Econometric Theory, **27**, 2011, 413–426.
doi:10.1017/S0266466610000290

PIVOTAL STRUCTURAL CHANGE TESTS IN LINEAR SIMULTANEOUS EQUATIONS WITH WEAK IDENTIFICATION

MEHMET CANER
North Carolina State University

Asymptotically pivotal structural change tests are developed for simultaneous equations with weakly identified parameters, extending the boundedly pivotal tests of Caner (2007, *Journal of Econometrics* 137, 28–67) by means of a simple reparametrization of the model.

1. INTRODUCTION

In recent years, inference with weakly identified parameters has been analyzed intensively. Staiger and Stock (1997) and Kleibergen (2002) introduce the Anderson–Rubin (1949) and Lagrange multiplier (LM) types of tests. Both Otsu (2006) and Guggenberger and Smith (2008) provide the Anderson–Rubin (1949) and Kleibergen (2005) types of test statistics with time series data.

One of the most important issues in empirical research is structural change in the parameters of the system. Recognizing this, structural change tests in weakly identified models with endogenous regressors have been analyzed recently. Caner (2007) proposes boundedly pivotal tests in a nonlinear model that are composed of the likelihood ratio, Anderson–Rubin, and LM type of tests. However, tests in Caner (2007) may suffer from loss of power due to being boundedly pivotal. In the linear case, these boundedly pivotal structural change tests are equivalent to subvector testing.

The present paper introduces asymptotically pivotal tests in the case of weakly identified parameters in a system of simultaneous equations with time series data. Andrews (1993) and Sowell (1996) provide structural change tests in simultaneous equations in the strongly identified case. When there are weakly identified parameters, both of these tests lead to limits with nuisance parameters.

A simple reparametrization is used to achieve a nuisance-parameter-free limit. Instead of the null hypothesis of the equality of the structural parameter vector under two regimes, the null is the difference between the parameter vectors.

I thank editor Peter C.B. Phillips, co-editor Richard J. Smith, and the two anonymous referees for their comments that improved the paper substantially. Address correspondence to Mehmet Caner, 4168 Nelson Hall, Department of Economics, North Carolina State University, Raleigh, NC 27695, USA; e-mail: mcaner@ncsu.edu.

© Cambridge University Press 2010

https://doi.org/10.1017/S0266466610000290 Published online by Cambridge University Press

Section 2 analyzes the model. Section 3 develops asymptotics. Section 4 concludes the paper, and the Appendix covers the proofs. The symbol $\Longrightarrow$ denotes weak convergence.

2. THE MODEL

Consider the structural equations

$$y_{i1} = \beta_1' Y_{i2} + \gamma' Y_{i3} + u_i, \qquad i \leq [nr],$$
$$y_{i1} = \beta_2' Y_{i2} + \gamma' Y_{i3} + u_i, \qquad i > [nr],$$

where $i = 1, 2, \ldots, n$, $r \in \Xi$. Here Ξ denotes a set whose closure lies within $(0, 1)$. In practice, Ξ is chosen to be $[0.15, 0.85]$, where $[\cdot]$ represents the integer part in any number. Let y_1, Y_2, and Y_3 represent endogenous variables. Control variables are not used because they may be projected out easily when the model only involves change in certain endogenous explanatory variables. When the control variables are subject to structural change, a joint test for structural change can be conducted.

Rewrite the model in single equation form as

$$y_{i1} = \delta' Y_{i2r} + \beta_2' Y_{i2} + \gamma' Y_{i3} + u_i,$$

where $Y_{i2r} = Y_{i2} 1_{\{i \leq [nr]\}}$ and it is of dimension $l_a \times 1$, where $1_{\{\cdot\}}$ is a scalar indicator function. Let $\delta = \beta_1 - \beta_2$ and the dimension of Y_{i3} be $l_b \times 1$. In matrix form,

$$y_1 = Y_{2r}\delta + Y_2\beta_2 + Y_3\gamma + u. \tag{1}$$

Equation (1) is the structural equation with the dimensions of $y_1 : n \times 1, Y_2 : n \times l_a, Y_{2r} : n \times l_a, Y_3 : n \times l_b$. The Y_{2r} matrix is obtained by stacking the vectors Y_{i2r}'.

The reduced form equations are

$$Y_{i2} = \Pi_{A1}' X_i + V_{iA2}, \qquad i \leq [nr],$$
$$Y_{i2} = \Pi_{A2}' X_i + V_{iA2}, \qquad i > [nr],$$
$$Y_{i3} = \Pi_B' X_i + V_{iB},$$

where $X_i : k \times 1$ is the instrument vector and $X : n \times k$ is the matrix form. The matrix X is of full column rank, and the matrix X_r is formed by stacking the vectors $X_{ir} = X_i 1_{\{i \leq [nr]\}}$. In matrix form, the reduced form equations can be represented as

$$[Y_2, Y_3] = [X_r\tilde{\Pi} + X\Pi_{A2}, X\Pi_B] + [V_{A2}, V_B], \tag{2}$$

where $\tilde{\Pi} = \Pi_{A1} - \Pi_{A2}$, Π_{A1}, Π_{A2} are $k \times l_a$ matrices and $k \geq l_a + l_b$. The dimension of the matrix Π_B is $k \times l_b$, and the dimensions of the matrices V_{A2}, V_B

https://doi.org/10.1017/S0266466610000290 Published online by Cambridge University Press

are $n \times l_a, n \times l_b$, respectively. The nature of the matrices Π_{A1}, Π_{A2} is discussed in Assumption 2(d) in Section 3. The parameter γ is considered as strongly identified and δ as weakly identified. A joint structural change test is introduced for the parameters δ, $\tilde{\Pi}$, and so the null hypothesis is $H_0 : [\delta', \text{vec}(\tilde{\Pi})']' = 0$ where vec$(\cdot)$ denotes the vector version of $\tilde{\Pi}$. Structural change in γ is not tested. The moment conditions are

$$\text{E}[x_i(y_{i1} - \delta' Y_{i2r} - \beta_2' Y_{i2} - \gamma' Y_{i3})] = 0,$$

for any r $\in \Xi$. The tests depend on the orthogonality restrictions. Specifically, equations (A.7) and (A.8) in the Appendix explain the form of the statistics.

3. HETEROSKEDASTICITY AND AUTOCORRELATION CORRECTED TEST STATISTICS

The heteroskedasticity-autocorrelation corrected versions of the Anderson–Rubin type and the LM type structural change tests are introduced here. Specifically, in the case of weak instruments with weakly dependent data, the Guggenberger and Smith (2008) and Otsu (2006) methods are used. Next, the assumptions are discussed.

Assumptions

1. $S_n = cn^{\alpha}$, where $0 < \alpha < \frac{1}{2}$, $c > 0$, and c is a constant.
2. (a) $\sup_i \text{E}\|x_i\|^{2\xi} < \infty$. (b) $\sup_i \text{E}\|V_{iA2}\|^{2\xi} < \infty, \sup_i \text{E}\|V_{iB2}\|^{2\xi} < \infty$. (c) $\sup_i \text{E}\|u_i\|^{2\xi} < \infty$, where $\xi > 2/(1-2\alpha)$. (d) Assume that δ is weakly identified and γ is strongly identified. Set $\Pi_{A1} = C_{A1}/n^{1/2}, \Pi_{A2} = C_{A2}/n^{1/2}$, $\tilde{\Pi} = (C_{A1} - C_{A2})/n^{1/2}$; Π_B is of full column rank, and C_{A1}, C_{A2} are matrices with constants.
3. Note that via $v_{i1} = \delta' V_{iA2} + \beta_2' V_{iA2} + \gamma' V_{iB} + u_i$,

$$\sum_{i=1}^{[nr]} \frac{x_i v_{i1}}{n^{1/2}} \Longrightarrow B_k(r) \equiv \Sigma_{xv1}^{1/2} W_k(r),$$

where $\Sigma_{xv1} = \lim_{n\to\infty} var[n^{-1/2} \sum_{i=1}^n x_i v_{i1}]$ is the variance matrix for $k \times 1$ Brownian motion and $W_k(r)$ represents standard Brownian motion. Assume also

$$\sum_{i=1}^{[nr]} \frac{\text{vec}(x_i V_{iA2}')}{n^{1/2}} \Longrightarrow B_{kl_a}(r) \equiv \Sigma_{xv2}^{1/2} W_{kl_a}(r),$$

where $\Sigma_{xv2} = \lim_{n\to\infty} var[n^{-1/2} \sum_{i=1}^n \text{vec}(x_i V_{ia2}')]$ is the variance matrix for $kl_a \times 1$ Brownian motion $B_{kl_a}(r)$ and $W_{kl_a}(r)$ is $kl_a \times 1$ for the standard Brownian motion.
4. (a) $\sup_{i,j} \text{E}\|x_i v_{i1} v_{j1} x_j'\| < \infty$, $\sup_{s\in Z} \text{E}\|1/nS_n \sum_{j=1}^n \sum_{i=s}^{s+S_n} x_{i+j} v_{(i+j)1} v_{j1} x_j'\| = o(1)$, where Z is the set of integer numbers and S_n is the bandwidth

https://doi.org/10.1017/S0266466610000290 Published online by Cambridge University Press

parameter in Assumption 1. (b) Set $k^*(j/S_n) = 1 - |j/S_n|$ if $|j/S_n| \leq 2$ and 0 otherwise. This gives the Bartlett kernel. Set both $h_{i,r} = x_{ir}v_{i1}$ and $m_{i,r} = x_{ir}v_{i1}$ as in the definition of $\tilde{J}(h,m,r)$ in equation (3). For $j \geq 0$, and $j < 0$, respectively, for each $r \in \Xi$

$$\tilde{J}(h,m,r) = \frac{1}{n}\sum_{j=-n+1}^{n-1} k^*(j/S_n) \sum_{i=j+1}^{n} x_{ir}v_{i1}v_{(i-j)1}x'_{(i-j),r} \quad (3)$$

$$= \frac{1}{n}\sum_{j=-n+1}^{n-1} k^*(j/S_n) \sum_{i=j+1}^{n} x_{(i+j)r}v_{(i+j)1}v_{i1}x'_{ir}. \quad (4)$$

Assume uniformly over $r \in \Xi$ using (3) and (4)

$$\tilde{J}(h,m,r) \xrightarrow{p} r\,\Sigma_{xv1}.$$

Define $V_{iA2r} = V_{iA2}1_{\{i \leq [nr]\}}$ and set $h_{i,r} = \text{vec}(x_{ir}V'_{iA2r}), m_{i,r} = \text{vec}(x_{ir}V'_{iA2r})$ in $\tilde{J}(h,m,r)$. Assume uniformly over $r \in \Xi$

$$\tilde{J}(h,m,r) \xrightarrow{p} r\,\Sigma_{xv2}.$$

The same analysis applies to covariance between terms $\text{vec}(x_{ir}V'_{iA2r})$ and $v_{i1}x'_{ir}$. (c) Uniformly over $r \in \Xi$, assume

$$\left(\frac{1}{n}\sum_{i=1}^{[nr]} x_i x_i'\right) - r\,\Sigma_{xx} \xrightarrow{p} 0,$$

where Σ_{xx} is positive definite and finite.

5. (a) $\sup_{i,j} \text{E}\|x_i x_i' x_j x_j'\| < \infty$; $\sup_{s \in Z} \text{E}\|\frac{1}{nS_n}\sum_{j=1}^{n}\sum_{i=s}^{s+S_n} x_{i+j}x'_{i+j}x_j x_j'\| = o(1)$. (b) Set both $h_{i,r}$ and $m_{i,r}$ as $x_i x_i'$ in the definition of $\tilde{J}(h,m,r)$ and evaluate the full sample, the $\tilde{J}(h,m,1) = \tilde{J}(h,m)$,

$$\tilde{J}(h,m) - \text{var}\left[\frac{1}{n^{1/2}}\sum_{i=1}^{n}\text{vec}(x_i x_i')\right] \xrightarrow{p} 0,$$

$$\lim_{n\to\infty}\text{var}\left[\frac{1}{n^{1/2}}\sum_{i=1}^{n}\text{vec}(x_i x_i')\right] < \infty,$$

where the limit variance is nonsingular.

Assumption 1 is a standard bandwidth assumption as used in Guggenberger and Smith (2008) for analyzing weakly identified parameters with time series data. Assumption 2 is basically sufficient for

$$\sup_r \max_i \|x_{ir}v_{i1}\| = o_p(S_n^{-1}n^{1/2}).$$

Assumption 3 is a multivariate invariance principle whose sufficient conditions are given in Corollary 29.19 in Davidson (1994) and Lemma A.4 in Andrews

https://doi.org/10.1017/S0266466610000290 Published online by Cambridge University Press

(1993) and that involves a triangular array of random variables that are L^2 near epoch dependent at size $-1/2$. Assumptions 4(a) and 4(b) are used in the weak instruments literature in time series data as described in Guggenberger and Smith (2008). Assumption 4(b) is an extension of the method to structural change models and is also used as Assumption 3 in Andrews (1993). Assumption 4(c) is Lemma A.3 of Andrews (1993) where the sufficient conditions are also shown. Assumption 5 is helpful in showing that some remainder terms in the variance-covariance matrix estimator converge to zero in probability.

The following Anderson–Rubin test is designed for each $r \in \Xi$ and is robust to identification problems. Specifically

$$AR_r = \frac{(y_1' X_r^*)^{w*} \hat{\Omega}_{11,r}^{-1} (X_r^{*'} y_1)^{w*}}{2n},$$

where

$$(X_r^{*'} y_1)^{w*} = (X_r' y_1)^w - (X_r' X)^w [(X'X)^w]^{-1} (X' y_1)^w \tag{5}$$

and

$$\hat{\Omega}_{11,r} = \frac{S_n}{n} \sum_{i=1}^{n} (x_{ir}^* y_{i1}^*)^{w*} (y_{i1}^* x_{ir}^{*'})^{w*},$$

with

$$(X_r' y_1)^w = \sum_{i=1}^{n} S_n^{-1} \sum_{j=i-n}^{i-1} k(j/S_n) x_{i-j,r} y_{i-j,1},$$

where $k(j/S_n)$ is the truncated kernel and $x_{ir} = x_i 1_{\{i \le [nr]\}}$. As in Guggenberger and Smith (2008) a truncated kernel is used, $k(x) = 1$ if $|x| \le 1$ and $k(x) = 0$ otherwise. The proofs follow for other kernels. This paper uses the same smoothing approach as used in Guggenberger and Smith (2008) and Otsu (2006). The other terms in (5) are smoothed in the same way. Note that a slightly different smoothing method does not affect the asymptotic analysis used in the present paper. Specifically, the individual terms on the right-hand side of (5) are not weighted. The right-hand-side term is formed first, and then rather than the individual terms in (5), the right-hand-side term is weighted as a whole. The terms of the variance-covariance matrix are described before the proof of Lemma 1 in the Appendix.

The Anderson–Rubin test statistic can be written as

$$AR_r = \bar{S}_{hac,r}' \bar{S}_{hac,r},$$

where

$$\bar{S}_{hac,r} = \hat{\Omega}_{11,r}^{-1/2} \frac{(X_r^{*'} y_1)^{w*}}{2^{1/2} n^{1/2}}. \tag{6}$$

https://doi.org/10.1017/S0266466610000290 Published online by Cambridge University Press

Next, a result is presented that is new to the weak identification literature. Define $C_{A1}, C_{A2} \in \Gamma_C$, $\beta_1, \beta_2 \in B$, and Γ_C, B as compact subsets of $R^{k \times l_a}$, R^{l_a} respectively.

LEMMA 1. *Under Assumptions 1–4, for some* $C_{A1}, C_{A2} \in \Gamma_C, \beta_1, \beta_2 \in B$,

(i) $$\frac{(X_r^{*'} y_1)^{w*}}{n^{1/2}} \Longrightarrow 2\Sigma_{xv1}^{1/2}[W_k(r) - rW_k(1)] + 2r(1-r)\Sigma_{xx}[C_{A1}\beta_1 - C_{A2}\beta_2],$$

where $W_k(r)$ *is the* $k \times 1$ *standard Brownian motion and* $W_k(1)$ *is the* $k \times 1$ *standard normal vector. The term* Σ_{xv1} *is explained in the Assumptions section.*

(ii) *Under the joint null hypotheses of* $\delta = 0, \tilde{\Pi} = 0$,

$$\frac{(X_r^{*'} y_1)^{w*}}{n^{1/2}} \Longrightarrow 2\Sigma_{xv1}^{1/2}[W_k(r) - rW_k(1)].$$

Next, the limit for the variance-covariance matrix estimator $\hat{\Omega}_{11,r}$ is derived. The result in Lemma 2 is different from that in Guggenberger and Smith (2008), the difference stemming from the analysis of unknown change points.

LEMMA 2. *Under Assumptions 1–5, uniformly over* $r \in \Xi$

$$\hat{\Omega}_{11,r} \xrightarrow{p} 2r(1-r)\Sigma_{xv1}.$$

Lemmas 1 and 2 are combined to derive the limit for the Anderson–Rubin test statistic. The following limit is one of the main theorems of the paper.

THEOREM 1. *Under Assumptions 1–5, for some* $C_{A1}, C_{A2} \in \Gamma_C, \beta_1, \beta_2 \in B$,

(i) $$\sup_{r \in \Xi} AR_r \xrightarrow{d} \sup_{r \in \Xi} \times \frac{[(W_k(r) - rW_k(1)) + M_{A1A2}(r)]'[(W_k(r) - rW_k(1)) + M_{A1A2}(r)]}{r(1-r)},$$

where $M_{A1A2}(r) = r(1-r)\Sigma_{xx}(C_{A1}\beta_1 - C_{A2}\beta_2)$.

(ii) *Under the null hypotheses of* $\delta = 0, \tilde{\Pi} = 0$ *(i.e.,* $\beta_1 = \beta_2, C_{A1} = C_{A2}$*)*

$$\sup_{r \in \Xi} AR_r \xrightarrow{d} \sup_{r \in \Xi} \frac{[(W_k(r) - rW_k(1))]'[(W_k(r) - rW_k(1))]}{r(1-r)}.$$

https://doi.org/10.1017/S0266466610000290 Published online by Cambridge University Press

Remarks.

1. Caner (2007) has previously provided an Anderson–Rubin based structural change test for nonlinear generalized method of moments with weakly identified parameters. However, that test is boundedly pivotal, whereas the limit in Theorem 1(ii) is pivotal and so will involve no loss of power on the latter. The critical values of the test can be obtained from Andrews (1993, Table 1).
2. The $\sup AR_r$ test here depends on equation (1), and so the null is basically $\delta = \beta_1 - \beta_2 = 0$, $\tilde{\Pi} = \Pi_1 - \Pi_2 = 0$, and a nuisance-parameter-free limit can be derived after some simple projection. The form of the test in Caner (2007) is different than that which is considered here. In that case, the model cannot be reparametrized as in (1) in the nonlinear case and cannot use the projection used here. Under the null of $H_0 : \beta_1 = \beta_2 = \beta$, the Caner (2007) test uses constrained estimates, and the following bounded limit is given:

$$\sup_r \frac{[W_k(r) - rW_k(1)]'[W_k(r) - rW_k(1)]}{r(1-r)} + \chi_k^2.$$

   Therefore the bound has an extra χ_k^2 term compared to Theorem 1 here, which is independent of the Brownian bridge term. It can be seen that loss of power may occur with a large number of instruments.
3. The $\sup LM$ test in Andrews (1993) is not asymptotically pivotal in the case of weak identification as has been documented in Caner (2007).
4. Theorem 1 extends the literature on weak instruments from standard inference to testing for regime shifts.
5. Note that when there is no structural change in the reduced form ($C_{A1} = C_{A2} = C_A$), but there is a structural change in the structural equation ($\beta_1 \neq \beta_2$), the M_{A1A2} expression will have the following term: $C_A(\beta_1 - \beta_2)$. It is clear from Theorem 1(i) that the test is not consistent, but the power may increase with a large structural change.

Next the LM test is considered. For each $r \in \Xi$, $P_{\bar{T}_{hac,r}} = \bar{T}_{hac,r}(\bar{T}'_{hac,r}\bar{T}_{hac,r})^{-1}\bar{T}'_{hac,r}$

$$LM_r = \bar{S}'_{hac,r} P_{\bar{T}_{hac,r}} \bar{S}_{hac,r},$$

where

$$\bar{T}_{hac,r} = \hat{\Omega}_{11,r}^{-1/2} J_r, \tag{7}$$

$$J_r = \left[n^{-1/2} \sum_{i=1}^{n} (x_{ir}^* Y'_{i2r})^{w*} \right] - \hat{\Omega}_{21,r}\hat{\Omega}_{11,r}^{-1} \left[n^{-1/2} \sum_{i=1}^{n} (x_{ir}^* y_{i1})^{w*} \right],$$

https://doi.org/10.1017/S0266466610000290 Published online by Cambridge University Press

and

$$\hat{\Omega}_{21,r} = \frac{S_n}{n}\sum_{i=1}^{n} \mathrm{vec}(x_{ir}^{*} Y_{i2r}^{*'})^{w*}(y_{i1}^{*} x_{ir}^{*'})^{w*},$$

$$Y_{i2r}^{*} = Y_{i2r}' - x_i' \left(\sum_{i=1}^{n} x_i x_i'\right)^{-1} \left(\sum_{i=1}^{n} x_i Y_{i2r}'\right). \tag{8}$$

Note that Y_{i2r} corresponds to the vector in (2). On the other hand, Y_{i2r}^{*} is different and is defined in (8). The term Y_{i2r}^{*} is used only in the variance-covariance matrix estimation. Note that $X_{ir}^{*} Y_{i2r}^{*'} = X_{ir}^{*} Y_{i2r}'$. The limits of the terms in the LM test are needed before the test statistic limit can be obtained, and so the limit for $n^{-1/2}\sum_{i=1}^{n}(x_{ir}^{*} Y_{i2r}')^{w*}$ has to be considered first.

LEMMA 3. *Under Assumptions 1–5, for some* $C_{A1} \in \Gamma_C$

$$n^{-1/2}\sum_{i=1}^{n} \mathrm{vec}(x_{ir}^{*} Y_{i2r}')^{w*} \Longrightarrow 2(1-r)[\Sigma_{xv2} W_{kl_a}(r) + r\,\mathrm{vec}(\Sigma_{xx} C_{A1})],$$

where $W_{kl_a}(r)$ *is* $kl_a \times 1$ *dimensional standard Brownian motion and* Σ_{xv2} *is the limit for* $\mathrm{var}[n^{-1/2}\sum_{i=1}^{n}\mathrm{vec}(x_i V_{i2})]$.

Next the limit for the variance-covariance estimator, $\hat{\Omega}_{21,r}$, is shown. Specifically, Σ_{xv21} is the limit for $\mathrm{covar}[n^{-1/2}\sum_{i=1}^{n}([\mathrm{vec}(x_i V_{i2}')], [v_{i1} x_i'])]$, where $\mathrm{covar}(.,.)$ represents the covariance between the two terms.

LEMMA 4. *Under Assumptions 1–5, uniformly over* $r \in \Xi$,

$$\hat{\Omega}_{21,r} \xrightarrow{p} 2r(1-r)^2 \Sigma_{xv21}.$$

The limit of the Kleibergen (2002) type test is introduced in the next theorem, which is one of the main results of this paper.

THEOREM 2. *Under Assumptions 1–5,*

(i) For some $C_{A1}, C_{A2} \in \Gamma_C$, $\beta_1, \beta_2 \in B$

$$\sup_{r\in\Xi} LM_r \xrightarrow{d} \sup_{r\in\Xi} \times \frac{[(W_{l_a}(r) - r W_{l_a}(1)) + M_J(r)]'[(W_{l_a}(r) - r W_{l_a}(1)) + M_J(r)]}{r(1-r)},$$

where

$$M_J(r) = r(1-r)(J_{l_a}(r)' \Sigma_{xv1}^{-1} J_{l_a}(r))^{-1/2} J_{l_a}(r)' \Sigma_{xv1}^{-1} \Sigma_{xx}(C_{A1}\beta_1 - C_{A2}\beta_2),$$

$J_{l_a}(r)$ is a functional of kl_a and k dimensional Brownian motions, which is described at the end of the proof of Theorem 2(i) in the Appendix, and $M_J(r)$ and $(W_{l_a}(r) - r W_{l_a}(1))$ are independent.

https://doi.org/10.1017/S0266466610000290 Published online by Cambridge University Press

(ii) Under the null hypotheses of $C_{A1} = C_{A2}$, $\beta_1 = \beta_2$ *(i.e.,* $\tilde{\Pi} = 0$, $\delta = 0$*)*

$$\sup_{r\in\Xi} LM_r \xrightarrow{d} \sup_{r\in\Xi} \frac{[(W_{l_a}(r) - rW_{l_a}(1))]'[(W_{l_a}(r) - rW_{l_a}(1))]}{r(1-r)}.$$

Remarks.

1. Remarks 1, 3, 4, and 5 for Theorem 1 apply here also.
2. The limit depends on "l_a" and may have better small-sample properties than the Anderson–Rubin test when the number of instruments is large.
3. A different type of LM test is proposed in Caner (2007) where an estimate of the parameter is found by minimizing the Kleibergen type of test, given a change point. Then the sup of that Kleibergen test is evaluated at this minimum, and the end result is not asymptotically but only boundedly pivotal. The distribution of the bound given by Theorem 2 in Caner (2007) is

$$\sup_r \frac{[W_{l_a}(r) - rW_{l_a}(1)]'[W_{l_a}(r) - rW_{l_a}(1)]}{r(1-r)} + \chi^2_{l_a}.$$

Therefore the preceding limit is just $\chi^2_{l_a}$ larger than the limit in Theorem 2(i). The loss of power may be important when the number of parameters is large.

4. CONCLUSION

This paper introduces tests for structural change for the structural and reduced form parameters in a system of simultaneous equations. When there is weak identification in the parameters of the reduced form system, the existing tests are not asymptotically pivotal, as shown recently by Caner (2007), who also introduced structural change tests that are robust to identification problems but that are boundedly pivotal. The tests here differ in that they are asymptotically pivotal and also take the time series nature of the data better into account.

REFERENCES

Anderson, T.W. & H. Rubin (1949) Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics* 20, 46–63.

Andrews, D.W.K. (1993) Tests for parameter instability and structural change with unknown change point. *Econometrica* 61, 821–856.

Caner, M. (2007) Boundedly pivotal structural change tests in continuous updating GMM with strong, weak identification and completely unidentified cases. *Journal of Econometrics* 137, 28–67.

Davidson, J. (1994) *Stochastic Limit Theory*. Oxford University Press.

Guggenberger, P. & R.J. Smith (2008) Generalized empirical likelihood tests in time series models with potential identification failure. *Journal of Econometrics* 142, 134–162.

Kleibergen, F. (2002) Pivotal statistics for testing structural parameters in instrumental variables regression. *Econometrica* 70, 1781–1805.

https://doi.org/10.1017/S0266466610000290 Published online by Cambridge University Press

Kleibergen, F. (2005) Testing parameters in GMM without assuming that they are identified. *Econometrica* 73, 1103–1124.
Otsu, T. (2006) Generalized empirical likelihood inference for nonlinear and time series models under weak identification. *Econometric Theory* 22, 513–529.
Sowell, F. (1996) Optimal tests for parameter instability in the generalized method of moments framework. *Econometrica* 64, 1085–1109.
Staiger, D. & J.H. Stock (1997) Instrumental variables regression with weak instruments. *Econometrica* 65, 557–587.

APPENDIX

The following expressions are needed before the proofs can be undertaken. From (A.6) later in this Appendix, the smoothed version of the vector that is used in variance-covariance estimation is

$$(x_{ir}^{*}y_{i1}^{*})^{w*} = (x_{ir}^{*}x_{ir}^{*'})^{w*}\left(\frac{C_{A1}\beta_1 - C_{A2}\beta_2}{n^{1/2}}\right) + (x_{ir}^{*}v_{i1})^{w}, \tag{A.1}$$

where

$$(x_{ir}^{*}v_{i1})^{w*} = (x_{ir}v_{i1})^{w} - (x_{ir}x_i')^{w}\bar{B}^{w} - \bar{A}^{w}(x_iv_{i1})^{w} + \bar{A}^{w}(x_ix_i')^{w}\bar{B}^{w} \tag{A.2}$$

and

$$\bar{A}^{w} = \left[\sum_{i=1}^{n}(x_{ir}x_i')^{w}\right]\left[\sum_{i=1}^{n}(x_ix_i')^{w}\right]^{-1}, \tag{A.3}$$

$$\bar{B}^{w} = \left[\sum_{i=1}^{n}(x_ix_i')^{w}\right]^{-1}\left[\sum_{i=1}^{n}(x_iv_{i1})^{w}\right], \tag{A.4}$$

where $x_{ir}^{*} = x_{ir} - \left(\sum_{i=1}^{n}x_{ir}x_i'\right)\left(\sum_{i=1}^{n}x_ix_i'\right)^{-1}x_i$.

Then define

$$(x_{ir}^{*}x_{ir}^{*'})^{w*} = [(x_{ir}x_{ir}')^{w} - (x_{ir}x_i')^{w}\bar{A}^{w'} - \bar{A}^{w}(x_ix_{ir}')^{w} + A^{w}(x_ix_i')^{w}\bar{A}^{w'}] \times \left(\frac{C_1\beta_1 - C_2\beta_2}{n^{1/2}}\right). \tag{A.5}$$

First, set $X_r^{*} = M_xX_r$, $M_x = I - X(X'X)^{-1}X'$. Next, substitute the Y_2, Y_3, Y_{2r} expressions in (2) into the structural equation (1) to produce

$$y_1 = X_r\Pi_{A1}\delta + X_r\tilde{\Pi}\beta_2 + X\Pi_{A2}\beta_2 + X\Pi_B\gamma + v_{1r}, \tag{A.6}$$

where $v_{1r} = V_{A2r}\delta + V_{A2}\beta_2 + V_B\gamma + u$. Note that X_r is orthogonal to each V term for each r. For fixed $\delta, \tilde{\Pi}$

$$X_r^{*'}y_1 = X_r^{*'}X_r\Pi_{A1}\delta + X_r^{*'}X_r\tilde{\Pi}\beta_2 + X_r^{*'}v_{1r}. \tag{A.7}$$

Under the null of $\delta = 0, \tilde{\Pi} = 0$,

$$X_r^{*'}y_1 = X_r^{*'}v_{1r} = X_r^{*'}V_{A2}\beta_2 + X_r^{*'}V_B\gamma + X_r^{*'}u, \tag{A.8}$$

https://doi.org/10.1017/S0266466610000290 Published online by Cambridge University Press

where $X_r^{*'} X = 0$. Note also that

$$\Pi_{A1}\delta + \tilde{\Pi}\beta_2 = \Pi_{A1}\beta_1 - \Pi_{A2}\beta_2.$$

Proof of Lemma 1(i). Following (A.7)

$$(X_r^{*'} y_1)^{w*} = (X_r^{*'} X_r)^{w*}(\Pi_{A1}\delta + \tilde{\Pi}\beta_2) + (X_r^{*'} v_{1r})^{w*},$$

where

$$(X_r^{*'} v_{1r})^{w*} = (X_r' v_{1r})^w - (X_r' X)^w [(X'X)^w]^{-1}(X' v_{1r})^w \tag{A.9}$$

and

$$(X_r^{*'} X_r)^w = (X_r' X)^w - (X_r' X)^w [(X'X)^w]^{-1}(X' X_r)^w. \tag{A.10}$$

Each term is considered in detail. Note that from Assumption 3

$$X_r' v_{1r} = X_r' v_1 = X_r'(V_{A2}\delta + V_{A2}\beta_2 + V_B\gamma + u).$$

The proof here follows from the proof of Lemma 1 of Guggenberger and Smith (2008, eqn. (A.6)) via Assumption 3, so that

$$n^{-1/2}\sum_{i=1}^{n}(x_{ir} v_{i1})^w \Longrightarrow 2\Sigma_{xv1}^{1/2} W_k(r). \tag{A.11}$$

Next given Assumptions 2a, 4b, and 4c and by applying Lemma 1 of Guggenberger and Smith (2008) uniformly in r,

$$n^{-1}\sum_{i=1}^{n}(x_{ir} x_i')^w \xrightarrow{p} 2r\Sigma_{xx}. \tag{A.12}$$

Then combine these in (A.9) with full-sample versions of (A.11) and (A.12) to provide

$$\left(\frac{X_r^{*'} v_{1r}}{n^{1/2}}\right)^{w*} \Longrightarrow 2\Sigma_{xv1}^{1/2}[W_k(r) - rW_k(1)].$$

Via (A.10), (A.12), and Assumption 2(d),

$$\left(\frac{X_r^{*'} X_r}{n^{1/2}}\right)^{w*} [\Pi_{A1}\delta + \tilde{\Pi}\beta_2] \xrightarrow{p} 2r(1-r)\Sigma_{xx}[C_{A1}\beta_1 - C_{A2}\beta_2].$$

Combine the last two equations to get the desired result. ■

Proof of Lemma 1(ii). When $\delta = 0, \tilde{\Pi} = 0$, which gives $\beta_1 = \beta_2, C_{A1} = C_{A2}$. The result can then be derived from Lemma 1(i). ■

Proof of Lemma 2. First analyze $\bar{A}^w, \bar{B}^w$ in (A.3) and (A.4). Then by (A.12) and the full-sample version of (A.12) (i.e., the $r = 1$ case), uniformly in r,

$$\bar{A}^w \xrightarrow{p} rI_k. \tag{A.13}$$

https://doi.org/10.1017/S0266466610000290 Published online by Cambridge University Press

Because $Ex_i v_{i1} = 0$ and using Lemma 1 in Guggenberger and Smith (2008),

$$n^{-1}\sum_{i=1}^{n}(x_i v_{i1})^w \xrightarrow{p} 0.$$

Then

$$\bar{B}^w \xrightarrow{p} 0. \tag{A.14}$$

Equations (A.13) and (A.14) provide two very useful results that will be used in the following proofs. Then use (A.1), (A.2), and (A.5) to have

$$\begin{aligned}\hat{\Omega}_{11,r} = \frac{S_n}{n}\sum_{i=1}^{n}&[(x_{ir}v_{i1})^w - (x_{ir}x_i')^w\bar{B}^w - \bar{A}^w(x_i v_{i1})^w + \bar{A}^w(x_i x_i')^w\bar{B}^w]\\ &\times[(x_{ir}v_{i1})^w - (x_{ir}x_i')^w\bar{B}^w - \bar{A}^w(x_i v_{i1})^w + \bar{A}^w(x_i x_i')^w\bar{B}^w]' + o_p(1).\end{aligned} \tag{A.15}$$

The asymptotically negligible remainder term in (A.15) is due to weak instrument asymptotics. Equation (A.15) can also be decomposed into several terms:

$$\begin{aligned}\hat{\Omega}_{11,r} = \frac{S_n}{n}\sum_{i=1}^{n}(x_{ir}v_{i1})^w(v_{i1}x_{ir}')^w - \left[\frac{S_n}{n}\sum_{i=1}^{n}(x_{ir}v_{i1})^w(v_{i1}x_i')^w\right]\bar{A}^{w'}\\ - \bar{A}^w\left[\frac{S_n}{n}\sum_{i=1}^{n}(x_i v_{i1})^w(v_{i1}x_{ir}')^w\right] + \bar{A}^w\left[\frac{S_n}{n}\sum_{i=1}^{n}(x_i v_{i1})^w(v_{i1}x_i')^w\right]\bar{A}^{w'}\\ + o_p(1) + o_p(1).\end{aligned} \tag{A.16}$$

In (A.16), the second asymptotically negligible term is obtained from Lemma 2 of Guggenberger and Smith (2008) via Assumptions 4 and 5. Now the first term on the right-hand side of (A.16) is considered. Via Assumptions 4a, 4b, and 4c and Lemma 2 of Guggenberger and Smith (2008)

$$\frac{S_n}{n}\sum_{i=1}^{n}(x_{ir}v_{i1})^w(v_{i1}x_{ir}')^w \xrightarrow{p} 2r\Sigma_{xv1}.$$

Then in the same manner, but also benefiting from (A.13),

$$\bar{A}^w\left[\frac{S_n}{n}\sum_{i=1}^{n}(x_i v_{i1})^w(v_{i1}x_{ir}')^w\right] \xrightarrow{p} 2r^2\Sigma_{xv1}.$$

Therefore with (A.13) and the full-sample result from Guggenberger and Smith (2008)

$$\bar{A}^w\left[\frac{S_n}{n}\sum_{i=1}^{n}(x_i v_{i1})^w(v_{i1}x_i')^w\right]\bar{A}^{w'} \xrightarrow{p} 2r^2\Sigma_{xv1}.$$

Substituting these into (A.16), uniformly in r, the following result is derived:

$$\hat{\Omega}_{11,r} \xrightarrow{p} 2r(1-r)\Sigma_{xv1}. \qquad \blacksquare$$

https://doi.org/10.1017/S0266466610000290 Published online by Cambridge University Press

Proof of Lemma 3. First analyze the following equation:

$$n^{-1/2}\sum_{i=1}^{n}(x_{ir}^{*}Y_{i2r}')^{w*}=n^{-1/2}\sum_{i=1}^{n}(x_{ir}^{*}x_{ir}')^{w*}\Pi_{A1}+n^{-1/2}\sum_{i=1}^{n}(x_{ir}^{*}V_{iA2r}')^{w*}. \tag{A.17}$$

In (A.17) consider

$$\begin{aligned}n^{-1/2}\sum_{i=1}^{n}(x_{ir}^{*}V_{iA2r}')^{w*}=&\left[n^{-1/2}\sum_{i=1}^{n}(x_{ir}V_{iA2r}')^{w}\right]\\&-\left[n^{-1}\sum_{i=1}^{n}(x_{ir}x_{i}')^{w}\right]\left[n^{-1}\sum_{i=1}^{n}(x_{i}x_{i}')^{w}\right]^{-1}\\&\times\left[n^{-1/2}\sum_{i=1}^{n}(x_{i}V_{iA2r}')^{w}\right].\end{aligned}$$

Note that the first and the last terms are the same because $x_{ir}V_{iA2r}'=x_{i}V_{iA2r}$. Then the same analysis as in the proof of Lemma 1 is used by benefiting from equations (A.11) and (A.12) and Assumption 3:

$$n^{-1/2}\sum_{i=1}^{n}\mathrm{vec}(x_{ir}^{*}V_{iA2r})^{w*}\Longrightarrow 2(1-r)\Sigma_{xv2}^{1/2}W_{kl_a}(r), \tag{A.18}$$

where $W_{kl_a}(r)$ is $kl_a\times 1$ dimensional standard Brownian motion and Σ_{xv2} is the limit for $\mathrm{var}[n^{-1/2}\sum_{i=1}^{n}\mathrm{vec}(x_iV_{iA2})]$.

Analyze the following term in (A.17) by assuming weak instrument asymptotics $\Pi_{A1}=C_{A1}/n^{1/2}$:

$$\begin{aligned}n^{-1/2}\sum_{i=1}^{n}(x_{ir}^{*}x_{ir}')^{w*}\Pi_{A1}&=n^{-1}\sum_{i=1}^{n}(x_{ir}^{*}x_{ir}')^{w*}C_{A1}\\&=n^{-1}\sum_{i=1}^{n}(x_{ir}x_{ir}')^{w}C_{A1}-\left(n^{-1}\sum_{i=1}^{n}(x_{ir}x_{i}')^{w}\right)\\&\quad\times\left(n^{-1}\sum_{i=1}^{n}(x_{i}x_{i}')^{w}\right)^{-1}\left(n^{-1}\sum_{i=1}^{n}(x_{i}x_{ir}')^{w}\right)C_{A1}\\&\xrightarrow{p}2r(1-r)\Sigma_{xx}C_{A1},\end{aligned}$$

by (A.3), (A.12), and (A.13). Therefore

$$n^{-1/2}\sum_{i=1}^{n}\mathrm{vec}(x_{ir}^{*}x_{ir}')^{w*}\Pi_{A1}\xrightarrow{p}2r(1-r)\mathrm{vec}(\Sigma_{xx}C_{A1}). \tag{A.19}$$

Then substitute (A.18) and (A.19) in (A.17) to gain the desired result. ■

Proof of Lemma 4. Lemma 4 considers one of the most important terms in the variance covariance matrix estimator. By (A.5),

$$\begin{aligned}(x_{ir}^{*}Y_{i2r}')^{w*}=&(x_{ir}Y_{i2r}')^{w}-(x_{ir}x_{i}')^{w}\bar{F}^{w}\\&-\bar{A}^{w}(x_{i}Y_{i2r}')^{w}+\bar{A}^{w}(x_{i}x_{i}')^{w}\bar{F}^{w},\end{aligned}$$

https://doi.org/10.1017/S0266466610000290 Published online by Cambridge University Press

where by Lemma 3 and the full-sample version of (A.12)

$$\bar{F}^w = \left[\frac{\sum_{i=1}^n (x_i x_i')^w}{n}\right]^{-1}\left[\frac{\sum_{i=1}^n (x_i Y_{i2r}')^w}{n}\right] \xrightarrow{p} 0.$$

Next use the definition of $\hat{\Omega}_{21,r}$ and proceed in the same way as in the proof of Lemma 2 to get the result. ■

Proof of Theorem 2.

(i) The proof relies on showing that $(X_r^{*'} y_1)^{w*}$ and J_r are asymptotically independent. Note that under the weak instruments assumption $\Pi_{A1} = C_{A1}/n^{1/2}$, where C_{A1} is a $k \times l_a$ constant matrix.
Note that by Lemmas 3 and 4 and (A.19),

$$\begin{aligned}\text{vec}(J_r, (n^{-1/2}(X_r^{*'} y_1)^{w*})) &= r(1-r)\text{vec}(\Sigma_{xx} C_{A1}, 0) \\ &\quad + \begin{bmatrix} I_{kl_a} & -(1-r)\Sigma_{xv21}\Sigma_{xv1}^{-1} \\ 0 & I_k \end{bmatrix} \\ &\quad \times \begin{bmatrix} n^{-1/2}\text{vec}(X_r^{*'} V_{A2r})^{w*} \\ n^{-1/2}(X_r^{*'} y_1)^{w*} \end{bmatrix} + o_p(1). \qquad \textbf{(A.20)}\end{aligned}$$

By Lemma 1 and (A.18) the sample variance covariance converges to the following limit uniformly in $r \in \Xi$:

$$\text{var}\begin{bmatrix} n^{-1/2}\text{vec}(X_r^{*'} V_{A2r})^{w*} \\ n^{-1/2}(X_r^{*'} y_1)^{w*} \end{bmatrix} \to \begin{bmatrix} r(1-r)^2\Sigma_{xv2} & r(1-r)^2\Sigma_{xv21} \\ r(1-r)^2\Sigma_{xv21}' & r(1-r)\Sigma_{xv1} \end{bmatrix}. \qquad \textbf{(A.21)}$$

Using (A.21) in (A.20) gives the limit variance-covariance matrix for

$$\begin{aligned}&\text{var}(\text{vec}(J_r, n^{-1/2}(X_r^{*\prime} y_1)^{w*})) \\ &\quad \to \begin{bmatrix} r(1-r)^2[\Sigma_{xv2} - (1-r)\Sigma_{xv21}\Sigma_{xv1}^{-1}\Sigma_{xv21}'] & 0 \\ 0 & r(1-r)\Sigma_{xv1} \end{bmatrix},\end{aligned}$$

which shows the asymptotic independence of J_r and $(X_r^{*\prime} y_1)^{w*}$. Rewrite the LM test statistic by referring to (6) and (7) as $LM_r = \frac{1}{2n}(y_1' X_r^*)^{w*}\hat{\Omega}_{11,r}^{-1/2} P_{\hat{\Omega}_{11,r}^{-1/2} J_r}$ $\hat{\Omega}_{11,r}^{-1/2}(X_r^{*\prime} y_1)^{w*}$. Note that $P_{\hat{\Omega}_{11,r}^{-1/2} J_r} = \hat{\Omega}_{11,r}^{-1/2} J_r (J_r' \hat{\Omega}_{11,r}^{-1} J_r)^{-1} J_r' \hat{\Omega}_{11,r}^{-1/2}$. Following Lemmas 1(i), 2, and 3

$$\begin{aligned}J_r \Longrightarrow\ & 2(1-r)[\Sigma_{xv2} W_{kl_a}(r) + r\,\text{vec}(\Sigma_{xx} C_{A1})] \\ & - 2(1-r)\Sigma_{xv1}^{-1}\left[\Sigma_{xv1}^{1/2}(W_k(r) - r W_k(1)) + 2r(1-r)\Sigma_{xx}(C_{A1}\beta_1 - C_{A2}\beta_2)\right] \\ & \equiv J_{l_a}.\end{aligned}$$

Then in a simple rotation, apply Lemmas 1 and 2, and the form of the test provides the result. ■

(ii) Imposing the null of $\beta_1 = \beta_2, C_{A1} = C_{A2}$ provides the result.

https://doi.org/10.1017/S0266466610000290 Published online by Cambridge University Press