THE EFFECT OF LINEAR TIME TRENDS ON THE KPSS TEST FOR COINTEGRATION

By Uwe Hassler

Free University of Berlin

First version received January 1999

Abstract. The so-called KPSS test for the null hypothesis of cointegration builds on residuals from single equation regressions. Critical values have been provided for regressions with and without detrending. Here it is shown that the latter are not appropriate if the series display linear trends, although this does not mean that detrending is required. In this paper adequate percentiles are suggested for series that follow linear time trends, and tests are based on regressions without detrending. These percentiles are readily available from the literature.

Keywords. Integrated series with drifts; effect of not detrending.

1. INTRODUCTION

Many non-stationary time series are not only considered as integrated of order 1 but also eventually display an approximately linear time trend after taking logarithms. Nevertheless, many researchers prefer regressions without detrending (see also the discussion in Hassler, 1999). It is common practice when testing for cointegration to use critical values that are designed for regressions without detrending (and are simulated under the assumption of no linear trends). This is not correct because 'the deterministic trends in the data affect the limiting distributions of the test statistics whether or not we detrend the data' (Hansen, 1992, p. 103). With respect to the residual-based Dickey–Fuller test introduced by Engle and Granger (1987) and Phillips and Ouliaris (1990), Hansen (1992) found that the asymptotic effect of linear trends is very small and in fact negligible. Consequently to the best of my knowledge it has never been taken into account in empirical work. With the residual-based KPSS test, however, the effect of linear trends turns out to be far from negligible.

Kwiatkowski, Phillips, Schmidt and Shin (1992) introduced the KPSS test for the null hypothesis that a univariate time series is (trend) stationary. It is a Lagrange multiplier test constructed against the alternative of a unit root (with drift). The idea of applying the KPSS test to residuals from a cointegrating regression in order to test the null hypothesis of cointegration suggests itself, and was pursued by Harris and Inder (1994), Leybourne and McCabe (1994) and Shin (1994). A similar approach was considered by McCabe *et al*. (1997).

However, contrary to the residual-based Dickey–Fuller cointegration test, the KPSS test will in general depend on nuisance parameters if applied to ordinary least squares (OLS) residuals. To circumvent this problem it must be applied to residuals from regressions that are estimated efficiently in the sense of Saikkonen (1991). I refer to Shin (1994) for a rigorous development of the limiting distribution of the residual-based KPSS test. He also provides extensive tables with simulated critical values for the case when the efficient cointegrating regression is detrended or not.

This paper supplements the above-mentioned work on the KPSS test for cointegration. I study the effect of linear time trends on the tests applied without detrending. It turns out that the critical values without detrending are not valid if the time series display linear trends. The use of regressions and critical values without detrending in the presence of linear time trends—which is common practice in econometrics—will lead to size distortions: the null hypothesis of cointegration is rejected far too seldom. However, this does not mean that detrending is required in order to obtain correct critical values. Appropriate critical values in the case of linear trends without detrending are suggested here; these values are readily available from the literature. As discovered by Hansen (1992) for the Dickey–Fuller cointegration test, tests based on n integrated regressors with drifts but without detrending require the critical values for $n - 1$ detrended regressors. In the case of simple regressions, $n = 1$, the test without detrending requires the percentiles from the detrended univariate KPSS test by Kwiatkowski *et al*. (1992).

The outline of the paper is as follows. The next section describes the cointegration model and the assumptions. Section 3 presents the residual-based KPSS test with the main result. Section 4 is reserved for the proof. The final section summarizes and describes in simple words how and where appropriate critical values can be found.

2. MODEL AND ASSUMPTIONS

I briefly present the cointegration model, the underlying assumptions and the notation that has become standard since the pioneering work by Phillips (1986), Phillips and Durlauf (1986) and Park and Phillips (1988). The vector $x_t^{(n)}$ contains n regressors x_{it} integration of order 1, I(1), with drifts:

$$x_t^{(n)} - x_{t-1}^{(n)} = \mu + u_t^{(n)} \qquad \mu \neq 0, \quad t = 1, 2, \ldots, T \tag{1}$$

where μ does not equal the zero vector; this does not require that all components are non-zero. The regressors $x_t^{(n)}$ alone are not cointegrated, but they cointegrate with the scalar y_t,

$$y_t = b' x_t^{(n)} + e_t \qquad b \neq 0, \quad t = 1, 2, \ldots, T. \tag{2}$$

For simplicity no constant term is included in (2). The $(n+1)$-dimensional vector of innovations

$$w_t := \begin{pmatrix} e_t \\ u_t^{(n)} \end{pmatrix} \qquad t = 1, 2, \ldots, T$$

is assumed to be zero mean and, for convenience, stationary with

$$T^{-1} \sum_{t=1}^{T-\tau} w_t w'_{t+\tau} \xrightarrow{\text{p}} \Gamma_\tau = (\Gamma_{\tau,ij})_{i,j=0,1} \qquad \tau = 0, 1, \ldots$$

where $\xrightarrow{\text{p}}$ stands for convergence in probability. Here and in what follows $i, j = 0, 1$ denotes a partitioning conformable with w_t. The covariance matrices are used to define

$$\Lambda = \sum_{\tau=1}^{\infty} \Gamma_\tau = (\Lambda_{ij})_{i,j=0,1}$$

$$\Delta = \Gamma_0 + \Lambda = (\Delta_{ij})_{i,j=0,1}$$

$$\Omega = \Delta + \Lambda' = (\Omega_{ij})_{i,j=0,1}$$

where Ω is proportional to the spectral matrix of w_t at frequency zero. Ω, and in partiuclar Ω_{11}, are positive definite so that $x_t^{(n)}$ is indeed not cointegrated. Furthermore, w_t is assumed to obey the following functional central limit theorem for $0 \leqslant r \leqslant 1$:

$$T^{-0.5} \sum_{j=1}^{[rT]} w_j \Rightarrow B(r) = \begin{pmatrix} B_0(r) \\ B_1(r) \end{pmatrix} = \Omega^{0.5} \begin{pmatrix} W_0(r) \\ W_1(r) \end{pmatrix} \tag{3}$$

where $\Rightarrow$ denotes weak convergence, [.] stands for integer part, and $B(r)$ is a Brownian motion with covariance matrix Ω, i.e., $W'(r) = (W_0(r), W_1'(r))$ is an $(n+1)$-dimensional standard Wiener process with independent components. For reasons that will be explained later, it is useful to suppose that $x_t^{(n)}$ is exogenous in the sense that

$$\Delta_{10} = 0 \qquad \Omega_{10} = 0. \tag{4}$$

This condition implies of course that Ω is block diagonal and hence, by (3),

$$B_0(r) = \Omega_{00}^{0.5} W_0(r) \qquad B_1(r) = \Omega_{11}^{0.5} W_1(r) \tag{5}$$

so that $B_0(r)$ is stochastically independent of $B_1(r)$.

The invariance principle (3) implies

$$x_t^{(n)} = x_0^{(n)} + \mu t + \sum_{j=1}^{t} u_j^{(n)}$$

$$= \text{O}(1) + \text{O}(T) + \text{O}_\text{p}(T^{0.5}).$$

Assume for the moment that $n = 1$. In this case the linear trend clearly dominates the stochastic trend, and limiting distributions will not depend on $u_t^{(1)}$. For $n > 1$ the series $x_t^{(n)}$ are driven by one and the same common linear time trend t. At the same time $x_t^{(n)}$ is not cointegrated and hence is driven by n common stochastic trends $\sum_{j=1}^{t} u_j^{(n)}$. However, the linear trend dominates one stochastic trend, and that is why $x_t^{(n)}$ from (1) behaves like $n - 1$ stochastic trends plus one linear trend. This is the reason why, asymptotically, the regression on $x_t^{(n)}$ without detrending provides limiting distributions identical to those arising from detrended regressions on $n - 1$ integrated regressors only. This is the intuition behind the findings of Hansen (1992) and the proposition below.

As has been intimated, we have to consider a second cointegration model with only $n - 1$ integrated regressors

$$\xi_t^{(n-1)} - \xi_{t-1}^{(n-1)} = m + v_t^{(n-1)} \tag{1'}$$

that are not cointegrated but cointegrate with the scalar z_t,

$$z_t = (\beta_1, \ldots, \beta_{n-1})\xi_t^{(n-1)} + \epsilon_t \qquad (\beta_1, \ldots, \beta_{n-1})' \neq 0 \tag{2'}$$

where ϵ_t is again stationary. It is assumed that the n-dimensional vector $(\epsilon_t, v_t^{(n-1)})'$ has analogous properties to w_t. In particular, analogous results to (3), (4) and (5) are assumed but not stated explicitly in order to save space.

3. RESIDUAL-BASED KPSS TESTS

Without (4) the KPSS test based on OLS residuals is plagued by the nuisance parameters Δ_{10} and Ω_{10}. Therefore, the violation of (4) requires efficient modifications of OLS to obtain asymptotically valid cointegration tests. Efficient estimators correcting for the influence of Δ_{10} and Ω_{10} result in mixed normal distributions being most concentrated in the sense of Saikkonen (1991, p. 6), and allow for asymptotic standard normal inference by means of appropriately modified t statistics. Efficient single equation estimators have been suggested by Phillips and Hansen (1990) and Hansen (1992) (fully modified OLS), by Saikkonen (1991) and Stock and Watson (1993) (dynamical OLS) and by Park (1992) (canonical cointegrating regression). Finally, condition (4) renders the OLS regression of (2) efficient. Instead of considering a generic efficient estimation of (2) one may just as well investigate OLS and assume (4), and that is how I shall proceed.

Consider the estimation of (2),

$$\begin{aligned} y_t &= \hat{b}' x_t^{(n)} + \hat{e}_t \\ &= \hat{b}_1 x_{1t} + \cdots + \hat{b}_n x_{nt} + \hat{e}_t \end{aligned} \tag{6}$$

as a detrended cointegrating regression of (2′) with only $n-1$ integrated regressors,

$$z_t = \hat{\beta}'\xi_t + \hat{\epsilon}_t$$
$$= \hat{\beta}_n t + (\hat{\beta}_1, \ldots, \hat{\beta}_{n-1})\xi_t^{(n-1)} + \hat{\epsilon}_t \tag{6′}$$

where $\hat{\beta}$ and ξ_t have been defined implicitly. For $n = 1$, let $\xi_t^{(0)}$ mean that there is no integrated regressor entering (2′) or (6′), so that (6′) simply detrends the stationary process z_t given in (2′). The derivation of the result below draws heavily on Hansen (1992) and Park (1992). Like these authors I neglect an intercept in (6) and (6′) that would change the limiting results only trivially.

Following Shin (1994) one computes the partial sum process $S_t = \sum_{j=1}^{t}\hat{e}_j$ or $\sum_{j=1}^{t}\hat{\epsilon}_j$ based on efficient residuals from (6) or (6′) in order to test the null hypothesis of cointegration. The core of the KPSS statistic proposed by Kwiatkowski *et al.* (1992) is $T^{-2}\sum_{t=1}^{T}S_t^2$. This term still has to be normalized with some variance term (see the next section for details). The null hypothesis of no cointegration is rejected for values that are too large. Critical values are provided in Shin (1994, Table 1). The proof of the following result is given in the next section.

PROPOSITION. Given the cointegration model with drifts, (1) and (2), the KPSS statistic based on efficient residuals from a cointegrating regression on the n I(1) regressors without detrending, (6), converges to the limiting distribution of the KPSS statistic from a detrended cointegrating regression on $n-1$ I(1) regressors, (6′).

The proposition can be stated as follows. If any of n regressors from an efficient cointegrating regression without detrending displays a linear time trend, the percentiles from the middle panel of Table 1 in Shin (1994) are not correct. The correct critical values are in the lower panel (detrended regression), however, in the column for $n-1$ given $n > 1$; for $n = 1$ the critical values from the lower panel of Table 1 in Kwiatkowski *et al.* (1992) stemming from the univariate detrended KPSS test are asymptotically correct. For $n = 1$ the wrong 5% critical value (which is the one used in standard practice) is 0.314 and hence more than twice as large as the correct 5% percentile 0.146 according to the proposition. With increasing n the difference between the corresponding percentiles decreases. But in any case the standard practice values are considerably larger than the correct values. As the KPSS test rejects for too large values, the standard use of the wrong critical values in the case of regressions without detrending will reject the null hypothesis of cointegration far too seldom.

McCabe, Leybourne and Shin (1997) suggested a modification of the residual-based test for the null of cointegration. Their test is affected by the presence of linear time trends in exactly the same way as the KPSS test because it also builds on the cumulation of a squared partial sum residual process. The two tests differ only with respect to the treatment of residual autocorrelation. The McCabe–Leybourne–Shin test applies a parametric procedure wheres the KPSS test relies on a spectral estimator ($\hat{\omega}^2$ in the next section).

4. PROOF

In order to prove the proposition, the limiting distributions of the regressions (6) and (6′) are required. Let us start with the efficient regression without detrending, i.e. with the estimation of (2). M is the $n \times (n-1)$ matrix of full rank spanning the null space of μ from (1):

$$M'\mu = 0 \qquad n > 1.$$

Due to the multicollinearity of $x_t^{(n)}$ for $n > 1$, a transformation is necessary that concentrates the linear trends of $x_t^{(n)}$ in one component:

$$C := (\mu(\mu'\mu)^{-1},\ M(M'\Omega_{11}M)^{-0.5}). \tag{7}$$

In the case of simple regressions, $n = 1$, μ cannot be eliminated and C is understood to be scalar, $C = 1/\mu$. Upon transformation $C'x_t^{(n)}$ requires the following diagonal weighting matrix:

$$D_T = \mathrm{diag}(T^{1.5},\ TI_{n-1}) \tag{8}$$

where I_{n-1} is the $(n-1)$-dimensional identity matrix for $n > 1$; if $n = 1$, D_T reduces simply to $T^{1.5}$. The transformed regressors obey the following functional central limit theorem for $s\epsilon[0, 1]$:

$$T^{0.5}D_T^{-1}C'x_{[sT]}^{(n)} \Rightarrow \begin{pmatrix} s \\ W_{(n-1)}(s) \end{pmatrix} \tag{9}$$

where

$$W_{(n-1)}(s) := (M'\Omega_{11}M)^{-0.5}M'\Omega_{11}^{0.5}W_1(s)$$

is an $(n-1)$-dimensional standard Wiener process independent of $B_0(s)$ by assumption (4). If $n = 1$, $W_{(0)}(s)$ means that there is no Wiener process entering the scalar right-hand side of (9). Now, it can be proved (see Hansen 1992, theorem 1; Park, 1992, Lemma 3.1) that

$$D_T C^{-1}(\hat{b} - b) \Rightarrow \Psi_n^{-1}\int_0^1 \begin{pmatrix} r \\ W_{(n-1)}(r) \end{pmatrix} dB_0(r) =: \overset{\infty}{b}_n \tag{10}$$

where

$$\Psi_n = \int_0^1 \begin{pmatrix} r^2 & rW_{(n-1)}{}'(r) \\ rW_{(n-1)}(r) & W_{(n-1)}(r)W_{(n-1)}{}'(r) \end{pmatrix} dr.$$

If $n = 1$, this simplifies to $\Psi_1 = 1/3$.

After the treatment of (6) let us now turn to the asymptotic analysis of the detrended regression (6′). The appropriate weighting matrix becomes

$$\overline{D}_T = \begin{pmatrix} T^{1.5} & 0' \\ \mu T^{1.5} & \Omega_{1(n-1)}^{0.5} T \end{pmatrix} = \begin{pmatrix} T^{-1.5} & 0' \\ -\Omega_{1(n-1)}^{-0.5}\mu T^{-1} & \Omega_{1(n-1)}^{-0.5} T^{-1} \end{pmatrix}^{-1}. \qquad (11)$$

In (11), $\Omega_{1(n-1)}$ represents the spectral density matrix of $v_t^{(n-1)}$ from (1′) analogous to Ω_{11} entering (5) and (3):

$$T^{-0.5}\sum_{j=1}^{[rT]} \begin{pmatrix} \epsilon_j \\ v_j^{(n-1)} \end{pmatrix} \Rightarrow \begin{pmatrix} \overline{B}_0(r) \\ \Omega_{1(n-1)}^{0.5}\overline{W}_{(n-1)}(r) \end{pmatrix} \qquad (3')$$

where $\overline{W}_{(n-1)}(r)$ is again an $(n-1)$-dimensional standard Wiener process independent of $\overline{B}_0(r)$. Again, for $n = 1$ we simply have $\overline{D}_T = T^{1.5}$. The functional central limit theorem analogous to (9) is

$$T^{0.5}\overline{D}_T^{-1}\xi_{[sT]} = T^{0.5}\overline{D}_T^{-1}\begin{pmatrix} [sT] \\ \xi_{[sT]}^{(n-1)} \end{pmatrix} \Rightarrow \begin{pmatrix} s \\ \overline{W}_{(n-1)}(s) \end{pmatrix}. \qquad (12)$$

From Hansen (1992, Theorem 1) it follows (for $n > 1$) that[1]

$$\overline{D}_T'(\hat{\beta} - \beta) \Rightarrow \overline{\Psi}_n^{-1}\int_0^1 \begin{pmatrix} r \\ \overline{W}_{(n-1)}(r) \end{pmatrix} d\overline{B}_0(r) =: \overset{\infty}{\beta}_n \qquad (13)$$

where $\overline{\Psi}_n$ is defined like Ψ_n but in terms of $\overline{W}_{(n-1)}$. It is important to note that the limiting distribution in (13) is the same as that from (10), apart from the different variances of B_0 and $\overline{B}_0$. In other words, the limiting theory of the efficient cointegrating regression on n regressors with drifts amounts to the limiting theory of a detrended cointegrating regression on $n - 1$ regressors. This provides the intuition for the proposition stated above.

Now, I turn to the analysis of the KPSS statistic. Given the partial sum process $S_t = \sum_{j=1}^{t}\hat{e}_j$ or $\sum_{j=1}^{t}\hat{\epsilon}_j$ based on efficient residuals from (6) or (6′), the core of the KPSS statistic is $T^{-2}\sum_{t=1}^{T}S_t^2$. This statistic still has to be normalized by $\hat{\omega}^2$, which is a consistent estimator of the variance of $B_0(r)$ or $\overline{B}_0(r)$, respectively. A consistent estimator can be constructed from regression residuals. Given some limit of the normalized partial sum process under the null,

$$(\hat{\omega}^2 T)^{-0.5} S_{[rT]} \Rightarrow S_\infty(r) \qquad (14)$$

say, it then follows by the continuous mapping theorem that

$$(\hat{\omega}^2 T)^{-2}\sum_{t=1}^{T} S_t^2 \Rightarrow \int_0^1 S_\infty^2(r)\,dr.$$

In order to establish the proposition one only has to show that $S_\infty(r)$ in (14) is one and the same irrespective of residuals from (6) or (6′). Let us start with residuals from the regression without detrending:

$$T^{-0.5}\sum_{j=1}^{[rT]}\hat{e}_j = T^{-0.5}(b-\hat{b})'\sum_{j=1}^{[rT]}x_j^{(n)} + T^{-0.5}\sum_{j=1}^{[rT]}e_j$$

$$= (b-\hat{b})'C'^{-1}D_T T^{-0.5}\sum_{j=1}^{[rT]}D_T^{-1}C'x_j^{(n)} + T^{-0.5}\sum_{j=1}^{[rT]}e_j$$

$$\Rightarrow -\overset{\infty}{b}{}'_n\int_0^r\begin{pmatrix} s \\ W_{(n-1)}(s)\end{pmatrix}ds + B_0(r)$$

by (10), (9) and the continuous mapping theorem, where $W_{(0)}(s)$ means that no Wiener process enters the scalar integral. Given a consistent estimator $\hat{\omega}^2$ this implies

$$(\hat{\omega}^2T)^{-0.5}\sum_{j=1}^{[rT]}\hat{e}_j \Rightarrow -\left[\Psi_n^{-1}\int_0^1\begin{pmatrix} s \\ W_{(n-1)}(s)\end{pmatrix}dW_0(s)\right]'\int_0^r\begin{pmatrix} s \\ W_{(n-1)}(s)\end{pmatrix}ds$$

$$+\; W_0(r) =: S_\infty(r)$$

by (5) and (10). With (13) and (12) one analogously shows for the residuals from (6′)

$$T^{-0.5}\sum_{j=1}^{[rT]}\hat{\epsilon}_j = (\beta-\hat{\beta})'\overline{D}_T T^{-0.5}\sum_{j=1}^{[rT]}\overline{D}_T^{-1}\xi_j + T^{-0.5}\sum_{j=1}^{[rT]}\epsilon_j$$

$$\Rightarrow -\overset{\infty}{\beta}{}'_n\int_0^r\begin{pmatrix} s \\ \overline{W}_{(n-1)}(s)\end{pmatrix}ds + \overline{B}_0(r).$$

Again, $\overline{W}_{(n-1)}$ is an $(n-1)$-dimensional standard Brownian motion independent of $\overline{B}_0$. Therefore, given a consistent estimator $\hat{\omega}^2$ of the variance of $\overline{B}_0$ we have

$$(\hat{\omega}^2T)^{-0.5}\sum_{j=1}^{[rT]}\hat{\epsilon}_j \Rightarrow S_\infty(r)$$

where the limit is the same as above, which amounts to a proof of the proposition.

5. SUMMARY

This paper deals with the KPSS test for the null hypothesis that y_t and the n-dimensional vector $x_t^{(n)}$ are cointegrated. The test builds on efficient residuals of a single equation regression. Critical values have been provided by Shin (1994) for two versions, one based on a detrended regression, and the other on a regression without detrending. The percentiles for the latter are simulated under the assumption that the series of interest are integrated without drifts, i.e. they are not designed for variables with linear time trends. Nevertheless those critical values are applied in practice with time series with linear trends entering a single equation regression without detrending.

The purpose of this paper is twofold. First, it is shown that the standard practice of using critical values without detrending in the presence of linear trends implies rejection rates deviating from the alleged nominal level. The KPSS test rejects too seldom, indicating cointegration too often. However, detrending is not the only route to correct percentiles, because second, I suggest appropriate critical values for the case of linear trends without detrending. Fortunately, they are readily available from the literature, and in particular from Shin (1994) for $n > 1$. With n I(1) variables x_t with drift but without detrending the adequate percentiles are from Shin's Table 1 with detrending for $n - 1$ x variables; if $n = 1$ the percentiles from the univariate detrended KPSS test are correct.

NOTE

1. For $n = 1$ it is elementary to verify (13) because (6′) reduces to detrending a stationary process.

REFERENCES

Engle, R. F. and Granger, C. W. J. (1987) Co-integration and error correction: representation, estimation, and testing. *Econometrica*, 55, 251–76.

Hansen, B. E. (1992) Efficient estimation and testing of cointegrating vectors in the presence of deterministic trends. *J. Economet*. 53, 87–121.

Harris, D. and Inder, B. (1994) A test of the null hypothesis of cointegration. In *Nonstationary Time Series Analysis and Cointegration* (ed. C. P. Hargreaves). Oxford: Oxford University.

Hassler, U. (1999) (When) should cointegration regressions be detrended? The case of a German money demand function. *Empirical Econ*. 24, 155–72.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1992) Testing the null hypothesis of stationarity against the alternative of a unit root. *J. Econ*. 54, 159–78.

Leybourne, S. J. and McCabe, B. P. M. (1994) A simple test for cointegration. *Oxford Bull. Econ. Stat.* 56, 97–103.

McCabe, B. P. M., Leybourne, S. J. and Shin, Y. (1997) A parametric approach to testing the null of cointegration. *J. Time Ser. Anal*. 18, 395–413.

Park, J. Y. (1992) Canonical cointegrating regressions. *Econometrica* 60, 119–43.

—— and Phillips, P. C. B. (1988) Statistical inference in regressions with integrated processes: Part I. *Economet. Theory* 4, 468–97.

Phillips, P. C. B. (1986) Understanding spurious regressions in econometrics. *J. Economet*. 33, 311–40.

—— and Durlauf, S. N. (1986) Multiple time series regressions with integrated processes. *Rev. Econ. Stud*. 53, 473–95.

—— and Hansen, B. E. (1990) Statistical inference in instrumental variables regression with I(1) processes. *Rev. Econ. Stud*. 57, 99–125.

—— and OULIARIS, S. (1990) Asymptotic properties of residual based tests for cointegration. *Econometrica* 58, 165–93.

SAIKKONEN, P. (1991) Asymptotic efficient estiamtion of cointegration regressions. *Economet. Theory* 7, 1–21.

SHIN, Y. (1994) A residual-based test of the null of cointegration against the alternative of no cointegration. *Economet. Theory* 10, 91–115.

STOCK, J. H. and WATSON, M. W. (1993) A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* 61, 783–820.