Estimating and Testing Multiple Structural Changes in Models with Endogenous Regressors*

Pierre Perron[†] **Yohei Yamamoto**[‡]

Boston University **Boston University**

June 16, 2008; Revised: October 9, 2008

Abstract

We consider the problem of estimating and testing for multiple breaks in a single equation framework with regressors that are endogenous, i.e., correlated with the errors. First, we show based on standard assumptions about the regressors, instruments and errors that the second stage regression of the instrumental variable (IV) procedure involves regressors and errors that satisfy all the assumptions in Perron and Qu (2006) so that the results about consistency, rate of convergence and limit distributions of the estimates of the break dates, as well as the limit distributions of the tests, are obtained as simple consequences. More importantly from a practical perspective, we show that even in the presence of endogenous regressors, it is still preferable to simply estimate the break dates and test for structural change using the usual ordinary least-squares (OLS) framework. It delivers estimates of the break dates with higher precision and tests with higher power compared to those obtained using an IV method. To illustrate the relevance of our theoretical results, we consider the stability of the New Keynesian hybrid Phillips curve. IV-based methods do not indicate any instability. On the other hand, OLS-based ones strongly indicate a change in 1991:1 and that after this date the model looses all explanatory power.

JEL Classification Number: C22.

Keywords: structural change, instrument variables, two-stage least-squares, parameter variations.

*Perron acknowledges financial support for this work from the National Science Foundation under Grant SES-0649350. We are grateful to Zhongjun Qu for useful comments.

[†]Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215 (perron@bu.edu).

[‡]Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215 (yyamam@bu.edu).

1 Introduction

Both the statistics and econometrics literature contain a vast amount of work on issues related to structural changes with unknown break dates, most of it specifically designed for the case of a single change (see, Perron, 2006, for a detailed review). However, the problem of multiple structural changes has received more attention recently. Bai and Perron (1998, 2003) provided a comprehensive treatment of various issues in the context of multiple structural change models: consistency of estimates of the break dates, tests for structural changes, confidence intervals for the break dates, methods to select the number of breaks and efficient algorithms to compute the estimates. Related contributions include Hawkins (1976) who presents a comprehensive treatment of estimation based on a dynamic programming algorithm. Perron and Qu (2006) extended the analysis to the case where arbitrary linear restrictions are imposed on the coefficients of the model. In doing so, they also considerably relaxed the assumptions used in Bai and Perron (1998). Bai, Lumsdaine and Stock (1998) considered asymptotically valid inference for the estimate of a single break date in multivariate time series allowing stationary or integrated regressors as well as trends with estimation carried using a quasi maximum likelihood (QML) procedure. Also, Bai (2000) considered the consistency, rate of convergence and limiting distribution of estimated break dates in a segmented stationary VAR model estimated again by QML when the break can occur in the parameters of the conditional mean, the variance of the error term or both. Qu and Perron (2007) considered a multivariate system estimated by quasi maximum likelihood which provides methods to estimate models with structural changes in both the regression coefficients and the covariance matrix of the errors. Kejriwal and Perron (2006a,b) provide a comprehensive treatment of issue related to testing and inference with multiple structural changes in a single equation cointegrated model. Zhou and Perron (2007) considered the problems of testing jointly for changes in regression coefficients and variance of the errors.

More recently, Hall and Han (2006, 2007) considered a single equation framework with regressors that are endogenous, i.e., correlated with the errors. They considered using an instrumental variable (IV) procedure. Hall and Han (2006) provided a very detailed proof for the consistency and rate of convergence of the estimates of the break fractions and the limit distributions of the tests for multiple structural changes. Hall and Han (2007) provided a detailed proof of the limit distribution of the estimate of the break date in a single break model assuming a stable reduced form equation. In all cases, the results are similar to those in Bai and Perron (1998).

This paper follows up on the work of Hall and Han (2006, 2007) in two directions. First, we provide a very simple proof of the results they derived by showing that using generated regressors, the projection of the regressors of the space spanned by the instruments, to account for potential endogeneity implies that all the assumptions of Perron and Qu (2006) (or those of Bai and Perron, 1998) obtained with original regressors contemporaneously uncorrelated with the errors, are satisfied. Hence, all results derived in those papers continue to hold. There is no need for a separate detailed analysis as done in Hall and Han (2006, 2007). This also allows us to have even more general results, in particular about the limit distributions of the estimates of the break dates for which we do not require a stable reduced form equation linking the regressors and the instruments.

The second contribution is more important from a practical perspective. We show that even in the presence of endogenous regressors, it is still preferable to simply estimate the break dates and test for structural change using the usual ordinary least-squares (OLS) framework. The idea is simple yet compelling. First, changes in the true parameters of the model imply a corresponding change in the probability limits of the OLS parameter estimates, which is equivalent in the leading case of regressors and errors that have a homogenous distribution across segments. Second, one can reformulate the model with those probability limits as the basic parameters in a way that the regressors and errors are contemporaneously uncorrelated. We are then simply back to the framework of Bai and Perron (1998) or Perron and Qu (2006) and we can use their results directly to obtain the relevant limit distributions. What is more important is that the OLS framework involves the original regressors, while the IV framework involves as regressors the projection of these original regressors on the space spanned by the instruments. As is well known, this implies that the generated regressors in the IV procedure have less quadratic variation than the original regressors. Hence, a given change in the true parameters will cause a larger change in the conditional mean of the dependent variable in the OLS framework compared to the corresponding change in an IV framework. Accordingly, using OLS not only delivers consistent estimates of the break fractions and tests with the usual limit distributions, it also improves on the efficiency of the estimates and the power of the tests. This is shown theoretically and also via simulations.

To illustrate the relevance of our theoretical results, we consider the stability of the New Keynesian hybrid Phillips curve as put forward by Gali and Gertler (1999). The results show that IV-based methods do not provide any evidence of structural instability. On the other hand, an OLS-based sup-Wald test for one break is highly significant, with 1991:1 as the estimate of the break date (with a very tight confidence interval (1990:4 to 1991:2)). The higher discriminatory power of OLS-based methods over IV-based ones occurs despite the fact that the instruments are highly correlated with future inflation. The estimates for the period 1960:1-1991:1 are close to the full sample estimates reported earlier and support Gali and Gertler's (1999) conclusion. On the other hand, the estimates for the period 1991:2-1997:4 are all very small and insignificantly different from zero. Hence, the hybrid New Keynesian Phillips curve specification has lost any explanatory power for inflation. This forecast breakdown of the New Keynesian Phillips curve for inflation is interesting and can be traced back to the change in the behavior of inflation (e.g., Stock and Watson, 2007).

The structure of the paper is as follows. Section 2 presents the model and the assumptions on the regressors, instruments and errors. We adopt those in Perron and Qu (2006) since they are the most general available and also allows for structural change models with restrictions on the parameters, including partial structural change models. Section 3 shows that these assumptions imply that the second stage regression of the instrumental variable procedure involves regressors and errors that satisfy all the assumptions in Perron and Qu (2006) so that the results about consistency, rate of convergence and limit distribution of the estimates of the break dates, as well as the limit distribution of the tests, are obtained as simple consequences. Section 4 shows how using the OLS framework is not only valid but also preferable in that it delivers estimates of the break dates with higher precision and tests with higher power. Section 5 substantiates our theoretical results via simulations and shows their practical importance. Section 6 presents our empirical illustration related to the New Keynesian hybrid Phillips curve. Section 7 provides brief concluding remarks and an appendix contains some technical derivations.

2 The model and assumptions

Consider a general multiple linear regression model with m breaks or $m+1$ regimes. There are T observations and m is assumed known. The break dates occur at $\{T_1, ..., T_m\}$. Let $y = (y_1, ..., y_T)'$ be the dependent variable and X a T by p matrix of regressors. Define $\bar{X} = diag(X_1, ..., X_{m+1})$, a T by $(m+1)p$ matrix with $X_i = (x_{T_{i-1}+1}, ..., x_{T_i})'$ for $i = 1, ..., m+1$, with the convention that $T_0 = 1$ and $T_{m+1} = T$ (each matrix X_i is a subset of the regressor matrix X corresponding to regime i). The matrix $\bar{X}$ is a diagonal partition of X, the partition being taken with respect to the set of break points $\{T_1, ..., T_m\}$. The vector $u = (u_1, ..., u_T)'$ is the set of disturbances and $\delta = (\delta_1', ..., \delta_{m+1}')'$ is the $(m+1)p$ vector of coefficients.

Following Perron and Qu (2006), we consider the general pure structural change model with restrictions on the coefficients, i.e.,

$$y = \bar{X}\delta + u, \tag{1}$$

where

$$R\delta = r, \tag{2}$$

with R a k by $(m+1)p$ matrix with rank k and r a k dimensional vector of constants. Note that this framework includes the case of a partial structural change model by an appropriate choice of the restrictions on the parameters. Since some regressors may be correlated with the errors, we assume that there exists a set of q variables z_t that can serve as instruments, and we define the T by q matrix $Z = (z_1, ..., z_T)'$.

The goal here is to estimate the unknown break dates whose true values are denoted with a 0 superscript, i.e., $(T_1^0, ..., T_m^0)$, using the observables (y, X, Z) and the restrictions on the coefficients. The relevant IV regression is then

$$y = \bar{X}^*\delta + u, \tag{3}$$

subject to the restrictions given by (2), where $\bar{X}^* = diag(\hat{X}_1, ..., \hat{X}_{m+1})$, a T by $(m+1)p$ matrix with $\hat{X}_i = (\hat{x}_{T_{i-1}+1}, ..., \hat{x}_{T_i})'$ for $i = 1, ..., m+1$, and $\hat{X} = (\hat{x}_1, ..., \hat{x}_T)' = P_Z X$ where $P_Z = Z(Z'Z)^{-1}Z'$ is the orthogonal projection matrix onto the space spanned by the columns of Z. The estimates of the break dates are then

$$(\hat{T}_1, ..., \hat{T}_m) = \arg \min_{T_1, ..., T_m} SSR_T^R(T_1, ..., T_m), \tag{4}$$

where $SSR_T^R(T_1, ..., T_m)$ is the sum of square residuals from the restricted OLS regression (3) evaluated at the partition $\{T_1, ..., T_m\}$. It will be useful also to define the break fractions $(\lambda_1^0, ..., \lambda_m^0) = (T_1^0/T, ..., T_m^0/T)$ with corresponding estimates $(\hat{\lambda}_1, ..., \hat{\lambda}_m) = (\hat{T}_1/T, ..., \hat{T}_m/T)$.

We impose the following assumptions on the data, the errors and the break dates, which are trivial extensions of those in Perron and Qu (2006).

- **Assumption A1**: Let $w_t = (x_t, z_t)$. For each $i = 1, ..., m+1$, let $l_i = (T_{i+1}^0 - T_i^0)$, then

$$(1/l_i) \sum_{t=T_i^0+1}^{T_i^0+[l_i v]} w_t w_t' \rightarrow_p Q^i(v) \equiv \begin{bmatrix} Q_{XX}^i(v) & Q_{XZ}^i(v) \\ Q_{ZX}^i(v) & Q_{ZZ}^i(v) \end{bmatrix},$$

a non-random positive definite matrix uniformly in $v \in [0, 1]$.

- **Assumption A2**: There exists an $l_0 > 0$ such that for all $l > l_0$, the minimum eigenvalues of $(1/l) \sum_{t=T_i^0+1}^{T_i^0+l} z_t z_t'$ and of $(1/l) \sum_{t=T_i^0-l}^{T_i^0} z_t z_t'$ are bounded away from zero ($i = 1, ..., m$).

• **Assumption A3**: The matrix $\sum_{t=k}^{l} z_t z_t'$ is invertible for $l - k \geq \epsilon T$ for some $\epsilon > 0$.

• **Assumption A4**: Let the L_r-norm of a random matrix A be defined by $\|A\|_r = (\sum_i \sum_j E|A_{ij}|^r)^{1/r}$ for $r \geq 1$. (Note that $\|A\|$ is the usual matrix norm or the Euclidean norm of a vector.) With $\{\mathcal{F}_i : i = 1, 2, ..\}$ a sequence of increasing σ-fields, we assume that $\{z_i u_i, \mathcal{F}_i\}$ forms a L^r-mixingale sequence with $r = 2 + \varepsilon$ for some $\varepsilon > 0$. That is, there exist nonnegative constants $\{c_i : i \geq 1\}$ and $\{\psi_j : j \geq 0\}$ such that $\psi_j \downarrow 0$ as $j \to \infty$ and for all $i \geq 1$ and $j \geq 0$, we have: (a) $\|E(z_i u_i | \mathcal{F}_{i-j})\|_r \leq c_i \psi_j$, (b) $\|z_i u_i - E(z_i u_i | \mathcal{F}_{i+j})\|_r \leq c_i \psi_{j+1}$. Also assume (c) $\max_i c_i \leq K < \infty$, (d) $\sum_{j=0}^{\infty} j^{1+k} \psi_j < \infty$, (e) $\|z_i\|_{2r} < M < \infty$ and $\|u_i\|_{2r} < N < \infty$ for some $K, M, N > 0$.

• **Assumption A5**: $T_i^0 = [T\lambda_i^0]$, where $0 < \lambda_1^0 < ... < \lambda_m^0 < 1$.

• **Assumption A6**: The minimization problem defined by (4) is taken over all possible partitions such that $T_i - T_{i-1} \geq \epsilon T$ for some $\epsilon > 0$.

Assumption A1 basically rules out unit root regressors; otherwise the particular scaling used is not important and could be relaxed at the expense of substantial technical complications [1]. Assumption A2 imposes restrictions on the instruments in a local neighborhood of the break points. They ensure that there is no local collinearity problem so the break points can be identified. Assumption A3 is a standard invertibility requirement to have well defined estimates. Assumption A4 imposes mild restrictions on the vector $z_t u_t$. They permit a wide class of potential correlation and heterogeneity (including conditional heteroskedasticity) and also allow lagged dependent variables. Finally, Assumption A5 is a standard requirement to have asymptotically distinct break dates and A6 requires that the search for breaks precludes candidates which are too close. This, however, is not constraining in practice since ϵ can be chosen arbitrarily small.

For future references, note that A1 implies that

$$T^{-1} \sum_{t=1}^{T} z_t z_t' = \sum_{i=1}^{m+1} \frac{l_i}{T} [l_i^{-1} \sum_{t=T_i^0+1}^{T_{i+1}^0} z_t z_t'] \to_p \sum_{i=1}^{m+1} (\lambda_i - \lambda_{i-1}) Q_{ZZ}^i(1) \equiv Q_{ZZ},$$

with Q_{ZZ} a non-random positive definite matrix. Similarly,

$$T^{-1} \sum_{t=1}^{T} z_t x_t' \to_p \sum_{i=1}^{m+1} (\lambda_i - \lambda_{i-1}) Q_{ZX}^i(1) \equiv Q_{ZX}, \quad (5)$$

[1] For example, allowing for trending data would imply a different scaling but it would not change results related to the consistency and rate of convergence of the estimates of the break dates or the fact that the limit distribution of the other estimates is the same as in the known break date case. For the limit distribution of the estimates of the break dates and the test statistics, this assumption matters and will be strenghtenned.

with Q_{ZX} is a non-random matrix with rank equal to p, and

$$T^{-1}\sum_{t=1}^{T} x_t x_t' \rightarrow_p \sum_{i=1}^{m+1} (\lambda_i - \lambda_{i-1}) Q_{XX}^i(1) \equiv Q_{XX},$$

a non-random matrix with rank p.

Note that unlike Hall and Han (2006), we make no assumption about the relation between the instruments Z and the regressors X other than they are correlated. That is, we do not specify a reduced form linking them. If there is a stable reduced form equation of the form

$$X = Z\theta^0 + v, \tag{6}$$

with

$$E\left[(u_t, v_t)'(u_t, v_t)\right] = \begin{bmatrix} \sigma^2 & \gamma' \\ \gamma & \Sigma \end{bmatrix},$$

then, $Q_{ZX} = Q_{ZZ}\theta^0$ and $Q_{XX} = \theta^{0\prime} Q_{ZZ}\theta^0 + \Sigma$. The framework can also accommodate a reduced form that exhibits structural changes which are estimated in the usual way. As in Hall and Han (2006), suppose that

$$X = \bar{Z}^0\theta^0 + v, \tag{7}$$

with $\bar{Z}^0 = diag(Z_1^0, ..., Z_n^0)$, the diagonal partition of Z at the (possibly different) break dates $(T_1^z, ..., T_n^z)$ and θ^0 re-defined accordingly. Suppose that we have estimates $(\hat{T}_1^z, ..., \hat{T}_n^z)$ and we use as instruments $\hat{Z} = diag(\hat{Z}_1, ..., \hat{Z}_n)$, a T by $(n+1)q$ matrix with $\hat{Z}_i = (z_{\hat{T}_{i-1}^z+1}, ..., z_{\hat{T}_i^z})'$ for $i = 1, ..., n+1$. This fits in the framework discussed above by augmenting the number of instruments so that there are n such set of instruments. Then, provided probability limit values for the estimates of the break fractions $(\hat{\lambda}_1^z, ..., \hat{\lambda}_n^z) = (\hat{T}_1^z/T, ..., \hat{T}_n^z/T)$ exists, Assumptions A1 still holds with the limit values properly re-defined. Note that the estimates of the break fractions need not be consistent. In fact, even ignoring breaks in the reduced form when some are present is a case that still fits in our framework. There is no need for a separate treatment as far as consistency and the rate of convergence is concerned. However, a proper treatment of breaks in the reduced form can lead to more precise estimates of the break dates, but Hall and Han (2007) do not cover this case.

3 Limit results for estimates and tests obtained using the IV method

Hall and Han (2006) provided a very detailed proof for the consistency and rate of convergence of the estimates of the break fractions and the limit distributions of the tests for multiple structural changes. They showed that all the results in Bai and Perron (1998) go through. Hall and Han (2007) provided a detailed proof of the limit distribution of the estimate of the break date in a single break model assuming a stable reduced form equation of the form (6). We shall prove all those results by simply showing that the assumptions imposed here, which are less restrictive than those of Hall and Han (2006, 2007), imply that the conditions stated in Perron and Qu (2006), for regressors and errors that are contemporaneously uncorrelated, continue to hold for the IV regression with $\hat{x}_t$ when the original regressors are contemporaneously correlated with the errors.

We first consider the consistency of the estimates of the break fractions. We have the following lemma proved in the appendix.

Lemma 1 *Assumptions A1-A4 hold if we replace z_t with $\hat{x}_t$.*

Lemma 1 in conjunction with A5 and A6 directly imply that the estimates of the break fractions are consistent. To obtain a rate of convergence, we adopt as is common in this literature a framework in which the magnitude of the shifts may decrease as the sample size increases. This is stated in the following assumption:

Assumption A7. Let $\Delta_{T,i} = \delta_{i+1}^0 - \delta_i^0$. Assume $\Delta_{T,i} = v_T \Delta_i$, for some Δ_i independent of T where $v_T > 0$ is a scalar satisfying either a) v_T is fixed, or b) $v_T \to 0$ and $T^{1/2-\eta} v_T \to \infty$ for some $\eta \in (0, 1/2)$.

Lemma 1 in conjunction with A5-A7 directly imply the following result.

Proposition 1 *Under Assumptions A1-A7: for every $\epsilon > 0$, there exists a $C < \infty$, such that for all large T, $P(|T v_T^2 (\hat{\lambda}_k - \lambda_k^0)| > C) < \epsilon$ for every $k = 1, ..., m$.*

To analyze the limit distributions of the estimates, we need, as in Bai and Perron (1998) and Perron and Qu (2006), some additional assumptions. These are stated as follows.

Assumption A8. Let $\Delta T_i^0 = T_i^0 - T_{i-1}^0$, for $i = 1, ..., m$, and $w_t = (x_t, z_t)$. Then, as $\Delta T_i^0 \to \infty$, uniformly in $s \in [0, 1]$: .

1. $(\Delta T_i^0)^{-1} \sum_{t=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} w_t w_t' \to_p sQ^i \equiv s \begin{bmatrix} Q_{XX}^i & Q_{XZ}^i \\ Q_{ZX}^i & Q_{ZZ}^i \end{bmatrix}$,
2. $(\Delta T_i^0)^{-1} \sum_{t=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} u_t u_t' \to_p s\sigma_i^2$,
3. $(\Delta T_i^0)^{-1} \sum_{t=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} \sum_{r=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} (z_r z_t' u_r u_t)' \to_p s\Omega_{ZU}^i$,

4. $(\Delta T_i^0)^{-1/2} \sum_{t=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} z_t u_t \Rightarrow B_{ZU}^i(s)$ with $B_{ZU}^i(s)$ a multivariate Gaussian process on $[0,1]$ with mean zero and covariance $E[B_{ZU}^i(s)B_{ZU}^i(r)'] = \min\{s,r\}\Omega_{ZU}^i$.

These assumptions are the same as those used in Perron and Qu (2006) for regressors that are uncorrelated with the errors. Part (1) compared to A1 implies, in particular, that the regressors are non-trending. Similarly, part (2) implies that the variance of the errors is fixed within each segment. Parts (3) and (4) are standard. These assumptions imply the following counterparts for the case with $\hat{x}_t$ as the regressors used in the instrumental variable regression (see the appendix for a short proof).

Lemma 2 *Let $\Delta T_i^0 = T_i^0 - T_{i-1}^0$, for $i = 1, ..., m$. Then, as $\Delta T_i^0 \to \infty$, uniformly in $s \in [0,1]$:*

1. $(\Delta T_i^0)^{-1} \sum_{t=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} \hat{x}_t \hat{x}_t' \to_p sQ_{ZX}'Q_{ZZ}^{-1}Q_{ZZ}^iQ_{ZZ}^{-1}Q_{ZX} \equiv sQ_{HH}^i$,
2. $(\Delta T_i^0)^{-1} \sum_{t=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} \sum_{r=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} (\hat{x}_r \hat{x}_t' u_r u_t)' \to_p sQ_{ZX}'Q_{ZZ}^{-1}\Omega_{ZU}^iQ_{ZZ}^{-1}Q_{ZX} \equiv s\Omega_{HU}^i$,
3. $(\Delta T_i^0)^{-1/2} \sum_{t=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} \hat{x}_t u_t \Rightarrow Q_{ZX}'Q_{ZZ}^{-1}B_{ZU}^i(s) \equiv B_{HU}^i(s)$ with $B_{HU}^i(s)$ a multivariate Gaussian process on $[0,1]$ with mean zero and covariance $E[B_{HU}^i(s)B_{HU}^i(u)'] = \min\{s,u\}\Omega_{HU}^i$.

The following result then follows immediately from Perron and Qu (2006).

Proposition 2 *Under Assumptions A1-A8 and that $(\delta_i^{0\prime}\hat{x}_t)^2 \pm \delta_i^{0\prime}\hat{x}_t' u_t$ has a continuous distribution, we have, for $i = 1, ..., m$:*

$$\frac{(\Delta_i' Q_{HH}^i \Delta_i)^2}{\Delta_i' \Omega_{HU}^i \Delta_i} v_T(\hat{T}_i - T_i^0) \to^d \arg\max_s V_H^{(i)}(s),$$

where

$$\begin{aligned} V_H^{(i)}(s) &= W_1^{(i)}(-s) - |s|/2 \text{ if } s \leq 0, \\ V_H^{(i)}(s) &= \sqrt{\xi_H^i}(\phi_H^{i,2}/\phi_H^{i,1})W_2^{(i)}(s) - \xi_H^i |s|/2 \text{ if } s > 0, \end{aligned}$$

with $(\phi_H^{i,1})^2 = \Delta_i'\Omega_{HU}^i\Delta_i/\Delta_i'Q_{HH}^i\Delta_i$, $(\phi_H^{i,2})^2 = \Delta_i'\Omega_{HU}^{i+1}\Delta_i/\Delta_i'Q_{HH}^{i+1}\Delta_i$ and

$$\xi_H^i = \Delta_i' Q_{HH}^{i+1}\Delta_i / \Delta_i' Q_{HH}^i \Delta_i.$$

Note that, unlike the case for the consistency and the rate of convergence, the assumptions about the reduced form and how it is analyzed if non-stable (e.g., using additional instruments based on the identified sub-samples) will affect the limit distribution. In the case of a stable reduced as specified by (7), we have

$$\frac{(\Delta_i'\theta^{0\prime}Q_{ZZ}^i\theta^0\Delta_i)^2}{\Delta_i'\theta^{0\prime}\Omega_{ZU}^i\theta^0\Delta_i}\upsilon_T(\hat{T}_i - T_i^0) \to^d \arg\max_s V_H^{(i)}(s),$$

where

$$\begin{aligned} V_H^{(i)}(s) &= W_1^{(i)}(-s) - |s|/2 \text{ if } s \leq 0, \\ V_H^{(i)}(s) &= \sqrt{\xi_H^i}(\phi_H^{i,2}/\phi_H^{i,1})W_2^{(i)}(s) - \xi_H^i |s|/2 \text{ if } s > 0, \end{aligned}$$

with $(\phi_H^{i,1})^2 = \Delta_i'\theta^{0\prime}\Omega_{ZU}^i\theta^0\Delta_i/\Delta_i'\theta^{0\prime}Q_{ZZ}^i\theta^0\Delta_i$, $(\phi_H^{i,2})^2 = \Delta_i'\theta^{0\prime}\Omega_{ZU}^{i+1}\theta^0\Delta_i/\Delta_i'\theta^{0\prime}Q_{ZZ}^{i+1}\theta^0\Delta_i$ and $\xi_H^i = \Delta_i'\theta^{0\prime}Q_{ZZ}^{i+1}\theta^0\Delta_i/\Delta_i'\theta^{0\prime}Q_{ZZ}^i\theta^0\Delta_i$. The result for this special case is the same as that in Theorem 3 of Hall and Han (2007).

Despite the fact that our result, as stated in Proposition 2, is simpler and involves less parameters to estimate, it is indeed more general since it does not presumes the existence of a stable reduced form equation. All that is required is a consistent estimate of the limit moment matrix of (x_t, z_t) which is trivial to obtain.

Consider now the problem of testing the null hypothesis of no structural change. In the appendix, we show that the following results hold given Assumptions A1-A4 when no structural change is present.

Lemma 3 *Assume that u_t is serially uncorrelated, then, uniformly in s,*

$$a)\ T^{-1}\sum_{t=1}^{[Ts]} \hat{x}_r\hat{x}_t' \to_p sQ_{ZX}'Q_{ZZ}{}^{-1}Q_{ZX} \equiv Q_{HH};$$

and

$$b)\ E(u_t^2) = \sigma^2 \text{ for all } t \text{ and } T^{-1/2}\sum_{t=1}^{[Ts]} \hat{x}_r u_t \Rightarrow \sigma Q_{HH}^{1/2}W_q(s).$$

Hence, assumption A10 in Perron and Qu (2006) is satisfied so that all results pertaining to hypothesis testing remain valid. Relaxing the assumption of serially uncorrelated errors is done by modifying the various tests using a heteroskedasticity and autocorrelation robust covariance matrix for the parameter estimates.

To summarize, we have shown that using the generated regressors $\hat{x}_t$, the projection of the regressors on the space spanned by the instruments, to account for potential endogeneity implies that all the assumptions of Perron and Qu (2006) (or those of Bai and Perron, 1998) obtained with original regressors contemporaneously uncorrelated with the errors, are satisfied. Hence, all results derived in those papers continue to hold. There is no need for a separate detailed analysis as done in Hall and Han (2006, 2007). But the question of interest is then whether using an instrumental variable approach is necessary or even useful. The next section addresses this issue and shows that simply using OLS is preferable.

4 Estimating and testing using OLS estimates

In this section, we show that it is preferable to estimate the break dates using the standard OLS method rather than an IV procedure even in the presence of endogenous regressors. The reasons for this are very simple. Assume for simplicity that the break dates are known and let $p\lim_{T\to\infty} E(X_i'u) = \phi_i$ for $i = 1, ..., m+1$, then, the probability limit of the OLS estimate $\hat{\delta}$ from (1) is given by, under A8,

$$\begin{aligned}\delta^* &= \delta^0 + p\lim_{T\to\infty}(\bar{X}_0'\bar{X}_0)^{-1}\bar{X}_0'u = \delta^0 + diag(Q_{XX}^1, ..., Q_{XX}^{m+1})^{-1}(\phi_1, ..., \phi_{m+1})' \\ &= \delta^0 + [(Q_{XX}^1)^{-1}\phi_1, ..., (Q_{XX}^{m+1})^{-1}\phi_{m+1})]'.\end{aligned}$$

Any change in the parameter δ^0 will imply an equivalent change in the limit value of the OLS estimate δ^*. Hence, one can still identify parameter changes using OLS estimates. The second feature is the well known inequality $||P_Z X|| \leq ||X||$ so that using an IV procedure leads to regressors that have less quadratic variation than when using OLS. This and the fact that the limit value of the OLS estimate changes in an equal way implies that the estimates of the break dates will be less precisely estimated using an IV procedure. The main cause for this is the fact that a change in the parameter δ^0 will cause a larger change in the conditional mean of the dependent variable in the OLS framework compared to the corresponding change in an IV regression.

To make the above arguments more precise, consider writing the DGP (1) as

$$\begin{aligned}y &= \bar{X}_0\delta^0 + P_{\bar{X}_0}u + (I - P_{\bar{X}_0})u \\ &= \bar{X}_0[\delta^0 + (\bar{X}_0'\bar{X}_0)^{-1}\bar{X}_0'u] + (I - P_{\bar{X}_0})u \\ &= \bar{X}_0\delta_T^* + u^*,\end{aligned}$$

where $u^* = (I - P_{\bar{X}_0})u$ and $\delta_T^* = [\delta^0 + (\bar{X}_0'\bar{X}_0)^{-1}\bar{X}_0'u]$ for which $\delta_T^* \to_p \delta^*$. So we can consider a regression in terms of the population value of the parameters, viz.

$$y = \bar{X}_0\delta^* + u^*.$$

It is clear that in this framework $\bar{X}_0$ is uncorrelated with u^* so that the OLS estimate, say $\hat{\delta}^*$, will be consistent for δ^*. This suggests estimating the break dates by minimizing the sum of squared residuals from the following regression

$$y = \bar{X}\delta^* + u^*.$$

Then, the estimates of the break dates are given by

$$(\hat{T}_1^*, ..., \hat{T}_m^*) = \arg\min_{T_1,...,T_m} SSR_T^*(T_1, ..., T_m), \tag{8}$$

where $SSR_T^*(T_1, ..., T_m) = (y - \bar{X}\delta^*)'(y - \bar{X}\delta^*)$. We then obtain directly from Perron and Qu (2006) that, under the same conditions as used above, the estimates of the break fractions are consistent and have the same convergence rate as in the usual OLS framework with regressors contemporaneously uncorrelated with the errors. The limit distribution of the estimates of the break dates is given in the following proposition.

Proposition 3 *Under Assumptions A1-A8 with A2-A4 and A8 stated with x_t instead of z_t, A4 and A8 stated with u_t^* instead of u_t (using the notation $\Omega_{XU^*}^i$ instead of Ω_{XU}^i), with A7 stated in term of $\Delta_i^* = \delta_{i+1}^* - \delta_i^*$ instead of $\Delta_i = \delta_{i+1}^0 - \delta_i^0$, and further assuming that $(\delta_i^{*'} z_t)^2 \pm \delta_i^{*'} z_t' u_t$ has a continuous distribution, we have, for $i = 1, ..., m$:*

$$\frac{(\Delta_i^{*\prime} Q_{XX}^i \Delta_i^*)^2}{\Delta_i^{*\prime} \Omega_{XU^*}^i \Delta_i^*} (\hat{T}_i - T_i^0) \to^d \arg\max_s V^{(i)}(s)$$

where

$$\begin{aligned} V^{(i)}(s) &= W_1^{(i)}(-s) - |s|/2 \quad \text{if } s \leq 0, \\ V^{(i)}(s) &= \sqrt{\xi_i}(\phi_{i,2}/\phi_{i,1}) W_2^{(i)}(s) - \xi_i |s|/2 \quad \text{if } s > 0, \end{aligned}$$

with $\phi_{i,1}^2 = \Delta_i^{*\prime} \Omega_{XU^*}^i \Delta_i^* / \Delta_i^{*\prime} Q_{XX}^i \Delta_i^*$, $\phi_{i,2}^2 = \Delta_i^{*\prime} \Omega_{XU^*}^{i+1} \Delta_i^* / \Delta_i^{*\prime} Q_{XX}^{i+1} \Delta_i^*$ and

$$\xi_i = \Delta_i^{*\prime} Q_{XX}^{i+1} \Delta_i^* / \Delta_i^{*\prime} Q_{XX}^i \Delta_i^*.$$

Note that the limit distribution depends on $\Delta_i^* = \delta_{i+1}^* - \delta_i^*$. But since the OLS estimates are consistent for δ^* this quantity can still be consistently estimated.

One can then compare the limit distributions of the estimates of the break dates using OLS or the IV procedure discussed in the previous section. For simplicity, consider the case where the instruments, regressors and errors have a constant distribution throughout the sample so the moment matrices do not change across regimes. Also, suppose that the errors are uncorrelated. Then, the limit distribution of the break fractions based on the IV procedure, denoted $\hat{\lambda}_{IV}$, is given by

$$\frac{\Delta_T' Q_{HH} \Delta_T}{\sigma^2}(\hat{\lambda}_{IV} - \lambda^0) \to^d \arg\max_s \left[W(s) - |s|/2\right],$$

and that of the break fractions based on the OLS procedure, denoted $\hat{\lambda}_{OLS}$, is

$$\frac{\Delta_T' Q_{XX} \Delta_T}{\sigma^{*2}}(\hat{\lambda}_{OLS} - \lambda^0) \to^d \arg\max_s \left[W(s) - |s|/2\right]$$

where $\sigma^{*2} = p\lim_{T\to\infty} T^{-1}\hat{u}^{*\prime}\hat{u}^* = p\lim_{T\to\infty} T^{-1}(y - \bar{X}\hat{\delta}^*)'(y - \bar{X}\hat{\delta}^*)$. Now, since $\|X\| \geq \|P_Z X\| = ||\hat{X}||$ and $\|u\| \geq \|(I - P_X)u\| = \|u^*\|$,

$$||Q_{HH}|| = ||p\lim_{T\to\infty} T^{-1}\hat{X}'\hat{X}|| \leq ||p\lim_{T\to\infty} T^{-1}X'X|| = ||Q_{XX}||,$$

and

$$\sigma^2 = p\lim_{T\to\infty} T^{-1}u'u \geq p\lim_{T\to\infty} T^{-1}\hat{u}^{*\prime}\hat{u}^* = \sigma^{*2},$$

so that

$$\frac{\Delta_T Q_{HH} \Delta_T}{\sigma^2} \leq \frac{\Delta_T Q_{XX} \Delta_T}{\sigma^{*2}}.$$

Therefore, the variance of the IV-based estimate $\hat{\lambda}_{IV}$ is greater than that of the OLS-based estimate $\hat{\lambda}_{OLS}$. This argument extends to more general cases. Hence, it is clearly preferable to estimate the break dates using the simple OLS-based method than using an IV procedure. As we shall see below, the loss in efficiency can be especially pronounced when the instruments are weak as is often the case in applications.

Of course, if the goal is not to get estimates of the break dates per se but estimates of the parameters within each regime, one should use an IV regression but conditioning on the estimates of the break dates obtained using the OLS-based procedure. Their limit distributions will, as usual, be the same as if the break dates were known since the estimates of the break fractions converge at a fast enough rate.

Using the same logic, we can expect the power of tests for structural change to be more powerful when based on the OLS regression rather than the IV regression. This will be investigated through simulations in the next section.

5 Simulation evidence

In this section, we provide simulation evidence to assess the advantages of estimating the break dates and constructing tests using an OLS regression compared to using an IV regression. We start in Section 5.1 with the leading case with regressors that have a homogenous distribution across segments and with the correlation between regressors and errors also invariant. In Section 5.2, we consider the case for which either can be changing, a case not covered by our theory.

5.1 Homogenous distributions across segments.

The data are generated by

$$y_t = x_t \delta_t + u_t,$$

where y_t, x_t and u_t are scalars. We consider, for simplicity, the case of a single break in the parameter δ_t occurring at mid-sample, so that

$$\delta_t = \begin{cases} c & \text{for } t \in [1, T/2], \\ -c & \text{for } t \in [T/2+1, T]. \end{cases} \tag{9}$$

First, define the following random variables:

$$\begin{aligned} \begin{pmatrix} u_t \\ v_t \end{pmatrix} &\sim i.i.d.\ N\left(0_{2\times 1}, \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix}\right), \\ \xi_t &\sim i.i.d.\ N(1,1), \\ \varepsilon_t &\sim i.i.d.\ N(0,1). \end{aligned}$$

The regressor x_t is kept the same throughout the specifications and is generated by

$$x_t = \xi_t + v_t,$$

and, again for simplicity, we consider the case of single instrument z_t generated by

$$z_t = \sqrt{\gamma}\xi_t + \sqrt{2-\gamma}\varepsilon_t,$$

with $0 \leq \gamma \leq 2$. Note that u_t and v_t are correlated so the regressors x_t are correlated with the errors u_t. The variable ξ_t is a component common to x_t and z_t and the parameter γ controls the extent of the correlation between the regressors and the instruments. The component ε_t is added in the generation of z_t to keep the variability of the instrument constant when varying γ. The R^2 in the reduced form linking x_t and z_t is a function of γ given by

$$R^2 = \sqrt{\gamma}\sigma_\xi / \sigma_x \sigma_z = \sqrt{\gamma}/2$$

with σ_ξ, σ_x and σ_z are the standard error of ξ_t, x_t and z_t, respectively. We report results for the following values $R^2 = .7$, $.5$, $.2$ and $.001$, and we consider three values for the size of the change in the parameter, $c = .25$, $.5$ and 1.0. In all cases, the sample size is $T = 100$ and the number of the replication is $1,000$. The results are presented in Figure 1 for the cumulative distribution functions.

The OLS-based estimates are always more precise than the instrumental variable-based estimates, even when the R^2 is as high as 0.7. The extent to which the IV-based estimates perform badly increases as R^2 decreases, and even for a value as high as 0.5, which is larger than what can be obtained in most applications, the IV-based estimates show much higher variability than the OLS-based estimates. In the case of very weak instruments, the IV-based estimates are totally uninformative. Also, when the magnitude of change increases, the precision of the OLS-based estimate increases noticeably. The increase in the precision of the IV-based estimate is, however, marginal as c increases. Hence, even with large breaks, the OLS-based estimates are far superior. The simulations show that it is indeed highly preferable to estimate the break dates using a standard OLS framework even when the regressors are contemporaneously correlated with the errors.

We also simulated the power of the sup-Wald test for a single break of Quandt (1958, 1960) as developed by Andrews (1993). Given our setup, it is given by

$$\sup F = \sup_{[\epsilon T] < T_b < [(1-\epsilon)T]} \frac{SSR_T^r - SSR(T_b)}{SSR(T_b)/(T-1)},$$

where SSR_T^r denotes the restricted sum of squared residuals obtained from a regression assuming no break, and $SSR(T_b)$ denotes the sum of squared residuals in a regression allowing for a break at date T_b. The trimming parameter is set to $\epsilon = 0.15$ and we report results for tests with a 5% nominal size. We also considered the UDmax test of Bai and Perron (1998), for which there is no need to pre-specify the number of breaks. The results were qualitatively similar and, hence, are not reported. To construct the power functions, we varied c between 0 and 1.

The results are presented in Figure 2. They show that the OLS-based test has the highest power. The power of the IV-based test is noticeably inferior and becomes uninformative as the instrument gets more weakly correlated with the regressors. Hence, testing should also be performed using the usual OLS-based methods.

5.2 Non-homogenous segments

We now consider the case for which the distribution of the regressors and the correlation between the regressors and errors can vary across segments. These cases are important since then the change in the pseudo limit values $\Delta_i^* = \delta_{i+1}^* - \delta_i^*$ need not equal the change in the true population values $\Delta_i = \delta_{i+1}^0 - \delta_i^0$ and neither is necessarily greater than the other. Given the nature of the limit distributions stated in Propositions 2 and 3, it is difficult to provide theoretical results; hence, we resort to simulations. The data generating process is similar to the one used in the previous section except that we allow some key parameters to change across regimes. More specifically, we now have

$$\begin{aligned} x_t &= \xi_t + v_t, \\ v_t &\sim i.i.d.\ N(0, \sigma_{tv}^2), \\ \xi_t &\sim i.i.d.\ N(\mu_t, 1), \end{aligned}$$

and

$$E(x_t u_t) = \phi_t.$$

We consider three cases: 1) a change in Q_{XX} induced by a change in the variance of the regressor such that

$$\sigma_{tv} = \begin{cases} \sigma_{1v} & \text{for } t \in [1, T/2], \\ \sigma_{2v} & \text{for } t \in [T/2+1, T]; \end{cases}$$

2) a change in the correlation between the regressor and the errors such that

$$\phi_t = \begin{cases} \phi_1 & \text{for } t \in [1, T/2], \\ \phi_2 & \text{for } t \in [T/2+1, T]; \end{cases}$$

and 3) a change in Q_{XX} induced by a change in the mean of the regressor such that

$$\mu_t = \begin{cases} \mu_1 & \text{for } t \in [1, T/2], \\ \mu_2 & \text{for } t \in [T/2+1, T]. \end{cases}$$

The other specifications are the same. For cases (1) and (3), a change in the distribution of the regressor across segments induces a change in the correlation between the regressor and the instrument if γ is fixed. In order to keep this correlation constant across segments, we change γ at date $[T/2]$ so that $R^2 = \sqrt{\gamma}\sigma_\xi/\sigma_x\sigma_z$ is held fixed throughout the sample. Again, in all cases, the sample size is $T = 100$ and the number of the replication is $1,000$ and we report results for $R^2 = .7$, $.5$, $.2$ and $.001$. The results reported are the cumulative distribution functions of the IV and OLS-based estimates. In all three cases, we consider changes such that the value of the relevant parameter is smaller or greater in the first regime compared to the second, since the distributions are not symmetric for equal changes occurring in the two segments. Also, the change in the parameter δ is given by (9) with $c = 0.25$.

Consider first case (1) with a smaller variance of the regressor in the first regime so that $\sigma_{2v} = 1$ and $\sigma_{1v} = a$ with $a = 0.8$, 0.5 and 0.25. The results are presented in Figure 3. Here Δ_1^* is greater in absolute value than Δ_1, so on that basis we can expect an increase in the efficiency of the estimate of the break date. When comparing with the base case in Figure 2, this is indeed the case and more so as a decreases. The results in Figure 4 pertain to the case of a smaller variance in the second regime so that Δ_1^* is smaller in absolute value than Δ_1. The specifications used are $\sigma_{1v} = 1$ and $\sigma_{2v} = b$ with $b = 0.8$, 0.5 and 0.25. One can see a reduction in the efficiency of the OLS-based estimate. It nevertheless remains more efficient than the IV-based estimate for all cases considered and again more so as the extent of the correlation between the regressor and instrument decreases.

Consider now a change in the correlation between the errors and the regressors across regimes. The first case is one for which the correlation is smaller in the first regime given by $\phi_2 = 1$ and $\phi_1 = a = 0.25$, 0.1 and 0.01, for which Δ_1^* is smaller in absolute value than Δ_1. As depicted in Figure 5, this translates into a reduction in the precision of the OLS-based estimate of the break date. Unlike case (1), we have some instances for which the IV-based estimate is more precise than the OLS-based one. This, however, requires a very high correlation between the regressor and instrument (e.g., an R^2 value of 0.7) and a very large change in the correlation between the regressor and the errors. The results for the case with a correlation higher in the first regime are presented in Figure 6 for the specifications $\phi_1 = 1$ and $\phi_2 = b = 0.25$, 0.1 and 0.01. Here Δ_1^* is greater in absolute value than Δ_1 and the results show that in all cases the OLS-based estimate is more precise than the IV-based one.

The last case considered is one for which the mean of the regressor changes across regimes. Figure 7 presents the results when the mean is higher in the first regime, specified by $\mu_2 = 1$ and $\mu_1 = a = 1.5$, 2.0 and 5.0. In this case, Δ_1^* is smaller in absolute value than Δ_1. The OLS-based estimate nevertheless remains more efficient unless the change in mean is very large ($a = 5$) and the correlation between the regressor and instrument is high (e.g., an R^2 value above 0.5). When the mean of the regressor is higher in the second regime Δ_1^* is greater in absolute value than Δ_1. The results presented in Figure 8 for the specifications $\mu_1 = 1$ and $\mu_2 = b = 1.5$, 2.0 and 5.0 indeed show that the OLS-based estimate is in such cases more efficient than the IV-based one and more so as the change in mean increases.

Overall, the simulation presented in this section showed that, in general, the OLS-based estimate is more precise than the IV-based one even when the distribution of the regressors or the correlation between the regressors and the errors are allowed to change across regimes.

There are cases for which this is not true but they occur for unrealistically high values of the correlation between the regressors and instruments and very large changes in either the correlation between the regressors and errors or the means of the regressors. Hence, for all practical purposes, it is more beneficial to use an OLS-based method.

6 Empirical example

To illustrate the relevance of our theoretical results, we consider the stability of the New Keynesian hybrid Phillips curve as put forward by Gali and Gertler (1999). The basic specification is given by

$$\pi_t = \mu + \kappa x_t + \beta E_t \pi_{t+1} + \gamma \pi_{t-1} \tag{10}$$

where π_t is the inflation rate and E_t is the expectation operator conditional on information available at time t. The variable x_t is a real determinant of inflation usually taken as a measure of real economic activity such as the output gap, though Gali and Gertler (1999) argue that using a measure of real marginal cost is preferable. The key ingredient is the component $E_t\pi_{t+1}$ which makes the Phillips-curve forward looking. The one-period lag of inflation is usually introduced on the ground that it improves the fit of the model but can be rationalized by supposing that a proportion of firms use backward-looking rules to set prices. In this specification, one expects β and γ to be positive and that β is substantially larger than γ so that the forward looking component dominates. Also, κ is expected to be positive so that an increase in real activity is associated with an increase in inflation. As argued by Gali and Gertler (1999), the estimate of κ is usually negative when using the output gap but is positive, though quite small, when using real marginal cost.

We follow the approach taken by Gali and Gertler (1999) and use the same specifications and data. The data was obtained from Andre Kurmann's website (and used in Kurmann, 2007). It is for the U.S.A. and quarterly for the period 1960:1-1997:4 as in Gali and Gertler (1999). The inflation rate π_t is the quarterly change in the GDP deflator and x_t is either 1) labor income share as a proxy of real unit labor cost (nonfarm business unit labor cost deflated by the GDP deflator), or 2) detrended real GDP assuming a quadratic linear trend as a measure of the output gap. We estimated the parameters in (10) by IV using, as in Gali and Gertler (1999), the following instruments: four lags of inflation, labor income share, the output gap, the interest rate spread (10 years US treasury bill rate minus the 3 years bill rate), wage inflation (quarterly change in the nonfarm business nominal wage rate) and the commodity price inflation (quarterly change in the spot market price index).

Gali and Gertler (1999) estimated the model using a non-linear GMM procedure given that their model imply restrictions across the coefficients which are functions of some basic parameters. The only difference from their approach is that we shall using a linear IV method of estimation. The results are, however, in line with those of Gali and Gertler (1999) for the full sample specification and doing so does not affect our conclusions. The parameter estimates for the full sample are presented in Table 1 for both cases in which the x_t is specified as the real marginal labor cost or as the output gap. The results are in close agreement with those of Gali and Gertler (1999). The coefficient β on future expected inflation is large and significant and more than twice as large as the coefficient γ on lagged inflation, indicating that the forward-looking behavior is dominant. The estimate of κ is positive when using the real marginal cost of labor and negative when using the output gap, though the values are small and insignificant. It is finally noteworthy to point out that the instruments are quite strongly correlated with future inflation, the R^2 of the first-stage regression being 0.812.

We now turn to the issue of the stability of this hybrid New-Keynesian Phillips curve. When using IV-based methods the sup-Wald test for one break is not significant in either case (it is 11.6 with the real marginal cost and 8.6 with the output gap). The estimate of the break date is at 1974:1 in both cases with a very large confidence interval (1963:3 to 1984:3 with real marginal cost and 1962:3 to 1985:3 with the output gap). Hence, IV-based methods do not provide any evidence of structural instability. Things are very different with OLS-based methods. In both cases, the sup-Wald test for one break is highly significant with a p-value less than 1% (28.2 with real marginal cost and 21.1 with the output gap). The estimate of the break date is at 1991:1 in both cases with a very tight confidence interval (1990:4 to 1991:2 in both cases). Hence, OLS-based methods provide strong evidence of a structural change in 1991:1. The higher discriminatory power of OLS-based methods over IV-based ones occurs despite the fact that instruments are highly correlated with future inflation.

It remains to be seen whether this change in 1991:1 is economically significant. To that effect we estimated the parameters of the model by IV using the same instruments allowing for a change in all parameters. The estimates are presented in Table 2. They are indeed very interesting. The estimates for the period 1960:1-1991:1 are close to the full sample estimates reported earlier and support Gali and Gertler's (1999) conclusion. On the other hand, the estimates for the period 1991:2-1997:4 are all very small and insignificantly different from zero in both cases. Hence, both of these hybrid New Keynesian Phillips curve specifications have lost any explanatory power for inflation.

This forecast breakdown of the New Keynesian Phillips curve for inflation is interesting and can be traced back to the change in the behavior of inflation. As argued by Stock and Watson (2007) even though inflation has become more stable after the mid-80's it also has become more difficult to forecast (see also Rossi and Sekhposyan, 2008). The evidence presented here is consistent with theirs, though stronger perhaps because of the different break date identified.

7 Conclusions

In this paper, we considered the problem of multiple structural changes in a single equation framework with regressors that are endogenous, i.e., correlated with the errors. We first analyzed the properties of the estimates of the break dates using an instrumental variable procedure. We provided a very simple proof for the consistency, rate of convergence and limit distributions of the estimates of the break fractions, as well as the limit distributions of the tests for multiple structural changes. We did so by showing that using generated regressors, the projection of the regressors of the space spanned by the instruments, to account for potential endogeneity implies that all the assumptions of Perron and Qu (2006) (or those of Bai and Perron, 1998) obtained with original regressors contemporaneously uncorrelated with the errors, are satisfied. Hence, all results derived in those papers continue to hold.

More importantly, we showed that even in the presence of endogenous regressors, it is still preferable to simply estimate the break dates and test for structural change using the usual ordinary least-squares framework. The reasons are simple. First, changes in the true parameters of the model imply a corresponding change in the probability limits of the OLS parameter estimates. Second, one can reformulate the model with those probability limits as the basic parameters in a way that the regressors and errors are contemporaneously uncorrelated. We are then simply back to the framework of Bai and Perron (1998) or Perron and Qu (2006) and we can use their results directly to obtain the relevant limit distributions. Since the OLS framework involves the original regressors, while the IV framework involves as regressors the projection of these original regressors on the space spanned by the instruments, this implies that the generated regressors in the IV procedure have less quadratic variation than the original regressors. Accordingly, using OLS not only delivers consistent estimates of the break fractions and tests with the usual limit distributions, it also improves on the efficiency of the estimates and the power of the tests.

Appendix

Proof of Lemma 1: To verify assumption 1, note that

$$\begin{aligned}
\frac{1}{l_i}\sum_{t=T_i^0+1}^{T_i^0+[l_i v]} \hat{x}_t\hat{x}_t' &= \frac{1}{l_i}\sum_{t=T_i^0+1}^{T_i^0+[l_i v]} (P_Z X)_t'(P_Z X)_t \\
&= (T^{-1}X'Z)(T^{-1}Z'Z)^{-1}(\frac{1}{l_i}\sum_{t=T_i^0+1}^{T_i^0+[l_i v]} z_t z_t')(T^{-1}Z'Z)^{-1}(T^{-1}Z'X) \\
&\rightarrow {}_p Q_{ZX}' {Q_{ZZ}}^{-1} Q_{ZZ}^i(v) Q_{ZZ}^{-1} Q_{ZX} \equiv Q_{HH}^i(v),
\end{aligned}$$

by A1. Since Q_{ZZ} is positive definite, so is $Q_{HH}^i(v)$, and it is non-random given that all elements are non-random. To verify assumption 2, let

$$M_T = (1/l)\sum_{t=T_i^0+1}^{T_i^0+l} \hat{x}_t\hat{x}_t' = (T^{-1}X'Z)'(T^{-1}Z'Z)^{-1}A_T^i(T^{-1}Z'Z)^{-1}(T^{-1}Z'X),$$

where $A_T^i = (1/l)\sum_{t=T_i^0+1}^{T_i^0+l} z_t z_t'$ and let $\Phi_T = (T^{-1}X'Z)'(T^{-1}Z'Z)^{-1}(T^{-1}Z'Z)^{-1}(T^{-1}Z'X)$. Also, let d be an arbitrary non-zero a $p\times 1$ vector and define $\xi_T = d'(T^{-1}X'Z)'(T^{-1}Z'Z)^{-1}$. By A3, Φ_T has rank p and, hence, is positive definite. Denote the minimum eigenvalue of Φ_T by $\omega_1(>0)$. Since A_T^i is symmetric, we have $\min_{\xi_T}(\xi_T A_T^i \xi_T'/\xi_T\xi_T') = \tau_1$ where τ_1 is the minimum eigenvalues of A_T^i which is bounded away from zero by A2. Now

$$\begin{aligned}
\min_d\left(\frac{d'M_T d}{d'd}\right) &= \min_d\left[\left(\frac{\xi_T A_T^i\xi_T'}{d'\Phi_T' d}\right)\left(\frac{d'\Phi_T d}{d'd}\right)\right] = \min_d\left[\left(\frac{\xi_T A_T^i\xi_T'}{\xi_T\xi_T'}\right)\left(\frac{d'\Phi_T d}{d'd}\right)\right] \\
&\geq \min_d\left(\frac{\xi_T A_T^i\xi_T'}{\xi_T\xi_T'}\right)\min_d\left(\frac{d'\Phi_T d}{d'd}\right) = \tau_1\omega_1.
\end{aligned}$$

Since τ_1 is bounded away from zero and $\omega_1 > 0$, $\tau_1\omega_1$ is bounded away from zero, which implies that the minimum eigenvalues of M_T are bounded away from zero. To verify assumption A3, note that

$$\sum_{t=k}^{l} \hat{x}_t\hat{x}_t' = (T^{-1}X'Z)'(T^{-1}Z'Z)^{-1}(\sum_{t=k}^{l} z_t z_t')(T^{-1}Z'Z)^{-1}(T^{-1}Z'X)$$

Since $(T^{-1}X'Z)'$, $(T^{-1}Z'Z)^{-1}$ and $(\sum_{t=k}^{l} z_t z_t')$ have full column rank by A3,

$$rank[(T^{-1}X'Z)'(T^{-1}Z'Z)^{-1}(\sum_{t=k}^{l} z_t z_t')(T^{-1}Z'Z)^{-1}(T^{-1}Z'X)] = rank\left[(T^{-1}X'Z)'\right] = p.$$

We now verify assumption A4. For any finite T, given a sample of $\{x_t\}$ and $\{z_t\}$, let $\gamma_T = (\sum_{t=1}^{T} z_t z_t')^{-1}(\sum_{t=1}^{T} z_t x_t')$, a $q\times p$ matrix. Then $\hat{x}_t u_t = \gamma_T' z_t u_t = \gamma z_t u_t + o_p(1)$, uniformly in t, where $\gamma = Q_{ZZ}^{-1}Q_{ZX}$. Now $\gamma z_t u_t$ is a linear combination of the L^r-mixingale sequence $\{z_t u_t\}$ (A4). Therefore all we need to show is that linear combinations of any finite numbers of L^r-mixingale processes are also L^r-mixingale processes. To prove this it is sufficient to consider two sequences $\{\eta_t^1\}$ and $\{\eta_t^2\}$ that satisfy A4 and show that linear combinations of these, say $a\eta_i^1 + b\eta_i^2$, satisfy properties (a), (b) and (e) in A4. For property (a), we have,

$$\begin{aligned}\left\|E(a\eta_i^1 + b\eta_i^2|\mathcal{F}_{i-j})\right\|_r &= \left\|E(a\eta_i^1|\mathcal{F}_{i-j}) + E(b\eta_i^2|\mathcal{F}_{i-j})\right\|_r \\ &\leq \left\|E(a\eta_i^1|\mathcal{F}_{i-j})\right\|_r + \left\|E(a\eta_i^2|\mathcal{F}_{i-j})\right\|_r \leq |a|_r c_i^1\psi_i^1 + |b|_r c_i^2\psi_i^2,\end{aligned}$$

while for property (b),

$$\begin{aligned}\left\|a\eta_i^1 + b\eta_i^2 - E(a\eta_i^1 + b\eta_i^2|\mathcal{F}_{i-j})\right\|_r &= \left\|a\eta_i^1 + b\eta_i^2 - E(a\eta_i^1|\mathcal{F}_{i-j}) - E(b\eta_i^2|\mathcal{F}_{i-j})\right\|_r \\ &\leq \left\|a\eta_i^1 - E(a\eta_i^1|\mathcal{F}_{i-j})\right\|_r + \left\|b\eta_i^2 - E(b\eta_i^2|\mathcal{F}_{i-j})\right\|_r \\ &\leq |a|c_i^1\psi_i^1 + |a|c_i^2\psi_i^2,\end{aligned}$$

and, for property (e),

$$\left\|a\eta_i^1 + b\eta_i^2\right\|_{2r} \leq |a| \left\|\eta_i^1\right\|_{2r} + |b| \left\|\eta_i^2\right\|_{2r} < (|a| + |b|)M < \infty.$$

Since $\gamma_T' z_t u_t$ differs from $\gamma z_t u_t$ by an $o_p(1)$ term uniformly in t, it has the same properties.

Proof of Lemma 2. For part 1, we have

$$\begin{aligned}&(\Delta T_i^0)^{-1} \sum_{t=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} \hat{x}_t\hat{x}_t' \\ =\;& (T^{-1}X'Z)(T^{-1}Z'Z)^{-1}[(\Delta T_i^0)^{-1} \sum_{t=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} z_t z_t'](T^{-1}Z'Z)^{-1}(T^{-1}X'Z) \\ \rightarrow\;& {}_p sQ_{ZX}'Q_{ZZ}^{-1}Q_{ZZ}^iQ_{ZZ}^{-1}Q_{ZX} \equiv sQ_{HH}^i.\end{aligned}$$

For part 2,

$$\begin{aligned}&(\Delta T_i^0)^{-1} \sum_{t=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} \sum_{r=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} (\hat{x}_r\hat{x}_t'u_ru_t) \\ =\;& (T^{-1}X'Z)(T^{-1}Z'Z)^{-1}[(\Delta T_i^0)^{-1} \sum_{t=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} \sum_{r=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} (z_rz_t'u_ru_t)]\; (T^{-1}Z'Z)^{-1}(T^{-1}Z'X) \\ \rightarrow\;& {}_p sQ_{ZX}'Q_{ZZ}^{-1}\Omega_Z^iQ_{ZZ}^{-1}Q_{ZX} \equiv s\Omega_{HU}^i.\end{aligned}$$

For part 3,

$$\begin{aligned}(\Delta T_i^0)^{-1/2} \sum_{t=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} \hat{x}_tu_t &= (T^{-1}X_T'Z_T)(T^{-1}Z_T'Z_T)^{-1}[(\Delta T_i^0)^{-1/2} \sum_{t=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} z_tu_t] \\ &\Rightarrow Q_{ZX}'Q_{ZZ}^{-1}B_{ZU}^i(s) \equiv B_{HU}^i(s)\end{aligned}$$

Proof of Lemma 3: Part (a) follows from the proof of Lemma 1(a) with trivial modifications using the fact that there is a single segment involved. Part (b) follows from the fact that with serially uncorrelated errors and a single regime, Ω_{HU}^i in lemma 2 (part 3) reduces to $\sigma^2 Q_{HH}$.

References

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

Bai, J. (2000): "Vector autoregressive models with structural changes in regression coefficients and in variance-covariance matrices," *Annals of Economics and Finance*, 1, 303-339.

Bai, J., Lumsdaine, R.L., and Stock, J.H. (1998): "Testing for and dating breaks in multivariate time series," *Review of Economic Studies*, 65, 395-432.

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

Bai, J., and Perron, P. (2003): "Computation and analysis of multiple structural change models," *Journal of Applied Econometrics*, 18, 1-22.

Gali, J., and Gertler, M. (1999): "Inflation dynamics: a structural econometric analysis," *Journal of Monetary Economics*, 44, 195-222.

Hall, A., and Han, S. (2006): "Inference regarding multiple structural changes in linear model estimated via two stage least squares," Unpublished manuscript, Department of Economics, North Carolina State University.

Hall, A., and Han, S. (2007): "Asymptotic distribution theory for break point estimators in models estimated via 2SLS," Unpublished manuscript, Department of Economics, North Carolina State University.

Hawkins, D.M. (1976): "Point estimation of the parameters of piecewise regression models," *Applied Statistics*, 25, 51-57.

Kejriwal, M, and Perron, P. (2006a): "Testing for multiple structural changes in cointegrated regression models," Unpublished Manuscript, Department of Economics, Boston University.

Kejriwal, M, and Perron, P. (2006b): "The limit distribution of the estimates in cointegrated regression models with multiple structural changes," Unpublished Manuscript, Department of Economics, Boston University.

Kurmann, A. (2007): "VAR-based estimation of Euler equations with an application to New Keynesian pricing," *Journal of Economic Dynamics and Control*, 31, 767-796.

Perron, P. (2006): " Dealing with structural breaks," in *Palgrave Handbook of Econometrics*, Vol. 1: Econometric Theory, K. Patterson and T.C. Mills (eds.), Palgrave Macmillan, 278-352.

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

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

Quandt, R.E. (1958): "The estimation of the parameters of a linear regression system obeying two separate regimes," *Journal of the American Statistical Association,* 53, 873-880.

Quandt, R.E. (1960): "Tests of the hypothesis that a linear regression system obeys two separate regimes," *Journal of the American Statistical Association*, 55, 324-330.

Rossi, B., and Sekhposyan T. (2008): "Has models' forecasting performance for US output growth and inflation changed over time, and when?" Economic Research Initiatives at Duke Working Paper No. 12, Duke University.

Stock, J.H., and M.W. Watson (2007): "Why has U.S. inflation become harder to forecast?" *Journal of Money, Credit and Banking*, 39, 3-33.

Zhou, J., amd Perron P. (2007): "Testing jointly for structural changes in the error variance and coefficients of a linear regression model," Unpublished Manuscript, Department of Economics, Boston University.

Figure 1 : Cumulative distribution functions of the estimates of the break date.

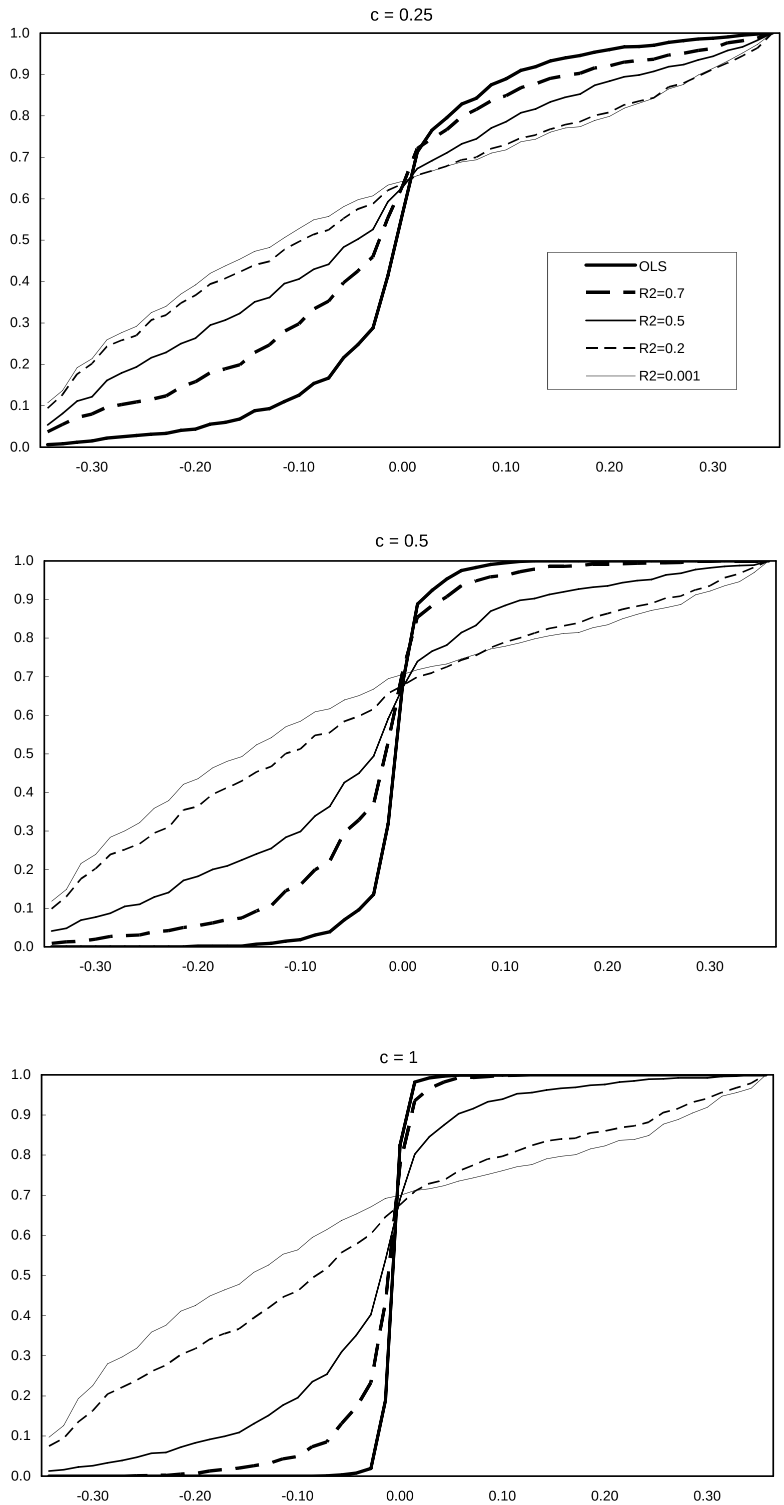

Figure 2 : Power functions of the sup-Wald structural change test.

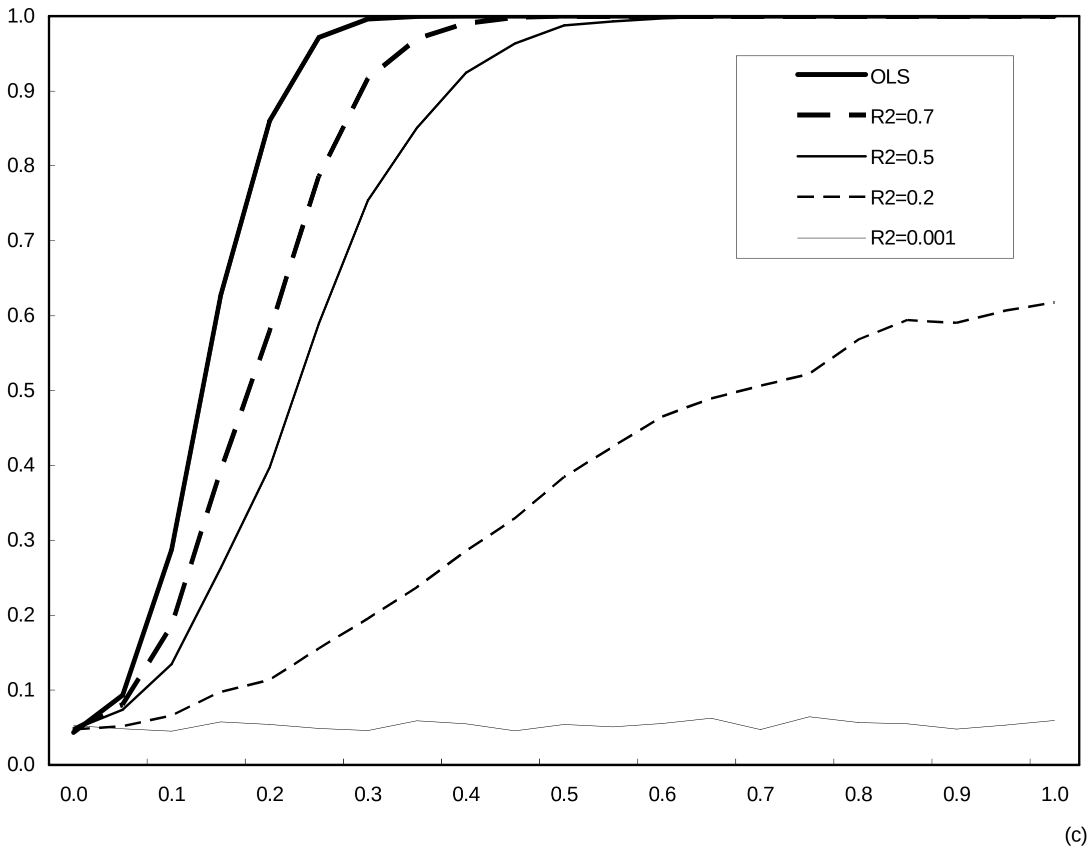

Figure 3 : Change in the variance of the regressor; $\sigma_{2v} = 1$ and $\sigma_{1v} = a$.

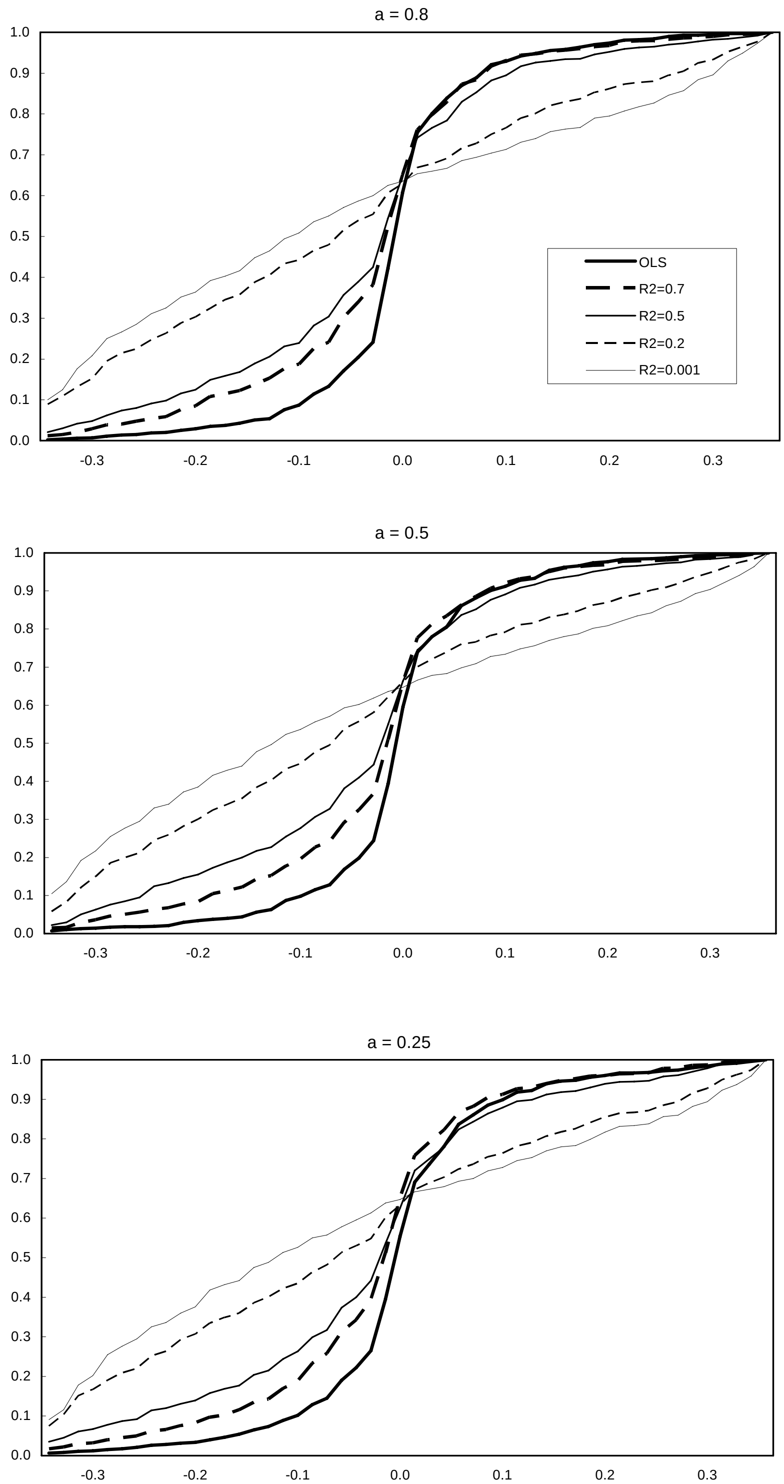

Figure 4. Change in the variance of the regressor; $\sigma_{1v} = 1$ and $\sigma_{2v} = b$.

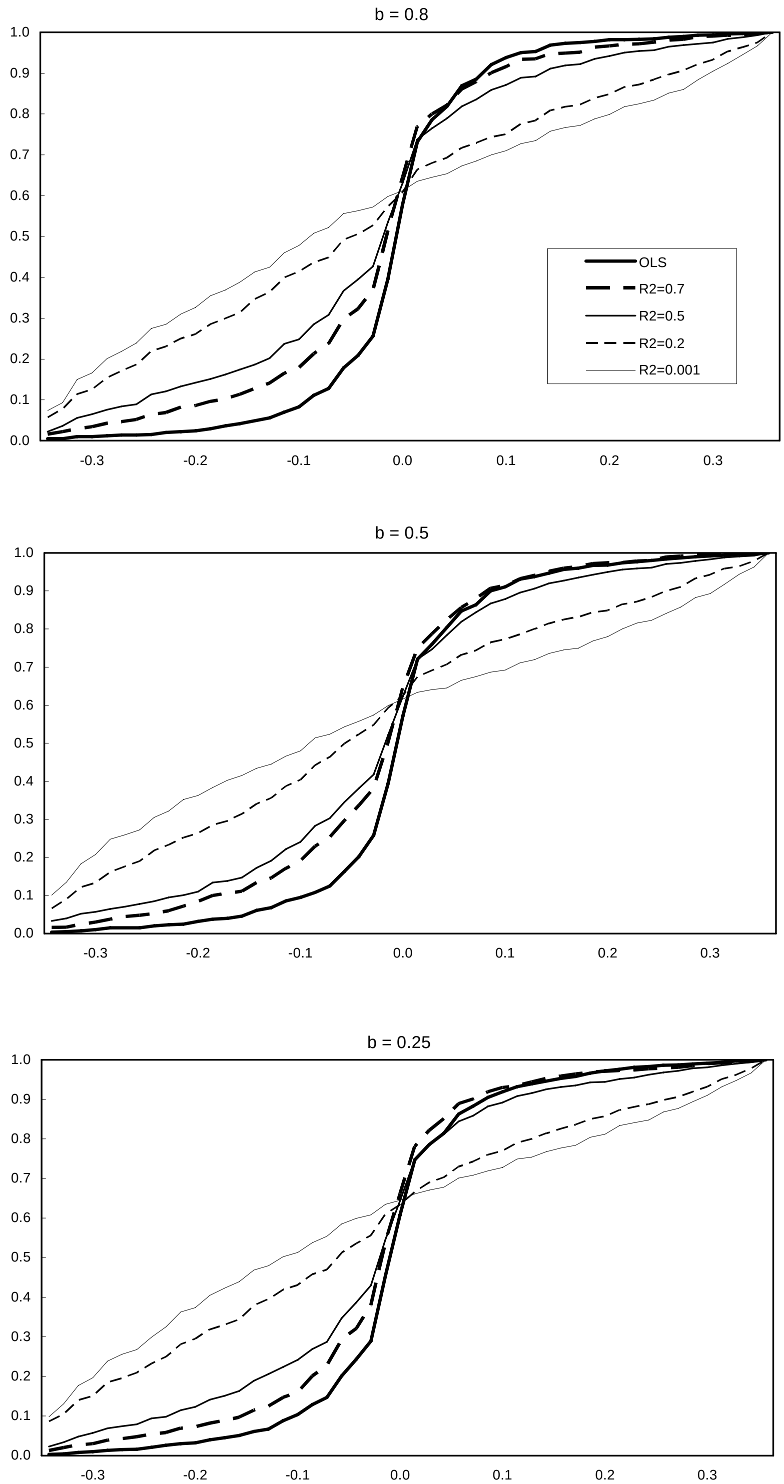

Figure 5 : Change in the correlation between the regressor and errors; $\phi_2 = 1$ and $\phi_1 = a$.

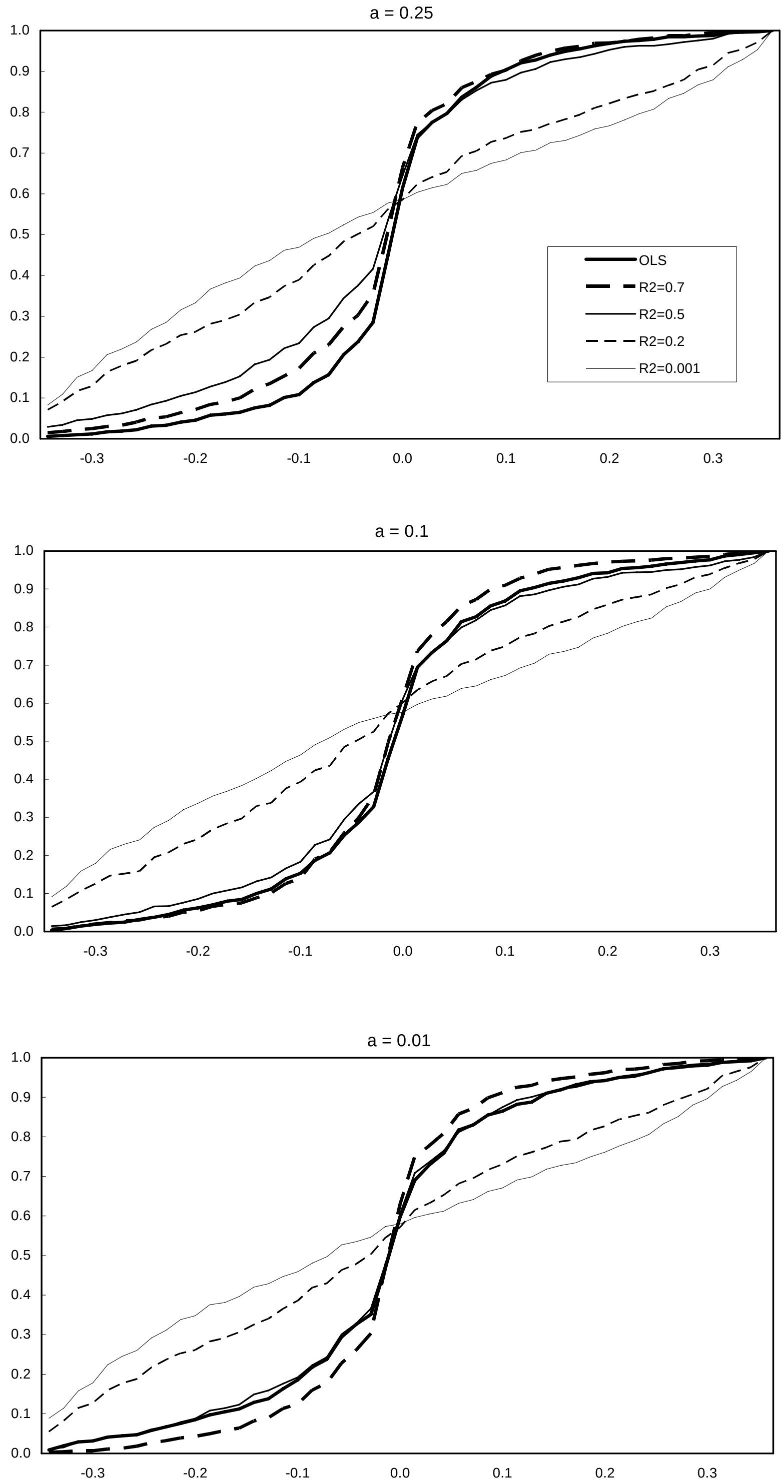

Figure 6 : Change in the correlation between the regressor and errors; $\phi_1 = 1$ and $\phi_2 = b$.

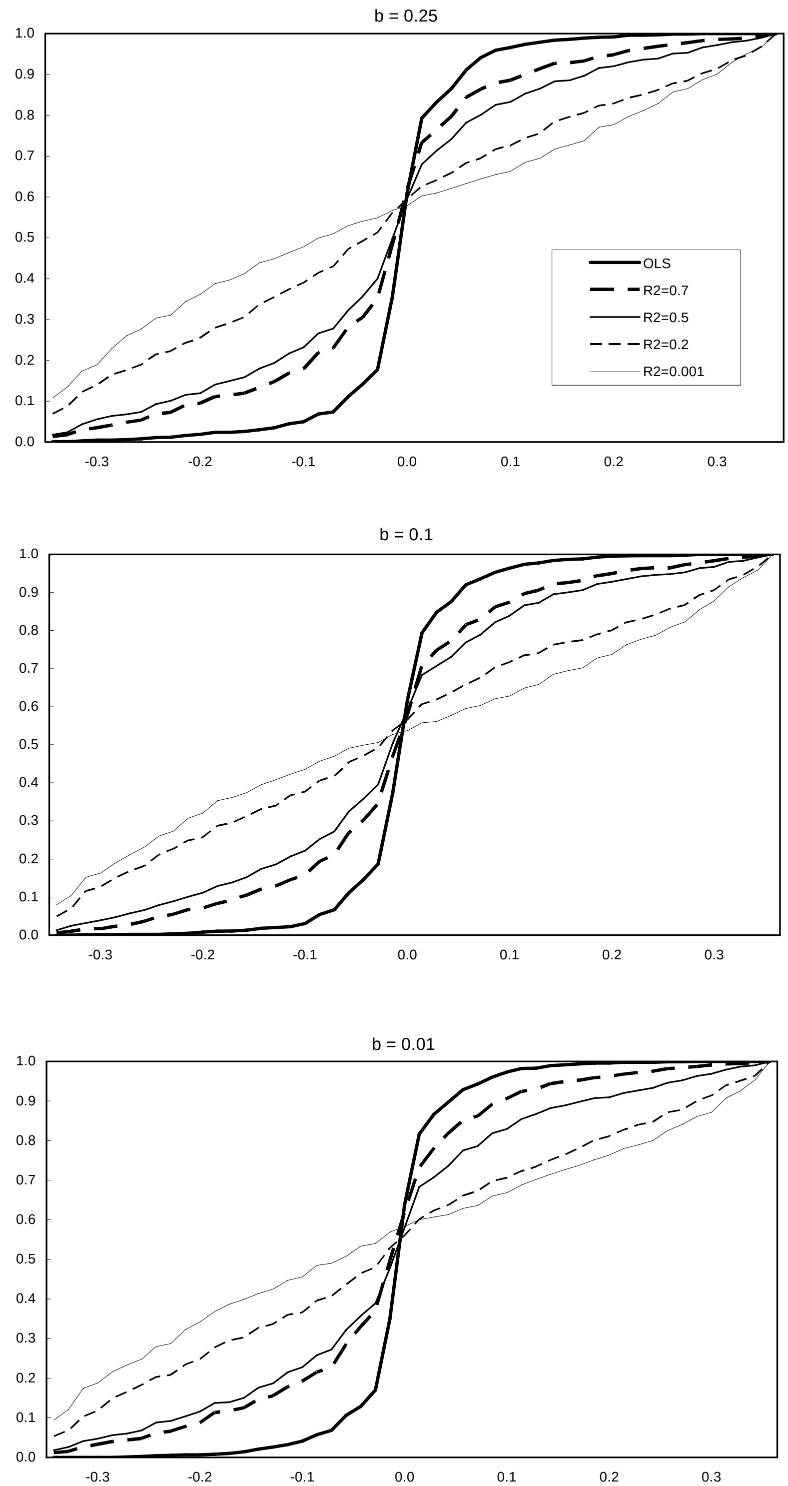

Figure 7 : Change in the mean of the regressor; $\mu_2 = 1$ and $\mu_1 = a$.

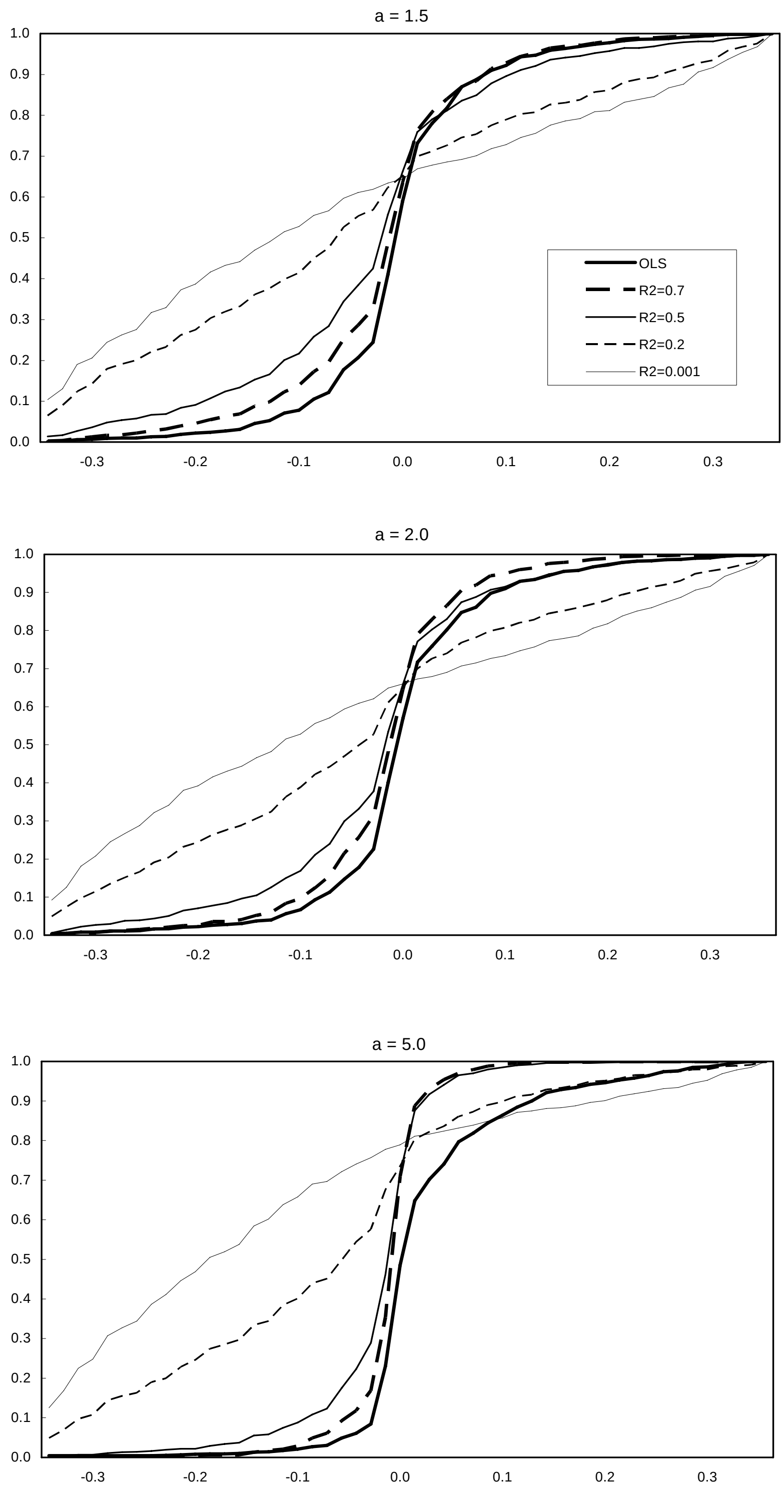

Figure 8 : Change in the mean of the regressor; $\mu_1 = 1$ and $\mu_2 = b$.

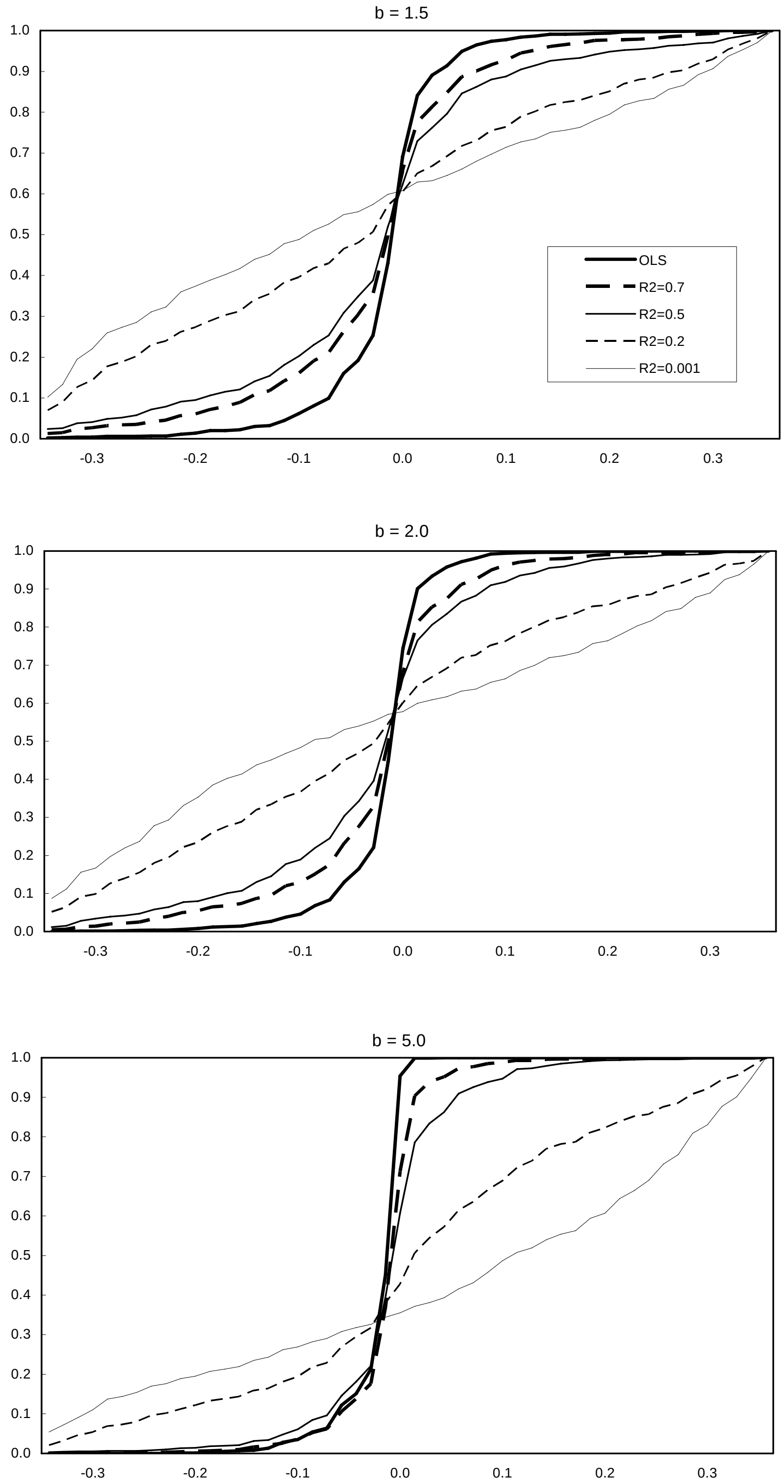

Table 1. Full sample estimates and tests for structural change for the hyblid Philips curve

(a) x_t is labor income share

	SupF	break date	95% C.I.		parameter estimates			
					const	π_{t-1}	x_t	$E(\pi_t)$
OLS	28.2***	1991:1	1990:4	1991:2	0.001	0.480	−0.002	0.479
					(0.001)	(0.060)	(0.007)	(0.063)
2SLS	11.6	1974:1	1963:3	1984:3	0.000	0.312	0.004	0.684
					(0.001)	(0.081)	(0.008)	(0.080)

(b) x_t is the GDP gap

	SupF	break date	95% C.I.		parameter estimates			
					const	π_{t-1}	x_t	$E(\pi_t)$
OLS	21.1***	1991:1	1990:4	1991:2	0.000	0.482	0.001	0.478
					(0.000)	(0.062)	(0.007)	(0.066)
2SLS	8.6	1974:1	1962:3	1985:3	0.000	0.293	−0.006	0.706
					(0.000)	(0.088)	(0.008)	(0.092)

Table 2. IV sub-sample estimates for the hybrid Phillips curve

(a) x_t is labor income share

	γ	α	x_t	$E(\pi_t)$
1960:1-1991:1	0.000	0.303	−0.001	0.679
	(0.002)	(0.086)	(0.009)	(0.084)
1991:2-1997:4	−0.002	−0.015	0.057	−0.062
	(0.006)	(0.150)	(0.041)	(0.233)

(b) x_t is the GDP gap

	γ	α	x_t	$E(\pi_t)$
1960:1-1991:1	0.000	0.286	−0.005	0.701
	(0.000)	(0.093)	(0.008)	(0.096)
1991:2-1997:4	0.006	−0.083	−0.046	−0.129
	(0.001)	(0.147)	(0.021)	(0.215)

Notes:

1. The Sup-F test uses a 15% trimming. In parentheses are the heteroskedasticity robust estimate of the standard errors. *, ** and *** indicates significance at the 10%, 5% and 1% level, respectively.

2. The confidence intervals of the estimates of the break date are based on the two-sided 95% nominal level symmetric method described in Bai (1997).