Some Properties of Tests for Possibly Unidentified Parameters

by

G. Forchini[1 2]

Monash University

September 2005

Abstract

It is well known that confidence intervals for weakly identified parameters are unbounded with positive probability (e.g. Dufour, *Econometrica* 65, pp. 1365-1387 and Staiger and Stock, *Econometrica* 65, pp. 557-586), and that the asymptotic risk of their estimators is unbounded (Pötscher, *Econometrica* 70, pp.1035-1065). In this note we extend these "impossibility results" and show that uniformly consistent tests for weakly identified parameters do not exist. We also show that all similar tests of size $\alpha < 1/2$ concerning possibly unidentified parameters have type II error probability that can be as large as $1-\alpha$.

JEL Classification: C12, C30

Key Words: similar tests, consistent tests, weak instruments, identification

---

[1] Address for correspondence: Giovanni Forchini, Department of Econometrics and Business Statistics, Monash University, Clayton, Victoria 3800, Australia.
E-mail: Giovanni.Forchini@BusEco.monash.edu.au

[2] I thank Grant Hillier, Don Poskitt and Richard Smith for useful and encouraging comments and discussions.

1. Introduction

Weak instruments have been the focus of much econometric research since the publication of the seminal paper by Staiger and Stock (1997). The topic has important practical applications because instruments have been found to be weak in much empirical research.

Weak instruments put forward a series of challenging non-standard inferential problems for theorists, too. In estimation theory, one of these challenging problems is that of measuring the *precision* of estimators of the weakly identified parameters. Dufour (1997) has shown that every confidence set of level α must be unbounded with positive probability. This suggests that the use of confidence intervals to measure precision of an estimator may be problematic (even when the sample size is very large). Staiger and Stock (1997) and Stock and Wright (2000) have shown that standard estimators of weakly identified parameters are not consistent. Moreover, no uniformly consistent estimator for weakly identified parameters exists (e.g. Pötscher (2002)).

Another challenging problem is the construction of tests which are, at least asymptotically, similar. Dufour (1997) shows that tests based on Wald confidence sets cannot be similar, and that the size of such tests cannot be bounded from above in a nontrivial way. Kleibergen (2002) and Moreira (2003) suggest similar tests for linear structural equations models, and Kleibergen (2005) generalizes these results to a GMM framework. Guggenberger and Smith (2005) construct generalized empirical likelihood tests, and find that they have good size properties under conditional heteroskedasticity. One common characteristic of the tests of Kleibergen (2002) and (2005), Moreira (2003) and Guggenberger and Smith (2005) is that they are inconsistent under weak identification (although only Guggenberger and Smith (2005) mention this property). On the one hand this is intuitive because data is not very informative about the weakly identified parameters (e.g. Guggenberger and Smith (2005)). On the other hand, one wonders whether tests procedures are affected by "impossibility results" analogous to those identified by Dufour (1997) and Pötscher (2002) for confidence intervals and point estimators.

In this paper we investigate the properties of tests on parameters that are possibly unidentified. With this expression we mean that the parameters can be unidentified or arbitrarily closed to being unidentified so that the weak instruments set-up is just a special case. We contribute two results to the existing literature. Firstly, we show that no uniformly consistent test exists for possibly unidentified parameters. Therefore, it is impossible to discriminate between the null and the alternative hypotheses even when the sample size is very large. Secondly, we prove that no test for which the size is bounded by a constant, α say, has power which is uniformly larger than α under the alternative hypothesis, and, thus, the probability of a type II error could be as large as $1-\alpha$.

The paper focuses on the limitations of tests for possibly unidentified parameters, and complements the results of Dufour (1997) and Pötscher (2002) on the properties of confidence intervals and point estimators in the sense that it gives "impossibility results" originating from the discontinuity of the functional defining the interest parameter. On the positive side, our work helps to understand what optimal properties a good test can be expected to have in these situations.

The remaining part of the paper is structured as follows. Section 2 explains the notation and gives some preliminary results showing that the problems associated with tests for possibly unidentified parameters arise from the discontinuity of the functional defining the interest parameters. These results are strongly related to the work of Pfanzagl (1998). Section 3 presents the main theorem about the non existence of uniformly consistent estimator and the properties of tests of bounded size for potentially unidentified parameters. Finally, Section 4 concludes.

2. Notation and preliminary results

We follow the notation of Pfanzagl (1998). Let $\mathfrak{P}$ be a family of probability measures on a measurable space $\left(X^n, \mathcal{A}^n\right)$ where n denotes the sample size. No assumption about the absolute continuity of the probability measures in $\mathfrak{P}$ is required (cf. Dufour (1997)). For the sake of simplicity, we assume that all "observations" take on values in the same set X with measurable sets in the same σ-algebra $\mathcal{A}$, even though the probability measures in $\mathfrak{P}$ could be defined on more general measurable spaces. The analysis is not restricted to the i.i.d. case.

For any two probability measures P and Q in $\mathfrak{P}$ define the total variation distance between them as $d(P,Q)=\sup\left\{\left|P(A)-Q(A)\right|: A\in\mathcal{A}^n\right\}$, and let $\kappa:\mathfrak{P}\to\mathbb{R}$ be the functional defining the interest parameter.

We now single out a probability measure P_0 on $\left(X^n,\mathcal{A}^n\right)$, which may or may not belong to $\mathfrak{P}$, and denote by $\mathfrak{P}_\varepsilon(P_0)$ a neighbourhood of P_0 in $\mathfrak{P}$, i.e.

$$\mathfrak{P}_\varepsilon(P_0)=\left\{P\in\mathfrak{P}: d(P,P_0)<\varepsilon\right\}.$$

Moreover, let

$$N_\varepsilon^\kappa(P_0)=\kappa\left(\mathfrak{P}_\varepsilon(P_0)\right)$$

and

$$N^\kappa(P_0)=\bigcap_{\varepsilon>0}N_\varepsilon^\kappa(P_0).$$

We wish to test the null hypothesis $H_0:\kappa(P)\in\mathcal{H}_0$ against the alternative $H_1:\kappa(P)\in\mathcal{H}_1$ where $\mathcal{H}_0$ and $\mathcal{H}_1$ are two disjoint ($\mathcal{H}_0\cap\mathcal{H}_1=\varnothing$) subsets of $\mathbb{R}$. Let $\hat{\varphi}_n(x)$ be a sequence of tests of the null hypothesis $H_0:\kappa(P)\in\mathcal{H}_0$ against the alternative $H_1:\kappa(P)\in\mathcal{H}_1$, that is $\hat{\varphi}_n(x)$ is a sequence of measurable functions $\hat{\varphi}_n(x)$ from X^n to $[0,1]$.

The sequence of tests $\hat{\varphi}_n(x)$ is uniformly consistent if $\hat{\varphi}_n(x)$ is a uniformly consistent estimator of the functional $\varphi(P)=0$ on $\mathcal{P}_0=\kappa^{-1}(\mathcal{H}_0)$ and $\varphi(P)=1$ on $\mathcal{P}_1=\kappa^{-1}(\mathcal{H}_1)$, and the convergence is uniform in P.

Since

$$\varphi(P)=\overline{\varphi}\left(\kappa(P)\right)=\begin{cases}0 \text{ if } \kappa(P)\in\mathcal{H}_0\\ 1 \text{ if } \kappa(P)\in\mathcal{H}_1\end{cases},$$

the properties of the functional $\kappa(P)$ are fundamental for the existence of uniformly consistent tests. The function $\overline{\varphi}$ is uniformly continuous on $\mathcal{P}_i$, $i=0,1$, so if κ is uniformly continuous then φ is also uniformly continuous on $\mathcal{P}_i$, $i=0,1$. One would expect from the work of Dufour (1997), Pfanzagl (1998) and Pötscher (2002), that problems arise when κ is not uniformly continuous.

Let $C_n(x) \subseteq [0,1]$ be a confidence interval for $\varphi(P)$. Such a confidence interval is *trivial* if $C_n(x) = [0,1]$ (and has coverage probability 1) or if $C_n(x) \subseteq (0,1)$ (and has coverage probability zero). A *nontrivial* confidence interval for $\varphi(P)$ must contain either 0 or 1 but not both. Let

$$N_\varepsilon^\kappa(P_0) \setminus \mathcal{P}_i = \kappa\left(\mathfrak{P}_\varepsilon(P_0) \cap \mathcal{P}_i\right)$$

and

$$N^\kappa(P_0) \setminus \mathcal{P}_i = \bigcap_{\varepsilon > 0}\left(N_\varepsilon^\kappa(P_0) \setminus \mathcal{P}_i\right)$$

for *i=0,1*. We will show that there is no uniformly consistent test for $H_0 : \kappa(P) \in \mathcal{H}_0$ against the alternative $H_1 : \kappa(P) \in \mathcal{H}_1$, where $\mathcal{H}_0$ and $\mathcal{H}_1$ are any two disjoint subsets of $\mathbb{R}$ by proving that a nontrivial confidence interval for $\varphi(P)$, has zero coverage probability when the functional κ is discontinuous. This is achieved with the help of two intermediate results.

Lemma 1. If $C_n(x)$ is *nontrivial* and $N^\varphi(P_0) \setminus \mathcal{P}_i = \{0,1\}$ then

$$\sup_{\varepsilon > 0} \inf_{P \in \mathfrak{P}_\varepsilon(P_0) \cap \mathcal{P}_i} P\left\{x \in X^n : \varphi(P) \in C_n(x)\right\} = 0.$$

Lemma 2. If $\mathcal{H}_j \subseteq N^\kappa(P_0) \setminus \mathcal{P}_i$ then $N^\varphi(P_0) \setminus \mathcal{P}_i = \{0,1\}$ for $i, j = 0,1$ $i \neq j$.

Lemma 1 shows that the coverage probability of any nontrivial confidence interval for $\varphi(P)$ is zero if there is a probability measure P_0 in a neighbourhood of which $\varphi(P)$ can take on both the value zero and one. Lemma 2, shows that this condition occurs whenever the parameter being tested is discontinuous at P_0 and the image of all sufficiently small neighbourhoods of P_0 contains both $\mathcal{H}_0$ and $\mathcal{H}_1$. Combining Lemmas 1 and 2 we have the following result.

Theorem 1. Let $\hat{\varphi}_n(x)$ be an estimator of $\varphi(P)$, and $0 < t < 1$. Then, if $\mathcal{H}_1 \subseteq N^\kappa(P_0) \setminus \mathcal{P}_0$,

$$\sup_{\varepsilon>0} \inf_{P\in\mathfrak{P}_\varepsilon(P_0)\cap\mathcal{P}_0} P\left\{x\in X^n : 0\le \hat{\varphi}_n(x) < t\right\} = 0\,,$$

and if $\mathcal{H}_0 \subseteq N^\kappa(P_0)\setminus\mathcal{P}_1$

$$\sup_{\varepsilon>0} \inf_{P\in\mathfrak{P}_\varepsilon(P_0)\cap\mathcal{P}_1} P\left\{x\in X^n : 0\le 1-\hat{\varphi}_n(x) < t\right\} = 0\,.$$

Theorem 1 applies for any sample size and thus it holds for *n* tending to infinity. It rules out the existence of uniformly consistent tests for situations where $\mathcal{H}_1 \subseteq N^\kappa(P_0)\setminus\mathcal{P}_0$ or $\mathcal{H}_0 \subseteq N^\kappa(P_0)\setminus\mathcal{P}_1$. This is certainly the case when the parameters are possibly unidentified as we will see in the next section.

Corollary 1. Let $\mathcal{H}_0$ and $\mathcal{H}_1$ be any two disjoint subsets of $\mathbb{R}$. If either $\mathcal{H}_1 \subseteq N^\kappa(P_0)\setminus\mathcal{P}_0$ or $\mathcal{H}_0 \subseteq N^\kappa(P_0)\setminus\mathcal{P}_1$ then no uniformly consistent test of $H_0 : \kappa(P)\in\mathcal{H}_0$ against the alternative $H_1 : \kappa(P)\in\mathcal{H}_1$ exists.

Theorem 1 and its corollary clearly characterize the problem affecting tests about discontinuous functionals: it originates from the *closeness* to a probability measure P_0 where the functional of interest can take on both values under the null and under the alternative hypotheses. This property continues to hold in any neighbourhood of P_0, even if we exclude P_0 from our family of probability measures $\mathfrak{P}$.

We now weaken our optimality requirements and focus on tests for which the size is fixed and known (similar tests) or can be bounded from above by a known constant. These are non-similar tests in the sense of Lehmann and Stein (1948). The starting point for the analysis is the following result.

Lemma 3. Let P_0 be a probability measure on $\left(X^n,\mathcal{A}^n\right)$ for which $\mathcal{H}_1 \subseteq N^\kappa(P_0)\setminus\mathcal{P}_0$ and $\mathcal{H}_0 \subseteq N^\kappa(P_0)\setminus\mathcal{P}_1$ (i.e. $N^\varphi(P_0)\setminus\mathcal{P}_i = \{0,1\}$, *i=0,1*). Then, for any estimator $\hat{\varphi}_n(x)$ of $\varphi(P)$ taking on only the values zero or one,

$$\sup_{\varepsilon>0} \inf_{P\in\mathfrak{N}_\varepsilon(P_0)\cap\mathcal{P}_i} P\left\{x\in X^n : \hat{\varphi}_n(x)=\varphi(P)\right\} \le \min\left\{P_0\left\{x\in X^n : \hat{\varphi}_n(x)=0\right\}, P_0\left\{x\in X^n : \hat{\varphi}_n(x)=1\right\}\right\}$$

The characteristic of Lemma 3 is that the left hand side depends on $\mathcal{P}_i$ but the right hand side does not. Therefore, by finding an upper bound for the right hand side we can bound the quantity $\sup_{\varepsilon>0} \inf_{P\in\mathfrak{N}_\varepsilon(P_0)\cap\mathcal{P}_i} P\left\{x\in X^n : \hat{\varphi}_n(x)=\varphi(P)\right\}$ under both the null and the alternative hypotheses.

Theorem 2. Let $\hat{\varphi}_n$ be a test for $H_0 : \kappa\in\mathcal{H}_0$ against the alternative $H_1 : \kappa\in\mathcal{H}_1$ where $\mathcal{H}_0$ and $\mathcal{H}_1$ are any two disjoint subsets of $\mathbb{R}$ having the property that $\sup_{P\in\mathcal{P}_0} P\left\{x\in X^n : \hat{\varphi}_n(x)=1\right\}\le\alpha$ and $0<\alpha<1/2$. Let P_0 be a probability measure on $\left(X^n,\mathcal{A}^n\right)$ for which $\mathcal{H}_1\subseteq N^\kappa(P_0)\setminus\mathcal{P}_0$ and $\mathcal{H}_0\subseteq N^\kappa(P_0)\setminus\mathcal{P}_1$. Then

(i) $\max\left\{P_0\left\{x\in X^n : \hat{\varphi}_n(x)=0\right\}, P_0\left\{x\in X^n : \hat{\varphi}_n(x)=1\right\}\right\}\le\alpha$;

(ii) the power of the test satisfies $\sup_{\varepsilon>0}\inf_{P\in\mathcal{P}_1} P\left\{x\in X^n : \hat{\varphi}_n(x)=1\right\}\le\alpha$

Corollary 2. If $\mathcal{H}_1\subseteq N^\kappa(P_0)\setminus\mathcal{P}_0$ and $\mathcal{H}_0\subseteq N^\kappa(P_0)\setminus\mathcal{P}_1$, the probability of a type II error for a similar test $\hat{\varphi}_n(x)$ of size $\alpha<1/2$ could be larger than $1-\alpha$:

$$\sup_{P\in\mathcal{P}_1} P\left\{x\in X^n : \hat{\varphi}_n(x)=0\right\}\ge 1-\alpha.$$

Theorem 2 shows two important properties of tests for discontinuous functionals. Firstly, in order to find a test having size which is bounded above by a constant α, we need to make sure that the test has at most size α for the probability measures where the functional of interest is continuous. So, for example, in possibly unidentified structural equations models if a test has size α in the identified case, then it will have size smaller than α in the totally or partially unidentified case.

Secondly, suppose that the size of $\hat{\varphi}_n$ is bounded by α, then, in the worst case scenario, the power of the test is less or equal to the size. Thus, no similar test can be uniformly consistent under the alternative if the functional κ is discontinuous. Moreover, in such situation, no test can be uniformly unbiased in the sense that $\sup_{\varepsilon>0} \inf_{P \in \mathfrak{P}_\varepsilon(P_0)\cap \mathcal{P}_1} P\left\{x \in X^n : \hat{\varphi}_n(x)=1\right\} > \alpha$. The intuition for this result is that at P_0 the null and the alternative hypotheses are indistinguishable because of the discontinuity of the functional κ. One cannot rule out a situation where the probability of a type I error is α and the probability of a type II error is as large as $1-\alpha$.

Pötscher (2002) briefly discusses hypotheses test of trend-stationarity against difference-stationarity and concludes that not test can have power larger than size (remark (ii) p. 1053). This result is a special case of those derived above since in his case the interest functional (i.e. the spectral density at zero) is not continuous at a point P_0 arbitrarily closed to both trend and difference stationary models (see also Faust (1996)).

3. Main result

We now specialise the results of Section 2 to a situation where the interest parameters are possibly unidentified in the sense that they can be arbitrarily closed to being unidentified. Define a general family of probability measures $\mathfrak{P} = \left\{P_{\kappa,\phi} : \kappa \in \mathbb{R}, \phi \in \Phi\right\}$, where Φ is a subset of a Hilbert space. Note that $\mathfrak{P}$ can be either a family of fully parametric models if Φ is an Euclidean space, or a family of semi-parametric models. Define the functional of interest as $\kappa\left(P_{\kappa,\phi}\right) = \kappa$ and suppose that the parameter κ is *identified*, i.e. $P_{\kappa_1,\phi_1} = P_{\kappa_2,\phi_2}$ implies $\kappa_1 = \kappa_2$. Moreover assume that κ is *possibly unidentified* in the sense that there exists $\phi_0 \in \Phi$ and a probability measure P_0 on $\left(X^n, \mathcal{A}^n\right)$, not necessarily in $\mathfrak{P}$, such that

$$\lim_{\phi \to \phi_0} d\left(P_{\kappa,\phi}, P_0\right) = 0$$

for all $\kappa \in \mathbb{R}$. This definition of a possibly unidentified parameter corresponds to the one used by Dufour (1997) and Pfanzagl (1998), and is slightly more general than the notion of weakly identified parameters used by Staiger and Stock (1997), which in our notation would translate as $\phi = \phi_0 + n^{-1/2}c$ where c is a fixed arbitrary element of H, and *n* is the sample size.

The main result of this paper is the following theorem.

Theorem 3. Let the parameter κ be possibly unidentified in the sense defined above. Let $\mathcal{H}_0$ and $\mathcal{H}_1$ be any two disjoint subsets of $\mathbb{R}$. Then

i. There is no uniformly consistent test of the null hypothesis $H_0 : \kappa(P) \in \mathcal{H}_0$ against the alternative hypothesis $H_1 : \kappa(P) \in \mathcal{H}_1$

ii. For any sample size *n*, no similar test of size $\alpha < 1/2$ for the null hypothesis $H_0 : \kappa \in \mathcal{H}_0$ against the alternative hypothesis $H_1 : \kappa \in \mathcal{H}_1$ can have power uniformly larger than α. Moreover, the power of a type II error could be larger than $1-\alpha$.

If $\mathcal{H}_1 \not\subset N_\varepsilon^\kappa(P_0) \setminus \mathcal{P}_0$ and $\mathcal{H}_0 \not\subset N_\varepsilon^\kappa(P_0) \setminus \mathcal{P}_1$ for every measure P_0 (not necessarily in $\mathfrak{P}$), we could identify classes of hypotheses for which uniformly consistent tests would exist. For example, Theorem 1 of LeCam and Schwartz (1960) and the continuity of $\kappa(P_{\kappa,\phi}) = \kappa$ would imply the existence of a uniformly consistent test of $H_0 : \kappa(P) = \kappa_0$ against $H_1 : \kappa(P) \in \mathcal{H}_1$ provided κ_0 is not a cluster point of $\mathcal{H}_1$, i.e. κ under the alternative hypothesis is not arbitrarily close to κ_0 (see also Berger (1951). However, Theorem 3 shows that there is no set of hypotheses $H_0 : \kappa(P) \in \mathcal{H}_0$ and $H_1 : \kappa(P) \in \mathcal{H}_1$ for which uniformly consistent tests exist in a neighbourhood of P_0. Therefore, even asymptotically it is not possible to discriminate between the null and alternative hypotheses concerning possibly unidentified parameters. The sets $\kappa^{-1}(\mathcal{H}_0)$ and $\kappa^{-1}(\mathcal{H}_1)$ have a cluster point in common, so that these two sets cannot be clearly separated.

Dufour (1997) considers tests of $H_0 : \kappa = \kappa_0$ against $H_1 : \kappa \neq \kappa_0$ based on Wald confidence intervals by defining a test as a function $\hat{\psi}_n(x)$ which equals zero if κ belongs to a Wald confidence interval $C_n(x)$ and zero otherwise (i.e. $\hat{\psi}_n(x) = 0$ means that H_0 is accepted, and $\hat{\psi}_n(x) = 1$ that H_0 is rejected). Dufour (1997) shows that the distribution of $\hat{\psi}_n(x)$ depends on the "nuisance parameter" ϕ and that there is no way of bounding the size of the test, $P\left\{\hat{\psi}_n\left(x\right)=1\right\}$, over ϕ. This negative result has encouraged econometricians to look for tests that are not based on the Wald confidence sets.

Dufour (1997) recommends the use of the Anderson-Rubin test because it is similar. Earlier results of Hillier (1987) show that the Anderson-Rubin test is optimal in the sense that it maximises a weighted average power under the assumption of normal errors. Kleibergen (2002) and Moreira (2003) have suggested tests on weakly identified parameters in structural equations models that are similar. Kleibergen (2005) has even suggested tests for parameters without assuming that they are identified. Guggenberger and Smith (2005) have discussed the construction of asymptotically similar tests in the context of generalized empirical likelihood tests. They also investigate the power properties of their test, and find that it is not consistent when instruments are weak.

Our Theorem 3 gives some information about the power of similar tests in possibly unidentified models. It shows that tests for which the size is bounded have very poor power properties. For example, if the size of the test is chosen to be equal to 0.01 the probability of a type II error can be higher than $.99$ both in finite samples and asymptotically. The situation does not improve when one focuses on GMM-type tests, because the cause of the problem (i.e. the proximity to a probability P_0 where the interest parameter is unidentified) is unchanged. This implies that Kleibergen (2005)'s GMM-M test may have a very large type II error probability.

4. Conclusions

In this paper, we have shown that tests on possibly unidentified parameters cannot satisfy some of the conditions which are usually satisfied by tests in a standard set-up. These results complement those of Dufour (1997) and Pötscher (2002) on confidence intervals and point estimators.

We have shown that the problem of testing possibly unidentified parameters is very difficult because (i) it is not possible to discriminate between null and alternative hypothesis even when the sample size is infinitely large; and (ii) any test with size bounded from above by a known constant has potentially very low power and a large type II error. Since models with weakly identified parameters seem to be very frequent in practical applications, our main result, Theorem 3, suggests that standard optimality criteria for tests may be inadequate tools to deal with these situations.

Appendix: Proofs

Proof of Lemma 1. Lemma 2.1 of Pfanzagl (1998) with $\mathfrak{P}_\varepsilon(P_0)$ restricted to $\mathfrak{P}_\varepsilon(P_0)\cap\mathcal{P}_i$ implies that for every probability measure P_0 on $(X^n,\mathcal{A}^n)$

$$\sup_{\varepsilon>0}\inf_{P\in\mathfrak{P}_\varepsilon(P_0)\cap\mathcal{P}_i} P\left\{x\in X^n:\varphi(P)\in C_n(x)\right\}\le\min\left\{P_0\left\{x\in X^n:\inf C_n(x)\le\sup_{\varepsilon>0}\inf\left\{N_\varepsilon^\varphi(P_0)\setminus\mathcal{P}_i\right\}\right\},\right.$$
$$\left.P_0\left\{x\in X^n:\sup C_n(x)\ge\inf_{\varepsilon>0}\sup\left\{N_\varepsilon^\varphi(P_0)\setminus\mathcal{P}_i\right\}\right\}\right\}.$$

The fact that $C_n(x)$ is nontrivial means that (i) when $\inf C_n(x)=0$ then $\sup C_n(x)<1$ (so that $P_0\left\{x\in X^n:\sup C_n(x)\ge 1\right\}=0$) and (ii) when $\sup C_n(x)=1$ then $\inf C_n(x)>0$ (so that $P_0\left\{x\in X:\inf C_n(x)\le 0\right\}=0$.

Proof of Lemma 2. By definition $N_\varepsilon^\kappa(P_0)\setminus\mathcal{P}_i=\kappa\left(\mathfrak{P}_\varepsilon(P_0)\cap\mathcal{P}_i\right)$ for any ε, and

$$\begin{aligned}N_\varepsilon^\varphi(P_0)\setminus\mathcal{P}_i&=\varphi\left(\mathfrak{P}_\varepsilon(P_0)\cap\mathcal{P}_i\right)=\bar{\varphi}\left(\kappa\left(\mathfrak{P}_\varepsilon(P_0)\cap\mathcal{P}_i\right)\right)\\&=\bar{\varphi}\left(N_\varepsilon^\kappa(P_0)\setminus\mathcal{P}_i\right).\end{aligned}$$

Since $\mathcal{H}_j\subseteq N_\varepsilon^\kappa(P_0)\setminus\mathcal{P}_i$, $j\ne i$ for every ε the statement of the lemma follows.

Proof of Theorem 1. The confidence intervals $0\le\hat{\varphi}_n<t$ and $0\le 1-\hat{\varphi}_n<t$ are nontrivial. The result follows from Lemmas 1 and 2.

Proof of Lemma 3. This can be proved along the lines of Lemma 2.1 of Pfanzagl (1998) with $\mathfrak{P}_\varepsilon(P_0)$ restricted to $\mathfrak{P}_\varepsilon(P_0)\cap\mathcal{P}_i$ and a degenerate confidence interval for which $\hat{\varphi}_n(P)$ equals zero or one.

Proof of Theorem 2. Note that

$$\begin{aligned}\alpha &\ge \sup_{P\in\mathcal{P}_0} P\left\{x\in X^n : \hat{\varphi}_n(x)=1\right\}\\ &\ge \inf_{\varepsilon>0} \sup_{P\in\mathfrak{P}_\varepsilon(P_0)\cap\mathcal{P}_0} P\left\{x\in X^n : \hat{\varphi}_n(x)=1\right\}\\ &= 1-\sup_{\varepsilon>0} \inf_{P\in\mathfrak{P}_\varepsilon(P_0)\cap\mathcal{P}_0} P\left\{x\in X^n : \hat{\varphi}_n(x)=0\right\}.\end{aligned}$$

Then Lemma 3 implies that

$$\sup_{\varepsilon>0} \inf_{P\in\mathfrak{P}_\varepsilon(P_0)\cap\mathcal{P}_0} P\left\{x\in X^n : \hat{\varphi}_n(x)=0\right\} \le \min\left\{P_0\left\{x\in X^n : \hat{\varphi}_n(x)=0\right\}, P_0\left\{x\in X^n : \hat{\varphi}_n(x)=1\right\}\right\}$$

so

$$\alpha \ge 1-\min\left\{P_0\left\{x\in X^n : \hat{\varphi}_n(x)=0\right\}, P_0\left\{x\in X^n : \hat{\varphi}_n(x)=1\right\}\right\},$$

and part (i) follows easily. Then, using Lemma 3 and part (i) of the Theorem

$$\begin{aligned}\inf_{P\in\mathcal{P}_1} P\left\{x\in X^n : \hat{\varphi}_n(x)=1\right\} &\le \sup_{\varepsilon>0} \inf_{P\in\mathfrak{P}_\varepsilon(P_0)\cap\mathcal{P}_1} P\left\{x\in X^n : \hat{\varphi}_n(x)=1\right\}\\ &\le \min\left\{P_0\left\{x\in X^n : \hat{\varphi}_n(x)=0\right\}, P_0\left\{x\in X^n : \hat{\varphi}_n(x)=1\right\}\right\}\\ &\le \alpha\end{aligned}$$

and part (ii) is also proved.

Proof of Corollary 2. The probability of a type II error is $P\left\{x\in X : \hat{\varphi}_n(x)=0\right\}$ for $P\in\mathcal{P}_1$. So

$$\begin{aligned}\sup_{P\in\mathcal{P}_1} P\left\{x\in X^n : \hat{\varphi}_n(x)=0\right\} &\ge \inf_{\varepsilon>0} \sup_{P\in\mathfrak{P}_\varepsilon(P_0)\cap\mathcal{P}_1} P\left\{x\in X^n : \hat{\varphi}_n(x)=0\right\}\\ &= 1-\sup_{\varepsilon>0} \inf_{P\in\mathfrak{P}_\varepsilon(P_0)\cap\mathcal{P}_1} P\left\{x\in X^n : \hat{\varphi}_n(x)=1\right\}\\ &\ge 1-\alpha.\end{aligned}$$

and the result is proved.

Proof of Theorem 3. Consider a functional $\kappa\left(P_{\kappa,\phi}\right)=\kappa$, and let

$$\mathfrak{P}_\varepsilon(P_0)=\left\{P_{\kappa,\phi} : \kappa\in\mathbb{R}, \phi\in\Phi, d\left(P_{\kappa,\phi}, P_0\right)<\varepsilon\right\},$$

since $\lim_{\phi\to\phi_0} d\left(P_{\kappa,\phi}, P_0\right)=0$, one has that $N_\varepsilon^\kappa(P_0)\setminus\mathcal{P}_i=\kappa\left(\mathfrak{P}_\varepsilon(P_0)\cap\mathcal{P}_i\right)=\kappa(\mathfrak{P})$ for every $\varepsilon>0$. Part (i) follows from Theorem 1 and part (ii) from Theorem 2.

References

Berger, A. (1951), On Uniformly Consistent Tests, *The Annals of Mathematical Statistics*, 22, 289-293.

Dufour, J.-M. (1997), Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models, *Econometrica* 65, 1365-1387.

Faust, J. (1996), Near Observational Equivalence and Theoretical Size Problems sith Unit Root Tests”, *Econometric Theory* 12, 724-731.

Guggenberger, P. and R.J. Smith (2005), Generalized Empirical Likelihood Estimators and Tests under Partial, Weak and Strong Identification, *Econometric Theory*, Forthcoming.

Hillier, G. H. (1987), Classes of Similar Regions and Their Power Properties for some Eonometric Testing Problems, *Econometric Theory* 3, 1-44.

Kleibergen, F. (2002), Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression, *Econometrica*, 70, 1781-1803.

Kleibergen, F. (2005), Testing Parameters in GMM without Assuming that They Are Identified, *Econometrica*, Forthcoming.

LeCam, L. and L. Schwartz (1960), A Necessary and Sufficient Condition for the Existence of Consistent Estimates, *The Annals of Mathematical Statistics*, 31, 140-150

Lehmann, E.L. and C. Stein (1948), Most Powerful Tests of Composite Hypotheses I. Normal Distributions, *The Annals of Mathematical Statistics*, Vol. 19, 495-516.

Moreira, M. J. (2003), A Conditional Likelihood Ratio Test for Structural Models, *Econometrica*, 71, 1027-1048.

Pfanzagl, J. (1998), The Nonexistence of Confidence Sets for Discontinuous Functionals, *Journal of Statistical Planning and Inference*, 75, 9-20.

Pötscher, B.M. (2002), Lower Risk Bounds and Properties of Confidence Sets for Ill-Posed Estimation Problems with Applications to Spectral Density and Persistence Estimation, Unit Roots, and Estimation of Long Memory Parameters, *Econometrica*, 70, 1035-1065

Staiger, D. and J.H. Stock (1997), Instrumental Variables Regression with Weak Instruments, *Econometrica*, 65, 557-586.

Stock, J.H and J.H. Wright (2000), GMM with Weak Identification, *Econometrica*, 68, 1055-1096.

Some further details about the proofs: not intended for publication

Detailed proof of Lemma 3. Consider an arbitrary $\delta > 0$ and a probability measure $P \in \mathfrak{P}_\delta(P_0) \cap \mathcal{P}_i$. This means that $\sup\left\{\left|P(A) - Q(A)\right| : A \in \mathcal{A}^n\right\} < \delta$ i.e. for $r \in \{0,1\}$

$$P\left\{x \in X^n : \hat{\varphi}_n(x) = r\right\} \le P_0\left\{x \in X^n : \hat{\varphi}_n(x) = r\right\} + \delta .$$

If $r \in N_\varepsilon^\varphi(P_0) \setminus \mathcal{P}_i$ then there is a probability measure $P_r \in \mathfrak{P}_\delta(P_0) \cap \mathcal{P}_i$ such that $\varphi(P) = r$. So by choosing $\varepsilon \le \delta$ we have

$$\begin{aligned}\inf_{P \in \mathfrak{P}_\delta(P_0) \cap \mathcal{P}_i} P\left\{x \in X^n : \hat{\varphi}_n(x) = \varphi(P)\right\} &\le P_r\left\{x \in X^n : \hat{\varphi}_n(x) = r\right\} \\ &\le P_0\left\{x \in X^n : \hat{\varphi}_n(x) = r\right\} + \delta\end{aligned}$$

and this implies that

$$\inf_{P \in \mathfrak{P}_\delta(P_0) \cap \mathcal{P}_i} P\left\{x \in X^n : \hat{\varphi}_n(x) = \varphi(P)\right\} \le \inf_{r \in N_\varepsilon^\varphi(P_0) \setminus \mathcal{P}_i} P_0\left\{x \in X^n : \hat{\varphi}_n(x) = r\right\} + \delta .$$

Since $\delta > 0$ is arbitrary and r can take on only the value zero or one, this implies that

$$\sup_{\varepsilon > 0} \inf_{P \in \mathfrak{P}_\varepsilon(P_0) \cap \mathcal{P}_i} P\left\{x \in X^n : \hat{\varphi}_n(x) = \varphi(P)\right\} \le \min\left\{P_0\left\{x \in X^n : \hat{\varphi}_n(x) = 0\right\}, P_0\left\{x \in X^n : \hat{\varphi}_n(x) = 1\right\}\right\}$$