Universidad de la República
Facultad de Ciencias Sociales
DEPARTAMENTO DE ECONOMIA

Documentos de trabajo

On the power of tests for superexogeneity and structural invariance

Z. Psaradakis, M. Solá

Documento No. 09/93
Diciembre, 1993

On the power of tests for superexogeneity and structural invariance

Zacharias Psaradakis
Martín Solá

Documento No. 9/93 Septiembre, 1993

Z. Psaradakis is a researcher at the Department of Economics of Bristol University and M. Solá is a researcher at the Department of Economics of Birbeck College and at the Centre for Economic Forecasting of the London Business School.

Abstract

This paper examines the finite-sample power of tests of structural invariance and superexogeneity hypotheses in econometric models with contemporaneous conditioning variables. We consider both direct parametric tests of superexogeneity, as well as indirect procedures based on temporal stability tests for the parameters of interest. Our Monte Carlo analysis reveals that both types of tests may lack power in interesting classes of models. An empirical illustration investigates the superexogeneity of the short-term interest rate in a dynamic specification for the U.S. term structure.

JEL Classification Numbers: C15, C22, C52, E43
Keywords: Parameter Constancy; Structural Invariance; Lucas Critique; Monte Carlo Simulation; Power; Superexogeneity; Term Structure of Interest Rates

1. INTRODUCTION

The dangers of using models with parameterizations which are not invariant relative to interesting classes of changes in the economic structure have been long acknowledged in econometrics [see Aldrich (1989) and Morgan (1990) for useful historical perspectives]. In environments of structural changes and regime shifts, such models provide a poor basis for analysis, as the conditional forecasts of the variables modelled are affected by changes (or "interventions") in the laws of motion of the contemporaneous conditioning variables. Conversely, econometric equations that remain constant in response to changes embodied in the conditioning process are serious candidates for useful descriptions of the invariants of economic behaviour, and are capable of sustaining meaningful forecast and policy simulation exercises. As argued in Engle, Hendry, and Richard (1983), such invariance characteristics are indeed guaranteed for a given set of interventions when the conditioning variables are superexogenous relative to the interventions under consideration. This also implies that the conditional specification is not interpretable as a reparameterized forward-looking model involving model-based expectations, and the Lucas critique [Lucas(1976)] does not apply for the relevant class of interventions [cf. Hendry(1988), Ericsson and Hendry(1989), and Favero and Hendry(1992)]."

In spite of the obvious importance of superexogeneity in econometric modelling, little attention has been devoted to the development of formal procedures for empirically assessing the validity of superexogeneity and invariance claims. The latter have been typically tested only indirectly via tests for predictive failure or for constancy of a set of parameters of interest. Recently, however, Engle and Hendry (1993) showed that it is also possible to determine (testable) parametric conditions under which conditioning variables in linear regression models are superexogenous with respect to a specified class of interventions (e.g., those that have occurred in a given sample period). The objective of this paper is to investigate, using Monte Carlo methodology, the finite-sample performance of such direct parametric tests, and in particular to evaluate their power against false superexogeneity assertions. The detectability of parameter non-constancy induced by shifts in the generating process of the conditioning variables is also examined, by various tests for parameter instability and structural change.

The paper is organized as follows. Section 2 describes procedures for testing invariance and superexogeneity hypotheses. Section 3 provides a Monte Carlo analysis of the power of these procedures in the case of regime changes in the marginal distribution of the conditioning variables. Section 4 examines the superexogeneity of short-term interest rates for the parameters of a dynamic specification for the U.S. term structure based on the linearized expectations model. Some concluding remarks are made in Section 5.

2. TESTING SUPEREXOGENEITY HYPOTHESES

Consider a bivariate random sequence $\{\mathbf{w}_t \equiv (y_t, x_t)': t \geq 1\}$ with associated conditional density functions $\{D(\mathbf{w}_t \mid \mathcal{D}_t; \psi)\}$, where $\mathcal{D}_t$ is the σ-field generated by $\mathbf{W}_t^1 \equiv (\mathbf{z}_t, \mathbf{z}_{t-1}, \mathbf{w}_{t-1}, \ldots, \mathbf{z}_1, \mathbf{w}_1)$, $\{\mathbf{z}_t\}$ is a sequence of vectors of valid conditioning variables, and ψ is a parameter vector. Let (ψ_1, ψ_2) be a reparameterization of ψ which supports the factorization:

$$D(\mathbf{w}_t \mid \mathcal{D}_t; \psi) = D_{y|x}(y_t \mid x_t, \mathcal{D}_t; \psi_1) \cdot D_x(x_t \mid \mathcal{D}_t; \psi_2).$$

Then Engle, Hendry, and Richard (1983) define x_t to be *superexogenous* for a set of parameters of interest ϕ if:

(*a*) x_t is *weakly exogenous* for ϕ, that is ϕ can be uniquely determined from ψ_1 alone, and ψ_1 and ψ_2 are variation free (i.e. not subject to cross-restrictions); and

(*b*) $D_{y|x}(y_t | x_t, \mathcal{D}_t; \psi_1)$ is *structurally invariant*, i.e. ψ_1 is invariant to a class of interventions which affect ψ_2.

It is therefore evident that formal testing of the null hypothesis of superexogeneity of a set of conditioning variables requires formulation of an alternative which allows for failures of both weak exogeneity and invariance and which explicitly specifies the way in which such failures occur.

To outline the testing procedure advocated in Engle and Hendry (1993), let $\{\mathbf{w}_t\}$ have the Gaussian probabilistic structure:

$$(\mathbf{w}_t | \mathcal{D}_t) \sim N(\mu_t, \Sigma_t), \qquad t \geq 1 \tag{1}$$

where

$$\mu_t \equiv (\mu_t^y, \mu_t^x)' = E(\mathbf{w}_t | \mathcal{D}_t),$$

$$\Sigma_t \equiv [\sigma_t^{ij}] = E\{(\mathbf{w}_t - \mu_t)(\mathbf{w}_t - \mu_t)' | \mathcal{D}_t\} \qquad (i, j = y, x).$$

The parameter of interest for the analysis is assumed to be the vector (δ, ρ') in the theoretical behavioural relationship:

$$\mu_t^y = \delta_t(I_t)\mu_t^x + \rho'\mathbf{z}_t, \tag{2}$$

where δ_t is allowed to vary under a set of possible interventions I_t

affecting the distribution of $\{x_t | \mathscr{D}_t\}$.

From the properties of the bivariate normal distribution, it follows that the conditional expectation of y_t given $(x_t, \mathscr{D}_t)$ is:

$$E(y_t | x_t, \mathscr{D}_t) = \beta_t(x_t - \mu_t^x) + \mu_t^y, \tag{3}$$

where $\beta_t = \sigma_t^{yx}(\sigma_t^{xx})^{-1}$. Letting $\varepsilon_t \equiv y_t - E(y_t | x_t, \mathscr{D}_t)$, substitution of (2) into (3) and rearranging yields:

$$y_t = \delta_t(I_t)x_t + \rho' \mathbf{z}_t + [\beta_t - \delta_t(I_t)](x_t - \mu_t^x) + \varepsilon_t. \tag{4}$$

Note that, by construction, $\{\varepsilon_t\}$ defines a Gaussian martingale difference sequence relative to the σ-field generated by $(x_{t+1}, \mathbf{W}^1_{t+1})$, with $E(\varepsilon_t^2) = \sigma_t^{yy} - (\sigma_t^{yx})^2(\sigma_t^{xx})^{-1} \equiv \omega_t$.

It is easy to see now that the conventional linear regression model:

$$y_t = \delta x_t + \rho' \mathbf{z}_t + \varepsilon_t, \qquad \varepsilon_t \sim IN(0, \omega) \tag{5}$$

is a valid basis for inference about (δ, ρ') only if:

(*a*) x_t is weakly exogenous for the parameters of interest; a necessary condition for this is $\beta_t = \delta_t(I_t)$, ensuring that σ_t^{xx}, σ_t^{yx} and μ_t^x do not enter the conditional model;

(*b*) δ_t is invariant to the events in the intervention set I_t, so that $\delta_t(I_t) = \delta_t$, $\forall t$;

(*c*) β_t is constant over time, so that $\beta_t = \beta$, $\forall t$; if, in addition, $\sigma_t^{yy} = \omega + \beta\sigma_t^{xx}$, then $\{\varepsilon_t\}$ is a homoscedastic process with variance ω.

These requirements entail that $\beta = \delta$, and provide a set of necessary conditions for model (5) to constitute a statistically adequate, temporally stable and structurally invariant parametric representation of the conditional mean $E(y_t | x_t, \mathcal{D}_t)$.

To derive a class of tests for the null hypothesis of superexogeneity, Engle and Hendry (1993) let $\delta_t(I_t)$ be a time-invariant function of the moments of $\{x_t | \mathcal{D}_t\}$. The parameter $\delta_t(\mu_t^x, \sigma_t^{xx})$ is thus allowed to vary under the alternative hypothesis according to the approximation:

$$\delta_t(\mu_t^x, \sigma_t^{xx}) = \lambda_0 + \lambda_1\mu_t^x + \lambda_2\sigma_t^{xx} + \lambda_3\sigma_t^{xx}(\mu_t^x)^{-1}, \tag{6}$$

on the assumption that $\mu_t^x \neq 0, \forall t$.

When the coefficient β_t is time-varying (under the null or the alternative), an expansion of the form:

$$\beta_t = \beta_0 + \beta_1\sigma_t^{xx}, \tag{7}$$

combined with (6) and (4) gives rise to the regression model:

$$\begin{aligned} y_t = \lambda_0 x_t + (\beta_0 - \lambda_0)(x_t - \mu_t^x) + \beta_1\sigma_t^{xx}(x_t - \mu_t^x) + \lambda_1(\mu_t^x)^2 \\ + \lambda_2\mu_t^x\sigma_t^{xx} + \lambda_3\sigma_t^{xx} + \rho' z_t + \varepsilon_t. \end{aligned} \tag{8}$$

Under the null of superexogeneity of x_t for (δ, ρ') in (5), the following (testable) conditions are satisfied: $\beta_0 - \lambda_0 = 0$ (weak exogeneity of x_t); $\lambda_i = 0$ (i = 1, 2, 3), taking σ_t^{xx} to have distinct values over different regimes (invariance of δ); $\beta_1 = 0$ (constancy of β_t).[1]

---

[1]Weak exogeneity and invariance alone obviously ensure superexogeneity of the conditioning variable, but as Engle and Hendry (1993, p. 124) point out it is

If, on the other hand, it is maintained that Σ_t is time-invariant, and hence $\beta_t = \beta$, a test for superexogeneity of x_t reduces to a test of $\beta - \lambda_0 = \lambda_1 = 0$ in (9):

$$y_t = \lambda_0 x_t + (\beta - \lambda_0)(x_t - \mu_t^x) + \lambda_1(\mu_t^x)^2 + \rho' z_t + \varepsilon_t. \qquad (9)$$

In what follows $\delta_t(I_t)$ is also allowed to vary with the first two moments of $\{x_t | \mathcal{D}_t\}$ via a linear as well as a quadratic approximation to the function $\delta_t(\mu_t^x, \sigma_t^{xx})$, i.e.:

$$\delta_t(\mu_t^x, \sigma_t^{xx}) = \lambda_0 + \lambda_1 \mu_t^x + \lambda_2 \sigma_t^{xx}, \qquad (10)$$

and

$$\delta_t(\mu_t^x, \sigma_t^{xx}) = \lambda_0 + \lambda_1 \mu_t^x + \lambda_2 \sigma_t^{xx} + \lambda_3 (\mu_t^x)^2 + \lambda_4 (\sigma_t^{xx})^2 + \lambda_5 \mu_t^x \sigma_t^{xx}. \qquad (11)$$

When time heterogeneity in the second moments of $\{\mathbf{w}_t | \mathcal{D}_t\}$ is allowed for, substitution of (10) into (4) and use of expansion (7) yields the test regression:

$$y_t = \lambda_0 x_t + (\beta_0 - \lambda_0)(x_t - \mu_t^x) + \lambda_1(\mu_t^x)^2 + \lambda_2 \mu_t^x \sigma_t^{xx}$$
$$+ \beta_1 \sigma_t^{xx}(x_t - \mu_t^x) + \rho' z_t + \varepsilon_t. \qquad (12)$$

Similarly, the regression model corresponding to (11) is:

---

not clear how the condition for weak exogeneity is to be satisfied if the coefficient β_t varies in unknown ways.

$$y_t = \lambda_0 x_t + (\beta_0 - \lambda_0)(x_t - \mu_t^x) + \lambda_1(\mu_t^x)^2 + \lambda_2\mu_t^x\sigma_t^{xx} + \lambda_3(\mu_t^x)^3$$

$$+ \lambda_4\mu_t^x(\sigma_t^{xx})^2 + \lambda_5\sigma_t^{xx}(\mu_t^x)^2 + \beta_1\sigma_t^{xx}(x_t - \mu_t^x) + \rho' z_t + \varepsilon_t. \qquad (13)$$

Superexogeneity of x_t for (δ, ρ') requires that $\beta_0 - \lambda_0 = 0$, $\beta_1 = 0$ and $\lambda_i = 0$ $(i = 1, ..., 5)$.[2]

An operational test of superexogeneity for historical interventions which have affected the distribution of $\{x_t | \mathcal{D}_t\}$ may be constructed along the lines described above by parameterizing the conditional mean μ_t^x through an appropriate set of instruments which allows for past changes in the marginal process. Inferences about the conditional variance σ_t^{xx} may also be based on such a model, using, for instance, a heteroscedasticity function or the autoregressive conditional heteroscedasticity (ARCH) formulation of Engle (1982) and its various extensions.

A related class of tests examines structural invariance and superexogeneity hypotheses indirectly via tests for parameter constancy and predictive failure. In cases where the marginal process $\{x_t | \mathcal{D}_t\}$ is subject to changes over the sample period, failure of superexogeneity is liable to induce parameter non-constancy in a model of $E(y_t | x_t, \mathcal{D}_t)$. Hence, temporal stability of (5) in the face of a time-varying marginal representation provides a necessary condition for superexogeneity of x_t under the class of interventions which have occurred in the sample [see, e.g., Engle, Hendry, and Richard (1983) and Anderson and Mizon (1989) for further discussion].

---

[2] If Σ_t is taken to be constant over time, the test regressions (12) and (13) do not include terms involving σ_t^{xx}.

3. MONTE CARLO EVIDENCE

To examine the power of the tests discussed in the previous section to detect false superexogeneity claims, a series of simulation experiments is performed for alternative regime shifts in a two-equation model derived from forward-looking behavioural assumptions. Below we first outline the Monte Carlo design, describing the data-generating mechanism chosen and the tests to be analysed. The results of the simulation analysis follow.

3.1 Experimental Design and Simulation

The data-generating process (DGP) used in the Monte Carlo sampling experiments is the following specific, but theoretically relevant, mechanism:

$$y_t = \delta_0 + \delta_1 E(x_{t+1} \mid \mathcal{I}_t) + \varepsilon_{1t}, \tag{14}$$

$$(1 - \theta_1 B - \theta_2 B^2)x_t = \theta_0 + (\theta_3 B + \theta_4 B^2)y_t + \varepsilon_{2t}, \tag{15}$$

$$\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \sim IN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} \right), \qquad t = 1, 2, \ldots, T,$$

where $\mathcal{I}_t$ is the σ-field generated by $\{(y_i, x_i): 1 \le i \le t\}$, and B denotes the lag operator defined by the transformation $B^k w_t = w_{t-k}$. Equation (14) is a typical rational expectations efficient markets model: it could represent, *inter alia*, a present value relation for stock prices, or a model of the term structure of interest rates, or a relation between forward and spot exchange rates.

The regression model of interest is the relationship between y_t and x_t

obtained by substitution of $E(x_{t+1} | \mathcal{D}_t)$ from (15) into (14). This is given by:

$$(1 - \phi_1 B)y_t = \phi_0 + (\phi_2 + \phi_3 B)x_t + \varepsilon^*_{1t}, \tag{16}$$

where

$$\phi_0 = \alpha(\delta_0 + \delta_1\theta_0), \qquad \phi_1 = \alpha\delta_1\theta_4, \qquad \phi_2 = \alpha\delta_1\theta_1, \qquad \phi_3 = \alpha\delta_1\theta_2,$$

$$\begin{bmatrix} \varepsilon^*_{1t} \\ \varepsilon_{2t} \end{bmatrix} \sim IN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \alpha^2\sigma_{11} & \alpha\sigma_{12} \\ \alpha\sigma_{12} & \sigma_{22} \end{bmatrix} \right),$$

with $\alpha = (1 - \delta_1\theta_3)^{-1}$.

From (14)-(15), the conditional expectation of y_t given $(x_t, \mathcal{D}_{t-1})$ implies the regression:

$$(1 - \tilde{\phi}_1 B - \tilde{\phi}_2 B^2)y_t = \tilde{\phi}_0 + (\tilde{\phi}_3 + \tilde{\phi}_4 B + \tilde{\phi}_5 B^2)x_t + u_t, \tag{17}$$

where

$$\tilde{\phi}_0 = \phi_0 - \alpha\theta_0\sigma_{12}\sigma_{22}^{-1}, \qquad \tilde{\phi}_1 = \phi_1 - \alpha\theta_3\sigma_{12}\sigma_{22}^{-1}, \qquad \tilde{\phi}_2 = -\alpha\theta_4\sigma_{12}\sigma_{22}^{-1},$$

$$\tilde{\phi}_3 = \phi_2 + \alpha\sigma_{12}\sigma_{22}^{-1}, \qquad \tilde{\phi}_4 = \phi_3 - \alpha\theta_1\sigma_{12}\sigma_{22}^{-1}, \qquad \tilde{\phi}_5 = -\alpha\theta_2\sigma_{12}\sigma_{22}^{-1},$$

$$u_t \sim IN\big(0, \alpha^2(\sigma_{11} - \sigma_{12}^2\sigma_{22}^{-1})\big).$$

It is now evident that when $\sigma_{12} \neq 0$, x_t is not weakly exogenous for $(\phi_0, \phi_1, \phi_2, \phi_3, \alpha^2\sigma_{11})$ because these parameters are not recoverable from the parameters of the conditional process (17) alone. Weak exogeneity of x_t also fails when $\sigma_{12} = 0$, as can be seen directly from (16) where the ϕ's are not variation free with the parameters of the marginal model (15). In this case however, given that (16) coincides with the conditional expectation

$E(y_t | x_t, \mathscr{D}_{t-1})$, it is unlikely that weak exogeneity be rejected on the basis of a conventional limited-information "exogeneity test" which seeks to detect the presence of $x_t - E(x_t | \mathscr{D}_{t-1})$ in model (16) [see, e.g., Holly (1987) and Pesaran and Smith (1990)]. Finally, since $(\phi_0, \phi_1, \phi_2, \phi_3, \alpha^2\sigma_{11})$ directly depends on the θ's, (16) is not structurally invariant to interventions which alter the parameters of (15). Hence, the autoregressive distributed lag specification in (16) is by construction subject to the Lucas critique due to lack of superexogeneity of x_t.

As observed earlier, superexogeneity claims are testable if it is known that the parameters of the distribution of the conditioning variables have changed over the sample period. Thus, in the experiments conducted, the marginal process $\{x_t | \mathscr{D}_{t-1}\}$ is allowed to undergo a single regime shift, the timing and magnitude of which are deterministically fixed. We consider three different types of regime shifts, namely a change in the intercept θ_0, a change in the variance σ_{22}, and a change in θ_3. Letting n $(0 < n < 1)$ denote the proportion of the sample that precedes the break, structural changes are parameterized as:

$$\theta_{0t} = \theta_0 + \nabla\theta_0 \cdot d_t, \qquad \theta_{3t} = \theta_3 + \nabla\theta_3 \cdot d_t, \qquad \sigma_{22t} = \sigma_{22} + \nabla\sigma_{22} \cdot d_t,$$

when $\nabla\theta_0$, $\nabla\theta_3$ and $\nabla\sigma_{22}$ are respectively the changes in θ_0, θ_3 and σ_{22}, $d_t = 0$ for $1 \le t \le [Tn]$, $d_t = 1$ for $[Tn] + 1 \le t \le T$, and $[Tn]$ is the integer part of Tn.

In the sampling experiments, we consider variable-addition tests for superexogeneity of x_t in (16) based on the Engle-Hendry expansion (6), as well as tests derived on the basis of linear and quadratic approximations of the form (10) and (11). These tests, referred to as *EHT*, *LAT* and *QAT* below, are conventional F-type tests for the joint significance of all terms involving λ in the ordinary least-squares (OLS) regressions (8), (12) and (13) respectively, when $\mathbf{z}_t \equiv (y_{t-1}, x_{t-1})'$. In addition, since in many cases the conditional variance $E\{y_t - E(y_t|x_t, \mathcal{Ø}_{t-1})\}^2 = (1 - \delta_1\theta_3)^{-2}(\sigma_{11} - \sigma_{12}^2\sigma_{22}^{-1})$ is non-constant, we also compute robust superexogeneity Wald tests, utilizing White's (1980) heteroscedasticity-consistent covariance matrix estimator with a degrees-of-freedom correction [see MacKinnon and White (1985)]. Wald statistics are divided by the number of restrictions under the null to obtain tests based upon the central F distribution. Abstracting from problems associated with misspecification of the marginal model for $\{x_t | \mathcal{Ø}_{t-1}\}$,[3] $x_t - \varepsilon_{2t}$ is used as a measure of the conditional mean $\mu_t^x = E(x_t | \mathcal{Ø}_{t-1})$ for all tests; the conditional variance $\sigma_t^{xx} = E\{(x_t - \mu_t^x)^2 | \mathcal{Ø}_{t-1}\}$ is proxied by its natural estimator ε_{2t}^2.[4]

We also examine the power of temporal stability tests to detect the non-constancy of (16) and hence reject structural invariance. A widely used test that might be employed to check for a one-time shift in the regression coefficients at a pre-specified point $[Tn]$ is the split-sample analysis-of-covariance F-test (denoted below by $CF(n)$) [see Chow (1960)].[5] However, since the number and location of break-points are typically unknown in practice, researchers may conduct tests designed to detect alternative hypotheses more general than that of the F-test. In this study we investigate the behaviour

---

[3] Specification errors in the formulation of the marginal model are likely to reduce the power of superexogeneity tests.

[4] Almost identical simulation results are obtained when the predictable first-order ARCH component of ε_{2t} is used as a measure of σ_t^{xx}.

[5] The power of Chow's (1960) prediction error test in a context similar to the one considered here is investigated in Favero and Hendry (1992). However, as Hansen (1991) shows, such tests have no asymptotic local power against changes in parameters other than the regression error variance.

of several Lagrange multiplier tests for the null of constancy of the parameter vector $(\phi_0, \phi_1, \phi_2, \phi_3)$. These include: (*i*) a test for a single shift at time $[Tn]$ (denoted by $LM(n)$); (*ii*) the supremum over $n \in \mathcal{N}$ of the $LM(n)$ test (denoted by *supLM*), when $\mathcal{N} \subset (0, 1)$ and $\mathcal{N}$ is bounded away from zero and one; (*iii*) the average of $LM(n)$ tests over $n \in \mathcal{N}$ (denoted by *meanLM*); and (*iv*) a test based upon the average of the squared forward cumulative scores (denoted by L_c) [for details see Hansen (1990)].

The last four tests are all tests of the same null, but differ in their choice of alternative hypothesis. For $LM(n)$ and *supLM* the alternative is that of a single shift in regime (with known and unknown location, respectively), whereas *meanLM* and L_c treat the parameters as martingale processes (with respectively varying and constant hazard of instability across the sample). Under the hypothesis of constant parameters, $LM(n)$ is asymptotically distributed as a central chi-squared variate with degrees of freedom equal to the number of parameters tested for constancy. The limiting distributions and critical values for the remaining test statistics can be found in Andrews (1993) and Hansen (1990). Following Andrews' (1993) suggestion, we select the trimming region as $\mathcal{N} = [0.15, 0.85]$ in the experiments.

Our Monte Carlo design sets 11 parameters as fixed, i.e.:[6]

$$\delta_0 = 0.3, \quad \delta_1 = 0.5, \quad \theta_0 = -0.14, \quad \theta_1 = 0.06, \quad \theta_2 = -0.01,$$
$$\theta_3 = 0.98, \quad \theta_4 = -0.13, \quad \sigma_{11} = 0.31, \quad \sigma_{22} = 0.14,$$
$$T = 100, \quad n = 0.5.$$

[6]To ensure the relevance of our simulations, these parameter values are obtained from estimation of a model for the spread between six-month and three-month U.S. Treasury Bill rates and the first-differences of the three-month rate.

The error covariance and the magnitude of structural breaks are, on the other hand, chosen by the grid:

$$\sigma_{12} \in \{0.15, 0\},$$
$$\nabla\theta_0 \in \{0.35, 0.7, 1.4, 2.8, 4.2\},$$
$$\nabla\theta_3 \in \{-0.35, -0.7, -1.4, -2.8, -4.2\},$$
$$\nabla\sigma_{22} \in \{0.35, 0.7, 1.4, 2.8, 4.2\}.$$ [7]

In all experiments, 5,000 samples of $51 + T$ observations on (y_t, x_t) are generated from (14)-(15), starting with $y_{-1} = x_{-1} = y_0 = x_0 = 0$. However, in order to attenuate the effect of the choice of initial values, only the last T observations of each sample are used for the calculation of the test statistics of interest. Pseudo-random samples of values of $(\varepsilon_{1t}, \varepsilon_{2t})$ from the bivariate normal distribution are obtained using the RNDN function of the *GAUSS-386i 3.0* matrix programming language.

3.2 Post-simulation Analysis

Table 1 reports rejection frequencies of variable-addition superexogeneity tests in the case of a change in θ_0. These are estimated as the proportion of rejections in 5,000 Monte Carlo replications, using 5% critical values from the upper tail of the central F distribution with appropriate degrees of freedom.[8] When $\sigma_{12} = 0.15$, test rejection frequencies are low for changes

[7] In the model of the U.S. term structure, the equation for the first-differences of the three-month rate had an error variance of 0.14 before 1979:3, which increased to 1.54 during 1979:4-1982:3; changes in the equation coefficients were of much smaller magnitude.

[8] The accuracy of an estimate P of a test rejection frequency can be assessed

approximately equal to one error standard deviation, but rapidly approach unity as the magnitude of the structural break increases. However, when $\sigma_{12} = 0$ all three tests have rejection frequencies which hardly exceed 50%, in spite of immense changes (larger than thirty-fold) in the intercept.

Such results are perhaps more easily understood if we note that, for the parameterization with $\sigma_{12} \neq 0$, the component of the null hypothesis which appears to be largely responsible for rejection is that of weak exogeneity. More specifically, whereas a standard *t*-test for the separate hypothesis of a zero effect from $x_t - \mu_t^x$ has a high rejection frequency, estimates of λ's are rarely statistically different from zero. As a summary statistic, the row of Table 1 labelled *t-inv* shows the maximum rejection frequency of two-sided *t*-type tests for $\lambda_i = 0$ $(i = 1, 2, ...)$ over all experiments (using a two-tailed standard normal critical region of size 0.05). This lack of power of tests for invariance is also apparent for the DGP with $\sigma_{12} = 0$. However, as anticipated from the discussion in Section 3.1, neither do statistics for the significance of the component $x_t - \mu_t^x$ provide powerful tests of weak exogeneity in this case, so the rejection frequency of tests for the joint hypothesis of superexogeneity is reduced dramatically.

When heteroscedasticity-robust tests are employed, the probability of rejecting the null of superexogeneity increases considerably, especially when using *QAT*. Nevertheless, such results must be treated with caution since the exact finite-sample size of heteroscedasticity-robust tests in homoscedastic models tends to be substantially larger than its nominal value [see MacKinnon and White (1985), Chesher (1989), and Chesher and Austin (1991), *inter*

---

by noting that its variance is consistently estimated by $M^{-1}P(1 - P)$, where M is the number of Monte Carlo replications.

alia].[9]

The estimated rejection frequencies of *EHT*, *LAT* and *QAT* in the case of a change in θ_3 are recorded in Table 2. The performance of the tests for the DGP with $\sigma_{12} = 0.15$ appears very impressive, but the rejection frequency never exceeds 20% when $\sigma_{12} = 0$. Moreover, accounting for the structural break in $E\{y_t - E(y_t|x_t, \mathcal{D}_{t-1})\}^2$ induced by $\nabla\theta_3 \neq 0$ by means of heteroscedasticity-robust tests does not improve matters substantially, except perhaps for *QAT*. Note, however, that these figures are probably spuriously high, as our earlier remark about the true rejection frequencies of heteroscedasticity-robust tests under the null hypothesis remains true even for heteroscedastic models.

The rejection frequency of the tests for changes in the variance σ_{22} is even lower, as the first block of Table 3 reveals, presumably partly because of the large error variances induced by $\nabla\sigma_{22} > 0$. When $\sigma_{12} \neq 0$, rejection rarely exceeds 25% for conventional *F*-tests, or 50% for heteroscedasticity-robust tests. For the parameterization with $\sigma_{12} = 0$, on the other hand, the estimated rejection frequencies of OLS-based tests are rarely statistically different from 5%. This should be no surprise since if $\sigma_{12} = 0$, OLS on (16) gives a consistent (albeit inefficient) estimate, and $(\phi_0, \phi_1, \phi_2, \phi_3, \alpha^2\sigma_{11})$ is invariant to changes in σ_{22} (i.e., all test variables ought to enter (16) with zero coefficients). The rejection frequencies of heteroscedasticity-

[9]Chesher (1989) and Chesher and Austin (1991) also show that the magnitude of size distortion is very sensitive to the degree of balance in the regression design. Since designs like (13) were typically found in our experiments to contain points of very high leverage (as measured by the corresponding diagonal element of the matrix that projects orthogonally onto the column space of the regressor matrix - see, e.g., Cook and Weisberg (1982)), the rejection frequencies of *QAT* are highly suspect.

robust tests in this case confirm our earlier reservations about the reliability of tests based on the White covariance matrix estimator when heteroscedasticity is in fact absent.

Finally, in order to examine the effects on the power of the test procedure of ignoring potential non-constancy of the regression coefficient of y_t on x_t conditional on $\mathcal{I}_{t-1}$, we investigate the performance of superexogeneity tests under the false maintained assumption of time-invariance of the second moments of $\{(y_t, x_t)' | \mathcal{I}_{t-1}\}$. The lower block of Table 3 reports estimated rejection frequencies of Wald-type tests for the joint significance of $[x_t - \mu_t^x, (\mu_t^x)^2]$ and $[x_t - \mu_t^x, (\mu_t^x)^2, (\mu_t^x)^3]$ when added to (16), in the case that $\nabla\sigma_{22} \neq 0$ and $\sigma_{12} = 0.15$; the tests are respectively denoted by *EHT.C* and *QAT.C*.[10] Comparison of these rejection frequencies with the corresponding figures in the top block of Table 3 reveals that substantial loss in power can indeed occur as result of failing to allow for a changing regression coefficient under the alternative hypothesis.

We now turn our attention to the power of tests for parameter instability and structural change to detect the non-constancy of (16) induced by regime shifts in the marginal process $\{x_t | \mathcal{I}_{t-1}\}$. Table 4 records the estimated test rejection frequencies for changes in the intercept θ_0. As before, these are obtained from 5,000 Monte Carlo replications, using asymptotic critical values at the 5% significance level. First note that all tests, except for *CF*(0.5), have null rejection frequencies which differ from their nominal level by more than 2.576 Monte Carlo standard errors, a result which should

[10] Note that expansions of the form (6) and (10) lead to the same test regression, when Σ_t is constant.

be borne in mind when comparing the power of different tests. When the null of constancy is false, tests which utilize information on the exact date of the break (*CF*(0.5), *LM*(0.5)) outperform substantially tests designed for alternatives which involve breaks of unknown location (*supLM*) or random walk parameters (*meanLM*, L_c). Even so, for the parameterization with $\sigma_{12} = 0$ *CF*(0.5) and *LM*(0.5) have rejection frequencies which rarely exceed 50%, despite radical changes in θ_0. The performance of L_c is particularly disappointing, the test often being biased for substantial intercept shifts. However, this is not perhaps surprising since L_c is specifically designed to detect gradual shifts in parameters (when the likelihood of instability is constant throughout the sample period).

When $\sigma_{12} = 0.15$ and non-constancy is due to a change in θ_3 in the marginal process, the probability of detecting a break in (16) increases substantially, as the figures in Table 5 reveal: rejection frequencies are nearly 100% even for relatively small changes in θ_3. However, as with variable-addition superexogeneity tests, shifts in the conditional model are much harder to detect when $\sigma_{12} = 0$. Also note that when $\nabla\theta_3 \neq 0$, the estimated rejection frequencies of *CF*(0.5) must be viewed with scepticism since, unlike the Lagrange multiplier tests, *CF*(0.5) is not robust to time-varying regression error variances.

Rejection frequencies of the tests for changes in σ_{22} are given in Table 6. As before, parameter constancy tests perform very well when $\sigma_{12} \neq 0$. In the case of the DGP with $\sigma_{12} = 0$, the estimated rejection frequencies merely reflect the invariance of (16) to changes in the error variance of the marginal process.

Finally, to investigate the possibility that non-constancy in an invalid conditional model induced by shifts in the generating mechanism of the conditioning variables is more difficult to detect than breaks in the marginal process [see Favero and Hendry (1992)], we examine the power of temporal stability tests when applied to (15). To save space, Table 7 records estimated test rejection frequencies only for the DGP which is least favourable to tests applied to (16) (similar results were obtained for other DGP's). In contrast to the results in Table 5, the structural break in the marginal process is always detected with a high probability, even when using tests not specifically designed for alternatives incorporating swift changes in regime.

4. THE TERM STRUCTURE OF INTEREST RATES

In this section, the rational expectations model of the term structure of interest rates is used to illustrate empirically some of the difficulties encountered when testing structural invariance and superexogeneity assertions. As observed earlier, the expectations model (linearized around equilibrium) may be expressed by an equation like (14), i.e.:

$$R_t - r_t \equiv s_t = \delta_0 + \delta_1 E(\Delta r_{t+1} \mid \mathcal{I}_t) + \varepsilon_{1t}. \tag{18}$$

In (18), R_t is the yield to maturity on a two-period pure discount bond, r_t is the yield to maturity on a one-period bond, $\mathcal{I}_t$ is the information set available at time t, ε_{1t} is a zero-mean white-noise innovation orthogonal to $E(\Delta r_{t+1} \mid \mathcal{I}_t)$, $\Delta \equiv 1 - B$, and $\delta_1 = 0.5$.[11] As well as being theoretically

[11] The expectations model has been formulated in terms of the yield spread and the first-difference of the short-term rate in order to ensure that no

appealing, such a model has been found to provide a statistically adequate representation of the short end of the U.S. term structure, once the 1979-1982 regime shift to monetary base control is accounted for [see Driffill, Psaradakis, and Sola (1992) and Sola and Driffill (1992)].

In what follows we investigate the applicability of the Lucas critique in the context of a dynamic model for the spread between long-term and short-term U.S. interest rates. Such a model may be obtained from (18) by imposing a dynamic structure like (15) on the process generating Δr_t, and is therefore intrinsically subject to the Lucas critique [see the discussion on model (14)-(17) in Section 3.1]. The measure for the short-term rate is the yield on three-month Treasury Bills, while the long-term rate is the yield on six-month bills. Both rates are expressed as decimals at annual rates, and the sample consists of quarterly observations for the period 1962:1 to 1987:3.

The following model for the yield spread was developed by sequential reduction of an autoregressive distributed lag specification allowing for up to six lags on (s_t, Δr_t) and for an intercept (estimation is by OLS):[12]

$$\underset{}{\hat{s}_t} = \underset{(0.088)}{0.225} + \underset{(0.216)}{0.632}\, \Delta r_t - \underset{(0.252)}{0.662}\, \Delta r_{t-2} - \underset{(0.292)}{0.924}\, s_{t-1} + \underset{(0.220)}{0.814}\, s_{t-3} \qquad (19)$$

$T = 93$ [1964:3-1987:3], $R^2 = 0.4200$, $\hat{\sigma} = 0.8190$, mean = 0.2025, SD = 1.0519, DW = 2.092, NOR[$\chi^2(2)$] = 32.888, HET[$F(8,79)$] = 4.046, ARCH[$F(4,80)$] = 4.950, SC[$\chi^2(5)$] = 2.109, OV[$\chi^2(1)$] = 2.012.

---

stochastic trend components are present in the variables involved.

[12] All of the empirical results reported in this section were obtained using *PcGive 7* [see Doornik and Hendry (1992)].

Numbers in parentheses beneath coefficients are heteroscedasticity-consistent standard errors (computed as in White (1980), with a degrees-of-freedom correction); R^2 is the squared multiple correlation coefficient; $\hat{\sigma}$ is the equation standard error; mean and SD respectively denote the mean and unconditional standard deviation of the dependent variable; DW is the Durbin-Watson statistic; NOR[·] is the Jarque and Bera (1987) test for non-normal residual skewness and kurtosis; HET[·] is White's (1980) test for residual heteroscedasticity; ARCH[·] is Engle's (1982) test for fourth-order ARCH residuals; SC[·] is a Godfrey (1978)-type test for residual serial correlation up to order five; and OV[·] is a Ramsey (1969) first-order RESET-type test for omitted variables and incorrect functional form. The last two tests are performed by means of the robust artificial regression procedure discussed in Davidson and MacKinnon (1985) and Wooldridge (1990, 1991), thus ensuring that the hypotheses of interest are tested in a manner which is asymptotically valid in the presence of heteroscedasticity of unknown form. The computed statistics indicate substantial residual non-normality and lack of homogeneity in the residual variance, evidence which can be consistent with the presence of outliers and/or unaccounted-for structural shifts.

In order to examine the superexogeneity of Δr_t in (19), we develop an empirical model for the marginal process generating the short-term rate. The following specification was obtained by sequential simplification of a model with six lags on $(s_t, \Delta r_t)$ and various dummy variables designed to allow for and define potential structural breaks/regime shifts that occurred during the period of implementation of the Federal Reserve System's New Operating Procedures:

$$\Delta\hat{r}_t = \underset{(0.040)}{-0.193} + \underset{(0.035)}{0.081}\ \Delta r_{t-5} + \underset{(0.078)}{1.001}\ s_{t-1} - \underset{(0.081)}{0.162}\ D^*79.4_t \cdot s_{t-1}$$

$$+ \underset{(0.073)}{0.300}\ D^*79.4_t \cdot s_{t-3} + \underset{[0.379]}{1.634}\ D79.4_t - \underset{[0.420]}{1.260}\ D80.2_t$$

$$+ \underset{[0.478]}{0.756}\ D80.4_t + \underset{[0.633]}{2.627}\ D81.1_t - \underset{[0.391]}{2.033}\ D82.4_t \qquad (20)$$

$T = 97$ [1963:3-1987:3], $R^2 = 0.8979$, $\hat{\sigma} = 0.3722$, mean = 0.0288, SD = 1.1089, DW = 1.820, NOR[$\chi^2(2)$] = 5.764, HET[$F(13,73)$] = 0.696, ARCH[$F(4,79)$] = 4.062, SC[$\chi^2(5)$] = 2.203, OV[$\chi^2(1)$] = 1.767.

In (20), [·] below coefficient estimates are conventionally calculated OLS standard errors,[13] $D^*79.4_t$ is a step-change dummy variable which takes the value unity over the period 1979:4-1982:3 and zero otherwise, and $Dp.q_t$ denotes an impulse dummy variable which is equal to unity at the q^{th} quarter of year p and zero otherwise. The misspecification test statistics indicate that, apart from some evidence of ARCH effects, (20) provides an adequate characterization of the data. Furthermore, the significance of the terms involving regime-shift dummy variables clearly reflects the importance of changes in monetary policy that took place during the period 1979-1982.

Model (20) is used to obtain proxies for the first two conditional moments of Δr_t. The conditional mean is approximated by $\Delta\hat{r}_t$, whilst the predictable

[13] In the absence of severe heteroscedasticity, tests for the significance of the dummy variables based on conventional estimated standard errors are likely to be more reliable than heteroscedasticity-robust tests, given the tendency of the latter to be adversely affected by the leverage points which isolate the non-zero values of the dummies from the body of the data.

fourth-order ARCH component of $(\Delta r_t - \Delta\hat{r}_t)^2$ (denoted by $\hat{\sigma}_t^{\pi}$) is used as a measure of the conditional variance. Following Hendry and Ericsson (1991), Favero and Hendry (1992) and Engle and Hendry (1993), $\Delta r_t - \Delta\hat{r}_t$, $\hat{\sigma}_t^{\pi}$ and $(\Delta r_t - \Delta\hat{r}_t)^2 - \hat{\sigma}_t^{\pi}$, as well as their one-period lagged values, were then included in (19) and their coefficients were tested for significance. However, none of these variables appeared to have a non-zero effect (either in the most general model or in simplifications thereof), indicating that there is little evidence against superexogeneity of Δr_t in the conditional model.

We also examined the significance in (19) of test variables like those implied by approximations of the form of (6), (10) and (11). In particular, the conditional model was augmented as in equation (13) and was subsequently simplified by sequentially deleting test variables with insignificant coefficients. The final specification included $\hat{\sigma}_t^{\pi}(\Delta r_t - \Delta\hat{r}_t)$ with a coefficient of -3.295 and a heteroscedasticity-consistent *t*-ratio of -2.147, so superexogeneity of Δr_t may now be rejected. It is worth noting however that the result is not robust to alternative procedures for performing a test that remains valid under heteroscedasticity. More specifically, the Davidson and MacKinnon (1985) analogue of a *t*-test for the significance of $\hat{\sigma}_t^{\pi}(\Delta r_t - \Delta\hat{r}_t)$ yields -1.823, while the value of a *t*-statistic based on the jackknife covariance matrix estimator described in MacKinnon and White (1985) is only -1.593. Given that such tests generally outperform those based on the White variance estimator, it would appear that the empirical evidence for superexogeneity of Δr_t in (19) is not particularly strong.

Further, since structural invariance implies that the determinants of parameter non-constancy in the marginal process ought not to affect the conditional model, we tested the significance of the dummy variables in (20) when added to (19). The resulting equation is:

$$\hat{s}_t = \underset{(0.133)}{0.247} + \underset{(0.218)}{0.464}\,\Delta r_t - \underset{(0.189)}{0.265}\,\Delta r_{t-2} - \underset{(0.246)}{0.444}\,s_{t-1} + \underset{(0.261)}{0.353}\,s_{t-3}$$

$$- \underset{[0.789]}{0.835}\,D79.4_t - \underset{[0.867]}{4.211}\,D80.2_t + \underset{[0.914]}{2.128}\,D80.4_t - \underset{[1.323]}{1.051}\,D81.1_t$$

$$+ \underset{[0.834]}{0.398}\,D82.4_t - \underset{(0.176)}{0.275}\,D^{*}79.4_t \cdot s_{t-1} + \underset{(0.224)}{0.015}\,D^{*}79.4_t \cdot s_{t-3} \qquad (21)$$

$T = 93$ [1964:3-1987:3], $R^2 = 0.6145$, $\hat{\sigma} = 0.6960$, DW = 1.90, NOR[$\chi^2(2)$] = 26.030, HET[$F(17,63)$] = 0.304, ARCH[$F(4,73)$] = 0.195, SC[$\chi^2(5)$] = 2.907, OV[$\chi^2(1)$] = 3.327.

Contrary to (19), (21) satisfies the diagnostic checks for heteroscedasticity reported, and it also variance-dominates the original model (the significant non-normality statistic is due to an outlier at 1980:1). Moreover, (19) fails to parsimoniously encompass (21) on an F-test [$F(7,81) = 5.838$], thereby rejecting superexogeneity.

Finally, we investigate the superexogeneity of Δr_t via an examination of the temporal stability properties of the conditional and marginal model. Figure 1 records one-step residuals from a recursive least-squares regression of Δr_t on a constant, Δr_{t-5}, s_{t-1} and s_{t-3}, bordered with $0 \pm 2\hat{\sigma}$ for every increasing sub-sample [cf. Dufour (1982)]. It is visually apparent that, when regime shifts are not accounted for, the equation standard error of the marginal model is non-constant, almost doubling over the sample period. The most notable breaks appear to have occurred during the period 1979-1982.[14] Moreover, the sequence of one-step residuals from (19) shown in Figure 2 indicates that the instability episode in the marginal process is transferred to the conditional model. This is also apparent in Figure 3 which records the recursive estimates of the coefficient of Δr_t in (19), together with a confidence region based upon plus-or-minus twice the estimated coefficient standard error at each sample size. A formal analysis-of-covariance test of parameter constancy over 1964:3-1979:3 versus 1979:4-1987:3 yields $F(5,83) = 6.352$, strongly rejecting stability of (19).[15] Such evidence is clearly inconsistent with superexogeneity of Δr_t in (19) for the class of interventions that have occurred during the sample period. Also note that these results are in full accordance with the Monte Carlo evidence in Section 3. As the latter suggested, when the generating process of the conditioning variables is subject to changes in the error variance, parameter constancy tests are more likely to reject false invariance assertions than direct superexogeneity tests (cf. Tables 3 and 6).

5. CONCLUDING REMARKS

Given the significance and important implications of structural invariance and superexogeneity for econometric modelling, it is essential that such hypotheses be carefully and reliably tested. The procedures proposed by Engle

---

[14] This is confirmed by a variance-ratio test for a break at 1979:3 (which yields $F(28,61) = 3.524$), as well as by Hansen's (1992) test for variance instability (which yields a value of 0.605). Note, however, that there is no significant evidence of a structural change in the coefficients of the marginal model.

[15] Following Goldfeld and Quandt (1978), the test is made asymptotically robust to heteroscedasticity by deflating the data with an estimate of the appropriate sub-period error standard deviation.

and Hendry (1993) provide an ingenious and simple way of doing so, but may yield incorrect information in certain cases. As our Monte Carlo analysis revealed, non-rejection of superexogeneity is indeed compatible with invalid conditioning and low test power, especially so in circumstances where failures of weak exogeneity do not violate the standard orthogonality condition required for consistent least-squares estimation of a conditional model.

In the case of direct superexogeneity tests, test power does not appear to be sensitive to the specification of the relationship between the parameters of interest and the moments of the generating process of the conditioning variables that is contemplated under the alternative hypothesis. Nevertheless, since the test variables implied by such expansions are likely to be collinear, there is much to gain from using the most parsimonious approximation believed to be relevant for the problem under scrutiny.

In connection with indirect tests of superexogeneity, it is worth pointing out that although tests for parameter constancy which utilize correct prior information about the point at which structural change occurs tend to be more powerful than tests that treat the location of the break as unknown, their outcome must be tempered with caution. Since the choice of a candidate break-point is rarely independent of the data on which constancy tests are applied -it is typically determined by historical evidence for shifts in the process of the conditioning variables- inferences based on asymptotic distributional theory may be misleading. Estimates of the parameters of interest generated by rolling and recursive sampling schemes can provide valuable information about the structural invariance of a model, and thus assist considerably in the interpretation of evidence provided by formal tests for temporal stability.

Finally, let us end by noting that the findings of our analysis reenforce earlier evidence against the usefulness of superexogeneity tests as a means of assessing the empirical relevance of forward-looking model-based expectational mechanisms [cf. Favero and Hendry (1992)]. The observation that structural non-invariance fails to hold in an empirical setting involving changing expectations does not necessarily entail rejection of forward-looking behaviour: it may merely reflect low test power due to zero off-diagonal elements in the covariance matrix of the errors of the structural and expectations-generating equations. Since many important economic phenomena are indeed consistent with forward-looking models with disturbances which are contemporaneously uncorrelated with the innovation on the relevant forcing variables, insignificant superexogeneity and invariance tests should be viewed with much caution. A prudent startegy would appear to involve testing of the relevant hypotheses in all possible ways and via a variety of different tests, in an attempt to maximize the possible evidence against maintained superexogeneity and invariance assumptions.

REFERENCES

Aldrich, J. (1989), "Autonomy", *Oxford Economic Papers*, **41**, 15-34; reprinted in de Marchi, N., and C.L. Gilbert (eds.), *History and Methodology of Econometrics*, Clarendon Press, Oxford, 1989.

Anderson, G.J., and G.E. Mizon (1989), "What Can Statistics Contribute to the Analysis of Economic Structural Change?", Chapter 1 in Hackl, P. (ed.), *Statistical Analysis and Forecasting of Economic Structural Change*, Springer-Verlag, Berlin.

Andrews, D.W.K. (1993), "Tests for Parameter Instability and Structural Change with Unknown Change Point", *Econometrica*, **61**, 821-856.

Chesher, A. (1989), "Hájek Inequalities, Measures of Leverage and the Size of Heteroskedasticity Robust Wald Tests", *Econometrica*, **57**, 971-977.

Chesher, A., and G. Austin (1991), "The Finite-Sample Distributions of Heteroskedasticity Robust Wald Statistics", *Journal of Econometrics*, **47**, 153-173.

Chow, G.C. (1960), "Tests of Equality Between Sets of Coefficients in Two Linear Regressions", *Econometrica*, **28**, 591-605.

Cook, R.D., and S. Weisberg (1982), *Residuals and Influence in Regression*, Chapman and Hall, London.

Davidson, R., and J.G. MacKinnon (1985), "Heteroskedasticity-Robust Tests in Regression Directions", *Annales de l'INSEE*, **59/60**, 183-218.

Doornik, J.A., and D.F. Hendry (1992), *PcGive Version 7: An Interactive Econometric Modelling System*, Institute of Economics and Statistics, University of Oxford.

Driffill, J., Z. Psaradakis, and M. Sola (1992), "A Reconciliation of Some Paradoxical Empirical Results on the Expectations Model of the Term Structure", Discussion Paper No. 93/345, Department of Economics, University of Bristol.

Dufour, J.-M. (1982), "Recursive Stability Analysis of Linear Regression Relationships: An Exploratory Methodology", *Journal of Econometrics*, **19**, 31-76.

Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, **50**, 987-1007.

Engle, R.F., and D.F. Hendry (1993), "Testing Superexogeneity and Invariance in Regression Models", *Journal of Econometrics*, **56**, 119-139.

Engle, R.F., D.F. Hendry, and J.-F. Richard (1983), "Exogeneity", *Econometrica*, **51**, 277-304; reprinted as Chapter 15 in Hendry, D.F. (1993), *Econometrics: Alchemy or Science? Essays in Econometric Methodology*, Blackwell, Oxford.

Ericsson, N.R., and D.F. Hendry (1989), "Encompassing and Rational

Expectations: How Sequential Corroboration Can Imply Refutation", International Finance Discussion Paper No. 354, Board of Governors of the Federal Reserve System, Washington, D.C..

Favero, C., and D.F. Hendry (1992), "Testing the Lucas Critique: A Review", *Econometric Reviews*, **11**, 265-318 (with discussion).

Godfrey, L.G. (1978), "Testing for Higher Order Serial Correlation in Regression Equations when the Regressors Include Lagged Dependent Variables", *Econometrica*, **46**, 1303-1310.

Goldfeld, S.M., and R.E. Quandt (1978), "Asymptotic Tests for the Constancy of Regressions in the Heteroskedastic Case", Research Memorandum No. 229, Econometric Research Program, Princeton University.

Hansen, B.E. (1990), "Lagrange Multiplier Tests for Parameter Instability in Non-linear Models", Unpublished Manuscript, Department of Economics, University of Rochester.

Hansen, B.E. (1991), "A Comparison of Tests for Parameter Instability: An Examination of Asymptotic Local Power", Unpublished Manuscript, Department of Economics, University of Rochester.

Hansen, B.E. (1992), "Testing for Parameter Instability in Linear Models", *Journal of Policy Modeling*, **14**, 517-533.

Hendry, D.F. (1988), "The Encompassing Implications of Feedback versus Feedforward Mechanisms in Econometrics", *Oxford Economic Papers*, **40**, 132-149.

Hendry, D.F., and N.R. Ericsson (1991), "Modeling the Demand for Narrow Money in the United Kingdom and the United States", *European Economic Review*, **35**, 833-886.

Holly, A. (1987), "Testing for Exogeneity: A Survey", *Economic Notes*, **3**, 108-130.

Jarque, C.M., and A.K. Bera (1987), "An Efficient Large-Sample Test for Normality of Observations and Regression Residuals", *International Statistical Review*, **55**, 163-172.

Lucas, R.E. Jr. (1976), "Econometric Policy Evaluation: A Critique", in Brunner, K., and A.H. Meltzer (eds.), *The Phillips Curve and Labor Markets*, Carnegie-Rochester Conference Series on Public Policy, Volume 1, North-Holland, Amsterdam, 19-46.

MacKinnon, J.G., and H. White (1985), "Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties", *Journal of Econometrics*, **29**, 305-325.

Morgan, M.S. (1990), *The History of Econometric Ideas*, Cambridge University Press, Cambridge.

Pesaran, M.H., and R.J. Smith (1990), "A Unified Approach to Estimation and Orthogonality Tests in Linear Single-Equation Econometric Models", *Journal of Econometrics*, **44**, 41-66.

Ramsey, J.B. (1969), "Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis", *Journal of the Royal Statistical Society*, Series B, **31**, 350-371.

Sola, M., and J. Driffill (1992), "Testing the Term Structure of Interest Rates Using a Stationary Vector Autoregression with Regime Switching", *Journal of Economic Dynamics and Control*, forthcoming.

White, H. (1980), "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrica*, **48**, 817-838.

Wooldridge, J.M. (1990), "A Unified Approach to Robust, Regression-Based Specification Tests", *Econometric Theory*, **6**, 17-43.

Wooldridge, J.M. (1991), "On the Application of Robust, Regression-Based Diagnostics to Models of Conditional Means and Conditional Variances", *Journal of Econometrics*, **47**, 5-46.

Table 1

Rejection Frequencies of Superexogeneity Tests at Nominal 5% Level, when $\nabla\theta_0 \neq 0$

σ_{12}	$\nabla\theta_0$	*EHT*	*LAT*	*QAT*
0.15	0.35	0.439 (0.595)	0.475 (0.588)	0.389 (0.690)
	0.70	0.772 (0.865)	0.807 (0.864)	0.708 (0.887)
	1.40	0.910 (0.952)	0.929 (0.951)	0.870 (0.953)
	2.80	0.954 (0.976)	0.961 (0.978)	0.923 (0.976)
	4.20	0.968 (0.985)	0.974 (0.985)	0.943 (0.982)
	t-inv	0.057 (0.105)	0.057 (0.103)	0.057 (0.132)
0.00	0.35	0.148 (0.333)	0.159 (0.300)	0.125 (0.505)
	0.70	0.295 (0.510)	0.327 (0.492)	0.243 (0.634)
	1.40	0.432 (0.642)	0.471 (0.628)	0.374 (0.729)
	2.80	0.510 (0.705)	0.554 (0.692)	0.434 (0.768)
	4.20	0.546 (0.739)	0.590 (0.726)	0.469 (0.800)
	t-inv	0.054 (0.102)	0.053 (0.100)	0.055 (0.131)

Notes: Numbers in parentheses are the estimated rejection frequencies of heteroscedasticity-robust tests. Critical values for *EHT*, *LAT* and *QAT* are taken from the $F(5,91)$, $F(4,92)$ and $F(7,89)$ distributions, respectively.

Table 2

Rejection Frequencies of Superexogeneity Tests at Nominal 5% Level, when $\nabla\theta_3 \neq 0$

σ_{12}	$\nabla\theta_3$	*EHT*	*LAT*	*QAT*
0.15	-0.35	0.885 (0.944)	0.912 (0.948)	0.852 (0.962)
	-0.70	0.996 (0.998)	0.997 (0.999)	0.993 (0.999)
	-1.40	0.999 (1.000)	0.999 (1.000)	0.999 (1.000)
	-2.80	0.999 (1.000)	1.000 (1.000)	0.999 (1.000)
	-4.20	0.998 (1.000)	1.000 (1.000)	0.998 (1.000)
	t-inv	0.163 (0.161)	0.162 (0.161)	0.162 (0.179)
0.00	-0.35	0.062 (0.244)	0.067 (0.207)	0.075 (0.494)
	-0.70	0.090 (0.279)	0.095 (0.239)	0.135 (0.567)
	-1.40	0.119 (0.291)	0.121 (0.252)	0.187 (0.581)
	-2.80	0.112 (0.261)	0.108 (0.222)	0.152 (0.497)
	-4.20	0.086 (0.227)	0.079 (0.301)	0.115 (0.438)
	t-inv	0.137 (0.138)	0.137 (0.138)	0.149 (0.166)

Notes: See Table 1.

Table 3

Rejection Frequencies of Superexogeneity Tests at Nominal 5% Level, when $\nabla\sigma_{22} \neq 0$

σ_{12}	$\nabla\sigma_{22}$	*EHT*	*LAT*	*QAT*
0.15	0.35	0.227 (0.455)	0.226 (0.401)	0.215 (0.612)
	0.70	0.277 (0.501)	0.279 (0.443)	0.259 (0.645)
	1.40	0.251 (0.460)	0.251 (0.405)	0.244 (0.625)
	2.80	0.179 (0.400)	0.188 (0.335)	0.175 (0.585)
	4.20	0.138 (0.359)	0.145 (0.304)	0.139 (0.562)
	t-inv	0.084 (0.146)	0.085 (0.143)	0.113 (0.168)
0.00	0.35	0.045 (0.225)	0.049 (0.180)	0.046 (0.453)
	0.70	0.043 (0.234)	0.046 (0.189)	0.047 (0.464)
	1.40	0.046 (0.239)	0.047 (0.189)	0.048 (0.468)
	2.80	0.047 (0.241)	0.047 (0.191)	0.048 (0.473)
	4.20	0.046 (0.243)	0.047 (0.190)	0.048 (0.471)
	t-inv	0.051 (0.121)	0.054 (0.125)	0.057 (0.145)

σ_{12}	$\nabla\sigma_{22}$	*EHT.C*		*QAT.C*
0.15	0.35	0.066 (0.097)		0.063 (0.153)
	0.70	0.055 (0.086)		0.057 (0.147)
	1.40	0.049 (0.078)		0.053 (0.143)
	2.80	0.046 (0.076)		0.047 (0.141)
	4.20	0.044 (0.076)		0.046 (0.140)

Notes: Critical values for *EHT.C* and *QAT.C* are taken from the $F(2,94)$ and $F(3,93)$ distributions, respectively. See also Table 1.

Table 4

Rejection Frequencies of Parameter Constancy Tests for (16) at Nominal 5% Level, when $\nabla\theta_0 \neq 0$

σ_{12}	$\nabla\theta_0$	$CF(0.5)$	$LM(0.5)$	$supLM$	$meanLM$	L_c
0.15	0.00	0.053	0.076*	0.022*	0.119*	0.050
	0.35	0.417	0.447	0.217	0.449	0.239
	0.70	0.780	0.770	0.447	0.673	0.261
	1.40	0.919	0.896	0.580	0.763	0.084
	2.80	0.956	0.939	0.620	0.784	0.022
	4.20	0.968	0.961	0.639	0.784	0.013
0.00	0.00	0.049	0.070*	0.016*	0.103*	0.037*
	0.35	0.147	0.173	0.048	0.194	0.053
	0.70	0.309	0.323	0.091	0.280	0.032
	1.40	0.463	0.451	0.127	0.314	0.009
	2.80	0.544	0.520	0.135	0.306	0.004
	4.20	0.582	0.559	0.143	0.308	0.003

Notes: An asterisk indicates that the estimated rejection frequency is different at the 1% level from the value implied by the relevant asymptotic null distribution. Critical values are taken from the $F(4,92)$ distribution for $CF(0.5)$, from the $\chi^2(4)$ distribution for $LM(0.5)$, from Andrews (1993, Table I) for *supLM*, from Hansen (1990, Table 2) for *meanLM*, and from Hansen (1990, Table 1) for L_c.

Table 5

Rejection Frequencies of Parameter Constancy Tests for (16) at Nominal 5% Level, when $\nabla\theta_3 \neq 0$

σ_{12}	$\nabla\theta_3$	*CF*(0.5)	*LM*(0.5)	*supLM*	*meanLM*	L_c
0.15	-0.35	0.926	0.936	0.675	0.894	0.663
	-0.70	0.998	0.999	0.972	0.995	0.881
	-1.40	1.000	1.000	0.998	1.000	0.895
	-2.80	1.000	1.000	0.999	1.000	0.798
	-4.20	1.000	1.000	1.000	1.000	0.711
0.00	-0.35	0.056	0.098	0.024	0.134	0.033
	-0.70	0.080	0.169	0.050	0.194	0.049
	-1.40	0.142	0.370	0.155	0.377	0.104
	-2.80	0.238	0.762	0.486	0.703	0.198
	-4.20	0.301	0.906	0.744	0.863	0.258

Notes: See Table 4.

Table 6

Rejection Frequencies of Parameter Constancy Tests for (16) at Nominal 5% Level, when $\nabla\sigma_{22} \neq 0$

σ_{12}	$\nabla\sigma_{22}$	*CF*(0.5)	*LM*(0.5)	*supLM*	*meanLM*	L_c
0.15	0.35	0.863	0.874	0.498	0.807	0.443
	0.70	0.965	0.963	0.648	0.897	0.424
	1.40	0.990	0.982	0.700	0.917	0.298
	2.80	0.995	0.987	0.692	0.909	0.158
	4.20	0.997	0.987	0.674	0.902	0.099
0.00	0.35	0.053	0.071	0.018	0.106	0.035
	0.70	0.052	0.071	0.017	0.104	0.033
	1.40	0.053	0.067	0.018	0.099	0.029
	2.80	0.051	0.068	0.019	0.096	0.024
	4.20	0.050	0.066	0.020	0.094	0.020

Notes: See Table 4.

Table 7

Rejection Frequencies of Parameter Constancy Tests for (15) at Nominal 5% Level, when $\nabla\theta_3 \neq 0$

σ_{12}	$\nabla\theta_3$	*CF*(0.5)	*LM*(0.5)	*supLM*	*meanLM*	L_c
0.00	0.00	0.047	0.075*	0.014*	0.115*	0.037*
	-0.35	0.937	0.928	0.602	0.878	0.650
	-0.70	1.000	1.000	0.980	0.997	0.939
	-1.40	1.000	1.000	0.998	1.000	0.959
	-2.80	1.000	1.000	0.999	1.000	0.876
	-4.20	1.000	1.000	0.999	1.000	0.751

Notes: Critical values for *CF*(0.5) and *LM*(0.5) are taken from the *F*(5,90) and

$\chi^2(5)$ distributions, respectively. See also Table 4.

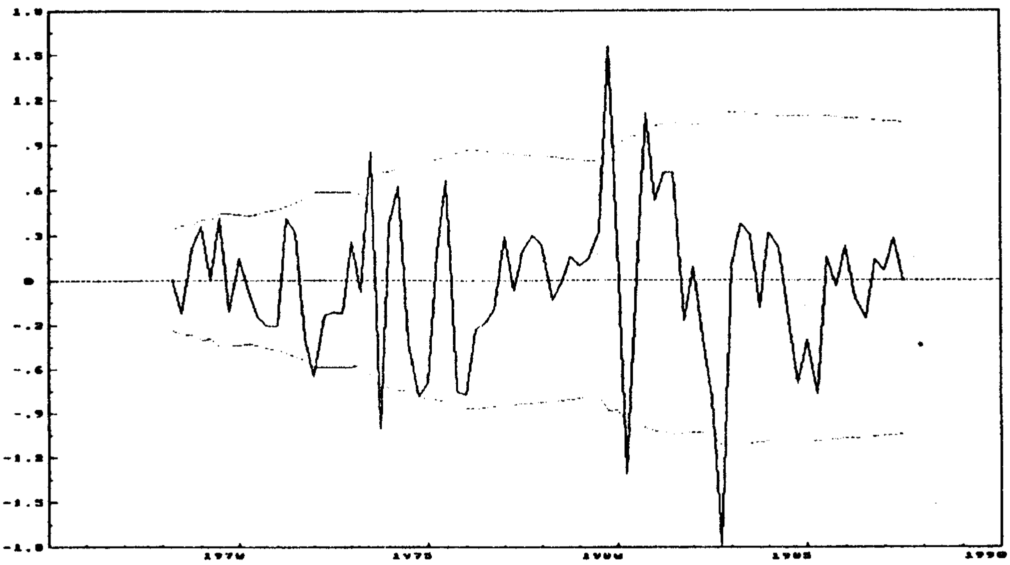

FIGURE 1: One-step residuals and corresponding estimated equation standard errors for the marginal model.

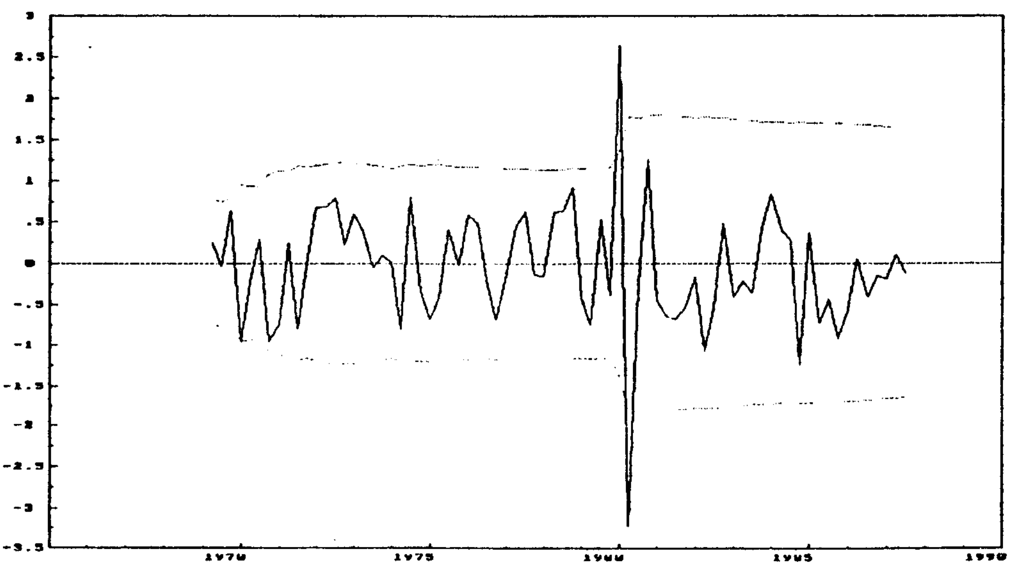

FIGURE 2: One-step residuals and corresponding estimated equation standard errors for the conditional model (19).

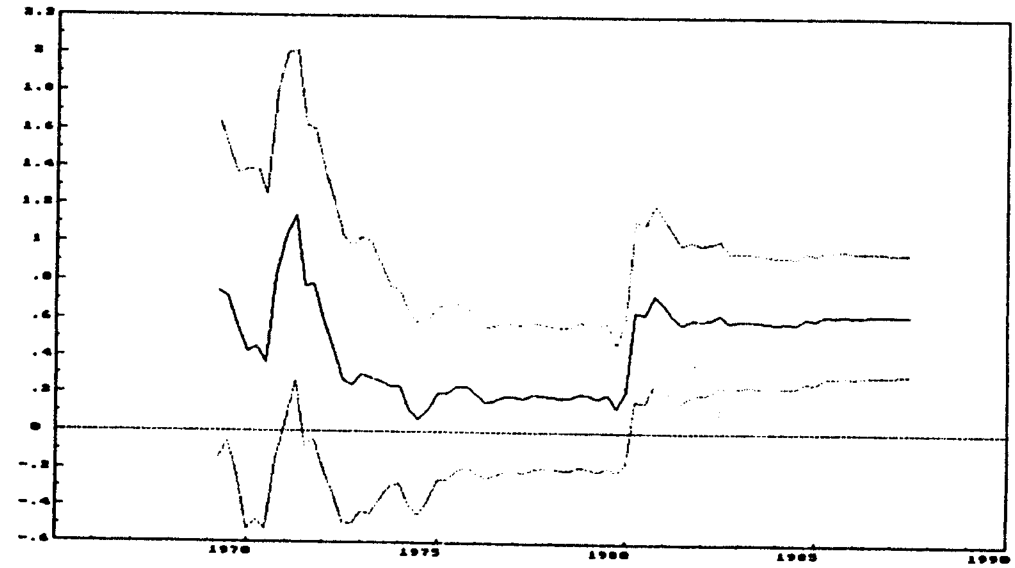

FIGURE 3: Recursive estimates of the coefficient of Δr_t in model (19), with ±2 estimated standard errors.