Inference Regarding Multiple Structural Changes in Linear Models Estimated via Two Stage Least Squares[1]

Alastair R. Hall

University of Manchester and North Carolina State University[2]

Sanggohn Han

Hyundai Research Institute

and

Otilia Boldea

Tilburg University

November 5, 2007

[1]We are grateful to Denise Osborn and Eric Renault for valuable comments, and also for the comments of participants at the presentation of this paper at the World Congress of the Econometric Society, London, August 19-24, 2005, the Triangle Econometrics Conference, RTP, NC, December 2, 2005, ESRC Econometrics group seminar at the Institute of Fiscal Studies, London and at seminars at the Universities of Birmingham and Warwick. We are also very grateful to Chengsi Zhang and Denise Osborn for providing us with the data used in the empirical example.

[2]Corresponding author. Economics, SoSS, University of Manchester, Manchester M13 9PL, UK. Email: alastair.hall@manchester.ac.uk

Abstract

In this paper, we extend Bai and Perron's (1998, *Econometrica*, p.47-78) framework for multiple break testing to linear models estimated via Two Stage Least Squares (2SLS). Within our framework, the break points are estimated simultaneously with the regression parameters via minimization of the residual sum of squares on the second step of the 2SLS estimation. We establish the consistency of the resulting estimated break point fractions. We show that various F-statistics for structural instability based on the 2SLS estimator have the same limiting distribution as the analogous statistics for OLS considered by Bai and Perron (1998). This allows us to extend Bai and Perron's (1998) sequential procedure for selecting the number of break points to the 2SLS setting. Our methods also allow for structural instability in the reduced form that has been identified *a priori* using data-based methods. As an empirical illustration, our methods are used to assess the stability of the New Keynesian Phillips curve.

JEL classification: C12, C13

Keywords: Structural Change, Multiple Break Points, Instrumental Variables Estimation.

1 Introduction

Linear models are widely applied in the analysis of macroeconomic time series. In many cases, at least some of the explanatory variables are correlated with the error and so the model is estimated via Instrumental Variables (IV). While it is routine to assume in estimation that the parameters of these models are constant over time, there are reasons why this assumption may be questionable. In particular, it can be argued that policy changes and/or exogenous shifts may cause realignments in the relationship between economic variables which are reflected in changes in the parameters. Therefore, it is important to develop methods for both detecting parameter instability and also for building models that incorporate this behaviour.

Considerable attention has focused on developing tests for structural instability within the IV or more generally the Generalized Method of Moments (GMM) framework.[1] The majority of this literature has focused on the design of tests against alternatives in which there is structural instability at a single breakpoint in the sample. Although these tests are also shown to have non-trivial power against other alternatives, it is clearly desirable to develop procedures that can discriminate between various forms of instability.

An important step in this direction is taken by Bai and Perron (1998).[2] They develop methods that are designed to test for discrete shifts in the parameters at potentially multiple and unknown break points in the sample. Their analysis is in the context of linear regression models estimated via Ordinary Least Squares (OLS). Within their framework, the break points are estimated simultaneously with the regression parameters via minimization of the residual sum of squares. Bai and Perron (1998) establish the consistency and the limiting distribution of the resulting break point fractions. They also propose a sequential procedure for selecting the number of break points in the sample based on various F-statistics for parameter constancy.

While not the only possible form for structural instability, the model with the discrete shifts at multiple unknown break points has some appeal in macroeconometric applications because it captures the case where relationships change due to changes in policy regime or exogenous shifts. However, since Bai and Perron's (1998) analysis is predicated on the assumption that

[1]See *inter alia* Andrews and Fair (1988), Ghysels and Hall (1990), Andrews (1993), Sowell (1996) and Hall and Sen (1999).

[2]Bai and Perron's (1998) paper also contributes to the literature in statistics on change point estimation in time series. See *inter alia* Picard (1985), Hawkins (1986), Bhattacharya (1987), Yao (1987) and Bai (1994).

all explanatory variables are exogenous, their methods can not be applied to the types of linear macroeconometric models mentioned above.

In this paper, we extend Bai and Perron's (1998) framework to linear models estimated via Two Stage Least Squares (2SLS) and thereby provide a methodology for estimating linear models with endogenous regressors that exhibit discrete shifts in the parameters at multiple unknown points in the sample. Within our framework, the break points are estimated simultaneously with the regression parameters via minimization of the residual sum of squares on the second step of the 2SLS estimation. We establish the consistency of the resulting break point fractions. We show that the various F-statistics for testing parameter constancy based on the 2SLS estimator have the same limiting distribution as the analogous statistics for OLS considered by Bai and Perron (1998). This allows us to extend Bai and Perron's (1998) sequential procedure for selecting the number of break points to the 2SLS setting.

As can be seen from the above summary, our focus is on the stability of the parameters in the second stage regression or, in other words, in the structural equation of interest. However to implement 2SLS, it is necessary in the first stage regression to estimate the reduced form for the endogenous regressors in the structural equation of interest and this, of course, requires an assumption about the constancy or lack thereof of these reduced form parameters. In this paper, we establish the aforementioned results under two scenarios of interest, namely: (i) the parameters in the first stage regression are constant; (ii) the parameters in the first stage regression are subject to discrete shifts within the sample period and the locations of these shifts are estimated *a priori* via a data-based method that satisfies certain conditions. The latter conditions allow the case in which the location of the instability is estimated via an application of Bai and Perron's (1998) methods to the appropriate reduced form equations on an equation by equation basis.

To illustrate our methods, we consider the stability of the New Keynesian Phillips curve (NKPC) estimated using quarterly data for the US over the period 1968.3-2001.4. The NKPC is of considerable theoretical importance in monetary policy analysis and its estimation has received considerable attention in the literature. Zhang, Osborn, and Kim (2007) observe that empirical studies of the NKPC often reach conflicting conclusions about the importance of key variables in the determination of inflation, and argue this may be due to neglected parameter variation. Zhang, Osborn, and Kim (2007) argue that changes in monetary policy regimes may cause changes in the parameters of the NKPC; if true, this would mean that the parameters of the NKPC would exhibit discrete shifts at potentially multiple points in the sample. Zhang, Osborn, and Kim (2007) investigate this issue using a methodology based on uncovering break points in the sample via the maximization of Wald statistics for parameter change associated with 2SLS estimation. However, while their methodology has an intuitive appeal, there is no theoretical justification for their methods as they note; it is, therefore, unclear exactly how to interpret their results. In contrast, our methods can be applied to this model under plausible assumptions about the data. Our analysis indicates that there are shifts in the parameters of both the appropriate reduced forms and also in the NKPC itself.

It is useful to compare our results to two other recent extensions of Bai and Perron's (1998) framework. Qu and Perron (2007) extend Bai and Perron's (1998) framework to systems of regression equations and consider the case in which estimation and inference are based on quasi-maximum likelihood techniques under normality. Perron and Qu (2006) consider the case of a regression equation in which the least squares estimation imposes cross-regime restrictions, such as the equality of parameters in two non-adjacent regimes. While both these papers expand the set of available techniques in important ways, both sets of results are predicated on the assumption that the explanatory variables are uncorrelated with the error(s). To our knowledge, our paper is the first to consider estimation and inference about multiple structural changes in a linear model with endogenous regressors.

An outline of the paper is as follows. Section 2 lays out the model, illustrates it via the NKPC and also explains details of the estimation. Section 3 presents results on the limiting behaviour of the break fraction estimators associated with the 2SLS estimation of the structural equation of interest. It is shown that the break fraction estimators are consistent and deviate from the true break fractions by a term of large order in probability T^{-1}, where T is the sample size. The import of this result is that inference regarding the parameters of the structural equation can be conducted as if the the true break fractions are known *a priori*. In the remainder of the paper, we consider the limiting behaviour of the 2SLS estimator and various associated inference procedures. Section 4 presents the limiting distribution of the 2SLS estimator. Section 5 presents the limiting distributions of the various F-statistics. The simulation evidence is reported in Section 6. Section 7 presents our empirical application and some concluding remarks are offered in Section 8. All proofs are relegated to a mathematical appendix.

2 The Model and The Estimation

2.1 The model

We consider the case in which the equation of interest is a multiple linear regression model with m breaks (*i.e.* $m+1$ regimes), that is

$$y_t \;=\; x_t'\beta^0_{x,i} \,+\, z_{1,t}'\beta^0_{z_1,i} \,+\, u_t, \qquad i = 1,...,m+1, \qquad t = T^0_{i-1}+1,...,T^0_i \tag{1}$$

where $T^0_0 = 0$ and $T^0_{m+1} = T$. In this model, y_t is the dependent variable, x_t is a $p_1 \times 1$ vector of explanatory variables that are correlated with the error u_t and $z_{1,t}$ is a $p_2 \times 1$ vector of explanatory variables that are uncorrelated with u_t and includes the intercept. We define $p = p_1 + p_2$. The error term, u_t, is assumed to have a mean of zero.

Following the convention in the literature, we index the break points $\{T^0_i\}$ by break fractions $\{\lambda^0_i\}$. These break fractions must satisfy the following:[3]

Assumption 1 $T^0_i = [T\lambda^0_i]$, *where* $0 < \lambda^0_1 < ... < \lambda^0_m < 1$.

Assumption 1 requires the break points to be asymptotically distinct.

In view of the correlation between x_t and u_t, OLS estimation of (1) would yield inconsistent estimators of the regression parameters. We therefore consider the case in which (1) is estimated via 2SLS. To implement 2SLS, it is necessary to specifiy the reduced form for x. As noted in the introduction, we consider two scenarios: (i) the reduced form for x_t is structurally stable; (ii) the reduced form for x_t exhibits parameter variation. We elaborate on these two scenarios in turn.

Scenario (i): stable reduced form.

The reduced form for x_t is assumed to be as follows:

$$x_t' \;=\; z_t'\Delta_0 \,+\, v_t' \tag{2}$$

where $z_t = (z_{t,1}, z_{t,2}, ..., z_{t,q})'$ is a $q \times 1$ vector of instruments that is uncorrelated with both u_t and v_t, $\Delta_0 = (\delta_{1,0}, \delta_{2,0}, ..., \delta_{p_1,0})$ with dimension $q \times p_1$ and each $\delta_{j,0}$ for $j = 1,...,p_1$ has dimension $q \times 1$. We assume that z_t contains $z_{1,t}$. Under the assumption that $E[{u_t}^2|z_t] = \sigma^2$, the optimal IV estimator is the 2SLS estimator.[4] Our analysis is confined to the 2SLS estimator,

[3] $[\cdot]$ denotes the integer part of the quantity in the brackets.

[4] See, for example, Hall (2005)[p.44].

although we wish to emphasize that the aforementioned conditional homoscedasticity restriction is only imposed in certain parts of the analysis. $\diamond$.

Scenario (ii): unstable reduced form.

The reduced form for x_t is:

$$x_t^{'} \;=\; z_t^{'}\Delta_0^{(i)} \;+\; v_t^{'}, \qquad i = 1, 2, \ldots, h+1, \qquad t = T_{i-1}^{*} + 1, \ldots, T_i^{*} \tag{3}$$

where $T_0^* = 0$ and $T_{h+1}^* = T$. The points $\{T_i^*\}$ are assumed to be generated as follows.

Assumption 2 $T_i^* = [T\pi_i^0]$, *where* $0 < \pi_1^0 < \ldots < \pi_h^0 < 1$.

Note that the break fractions $\{\pi_i^0\}$ may or may not coincide with $\{\lambda_i^0\}$. Let $\pi^0 = [\pi_1^0, \pi_2^0, \ldots, \pi_h^0]'$. Within our analysis, it is assumed that the break points in the reduced form are estimated prior to estimation of the structural equation in (1). For our analysis to go through, the estimated break fractions in the reduced form must satisfy certain conditions that are detailed below; these conditions would hold, for instance, if Bai and Perron's (1998) methodology is applied equation by equation to the reduced form.

Equation (3) can be re-written as follows

$$x_t^{'} \;=\; \tilde{z}_t(\pi^0)^{'}\Theta_0 \;+\; v_t^{'}, \qquad\qquad t = 1, 2, \ldots, T \tag{4}$$

where $\Theta_0 = [\Delta_0^{(1)'}, \Delta_0^{(2)'}, \ldots, \Delta_0^{(h+1)'}]'$, $\tilde{z}_t(\pi^0) = \iota(t,T) \otimes z_t$, $\iota(t,T)$ is a $(h+1) \times 1$ vector with first element $\mathcal{I}\{t/T \in (0, \pi_1^0]\}$, $h+1^{th}$ element $\mathcal{I}\{t/T \in (\pi_h^0, 1]\}$, k^{th} element $\mathcal{I}\{t/T \in (\pi_{k-1}^0, \pi_k^0]\}$ for $k = 1, 2, \ldots, h$ and $\mathcal{I}\{\cdot\}$ is an indicator variable that takes the value one if the event in the curly brackets occurs. Notice that (4) fits the generic constant parameter form of (2). $\diamond$

To illustrate the potential interest in our framework, we consider the case of the NKPC. For ease of exposition, it suffices here to consider the following stylized version of the NKPC,

$$inf_t \;=\; c_0 \;+\; \alpha_f inf_{t+1|t}^e \;+\; \alpha_b inf_{t-1} \;+\; \alpha_{og} og_t \;+\; u_t \tag{5}$$

where inf_t is inflation in (time) period t, $inf_{t+1|t}^e$ denotes expected inflation in period $t+1$ given information available in period t, og_t is the output gap in period t, u_t is an unobserved error term and $\theta = (c_0, \alpha_f, \alpha_b, \alpha_y)'$ are unknown parameters. The variables $inf_{t+1|t}^e$ and og_t are anticipated to be correlated with the error u_t, and so (5) is commonly estimated via IV; *e.g.*

see Zhang, Osborn, and Kim (2007) and the references therein. Suitable instruments must be both uncorrelated with u_t and correlated with $inf^e_{t+1|t}$ and og_t. In this context, the instrument vector z_t commonly includes such variables as lagged values of expected inflation, the output gap, the short-term interest rate, unemployment, money growth rate and inflation. This model fits within our framework with (5) as the structural equation of interest provided the reduced forms for $inf^e_{t+1|t}$ and og_t are assumed to be given by either (2) or (3). We return to this example in Section 7.

2.2 The estimation

To describe the estimation of the model, it is assumed that the number of break points m is known but their location is not. Therefore the researcher must estimate both the break points and regression parameters. This estimation proceeds as follows. On the first stage, the reduced form for x_t is estimated via OLS using - as appropriate - either (2) or a version of (4) with estimated break fractions substituted for π^0. Let $\hat{x}_t$ denote the resulting predicted value for x_t. The second stage of the 2SLS estimation is itself divided into a number of steps because of the need to estimate both the break points and the regression parameters. The first step of the second stage is to estimate the model

$$y_t \;=\; \hat{x}_t^{'}\beta^*_{x,i} \;+\; z_{1,t}'\beta^*_{z_1,i} \;+\; \tilde{u}_t, \quad i = 1, ..., m+1; \quad t = T_{i-1}+1, ..., T_i \tag{6}$$

via OLS for each possible m-partition of the sample, denoted by $\{T_j\}_{j=1}^m$, such that $T_i - T_{i-1} \geq q$. Letting $\beta_i^{*\prime} = ({\beta^*_{x,i}}', {\beta^*_{z_1,i}}')'$, the resulting estimates of $\beta^* = ({\beta_1^*}', {\beta_2^*}', ..., {\beta^*_{m+1}}')'$ are obtained by minimizing the sum of squares of the residuals

$$S_T(T_1, ..., T_m) = \sum_{i=1}^{m+1} \sum_{t=T_{i-1}+1}^{T_i} (y_t - \hat{x}_t'\beta_{x,i} - z_{1,t}'\beta_{z_1,i})^2 \tag{7}$$

with respect to $\beta = ({\beta_1}', {\beta_2}', ..., {\beta_{m+1}}')'$. We denote these estimators by $\hat{\beta}(\{T_i\}_{i=1}^m)$.

The second step of the second stage involves constructing the minimized sum of squares associated with (6) for each partition, that is

$$S_T(T_1, ..., T_m;\ \hat{\beta}(\{T_i\}_{i=1}^m) \;=\; \sum_{i=1}^{m+1} \sum_{t=T_{i-1}+1}^{T_i} (y_t - \hat{x}_t'\beta_i - z_{1,t}'\beta_{z_1,i})^2 \Big|_{\beta=\hat{\beta}(\{T_i\}_{i=1}^m)} \tag{8}$$

The estimates of the break points, $(\hat{T}_1, ..., \hat{T}_m)$, are defined as

$$(\hat{T}_1, ..., \hat{T}_m) = \arg\min_{T_1,...,T_m} S_T(T_1, ..., T_m;\ \hat{\beta}(\{T_i\}_{i=1}^m)) \tag{9}$$

where the minimization is taken over all partitions, $(T_1, ..., T_m)$ such that $T_i - T_{i-1} \geq q$. The 2SLS estimates of the regression parameters, $\hat{\beta}(\{\hat{T}_i\}_{i=1}^m) = (\hat{\beta}_1', \hat{\beta}_2', ..., \hat{\beta}_{m+1}')'$, are the regression parameter estimates associated with the estimated partition, $\{\hat{T}_i\}_{i=1}^m$.

3 Limiting behaviour of the break fraction estimators

In this section we analyze the limiting behaviour of the break point fraction estimators $\{\hat{\lambda}_i = \hat{T}_i/T\}$. Two properties are established: consistency and that the estimated break fractions deviate from the true break fractions by an $O_p(T^{-1})$ term. These results are established for both the scenarios regarding the parameters of the reduced form for x_t described in Section 2. We take each of these scenarios in turn.

3.1 Stable reduced form

In this case, the predicted value for x_t is given by

$$\hat{x}_t' \ = \ z_t'\hat{\Delta}_T \ = \ z_t'(\sum_{t=1}^{T} z_t z_t')^{-1} \sum_{t=1}^{T} z_t x_t' \tag{10}$$

To facilitate the analysis of this version of the model, we impose the following conditions.

Assumption 3 *Let* $b_t = (u_t, v_t')'$ *and* $\mathcal{F} = \sigma - field\{\ldots, z_{t-1}, z_t, \ldots, b_{t-2}, b_{t-1}\}$. *Assume* b_t *is a martingale difference relative to* $\{\mathcal{F}_t\}$ *and* $sup_t E[\|b_t\|^4] < \infty$.

Assumption 4 $rank\{\Delta_0\} = p_1$.

Assumption 5 *There exists an* $l_0 > 0$ *such that for all* $l > l_0$, *the minimum eigenvalues of* $A_{il} = (1/l)\sum_{t=T_i^0+1}^{T_i^0+l} z_t z_t'$ *and of* $A_{il}^* = (1/l)\sum_{t=T_i^0-l}^{T_i^0} z_t z_t'$ *are bounded away from zero for all* $i = 1, ..., m+1$.

Assumption 6 $T^{-1}\sum_{t=1}^{[Tr]} z_t z_t' \xrightarrow{p} Q_{ZZ}(r)$ *uniformly in* $r \in [0,1]$ *where* $Q_{ZZ}(r)$ *is positive definite for any* $r > 0$ *and strictly increasing in* r.

Assumption 7 *The minimization in (9) is over all partitions* $(T_1, ..., T_m)$ *such that* $T_i - T_{i-1} > \epsilon T$ *for some* $\epsilon > 0$ *and* $\epsilon < inf_i(\lambda^0_{i+1} - \lambda^0_i)$.

A few comments on these assumptions are in order. Assumption 3 includes the restrictions that b_t is a serially uncorrelated process, and hence the errors in both the structural equation and reduced form exhbit this property. This assumption also includes the restriction that $E[z_t b_t'] = 0_{q\times(p_1+1)}$ which implies both the implicit population moment condition in 2SLS is valid - that is $E[z_t u_t] = 0$ - and also that the conditional mean of the reduced form is correctly specified. However, note that this assumption does allow z_t to contain lagged values of y_t. Assumption 4 implies the standard rank condition for identification in IV estimation in the linear regression model[5] because Assumptions 3, 4 and 6 together imply that

$$T^{-1}\sum_{t=1}^{[Tr]} z_t {x_t}' \;\Rightarrow\; Q_{ZZ}(r)\Delta_0 \;=\; Q_{ZX}(r) \text{ uniformly in } r \in [0,1]$$

where $Q_{ZX}(r)$ has rank equal to p_1 for any $r > 0$. Assumption 5 requires that there be enough observations near the true break points so that they can be identified. This condition is analagous to Bai and Perron's (1998) Assumption A2 and the interested reader is refered to this source for further discussion of this condition. Assumption 7 requires that each segment considered in the minimization contains a positive fraction of the sample asymptotically; in practice ϵ is chosen to be small in the hope that the last part of the assumption is valid.

The proof strategy for consistency is identical to that used by Bai and Perron (1998) in their proof of the corresponding results for OLS estimators. The proof builds from the following two properties of the error sum of squares on the second stage of the 2SLS esimation.

- Since the 2SLS estimators minimize the error sum of squares in (7), it follows that

$$(1/T)\sum_{t=1}^{T}\hat{u}_t^2 \;\le\; (1/T)\sum_{t=1}^{T}\tilde{u}_t^2 \tag{11}$$

  where $\hat{u}_t = y_t - \hat{x}_t'\hat{\beta}_{x,j} - z_{1,t}'\hat{\beta}_{z_1,j}$ denotes the estimated residuals for $t \in [\hat{T}_{j-1}+1, \hat{T}_j]$ in the second stage regression of 2SLS estimation procedure and $\tilde{u}_t = y_t - \hat{x}_t'\beta^0_{x,i} - z_{1,t}'\beta^0_{z_1,i}$ denotes the corresponding residuals evaluated at the true parameter value for $t \in [T^0_{i-1}+1, T^0_i]$.

- Using $d_t = \tilde{u}_t - \hat{u}_t = \hat{x}_t'(\hat{\beta}_{x,j} - \beta^0_{x,i}) - z_{1,t}'(\hat{\beta}_{z_1,j} - \beta^0_{z_1,i})$ over $t \in [\hat{T}_{j-1}+1, \hat{T}_j] \cap [T^0_{i-1}+1, T^0_i]$,

[5]See *e.g.* Hall (2005)[p.35].

it follows that

$$T^{-1}\sum_{t=1}^{T}\hat{u}_t^2 \;=\; T^{-1}\sum_{t=1}^{T}\tilde{u}_t^2 \;+\; T^{-1}\sum_{t=1}^{T}{d_t}^2 \;-\; 2T^{-1}\sum_{t=1}^{T}\tilde{u}_t d_t \tag{12}$$

Consistency is established by proving that if at least one of the estimated break fractions does not converge in probability to a true break fraction then the results in (11)-(12) contradict each other. This conflict is established using the results in the following lemma.

Lemma 1 *Let y_t be generated by (1), x_t be generated by (2), $\hat{x}_t$ be generated by (10) and Assumptions 1, 3-7 hold.*

(i) $T^{-1}\sum_{t=1}^{T}\tilde{u}_t d_t = o_p(1)$.

(ii) If $\hat{\lambda}_j \overset{p}{\nrightarrow} \lambda_j^0$ for some j, then

$$\limsup_{T\to\infty} P\left(T^{-1}\sum_{t=1}^{T}{d_t}^2 > C\left\{\,\|\Delta_0(\beta^0_{x,j}\;-\;\beta^0_{x,j+1})\|^2\;+\;\|\beta^0_{z_1,j}\;-\;\beta^0_{z_1,j+1}\|^2\,\right\}\;+\;\xi_T\right) > \bar{\epsilon}$$

for some $C > 0$ and $\bar{\epsilon} > 0$, where $\xi_T = o_p(1)$.

Using (11)-(12) and Lemma 1, consistency is established along the lines anticipated above.

Theorem 1 *Let y_t be generated by (1), x_t be generated by (2), $\hat{x}_t$ be generated by (10) and Assumptions 1, 3-7 hold, then $\hat{\lambda}_j \overset{p}{\to} \lambda_j^0$ for all $j = 1, 2, ..., m$.*

For the development of inference procedures for determining the number of breaks, it is important to know not only that the break fraction estimators are consistent but also the order of magnitude of their deviation from the true break fraction. This is established in the following theorem.

Theorem 2 *Let y_t be generated by (1), x_t be generated by (2), $\hat{x}_t$ be generated by (10) and Assumptions 1, 3-7 hold then, for every $\eta > 0$, there exists C such that for all large T, $P(T|\hat{\lambda}_j - \lambda_j^0| > C) < \eta$, for $j = 1, ..., m$.*

Therefore, the break fraction estimators deviate from the true break fractions by a term of order in probability T^{-1}.

3.2 Unstable reduced form

Recall that the reduced form exhibits discrete parameter changes at unknown points in the sample and these points are indexed by the break fraction vector, π^0. We suppose that π^0 is estimated by $\hat{\pi}$ and that these estimated break fractions satisfy the following condition.

Assumption 8 $\hat{\pi} \ = \ \pi^0 \ + \ O_p(T^{-1})$

Note that Assumption 8 implies $\hat{\pi}$ is consistent for π^0 and $T(\hat{\pi} - \pi^0)$ is bounded in probability. Such an estimator might be obtained by applying Bai and Perron (1998)'s methodology equation by equation and then pooling the resulting estimates of the break fractions. For our purposes, it only matters that Assumption 8 holds and not how $\hat{\pi}$ is obtained. The latter is, of course, a matter of practical importance but we do not address it here.

These estimated breaks are imposed on the the reduced form for x_t. Let $\hat{\Theta}_T$ be the OLS estimator of Θ_0 from the model

$$x_t' \ = \ \tilde{z}_t(\hat{\pi})'\Theta_0 + error \qquad t = 1, 2, \cdots, T \tag{13}$$

where $\tilde{z}_t(\hat{\pi})$ is defined analogously to $\tilde{z}_t(\pi^0)$, and now define $\hat{x}_t$ to be

$$\hat{x}_t' \ = \ \tilde{z}_t(\hat{\pi})'\hat{\Theta}_T \ = \ \tilde{z}_t(\hat{\pi})'\{\sum_{t=1}^{T} \tilde{z}_t(\hat{\pi})\tilde{z}_t(\hat{\pi})'\}^{-1} \sum_{t=1}^{T} \tilde{z}_t(\hat{\pi})x_t' \tag{14}$$

For the analysis in the case, the regularity conditions need to be altered. Assumption 4 is replaced by:

Assumption 9 $rank\{\Delta_0^{(i)}\} \ = \ p \quad for \ \ i = 1, 2, \cdots, h+1.$

It is also necessary to modify Assumption 7.

Assumption 10 *The minimization in (9) is over all partitions* $(T_1, ..., T_m)$ *such that* $T_i - T_{i-1} > \epsilon T$ *for some* $\epsilon > 0$ *and* $\epsilon < inf_i(\lambda_{i+1}^0 - \lambda_i^0)$ *and* $\epsilon < inf_j(\pi_{j+1}^0 - \pi_j^0)$.

The following theorem establishes the consistency of the break fraction estimators.

Theorem 3 *If Assumptions 1-3, 5-10 hold,* y_t *is generated via (1),* x_t *is generated via (4) and* $\hat{x}_t$ *is calculated via (14), then*

$$\hat{\lambda}_j \xrightarrow{p} \lambda_j^0 \qquad for\ all\ j = 1, 2, \cdots, m.$$

In order to extend Theorem 2, we impose one final condition.

Assumption 11 *There exists an $l_* > 0$ such that for all $l > l_*$, the minimum eigenvalues of $B_{il} = (1/l)\sum_{t=T_i^*+1}^{T_i^*+l} z_t z_t'$ and of $B_{il}^* = (1/l)\sum_{t=T_i^*-l}^{T_i^*} z_t z_t'$ are bounded away from zero for all $i = 1, ..., h+1$.*

Assumption 11 is similar to Asssumption 5 above but refers to the break points in the reduced form. The order in probability of the estimated break fractions is given in the following theorem.

Theorem 4 *If Assumptions 1-3, 5-11 hold, y_t is generated via (1), x_t is generated via (4) and $\hat{x}_t$ is calculated via (14), then, for every $\eta > 0$, there exists C such that for all large T, $P(T|\hat{\lambda}_j - \lambda_j^0| > C) < \eta$, for $j = 1, ..., m$.*

3.3 Discussion

At this stage, it is useful to comment on the nature of the foregoing analysis. First consider the case where the reduced form is structurally stable. In this case, Theorems 1-2 establish that the break fraction estimators, $\{\hat{\lambda}_j\}$, are consistent and $\hat{\lambda}_j - \lambda_j^0 = O_p(T^{-1})$. Now consider the case where the reduced form exhibits parameter variation. If the location of the breaks in the reduced form are known *a priori* then, as noted above, the reduced form can be re-written as a structurally stable regression equation involving the augmented parameter vector.[6] Therefore, in this case, the limiting behaviour of the break fraction estimators associated with the structural equation is covered by Theorems 1-2. However, in most cases, the locations of the breaks in the reduced form are unknown and so must be estimated *a priori.* In this case, Theorems 3-4 provide conditions on the estimators of the reduced form break fractions, $\{\hat{\pi}_i\}$, under which the break fraction estimators associated with the structural equation, $\{\hat{\lambda}_j\}$, are consistent and $\hat{\lambda}_j - \lambda_j^0 = O_p(T^{-1})$.

Of the scenarios described above, the most empirically relevant is likely to be the one involving estimation of break fractions in both reduced form and structural equations. Under our assumptions, the estimators of the break fractions in both reduced form and structural equations converge at rate T to the true break fractions. It emerges below that this rate is sufficiently fast that the estimation of the break fractions can be ignored in the asymptotic analysis of the

[6]See equation (4).

2SLS estimators and its associated statistics.[7] In other words, for the purposes of the asymptotic analysis of the 2SLS estimator and its associated statistics, we can essentially proceed as if the break fractions in both equations are known. Since, as noted above, the reduced form with known break points can be rewritten as a constant parameter regression model, we focus exclusively for the remainder of the paper on the case in which the reduced form is structurally stable. The analagous results for the model with parameter variation in the reduced form can be deduced from the results presented with an appropriate redefinition of the regressor vector in the reduced form.

4 The limiting distribution of the 2SLS estimators

Once the break fractions are estimated, it is clearly desirable to perform inference about the structural parameters $\{\beta_i^0\}$. If the break fractions are known *a priori* then standard arguments can be employed to show the root T asymptotic normality of the 2SLS estimator. Since the estimated break fractions converge at rate T, this standard asymptotic distribution theory can be extended to the 2SLS estimates based on the estimated break fractions.

Theorem 5 *Let y_t be generated by (1), x_t be generated by (2), $\hat{x}_t$ be generated by (10) and Assumptions 3-6 hold, then*

$$T^{1/2}\left(\hat{\beta}(\{\hat{T}_i\}_{i=1}^m) - \beta^0\right) \Longrightarrow N\left(0_{p(m+1)\times 1}, V_\beta\right)$$

where $\beta^0 = [\beta_1^{0\prime}, \beta_2^{0\prime}, \ldots, \beta_{h+1}^{0}{}']'$, $\beta_i^0 = [\beta_{x,i}^0{}', \beta_{z_1,i}^0{}']'$,

$$\begin{aligned} V_\beta &= \begin{pmatrix} V_\beta^{(1,1)} & \cdots & V_\beta^{(1,m+1)} \\ \vdots & \ddots & \vdots \\ V_\beta^{(m+1,1)} & \cdots & V_\beta^{(m+1,m+1)} \end{pmatrix} \\ V_\beta^{(i,j)} &= R_i S_{(i,j)} R_j', \quad \textit{for } i,j = 1,2,\ldots m+1 \\ R_i &= \left(A(1)Q_{ZZ}(1)^{-1}Q_i Q_{ZZ}(1)^{-1}A(1)'\right)^{-1} A(1)Q_{ZZ}(1)^{-1} \end{aligned}$$

and $Q_i = Q_{ZZ}(\lambda_i^0) - Q_{ZZ}(\lambda_{i-1}^0)$, $A(r)' = [Q_{ZX}(r), Q_{Z_1Z}(r)']$, $Q_{Z_1Z}(r)$ *is the probability limit of* $T^{-1}\sum_{t=1}^{[Tr]} z_{1,t}z_t'$ *(defined in Assumption 6),* $S_{(i,j)} = lim_{T\to\infty} Cov[T^{-1/2}\sum_{i_0} z_t\tilde{u}_t, T^{-1/2}\sum_{j_0} z_t\tilde{u}_t]$, $\sum_{i_0}$ *denotes the summation over* $t = [T\lambda_{i-1}^0]+1, \ldots [T\lambda_i^0]$, *and we set* $\lambda_0^0 = 0$, $\lambda_{m+1}^0 = 1$.

[7] A similar finding is reported by Bai and Perron (1998) in their analysis of OLS estimators.

Note that $S_{(i,j)}$ is non-zero in general because the first stage regression pools observations across regimes and this creates a connection between the aforementioned sums from different regimes. However, if the reduced form is also unstable then the connection across regimes is broken in one leading case. If the breaks in the structural equation also occur in the reduced form then the predictions are only based on the observations in the sub-sample in question and so V_β is block diagonal. Specifically, if $h \geq m$ and $\lambda_i^0 = \pi_j^0$ for some j for each i then

$$V_\beta \ = \ diag(\tilde{V}_\beta^{(1,1)}, \tilde{V}_\beta^{(2,2)}, \ldots, \tilde{V}_\beta^{(m+1,m+1)}) \tag{15}$$

where $\tilde{V}_\beta^{(i,i)} = \tilde{R}_i \tilde{S}_{(i,i)} \tilde{R}_i'$, $\tilde{R}_i = \left(A_i Q_i^{-1} A_i'\right)^{-1} A_i Q_i^{-1}$, $A_i = A(\lambda_i^0) - A(\lambda_{i-1}^0)$, and $\tilde{S}_{(i,i)} = lim_{T\to\infty} T^{-1} \sum_{i_0} Var[z_t u_t]$. Notice that $\tilde{V}_\beta^{(i,i)}$ is just the variance of the 2SLS estimator based on the i^{th} sub-sample allowing potentially for breaks in the reduced fome within that sub-sample.

5 Test statistics for multiple breaks

The sup-F type test of no structural break ($m = 0$) versus the alternative hypothesis that there is $m = 1$ break has been considered by Andrews (1993). Bai and Perron (1998) generalize Andrew's sup-F type test to the hypothesis $m = k$ for linear models estimated via OLS. In this section, we extend Bai and Perron's results to linear models estimated via $2SLS$.

For this part of the analysis, we impose the following restrictions.

Assumption 12 *(i)* $T^{-1} \sum_{t=1}^{[Tr]} z_t z_t' \xrightarrow{p} r Q_{ZZ}$ *uniformly in* $r \in [0,1]$ *where* Q_{ZZ} *is a positive definite matrix of constants;*

(ii) the conditional variance of the errors is independent of t*, that is*

$$Var\left[\begin{pmatrix} u_t \\ v_t \end{pmatrix} \Big| z_t \right] \ = \ \Omega \ = \begin{bmatrix} \sigma^2 & \gamma' \\ \gamma & \Sigma \end{bmatrix}$$

where Ω *is a constant, positive definite matrix,* σ^2 *is a scalar and* Σ *is a* $p_1 \times p_1$ *matrix;*

The restrictions in Assumption 12 are analogous to that imposed by Bai and Perron (1998) in their Assumptions A8 and A9 which underpin their analysis of various F-statistics for testing for multiple breaks within the OLS framework.[8]

[8] Although note that the conditional variance restriction in Assumption 12 involves both u_t and v_t whereas Bai and Perron (1998) need only restrict the conditional variance of u_t.

Assumptions 3 and 12 together ensure that a uniform version of the multivariate functional central limit theorem in de Jong and Davidson (2000) holds:

$$T^{-1/2} \sum_{t=1}^{[Tr]} \begin{pmatrix} u_t \\ v_t \end{pmatrix} \otimes z_t \implies (\Omega^{1/2} \otimes Q_{ZZ}^{1/2}) B(r) \tag{16}$$

where $B_n(r)$ is a $n \times 1$ standard Brownian motion with $n = q \times (p_1 + 1)$ and "$\Longrightarrow$" denotes weak convergence in the space $D[0,1]$ under the skorohod metric.

The sup-F type test statistic can be defined as follows. Let $(T_1, ..., T_k)$ be a partition such that $T_i = [T\lambda_i]$ $(i = 1, ..., k)$. Define

$$F_T(\lambda_1, ..., \lambda_k; p) = \left\{ \frac{T - (k+1)p}{kp} \right\} \left\{ \frac{SSR_0 - SSR_k}{SSR_k} \right\} \tag{17}$$

where SSR_0 and SSR_k are the sum of squared residuals based on the fitted X under null and alternative hypothesis, respectively. Recall from Assumption 7 that the minimzation is performed over partitions which are asymptotically large and the size of the partitions is controlled by ϵ, a non-negative constant. Accordingly, we define $\Lambda_\epsilon = \{(\lambda_1, ..., \lambda_k) : |\lambda_{i+1} - \lambda_i| \geq \epsilon, \lambda_1 \geq \epsilon, \lambda_k \leq 1 - \epsilon\}$. Finally, the sup-F type test statistic is defined as

$$Sup - F_T(k; p) = Sup_{(\lambda_1, ..., \lambda_k) \in \Lambda_\epsilon} F_T(\lambda_1, .., \lambda_k; p) \tag{18}$$

Theorem 6 *If the data are generated by (1)-(2) with $m = 0$, $\hat{x}_t$ is generated by (10) and Assumptions 1, 3-7 and 12 hold then $Sup - F_T(k; p) \Rightarrow Sup - F_{k,p} \equiv Sup_{(\lambda_1, ..., \lambda_k) \in \Lambda_\epsilon} F(\lambda_1, .., \lambda_k; p)$ where*

$$F(\lambda_1, ..., \lambda_k; p) \equiv \frac{1}{kp} \sum_{i=1}^{k} \frac{||\lambda_{i+1} W_i - \lambda_i W_{i+1}||^2}{\lambda_i \lambda_{i+1} (\lambda_{i+1} - \lambda_i)}$$

where k is the number of break points under the alternative hypothesis, and $W_i \equiv W(\lambda_i)$, a $p \times 1$ vector standard Brownian motion process.

We note that the limiting distribution in Theorem 6 is exactly the same as the one in Bai and Perron's (1998) analogous result for the sup-F test based on OLS estimators when the regressors are exogenous. Percentiles for this distribution can be found in Bai and Perron (1998)[Table I] for $\epsilon = 0.05$ and in Bai and Perron (2001) for other values of ϵ.

The $Sup - F_T(k; p)$ statistic is used to test the null hypothesis of structural stability against the k-break model, and so is designed for the case in which a particular choice of k is of interest.

In many circumstances, a researcher is unlikely to know *a priori* the appropriate choice of k for the alternative hypothesis. To circumvent this problem, Bai and Perron (1998) propose so called "Double Maximum tests" that combine information from the $Sup-F_T(k;p)$ statistics for different values of k running from one to some ceiling K. We consider here only the following example of Double Maximum test,[9]

$$UDmaxF_T(K;p) \;=\; \max_{1\leq k\leq K} \sup_{(\lambda_1,...,\lambda_k)\in\Lambda_\epsilon} F_T(\lambda_1,...,\lambda_k;p) \tag{19}$$

The limiting distribution of this statistic follows directly from Theorem 6.

Corollary 1 *Under the conditions of Theorem 6, it follows that*

$$UDmaxF_T(K;p) \implies \max_{1\leq k\leq K}\{Sup-F_{k,p}\}$$

Critical values for the limiting distribution in Corollary 1 are presented in Bai and Perron (1998)[Table 1] for $\epsilon = 0.05$ and in Bai and Perron (2001) for other values of ϵ.

The $Sup-F_T(k;p)$ and $UDmaxF_T(K;p)$ statistics are used to test the null hypothesis of no breaks. It is also of interest to develop statistics for testing the null hypothesis of l breaks against the alternative of $l+1$ breaks. Following Bai and Perron (1998), a suitable statistic can be constructed as follows. For the model with l breaks, the estimated break points, denoted by $\hat{T}_1,...,\hat{T}_l$, are obtained by a global minimization of the sum of the squared residuals as in (9). For the model with $l+1$ breaks, l of the breaks are fixed at $\hat{T}_1,...,\hat{T}_l$ and then the location of the $(l+1)^{th}$ break is chosen by minimizing the residual sum of squares. The test statistic is given by

$$F_T(l+1|l) = \max_{1\leq i\leq l+1}\{\frac{SSR_l(\hat{T}_1,...,\hat{T}_l) - \inf_{\tau\in\Lambda_{i,\eta}} SSR_{l+1}(\hat{T}_1,...,\hat{T}_{i-1},\tau,\hat{T}_i,...,\hat{T}_l)\}}{\hat{\sigma}_i^2}\} \tag{20}$$

where

$$\begin{aligned}\hat{\sigma}_i^2 &= \sum_{t=\hat{T}_{i-1}+1}^{\hat{T}_i} (y_t - \hat{x}_t'\hat{\beta}_{x,i} - z_{1,t}'\hat{\beta}_{z_1,i})^2/(\hat{T}_i - \hat{T}_{i-1} - p)\\ \Lambda_{i,\eta} &= \{\tau : \hat{T}_{i-1} + (\hat{T}_i - \hat{T}_{i-1})\eta \leq \tau \leq \hat{T}_i - (\hat{T}_i - \hat{T}_{i-1})\eta\}\end{aligned}$$

and $\hat{\beta}_i$ is the 2SLS estimator calculated using the sample $\hat{T}_{i-1}+1,\ldots,\hat{T}_i$ on the second stage.

[9] *UDmax* denotes Unweighted Double maximum. Bai and Perron (1998) also consider a *WDmax* statistic in which the the maximum is taken over weighted values of the $Sup-F_T(k;p)$ statistics. Analogous *WDmax* statistics can be developed within our framework, but for brevity we do not explore them here.

The following theorem gives the limiting distribution of this statistic under the null hypothesis of l breaks.

Theorem 7 *If the data are generated by (1)-(2) with* $m = l$, $\hat{x}_t$ *is generated by (10) and Assumptions 1, 3-7 and 12 hold then then* $\lim_{T\to\infty} P(F_T(l+1|l) \leq x) = G_{p,\eta}(x)^{l+1}$ *where* $G_{p,\eta}(x)$ *is the distribution function of* $\sup_{\eta\leq\mu\leq 1-\eta} \|W(\mu) - \mu W(1)\|^2/\mu(1-\mu)$.

Once again, the limiting behaviour of the test statistic is the same as that of the analogous statistic proposed by Bai and Perron (1998) for the OLS case. Critical values can be found in Bai and Perron (1998)[Table II] for the case in which calculated with $\eta = .05$ and in Bai and Perron (2001) for other values of η.

Following Bai and Perron (1998), the statistics described in this section can be used to determine the estimated number of breakpoints, $\hat{k}_T$ say, via the following sequential strategy. On the first step, use either $Sup - F_T(1;p)$ or $UDmaxF_T(K,p)$ to test the null hypothesis that there are no breaks. If this null is not rejected then $\hat{k}_T = 0$; else proceed to the next step. On the second step $F_T(2|1)$ is used to test the null hypothesis that there is only one break against the alternative hypothesis of two breaks. If $F_T(2|1)$ is insignificant then $\hat{k}_T = 1$; else proceed to the next step. On the l^{th} step $F_T(l+1|l)$ is is used to test the null hypothesis that there are l breaks against the alternative hypothesis of $l+1$ breaks. If $F_T(l+1|l)$ is insignificant then $\hat{k}_T = l$; else proceed to the next step. This sequence is continued until some preset ceiling for the number of breaks, L say, is reached. If all statistics in the sequence are significant then the conclusion is that there are at least L breaks. We evaluate the finite sample performance of this strategy as part of the simulation study reported in the following section.

To conclude our discussion of these F-statistics, we return to the issue of the assumptions on the errors. Assumption 12 requires the errors to be homoscedastic and serially uncorrelated. It is, however, possible to relax this assumption to some extent as we now discuss. Suppose that it is assumed that a regime is characterized by both a change in the regression parameter vector and also a change in the conditional variance matrix of the errors, that is Ω in Assumption 12 is replaced by Ω_i for $t \in \left([T\lambda_{i-1}^0]+1, [T\lambda_i^0]\right)$. Since the calculation of $F_T(l+1|l)$ only involves sub-sample covariance matrix estimators, it follows that the limiting distribution of the test statistic is unaffected by heteroscedasticity of this type. It is therefore possible to use the the test statistics described above to develop a sequential strategy to determine the number of breaks for the case where the no break model is homoscedastic and the l break models involve a conditional error variance that is constant within a regime but varies across regimes.

6 Finite sample behaviour

In this section, we evaluate the finite sample behaviour of the various statistics discussed in the previous sections via a small simulation study. The simulation design involves models with zero, one or two breaks. Since our analysis of the break fractions is premised on the existence of a break, we begin by discussing the one break and two break models. We then conclude the sections by considering the behaviour of the test statistics in the no break model.

6.1 One break model

The data generating process for the structural equation is:

$$\begin{aligned} y_t &= \beta_1^0 x_t + u_t, \qquad \text{for } t = 1, \ldots, [T/2] \\ &= \beta_2^0 x_t + u_t, \qquad \text{for } t = [T/2]+1, \ldots, T \end{aligned} \tag{21}$$

The reduced form equation for the scalar variable x_t is:

$$x_t = z_t'\delta + v_t, \qquad \text{for } t = 1, \ldots, T \tag{22}$$

where δ is $q \times 1$. The errors are generated as follows: $(u_t, v_t)' \sim IN(0_{2\times 1}, \Omega)$ where the diagonal elements of Ω are equal to one and the off-diagonal elements are equal to 0.5. The instrumental variables, z_t, are generated via: $z_t \sim i.i.d\ N(0_{q\times 1}, I_q)$. The specific parameter values are as follows: (i) $T = 60, 120, 240, 480$; (ii) $(\beta_1^0, \beta_2^0) = (0.1, -0.1), (1, -1)$;[10] (iii) $q = 2, 4, 8$; (iv) δ is chosen to yield the population $R^2 = 0.5$ for the regression in (22).[11] For each configuration, 1000 simulations are performed.

The results are presented in Tables 1-5. We first consider the behaviour of the break fraction estimator calculated under the assumption that there is only one break. Table 1 reports the

[10] We also ran simulations for $(\beta_1^0, \beta_2^0) = (5, -5)$. Since the results for this case are qualitatively the same as the $(1, -1)$ case, they are omitted for brevity. However, they are available from the authors upon request.

[11] For this model, $\delta = \sqrt{R^2/(q - q \times R^2)}$; see Hahn and Inoue (2002).

proportion of the simulations in which $|\hat{\lambda}_1 - \lambda_1^0| \leq c$ for $c = 0.01, 0.02, 0.03, 0.05, 0.1$. In Table 1, the change in the regression parameters is only $\beta_1^0 - \beta_2^0 = 0.2$, and it appears that this makes it difficult to locate the true break point. However, the proportions clearly increase with T and exhibit behaviour in line with the consistency result in Theorem 1. For the other parameter setting, the change in the regression parameters is larger, and the estimated break fractions are within 0.01 of the true break fraction in all replications (and so are not reported in tabular form for brevity). Tables 2-3 report the relative rejection frequencies of $Sup - F_T(k;1)$ (for $k = 1, 2$), $UDmaxF_T(5;1)$ and $F_T(l+1|l)$ (for $l = 1, 2$) statistics where, in both cases the nominal size is 0.05. Notice that the alternative hypothesis is true for the $Sup - F_T(k;1)$ statistic and so these relative frequencies are empirical powers for this statistic. Whereas, for $l = 1$, the null hypothesis is correct for $F_T(l+1|l)$ and so the relative frequencies are the empirical size, and for $l = 2$, the null assumes more breaks than there actually are. The evidence suggests that $Sup - F_T(k;1)$ and $UDmaxF_T(5;1)$ have similar power properties and reasonable power even in the case where $(\beta_1^0, \beta_2^0) = (0.1, -0.1)$. The $F_T(2|1)$ statistics is close to its nominal size; $F_T(3|2)$ tends to reject less frequently than the nominal size. Tables 4-5 report the results from using the sequential strategy based on these statistics that is described in Section 5 with a maximum number of breaks set equal to five. The results indicate that the strategy works well in most cases. There is only one case in which the estimated number of breaks has significant mass at zero and that is where $(\beta_1^0, \beta_2^0) = (0.1, -0.1)$, $q = 2$ and $T = 60$. In all bar two other cases, the estimated number of breaks has no mass at zero and a mass at one that is not less than 0.94.[12] The choice of statistic on the first step appears to have little impact on the empirical distribution of $\hat{k}_T$.

6.2 Two break model

The data generation process for the structural equation is:

$$\begin{aligned} y_t &= \beta_1^0 x_t + u_t, && \text{for } t = 1, \ldots, [T/3] \\ &= \beta_2^0 x_t + u_t, && \text{for } t = [T/3]+1, \ldots, [2T/3] \\ &= \beta_3^0 x_t + u_t, && \text{for } t = [2T/3]+1, \ldots, T \end{aligned}$$

[12] The other values with non-zero mass at zero are: $(\beta_1^0, \beta_2^0) = (0.1, -0.1)$ with $(q, T) = (2, 120), (4, 60)$.

Two choices for β^0 are considered: $(\beta_1^0, \beta_2^0, \beta_3^0) = (-0.1, 0.1, -0.1), (-1, 1, -1)$.[13] All other aspects of the design are the same as the one break model.

Again, we begin by considering the performance of the estimated break fractions. Table 6 reveals that, as in the one break model, the method has some problem locating the breaks in the parameterization involving the smallest change which in this case is $(\beta_1^0, \beta_2^0, \beta_3^0) = (-0.1, 0.1, -0.1)$; nevertheless the empirical distribution of the break fraction estimator does appear to be collapsing on the true fraction as T increases. For the other parameter setting, the change in the regression parameters is much larger, and the estimated break fractions are within 0.01 of the true break fraction in all replications (and so are not reported in tabular form for brevity). Tables 7-8 report the relative rejection frequencies of $Sup-F_T(k;1)$ (for $k = 1, 2$), $UDmaxF_T(5;1)$ and $F_T(l+1|l)$ (for $l = 1, 2$) statistics where, in both cases the nominal size is 0.05. As in the one break model, the statistics are applied with $k, l = 1, 2$. Notice that the alternative hypothesis is true for the $Sup-F_T(k;1)$, $UDmaxF_T(5;1)$ and $F_T(2|1)$ statistics and so these relative frequencies are empirical powers for this statistic. Whereas, the null hypothesis is correct for $F_T(3|2)$ and so the relative frequencies are the empirical size. Table 7 reports the results for the case in which there is the smallest change between regimes, that is $(\beta_1^0, \beta_2^0, \beta_3^0) = (-0.1, 0.1, -0.1)$. In this case, it can be see that the $UDmaxF_T(5,1)$ statistic has markedly higher power than $Sup-F_T(1;1)$, although both can have low power in small samples, a property shared with $F_T(2|1)$. However, in the other parameter configuration, all these statistics reject 100% of the time. The $F_T(3|2)$ statistic is close to its nominal size in most cases. Tables 9-10 report the results form using the sequential strategy for estimating the number of breaks. The results indicate that the strategy works well in the parameter configuration with relatively more change, that is $(\beta_1^0, \beta_2^0, \beta_3^0) = (-1, 1, -1)$ but less well in the case with the smaller change, $(\beta_1^0, \beta_2^0, \beta_3^0) = (-0.1, 0.1, -0.1)$. In the latter case, there is a non-negligible tendency to underfit at samples sizes $T = 60, 120$. This problem stems from the low power properties of the statistics $Sup-F_T(1;1)$ and $UDmax(5;1)$ that are used on the first step of the sequential stratgy. Recall that $UDmax(5;1)$ has better power properties and this translates to a reduced tendency to underfit although there is still a non-negligible mass for $\hat{k}_T$ at zero in the cases where $(q, T) = (2, 60), (2, 120), (4, 60)$.

[13] We also ran simulations for $(\beta_1^0, \beta_2^0, \beta_3^0) = (-5, 5, -5)$. Since the results for this case are qualitatively the same as the $(-1, 1, -1)$ case, they are omitted for brevity. However, they are available from the authors upon request.

6.3 No break model

The previous two designs involve cases where there is a change in the regression parameters of the structural equation. It is also of interest to explore how the test statistics perform in the case where there is no break and so the model is structurally constant. To this end, data are generated from (21) with $\beta_1^0 = \beta_2^0 = 1$. All other aspects of the design are the same as the one break model. Table 11 contains the empirical rejection frequencies for $Sup - F_T(k;1)$ $(k = 1, 2)$, $F_T(l+1|l)$ $(l = 1, 2)$ and $UDmaxF_T(5;1)$ statistics. Note that within this design, the null hypothesis is corrrect for the $Sup-F_T(1;1)$, $Sup-F_T(2;1)$, and $UDmaxF_T(5,1)$ statistics, and so the rejection frequency equals the empirical size. For $F_T(2|1)$ and $F_T(3|2)$ statistics, the null hypothesis involves more breaks than are present in the data. From Table 11, it can be seen that $Sup-F_T(1;1)$, $Sup-F_T(2;1)$, and $UDmaxF_T(5,1)$ exhibit empirical size close to the nominal level of 0.05; both $F_T(2|1)$ and $F_T(3|2)$ reject less frequently than the size. Table 12 presents the empirical distribution of $\hat{k}_T$ based on the sequential strategies using $Sup - F_T(1;1)$ and $UDmaxF_T(5,1)$. Both strategies indicate that no breaks are present in nearly every case.

7 Application

In this section we use our methods to explore the stability of the New Keynesian Phillips curve (NKPC). Zhang, Osborn, and Kim (2007) report that the stylized version of the NKPC in (5) does not have serially uncorrelated errors as required by our Assumption 3, and so we follow their practice and include lagged values of $\Delta inf_t = inf_t - inf_{t-1}$ to remove this dynamic structure from the errors. Accordingly, our analysis is based on

$$inf_t \;=\; c_0 \,+\, \alpha_f inf^e_{t+1|t} \,+\, \alpha_b inf_{t-1} \,+\, \alpha_{og} og_t \,+\, \sum_{i=1}^{3} \Delta inf_{t-i} \,+\, u_t \qquad (23)$$

The data is for the US and is quarterly spanning 1968.3-2001.4. The span of the data is slightly longer than Zhang, Osborn, and Kim (2007) but the definitions of the variables are the same and as follows: inf_t is the annualized quarterly growth rate of the GDP deflator, og_t is obtained from the estimates of potential GDP published by the Congressional Budget Office, $inf^e_{t+1|t}$ is the Greenbook one quarter ahead forecast of inflation prepared within the Fed.[14]

[14]One interesting aspect of Zhang, Osborn, and Kim's (2007) study is that they employ various different inflation forecasts in their estimation. We focus here on just one of their choices for brevity.

Both expected inflation and the output gap are taken to be endogenous and we model their reduced forms as

$$inf^e_{t+1|t} = z_t'\delta_1 + v_{1,t} \tag{24}$$

$$og_t = z_t'\delta_2 + v_{2,t} \tag{25}$$

where z_t contains all other explanatory variables on the righthand side of (23) along with the first lagged value of each of the short term interest rate, the unemployment rate, and the growth rate of the money aggreagte M2.

We first consider the stability of the reduced forms in (24)-(25) using Bai and Perron's (1998) methodology. [15] We assume that the maximum number of breaks is 5 and set $\epsilon = 0.1$. The results are reported in Table 13. First consider the reduced form for $inf^e_{t+1|t}$. There is clear evidence of parameter variation with all the sup-F statistics being significant at the 1% level. Using the sequential testing strategy, we identify two breaks: one at 1975.2 and the other at 1981.1. As a robustness check, we also use BIC to choose the break points and obtain the same estimates.[16] Now consider the reduced form for og_t. Again, there is evidence of parameter variation. The sequential strategy suggests a break at 1975.2. In contrast, BIC favours the model with no breaks. Given our purposes, it seems better to impose this break in our estimation of the reduced form.

We now consider the results for the NKPC. Given the evidence above, the predicted values of expected inflation are constructed allowing for breaks at 1975.2 and 1981.1, and the predicted value for the output gap is constructed allowing for a break at 1975.2. As with the reduced forms, we assume that the maximum number of breaks is 5 and set $\epsilon = 0.1$. The results from the 2SLS estimations of the NKPC are given in Table 14. As with the reduced forms, there is evidence of instability from the sup-F tests. Using the sequential strategy, we estimate there to be only one break located at 1975.1.[17] Parenthetically, we note that if the number of breaks is

[15] These calculations are made using the code available from http://people.bu.edu/perron/code.html. All hypotheses are tested with F-statistics which are the OLS analogs of those discussed in the text; further details can be found in Bai and Perron (1998).

[16] For ease of presentation, we define the BIC criterion below for 2SLS; the appropriate modification for OLS is then obvious.

[17] We note that it was not possible to calculate the test of the four break model against the five break model because the location of the breaks in the four break model meant certain sub-samples in the five break model were too small.

chosen by minimizing the BIC,

$$BIC(m) \; = \; ln[\min_{T_1,...,T_m} S_T(T_1,...,T_m;\; \hat{\beta}(\{T_i\}_{i=1}^m))/T] \; + \; m(p+1)ln(T)/T$$

then the estimated number is also one and the location is again 1975.1.

The estimated NKPC is as follows (omitting the error and with estimates to 2dp):

for 1969.1-1975.2:

$$inf_t \;=\; -4.45 + 0.52 inf^e_{t+1|t} + 1.48 inf_{t-1} + 0.39 og_t - 1.39 \Delta inf_{t-1} - 1.05 \Delta inf_{t-2} - 0.37 \Delta inf_{t-3}$$

for 1975.3-2001.4:

$$inf_t \;=\; -0.27 + 0.69 inf^e_{t+1|t} + 0.33 inf_{t-1} + 0.11 og_t - 0.16 \Delta inf_{t-1} - 0.13 \Delta inf_{t-2} - 0.28 \Delta inf_{t-3}$$

Of particular interest are the coefficients on expected and lagged inflation as they reflect the degree to which policy is forward or backward looking respectively. One most striking difference between the two periods is in the coefficient on lagged inflation. Our results suggest that this variable plays a far weaker role in the post-1975 sample. However, one important caveat is the small size of the pre-1975.2 subsample.

It is interesting to note that our results closely match Zhang, Osborn, and Kim's (2007) findings with regard to both the number of breaks and the location of the break.[18] However, we cannot directly compare our estimates as Zhang, Osborn, and Kim (2007) do not report the specific estimates associated with this sample break.

8 Concluding Remarks

In this paper, we extend Bai and Perron's (1998) framework for multiple break testing to linear models estimated via Two Stage Least Squares (2SLS). Within our framework, the break points are estimated simultaneously with the regression parameters via minimization of the residual sum of squares on the second step of the 2SLS estimation. We establish the consistency of the resulting estimated break point fractions. We show that various F-statistics for structural instability based on the 2SLS estimator have the same limiting distribution as the analogous

[18] We note that with other choices of inflation forecast series, Zhang, Osborn, and Kim (2007) find evidence of breaks at other points in the sample.

statistics for OLS considered by Bai and Perron (1998). This allows us to extend Bai and Perron's (1998) sequential procedure for selecting the number of break points to the 2SLS setting.

Our focus is on the stability of the parameters in the structural equation of interest. However to implement 2SLS, it is necessary in the first stage regression to estimate the reduced form for the endogenous regressors in the structural equation of interest and this, of course, requires an assumption about the constancy or lack thereof of these reduced form parameters. In this paper, we establish the aforementioned results under two scenarios of interest, namely: (i) the parameters in the first stage regression are constant; (ii) the parameters in the first stage regression are subject to discrete shifts within the sample period and the locations of these shifts are estimated *a priori* via a data-based method that satisfies certain conditions. The latter conditions allow the case in which the location of the instability is estimated via an application of Bai and Perron's (1998) methods to the appropriate reduced form equations on an equation by equation basis. We have illustrated the empirical relevance of our framework via an application to the New Keynesian Phillips curve. Most empirical investigations of the NKPC assume the parameters are constant. However, our results indicate that if estimated over 1968-2001 then this relationship is not stable.

In practice, a researcher may also be interested in performing inference about the timing of the structural changes. Hall, Han, and Boldea (2007) provide a distribution theory for the break fraction estimators in the case where the reduced form regression parameters are structurally stable. The extension of this theory to the case in which the reduced form exhibits parameter variation is complicated by the potential dependence on the limiting distribution of the estimated break fractions in the structural equations on that of the estimated break fractions from the reduced form. This extension is work in progress.

In two recent papers, Perron and Qu extend Bai and Perron's (1998) framework in a number of interesting ways. Qu and Perron (2007) consider estimation and inference of mutliple structural changes in systems of regression equations, and show that there are efficiency gains from estimation of the system rather than on an equation by equation basis. Perron and Qu (2006) show that there are also efficiency gains from imposing cross-regime restrictions, such as the equality of parameters in two non-adjacent regimes. It would be interesting to explore the potential for such efficiency gains within the context of our 2SLS framework; however, these extensions are beyond the scope of the current paper and are left to future research.

Mathematical Appendix

For brevity, the appendix is omitted from this version of the paper but is available from the authors upon request.

Table 1: Finite sample behavior of break fraction estimator

one break model with $(\beta_1^0, \beta_2^0) = (0.1, -0.1)$

q	T	Deviation from the True Break Fraction				
		1 %	2 %	3 %	5 %	10 %
2	60	.19	.37	.37	.58	.73
	120	.38	.51	.58	.74	.89
	240	.51	.69	.80	.90	.97
	480	.66	.84	.92	.98	1.00
4	60	.27	.50	.50	.71	.86
	120	.49	.65	.73	.86	.95
	240	.66	.82	.92	.97	1.00
	480	.82	.96	.98	.99	1.00
8	60	.37	.62	.62	.84	.95
	120	.66	.79	.88	.95	.99
	240	.78	.91	.97	.99	1.00
	480	.92	1.00	1.00	1.00	1.00

Notes: The column headed $100a\%$ gives the proportion of the simulations in which $|\hat{\lambda}_1 - \lambda_1^0| \leq a$; q is the number of instruments; T is the sample size.

Table 2: Relative rejection frequencies of F-statistics

one break model: $(\beta_1^0, \beta_2^0) = (0.1, -0.1)$

q	T	supF(k)		supF(l+1:l)		UDmax
		1	2	2:1	3:2	
2	60	.68	.53	.04	0	.67
	120	.96	.89	.04	0	.95
	240	1.00	1.00	.05	.01	1.00
	480	1.00	1.00	.03	0	1.00
4	60	.93	.87	.06	.01	.92
	120	1.00	1.00	.05	.01	1.00
	240	1.00	1.00	.06	.01	1.00
	480	1.00	1.00	.06	.01	1.00
8	60	1.00	.99	.07	.01	1.00
	120	1.00	1.00	.07	.01	1.00
	240	1.00	1.00	.08	.01	1.00
	480	1.00	1.00	.06	.01	1.00

Notes: supF(k) denotes the statistic $Sup - F_T(k; 1)$ and the second tier column heading under it denotes k; F(l+1:l) denotes the statistic $F_T(l + 1|l)$ and the second tier column beneath it denotes $l + 1 : l$; UDmax denotes the statistic $UDmaxF_T(5, 1)$; q is the number of instruments; T is the sample size.

Table 3: Relative rejection frequencies of F-statistics

one break model: $(\beta_1^0, \beta_2^0) = (1, -1)$

q	T	supF(k)		supF(l+1:l)		UDmax
		1	2	2:1	3:2	
2	60	1.00	1.00	.07	.02	1.00
	120	1.00	1.00	.04	.01	1.00
	240	1.00	1.00	.05	.01	1.00
	480	1.00	1.00	.04	.01	1.00
4	60	1.00	1.00	.09	.03	1.00
	120	1.00	1.00	.08	.02	1.00
	240	1.00	1.00	.07	.02	1.00
	480	1.00	1.00	.06	.01	1.00
8	60	1.00	1.00	.07	.03	1.00
	120	1.00	1.00	.05	.02	1.00
	240	1.00	1.00	.08	.02	1.00
	480	1.00	1.00	.06	.01	1.00

Notes: See Table 2 for definitions.

Table 4: Empirical distribution of the estimated number of breaks

one break model: $(\beta_1^0, \beta_2^0) = (0.1, -0.1)$

q	T	supF(1)				UDmax			
		0	1	2	3,4,5	0	1	2	3,4,5
2	60	.32	.66	.02	0	.33	.65	.02	0
	120	.05	.94	.01	0	.05	.93	.02	0
	240	0	.97	.03	0	0	.97	.03	0
	480	0	.99	.01	0	0	.99	.01	0
4	60	.07	.90	.03	0	.08	.89	.03	0
	120	0	.97	.03	0	0	.97	.03	0
	240	0	.97	.03	0	0	.97	.03	0
	480	0	.97	.03	0	0	.97	.03	0
8	60	0	.96	.04	0	0	.96	.04	0
	120	0	.96	.04	0	0	.96	.04	0
	240	0	.96	.04	0	0	.96	.04	0
	480	0	.96	.04	0	0	.96	.04	0

Notes: The figures in the block headed supF(1) give the empirical distribution of the estimated number of breaks, $\hat{k}_T$, obtained via the sequential strategy using $Sup-F_T(1;1)$ on the first step with the maximum number of breaks set equal to five. The figures in the block UDmax give the empirical distribution of the estimated number of breaks, $\hat{k}_T$,obtained via the sequential strategy using $UDmaxF_T(5,1)$ on the first step with the maximum number of breaks set equal to five.

Table 5: Empirical distribution of the estimated number of breaks

one break model: $(\beta_1^0, \beta_2^0) = (1, -1)$

q	T	supF(1)				UDmax			
		0	1	2	3,4,5	0	1	2	3,4,5
2	60	0	.96	.04	0	0	.96	.04	0
	120	0	.98	.02	0	0	.98	.02	0
	240	0	.97	.03	0	0	.97	.03	0
	480	0	.98	.02	0	0	.98	.02	0
4	60	0	.94	.06	0	0	.94	.06	0
	120	0	.95	.05	0	0	.95	.05	0
	240	0	.96	.04	0	0	.96	.04	0
	480	0	.97	.03	0	0	.97	.03	0
8	60	0	.96	.04	0	0	.96	.04	0
	120	0	.97	.03	0	0	.97	.03	0
	240	0	.95	.05	0	0	.95	.05	0
	480	0	.97	.03	0	0	.97	.03	0

Notes: See Table 4 for definitions.

Table 6: Finite sample behavior of break fraction estimator

two break model: $(\beta_1^0, \beta_2^0, \beta_3^0) = (-0.1, 0.1, -0.1)$

q	T	i-th Break	Deviation from the True Break Fraction 1 %	2 %	3 %	5 %	10 %
2	60	1st	.19	.37	.37	.52	.71
		2nd	.19	.38	.38	.59	.74
	120	1st	.36	.48	.56	.71	.87
		2nd	.37	.51	.58	.75	.89
	240	1st	.52	.67	.80	.89	.97
		2nd	.51	.65	.79	.88	.96
	480	1st	.68	.86	.94	.98	1.00
		2nd	.68	.86	.93	.97	1.00
4	60	1st	.26	.47	.47	.66	.87
		2nd	.26	.49	.49	.72	.83
	120	1st	.49	.61	.70	.84	.95
		2nd	.50	.60	.69	.84	.94
	240	1st	.67	.81	.91	.97	1.00
		2nd	.64	.81	.90	.96	.99
	480	1st	.80	.95	.98	.99	1.00
		2nd	.82	.94	.98	1.00	1.00
8	60	1st	.36	.64	.64	.79	.94
		2nd	.34	.60	.60	.81	.92
	120	1st	.66	.77	.84	.93	.99
		2nd	.61	.75	.85	.94	.99
	240	1st	.79	.91	.97	.99	1.00
		2nd	.78	.91	.97	1.00	1.00
	480	1st	.92	.99	1.00	1.00	1.00
		2nd	.91	.98	1.00	1.00	1.00

Notes: See Table 1 for definitions.

Table 7: Relative rejection frequencies of F-statistics

two break model: $(\beta_1^0, \beta_2^0, \beta_3^0) = (-0.1, 0.1, -0.1)$

q	T	supF(k)		supF(l+1:l)		UDmax
		1	2	2:1	3:2	
2	60	.18	.48	.38	.02	.38
	120	.42	.84	.71	.01	.76
	240	.87	1.00	.99	.01	1.00
	480	1.00	1.00	1.00	.02	1.00
4	60	.30	.77	.65	.03	.65
	120	.73	.99	.95	.02	.98
	240	.99	1.00	1.00	.03	1.00
	480	1.00	1.00	1.00	.03	1.00
8	60	.49	.96	.89	.05	.93
	120	.97	1.00	1.00	.05	1.00
	240	1.00	1.00	1.00	.03	1.00
	480	1.00	1.00	1.00	.04	1.00

Notes: See Table 2 for definitions.

Table 8: Relative rejection frequencies of F-statistics

two break model: $(\beta_1^0, \beta_2^0, \beta_3^0) = (-1, 1, -1)$

q	T	supF(k)		supF(l+1:l)		UDmax
		1	2	2:1	3:2	
2	60	1.00	1.00	1.00	.06	1.00
	120	1.00	1.00	1.00	.04	1.00
	240	1.00	1.00	1.00	.03	1.00
	480	1.00	1.00	1.00	.03	1.00
4	60	1.00	1.00	1.00	.05	1.00
	120	1.00	1.00	1.00	.04	1.00
	240	1.00	1.00	1.00	.04	1.00
	480	1.00	1.00	1.00	.04	1.00
8	60	1.00	1.00	1.00	.07	1.00
	120	1.00	1.00	1.00	.05	1.00
	240	1.00	1.00	1.00	.05	1.00
	480	1.00	1.00	1.00	.04	1.00

Notes: See Table 2 for definitions.

Table 9: Empirical distribution of the estimated number of breaks

two break model: $(\beta_1^0, \beta_2^0, \beta_3^0) = (-0.1, 0.1, -0.1)$

q	T	supF(1)				UDmax			
		0	1	2	3,4,5	0	1	2	3,4,5
2	60	.83	.11	.06	0	.63	.15	.21	.01
	120	.58	.13	.28	.01	.24	.16	.58	.02
	240	.13	.02	.82	.03	0	.02	.95	.03
	480	0	0	.96	.04	0	0	.96	.04
4	60	.70	.10	.19	.01	.35	.17	.45	.03
	120	.27	.05	.64	.04	.02	.06	.86	.06
	240	.01	0	.95	.04	0	0	.96	.04
	480	0	0	.94	.06	0	0	.94	.06
8	60	.51	.07	.39	.03	.07	.10	.78	.05
	120	.04	0	.90	.06	0	0	.93	.07
	240	0	0	.95	.05	0	0	.95	.05
	480	0	0	.96	.04	0	0	.96	.04

Notes: See Table 4 for definitions.

Table 10: Empirical distribution of the estimated number of breaks

two break model: $(\beta_1^0, \beta_2^0, \beta_3^0) = (-1, 1, -1)$

q	T	supF(1)				UDmax			
		0	1	2	3,4,5	0	1	2	3,4,5
2	60	0	0	.95	.05	0	0	.95	.05
	120	0	0	.97	.03	0	0	.97	.03
	240	0	0	.98	.02	0	0	.98	.02
	480	0	0	.98	.02	0	0	.98	.02
4	60	0	0	.96	.04	0	0	.96	.04
	120	0	0	.98	.02	0	0	.98	.02
	240	0	0	.97	.03	0	0	.97	.03
	480	0	0	.97	.03	0	0	.97	.03
8	60	0	0	.95	.05	0	0	.95	.05
	120	0	0	.96	.04	0	0	.96	.04
	240	0	0	.97	.03	0	0	.97	.03
	480	0	0	.97	.03	0	0	.97	.03

Notes: See Table 4 for definitions.

Table 11: Rejection frequency of test statistics

no break model

q	T	supF(k)		supF(l+1:l)		UDmax
		1	2	2:1	3:2	
2	60	.03	.03	.03	0	.03
	120	.03	.02	.03	0	.03
	240	.03	.02	.02	0	.02
	480	.02	.01	.02	0	.02
4	60	.03	.03	.04	.01	.04
	120	.04	.03	.04	0	.04
	240	.04	.03	.03	0	.03
	480	.04	.04	.03	0	.03
8	60	.04	.04	.04	.01	.04
	120	.04	.05	.04	0	.04
	240	.05	.05	.04	0	.04
	480	.06	.05	.04	0	.06

Notes: See Table 2 for definitions.

Table 12: Empirical distribution of the estimated number of breaks

no break model

q	T	supF(1)				UDmax			
		0	1	2	3	0	1	2	3
2	60	.97	.03	0	0	.97	.03	0	0
	120	.97	.03	0	0	.97	.03	0	0
	240	.97	.03	0	0	.98	.02	0	0
	480	.98	.02	0	0	.98	.02	0	0
4	60	.97	.03	0	0	.97	.03	0	0
	120	.96	.04	0	0	.96	.03	.01	0
	240	.96	.04	0	0	.97	.03	0	0
	480	.96	.04	0	0	.97	.03	0	0
8	60	.96	.04	0	0	.96	.03	.01	0
	120	.96	.04	0	0	.96	.04	0	0
	240	.96	.04	0	0	.96	.04	0	0
	480	.94	.05	.01	0	.95	.04	.01	0

Notes: See Table 4 for definitions.

Table 13: Application to NKPC - stability statistics for the reduced forms

Dep.var	k	sup-F	F(k+1:k)	BIC
$inf^e_{t+1\|t}$	0			-0.615
	1	43.6	41.7	-0.623
	2	67.0	10.4	-0.680
	3	176.5	34.3	-0.649
	4	80.5	46.8	-0.452
	5	70.2		-0.369
og_t				
	0			-0.663
	1	50.0	30.53	-0.552
	2	40.1	23.1	-0.497
	3	40.	11.3	-0.276
	4	34.91	11.3	-0.046
	5	31.9		0.255

Notes: Dep. Var. denotes the dependent variable in the reduced form; sup-F denotes the statistic for testing $H_0 : m = 0$ vs. $H_1 : m = k$; F(k+1:k) is the statistic for testing $H_0 : m = k$ vs. $H_1 : m = k + 1$; BIC is the BIC criterion. The percentiles for the statistics are for $k = 1, 2, \ldots$ respectively: (i) sup-F: (10%, 1%) significance level = (25.29, 32.8), (23.33, 28.24), (21.89, 25.63), (20.71, 23.83), (19.63,22.32); (ii) F(k+1:k): (10%, 1%) significance level =(25.29, 32.8), (27.59,34.81), (28.75, 36.32), (29.71,36.65).

Table 14: Application to NKPC - stability statistics for structural equation

k	sup-F	F(k+1:k)	BIC
0			0.021
1	41.3	9.55	0.017
2	25.0	7.83	0.240
3	21.4	12.8	0.427
4	17.4		0.664
5	13.4		0.942

Notes: Dep. Var. denotes the dependent variable in the reduced form; sup-F denotes the statistic for testing $H_0 : m = 0$ vs. $H_1 : m = k$; F(k+1:k) is the statistic for testing $H_0 : m = k$ vs. $H_1 : m = k + 1$; BIC is the BIC criterion. The percentiles for the statistics are for $k = 1, 2, \ldots$ respectively: (i) sup-F: (10%, 1%) significance level = (19.7, 26.71), (17.67, 21.87), (16.04, 19.42), (14.55, 17.44), (12.59,15.02); (ii) F(k+1:k): (10%, 1%) significance level =(19.7, 26.71), (21.79, 28.36), (22.87, 29.30).

References

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

Andrews, D. W. K., and Fair, R. (1988). 'Inference in econometric models with structural change', *Review of Economic Studies*, 55: 615–640.

Bai, J. (1994). 'Least squares estimation of a shift in linear processes', *Journal of Time Series Analysis*, 15: 453–472.

Bai, J., and Perron, P. (1998). 'Estimating and testing linear models with multiple structural changes', *Econometrica*, 66: 47–78.

——— (2001). 'Additional critical values for multiple structural change tests', Discussion paper, Department of Economics, Boston University, Boston, MA.

Bhattacharya, P. K. (1987). 'Maximum Likelihood estimation of a change-point in the distribution of independent random variables: general multiparameter case', *Journal of Multivariate Analysis*, 23: 183–208.

de Jong, R. M., and Davidson, J. (2000). 'THE Functional Central Limit Theorem and Weak Convergence to Stochastic Integrals I', *Econometric Theory*, 16: 621–642.

Ghysels, E., and Hall, A. R. (1990). 'A test for structural stability of Euler condition parameters estimated via the Generalized Method of Moments', *International Economic Review*, 31: 355–364.

Hahn, J., and Inoue, A. (2002). 'A Monte Carlo comparison of various asymptotic approximations to the distribution of instrumental variables estimators', *Econometric Reviews*, 21: 309–336.

Hall, A. R. (2005). *Generalized Method of Moments.* Oxford University Press, Oxford, U.K.

Hall, A. R., Han, S., and Boldea, O. (2007). 'A distribution theory for change point estimators in models estimated by Two Stage Least Squares', Discussion paper, Department of Economics, North Carolina State University, Raleigh, NC.

Hall, A. R., and Sen, A. (1999). 'Structural stability testing in models estimated by Generalized Method of Moments', *Journal of Business and Economic Statistics*, 17: 335–348.

Hawkins, D. L. (1986). 'A simple least square method for estimating a change in mean', *Communications in Statistics - Simulation*, 15: 655–679.

Perron, P., and Qu, Z. (2006). 'Estimating restricted structural change models', *Journal of Econometrics*, 134: 373–399.

Picard, D. (1985). 'Testing and estimating change points in time series', *Journal of Applied Probability*, 20: 411–415.

Qu, Z., and Perron, P. (2007). 'Estimating and testing structural changes in multivariate regressions', *Econometrica*, 75: 459–502.

Sowell, F. (1996). 'Optimal tests of parameter variation in the Generalized Method of Moments framework', *Econometrica*, 64: 1085–1108.

Yao, Y.-C. (1987). 'Approximating the distribution of the ML estimate of the chnage point in a sequence of independent r.v.'s', *Annals of Statistics*, 4: 1321–1328.

Zhang, C., Osborn, D. R., and Kim, D. H. (2007). 'The new Keynesian Phillips curve: from sticky inflation to sticky prices', *Journal of Money, Credit and Banking*, forthcoming.