ARTICLE IN PRESS

Economics Letters ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at SciVerse ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

A note on the (in)consistency of the test of overidentifying restrictions and the concepts of true and pseudo-true parameters

Patrik Guggenberger *

Department of Economics, UCSD, 9500 Gilman Dr., La Jolla, CA 92093-0508, United States

ARTICLE INFO

Article history:
Received 2 November 2010
Received in revised form
28 June 2012
Accepted 7 July 2012
Available online xxxx

JEL classification:
C1
C2
C12

Keywords:
Nonexogenous instruments
Pretest
Pseudo-true parameter vector
Test of overidentifying restrictions

ABSTRACT

In the linear instrumental variables model, we characterize fixed alternatives against which the test of overidentifying restrictions (OR) is inconsistent. When there is the notion of a "true parameter", we relate this inconsistency result to the literature on optimality properties of various versions of the test of OR.

1. Introduction and motivation

For motivation, consider the linear regression model

$$y_1 = y_2\theta^* + X\gamma + \widetilde{u} \in R^n, \qquad E(\widetilde{u}_i|y_{2i}, X_i) = 0, \tag{1}$$

where $\theta^* \in \Theta \subset R$, $\gamma \in R^d$, and by $\widetilde{u}_i$, y_{2i}, and X_i we denote the i-th rows of $\widetilde{u}$, y_2, and X, respectively, written as column vectors (or scalars) and similarly for other random variables. The parameter of interest θ^* in this model is the *expected change in* y_{1i} *of changing* y_{2i} *by one unit holding* X_i *constant*. We assume that the ultimate goal of the applied researcher is inference on this "true parameter" θ^*. For example, in a wage regression, the applied researcher might be interested in the expected change in wage of changing an individual's education by one year holding all other covariates fixed.

An applied researcher may be faced with various data problems that complicate inference in (1) and, in particular, rule out OLS inference. As one scenario, assume there is measurement error $\varepsilon_i = (\varepsilon_{yi}, \varepsilon_{Xi}')' \in R^{1+d}$ and instead of $(y_{2i}, X_i')'$, the applied researcher observes $(y_{2i}^*, X_i^{*\prime})' = (y_{2i}, X_i')' + (\varepsilon_{yi}, \varepsilon_{Xi}')'$. For example, the variable "years of education" may be mismeasured in a wage regression. The measurement error ε_i or $\widetilde{u}_i$ may be be correlated with $(y_{2i}^*, X_i^{*\prime})'$ leading to an endogeneity problem in the regression of observed variables $y_1 = y_2^*\theta^* + X^*\gamma + u$, for $u = \widetilde{u} - \varepsilon_y\theta^* - \varepsilon_X\gamma$.

As a second scenario, assume a component of the covariate vector X_i, $X_{2i} \in R^f$ for $f \leq d$ and $X_i = (X_{1i}', X_{2i}')'$, may not be observed by the applied researcher, for instance "ability" in a wage regression. If the variable X_{2i} is correlated with y_{2i} or X_{1i}, then an endogeneity problem is caused in the induced regression $y_1 = y_2\theta^* + X_1\gamma_1 + u$, for $u = \widetilde{u} + X_2\gamma_2$ and $\gamma = (\gamma_1', \gamma_2')'$. For a nonlinear example, consider Hansen and Singleton's (1982) model of stock prices. The unknown "true parameter vector" of interest consists of the marginal utility of consumption and the discount rate.

In these endogenous scenarios, OLS inference is inconsistent for the true parameter θ^* and instrumental variables (IVs) are needed to do inference on θ^*. Consider the linear IV model

$$y_1 = y_2\theta^* + u,$$
$$y_2 = Z\pi + v, \tag{2}$$

where $y_1, y_2 \in R^n$ are vectors of endogenous variables, $Z \in R^{n\times k}$ for $k \geq 2$ is a matrix of IVs, and $(\theta^*, \pi')' \in R^{1+k}$ are the unknown parameters.[1]

[1] For simplicity, we have excluded exogenous variables $X \in R^{n\times d}$ from the model, $y_1 = y_2\theta^* + X\gamma + u$, $y_2 = Z\pi + X\phi + v$. If there are exogenous variables in (2), then all variables in what follows should be interpreted as the residuals from projection onto the column space of X. With more complicated notation and assumptions we could also consider nonlinear models.

* Tel.: +1 858 534 3201.
E-mail address: pguggenberger@ucsd.edu.

0165-1765/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.econlet.2012.07.008

Please cite this article in press as: Guggenberger, P., A note on the (in)consistency of the test of overidentifying restrictions and the concepts of true and pseudo-true parameters. Economics Letters (2012), doi:10.1016/j.econlet.2012.07.008

Assume that $\{(u_i, v_i, Z_i')' : i \le n\}$ are i.i.d. with distribution F. Assuming fixed $\pi \ne 0$, identification of and inference on the parameter of interest θ^* hinges on the exogeneity assumption $E_F u_i Z_i = 0$, where E_F denotes expectation when the distribution of the random vector $(u_i, v_i, Z_i')'$ is F. If the instruments are nonexogenous,

$$E_F u_i Z_i \ne 0, \tag{3}$$

the two-stage least squares (2SLS) estimator $\widehat{\theta}$ is inconsistent and $\widehat{\theta} \to_p \theta^* + b$, where the "large sample bias" b is defined in (15), see Lemma 1(iii) in the Appendix, and a t test based on 2SLS testing a hypothesis involving θ^* is size distorted. Clearly, whether or not the instruments are exogenous does not affect the value of the true parameter θ^* and its interpretation above, as an expected change in the original linear regression model (1).

Exogeneity of the instruments is often a questionable assumption in applied work, see Guggenberger (2012) for examples. In an effort to prevent wrong inference on θ^* based on nonexogenous instruments, applied researchers typically implement a test of overidentifying restrictions (14), Sargan (1958) and Hansen (1982), in overidentified situations, where $k \ge 2$, prior to implementing a hypothesis test on θ^*. Consider the following null and alternative hypotheses:

$$H_0 : E_F u_i Z_i = 0 \quad \text{against} \quad H_1 : E_F u_i Z_i \ne 0. \tag{4}$$

The first objective of this note is to show that for this formulation of the null and alternative hypotheses, the test of overidentifying restrictions is inconsistent against certain fixed alternatives and to characterize these alternatives; also see Kadane and Anderson (1977) and Small (2007). To do so, we introduce the following notation. For a matrix A with k rows, let I_k be the identity matrix of dimension k, $P_A = A(A'A)^{-1}A'$, and $M_A = I_k - P_A$. Proposition 1 shows that in model (2), such that $M_{(E_F Z_i Z_i')^{1/2}\pi}(E_F Z_i Z_i')^{-1/2} E_F u_i Z_i = 0$ and an additional weak restriction hold, the test statistic to test overidentifying restrictions is $O_p(1)$ and, therefore, the rejection probability of the test of overidentifying restrictions does not converge to 1 as the sample size n goes off to infinity *even if* $E_F u_i Z_i \ne 0$. Therefore, in these cases, no matter how large the sample size n, the overidentifying restrictions pretest may not prevent the use of 2SLS inference methods in the second stage even though these methods will suffer from size distortion.

Despite the inconsistency, we recommend the use of a test of overidentifying restrictions as a pretest before testing a hypothesis on θ^*. Rejection of the null hypothesis in (4) by the test of overidentifying restrictions provides valuable information to the applied researcher who is interested in inference on θ^*. It may prevent the researcher from doing inference on θ^* using the – likely nonexogenous – instruments Z. However, nonrejection of the test of overidentifying restrictions should be interpreted carefully by the applied researcher given Proposition 1.[2] In particular, in that case Guggenberger and Kumar (forthcoming) advocate supplementing the result of a two-stage test on θ^* for a hypothesis on θ^*, i.e. a test of overidentifying restrictions followed by a t-test based on 2SLS, by the Anderson and Rubin (1949) test that is shown to be more robust in terms of size distortion to violations of the exogeneity assumption of the instruments if the two-stage test is implemented with same nominal size in the first and second stages. (The AR test is also fully robust to weak instruments, a scenario which is ruled out in the current note by Assumption 1.)

The second aspect of this note is the following discussion of the relation of the above statements about inconsistency of the test of overidentifying restrictions to the recent large literature in econometrics that derives optimality properties of the test of overidentifying restrictions under a different formulation of the null and alternative hypothesis than that in (4).[3] The original test of overidentifying restrictions (and its various modifications, for instance, versions based on empirical likelihood methods) *is* consistent against fixed alternatives if one considers the following formulation of the null and alternative hypotheses, namely

$$H_0 : F \in \mathcal{F}_0 \quad \text{versus} \quad H_1 : F \in \mathcal{F}_1 = \mathcal{F} \setminus \mathcal{F}_0, \tag{5}$$

where $W_i \in W$ for $i = 1, \ldots, n$ is an i.i.d. sequence of random variables with distribution $F \in \mathcal{F}$, $\mathcal{F}$ denotes a set of (properly restricted) distributions, $g : W \times \Theta \to R^k$ is a known (possibly nonlinear) measurable function, and

$$\mathcal{F}_0 = \{F \in \mathcal{F}; \exists \theta \in \Theta \text{ s.t. } E_F g(W_i, \theta) = 0\}. \tag{6}$$

In the context of the current paper, we can take $W_i = (u_i, v_i, Z_i')'$, $\mathcal{F}$ is the set of distributions for W_i restricted by Assumption 1, and

$$\begin{aligned} g(W_i, \theta) &\equiv (y_{1i} - y_{2i}\theta) Z_i \\ &= Z_i Z_i' \pi(\theta^* - \theta) + Z_i V_i(\theta^* - \theta) + Z_i u_i, \end{aligned} \tag{7}$$

where (2) is used for the second equality. We do not index $g(W_i, \theta)$ by π because π is assumed fixed. Under Assumption 1, we have $E_F g(W_i, \theta) = E_F Z_i Z_i' \pi(\theta^* - \theta) + E_F Z_i u_i$. Thus, for $\Theta = R$, the null in (5) holds iff $E_F Z_i Z_i' \pi$ and $E_F Z_i u_i$ are colinear, or, equivalently, if $M_{(E_F Z_i Z_i')^{1/2}\pi}(E_F Z_i Z_i')^{-1/2} E_F u_i Z_i = 0$. To relate (5) to (4), note that (4) can be equivalently formulated as follows[4]:

$$H_0 : F \in \mathcal{F}_0(\theta^*) \quad \text{versus} \quad H_1 : F \in \mathcal{F}_1(\theta^*) = \mathcal{F} \setminus \mathcal{F}_0(\theta^*), \tag{8}$$

where

$$\mathcal{F}_0(\theta^*) = \{F \in \mathcal{F}; E_F g(W_i, \theta^*) = 0\}. \tag{9}$$

Noting that $\mathcal{F}_0(\theta^*) \subset \mathcal{F}_0$, any test whose limiting null rejection probability does not exceed the nominal size of the test under $F \in \mathcal{F}_0$ does not do so either under $F \in \mathcal{F}_0(\theta^*)$. However, because $\mathcal{F}_1(\theta^*) \supset \mathcal{F}_1$, tests that are consistent against any fixed alternative in $\mathcal{F}_1$ are not necessarily consistent against any fixed alternative in $\mathcal{F}_1(\theta^*)$.

If the test of overidentifying restrictions is used as a pretest and the ultimate goal of the researcher is inference on the true parameter θ^* (as is often the case), then the consistency of the test of overidentifying restrictions for tests of (5) should not lead applied researchers to disregard the above warning about careful interpretation of the test of overidentifying restrictions. In the linear IV model (2), if $E_F u_i Z_i \ne 0$, (16) and (17) hold, easy calculations show that the "*pseudo-true*" *parameter* $\theta^+ = \theta^* + b$ satisfies the moment

[2] Of course, nonrejection by a test should always be interpreted with care because it could be that nonrejection occurs – even though the null is false – because the sample size is too small. However, here the situation is worse because for certain fixed alternatives, the rejection probability of the test does not converge to 1 as the sample size goes off to infinity.

See also the results in Guggenberger and Kumar (forthcoming) about size distortion of a two-stage test involving θ^* when a test of overidentifying restrictions is used as a pretest.

[3] See, among other important contributions, Hansen et al. (1996), Imbens et al. (1998), Kitamura (2001), Otsu (2009), Kitamura et al. (2012) and Canay and Otsu (forthcoming, Section 3.2). Several of these papers are concerned with large deviation optimality properties of tests of (5); for instance, Kitamura (2001) and Kitamura et al. (2012) investigate Hoeffding optimality of empirical likelihood, Otsu (2009) establishes Bahadur efficiency of empirical likelihood based tests, and Canay and Otsu (forthcoming) discuss Hodges–Lehmann efficiency.

[4] I would like to thank the referee for pointing out this connection. The equivalence of (4) and (8) follows directly from noting that $E_F g(W_i, \theta^*) = E_F Z_i u_i$.

Overidentifying restrictions can only be tested if the model is overidentified. Because θ^* is scalar that means we need to assume $k \ge 2$. Note that if $k = 1$, it follows that under Assumption 1, $\mathcal{F}_0 = \mathcal{F}$ and $\mathcal{F}_1 = \varnothing$ in (5), that is, a θ always exists that satisfies $E_F g(W_i, \theta) = 0$.

Please cite this article in press as: Guggenberger, P., A note on the (in)consistency of the test of overidentifying restrictions and the concepts of true and pseudo-true parameters. Economics Letters (2012), doi:10.1016/j.econlet.2012.07.008

condition $Eg(W_i, \theta^+) = 0$; however, obviously θ^+ differs from the "true parameter" θ^* by the – potentially large amount – b, and 2SLS estimation is inconsistent for θ^* and estimates the "pseudo true" parameter.[5] However, the rejection probability of the test of overidentifying restrictions does not converge to 1 in this case. If inference on the true θ^* is the ultimate objective, the optimality properties established for some versions of the test of overidentifying restrictions in the recent literature for alternatives as in (5) therefore have to be carefully interpreted. In particular, even though the various different tests under consideration are inconsistent against certain fixed alternatives $E_F u_i Z_i \neq 0$ as in (17), the tests may differ in their (finite sample) power properties against such alternatives. Also, for the ultimate goal of inference on θ^*, high power of the test of overidentifying restrictions in a pretest stage is particularly important for a parameter constellation where the second stage inference procedure on θ^* is likely to overreject the null hypothesis involving θ^*. In other words, if θ^* is the ultimate object of interest, the optimality properties of a pretest depend on the inference procedure in the second stage. These potentially important power differences of the pretest (for subsequent inference on θ^*) are not taken into account under the various optimality criteria if the hypothesis is formulated as in (5).

2. Theoretical results

In this section it is shown that the (standard) test statistic for overidentifying restrictions is $O_p(1)$ for certain fixed parameter constellations even if $E_F u_i Z_i \neq 0$. We write $g_i(\theta)$ as shorthand for $g(W_i, \theta)$ from now on. Let

$$\widehat{g}(\theta) = n^{-1} \sum_{i=1}^{n} g_i(\theta). \tag{10}$$

The test statistic under conditional homoskedasticity is given by

$$J_n = n\widehat{g}(\widehat{\theta})'(\widehat{\sigma}_u^2 n^{-1} Z'Z)^{-1} \widehat{g}(\widehat{\theta}), \tag{11}$$

where

$$\widehat{\sigma}_u^2 = n^{-1} \sum_{i=1}^{n} \widehat{u}_i^2 \quad \text{for } \widehat{u}_i = y_{1i} - y_{2i}\widehat{\theta} \tag{12}$$

is an estimator for $\sigma_u^2 = E_F u_i^2$ and

$$\widehat{\theta} = (y_2' P_Z y_2)^{-1} y_2' P_Z y_1 \tag{13}$$

is the 2SLS estimator. The test of overidentifying restrictions rejects the null hypothesis in (4) at nominal size α if

$$J_n > \chi^2_{k-1,1-\alpha}, \tag{14}$$

where $\chi^2_{k-1,1-\alpha}$ denotes the $1-\alpha$-quantile of a central chi-square distribution χ^2_{k-1} with $k-1$ degrees of freedom.

Assumption 1. $\pi \neq 0$ and F are fixed, such that $E_F u_i^2 Z_i Z_i' = \sigma_u^2 E_F Z_i Z_i'$, $E_F Z_i Z_i'$ has full rank, $\sigma_u^2 > 0$, $E_F Z_i v_i = 0$, and $\|E_F(\|Z_i u_i\|^2, \|Z_i Z_i'\|^2, \|Z_i v_i\|^2, \|u_i\|^2, \|u_i v_i\|^2, \|Z_i u_i^2 v_i\|)\| < \infty$, where $\|\cdot\|$ denotes the Euclidean norm.

Assumption 1 implies conditional homoskedasticity, strong instruments, and maintains the interpretation of the reduced form equation by assuming $E_F Z_i v_i = 0$. Define

$$b \equiv (\pi' E_F Z_i Z_i' \pi)^{-1} \pi' E_F Z_i u. \tag{15}$$

[5] White (1982) coins the term "minimum ignorance" estimator for the quasi maximum likelihood estimator of the "pseudo true" parameter in an MLE context under misspecification. Also see Chalak and White (2011, Theorem 3.1) for a related result.

Proposition 1. *Under Assumption 1 and assuming that F and π satisfy*

$$\sigma_u^2 - 2(\pi'(E_F Z_i u_i) + E_F v_i u_i)b + (E_F(\pi' Z_i)^2 + \sigma_v^2)b^2 > 0 \quad \text{and} \tag{16}$$

$$M_{(E_F Z_i Z_i')^{1/2}\pi}(E_F Z_i Z_i')^{-1/2} E_F u_i Z_i = 0, \tag{17}$$

the statistic J_n in (11) is $O_p(1)$ as $n \to \infty$, even if $E_F u_i Z_i \neq 0$.

The proposition, whose proof is given in the Appendix, states that, under fixed alternatives $E_F u_i Z_i \neq 0$ satisfying (16) and (17), the statistic J_n does not go off to infinity. For (small enough) significance levels α this then implies that the power of the test of overidentifying restrictions does not go to 1 as the sample size n increases. Condition (17) simply states that $(E_F Z_i Z_i')^{-1/2} E_F u_i Z_i$ and $(E_F Z_i Z_i')^{1/2}\pi$ are colinear. This result supplements Newey (1985, Propostion 1) who shows in nonlinear models that for this type of local alternatives the *local* power of the test of overidentifying restrictions does not exceed the nominal size of the test. Also, see Guggenberger and Kumar (forthcoming) for the discussion of local power of the test of overidentifying restrictions and the resulting consequences of using the test of overidentifying restrictions as a pretest.

Condition (16) is a weak technical condition that guarantees that $(\sigma_u/\widehat{\sigma}_u)$ is $O_p(1)$. It is easy to find distributions F of (u_i, v_i, Z_i) and π such that (16) and (17) and Assumption 1 are satisfied, e.g. see Guggenberger (2012, proof of Theorem 2). Proposition 1 could be generalized to nonlinear models at the expense of more complicated notation.

It is maybe worthwhile to mention that the statistic J_n is in general *not* asymptotically distributed as χ^2_{k-1} under the assumptions of Proposition 1 despite the fact that the pseudo-true parameter θ^+ satisfies the moment condition $E_F g_i(\theta^+) = 0$. The reason is that not all assumptions of Lemma 4.2 in Hansen (1982) are satisfied here. In particular, the weighting matrix $\widehat{\sigma}_u^2 n^{-1} Z'Z$ employed in (11) is not consistent for $E_F g_i(\theta^+) g_i(\theta^+)'$, see Hansen (1982, p. 1049, line 5 ↑). Consider instead a heteroskedasticity robust version of the statistic in (11),

$$J_n^{He} = n\widehat{g}(\widehat{\theta}^{He})'(\widehat{\Omega}_n)^{-1}\widehat{g}(\widehat{\theta}^{He}), \tag{18}$$

where

$$\widehat{\Omega}_n = n^{-1} \sum_{i=1}^{n} \widehat{u}_i^2 Z_i Z_i',$$

$$\widehat{\theta}^{He} = (n^{-1} y_2' Z(\widehat{\Omega}_n)^{-1} n^{-1} Z' y_2)^{-1} n^{-1} y_2' Z(\widehat{\Omega}_n)^{-1} n^{-1} Z' y_1, \tag{19}$$

and $\widehat{u}_i$ is defined in (12). Under (17), $\pi \neq 0$, $E_F Z_i v_i = 0$, and appropriate moment conditions, Lemma 4.2 in Hansen (1982) implies that $J_n^{He} \to_d \chi^2_{k-1}$ even if $E_F Z_i u_i \neq 0$ and without imposing conditional homoskedasticity.

Acknowledgments

I would like to thank a referee and the Editor for very insightful and helpful comments, the NSF under grant SES-1021101 for financial support, the Alfred P. Sloan Foundation for a 2009 fellowship, and Ivan Canay, Bruce Hansen, Yuichi Kitamura, Taisuke Otsu, Eric Renault, Andres Santos, Azeem Shaikh, Hal White, and Michael Wolf for helpful discussions and/or comments. The first version was drafted in 2010.

Appendix

This Appendix provides the proof of Proposition 1. The following lemma is helpful.

Please cite this article in press as: Guggenberger, P., A note on the (in)consistency of the test of overidentifying restrictions and the concepts of true and pseudo-true parameters. Economics Letters (2012), doi:10.1016/j.econlet.2012.07.008

Lemma 1. *Under Assumption 1 we have* (i) $n^{-1}Z'Z = E_F Z_i Z_i' + O_p(n^{-1/2})$, (ii) $n^{-1}y_2' P_Z y_2 = \pi' E_F Z_i Z_i' \pi + O_p(n^{-1/2})$, (iii) $\widehat{\theta} - \theta^* \to_p b$, (iv) $\widehat{\sigma}_u^2 \to_p \sigma_u^2 - 2(\pi'(E_F Z_i u_i) + E_F v_i u_i)b + (E_F(\pi' Z_i)^2 + \sigma_v^2)b^2$, *and, also assuming* (16) *and* (17), (v) $n^{1/2}(\widehat{\sigma}_u^2 n^{-1}Z'Z)^{-1/2}\widehat{g}(\widehat{\theta}) = O_p(1)$.

Note that Proposition 1 follows immediately from part (v) of the lemma because

$$J_n = (n^{1/2}(\widehat{\sigma}_u^2 n^{-1}Z'Z)^{-1/2}\widehat{g}(\widehat{\theta}))' n^{1/2}(\widehat{\sigma}_u^2 n^{-1}Z'Z)^{-1/2}\widehat{g}(\widehat{\theta}).$$

Proof of Lemma 1. Statement (i) follows from $n^{-1}Z'Z = (n^{-1}Z'Z - E_F Z_i Z_i') + E_F Z_i Z_i' = E_F Z_i Z_i' + O_p(n^{-1/2})$ using the central limit theorem (CLT) for the second equality. Statement (ii) follows from

$$\begin{aligned} & n^{-1}y_2' P_Z y_2 \\ &= (\pi' n^{-1}Z'Z + n^{-1}v'Z)(n^{-1}Z'Z)^{-1}(n^{-1}Z'Z\pi + n^{-1}Z'v) \\ &= (\pi' E_F Z_i Z_i' + O_p(n^{-1/2})) \\ &\quad \times (E_F Z_i Z_i' + O_p(n^{-1/2}))^{-1}(E_F Z_i Z_i' \pi + O_p(n^{-1/2})) \\ &= \pi' E_F Z_i Z_i' \pi + O_p(n^{-1/2}) \end{aligned}$$

using (i) and a CLT for $n^{-1}Z'v$. Statement (iii) follows from $\widehat{\theta} - \theta^* = (n^{-1}y_2' P_Z y_2)^{-1} n^{-1} y_2' P_Z u$, (ii), and

$$\begin{aligned} n^{-1}y_2' P_Z u &= (\pi' E_F Z_i Z_i' + O_p(n^{-1/2})) \\ &\quad \times (E_F Z_i Z_i' + O_p(n^{-1/2}))^{-1}(n^{-1}Z'u), \end{aligned}$$

using the law of large numbers (LLN) $n^{-1}Z'u \to_p E_F Z_i u_i$. Statement (iv) follows from

$$\begin{aligned} \widehat{\sigma}_u^2 &= n^{-1}\sum_{i=1}^n (y_{1i} - y_{2i}\widehat{\theta})^2 \\ &= n^{-1}\sum_{i=1}^n (u_i - y_{2i}(\widehat{\theta} - \theta^*))^2 \\ &= n^{-1}\sum_{i=1}^n (u_i^2 - 2y_{2i}u_i(\widehat{\theta} - \theta^*) + y_{2i}^2(\widehat{\theta} - \theta^*)^2) \\ &\to_p \sigma_u^2 - 2(\pi'(E_F Z_i u_i) + E_F v_i u_i)b + (E_F(\pi' Z_i)^2 + \sigma_v^2)b^2, \quad (20) \end{aligned}$$

using $y_{2i}u_i = \pi' Z_i u_i + v_i u_i$, the LLN, and (iii). For statement (v), note that

$$\begin{aligned} & n^{1/2}(\sigma_u^2 n^{-1}Z'Z)^{-1/2}\widehat{g}(\widehat{\theta}) \\ &= (\sigma_u^2 n^{-1}Z'Z)^{-1/2}[I_k - (y_2' P_Z y_2)^{-1}Z' y_2 \\ &\quad \times y_2' Z(Z'Z)^{-1}] n^{-1/2} Z'u \\ &= [I_k - P_{(E_F Z_i Z_i')^{1/2}\pi} + O_p(n^{-1/2})](\sigma_u^2 n^{-1}Z'Z)^{-1/2} n^{-1/2} Z'u \\ &= [M_{(E_F Z_i Z_i')^{1/2}\pi} + O_p(n^{-1/2})](\varphi_n + m_n), \end{aligned}$$

where

$$\begin{aligned} \varphi_n &= (\sigma_u^2 n^{-1}Z'Z)^{-1/2} n^{-1/2}(Z'u - E_F Z'u), \\ m_n &= (\sigma_u^2 n^{-1}Z'Z)^{-1/2} n^{1/2} E_F u_i Z_i, \end{aligned}$$

and φ_n is $O_p(1)$. Note that

$$\begin{aligned} & [M_{(E_F Z_i Z_i')^{1/2}\pi} + O_p(n^{-1/2})] m_n \\ &= [M_{(E_F Z_i Z_i')^{1/2}\pi} + O_p(n^{-1/2})]((\sigma_u^2 E_F Z_i Z_i')^{-1/2} \\ &\quad + O_p(n^{-1/2})) n^{1/2} E_F u_i Z_i \\ &= [M_{(E_F Z_i Z_i')^{1/2}\pi} + O_p(n^{-1/2})](\sigma_u^2 \\ &\quad \times E_F Z_i Z_i')^{-1/2} n^{1/2} E_F u_i Z_i + O_p(1) \\ &= O_p(1), \end{aligned}$$

where for the last equality we use (17). Clearly, $[M_{(E_F Z_i Z_i')^{1/2}\pi} + O_p(n^{-1/2})]\varphi_n = O_p(1)$ and thus $n^{1/2}(\sigma_u^2 n^{-1}Z'Z)^{-1/2}\widehat{g}(\widehat{\theta}) = O_p(1)$. Thus $n^{1/2}(\widehat{\sigma}_u^2 n^{-1}Z'Z)^{-1/2}\widehat{g}(\widehat{\theta}) = (\sigma_u/\widehat{\sigma}_u)\, n^{1/2}(\sigma_u^2 n^{-1}Z'Z)^{-1/2}\widehat{g}(\widehat{\theta}) = O_p(1)$ by condition (16) and (20). □

References

Anderson, T., Rubin, H., 1949. Estimation of the parameters of a single equation in a complete set of stochastic equations. The Annals of Mathematical Statistics 20, 46–63.

Canay, I., Otsu, T., 2012. Hodges–Lehmann optimality for testing moment conditions. Journal of Econometrics (forthcoming).

Chalak, K., White, H., 2011. An extended class of instrumental variables for the estimation of causal effects. Canadian Journal of Economics 44, 1–51.

Guggenberger, P., 2012. On the asymptotic size distortion of tests when instruments locally violate the exogeneity assumption. Econometric Theory 28, 387–421.

Guggenberger, P., Kumar, G., 2012. On the size distortion of tests after an overidentifying restrictions pretest. Journal of Applied Econometrics (forthcoming).

Hansen, L., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.

Hansen, L., Heaton, J., Yaron, A., 1996. Finite sample properties of some alternative GMM estimators. Journal of Business and Economic Statistics 14, 262–280.

Hansen, L., Singleton, K., 1982. Generalized instrumental variables estimation of nonlinear rational expectations models. Econometrica 50, 1269–1286.

Imbens, G., Spady, R., Johnson, P., 1998. Information theoretic approaches to inference in moment condition models. Econometrica 66, 333–357.

Kadane, J., Anderson, T., 1977. A comment on the test of overidentifying restrictions. Econometrica 45, 1027–1032.

Kitamura, Y., 2001. Asymptotic optimality of empirical likelihood for testing moment restrictions. Econometrica 69, 1661–1672.

Kitamura, Y., Santos, A., Shaikh, A., 2012. On the asymptotic optimality of empirical likelihood for testing moment restrictions. Econometrica 80, 413–423.

Newey, W., 1985. Generalized methods of moments specification testing. Journal of Econometrics 29, 229–256.

Otsu, T., 2009. On Bahadur efficiency of empirical likelihood. Journal of Econometrics 157, 248–256.

Sargan, J., 1958. The estimation of economic relationships using instrumental variables. Econometrica 26, 393–415.

Small, D., 2007. Sensitivity analysis for instrumental variables regression with overidentifying restrictions. Journal of the American Statistical Association 102, 1049–1058.

White, H., 1982. Maximum likelihood estimation of misspecified models. Econometrica 50, 1–25.

Please cite this article in press as: Guggenberger, P., A note on the (in)consistency of the test of overidentifying restrictions and the concepts of true and pseudo-true parameters. Economics Letters (2012), doi:10.1016/j.econlet.2012.07.008