Pivotal Structural Change Tests in Linear Simultaneous Equations with Weak Identification

Mehmet Caner *

North Carolina State University

August 31, 2007

Abstract

This paper develops asymptotically pivotal structural change tests in simultaneous equations with weakly identified parameters. In former literature, Caner (2007) proposes boundedly pivotal structural change tests when there are weakly identified parameters. The tests developed in this article are new, and benefit from reparameterizing the model. This results in asymptotically pivotal tests. Simulation exercise compares the pivotal and boundedly pivotal tests.

Keywords: Similar Tests, Regime Shift, Structural Parameters

*Comments in "Structural Change Tests Conference" in CASS Business School, London in December 2006 were helpful. email: mcaner@ncsu.edu.

1 Introduction

In recent years, testing in weakly identified framework has been analyzed intensively. Staiger and Stock (1997), Kleibergen (2002) introduce Anderson-Rubin (1949) and LM type of tests. Recently, there has been various test statistics suggested in weak identification literature with time series data. Both Otsu (2006), and Guggenberger and Smith (2006) provide Anderson-Rubin (1949) and Kleibergen (2005) type of test statistics. These are all asymptotically pivotal.

We know that one of the most important issues in empirical research is change in the parameters of the system. However, the structural change tests in the weakly identified models with endogenous regressors have been analyzed very recently. Caner (2007) proposes asymptotically boundedly pivotal tests in a nonlinear Continuous Updating framework. These are likelihood ratio, Anderson-Rubin and LM type of tests. However, they may suffer from power loss. Even in the linear case, the tests are still boundedly pivotal since this testing is not trivial and amounts to subvector testing.

In this article, we introduce asymptotically pivotal tests when there are weakly identified parameters in a simultaneous equations system with weakly dependent time series data. This is new in the literature and may be more useful in applied work. Structural change tests in simultaneous equations exist in the strongly identified case such as Andrews (1993), Sowell (1996). Both of these are subject to criticism of using inconsistent estimates in the tests and hence leading to limits with nuisance parameters when there are weakly identified parameters.

We use a simple reparameterization to achieve the nuisance parameter free limit. Instead of testing the null of the equality of the structural parameter vector under two regimes, we test the null hypotheses of the difference between the parameter vectors under two regimes are zero. We show that a version of Anderson-Rubin (1949) and Kleibergen (2002) type of tests are asymptotically pivotal in time series case with unknown change point. Next we conduct some simulations and show that the tests have reasonable small sample properties.

Section 2 analyzes the model. Section 3 develops asymptotics. Section 4 conducts some Monte Carlo exercise. Section 5 concludes the paper. For notation clarification: "$\Longrightarrow$" denotes weak convergence, "$\bigotimes$" denotes Kronecker product. Appendix covers the proofs.

2 The Model

Consider the structural equation:

$$y_{i1} = \beta_1' Y_{i2} + u_i, \quad i \leq [nr]$$

$$y_{i1} = \beta_2' Y_{i2} + u_i, \quad i > [nr]$$

where $i = 1, 2, \cdots, n$ are observations, $r \in \Xi, \Xi$ denotes a set whose closure lies in $(0, 1)$. In practice, Ξ is chosen to be $[.15, .85]$, for this point see Andrews (1993). [.] represents the integer part in any number. y_1 and Y_2 represent endogenous variables. We abstract from control variables. Exogenous control variables may be projected out easily when the model involves only change in the endogenous explanatory variables. If the model involves change in the control variables in the structural equation as well, a joint test for structural change can be conducted.

Then rewrite this in single equation:

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

where $Y_{i2r} = Y_{i2} 1_{\{i \leq [nr]\}}$, Y_{i2r} is $l \times 1$, $1_{\{.\}}$ is a scalar indicator function and $\delta = \beta_1 - \beta_2$.

In matrix form,

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

Equation(1) is our structural equation. $y_1 : n \times 1, Y_2 : n \times l, Y_{2r} : n \times l$. Y_{2r} matrix is obtained by stacking the vectors Y_{i2r}'.

Y_2 and instruments X are related in the following way:

$$Y_2 = X\Pi + V_2,$$

where X is $n \times k$ instrument matrix and is of full column rank , Π is $k \times l$ matrix, and V_2 is $n \times l$ matrix. Assume $k \geq l$.

If there is structural change in the reduced form equation, then the analysis here is not valid. In that case, the null distribution will depend on the value of the reduced form matrix in the two regimes. A different approach and different tests may be needed in that scenario.

The reduced form system can be rearranged as:

$$y_1 = X_r \Pi \delta + X \Pi \beta_2 + v_{1r}. \tag{2}$$

$$Y_{2r} = X_r \Pi + V_{2r}. \tag{3}$$

where X_r is formed by stacking up vectors : $x_i' 1_{\{i \leq [nr]\}}$. (same as Y_{2r} formation) so $v_{1r} = V_{2r}\delta + V_2\beta_2 + u$ and V_{2r} is formed in the same way as Y_{2r} above.

X_r is orthogonal to V_r for each r. We want to test $H_0 : \delta = 0$ against $H_1 : \delta \neq 0$, we treat Π, β_2 as nuisance parameters.

Furthermore, we can rewrite the reduced form in the matrix form

$$Y_r = X_r \Pi_* + X[\Pi\beta_2, 0] + V_r, \tag{4}$$

where $Y_r = [y_1, Y_{2r}], \Pi_* = \Pi A'$. $A = [\delta, I_l]'$.

3 Heteroskedasticity and Autocorrelation Corrected Test Statistics

In this part we analyze the heteroskedasticity-autocorrelation corrected versions of Anderson-Rubin type structural change test and LM like test statistic. We benefit from the methods of Guggenberger and Smith (2006), and Otsu (2006) in the case of weak instruments with weakly dependent data. First, we show the smoothing idea in a time series case. Then we provide a specific analysis for our case. Let θ denote the parameters z_i is the data and $g(z_i, \theta)$ represents the moments. We denote $g(z_i, \theta)$ by $g_i(\theta)$. The smoothed moments are

$$g_{in}(\theta) = S_n^{-1} \sum_{j=i-n}^{i-1} k(j/S_n) g_{i-j}(\theta),$$

where S_n is a bandwidth parameter. $S_n \to \infty$ as $n \to \infty$. $k(.)$ is a kernel as in Guggenberger and Smith (2006) we use a truncated kernel $k(x) = 1$ if $|x| \leq 1$ and $k(x) = 0$ otherwise. The crucial issue is what will be the smoothed moments in our specific case. Building also a heteroskedasticity and autocorrelation consistent $\bar{S}_r, \bar{T}_r$ terms in the case of Anderson-Rubin type of test and LM-like test will be helpful. Before providing the test statistics and limits we introduce assumptions and discuss them.

Assumptions:

1. $S_n = cn^\alpha$, where $0 < \alpha < 1/2$, $c > 0$, where c is a constant.

2. a)

$$\sup_i E\|x_i\|^{2\xi} < \infty,$$

b)

$$\sup_i E\|V_{2i}\|^{2\xi} < \infty,$$

c)

$$\sup_i E\|u_i\|^{2\xi} < \infty,$$

where $\xi > 2/(1-2\alpha)$.

3. Under the null hypothesis of $\delta = 0$, $v_{i1} = \beta_2' V_{i2} + u_i$,

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

where Σ_{xv1} is the variance matrix for $k \times 1$ Brownian Motion. This is also described in 4c below. $W_k(r)$ represents standard Brownian Motion.

Assume also

$$\sum_{i=1}^{[nr]} \frac{vec(x_i V_{i2}')}{n^{1/2}} \Longrightarrow B_{kl}(r) \equiv \Sigma_{xv2}^{1/2} W_{kl}(r),$$

where Σ_{xv2} is the variance matrix for $kl \times 1$ Brownian Motion $B_{kl}(r)$. This matrix is also explained in detail in Assumption 4c below. $W_{kl}(r)$ is $kl \times 1$ standard Brownian Motion.

4.

a) Uniformly over $r \in \Xi$

$$E[\frac{1}{n}\sum_{i=1}^{[nr]} x_i x_i'] \to r\Sigma_{xx},$$

Σ_{xx} is positive definite and finite.

b)Under the null hypotheses of $\delta = 0$, $v_{i1} = \beta_2' V_{i2} + u_i$,

$$\sup_{i,j} E\|x_i v_{i1} v_{j1} x_j'\| < \infty,$$

$$\sup_{s\in Z} E\|\frac{1}{nS_n}\sum_{j=1}^{n}\sum_{i=s}^{s+S_n} x_{i+j} v_{(i+j)1} v_{j1} x_j'\| = o(1).$$

c) In all the results in Assumption 4c here, the results are uniformly over r $\in \Xi$,

$$var[\frac{1}{n^{1/2}}\sum_{i=1}^{n} x_{ir} v_{i1}] = var[\frac{1}{n^{1/2}}\sum_{i=1}^{[nr]} x_i v_{i1}] \to r\Sigma_{xv1},$$

where Σ_{xv1} is nonsingular and finite. This is the long-run variance

$$\Sigma_{xv1} = \lim_{n\to\infty} var[\frac{1}{n^{1/2}}\sum_{i=1}^{n} x_i v_{i1}].$$

Specifically, also

$$var[\frac{1}{n^{1/2}}\sum_{i=1}^{n} vec(x_{ir}V'_{i2r})] = var[\frac{1}{n^{1/2}}\sum_{i=1}^{[nr]} vec(x_i V'_{i2})] \to r\Sigma_{xv2},$$

where Σ_{xv2} is nonsingular and finite. This is the long-run variance

$$\Sigma_{xv2} = \lim_{n\to\infty} var[\frac{1}{n^{1/2}}\sum_{i=1}^{n} vec(x_i V'_{i2})].$$

Furthermore let cov (a,b) represent the covariance between terms a and b,

$$cov[n^{-1/2}\sum_{i=1}^{n}(vec(x_{ir}V'_{i2r}),(v_{i1}x'_{ir}))] = cov[n^{-1/2}\sum_{i=1}^{[nr]}(vec(x_i V'_{i2}),(v_{i1}x'_i))] \to r\Sigma_{xv21},$$

where Σ_{xv21} is the limit for $covar[n^{-1/2}\sum_{i=1}^{n}([vec(x_i V'_{i2})],[v_{i1}x'_i])]$. This is the limit covariance between $vec(x_i V'_{i2})$ and $(v_{i1}x'_i)$, this is finite and has full rank.

d) Define for each r $\in \Xi$

$$\tilde{J}(h,m,r) = \frac{1}{n}\sum_{j=-n+1}^{n-1} k^*(j/S_n)\tilde{\Gamma}(h,m,r),$$

where $\tilde{\Gamma}(h,m,r) = \sum_{i=j+1}^{n} h_{ir}m'_{i-j,r}$ for $j \geq 0$, and $\sum_{i=-j+1}^{n} h_{i+j,r}m'_{i,r}$ for $j < 0$. "h" and "m" may denote different moments of data. $k^*(j/S_n) = 1 - |j/S_n|$ if $|j/S_n| \leq 2$ and 0 otherwise. This is the Bartlett kernel.

Set $h_{i,r} = x_{ir}v_{i1}, m_{i,r} = x_{ir}v_{i1}$ in $\tilde{J}(h,m,r)$. For $j \geq 0$, for $j < 0$ respectively

$$\tilde{J}(h,m,r) = \frac{1}{n}\sum_{j=-n+1}^{n-1} k^*(j/S_n)\sum_{i=j+1}^{n} x_{ir}v_{i1}v_{(i-j)1}x'_{(i-j),r} \quad (5)$$

$$= \frac{1}{n}\sum_{j=-n+1}^{n-1} k^*(j/S_n)\sum_{i=j+1}^{n} x_{(i+j)r}v_{(i+j)1}v_{i1}x'_{ir}. \quad (6)$$

Assume uniformly over $r \in \Xi$ using (5)(6)

$$\tilde{J}(h,m,r) - var[\frac{1}{n^{1/2}}\sum_{i=1}^{n} x_{ir}v_{i1}] \xrightarrow{p} 0.$$

Also setting $h_{i,r} = vec(x_{ir}V'_{i2r}), m_{i,r} = vec(x_{ir}V'_{i2r})$ in $\tilde{J}(h,m,r)$, assume uniformly over $r \in \Xi$

$$\tilde{J}(h,m,r) - var[n^{-1/2}\sum_{i=1}^{n} vec(x_{ir}V'_{i2r})] \xrightarrow{p} 0.$$

Same analysis applies to covariance between terms $vec(x_{ir}V'_{i2r})$ and $v_{i1}x'_{ir}$.

e) Uniformly over $r \in \Xi$, assume

$$\frac{1}{n}\sum_{i=1}^{[nr]} x_i x_i' - \frac{1}{n}E\sum_{i=1}^{[nr]} x_i x_i' \xrightarrow{p} 0.$$

5. Assume the following

a)

$$\sup_{i,j} E\|x_i x_i' x_j x_j'\| < \infty,$$

$$\sup_{s\in Z} E\|\frac{1}{nS_n}\sum_{j=1}^{n}\sum_{i=s}^{s+S_n} x_{i+j}x'_{i+j}x_j x_j'\| = o(1).$$

b) Set both h and m as $x_i x_i'$ in $\tilde{J}(h,m,r)$ definition and evaluate at full sample, $\tilde{J}(h,m,1) = \tilde{J}(h,m)$ to have

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

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

and nonsingular.

We discuss the assumptions here. Assumption 1 is a standard bandwidth assumption used in Guggenberger and Smith (2006). This is used for the analysis of weakly identified parameters with time series data (triangular array format is allowed). Assumption 2 is basically sufficient for (under the null hypotheses of $\delta = 0$, $v_{i1} = \beta_2' V_{i2} + u_i$)

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

Since $x_{ir} = x_i 1_{\{i\leq[nr]\}}$, then

$$\sup_r \sup_i \|x_{ir}v_{i1}\| \leq \sup_i \|x_i v_{i1}\|. \tag{7}$$

Via Assumption 2 we have

$$\sup_i E\|x_i v_{i1}\|^{\xi} < \infty,$$

then this is sufficient for

$$\sup_i \|x_i v_{i1}\| = o_p(S_n^{-1}n^{1/2}), \tag{8}$$

given Assumption 1. For this last point see p.17-18 of Guggenberger and Smith (2006).

Assumption 3 is a multivariate invariance principle. This invariance principle and sufficient conditions to obtain are given in Corollary 29.19 of Davidson (1994) or Lemma A.4 and p.849 of Andrews (1993). These involve triangular array of random variables that are L^2 NED (Near Epoch Dependent) of size -1/2. For these definitions of various time series data, Davidson (1994) is an excellent source.

Assumptions 4a and 4c are asymptotic covariance stationary conditions, used in structural change literature by Andrews (1993). Assumption 4b and 4d are used in weak instrument literature in time series data by p. 16-17 of Guggenberger and Smith (2006). Assumption 4d is an extension of that to structural change models and also used as Assumption 3 in Andrews (1993). Assumption 4e is Lemma A.3 of Andrews (1993). The sufficient conditions can be seen in that Lemma. Basically these are Assumption 2 here with data being L^2 NED of size -1/2.

Assumption 5 is an auxiliary assumption that is helpful in showing some remainder terms converge in probability to zero in variance-covariance matrix estimator.

We introduce the heteroskedasticity and autocorrelation consistent Anderson-Rubin test for structural change with weakly dependent data. This test is robust to identification problems, since it does not depend on Π. For each $r \in \Xi$,

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

where

$$(X_r^{*'} y_1)^w = (X_r' y_1)^w - (X_r' X)^w [(X'X)^w]^{-1} (X' y_1)^w. \tag{9}$$

Now we describe the terms in the above equation. Given $x_{ir} = x_i 1_{\{i \leq [nr]\}}$,

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

where $k(j/S_n)$ is the truncated kernel described above. The proofs follow for other kernels, but this is chosen for ease of notation. This kind of smoothing is used in Guggenberger and Smith (2006) and Otsu (2006). The other terms in (9) is smoothed in the same way including the matrix terms.

Note that we could have smoothed the moments in (9) differently. First, we could have calculated the terms on the right hand side, then we could have weighted the result. This results in the same asymptotics.

We define

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

We describe the terms in the variance-covariance matrix estimator now.

Under the null hypothesis of $\delta = 0$,

$$M_x y_1 = M_x v_1. \tag{10}$$

So

$$y_{i1} - x_i'(\sum_{i=1}^n x_i x_i')^{-1} \sum_{i=1}^n x_i y_{i1} = v_{i1} - x_i'(\sum_{i=1}^n x_i x_i')^{-1} \sum_{i=1}^n x_i v_{i1}. \tag{11}$$

Note also that $k \times 1$ vector in $M_x X_r$ is

$$x_{ir} - (\sum_{i=1}^n x_{ir} x_i')(\sum_{i=1}^n x_i x_i')^{-1} x_i. \tag{12}$$

First define the unsmoothed version, by (11)(12)

$$x_{ir}^* y_{i1} = x_{ir} v_{i1} - x_{ir} x_i' \bar{B} - \bar{A} x_i v_{i1} + \bar{A} x_i x_i' \bar{B}, \tag{13}$$

where

$$\bar{A} = (\sum_{i=1}^n x_{ir} x_i')(\sum_{i=1}^n x_i x_i')^{-1},$$

$$\bar{B} = (\sum_{i=1}^n x_i x_i')^{-1}(\sum_{i=1}^n x_i v_{i1}).$$

The smoothed version of the same vector is defined as

$$(x_{ir}^* y_{i1})^w = (x_{ir} v_{i1})^w - (x_{ir} x_i')^w \bar{B}^w - \bar{A}^w (x_i v_{i1})^w + \bar{A}^w (x_i x_i')^w \bar{B}^w, \tag{14}$$

where

$$\bar{A}^w = [\sum_{i=1}^n (x_{ir} x_i')^w][\sum_{i=1}^n (x_i x_i')^w]^{-1},$$

$$\bar{B}^w = [\sum_{i=1}^n (x_i x_i')^w]^{-1}[\sum_{i=1}^n (x_i v_{i1})^w].$$

Now we rewrite the Anderson-Rubin test statistic for structural change and analyze the limit behavior of the components of the test statistic.

Anderson-Rubin test statistic can be written as:

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

where

$$\bar{S}_{hac,r} = \hat{\Omega}_{11,r}^{-1/2} \frac{(X_r^{*\prime} y_1)^w}{2^{1/2} n^{1/2}}. \tag{15}$$

Note that we define

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

Before we derive the limit of the test we need the following results. The following is a new result in weak identification literature.

Lemma 1. *Under Assumptions 1-5, and under the null hypotheses of* $\delta = 0$,

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

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

Next we need to derive the limit for the variance-covariance matrix estimator $\hat{\Omega}_{11,r}$. The limit in the following is different than the limit found in Guggenberger and Smith (2006). This is due to analysis of unknown change points compared with standard variance covariance matrix estimation.

Lemma 2. *Under Assumptions 1-5, and under the null hypotheses of* $\delta = 0$, *uniformly over* $r \in \Xi$

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

We combine Lemma 1 and Lemma 2 in the AR test statistic to have the limit. This is one of the main theorems.

Theorem 1. *Under the null of no structural change and Assumptions 1-5,*

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

Remarks.

1. This result is new in the structural change literature. Formerly, Caner (2007) provide Anderson-Rubin based structural change test in nonlinear GMM with weakly identified parameters. However, that test is boundedly pivotal whereas this one is pivotal, so there will be no loss of power. This limit can be easily simulated. The limit critical values of our test can be obtained from Table 1 of Andrews (1993).

2. The reason that this version of the Anderson-Rubin test is pivotal but the one in Caner (2007) is not basically stems from the model and estimation. The sup AR test here depends on equation (1). So basically we test $\delta = \beta_1 - \beta_2 = 0$. Hence we can work under the null of $\delta = 0$. After some simple projection this results in a nuisance parameter free limit.

In the case of Caner (2007), we can not reparametrize the model as in (1) in the nonlinear case, and can not use the projection that we use. AR test in Caner (2007) uses constrained estimates (Under the null:$\beta_1 = \beta_2 = \beta$) $\tilde{\beta}$ which can be shown to be bound by a nuisance parameter free limit. The bound is given in Caner (2007) and it is

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

So the bound has an extra (independent from the Brownian Bridge) χ^2_k term compared to Theorem 1 here. With large number of instruments this may result in large losses of power.

3. The sup LM test of Andrews (1993) does not work in the case of weak identification. This is documented in Caner (2007).

4. Theorem 1 extends the weak instruments literature from standard inference to testing for regime shifts.

Now we handle the heteroskedasticity and autocorrelation corrected Lagrange multiplier test. The test statistic can be written as for each $r \in \Xi$,

$$LM = \bar{S}'_{hac,r} P_{\bar{T}_{hac,r}} \bar{S}_{hac,r}, \tag{16}$$

where

$$\bar{T}_{hac,r} = \hat{\Omega}_{11,r}^{-1/2} J. \tag{17}$$

where

$$J = n^{-1/2} \sum_{i=1}^{n} (x^*_{ir} Y'_{i2r})^w - \hat{\Omega}_{21,r} \hat{\Omega}_{11,r}^{-1} n^{-1/2} \sum_{i=1}^{n} (x^*_{ir} y_{i1})^w, \tag{18}$$

where

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

Before the test statistic we have to derive the limits of the terms in the test above. So first we want to obtain the limit for $n^{-1/2} \sum_{i=1}^{n} (x^*_{ir} Y'_{i2r})^w$. But we can rewrite that by (3) and linearity of smoothing operator that

$$n^{-1/2} \sum_{i=1}^{n} (x^*_{ir} Y'_{i2r})^w = n^{-1/2} \sum_{i=1}^{n} (x^*_{ir} x'_{ir})^w \Pi + n^{-1/2} \sum_{i=1}^{n} (x^*_{ir} V'_{i2r})^w. \tag{19}$$

We find the limit in the subsequent lemma.

Lemma 3. *Under Assumptions 1-5, and under the weak instrument asymptotics* $\Pi = C/n^{1/2}$,

$$n^{-1/2}\sum_{i=1}^{n} vec(x_{ir}^{*}Y_{i2r}')^{w} \Longrightarrow 2(1-r)[\Sigma_{xv2}W_{kl}(r) + rvec(\Sigma_{xx}C)].$$

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

Then we also want to analyze

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

First we want to clarify the expressions in the estimator. We simplify $x_{ir}^{*}Y_{i2r}'$ which is derived from the matrix $X_r' M_x M_x Y_{2r}$. See that $1 \times l$ row vector from the matrix $M_x Y_{2r}$ is

$$Y_{i2r}' - x_i'(\sum_{i=1}^{n} x_i x_i')^{-1}(\sum_{i=1}^{n} x_i Y_{i2r}').$$

Combine this with (12) to have

$$\begin{aligned} x_{ir}^{*}Y_{i2r}' &= x_{ir}Y_{i2r}' - x_{ir}x_i'(\sum_{i=1}^{n} x_i x_i')^{-1}(\sum_{i=1}^{n} x_i Y_{i2r}') \\ &- (\sum_{i=1}^{n} x_{ir}x_i')(\sum_{i=1}^{n} x_i x_i')^{-1} x_i Y_{i2r}' \\ &+ (\sum_{i=1}^{n} x_{ir}x_i')(\sum_{i=1}^{n} x_i x_i')^{-1} x_i x_i'(\sum_{i=1}^{n} x_i x_i')^{-1}(\sum_{i=1}^{n} x_i Y_{i2r}'). \end{aligned}$$

Define the following term:

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

The smoothed version of the above is:

$$\begin{aligned} (x_{ir}^{*}Y_{i2r}')^{w} &= (x_{ir}Y_{i2r}')^{w} - (x_{ir}x_i')^{w}\bar{F}^{w} \\ &- \bar{A}^{w}(x_i Y_{i2r}')^{w} + \bar{A}^{w}(x_i x_i')^{w}\bar{F}^{w}, \end{aligned} \tag{20}$$

where by Lemma 3, and (26)

$$\bar{F}^{w} = \left[\frac{\sum_{i=1}^{n}(x_i x_i')^{w}}{n}\right]^{-1}\left[\frac{\sum_{i=1}^{n}(x_i Y_{i2r}')^{w}}{n}\right] \xrightarrow{p} 0. \tag{21}$$

Now we provide the limit for the variance-covariance estimator, $\hat{\Omega}_{21,r}$.

Lemma 4. *Under Assumptions 1-5, uniformly over $r \in \Xi$, under weak instrument asymptotics*

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

where Σ_{xv21} is the limit for $covar[n^{-1/2} \sum_{i=1}^{n}([vec(x_i V_{i2}')], [v_{i1} x_i'])]$. This is the limit covariance between $vec(x_i V_{i2}')$ and $(v_{i1} x_i')$

We can now provide the Kleibergen (2002) type of LM test for structural change in the case of weak identification. This is one of the main results of the paper.

Theorem 2: *Under Assumptions 1-5, and under the null hypothesis of $\delta = 0$, with weak instrument asymptotics $\Pi = C/n^{1/2}$,*

$$\sup_{r \in \Xi} LM \xrightarrow{d} \sup_{r \in \Xi} \frac{[W_l(r) - rW_l(1)]'[W_l(r) - rW_l(1)]}{r(1-r)}.$$

Remarks.

1. The result is new, and provides nuisance parameter free limit in the case of structural change with weakly identified parameters.

2. This is the same limit that Andrews (1993) obtained when testing for structural change in standard identified models. However as explained in Theorem 1, that test statistic can not be used when there is an identification problem.

3. This is an extension of Guggenberger and Smith (2006) Kleibergen type of test to structural change test.

4. The limit depends on "l" number of parameters to be tested and may have better small sample properties than Anderson-Rubin test when the number of instruments are large.

5. In Caner (2007), a different version of LM type of test is proposed. This basically finds an estimate of the parameter by minimizing Kleibergen type of test given change point. Then sup of this Kleibergen test evaluated at this minimum is used (sup over unknown change points). This is not asymptotically pivotal but only boundedly pivotal. The distribution of the bound is given by Theorem 2 in Caner (2007) and

$$\sup_{r} \frac{[W_l(r) - rW_l(1)]'[W_l(r) - rW_l(1)]}{r(1-r)} + \chi_l^2.$$

So this is just χ_l^2 larger than our limit in Theorem 2. This loss of power may be important when the number of parameters are large. Table 2 of Caner (2007) compares the limit in Theorem 2 (same as Table 1 in Andrews (1993)) and the bound in an asymptotic exercise.

It has been found that even with two parameters to be tested this bound is conservative. We expect the limit in Theorem 2 to be useful compared to others that may be used.

6. Note that we could have smoothed the moments in (9) differently. First, we could have calculated the terms on the right hand side, then we could have weighted the result. This results in the same asymptotics.

4 Monte Carlo

In this section we setup a simple Monte Carlo exercise to measure size and power properties of the test statistics proposed in this article. We also compare them with boundedly pivotal Anderson-Rubin type and Kleibergen type of test proposed in Caner (2007). First, we consider the size of the test statistics. We generate the data according to a simple linear simultaneous equation model.

$$y_{i1} = Y_{i2}\beta + u_i, \tag{22}$$

$$Y_{i2} = X_i'\Pi + V_i,$$

where under the null of no structural change ($\beta_1 = \beta_2 = \beta$) $\beta = 0.5$. X_i is the $k \times 1$ instrument vector, Y_{i2} is the scalar endogenous variable. Also we have

$$u_i = 0.7u_{i-1} + \epsilon_i,$$

$$E\epsilon_i V_i = \Sigma_{\epsilon V} = \begin{bmatrix} 1 & 0.25 \\ 0.25 & 1 \end{bmatrix},$$

and ϵ_i, V_i are jointly normally distributed with zero mean and $\Sigma_{\epsilon V}$ as the covariance matrix. As u_i X_i follows the following AR(1) process

$$X_i = X_{i-1}'\rho + e_i,$$

where $\rho = (0.5, 0.5, \cdots, 0.5)'$ is $k \times 1$ vector. e_i is the standard normal distributed and independent of X_i, Y_{i2}, V_i, u_i.

This setup is very similar to the one considered in Guggenberger and Smith (2006). The sample size we consider is $n = 100$. We analyze several cases of interest such as $\Pi = .1, k = 2$ (weak identification with limited number of instruments), $\Pi = .1, k = 5$ (weak identification with many number of instruments). The others that we consider are $\Pi = 1, k = 2$ (standard identification, limited number of instruments), $\Pi = 1, k = 5$ (standard identification with limited number of instruments). Specifically, we analyze $\sup AR, \sup LM$ tests that are shown in section 3. The limits of these tests are in Theorems 1 and 2. Now we describe the boundedly pivotal tests in Caner (2007). We denote these by $\sup S, \sup K$. The boundedly pivotal Anderson-Rubin test can be described in several steps. First, define for each $r \in \Xi$

$$\tilde{\beta}(r) = arg \min_{\beta \in B} S_n(\beta, r),$$

where B is a compact subset of the real line in this exercise,

$$\begin{aligned} S_n(\beta,r) &= [n^{-1/2}\sum_{i=1}^{[nr]}\psi_i(\beta)]'[\frac{\hat{V}_1(\beta,r)^{-1}}{r}][n^{-1/2}\sum_{i=1}^{[nr]}\psi_i(\beta)] \\ &+ [n^{-1/2}\sum_{i=[nr]+1}^{n}\psi_i(\beta)]'[\frac{\hat{V}_2(\beta,r)^{-1}}{1-r}][n^{-1/2}\sum_{i=[nr]+1}^{n}\psi_i(\beta)], \end{aligned}$$

where $\psi_i(\beta) = (y_{i1} - Y_{i2}\beta)$. Then by $\bar{\psi}_1 = \sum_{i=1}^{[nr]} \psi_i(\beta)/[nr]$, $\bar{\psi}_2 = \sum_{i=[nr]+1}^{n} \psi_i(\beta)$. So define

$$\begin{aligned} \hat{V}_1(\beta,r) &= \sum_{v=0}^{[nr]-1} k(v/l_n)\frac{1}{[nr]}\sum_{i=v+1}^{[nr]}(\psi_i(\beta)-\bar{\psi}_1)(\psi_{i-v}(\beta)-\bar{\psi}_1)' \\ &+ \sum_{v=1}^{[nr]-1} k(v/l_n)\frac{1}{[nr]}\sum_{i=v+1}^{[nr]}(\psi_{i-v}(\beta)-\bar{\psi}_1)(\psi_i(\beta)-\bar{\psi}_1)', \end{aligned}$$

$$\begin{aligned} \hat{V}_2(\beta,r) &= \sum_{v=[nr]}^{n} k(v/l_n)\frac{1}{n-[nr]}\sum_{i=v+1}^{n}(\psi_i(\beta)-\bar{\psi}_2)(\psi_{i-v}(\beta)-\bar{\psi}_2)' \\ &+ \sum_{v=[nr]}^{n} k(v/l_n)\frac{1}{[nr]}\sum_{i=v+1}^{n}(\psi_{i-v}(\beta)-\bar{\psi}_2)(\psi_i(\beta)-\bar{\psi}_2)', \end{aligned}$$

and $k(v/l_n)$ is the Bartlett kernel. This is the one described in Newey and West (1987),

$$k(v/l_n) = 1 - v/(l_n + 1), \quad for \quad v \leq l_n,$$

and 0 otherwise. This is also used in all test statistics considered here. According to Andrews (1991), with $n = 100$ the optimal $l_n = 4$. This value is used for all test statistics analyzed.

The boundedly pivotal Anderson-Rubin type of test is:

$$\sup_r S_n(\tilde{\beta}(r), r).$$

This tests the null of $H_0 : \beta_1 = \beta_2$, in contrast to reparametrized but the same null in this article $H_0 : \delta = 0$.

To describe the boundedly pivotal Kleibergen test in Caner (2007) we need

$$\bar{\beta}(r) = arg \min_{\beta \in B} K_n(\beta, r),$$

where

$$\begin{aligned} K_n(\beta,r) &= [[nr]^{-1/2}\sum_{i=1}^{[nr]}\psi_i(\beta)]'[\hat{V}_1(\beta,r)]^{-1/2}P_{[\hat{V}_1(\beta,r)]^{-1/2}\tilde{D}_1}[\hat{V}_1(\beta,r)]^{-1/2}[[nr]^{-1/2}\sum_{i=1}^{[nr]}\psi_i(\beta)] \\ &+ [(n-[nr])^{-1/2}\sum_{i=[nr]+1}^{n}\psi_i(\beta)]'[\hat{V}_2(\beta,r)]^{-1/2}P_{[\hat{V}_2(\beta,r)]^{-1/2}\tilde{D}_2}[\hat{V}_2(\beta,r)]^{-1/2} \\ &\times [(n-[nr])^{-1/2}\sum_{i=[nr]+1}^{n}\psi_i(\beta)], \end{aligned}$$

where

$$\tilde{D}_1 = (n^{-1/2}\sum_{i=1}^{[nr]} X_i Y_{i2}) - \hat{V}_{q1}(\beta,r)[\hat{V}_1(\beta,r)]^{-1}(n^{-1/2}\sum_{i=1}^{[nr]}\psi_i(\beta)),$$

$$\begin{aligned} \hat{V}_{q1}(\beta,r) &= \sum_{v=0}^{[nr]-1} k(v/l_n)\frac{1}{[nr]}\sum_{i=v+1}^{[nr]}(X_iY_{i2}-\bar{\psi}_{x1})(\psi_{i-v}(\beta)-\bar{\psi}_1)' \\ &+ \sum_{v=1}^{[nr]-1} k(v/l_n)\frac{1}{[nr]}\sum_{i=v+1}^{[nr]}(\psi_{i-v}(\beta)-\bar{\psi}_1)(X_iY_{i2}-\bar{\psi}_{x1})', \end{aligned}$$

where $\bar{\psi}_{x1} = \sum_{i=1}^{[nr]} X_i Y_{i2}/[nr]$. $\tilde{D}_2, \hat{V}_{q2}$ is defined accordingly replacing the first part of the sample with second part of the sample. The test statistic is

$$\sup_r K_n(\bar{\beta}, r).$$

This test is used for $H_0 : \beta_1 = \beta_2$, rather than $H_0 : \delta = 0$.

We use 1000 iterations for the tests here. We report the rejection rate of the true null of $H_0 : \beta_1 = \beta_2 = 0.5$ (boundedly pivotal tests) or reparametrized version of $H_0 : \delta = 0$ (for tests that are considered in this article). The critical values for our sup AR, sup LM tests are 11.79, 8.85 at 5% level which are obtained from Table 1 of Andrews (1993) for $k = 2$, and one parameter respectively. When $k = 5$ the sup AR test critical value is 18.35 at 5% level. For the boundedly pivotal tests , the AR type test has critical values of 14.8, 25.4 at 5% level in Table 1 of Caner (2007) for $k = 2, k = 5$ respectively. For the boundedly pivotal Kleibergen type of test the 5% critical value is 10.4 from Table 1 of Caner (2007) corresponding to one parameter. The limits for the boundedly pivotal tests are explained in the remarks after Theorems 1 and 2. The results of the size exercise are in Table 1. When we look at the case of weak identification with limited number of instruments, boundedly pivotal Kleibergen test has 6.3% size. This comes closest to nominal size of 5%. The Anderson-Rubin type test that is proposed here underrejects severely at 0%. When we increase the number of instruments to 5 then we see that sup LM test gives the best result at 0.8%. The boundedly pivotal tests have massive size distortions. Their respective sizes are 54.1%, 32.5% for sup $S_n(\tilde{\beta}(r), r), supK_n(\bar{\beta}(r), r)$. This is puzzling since these tests are boundedly pivotal. However, these boundedly pivotal tests have distributions bounded in the limit. Caner (2007) also shows that in small samples they are oversized as documented in this article.

We also calculated power and size adjusted power of the tests analyzed in the cases of interest outlined above. The size adjusted power results are shown in Tables 2a-b. $\delta = \beta_1 - \beta_2$, and δ takes the values of 5,1,-1, -5 respectively. In other words, we generate half of the observations in regime 1, and the other in the second regime 2. For example, for $\delta = 5$ case, the first 50 observations are generated with $\beta_1 = 2.5$ and the next 50 observations are generated by using $\beta_2 = -2.5$. Equation (22) is used with different β in each regime. The reduced form equation is the same in the size exercise. When $\delta = 1, \beta_1 = 0.5, \beta_2 = -0.5$, for the negative δ numbers β's reverse. Each case in the size adjusted power are explained at the head of the Table, $\Pi = .1, k = 2$ represents weak identification with small number of instruments, and $\Pi = .1, k = 5$ represents the case with weak identification and large number of instruments. This is Table 2a. Table 2b is done in the same way but with strong identification.

In Table 2a we see that the power of the tests are very similar, they almost have no power. This is because of the weak identification of the parameters in that setup. In the strong identification case in Table 2b, the AR based tests do well compared to Kleibergen type of test.

5 Conclusion

This paper introduces tests for structural change for the structural parameters in a simultaneous equations. When there is weak identification in the parameters of the reduced form system as introduced in Staiger and Stock (1997) the existing tests are not asymptotically pivotal. This is shown recently by Caner (2007). Caner (2007) also introduced structural

Table 1: Size at 5% level

Tests	$\Pi = .1, k = 2$	$\Pi = 1, k = 2$	$\Pi = .1, k = 5$	$\Pi = 1, k = 5$
$\sup AR$	0.0	0.0	0.0	0.0
$\sup LM$	24.1	11.9	0.8	0.4
$\sup S_n(\tilde{\beta}(r), r)$	6.8	26.4	54.2	78.1
$\sup K_n(\bar{\beta}(r), r)$	6.3	18.9	32.5	54.3

Note: $n = 100$, *1000 iterations are used to generate the table. The critical values are explained in Monte Carlo section. sup AR and sup LM tests refer to tests in Theorems 1 and 2.* $supS_n(\tilde{\beta}(r), r), supK_n(\bar{\beta}(r), r)$ *are the boundedly pivotal test statistics that are explained in Caner (2007).*

Table 2a: Size Adjusted Power

	$\Pi = .1, k = 2$				$\Pi = .1, k = 5$			
Tests	$\delta = 1$	$\delta = 5$	$\delta = -5$	$\delta = -1$	$\delta = 1$	$\delta = 5$	$\delta = -5$	$\delta = -1$
$\sup AR$	3.9	7.1	6.8	4.7	5.1	7.0	7.0	5.8
$\sup LM$	4.6	10.3	9.0	7.0	5.2	8.1	7.5	5.6
$\sup S_n(\tilde{\beta}(r), r)$	5.7	12.3	11.4	6.6	13.1	16.6	12.8	13.2
$\sup K_n(\bar{\beta}(r), r)$	5.9	11.1	9.5	5.2	3.3	7.1	6.4	6.1

Note: $n = 100$, *1000 iterations are used to generate the table. The critical values can be obtained from the author on demand. sup AR and sup LM tests refer to tests in Theorems 1 and 2.* $supS_n(\tilde{\beta}(r), r), supK_n(\bar{\beta}(r), r)$ *are the boundedly pivotal test statistics that are explained in Caner (2007).* Π *is the reduced form coefficient vector, where 0.1 is the value in each cell of the vector. This is the weak identification case. K represents number of instruments.*

Table 2b: Size Adjusted Power

	$\Pi = 1, k = 2$				$\Pi = 1, k = 5$			
Tests	$\delta = 1$	$\delta = 5$	$\delta = -5$	$\delta = -1$	$\delta = 1$	$\delta = 5$	$\delta = -5$	$\delta = -1$
$\sup AR$	32.1	81.6	79.8	36.5	44.2	61.3	64.9	44.5
$\sup LM$	46.5	78.0	75.1	46.4	13.4	29.1	28.5	9.9
$\sup S_n(\tilde{\beta}(r), r)$	40.9	85.9	86.9	42.5	31.4	62.3	64.7	32.0
$\sup K_n(\bar{\beta}(r), r)$	41.0	79.8	78.9	41.5	11.8	28.8	31.8	12.1

Note: $n = 100$, *1000 iterations are used to generate the table. The critical values can be obtained from the author on demand. sup AR and sup LM tests refer to tests in Theorems 1 and 2.* $supS_n(\tilde{\beta}(r), r), supK_n(\bar{\beta}(r), r)$ *are the boundedly pivotal test statistics that are explained in Caner (2007).* Π *is the reduced form coefficient vector, where 0.1 is the value in each cell of the vector. This is the weak identification case. K represents number of instruments.*

change tests which are robust to identification problems. But these are boundedly pivotal and hence may suffer from low power. The tests in this paper are different from the tests in Caner (2007). First of all, they are asymptotically pivotal. Then also takes into account time series nature of the data better into account. They are also robust to identification problems. A reparametrization of the model provides the way for building similar tests.

APPENDIX

Proof of Lemma 1. We consider the limit behavior, under the null of $\delta = 0$,

$$\frac{(X_r^{*\prime} y_1)^w}{n^{1/2}} = \frac{(X_r' v_1)^w}{n^{1/2}} - \frac{(X_r' X)^w}{n}[\frac{(X'X)^w}{n}]^{-1}\frac{(X'v_1)^w}{n^{1/2}}. \quad (23)$$

In (23) we analyze the first term on the right-hand side. The proof follows closely Lemma 1 in Guggenberger and Smith (2006). First, we know that $x_{ir} = x_i 1_{\{i \leq [nr]\}}$. Then use the definition of the truncated kernel

$$n^{-1}\sum_{i=1}^{n}(x_{ir}v_{i1})^w = n^{-1}\sum_{i=1}^{n} S_n^{-1} \sum_{j=max(i-n,-S_n)}^{min(i-1,S_n)} x_{(i-j)r}v_{(i-j)1}.$$

Then following exactly the same steps in the proof of Lemma 1 (i.e. equation (A.6)) in Guggenberger and Smith (2006),

$$\begin{aligned} n^{-1}\sum_{i=1}^{n}(x_{ir}v_{i1})^{w} &= 2n^{-1}\sum_{i=1}^{n}x_{ir}v_{i1}+n^{-1}S_n^{-1}\sum_{i=S_n+1}^{n-S_n}x_{ir}v_{i1} \\ &+ n^{-1}S_n^{-1}\sum_{i=1}^{S_n}(i-S_n)x_{ir}v_{i1}+n^{-1}S_n^{-1}\sum_{i=n-S_n+1}^{n}(-S_n+n-i+1)x_{ir}v_{i1} \\ &= 2n^{-1}\sum_{i=1}^{n}x_{ir}v_{i1}+o_p(n^{-1/2}), \end{aligned}$$

where the remainder is obtained via Assumption 3 and (7)(8).

Then

$$n^{-1/2}\sum_{i=1}^{n}(x_{ir}v_{i1})^{w}=2n^{-1/2}\sum_{i=1}^{n}x_{ir}v_{i1}+o_p(1)=2n^{-1/2}\sum_{i=1}^{[nr]}x_{i}v_{i1}+o_p(1),$$

by definition of $x_{ir}=x_i 1_{\{i\leq[nr]\}}$. Then use Assumption 3 to have

$$n^{-1/2}\sum_{i=1}^{n}(x_{ir}v_{i1})^{w}\Longrightarrow 2\Sigma_{xv1}^{1/2}W_k(r). \tag{24}$$

Next given Assumption 2a and 4a, 4e by applying Lemma 1 in Guggenberger and Smith (2006) for nonzero mean random variables, uniformly in r

$$n^{-1}\sum_{i=1}^{n}(x_{ir}x_i')^{w}\xrightarrow{p}2r\Sigma_{xx}. \tag{25}$$

Then the full sample versions of (24)(25) provides

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

$$n^{-1}\sum_{i=1}^{n}(x_{i}x_i')^{w}\xrightarrow{p}2\Sigma_{xx}. \tag{26}$$

Then combine these in (23) to have, under the null hypotheses of $\delta=0$ to have the desired result.**Q.E.D**

Proof of Lemma 2. First we analyze $\bar{A}^w, \bar{B}^w$ in (14).

$$\bar{A}^{w}=[\sum_{i=1}^{n}(x_{ir}x_i')^{w}][\sum_{i=1}^{n}(x_{i}x_i')^{w}]^{-1},$$

Then by (25)(26), uniformly over r

$$\bar{A}^w \xrightarrow{p} rI_k. \tag{27}$$

Since

$$\bar{B}^w = [\sum_{i=1}^{n}(x_i x_i')^w]^{-1}[\sum_{i=1}^{n}(x_i v_{i1})^w],$$

given $Ex_i v_{i1} = 0$, and Lemma 1 in Guggenberger and Smith (2006) provides

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

and

$$\bar{B}^w \xrightarrow{p} 0. \tag{28}$$

Equations (27) and (28) provide very useful two results that will be used in subsequent proofs. Now, given (14) we rewrite

$$\begin{aligned}\hat{\Omega}_{11,r} &= \frac{S_n}{n}\sum_{i=1}^{n}[(x_{ir}v_{i1})^w - (x_{ir}x_i')^w\bar{B}^w - \bar{A}^w(x_i v_{i1})^w + \bar{A}^w(x_i x_i')^w\bar{B}^w] \\ &\times \; [(x_{ir}v_{i1})^w - (x_{ir}x_i')^w\bar{B}^w - \bar{A}^w(x_i v_{i1})^w + \bar{A}^w(x_i x_i')^w\bar{B}^w]'. \end{aligned} \tag{29}$$

Furthermore (29) can be decomposed into several terms.

$$\begin{aligned}\hat{\Omega}_{11,r} &= \frac{S_n}{n}\sum_{i=1}^{n}(x_{ir}v_{i1})^w(v_{i1}x_{ir}')^w - [\frac{S_n}{n}\sum_{i=1}^{n}(x_{ir}v_{i1})^w(v_{i1}x_i')^w]\bar{A}^{w\prime} \\ &- \; \bar{A}^w[\frac{S_n}{n}\sum_{i=1}^{n}(x_i v_{i1})^w(v_{i1}x_{ir}')^w] + \bar{A}^w[\frac{S_n}{n}\sum_{i=1}^{n}(x_i v_{i1})^w(v_{i1}x_i')^w]\bar{A}^{w\prime} + o_p(1). \end{aligned} \tag{30}$$

In (30), at the end of this proof we explain the asymptotically negligible remainder term. So we consider the first term on the right-hand side of (30). First note that we can apply Lemma 2 in Guggenberger and Smith (2006). However, this has to be done uniformly over r in addition to Lemma 2 in Guggenberger and Smith (2006). So following the proof of Lemma 2 in Guggenberger and Smith (2006) we need

$$\sup_{r\in\Xi}\sup_{i,j} E\|x_{ir}v_{i1}v_{j1}x_{jr}'\| < \infty,$$

and

$$\sup_{r\in\Xi}\sup_{s\in Z} E\|\frac{1}{nS_n}\sum_{j=1}^{n}\sum_{i=s}^{s+S_n} x_{(i+j)r}v_{(i+j)1}v_{j1}x_{jr}'\| = o(1).$$

One important fact to remember is $x_{ir} = x_i 1_{\{i \le [nr]\}}$ and then

$$\sup_{r\in\Xi}\sup_{i,j} E\|x_{ir}v_{i1}v_{j1}x'_{jr}\| \le \sup_{i,j} E\|x_i v_{i1}v_{j1}x'_j\|.$$

$$\sup_{r\in\Xi}\sup_{s\in Z} E\|\frac{1}{nS_n}\sum_{j=1}^{n}\sum_{i=s}^{s+S_n} x_{(i+j)r}v_{(i+j)1}v_{j1}x'_{jr}\| \le \sup_{s\in Z} E\|\frac{1}{nS_n}\sum_{j=1}^{n}\sum_{i=s}^{s+S_n} x_{(i+j)}v_{(i+j)1}v_{j1}x'_j\|.$$

Apply Assumption 4b and use Lemma 2 of Guggenberger and Smith (2006) to have

$$\frac{S_n}{n}\sum_{i=1}^{n}(x_{ir}v_{i1})^w(v_{i1}x'_{ir})^w - 2\tilde{J}_n(x_{ir}v_{i1}, x_{ir}v_{i1}, r) \xrightarrow{p} 0,$$

uniformly over r $\in \Xi$. $\tilde{J}_n(x_{ir}v_{i1}, x_{ir}v_{i1}, r)$ is described in (5)(6). Next apply Assumption 4d, to have uniformly over r

$$\tilde{J}_n(x_{ir}v_{i1}, x_{ir}v_{i1}, r) \xrightarrow{p} r\Sigma_{xv1} < \infty.$$

Combining the last two results

$$\frac{S_n}{n}\sum_{i=1}^{n}(x_{ir}v_{i1})^w(v_{i1}x'_{ir})^w \xrightarrow{p} 2r\Sigma_{xv1}. \tag{31}$$

Then in the same manner but also benefiting from (27)

$$\bar{A}^w[\frac{S_n}{n}\sum_{i=1}^{n}(x_i v_{i1})^w(v_{i1}x'_{ir})^w] \xrightarrow{p} 2r^2\Sigma_{xv1}. \tag{32}$$

$$\bar{A}^w[\frac{S_n}{n}\sum_{i=1}^{n}(x_i v_{i1})^w(v_{i1}x'_i)^w]\bar{A}^{w\prime} \xrightarrow{p} 2r^2\Sigma_{xv1}. \tag{33}$$

Substituting these into (30) uniformly over r

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

We can analyze how the asymptotically negligible remainder term is obtained in (30). We show that in (29) all the terms involving $\bar{B}^w$ are $o_p(1)$. We provide the proof now. We consider one of those cross-product terms in (29)

$$\frac{S_n}{n}\sum_{i=1}^{n}(x_{ir}v_{i1})^w\bar{B}^{w\prime}(x_i x'_{ir})^w \le [\frac{S_n}{n}\sum_{i=1}^{n}\|(x_{ir}v_{i1})^w\|^2]^{1/2}[\frac{S_n}{n}\sum_{i=1}^{n}\|(x_i x'_{ir})^w\|^2]^{1/2}\|\bar{B}^w\|. \tag{34}$$

First we know that by (28)

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

Then see that

$$\|(x_{ir}v_{i1})^w\| \leq \|(x_i v_{i1})^w\|. \tag{35}$$

$$\|(x_i x'_{ir})^w\| \leq \|(x_i x'_i)^w\|. \tag{36}$$

Then by Assumption 4b, we can obtain Lemma 2 in Guggenberger and Smith (2006) and then applying Assumption 4d for the full sample $r = 1$, we have

$$\frac{S_n}{n}\sum_{i=1}^{n}\|(x_{ir}v_{i1})^w\|^2 = O_p(1). \tag{37}$$

Next by Assumption 5a we obtain Lemma 2 of Guggenberger and Smith (2006). Then apply Assumptions 5b and 5c to have

$$\frac{S_n}{n}\sum_{i=1}^{n}\|(x_i x_{ir})^w\|^2 = O_p(1). \tag{38}$$

Use these results in (34) to have

$$\frac{S_n}{n}\sum_{i=1}^{n}(x_{ir}v_{i1})^w \bar{B}^{w\prime}(x_i x'_{ir})^w \xrightarrow{p} 0, \tag{39}$$

uniformly over r $\in \Xi$. The same method applies to all terms with $\bar{B}^w$ and they converge in probability to zero as well. So we have the asymptotically negligible remainder in (30).**Q.E.D**

Proof of Lemma 3. In (19) consider

$$n^{-1/2}\sum_{i=1}^{n}(x^*_{ir}V'_{i2r})^w = [n^{-1/2}\sum_{i=1}^{n}(x_{ir}V'_{i2r})^w] - [n^{-1}\sum_{i=1}^{n}(x_{ir}x'_i)^w][n^{-1}\sum_{i=1}^{n}(x_i x'_i)^w]^{-1}[n^{-1/2}\sum_{i=1}^{n}(x_i V'_{i2r})^w].$$

Note that the first and the last terms are the same since $x_{ir}V'_{i2r} = x_i V_{i2r}$. Then we use the same analysis in the proof of Lemma 1, equations (24)-(26) to have

$$n^{-1/2}\sum_{i=1}^{n}vec(x^*_{ir}V_{i2r})^w \Longrightarrow 2(1-r)\Sigma^{1/2}_{xv2}W_{kl}(r), \tag{40}$$

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

Analyze in (19), by assuming weak instrument asymptotics $\Pi = C/n^{1/2}$.

$$\begin{aligned} n^{-1/2}\sum_{i=1}^{n}(x_{ir}^{*}x_{ir}')^{w}\Pi &= n^{-1}\sum_{i=1}^{n}(x_{ir}^{*}x_{ir}')^{w}C \\ &= n^{-1}\sum_{i=1}^{n}(x_{ir}x_{ir}')^{w}C - (n^{-1}\sum_{i=1}^{n}(x_{ir}x_{i}')^{w})\bar{A}^{w} \\ &\times\ (n^{-1}\sum_{i=1}^{n}(x_{i}x_{i}')^{w})^{-1}(n^{-1}\sum_{i=1}^{n}(x_{i}x_{ir}')^{w})C \\ &\xrightarrow{p}\ 2r(1-r)\Sigma_{xx}C, \end{aligned}$$

by (27)(25)(26). So

$$n^{-1/2}\sum_{i=1}^{n}vec(x_{ir}^{*}x_{ir}')^{w}\Pi \xrightarrow{p} 2r(1-r)vec(\Sigma_{xx}C). \tag{41}$$

So substitute (41)(40) in (19) to have the desired result.**Q.E.D**.

Proof of Lemma 4. $\hat{\Omega}_{21,r}$ can be rewritten using (20)(14)

$$\begin{aligned} \hat{\Omega}_{21,r} &= \frac{S_n}{n}\sum_{i=1}^{n}vec(x_{ir}Y_{i2r}')^{w}(v_{i1}x_{ir}')^{w} - \frac{S_n}{n}\sum_{i=1}^{n}vec(x_{ir}Y_{i2r}')^{w}\bar{B}^{w\prime}(x_{i}x_{ir}')^{w} \\ &- \frac{S_n}{n}\sum_{i=1}^{n}vec(x_{ir}Y_{i2r}')^{w}(v_{i1}x_{i}')^{w}\bar{A}^{w\prime} + \frac{S_n}{n}\sum_{i=1}^{n}vec(x_{ir}Y_{i2r}')^{w}\bar{B}^{w\prime}(x_{i}x_{ir}')^{w}\bar{A}^{w} \\ &- \frac{S_n}{n}\sum_{i=1}^{n}vec[(x_{ir}x_{i}')^{w}\bar{F}^{w}](v_{i1}x_{ir})^{w} + \frac{S_n}{n}\sum_{i=1}^{n}vec[(x_{ir}x_{i}')^{w}\bar{F}^{w}]\bar{B}^{w\prime}(x_{i}x_{ir}')^{w} \\ &+ \frac{S_n}{n}\sum_{i=1}^{n}vec[(x_{ir}x_{i}')^{w}\bar{F}^{w}](v_{i1}x_{i}')^{w}\bar{A}^{w\prime} - \frac{S_n}{n}\sum_{i=1}^{n}vec[(x_{ir}x_{i}')^{w}\bar{F}^{w}]\bar{B}^{w\prime}(x_{i}x_{i}')^{w}\bar{A}^{w\prime} \\ &- \frac{S_n}{n}\sum_{i=1}^{n}vec[\bar{A}^{w}(x_{i}Y_{i2r}')^{w}](v_{i1}x_{ir}')^{w} + \frac{S_n}{n}\sum_{i=1}^{n}vec[\bar{A}^{w}(x_{i}Y_{i2r}')^{w}]\bar{B}^{w\prime}(x_{i}x_{ir}')^{w} \\ &+ \frac{S_n}{n}\sum_{i=1}^{n}vec[\bar{A}^{w}(x_{i}Y_{i2r}')^{w}](v_{i1}x_{i}')^{w}\bar{A}^{w\prime} - \frac{S_n}{n}\sum_{i=1}^{n}vec[\bar{A}^{w}(x_{i}Y_{i2r}')^{w}]\bar{B}^{w\prime}(x_{i}x_{i}')^{w}\bar{A}^{w\prime} \\ &+ \frac{S_n}{n}\sum_{i=1}^{n}vec[\bar{A}^{w}(x_{i}x_{i}')^{w}\bar{F}^{w}](v_{i1}x_{ir}')^{w} - \frac{S_n}{n}\sum_{i=1}^{n}vec[\bar{A}^{w}(x_{i}x_{i}')^{w}\bar{F}^{w}]\bar{B}^{w\prime}(x_{i}x_{ir}')^{w} \\ &- \frac{S_n}{n}\sum_{i=1}^{n}vec[\bar{A}^{w}(x_{i}x_{i}')^{w}\bar{F}^{w}](v_{i1}x_{i}')^{w}\bar{A}^{w\prime} + \frac{S_n}{n}\sum_{i=1}^{n}vec[\bar{A}^{w}(x_{i}x_{i}')^{w}\bar{F}^{w}]\bar{B}^{w\prime}(x_{i}x_{i}')^{w}\bar{A}^{w\prime}. \end{aligned}$$

Taking into account

$$Y_{2r} = X_r\Pi + V_{2r},$$

and under weak instrument asymptotics we have

$$Y_{2r} = X_r \frac{C}{n^{1/2}} + V_{2r}.$$

Benefiting from (21) and (28) proceeding exactly as in (35)-(39) (using Cauchy-Schwartz inequality) uniformly over $r \in \Xi$ we see that all terms involving $\bar{B}^w, \bar{F}^w$ converge in probability to zero. Then we have the following expression:

$$\begin{aligned}
\hat{\Omega}_{21,r} &= \frac{S_n}{n}\sum_{i=1}^{n} vec(x_{ir}Y'_{i2r})^w (v_{i1}x'_{ir})^w - \frac{S_n}{n}\sum_{i=1}^{n} vec(x_{ir}Y'_{i2r})^w (v_{i1}x'_{i})^w \bar{A}^{w'} \\
&- \frac{S_n}{n}\sum_{i=1}^{n} vec[\bar{A}^w (x_i Y'_{i2r})^w](v_{i1}x'_{ir})^w + \frac{S_n}{n}\sum_{i=1}^{n} vec[\bar{A}^w (x_i Y'_{i2r})^w](v_{i1}x'_{i})^w \bar{A}^{w'} + o_p(1).
\end{aligned}$$

We begin analyzing the first term on the right hand side of the equation above.

$$\frac{S_n}{n}\sum_{i=1}^{n} vec(x_{ir}Y'_{i2r})^w (v_{i1}x'_{ir})^w = \frac{S_n}{n}\sum_{i=1}^{n} vec(x_{ir}x'_{ir}\Pi)^w (v_{i1}x'_{ir})^w + \frac{S_n}{n}\sum_{i=1}^{n} vec(x_{ir}V'_{i2r})^w (v_{i1}x'_{ir})^w.$$

Under weak instrument asymptotics ($\Pi = C/n^{1/2}$)

$$\frac{S_n}{n}\sum_{i=1}^{n} vec(x_{ir}Y'_{i2r})^w (v_{i1}x'_{ir})^w = \frac{1}{n^{1/2}}\frac{S_n}{n}\sum_{i=1}^{n} vec(x_{ir}x'_{ir}C)^w (v_{i1}x'_{ir})^w + \frac{S_n}{n}\sum_{i=1}^{n} vec(x_{ir}V'_{i2r})^w (v_{i1}x'_{ir})^w.$$

First we consider the following term as in (31),

$$\frac{S_n}{n}\sum_{i=1}^{n} vec(x_{ir}V'_{i2r})^w (v_{i1}x'_{ir})^w \xrightarrow{p} 2r\Sigma_{xv21},$$

uniformly over $r \in \Xi$. Note that in the same manner, uniformly over $r \in \Xi$

$$\frac{1}{n^{1/2}}\frac{S_n}{n}\sum_{i=1}^{n} vec(x_{ir}x'_{ir}C)^w (v_{i1}x'_{ir})^w \xrightarrow{p} 0.$$

Combining these we have, uniformly over $r \in \Xi$

$$\frac{S_n}{n}\sum_{i=1}^{n} vec(x_{ir}Y'_{i2r})^w (v_{i1}x'_{ir})^w \xrightarrow{p} 2r\Sigma_{xv21}. \tag{42}$$

Then in the same manner as in (32)(33) (27), uniformly over $r \in \Xi$

$$\frac{S_n}{n}\sum_{i=1}^{n} vec(x_{ir}Y'_{i2r})^w (v_{i1}x'_{i})^w \bar{A}^{w'} \xrightarrow{p} 2r^2\Sigma_{xv21}.$$

$$\frac{S_n}{n}\sum_{i=1}^{n} vec[\bar{A}^w(x_{ir}Y'_{i2r})^w](v_{i1}x'_i)^w\bar{A}^{w'} \xrightarrow{p} 2r^3\Sigma_{xv21}.$$

In the results above we use

$$\frac{S_n}{n}\sum_{i=1}^{n} vec(x_{ir}Y'_{i2r})^w(v_{i1}x'_i)^w \xrightarrow{p} 2r\Sigma_{xv21},$$

which is basically obtained in the same way as (42). So combining these last three results we obtain the desired result.**Q.E.D**

Proof of Theorem 2. The proof relies on showing $(X_r^{*'}y_1)^w$ and J are asymptotically independent (i.e. $\bar{S}_{hac,r}$ and $\bar{T}_{hac,r}$ (15)(17) are asymptotically independent). Note that under the weak instruments assumption $\Pi = C/n^{1/2}$, where C is a $k \times l$ constant matrix.

We can prove the asymptotic independence of J and $(X_r^{*'}y_1)^w$. Note that by Lemma 3, 4 and (41)

$$\begin{aligned} vec(J,(n^{-1/2}(X_r^{*'}y_1)^w) &= r(1-r)vec(\Sigma_{xx}C,0) \\ &+ \begin{bmatrix} I_{kl} & -(1-r)\Sigma_{xv21}\Sigma_{xv1}^{-1} \\ 0 & I_k \end{bmatrix}\begin{bmatrix} n^{-1/2}vec(X_r^{*'}V_{2r})^w \\ n^{-1/2}(X_r^{*'}y_1)^w \end{bmatrix} + o_p(1) \end{aligned} \tag{43}$$

By Lemma 1, (40) the sample variance covariance converges to the following limit uniformly in $r \in \Xi$

$$var\begin{bmatrix} n^{-1/2}vec(X_r^{*'}V_{2r})^w \\ n^{-1/2}(X_r^{*'}y_1)^w \end{bmatrix} \rightarrow \begin{bmatrix} r(1-r)^2\Sigma_{xv2} & r(1-r)^2\Sigma_{xv21} \\ r(1-r)^2\Sigma'_{xv21} & r(1-r)\Sigma_{xv1} \end{bmatrix}. \tag{44}$$

Using (44) in (43) gives us the limit variance covariance matrix for

$$var(vec(J,n^{-1/2}(X_r^{*'}y_1)^w)) \rightarrow \begin{bmatrix} r(1-r)^2[\Sigma_{xv2}-(1-r)\Sigma_{xv21}\Sigma_{xv1}^{-1}\Sigma'_{xv21}] & 0 \\ 0 & r(1-r)\Sigma_{xv1} \end{bmatrix} \tag{45}$$

The asymptotic independence of J and $(X_r^{*'}y_1)^w$ is shown by the above equation. Note that we can rewrite the LM test statistic by (15)(17)

$$LM = \frac{1}{2n}(y'_1X_r^*)^w\hat{\Omega}_{11,r}^{-1/2}P_{\hat{\Omega}_{11,r}^{-1/2}J}\hat{\Omega}_{11,r}^{-1/2}(X_r^{*'}y_1)^w. \tag{46}$$

Note that $P_{\hat{\Omega}_{11,r}^{-1/2}J} = \hat{\Omega}_{11,r}^{-1/2}J(J'\hat{\Omega}_{11,r}^{-1}J)^{-1}J'\hat{\Omega}_{11,r}^{-1/2}$. Using Theorem 1 and the form of the test statistics in (46) we have the desired result.**Q.E.D**

REFERENCES

ANDERSON, T.W., AND 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. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica,* 59, 817-858.

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. AND R.J. SMITH (2006): "Generalized Empirical Likelihood Tests in Time Series Models with Potential Identification Failure," Working Paper. Department of Economics, UCLA.

KLEIBERGEN, F. (2002): "Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression," *Econometrica,* 70, 1781-1805.

KLEIBERGEN, F. (2005): "Testing Parameters in GMM without Assuming That They are Identified," *Econometrica,*, 73, 1103-1124.

NEWEY W.K. AND K.D. WEST (1987): "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica,* 55, 703-708.

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. AND J.H. STOCK (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica,* 65, 557-587.

STOCK, J.H. AND J.H.WRIGHT (2000): "GMM With Weak Identification," *Econometrica,* 68, 1055-1096.