned} A &= \left[\int \zeta_k(x)\zeta_l(x) \exp\left\{\sum_{k=1}^{K} \alpha_{k,t}\zeta_k(x)\right\} dx\right]\left[\int \exp\left\{\sum_{k=1}^{K} \alpha_{k,t}\zeta_k(x)\right\} dx\right] \\ &\quad - \left[\int \zeta_k(x) \exp\left\{\sum_{k=1}^{K} \alpha_{k,t}\zeta_k(x)\right\} dx\right]\left[\int \zeta_l(x) \exp\left\{\sum_{k=1}^{K} \alpha_{k,t}\zeta_k(x)\right\} dx\right] \\ B &= \left[\int \exp\left\{\sum_{k=1}^{K} \alpha_{k,t}\zeta_k(x)\right\} dx\right]^2. \end{aligned}$$

Using the constraint on $\alpha_{0,t}$ in (A.2), we can simplify the expressions to

$$\begin{aligned} A &= \left[\int \zeta_k(x)\zeta_l(x)\exp\left\{\sum_{k=1}^{K}\alpha_{k,t}\zeta_k(x)\right\}dx\right]\exp\{-a_0\} \\ &\quad -\left[\int \zeta_k(x)\exp\left\{\sum_{k=1}^{K}\alpha_{k,t}\zeta_k(x)\right\}dx\right]\left[\int \zeta_l(x)\exp\left\{\sum_{k=1}^{K}\alpha_{k,t}\zeta_k(x)\right\}dx\right] \\ B &= \exp\{-2a_{0,t}\}. \end{aligned}$$

We can therefore write the $K \times K$ Hessian as

$$\mathcal{H}(\alpha_t) = \left[\mathcal{L}^{(2)}_{kl}(\alpha_t|X^o_t)\right] \tag{A.6}$$

with elements

$$\begin{aligned} &\mathcal{L}^{(2)}_{kl}(\alpha_t|X^o_t) \\ &= -N\left(\int \zeta_k(x)\zeta_l(x)p^{(K)}(x|\alpha_t)dx - \int \zeta_k(x)p^{(K)}(x|\alpha_t)dx\int \zeta_l(x)p^{(K)}(x|\alpha_t)dx\right) \\ &= -N\int\left(\zeta_k(x) - \int \zeta_k(x)p^{(K)}(x|\alpha_t)dx\right)\left(\zeta_l(x) - \int \zeta_l(x)p^{(K)}(x|\alpha_t)dx\right)p^{(K)}(x|\alpha_t)dx. \end{aligned}$$

and

$$p^{(K)}(x|\alpha_t) = \exp\Big\{\zeta(x)\alpha_t\Big\}.$$

The Hessian can be evaluated at $\alpha_t = \hat{\alpha}^r_t$.

We subsequently make two adjustments to the coefficients. First, we remove a (seasonal) mean $\bar{\alpha}_t = \sum_{q=1}^{4} s_q(t)\bar{\alpha}_{q,k}$. Second, we compress the coefficients to eliminate perfect collinearities:

$$(\hat{\alpha}^r_t - \bar{\alpha}_t) = \Lambda\hat{\alpha}^o_t,$$

which implies that

$$\hat{\alpha}^o_t = (\Lambda'\Lambda)^{-1}\Lambda'(\hat{\alpha}^r_t - \bar{\alpha}_t).$$

Thus, we define

$$V_{\hat{\alpha}^o_t} = -\big(\hat{\Lambda}\hat{\Lambda}'\big)^{-1}\hat{\Lambda}\big[\mathcal{H}(\hat{\alpha}^r_t)\big]^{-1}\hat{\Lambda}'\big(\hat{\Lambda}\hat{\Lambda}'\big)^{-1}. \tag{A.7}$$

# B Shock Identification

Here we provide additional details on how to identify a shock that maximizes the contribution to the variance of variable $i$ at horizons $h = 1, \ldots, \bar{h}$. Define the matrix $M = \left[0_{n_z \times n_{\alpha^c}}, I_{n_z}\right]$ and the vector $e_i$ that has a one in position $i$ and zeros elsewhere such that we can write

$$w_{i,t+h} - \mathbb{E}[w_{i,t+h}] = \ldots + e_i' \sum_{j=0}^{h-1} \Phi_1^j \Sigma_{tr} M q_\alpha + \ldots.$$

We can now define $q_\alpha^*$ as the impact effect of the shock that maximizes the forecast error variance over horizons $h = 1, \ldots, \bar{h}$:

$$q_\alpha^* = \text{argmax } e_i' \left[ \sum_{h=1}^{\bar{h}} \sum_{j=0}^{h-1} \Phi_1^j \Sigma_{tr} M q_\alpha q_\alpha' M' \Sigma_{tr}' (\Phi_1^j)' \right] e_i. \tag{A.8}$$

Using the facts that $x'A'x = \text{tr}[xx'A]$ and $\text{tr}[AB] = \text{tr}[BA]$, we can rewrite the objective function as

$$\begin{aligned} & e_i' \left[ \sum_{h=1}^{\bar{h}} \sum_{j=0}^{h-1} \Phi_1^j \Sigma_{tr} M q_\alpha q_\alpha' M' \Sigma_{tr}' \Phi_1^{j'} \right] e_i \qquad (A.9) \\ = & \sum_{h=1}^{\bar{h}} \sum_{j=0}^{h-1} \text{tr}\left[ (e_i e_i')(\Phi_1^j \Sigma_{tr} M)(q_\alpha q_\alpha')(M' \Sigma_{tr}' \Phi_1^{j'}) \right] \\ = & \sum_{h=1}^{\bar{h}} \sum_{j=0}^{h-1} \text{tr}\left[ (q_\alpha q_\alpha')(M' \Sigma_{tr}' \Phi_1^{j'})(e_i e_i')(\Phi_1^j \Sigma_{tr} M) \right] \\ = & q_\alpha' \left[ \sum_{h=1}^{\bar{h}} \sum_{j=0}^{h-1} (M' \Sigma_{tr}' \Phi_1^{j'})(e_i e_i')(\Phi_1^j \Sigma_{tr} M) \right] q_\alpha \\ = & q_\alpha' S q_\alpha. \end{aligned}$$

The optimization problem can therefore be expressed as Lagrangian

$$\mathcal{L} = q_\alpha' S q_\alpha - \lambda (q_\alpha' q_\alpha - 1), \tag{A.10}$$

which leads to the first-order condition

$$S q_\alpha = \lambda q_\alpha. \tag{A.11}$$

At the first-order condition, we obtain that $\mathcal{L} = \lambda$. Thus, the solution is obtained by finding the eigenvector associated with the largest eigenvalue of the matrix $S$.

# C Supplemental Information About the Empirical Analysis

## C.1 Data Construction

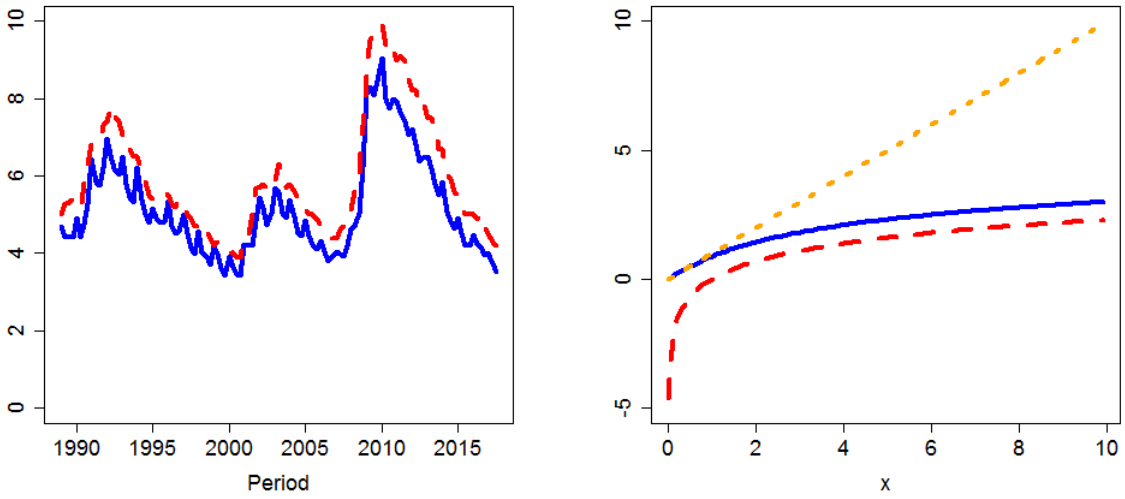
The observations on real per capita GDP, GDP deflator, and the unemployment rate are downloaded from the Federal Reserve Bank of St. Louis' FRED database:
https://fred.stlouisfed.org/.
The TFP series is available from the Federal Reserve Bank of San Francisco:
https://www.frbsf.org/economic-research/indicators-data/total-factor-productivity-tfp/.
The labor share series is available from the Bureau of Labor Statistics, labor productivity and cost measures: https://www.bls.gov/lpc/.
The CPS raw data are downloaded from
http://www.nber.org/data/cps_basic.html.
The raw data files are converted into STATA using the do-files available at:
http://www.nber.org/data/cps_basic_progs.html.
We use the series PREXPLF ("Experienced Labor Force Employment"), which is the same as in the raw data, and the series PRERNWA ("Weekly Earnings"), which is constructed as PEHRUSL1 ("Hours Per Week at One's Main Job") times PRHERNAL ("Hourly Earnings") for hourly workers, and given by PRWERNAL for weekly workers. STATA dictionary files are available at:
http://www.nber.org/data/progs/cps-basic/

## C.2 Data Transformations

The left panel of Figure 13 compares the aggregate unemployment rate to the unemployment rate computed from the CPS data. The levels of the two series are very similar, but the CPS unemployment rate exhibits additional high-frequency flucutations, possibly due to seasonals that have been removed from the aggregate unemployment rate.

Figure 13: CPS Unemployment and Earnings Transformation

Unemployment Rates Income Transformation




*Notes:* Left panel: CPS unemployment rate (blue, solid) and aggregate unemployment rate (red, dased). Right panel: inverse hyperbolic sine transformation (blue, solid) for $\theta = 1$ given in Eq. (A.12), logarithmic transformation (red, dashed), and 45-degree line (orange, dotted).

We transform the earnings-GDP ratio using the inverse hyperbolic sine transformation, which is given by

$$g(x|\theta) = \frac{\ln(\theta x + (\theta^2 x^2 + 1)^{1/2})}{\theta} = \frac{\sinh^{-1}(\theta x)}{\theta}. \tag{A.12}$$

The transformation is plotted in the right panel of Figure 13 for $\theta = 1$. Note that $g(0|\theta) = 0$ and $g^{(1)}(0|\theta) = 1$, that is, for small values of $x$ the transformation is approximately linear. For large values of $x$ the transformation is logarithmic:

$$g(x|\theta) \approx \frac{1}{\theta}\ln(2\theta x) = \ln 2 + \frac{1}{\theta}\ln(x).$$

# D Outline of Theory

The data generating process is given by (3), (4), (5), we we reproduce for convenience

$$\begin{aligned}
\ln Z_t^o &= \ln Z_* + \tilde{Z}_t \\
x_{it}^o &\sim \text{iid}\, p_t(x) = \frac{\exp\{\ell_*(x) + \tilde{\ell}_t(x)\}}{\int \exp\{\ell_*(x) + \tilde{\ell}_t(x)\} dx}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T \\
\tilde{Z}_t &= B_{zz}\tilde{Z}_{t-1} + \boldsymbol{B}_{zl}[\tilde{\ell}_{t-1}] + u_{z,t} \\
\tilde{\ell}_t(x) &= B_{lz}(x)\tilde{Z}_{t-1} + \boldsymbol{B}_{ll}[\tilde{\ell}_{t-1}](x) + u_{l,t}(x)
\end{aligned}$$

The first part of our theory focuses on the analysis of a least squares estimator that is generated as follows:

1. Compute period-by-period ML estimates $\hat{\alpha}_t^o$; see (11).
2. Let $\hat{\alpha}_* = \frac{1}{T}\sum_{t=1}^T \hat{\alpha}_t^o$ and $\tilde{\alpha}_t = \hat{\alpha}_t^o - \hat{\alpha}_*$.
3. Let $\widehat{\ln Z_*} = \frac{1}{T}\sum_{t=1}^T \ln Z_t^o$ and $\tilde{Z}_t = \ln Z_t^o - \widehat{\ln Z_*}$.
4. Use OLS to estimate
$$\begin{bmatrix} \tilde{Z}_t \\ \tilde{\alpha}_t \end{bmatrix} = \begin{bmatrix} B_{zz} & B_{zl} \\ B_{lz} & B_{ll} \end{bmatrix} \begin{bmatrix} \tilde{Z}_{t-1} \\ C_\alpha \tilde{\alpha}_{t-1} \end{bmatrix} + \begin{bmatrix} u_{z,t} \\ u_{\alpha,t} \end{bmatrix}, \quad C_\alpha = \int \xi'(\tilde{x})\zeta(\tilde{x})d\tilde{x}.$$

**Step 1:** Control "biases" from finite-dimensional approximations. Given $K$:

- Assuming that $N \longrightarrow \infty$, derive pseudo-true $\bar{\ell}_t^{(K)}$.
- Given $\bar{\ell}_t^{(K)}$ and assuming $T \longrightarrow \infty$, derive pseudo-true $\bar{\boldsymbol{B}}_{ll}^{(K)}(x, \tilde{x})$, etc., based on the definition of the OLS estimator.

**Step 2:** Control "variances" that result from replacing:

- pseudo-true $\bar{\ell}_t^{(K)}$ by an estimate $\hat{\ell}_t^{(K)}$ (sample size $N$);
- $\bar{\boldsymbol{B}}_{ll}^{(K)}(x, \tilde{x})$, etc., by estimates $\hat{\boldsymbol{B}}_{ll}^{(K)}(x, \tilde{x})$, etc., based on $\hat{\ell}_t^{(K)}$ (sample size $T$);

**Step 3:** Under suitable regularity conditions for DGP, derive bounds on "bias" and "variance" to obtain rates for the consistency of coefficients, forecasts, IRFs as $(K, N, T) \longrightarrow \infty$.