ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 67/68 – 2002

Policy Evaluation in Macroeconometric Doubly Stochastic Models

Marine CARRASCO, Stéphane GREGOIR *

ABSTRACT. – We consider descriptive macroeconometric models with random coefficients in order to capture the possible relationship between the coefficients in the equations of the variables under the (partial) control of the authorities and those in the behavioral equations as observed in the rational expectation framework. This relationship is the main element of the *Lucas Critique*. This model family allows us to test for the immunity to this critique and implicitly for the presence of contemporaneous breaks as in the approach proposed by HENDRY [1988] and ENGLE and HENDRY [1993]. When the immunity is rejected, we then can also propose a measure of the consequences of some particular changes in the coefficients involved in the policy equation on modeled macroeconomic variables. We illustrate this approach on French data. Starting from descriptive considerations, we specify a system of linear equations with random coefficients to study the response of the French household consumption to the changes in monetary policy linked to the entry in EMS and the Maastricht Treaty.

Évaluation des politiques et modèles macroéconomiques doublement stochastiques

RÉSUMÉ. – Nous considérons des modèles macroéconométriques descriptifs avec des coefficients aléatoires afin de capter les éventuelles relations qui peuvent exister entre les coefficients des équations qui déterminent les variables sous contrôle (partiel) du gouvernement et ceux des équations de comportement comme dans le cadre des modèles à anticipations rationnelles. Ce type de relations est l'élément principal de la *critique de Lucas*. Cette famille de modèles permet de tester si la *critique de Lucas* s'applique et implicitement la présence de ruptures contemporaines comme dans l'approche proposée par HENDRY [1988] et ENGLE et HENDRY [1993]. Quand la critique s'applique, nous pouvons aussi proposer une mesure des conséquences sur les données macroéconomiques de quelques changements de politiques associés à des modifications de coefficients particuliers des équations des variables sous contrôle (partiel) du gouvernement. Nous illustrons cette approche sur des données françaises. En nous appuyant sur des considérations descriptives, nous spécifions un système d'équations linéaires avec des coefficients aléatoires pour étudier la réponse de la consommation des ménages français à une modification de la politique monétaire associée à l'entrée dans le système monétaire européen et à la mise en place des conséquences du traité de Maastricht.

* M. CARRASCO: University of Rochester, Department of Economics; S. GREGOIR: INSEE.
We thank J.-P. FLORENS et P. FÈVE for helpful suggestions and the participants to the Conference "*Econometrics of Policy Evaluation*" Paris, 10-12 January 2000 and CREST, Ente Einaudi (Rome), Montréal University, and UQAM seminars for their comments. We have sole responsibility for any remaining mistakes or inaccuracy.

1 Introduction

In his famous 1976 paper, LUCAS suggested that using a simultaneous equations model for policy analysis and policy simulation is only valid when agents' expectation rules are embedded in its functional form since they are supposed to change with the changes in policy.[1] To state his critique, he used the rational expectation equilibrium (REE) framework in which this property is satisfied by the expectation rules. Nevertheless, the status of these changes has been the basis for critiques of a strict reading of the critique itself (SIMS [1982 and 1987], SARGENT [1984] and COOLEY, LEROY, and RAYMON [1984]) in the sense that if such changes can occur and agents are rational and know the model,[2] they should be able to associate probability measures with these events and they should take them into account when computing their forecasts. If the agents consider that the current value of the parameter associated with the current policy have a large probability to remain unchanged, the regime changes considered by LUCAS are rationally negligible. Nevertheless, the critique keeps its relevance and applied macroeconometricians are asked to evaluate the consequences of a change in policy on some key variables in situation in which agents' decision process may be affected by the change.

The strict reading of the critique led to the recommendation to separate the structural parameters such as those that measure tastes and technology from those associated with policies that are to be modeled as different drawings of random processes (COOLEY, LEROY, and RAYMON [1984]). The question of interest then becomes how to model the changes in policy with relevant stochastic processes based on the economic and political institutions. This is an ambitious task. Another reading of the *Lucas Critique* emphasizes the fact that models are defective approximations. But solving a REE entails establishing the existence of functional relationships between two sets of parameters: Parameters related to a policy choice and those related to agents' behavior. Since non linearities are generally present in the derivation of the forecasting rules, specification errors may have drastic consequences on these functional links that are rarely analytical. In empirical works, this may lead to flimsily measured effects of a policy change.

Considering the above arguments, we choose to neglect the RE framework. An alternative and modest approach consists in relying on more descriptive properties of the data. We introduce a model of which structural features[3] are based on the statistical properties derived from an economic analysis. We propose to use random coefficients in order to model the possible link between the coefficients in the behavioral equations and those related with, in our case, the monetary policy. We are then interested in testing for the presence of this link and in plotting it when it exists.

Econometricians have focused for the last twenty years on the problem of designing a statistical framework to test for the invariance of behavioral equa-

1. Policy changes means for us a change in the value of the structural parameters that define the authority preferences.

2. A basic assumption in rational expectation equilibrium literature.

3. In the sense of HARVEY.

tions to policy changes captured by shifts in the coefficients of the equation for the variables under (partial) control of the authorities (see *inter alia*, HENDRY [1988], ENGLE and HENDRY [1993], FAVERO and HENDRY [1992]). This testing approach only considers simultaneous modifications of the coefficients which is consistent with the rational expectation framework, but should be relaxed in a more realistic framework which would include for instance some learning. Moreover, when the null of absence of a relationship is rejected, it does not offer any guidance on how likely the coefficients are to change again in the future. In contrast, the use of a random coefficient model allows us to quantify the departure from stability and to perform simulations and forecasts under the assumption that simulated changes are close to those that have been observed in the past. If this assumption were not satisfied, any simulation would be meaningless.

The remainder of this paper is divided into four parts. In a second part, we introduce the doubly stochastic processes, characterize in this framework the invariance properties and present an approach for testing for the immunity to the *Lucas Critique*. In Section 3, a small Monte Carlo experiment is commented. The fourth part is devoted to a macroeconomic analysis of the source of breaks observed in the French household consumption function at the beginning of the nineties. A major modification is related to the independence of the French Central Bank and the possible modification of its preferences it implied. This helps us to specify a descriptive model that is estimated and a test for the immunity to the *Lucas Critique* is then performed. A brief conclusion ends the paper.

2 Doubly Stochastic Models

Linear models with random coefficients were the subject of numerous studies in the seventies and eighties (see, for instance, CHOW [1984]). Various estimation methods and test procedures were developed for handling models in which no lagged endogenous variables were present. These models have been used in macroeconomic studies (*inter alia* COOLEY and PRESCOTT [1973]) sometimes in the context of policy evaluation (SIMS [1982]) or forecasting (DOAN, LITTERMAN and SIMS [1984]).

The statistical treatment of models with lagged endogenous variables and random coefficients is more recent and burdensome. In the absence of exogenous variables, these models are called doubly stochastic time series. Their mathematical properties have been studied by *inter alia* POURAHMADI [1986, 1988] and TJOSTHEIM [1986] for univariate time series, and MEYN and GUO [1993] for multivariate time series.

We propose to use a similar set-up to describe simultaneously the dynamics of some well chosen macroeconomic policy-dependent and agent-dependent variables, these dynamics depend on random parameters. This allows us to describe the possible dependence that may exist between both sets of random parameters, those related to a policy choice and those related to agents' beha-

vior. In practice, we postulate linear relationships between endogenous variables (*ie,* policy variables, agents' decisions variables) and exogenous variables. Their coefficients are time varying and may be mutually dependent. For the sake of simplicity, we will assume that the coefficients follow themselves a time series model. This model family can be viewed as a particular kind of conditional heteroskedastic model, which specification is derived from a descriptive structural framework in the spirit of structural time series modeling proposed by HARVEY [1989]. This is consistent with the approach proposed by ENGLE and HENDRY [1993].

2.1 A Simple Framework

We start by motivating our approach and defining the property of invariance for a set of equations in a doubly stochastic set-up. We consider multivariate ergodic stochastic variable (z_t, y_t) whose data generating process is driven by an unobservable multivariate stochastic process α_t that corresponds to some key parameters of the observable DGP. These observable data can be decomposed into two sub-vectors z_t and y_t. z_t includes the variables under the (partial) control of the authorities and y_t represents variables chosen by the private agents. Similarly α_t can be decomposed into two sub-vectors $(\alpha_{z,t}, \alpha_{y,t})$ where $\alpha_{z,t}$ denotes the vector of random coefficients in the government equations and $\alpha_{y,t}$ denotes those in agents' equations. For instance the authorities choose the relative importance they attribute to unemployment and inflation and act on variables such as tax rates and interest rates to reach their aims. Households optimally decide on the level of their consumption knowing authorities' preferences. According to *Lucas Critique*, a change in the policy of the authorities (here a change in the relative weight of the two targets) implies a change in the coefficients in the households' decision equations. When these coefficients are not affected by such a change, the equation is said to be invariant and policy simulations consistent with this kind of changes can be performed.

A change in authorities' preferences could be captured by an intervention function whose general form is $\alpha_{z,t} = g\left(\underline{z_{t-1}}, \underline{y_{t-1}}, \underline{\alpha_{z,t-1}}, \underline{\alpha_{y,t-1}}, \eta_t\right)$ where $\underline{x_s}$ stands for $\{x_s, \ldots, x_0\}$ and η_t is a structural shock. It says that according to what they observed in the past, the authorities may change their preferences and then the coefficients in the equations that describe the way they fix the variables under their (partial) control. $\alpha_{y,t}$ is then said invariant to $g(.)$ if it remains unchanged when $\alpha_{z,t}$ is generated by $g(.)$, technically speaking:

$$\frac{\partial \alpha_{y,t}}{\partial \alpha_{z,t}}\bigg|_{\alpha_{z,t}=g\left(\underline{z_{t-1}}, \underline{y_{t-1}}, \underline{\alpha_{z,t-1}}, \underline{\alpha_{y,t-1}}, \eta_t\right)} = 0$$

For this framework to be operational, we have to complete it in specifying the intervention function and the link that may exist between $\alpha_{y,t}$ and $\alpha_{z,t}$. We must emphasize that without a theory on the way political decisions are taken, it is not possible to learn something about the invariance of $\alpha_{y,t}$ to interventions that did not occur in the past. In practice, we need to be able to

identify and interpret the important past changes from estimated quantities in order to perform informative simulations. In a pragmatic approach, we propose to restrict our attention to simple time-series coefficient generating processes that can mimic various functional forms and complete the model with a linear relationship between $\alpha_{y,t}$ and $\alpha_{z,t}$.

2.2 The Dynamics of Random Coefficients

The introduction of random coefficients may increase significantly the number of coefficients to estimate. Therefore, we propose to limit the use of random coefficients to some key coefficients in the set of equations chosen according to the question we want to answer. Moreover because changes will occur several times over a reasonable period but with unequal persistence and amplitude, we include two sources of fluctuations: the first one reflects rare and persistent changes of a large amplitude while the second one produces small but frequent shocks. We propose to limit our attention to the following family of Coefficient Generating Processes:

$$
\begin{aligned}
T(L)\begin{pmatrix}\alpha_{z,t}-\alpha_z\\ \alpha_{y,t}-\alpha_y\end{pmatrix} &= \Lambda(L)(\mu_t+\rho\eta_t) \Leftrightarrow \\
\begin{pmatrix}T_z(L) & 0\\ 0 & T_y(L)\end{pmatrix}\begin{pmatrix}\alpha_{z,t}-\alpha_z\\ \alpha_{y,t}-\alpha_y\end{pmatrix} &= \begin{pmatrix}\Lambda_z(L)\\ \Lambda_y(L)\end{pmatrix}(\mu_t+\rho\eta_t)
\end{aligned}
\tag{1}
$$

where $T(L), T_z(L)$ and $T_y(L)$ are diagonal polynomial matrices, $\Lambda(L)$, $\Lambda_z(L)$ and $\Lambda_y(L)$ are polynomial vectors, μ_t is a unidimensional dummy variable which splits the sample in sub-samples according to well-known institutional or political modifications and η_t is a unidimensional unit-variance white noise. We therefore focus on particular models in which changes in the time-varying coefficients are driven by a unique source of shocks. The different dynamics of the coefficients are due to different coefficient values and lag structures for each polynomial. More elaborate models can be specified in particular regarding the dimension of the shocks used to describe the policy changes, but this should result from theoretical a prioris. An alternative modeling would consist in assuming that the coefficient processes are integrated of order one.[4] This approach has been used several times in the past (COOLEY and PRESCOTT [1973] or MACNELIS and NEFTÇI [1982] *inter alia*) and assumes that each random coefficient process has his own innovation process independent from the other ones. On the contrary, in our approach, we propose to consider a model in which the source of stochastic shocks in the random coefficients is unique.

In this framework, as the coefficients are *a priori* random, a Bayesian treatment of exogeneity may be of interest, see FLORENS and MOUCHART [1985], but is beyond the scope of this paper. The random coefficient model is more flexible than other nonlinear models also popular in macroeconometrics. The deterministic break-point approach does not provide any tool to do forecasting. The *Markov*-switching model (HAMILTON [1989]) would consist in

4. In section 2.5.2, we give some technical reasons to explain why we do not follow this approach.

assuming that $\alpha_{z,t}$ and $\alpha_{y,t}$ are *Markov* chains that take on a finite number of values (corresponding to various regimes). The Threshold Autoregressive (TAR) model of TONG [1990] has the same disadvantage of offering only a finite number of regimes and similarly the Smooth Transition Autoregressive model (TERASVIRTA [1994]). In these models, the authorities would change their preferences by selecting in a finite set of possible choices according to the value taken by some endogenous or predetermined variable that should not be observable or predictable by the agents. A VAR specification for $\alpha_{y,t}$ and $\alpha_{z,t}$ is simpler. It allows them to take any value in **R**. Moreover, it can be viewed as a weak linear representation of stationary non linear processes such as stationary markov-switching or TAR models.

We need to impose some restrictions on Model (1). First, the model with the above parameterization is not identified, we cannot distinguish between the zero-degree coefficients in the polynomial $\Lambda(L)$ and the standard error of η_t. A possible identifiability condition is $\|\Lambda(0)\| = 1$. This allows for situations in which the zero-degree coefficient of a component of the polynomial $\Lambda(L)$ is zero, capturing the existence of a delay in the modification of the corresponding coefficient.

Second, we want to limit our attention to the canonical weak stationary representation of the stochastic coefficients. We assume that the roots of each couple of polynomials $\left(T_j(L), \Lambda_j(L)\right)$ related to the j^{th} component are different and of modulus strictly larger than one. The link between any couple of parameters $\left(\alpha_{z,i,t}, \alpha_{y,j,t}\right)$, $\alpha_{z,i,t}$ being the ith component of $\alpha_{z,t}$ and $\alpha_{y,j,t}$ being the jth component of $\alpha_{y,t}$, is approximated by:

$$\alpha_{y,j,t} = \alpha_{y,j} + T_{y,j}(L)^{-1} \Lambda_{y,j}(L) \Lambda_{z,i}(L)^{-1} T_{z,i}(L) \left(\alpha_{z,i,t} - \alpha_{z,i}\right)$$

This formula can mimic a large set of functional links.

2.3 Testing for the *Lucas Critique*

ERICSSON and IRONS [1995] find that the *Lucas Critique* does not seem to be an important empirical critique. Their finding may be related to the fact that most of the studies look for coincident breaks in coefficients of the variables under (partial) control of the government and those of the behavioral equation(s). This is consistent with the REE framework. Our testing strategy is slightly different from that proposed by ENGLE, HENDRY and RICHARD [1983], HENDRY [1988], ENGLE and HENDRY [1993] and FAVERO and HENDRY [1992]. Here, we consider a larger family of functional links between the two sets of coefficients, with possible lagged effects. This corresponds to a larger interpretation of the *Lucas Critique* that may give it a stronger empirical relevance. As REE assumptions do not seem satisfied in practice, it is cautious to depart from them. Furthermore, the use of random coefficients seems natural when the number of changes is large and when we are interested in approximating the relationship which may exist between the two sets of parameters. Notice that the co-breaking approach does not provide the tools to perform simulation studies.

In our setting, constancy of the parameters and invariance can be expressly tested for. The constancy of the parameters, $\alpha_{y,t} = \alpha_y$ and $\alpha_{z,t} = \alpha_z$, corresponds to $\rho = 0$ and $(\mu_t) = 0$ (under the initial conditions $\alpha_{y,0} = \alpha_y$ and

$\alpha_{z,0} = \alpha_z$.) The invariance of the α_y to some modifications of the α_z corresponds to the event $A \cup B$ (A or B) where $A = \{\Lambda_y(L) = 0 \text{ and } \Lambda_z(L) \neq 0\}$ and $B = \{\Lambda_z(L) = 0 \text{ and } \Lambda_y(L) \neq 0\}$. This states the non existence of (simultaneous) causality between the two sets of coefficients. It does not preclude the possibility that the $\alpha_{y,t}$ coefficients are varying, but these variations are not related to those of $\alpha_{z,t}$. In general, testing for hypotheses concerning unions of mutually exclusive events is not an easy task. Here, we have to take into account the identification problem evoked in the preceding section, so that under the selected parameterization, $\Lambda_y(L)$ and $\Lambda_z(L)$ cannot be simultaneously equal to 0. We then limit our attention to the null hypothesis $\Lambda_y(L) = 0$. When the invariance hypothesis is rejected, a graphical plot of the relationships between the coefficients can be easily done.

In conclusion, we want to test whether we can parsimoniously describe the changes in some coefficients with a unique source of noise. If this feature is not in the data, the above test strategy should reject this specification. Otherwise, this allows us to capture co-movements between the coefficients. We thus propose to reconsider the concept of invariance in a larger setup and to shift the problem of the possible non-constancy of the parameters in the expectation rule to that of possible propagation of changes in the way these rules are updated.

2.4 Limits and Advantages of this Approach

This approach has at least two limits. First, as said above, it is descriptive. If the changes in policy are not well identified or important enough, the estimation will not be precise. It may be difficult to provide meaningful interpretation of the results if we are not able to refer to a well-known event. Second, its cost in parameters is large. Introducing a random coefficient in a model implies the introduction of several coefficients for modeling its data generating process. The model coefficients that are to be stochastically modeled must be well chosen and the data generating process parsimonious. This choice is to be driven by structural considerations. A risk of overparameterization exists.

Its main advantage is that it provides a framework which allows us to control for the size of the test for immunity to *Lucas Critique* which appears to be the main drawback of the sequential procedure proposed by HENDRY and his co-authors. Nevertheless, the test procedure is not simple. In particular, non-identified nuisance parameters are present under the null hypothesis of parameter constancy. We illustrate this point in the following part.

2.5 Asymptotic Distribution of the Test Statistics

2.5.1 Statistical Framework

In the simplest case, we propose to introduce some random coefficients in a general autoregressive representation of a $(n \times 1)$ vector y_t.

$$y_t = \sum_{j=1}^{p} \psi_{t,j} y_{t-j} + \varepsilon_t$$

We work with the AR(1) representation which can always be derived from the AR(p) one. The data generating process (D.G.P.) we consider is thus as follows:

$$Y_t = \Phi_t Y_{t-1} + \Phi Y_{t-1} + \widetilde{\varepsilon}_t$$

where Y_t is a $np \times 1$ vector composed of the current and up to $(p-1)$ lagged values of the $n \times 1$ vector of observations y_t, Φ_t, Φ are two $np \times np$ matrices such that only the n first rows of $\Phi_t + \Phi$ are composed of (possibly random) matrix coefficients, the $n(p-1) \times np$ remaining rows are equal to $(I_{n(p-1)} \quad 0_{n(p-1)\times n})$. The coefficients in Φ_t are supposed to have zero mean. $\widetilde{\varepsilon}_t = \begin{pmatrix} \varepsilon_t \\ 0_{n(p-1)\times 1} \end{pmatrix}$ with $\varepsilon_t \sim iid\,\mathcal{N}(0,\Sigma)$. We rewrite this model and complete it by introducing the dynamics of the stochastic coefficients present in Φ_t that we stacked into a vector $\tilde{\alpha}_t$. We assume it satisfies an ARMA $(1,q)$ representation as introduced in Section 3.2.

$$\tilde{\alpha}_t = \widetilde{T}\tilde{\alpha}_{t-1} + \tilde{\Lambda}\sqrt{\rho^2}\eta_t + \Gamma_1\eta_{t-1} + ... + \Gamma_q\eta_{t-q}.$$

Let $\alpha_t = \left(\tilde{\alpha}_t, \eta_{t-q+1}, ..., \eta_t\right)'$. The model can be rewritten under the following state-space form:

$$\text{(A)} \qquad \begin{cases} Y_t = Z_t\alpha_t + \Phi Y_{t-1} + \widetilde{\varepsilon}_t \\ \alpha_t = T\alpha_{t-1} + \Lambda\sqrt{\rho^2}\eta_t \end{cases}$$

where α_t is a $k \times 1$ vector composed of the q random coefficients present in the model and the $k-q$ lagged values of the univariate noise η_t which is the only source of randomness of the random coefficients, $\eta_t \sim iid\,\mathcal{N}(0,1)$ and for identification purpose, $\|(I_q \quad 0_{(q\times k-q)})\,\Lambda\| = 1$ and $(0_{k-q\times q} \quad I_{k-q})\,\Lambda = \begin{pmatrix} 0_{k-q-1\times 1} \\ 1 \end{pmatrix}$. Z_t is a $np \times k$ matrix whose coefficients are null or equal to the appropriate components of Y_{t-1}.We stack the first order parameters in a vector denoted $vec\phi$. Similarly the coefficients of the variance-covariance matrix of the random terms Σ are set in a vector $vech\Sigma$. Let $\theta = \left(vec\phi', vech\Sigma', \quad \rho^2\right)' = \left(\underline{\theta}', \rho^2\right)'$. At last we denote $K = (I_n \quad 0_{(n(p-1)\times n)})$ such that $y_t = KY_t$ and $\varepsilon_t = K\widetilde{\varepsilon}_t$.

If in the spirit of Section 2.2, we now introduce some dummy variables in the Coefficient Generating Process, we get:

$$\text{(B)} \qquad \begin{cases} Y_t = Z_t\alpha_t + \Phi Y_{t-1} + \widetilde{\varepsilon}_t \\ \alpha_t = T\alpha_{t-1} + \Lambda\left(\mu_t + \sqrt{\rho^2}\widetilde{\eta}_t\right) \end{cases}$$

with

$$\mu_t = \sum_{j=1}^{s} \mu_j \mathbf{1}_{\{t\in[t_{j-1},\, t_j[\}}$$

and for identification purpose $t_0 > -\infty$ and $t_s = +\infty$. The coefficient vector θ becomes in this case $\left(\underline{\theta}', (\mu_j)_{j\in\{1,...s\}}, \rho^2\right)'$.

2.5.2 **Test Statistics**

We first are interested in testing that the use of random coefficients is relevant in Model (*A*). This amounts to test the null hypothesis H_0: $\rho^2 = 0$ against H_a: $\rho^2 > 0$. The parameter set is thus $\mathbf{R}^+$, (a convex cone). Under the null, the model is a standard autoregressive model (with possibly some constraints on the parameters):

$$Y_t = \Phi Y_{t-1} + \widetilde{\varepsilon}_t$$

The coefficients in Λ and T cannot be identified under the null, they are nuisance parameters we must deal with in computing the distribution of the test statistic. We set all the nuisance parameters non identified under the null in a vector v and assume that they belong to a compact subset N of the real space of appropriate dimension (in the sequel we denote $T = T_v$ and $\Lambda = \Lambda_v$ to remind of the dependence of T and Λ on v). Moreover, under the null, we are on the boundary of the parameter set for ρ^2. A growing literature exists on tests when the parameters are on the boundary under the null (see, for instance, GOURIEROUX, HOLLY and MONFORT [1982], GOURIEROUX and MONFORT [1989]). But in presence of nuisance parameters, the nature of the test procedure is more complicated and was recently tackled by ANDREWS [1999, 2001]. In such cases, the distribution may depend on the (unknown) parameter of the model. However, in the case of Model (*A*), the asymptotic distribution of the likelihood ratio (LR) test of H_0 is nuisance parameter free as stated below. Note that we consider a particular initialization of the state-space algorithm such that the likelihood function does not depend on the nuisance parameters under the null. Another initialization might lead to a more complicated asymptotic distribution. This is the reason why we do not use a random walk representation of the coefficient processes as it is frequently done.

PROPOSITION 1: *In Model (A), let* $\widehat{\theta}_0$ *be MLE under* H_0 *(therefore independent of the nuisance parameters* v*) and* $\widehat{\theta}_v$ *be the MLE estimator under the alternative* H_a*. Assume that the following assumptions hold: (i) there exists* $\delta > 0$*, such that* $\forall v \in N$*, the roots of the polynomial* $\det(I_k - xT_v)$ *have a modulus larger than* $1 + \delta$*, (ii) the roots of the polynomial* $\det(I_n - \Phi(x))$ *have modulus strictly larger than 1. Then, under* H_0

$$LR = -2\left(\ln l_T\left(\widehat{\theta}_0\right) - \sup_{v \in N} \ln l_T\left(\widehat{\theta}_v, v\right)\right) \Longrightarrow \frac{1}{2}\chi^2(0) + \frac{1}{2}\chi^2(1)$$

where $\chi^2(0)$ *is the Dirac distribution at the origin and* $\chi^2(1)$ *is a chi-square distribution with one degree of freedom.*

For significance levels 10%, 5%, and 1%, the critical values of the LR test are 1.642, 2.706, and 5.412. In presence of deterministic terms $(\mu_j)_{j \in \{1,\ldots s\}}$ as in Equation (B) above, the LR test for constancy can also be derived. Testing for parameter stability in Model (B) amounts to testing the null hypothesis $H_0' : \rho^2 = 0, \mu_j = 0, j = 1,\ldots,s$ against $H_a' : \rho^2 > 0$ or $\mu_j \neq 0$ for

some j. The parameter set is $\mathbf{R}^{+} \cup \mathbf{R}^{s}$. Note that the coefficients in Λ and T and the ratio $\left(\frac{t_j}{T}\right)_{j=1,\ldots s}$ are not identified under the null. For the sake of simplicity, we assume that the ratio $\left(\frac{t_j}{T}\right)_{j=1,\ldots s}$ are exogenously given. They correspond to well-known institutional changes and are not estimated. We can construct a similar approach to the previous one to get that:

PROPOSITION 2: *In Model (B), let* $\widehat{\theta_0}$ *be MLE under* H_0' *(therefore independent of the nuisance parameters* v*) and* $\widehat{\theta_v}$ *be the MLE estimator under* H_a'*. Assume that the following assumptions hold: (i) there exists* $\delta > 0$*, such that* $\forall v \in N$*, the roots of the polynomial* $\det(I_k - xT_v)$ *have a modulus larger than* $1+\delta$*, (ii) the roots of the polynomial* $\det(I_n - \Phi(x))$ *have modulus strictly larger than 1. Then, under* H_0'

$$LR = -2\left(\ln l_T\left(\widehat{\theta_0}\right) - \sup_{v \in N} \ln l_T\left(\widehat{\theta_v}, v\right)\right)$$
$$\Longrightarrow \frac{1}{2}\chi^2(0) + \frac{1}{2}\chi^2(1) + \chi^2(s).$$

In our application, $s = 1$. In this case, the critical values of the LR test are equal to 3.492, 4.630, and 7.382 at a 10%, 5% and 1% significance level respectively.

Remark 1: When more than one variance is under test (that is the dimension of ρ is greater than 1), the above test statistics are not nuisance parameter free and simulations are necessary to compute their critical values.

Remark 2: It is not possible to use first step OLS estimation to derive consistent estimates of the error terms (equal here to $K(Z_t\alpha_t + \varepsilon_t)$) in order to test for the presence of heteroskedasticity and detect which variables are involved in its functional form. Indeed, the OLS first order coefficient are not consistent as there exists autocorrelation between the lagged endogenous variables Y_{t-1} and the term $Z_t\alpha_t$.

3 Simulations

We propose to study the power and size properties of our test procedure in a simple monetary model introduced by CLARIDA, GALI, and GERTLER [1999].

3.1 A Simple Theoretical Model

We consider a simple model. Let us denote the output gap x_t (which is the difference between actual and potential output), the inflation rate π_t, and the nominal interest rate i_t. We represent the baseline model in terms of an IS curve:

$$x_t = E_t x_{t+1} - \varphi\left(i_t - E_t \pi_{t+1}\right) + \eta_t \tag{2}$$

and a *Phillips* curve:

$$\pi_t = \lambda x_t + \gamma E_t \pi_{t+1} + \varepsilon_t \tag{3}$$

where

$$\eta_t = \mu \eta_{t-1} + u_t,$$
$$\varepsilon_t = \nu \varepsilon_{t-1} + v_t.$$

and $0 \leqslant \nu, \mu \leqslant 1$ and u_t and v_t are i.i.d. with zero mean and variance σ_u^2 and σ_v^2 respectively. *Taking private agents' expectations as given*, the government chooses i_t so that it minimizes the intertemporal objective function:

$$\min_{i_t} E_t \sum_{\tau=1}^{\infty} \delta^\tau \frac{1}{2}\left(\frac{1}{1+z_\tau^2} x_\tau^2 + \frac{z_\tau^2}{1+z_\tau^2}\pi_\tau^2\right). \tag{4}$$

As the focus of the monetary policy changes over time, it is assumed that the relative weight given by the government to inflation and output can be described by a random process $(Z_t)_t$ that we will specify later. Conditional on z_t, the decision of the government leads to:

$$x_t = -\lambda z_t^2 \pi_t. \tag{5}$$

It is assumed that the agents do not observe z_t. The agents solve (2) and (3). To find the equilibrium solution, we first postulate an expression for π_t:

$$\pi_t = a_t \varepsilon_t. \tag{6}$$

with,

$$a_t = \beta a_{t-1} + w_t.$$

From (3) and (5), we obtain:

$$\begin{aligned}\pi_t &= -\lambda^2 z_t^2 \pi_t + \gamma E_t a_{t+1}\varepsilon_{t+1} + \varepsilon_t \\ &= -\lambda^2 z_t^2 \pi_t + \gamma\beta\nu a_t \varepsilon_t + \varepsilon_t.\end{aligned}$$

So that π_t satisfies:

$$\left(1 + \lambda^2 z_t^2\right)\pi_t = \left(1 + \gamma\beta\nu a_t\right)\varepsilon_t.$$

To be consistent with (6), we need:

$$a_t = \frac{1}{1 - \gamma\beta\nu + \lambda^2 z_t^2}.$$

The dynamics of a_t determines the dynamics of z_t^2. On the other hand, we have a constraint of positivity on z_t^2:

$$z_t^2 = \frac{1}{\lambda^2 a_t} - \frac{(1-\gamma\beta\nu)}{\lambda^2} > 0.$$

For this to be satisfied, a_t needs to lie in the interval $[0,\bar{a}]$ with $\bar{a} = 1/(1-\gamma\beta\nu)$. We therefore choose w_t such that its support is equal to $[0,(1-\beta)\bar{a}]$. More precisely in the simulations, we take w_t i.i.d. $U[0,(1-\beta)\bar{a}]$.

From Equation (2), we can infer the dynamic of i_t so that the general solution is given by:

$$\begin{aligned} &\pi_t = a_t\varepsilon_t, \\ &x_t = -\lambda z_t^2 a_t \varepsilon_t, \\ &i_t = \frac{\lambda}{\varphi} z_t^2 a_t \varepsilon_t - \frac{\nu}{\lambda\varphi} + \nu\beta\left[1 + (1-\gamma\beta\nu)\frac{1}{\lambda\varphi}\right] a_t\varepsilon_t + \frac{1}{\varphi}\eta_t. \end{aligned} \tag{7}$$

Model (2) to (4) is the basis for a simulation study presented in the next section. Note that this model is illustrative only and supposes some simplifying assumptions: (a) the agents have no knowledge whatsoever about the value of z_t, (b) the government objective function has a simple form meant to simplify the calculations.

3.2 Results from Monte Carlo Experiments

Consider two types of models:

Model VAR: We generate data under H_0:

$$\begin{cases} i_t = \phi_{ii} i_{t-1} + \phi_{ix} x_{t-1} + \sigma_i \varepsilon_{it} \\ x_t = \phi_{xx} x_{t-1} + \phi_{xi} i_{t-1} + \sigma_x \varepsilon_{xt} \end{cases} \tag{8}$$

where ε_{it}, and ε_{xt} are i.i.d. bivariate centered normal with correlation σ_{ix}.

Model RE: We generate data from the simple model presented in Equations (2) to (4) which plays the role of the alternative. Using these two sets of simulated data, we estimate a random coefficient model:

$$\begin{cases} i_t = \left(\phi_{ii} + \phi_{ii,t}\right) i_{t-1} + \phi_{ix} x_{t-1} + \sigma_i \varepsilon_{it} \\ x_t = \left(\phi_{xx} + \phi_{xx,t}\right) x_{t-1} + \phi_{xi} i_{t-1} + \sigma_x \varepsilon_{xt} \end{cases} \tag{9}$$

where ε_{it}, and ε_{xt} are supposed to be i.i.d. bivariate centered normal with correlation σ_{ix}. Let $\tilde{\phi}_t = (\phi_{iit}, \phi_{xxt})'$ where $\tilde{\phi}_t$ follows a bivariate stationary ARMA(1,1):

$$\tilde{\phi}_t = \widetilde{D}\tilde{\phi}_{t-1} + \Lambda_1 \rho u_t + \Lambda_2 \rho u_{t-1}$$

with u_t i.i.d. $\mathcal{N}(0,1)$. The matrix $\widetilde{D}$ is diagonal with diagonal element (d_1,d_2), $\Lambda_1 = (\Lambda_{11},\Lambda_{12})'$ and $\Lambda_2 = (\Lambda_{21},\Lambda_{22})'$. Note that the model RE for i_t and x_t is a VAR if the relative weights in the target are fixed constant ($z_t = z$ for all t). However, when z_t is random, Model RE cannot be written exactly as Model (9) because it generates conditional heteroskedasticity that is not taken into account in Model (9). Below we are going to discuss a LR test of (8) versus (9) and apply this test to data generated by RE, we will show that the LR test has power against RE in spite of the misspecification induced by the omitted heteroskedasticity.

Model (9) may be rewritten as:

$$\alpha_t = D\alpha_{t-1} + \Lambda\rho u_t$$

with $\alpha_t = \left(\tilde{\phi}_t', u_t\right)$, $\mu = \left(\tilde{\mu}', 0\right)'$, $\Lambda = (\Lambda_{11},\Lambda_{12},1)'$, and:

$$D = \begin{bmatrix} d_1 & 0 & \Lambda_{21}\rho \\ 0 & d_2 & \Lambda_{22}\rho \\ 0 & 0 & 0 \end{bmatrix}.$$

For identification, we need to impose (1) the sign of the components of Λ_2 and (2) $\|\Lambda_1\| = 1$. We want to test H_0: $\rho^2 = 0$ against H_1: $\rho^2 > 0$. This is not a standard problem for two reasons. First, under H_0, the parameter is at the boundary of its space. Second, under H_0, various nuisance parameters are not identified. Let v denote these parameters. ANDREWS [2001] proposes to use a $\sup_{v\in\mathcal{I}} LR(v)$ test and shows that the asymptotic distribution may depend on the parameters of interest. Here interestingly, the asymptotic distribution is nuisance parameter free and is a mixture of a mass point at zero and a Chi-square distribution as given in Proposition 1.

The simulations are performed with GAUSS. To initialize the optimization procedure 'optmum', we use the parameters obtained from a VAR model (model under H_0).

For all the simulations, the sample size is $n = 100$, and the number of simulations is $rep = 1000$. The empirical critical values are computed from 5000 replications of a sample of 100 generated by Model VAR with parameter values: $\phi_{ii} = 0.3, \phi_{ix} = 0.03, \phi_{xx} = 0.3, \phi_{xi} = 0.03, \sigma_i = 0.1$, $\sigma_x = 0.1$, $\sigma_{ix} = -0.01$ (first line of Table 1). They are equal to 11.237, 7.418, 5.609 for 1%, 5% and 10% significance level respectively. In Table 1, we report the rejection rate of the LR test when data are generated under H_0 (that is Model VAR). In Table 2, we give the same rejection rate when data are generated under the alternative Model RE. The empirical size/power corresponds to the rejection rates when the test statistic is compared with the theoretical critical values (given by the asymptotic distribution). The size-corrected size/power corresponds to the rejection rate when the test statistic is compared with the empirical critical values.

The LR test exhibits an important size distortion in small samples that can be corrected using empirical critical values. A good point is that the LR test exhibits a high power against the RE alternative.

TABLE 1
Size of LR Test

		1%	5%	10%
$\phi_{ii} = 0.3, \phi_{iy} = 0.03, \phi_{yy} = 0.3, \phi_{yi} = 0.03,$	emp. size	.109	.314	.455
$\sigma_i = 0.1, \sigma_y = 0.1, \sigma_{iy} = -0.01$	size corr. size	.010	.050	.10
$\phi_{ii} = 0.3, \phi_{iy} = 0.03, \phi_{yy} = 0.3, \phi_{yi} = 0.03,$	emp. size	.309	.782	.988
$\sigma_i = 0.4, \sigma_y = 0.4, \sigma_{iy} = -0.01$	size corr. size	.030	.146	.293
$\phi_{ii} = 0.6, \phi_{iy} = 0.03, \phi_{yy} = 0.6, \phi_{yi} = 0.03,$	emp. size	.111	.300	.462
$\sigma_i = 0.1, \sigma_y = 0.1, \sigma_{iy} = -0.01$	size corr. size	.010	.041	.099
$\phi_{ii} = 0.3, \phi_{iy} = 0.03, \phi_{yy} = 0.3, \phi_{yi} = 0.03,$	emp. size	.179	.475	.674
$\sigma_i = 0.2, \sigma_y = 0.2, \sigma_{iy} = -0.1$	size corr. size	.022	.085	.169

TABLE 2
Power of LR Test

		1%	5%	10%
$\nu = 0.2$, $\mu = 0.5$, $\beta = 0.8$, $\gamma = 0.6$,	emp. power	1	1	1
$\lambda = 0.4, \varphi = 0.05$, $\sigma_v = 0.2$, $\sigma_u = 0.2$.	size corr. power	.812	1	1
$\nu = 0.5$, $\mu = 0.5$, $\beta = 1$, $\gamma = 0.6$,	emp. power	1	1	1
$\lambda = 0.4, \varphi = 0.05$, $\sigma_v = 0.2$, $\sigma_u = 0.2$.	size corr. power	.723	1	1
$\nu = 0.2$, $\mu = 0.5$, $\beta = 0.8$, $\gamma = 0.6$,	emp. power	1	1	1
$\lambda = 0.4, \varphi = 0.05$, $\sigma_v = 0.2$, $\sigma_u = 0.2$.	size corr. power	.812	1	1
$\nu = 0.2$, $\mu = 0.5$, $\beta = 0.8$, $\gamma = 0.6$,	emp. power	1	1	1
$\lambda = 0.4, \varphi = 0.05$, $\sigma_v = 0.2$, $\sigma_u = 0.2$.	size corr. power	.812	1	1
$\nu = 0.2$, $\mu = 0.5$, $\beta = 0.8$, $\gamma = 0.8$,	emp. power	1	1	1
$\lambda = 0.4, \varphi = 0.05$, $\sigma_v = 0.2$, $\sigma_u = 0.2$.	size corr. power	.823	1	1
$\nu = 0$, $\mu = 0$, $\beta = 0.8$, $\gamma = 0.6$,	emp. power	1	1	1
$\lambda = 0.4, \varphi = 0.05$, $\sigma_v = 0.2$, $\sigma_u = 0.2$.	size corr. power	0.839	1	1

4 French Household Consumption and the Independence of the Central Bank

4.1 Some Possible Economic Mechanisms

4.1.1 A Basic Framework

As an illustration, we propose to consider the influence of the expected changes in monetary policy on household consumption decisions. We draw on a now standard literature devoted to the effects of monetary policy in presence

of temporary nominal price rigidities.[5] These rigidities create a short term dilemma between the inflation and the output gap and may generate unemployment. Hence, the monetary authorities select their control variable –the nominal short term interest rate– in order to satisfy at best their conflicting objectives regarding the output gap and the inflation rate.

Our illustration deals with French economy. In a small open economy with a large index of openness such as France, we expect the nominal exchange rate to have also a direct influence on the inflation rate. Moreover, the nominal exchange rate is very likely one of the monetary authorities' objectives as the country is involved in a monetary agreement on quasi-fixed parities with some foreign countries within the European Monetary System (EMS). Meeting this obligation contributes to the credibility of the monetary policy. A measure of the output gap, the inflation rate and the nominal exchange rate seem to be good candidates to play the role of the conflicting objectives of the monetary policy in a small open economy such as France.

On the one hand, empirical studies have illustrated at a micro and an aggregate level the influence of real interest rate on the household consumption (WICKENS and MOLANA [1984], ATTANAZIO and WEBER [1993]). On the other hand, recent papers have illustrated empirically and theoretically the importance of income uncertainty on household consumption (see *inter alia* CARROLL [1997], CARROLL and KIMBALL [1996], DEATON [1991], ZELDES [1989]). A change in the relative weights in the objective function of the central bank will result in different levels of real interest rate and of unemployment. Once aware of this change, the households will revise their consumption path consequently. Indeed, this change modifies their (perception of the) probability of been unemployed and the real return of assets, modifications which alter the mean and the variance of their real wealth path. We would expect that such an uncertainty may imply an increase of their savings and a reduction of their consumption. We propose to study if it is statistically possible to detect such a modification in some coefficients of a set of descriptive equations linking household consumption and macroeconomic variables monitored by monetary authorities. We particularly focus on a date corresponding to an institutional modification that may have induced a possible change in the way monetary policy is set.

4.1.2 A Brief Historical Overview

Historically, French Monetary policy went through various regimes. Capital control accompanied by credit control and price control was the main tool of the French monetary policy up to the beginning of the eighties. In 1973, France entered the European currency snake, but left it in January 1974. In 1979, France entered the European Monetary System. Three devaluations occurred between May 1981 and March 1983. After 1983, the monetary policy aimed at controlling the inflation rate and ensuring the stability of the French Franc parity inside the European monetary system. This was accompanied by a break in the wage setting rule. This led to swift adjustments of the French nominal interest rates to the changes of those of its European partners, mainly Germany. French monetary authorities have been clearly and heavily

5. See *inter alia* GOODFRIEND and KING [1997], WALSH [1998].

influenced by the German monetary policy since 1979 in order to maintain the French Franc parity in a long term concern of building the European community. In particular, parities were not adjusted in the EMS (except for the Italian Lira) after the Basles-Nyborg agreements at the beginning of 1987. CLARIDA, GALI and GERTLER [1998] have illustrated the influence of German monetary policy on its European partners and concluded that this led to interest rates higher than warranted by domestic macroeconomic conditions. The Maastricht Treaty was concluded on December 10, 1991 by the leaders of EC nations in Maastricht, the Netherlands, and formally signed on February 7, 1992. In the referendum of September 1992, the French voted only narrowly in favor of it. As a consequence of the Treaty, French Central Bank got its independence at the end of 1993.

On the other hand, a deregulation of the French financial market took place in late 1985, early 1986. This induced a decrease of firms' credit demand to the banks. In response, the latter increased their supply to households. This contributed significantly to the increase in household consumption up to the end of the eighties. The international slowdown at the beginning of the nineties dampened the growth of French household disposable income and put in economic difficulty a large number of households heavily in debt. An adjustment was necessary and was implemented through a set of laws (lois Neiertz in 1993). Empirically, we observed in France at the beginning of the nineties, a break in the standard empirical household consumption function. Some empirical works have tried to find an omitted variable which would make this relationship stable (BONNET and DUBOIS [1995] *inter alia*). A good candidate was the unemployment rate, although the relevance of its interaction with the household consumption decision at the beginning of the nineties and not before this period, was not really analyzed. We think that the above arguments about uncertainty give a rationale to the presence of unemployment rate in the household consumption equation. An additional argument should be given to explain the introduction of this variable only at the beginning of the nineties. The source of the break in the consumption equation may be related to the uncertainty associated with the possible institutional changes implied by the European construction and the Maastricht Treaty. The French Central Bank was to become independent and its objectives were to be restricted to the level of inflation and the parity French Franc against Deutschemark. This might have distorted the relationship between inflation and unemployment and might justify the introduction of the unemployment rate in the consumption equation.

As a starting point, without structural model that would allow us to distinguish between different possible explanations, we propose simply to investigate if the breaks in the household consumption equation coefficients can be related to a break in the way interest rate was set by the monetary authorities at the beginning of the nineties.

4.2 A Simple Multivariate Model

Our model is a simple descriptive VARX model of six variables whose choice is induced by the preceding descriptive section. These are the real exchange rate, the interest rate, the unemployment rate, the inflation rate, the

TABLE 3
The Data

i_t	real interest rate quarterly average	Banque de France
u_t	unemployment rate	INSEE
$\pi_t = \log(\frac{P_t}{P_{t-1}})$	logarithm of CPI	INSEE
c_t	logarithm of real consumption per capita	INSEE
e_t	logarithm of real exchange rate	OECD
y_t	logarithm of real gross disposable income per capita	INSEE

household consumption per capita and the real disposable income per capita that are observed on quarterly basis from 1970 I to 1998 III. Their sources are given in Table 3. We add the following assumptions in our descriptive model: *(i)* we limit the dimension of the source of policy shocks to be one as we did so far, *(ii)* we consider the exchange rate as an exogenous variable that we do not try to model, *(iii)* the monetary authorities try to offset the observed value of supply shocks, conditionally on that of the exchange rate. This modeling actually permits to quantify the correlation between interest rate selected by the authorities and the monetary policy indicators, namely output or unemployment rate and inflation, which are likely to vary with changes in the policy rule, and some parameters in the consumption equation.

The vector of endogenous variables denoted Y_t is composed on the one hand of variables directly influenced by the monetary policy, namely i_t, Δu_t, and π_t, and on the other hand of the real consumption and the real disposable income.

$$Y_t = \begin{pmatrix} i_t \\ \Delta u_t \\ \pi_t \\ \Delta c_t \\ \Delta y_t \end{pmatrix}$$

We consider a simple constrained AR representation. The choice of the lag order in the Coefficient Generating Process and Conditional Data Generating process should be based on statistical criteria. In the absence of an appropriate theory and criteria, we choose to fix the lag order to 2. We assume that there exists a cointegration relationship between income and consumption with a cointegrating vector equal to $(1 \quad -1)$ as was observed in various studies (see for instance BLOCH and MAUREL [1992]). This amounts to the introduction of a variable related to saving per head in the consumption equation.

Without a structural framework and *a priori* identification assumptions, we are unable to distinguish between a change in the target values and a change in the relative weights of the conflicting objectives; changes in the weights will induce changes in the first and second order parameters of the model (this is consistent with ENGLE and HENDRY [1993] approach who suggests that heteroskedastic models are generally to be considered). We read our

VARX specification as a set of behavioral equations. Two equations will include random coefficients: the interest rate and the household consumption equations. The first one describes the rule followed on average by the authorities to set the interest rate according to the past observed values of the unemployment rate and inflation, the current value of the exchange rate and the preceding selected values of the interest rate. These last variables capture their smoothing practice. Stochastic shocks correspond to innovative departure to the usual rule. We assume that four coefficients are random in this equation. These are *(i)* the intercept term, *(ii)* the lagged interest rate coefficient, *(iii)* the lagged unemployment rate coefficient and *(iv)* the lagged inflation coefficient. They are supposed to describe the consequences on the rule of possible changes in authorities' preferences. The second equation, that includes random coefficients, describes household consumption choice according to the last changes in income, inflation, unemployment rate and interest rate. Furthermore, it is assumed that this decision is taken conditionally on the current departure from the usual interest rate setting rule. Two of its coefficients are considered as random: one is related to the lagged value of the interest rate and the other one to the cointegration relationship, *ie*, their past saving decision. Both are supposed to be the channel through which households adapt their behaviors to changes in authorities' preferences.

In short, we work on:

$$
(10)\begin{cases}
i_t = & \left(m_i + m_{i,t}\right) + \delta_i \Delta e_t + \left(\phi_{ii} + \phi_{ii,t}\right) i_{t-1} + \phi_{ii,2} i_{t-2} \\
& + \left(\phi_{iu} + \phi_{iu,t}\right) \Delta u_{t-1} + \left(\phi_{i\pi} + \phi_{i\pi,t}\right) \pi_{t-1} + \Lambda_{3,i} v_t \xi_t \\
\Delta u_t = & m_u + \delta_u \Delta e_t + \phi_{ui} i_{t-1} + \phi_{uu,1} \Delta u_{t-1} \\
& + \phi_{uu,2} \Delta u_{t-2} + \phi_{u\pi} \pi_{t-1} + \Lambda_{3u} v_t \xi_t + \varepsilon_{u,t} \\
\pi_t = & m_\pi + \delta_\pi \Delta e_t + \phi_{\pi i} i_{t-1} + \phi_{\pi u} \Delta u_{t-1} \\
& + \phi_{\pi\pi,1} \pi_{t-1} + \phi_{\pi\pi,2} \pi_{t-2} + \Lambda_{3,\pi} v_t \xi_t + \varepsilon_{\pi,t} \\
\Delta c_t = & \phi_{cc,1} \Delta c_{t-1} + \phi_{cc,2} \Delta c_{t-2} + \phi_{cy} \Delta y_{t-1} + \left(\phi_c + \phi_{c,t}\right) (c_{t-1} - y_{t-1}) \\
& + \left(\phi_{ci} + \phi_{ci,t}\right) i_{t-1} + \phi_{cu} \Delta u_{t-1} + \phi_{c\pi} \pi_{t-1} + \Lambda_{3,c} v_t \xi_t + \varepsilon_{c,t} \\
\Delta y_t = & \phi_{yc} \Delta c_{t-1} + \phi_{yy,1} \Delta y_{t-1} + \phi_{yy,2} \Delta y_{t-2} + \phi_y (c_{t-1} - y_{t-1}) \\
& + \phi_{yi} i_{t-1} + \phi_{yu} \Delta u_{t-1} + \phi_{y\pi} \pi_{t-1} + \Lambda_{3,y} v_t \xi_t + \varepsilon_{y,t}
\end{cases}
$$

We complete these equations with the definition of the deterministic functions μ_t and v_t and the data generating process of the random coefficients. μ_t and v_t are such that:

$$
(11) \qquad \mu_t = \begin{cases} 0 & \text{if} \quad t \leqslant T_1 \\ \mu_1 & \text{if} \quad T_1 < t \end{cases}
$$

and,

$$
v_t = \begin{cases} 1 & \text{if} \quad t \leqslant T_1 \\ v_1 & \text{if} \quad T_1 < t \leqslant T_2 \\ v_2 & \text{if} \quad T_2 < t \end{cases}
$$

The date T_1 corresponds to France joining EMS (March 1979) and T_2 to the deregulation of the French financial market (October 1985). They permit to take into account these institutional changes. ξ_t is a scalar white noise $\mathcal{N}(0,1)$

and corresponds to the structural innovation in the interest rate policy used to offset inflationary shocks. $\varepsilon_t = (\varepsilon_{ut}, \varepsilon_{\pi t}, \varepsilon_{ct}, \varepsilon_{yt})'$ is i.i.d. $\mathcal{N}(0, H)$ with H a 4×4 – symmetric full-rank matrix.

Let α_t denote the vector of random coefficients in (10). We assume that it satisfies the following DGP:

$$\alpha_t = D\alpha_{t-1} + \Lambda_1 (\mu_t + \rho\eta_t) + \Lambda_2 \left(\mu_{t-1} + \rho\eta_{t-1}\right) \tag{12}$$

where D is a 6×6 – diagonal matrix, we restrain these parameters to be positive, Λ_1 is a 6×1 – vector such that $\|\Lambda_1\| = 1$, η_t is a scalar white noise $\mathcal{N}(0,1)$. (10) and (12) can be estimated by *Kalman* filter. See HARVEY [1989] for a review of this method. We initialize the algorithm by the result of the estimation of a VAR model based on (10) assuming that the coefficients are constant and $\mu_1 = 0$ and $v_t = 1$, for all t.

Estimation result summary: Regarding the random coefficient generating process, the estimation results are summarized in the following set of equations:

$$\begin{cases} m_{i,t} = \underset{(0.53)}{0.26}\ m_{i,t-1} + \underset{(\mathbf{11.57})}{0.85}\ (\rho\,\eta_t + \mu_t) - \underset{(\mathbf{-2.22})}{0.47}\ \left(\rho\,\eta_{t-1} + \mu_{t-1}\right) \\ \phi_{ii,t} = \underset{(\mathbf{5.43})}{0.63}\ \phi_{ii,t-1} - \underset{(\mathbf{-22.50})}{0.39}\ (\rho\,\eta_t + \mu_t) + \underset{(\mathbf{9.20})}{0.39}\ \left(\rho\,\eta_{t-1} + \mu_{t-1}\right) \\ \phi_{iu,t} = \underset{(\mathbf{2.87})}{0.69}\ \phi_{iu,t-1} + \underset{(0.00)}{0.00}\ (\rho\,\eta_t + \mu_t) - \underset{(-0.55)}{0.37}\ \left(\rho\,\eta_{t-1} + \mu_{t-1}\right) \\ \phi_{i\pi,t} = \underset{(\mathbf{3.00})}{0.32}\ \phi_{i\pi,t-1} + \underset{(1.66)}{0.32}\ (\rho\,\eta_t + \mu_t) - \underset{(\mathbf{-7.36})}{0.82}\ \left(\rho\,\eta_{t-1} + \mu_{t-1}\right) \\ \phi_{ci,t} = \underset{(\mathbf{15.26})}{0.93}\ \phi_{ci,t-1} + \underset{(\mathbf{2.25})}{0.09}\ (\rho\,\eta_t + \mu_t) + \underset{(0.58)}{0.03}\ \left(\rho\,\eta_{t-1} + \mu_{t-1}\right) \\ \phi_{c,t} = \underset{(\mathbf{4.83})}{0.72}\ \phi_{c,t-1} + \underset{(0.94)}{0.07}\ (\rho\,\eta_t + \mu_t) + \underset{(1.68)}{0.14}\ \left(\rho\,\eta_{t-1} + \mu_{t-1}\right) \end{cases}$$

where $\widehat{\rho^2} = \underset{(\mathbf{3.05})}{0.36}$, $\widehat{\mu_1} = \underset{(-0.02)}{-0.01}$ and μ_t is defined in (11). The values in parentheses are t-statistics (in bold if significant at a 5% level). In the case of ρ^2, this value (3.05) should be compared with the percentile of a half normal, it is significant at a 1% level (the critical values are 1.28, 1.64, and 2.32 at 10%, 5% and 1% level). Hence we conclude that random coefficients are present in the model. But as the t-statistics associated with $\widehat{\mu_1}$ is – 0.02, the presence of a deterministic break at date T_1 is rejected. All the random coefficient equations contain estimates that are significantly different from 0 in a 5% level test. $\left(\phi_{ii,t}\right)_t$ seems well described by a ARMA(1,1) model with a root close to unity in the MA part. $\left(m_{i,t}\right)_t$ seems to follow an MA model and the other coefficients have an AR component. When using the test introduced in Section 2 to test for the presence of random coefficients, we get the following likelihood ratio tests:

• testing for H_0: $\rho^2 = 0$

$$-2\left(\ln l_T\left(\widehat{\theta_0}\right) - \sup_{v \in N} \ln l_T\left(\widehat{\theta_v}, v\right)\right) = -2(-388.09 - (-325.33)) = 125.52$$

• testing for H_0: $\rho^2 = 0$ and $\mu_1 = 0$

$$-2\left(\ln l_T\left(\widehat{\theta}_0\right) - \sup_{v \in N} \ln l_T\left(\widehat{\theta}_v, v\right)\right) = -2(-408.91 - (-325.33)) = 167.16$$

In both cases, we reject the null. We therefore reject the absence of random coefficients in the model, but clearly the deterministic break is not to be considered. We now turn to the invariance properties of $\phi_{ci,t}$ and $\phi_{c,t}$ to changes in $m_{i,t}$, $\phi_{ii,t}$, $\phi_{iu,t}$ and $\phi_{i\pi,t}$. Here, it corresponds with obvious notations to the test for $\Lambda_{\phi_c,1} = 0$, $\Lambda_{\phi_c,2} = 0$, $\Lambda_{\phi_{ci},1} = 0$ and $\Lambda_{\phi_{ci},2} = 0$. Technically speaking, the two associated autoregressive coefficients are not identified under the null. We are in a similar situation to that of testing for the presence of random coefficients in the model. The usual asymptotic results may not be valid and in the test decision, the use of a χ^2 percentile may be wrong. Simulations of the asymptotic law for a given couple of values of the autoregressive parameters and maximization over a grid of values for these nuisance parameters may be necessary. We get the following likelihood ratio test:

• when testing for H_0: $\Lambda_{\phi_c,1} = \Lambda_{\phi_c,2} = \Lambda_{\phi_{ci},1} = \Lambda_{\phi_{ci},2} = 0$

$$-2\left(\ln l_T\left(\widehat{\theta}_0\right) - \ln l_T\left(\widehat{\theta}\right)\right) = -2(-355.82 - (-325.33)) = 60.98$$

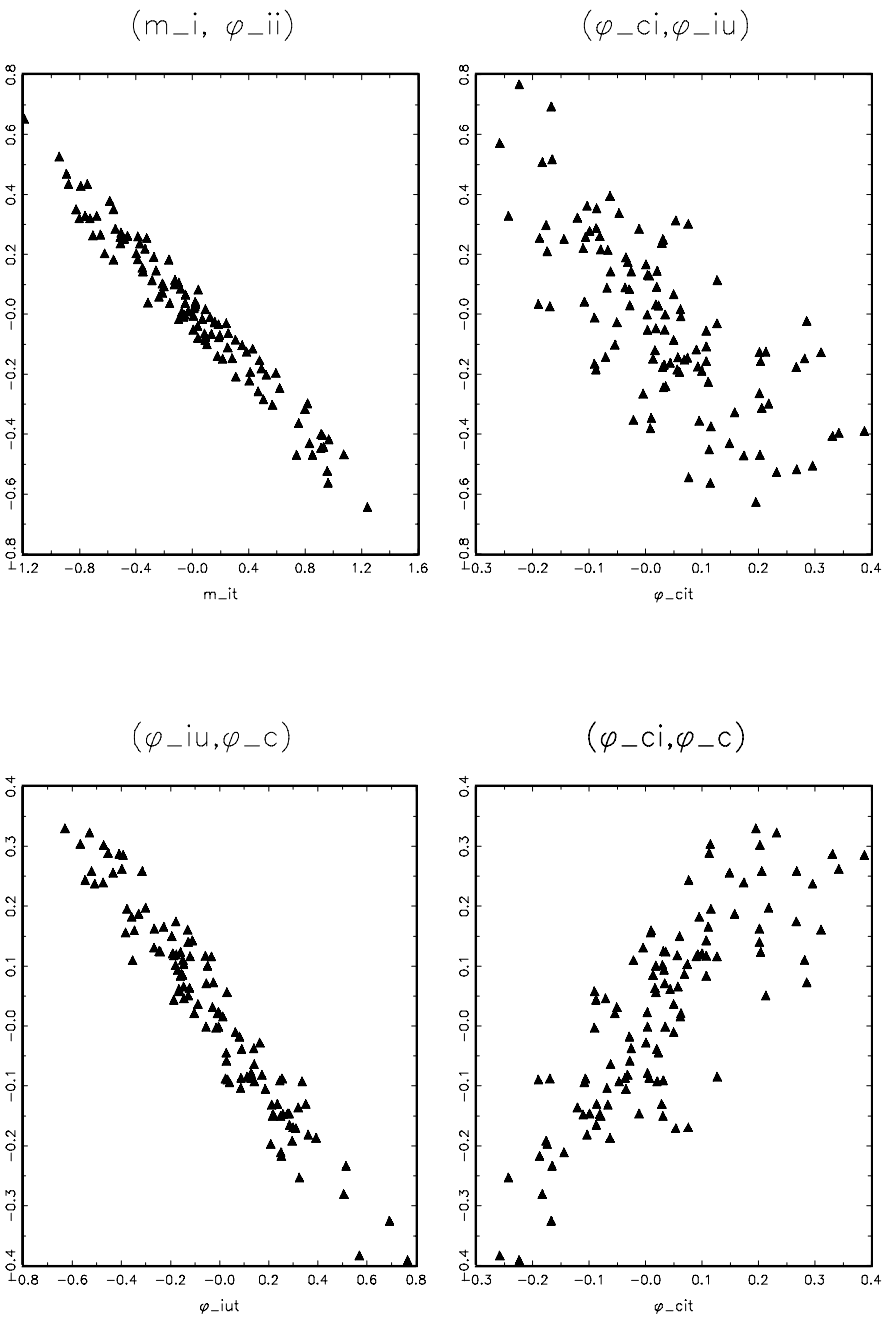
This test statistic is very large in comparison with standard $\chi^2(4)$ percentile. It seems reasonable to think that we reject the invariance property. At least, one of the parameters is clearly significantly different from 0 ($\Lambda_{\phi_{ci},1}$.) In other words, in the past, the hypothesis that a co-movement of the coefficients in the household consumption equation and in the interest rate setting rule cannot be rejected. We *interpret* this property in the following sense: changes in the authorities' preferences that implied changes in their interest rate setting rule may have induced changes in the household consumption decision process.

The change in the variance of the interest rate shocks $(\xi_t)_t$ is given by $\widehat{\nu_1} = 0.84$ (t statistics equal to 4.61) and $\widehat{\nu_2} = 0.33$ (t statistics equal to 5.07). It decreases during the period, which is consistent with what can be observed on the inflation series. Various reasons can be given to explain this result, two of them seems to be worth mentioning. First, it may correspond to the fact that Banque de France implemented during this time period more and more elaborate strategies playing with various interest rates and terms. This allows it to be more efficient. Second, during the same period, *Banque de France*'s objectives became more and more publicly known and its actions foreseeable.

Model contributions to the data analysis: We propose to illustrate three aspects of the data analysis obtained through the estimated model. We examine sequentially the relationships between the various random coefficients, the forecast performance of the Random Coefficient VARX model relative to those of a standard VARX model and the existence of a large change in the coefficients during the first part of the nineties.

First, as emphasized earlier, we can plot the relationships between some of the random coefficients. In Figure 1, we show the relationships between $m_{i,t}$, $\phi_{ii,t}$, $\phi_{ci,t}$, $\phi_{iu,t}$, and $\phi_{c,t}$.

FIGURE 1
Plot of Some Random Coefficients



$\phi_{ii,t}$ is related to the speed of adjustment of the interest rate as an implementing tool of the monetary policy. When it increases, it means that the past values of the interest rate play a more important role in the setting rule, the interest rate adjusts more smoothly to its desired value. $m_{i,t}$ affects this desired value of the interest rate. As $m_{i,t}$ and $\phi_{ii,t}$ are negatively correlated, a choice of a more persistent setting rule implies a lesser desired value. Over a given time period, the global average effect results from both parameters that more or less cancel each other out.

$\left(\phi_{iu} + \phi_{iu,t}\right)$ measures the sensitivity of the interest rate setting rule to the changes in the unemployment rate, $\left(\phi_{ci} + \phi_{ci,t}\right)$ the sensitivity of change in consumption decision to the past interest rate and $\left(\phi_c + \phi_{c,t}\right)$ to the past saving decision.[6] The random coefficients are all centered around zero so that the average effects of $\left(\phi_{iu} + \phi_{iu,t}\right)$, $\left(\phi_{ci} + \phi_{ci,t}\right)$ and $\left(\phi_c + \phi_{c,t}\right)$ are given by the constants: ϕ_{iu}, ϕ_{ci}, and ϕ_c. The estimates of ϕ_{iu}, ϕ_{ci}, and ϕ_c are respectively 0.02 (t-statistic 0.02), 0.13 (t-statistic 0.43) and -3.24 (t-statistic -12.73). When $\phi_{iu,t}$ is negative (or equivalently $\phi_{iu} + \phi_{iu,t} < 0$ since $\phi_{iu} \simeq 0$), the interest rate is *ceteris paribus* set at a lower value when unemployment increases. As $\phi_{iu,t}$ and $\phi_{c,t}$ are negatively correlated (Figure 1), $\phi_{c,t}$ will be positive and $\phi_{c,t} + \phi_c$ will be smaller in absolute value (remember that $\phi_{c,t} + \phi_c < 0$ and $(c_t - y_t) < 0$). The first conclusion we draw is that $\phi_{iu,t}$ negative should result in a lower consumption increase. This conclusion is negated by an opposite effect due to the fact that $\phi_{iu,t}$ and $\phi_{ci,t}$ are negatively correlated. When $\phi_{iu,t}$ is negative, $\phi_{ci,t}$ is positive. This implies, from Model (10), a larger increase in current consumption for a given interest rate. Reconciling these features may be done in the following way. *Ceteris paribus*, active monetary policy that takes into account changes in unemployment is implemented by gradually reducing the interest rate. It is supposed to reduce the probability of being unemployed. Households will increase their consumption for two reasons: (a) they are less likely to become unemployed and (b) borrowing is cheaper.

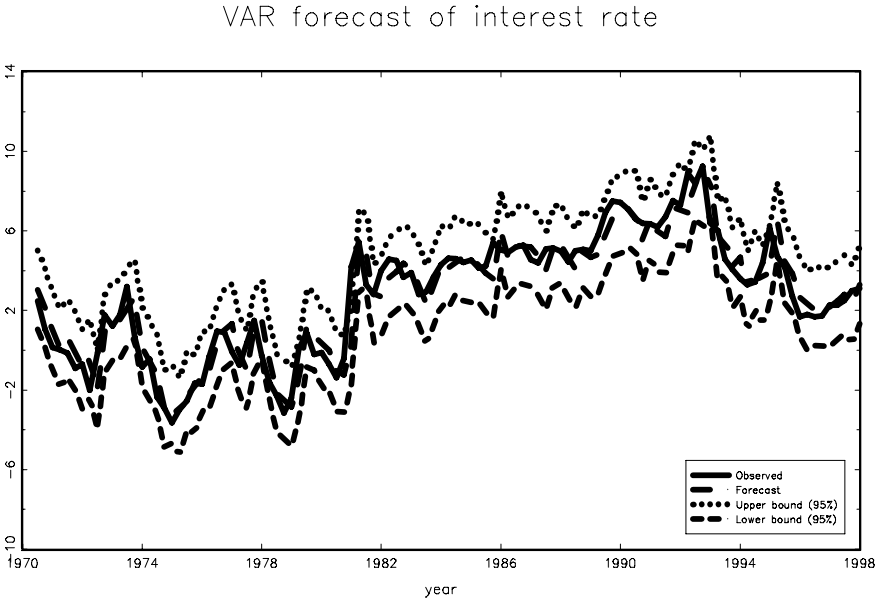
Second, to illustrate the performance of the model introduced in this section, we proceed to the estimation of an unconstrained VARX (p,l) model without random coefficients, of the type:

$$\begin{aligned} Y_t =& C + A_1 Y_{t-1} + A_2 Y_{t-2} + ... + A_p Y_{t-p} + b_0 \Delta e_t \\ &+ b_1 \Delta e_{t-1} + ... + b_l \Delta e_{t-l} + \varepsilon_t \end{aligned}$$

where Δe_t is the growth rate of the exchange rate (considered exogenous) and $Y_t = (i_t, \Delta u_t, \pi_t, \Delta c_t, \triangle y_t)'$. First we discuss the choice of l and p. By doing successive LR test, we find evidence in favor of $l = 0$. To determine the number of lags, p, in the VARX model, we perform several tests. The AIC criterion gives $p = 2$ and BIC criterion gives $p = 1$. We test H_0: $A_p = 0$ using a LR test statistic, this test suggests $p = 2$. Finally, we perform a

6. Savings evolve negatively with the variables introduced in the equation $(c_t - y_t)$.

FIGURE 2
In-the-sample Forecast of Real Interest Rates



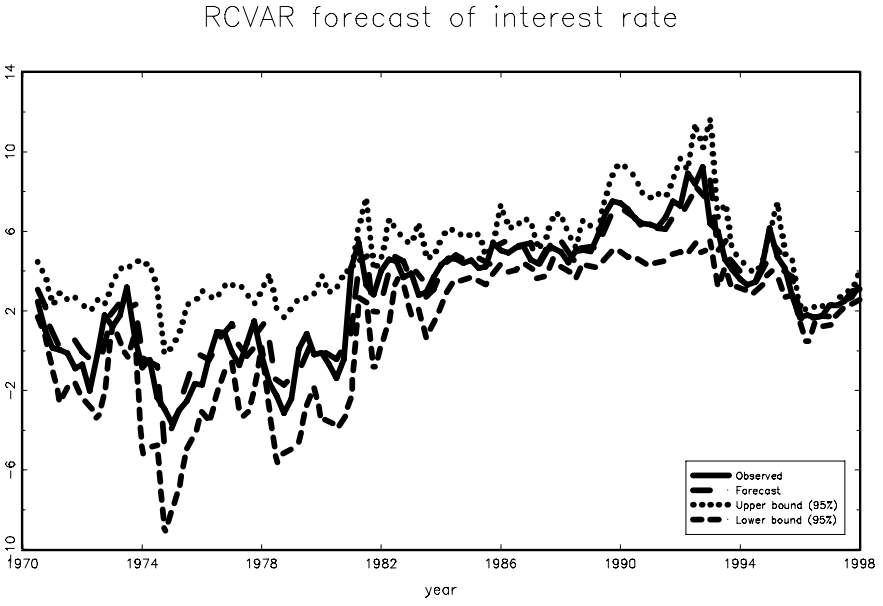
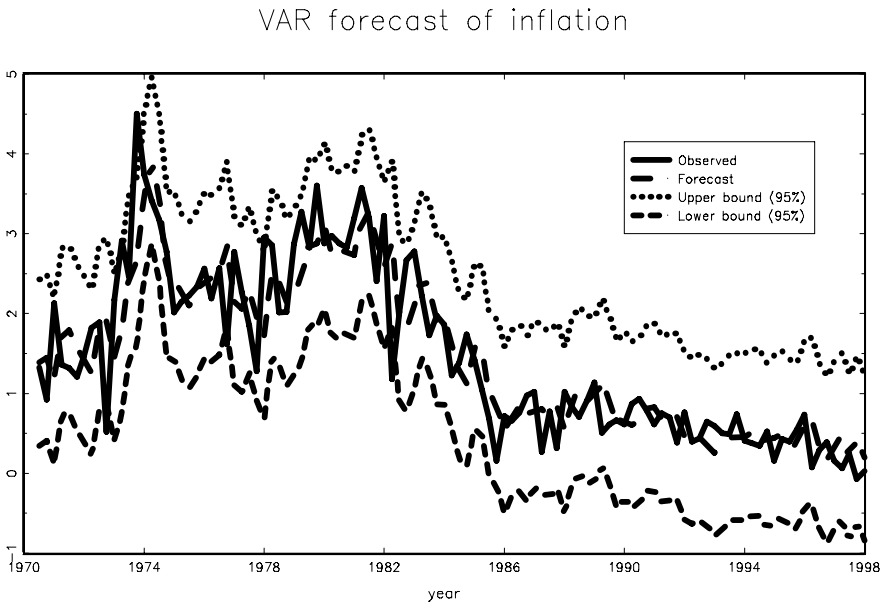


FIGURE 3
In-the-sample Forecast of Inflation



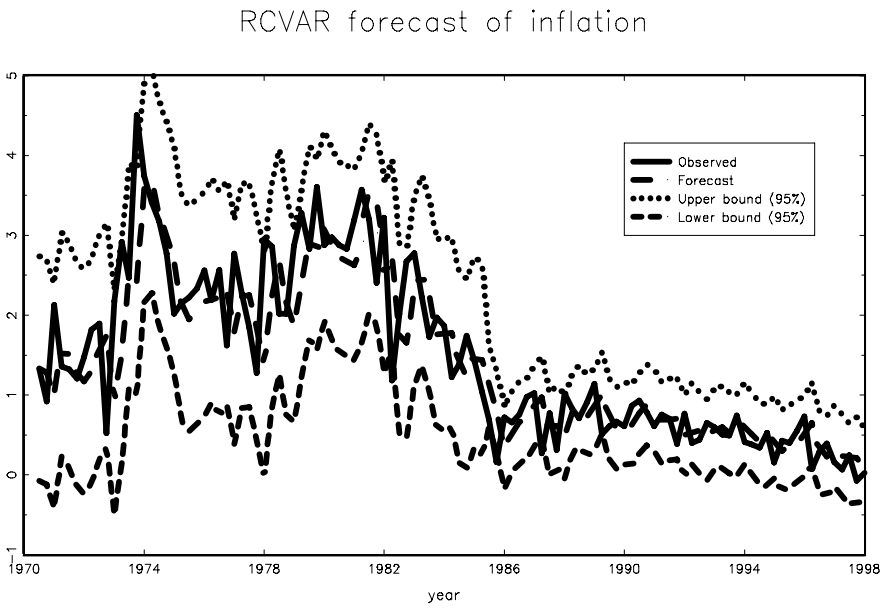
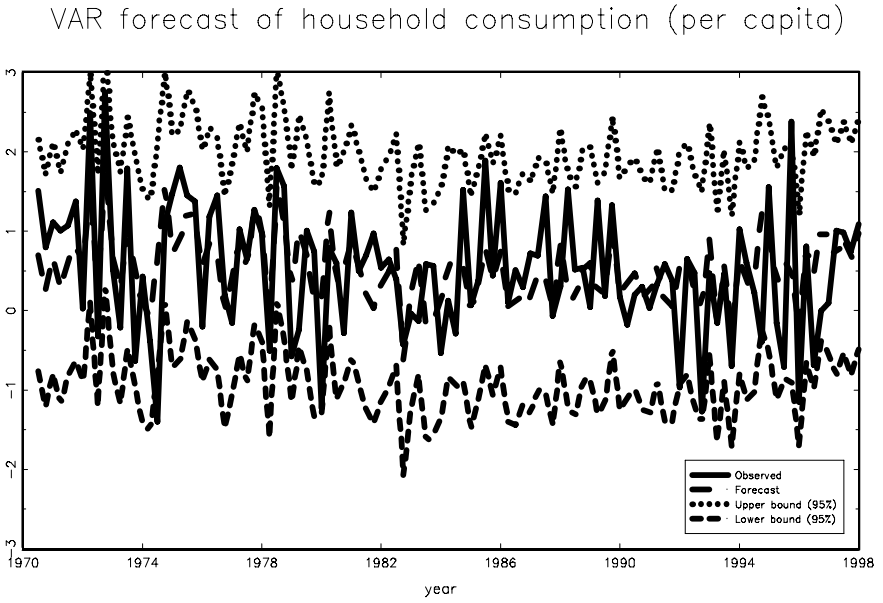


FIGURE 4
In-the-sample Forecast of Change in Household Consumption



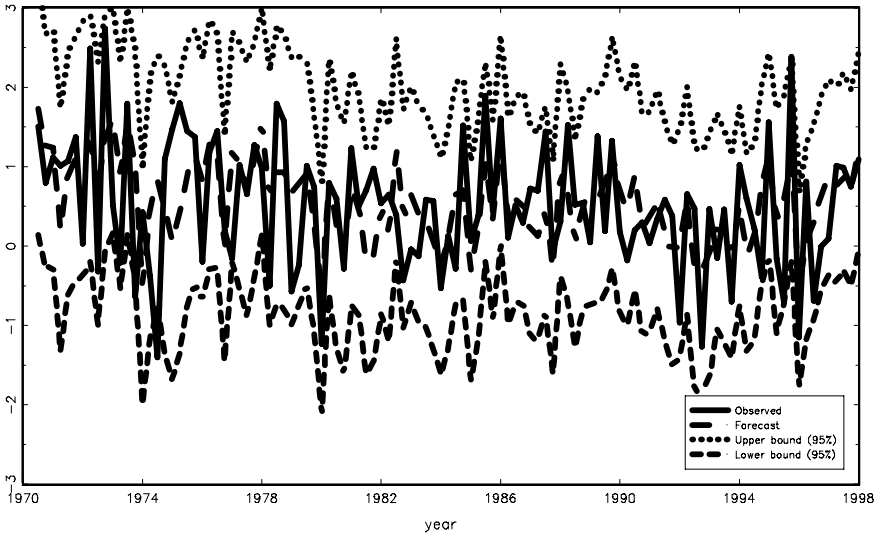


FIGURE 5
RCVAR and VAR Forecast Conditional Standard Errors

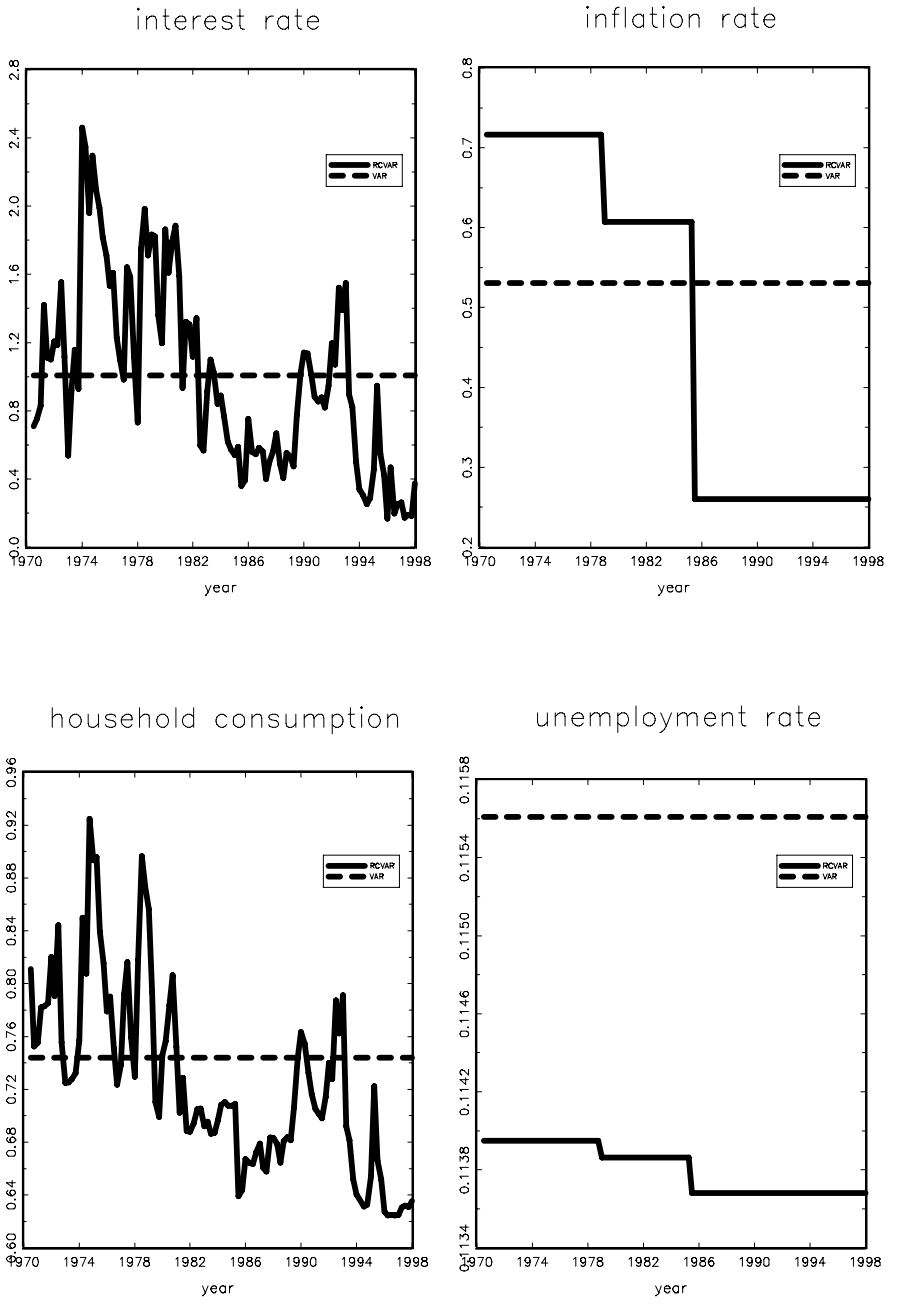
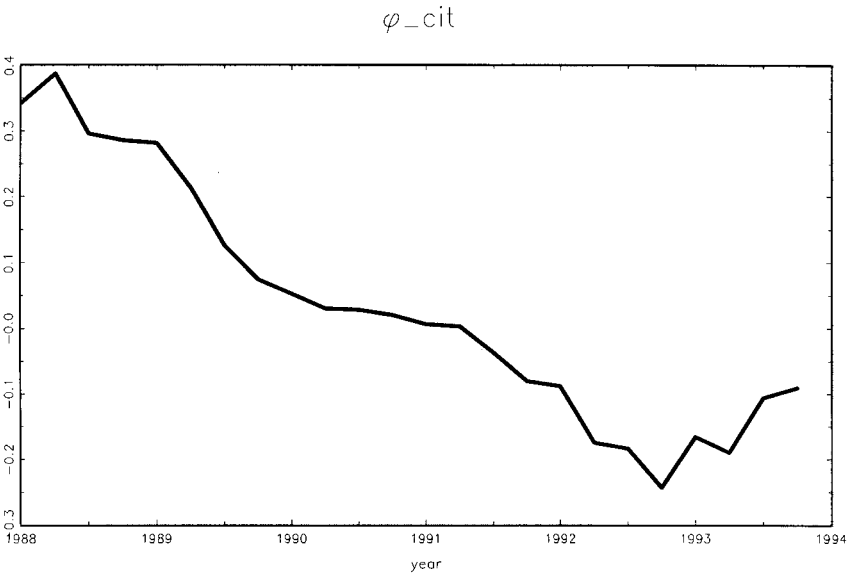
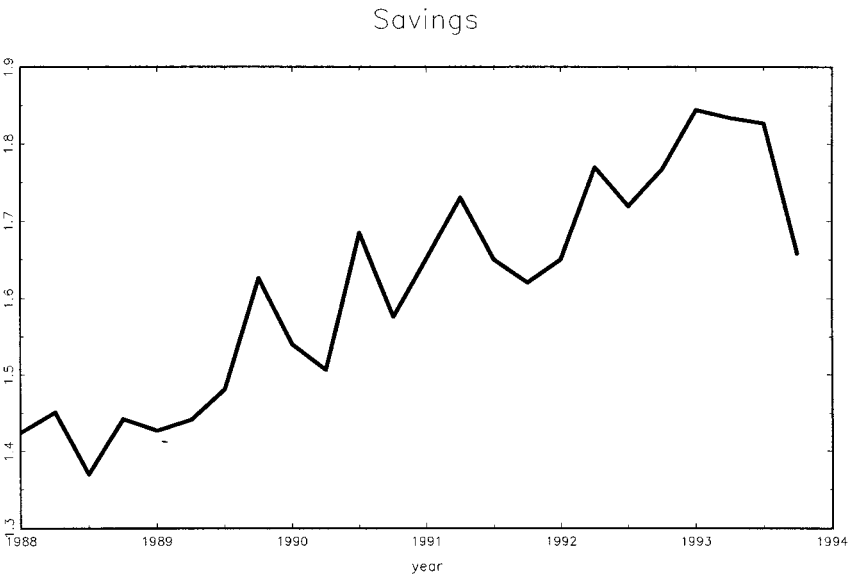


FIGURE 6
$\phi_{ci,t}$ and Saving between 1988 and 1994





Portmanteau test, Q, to test whether the first h residual autocorrelations are zero, for up to 24 lags (see LÜTKEPOHL [1993], Formula (4.4.23)). This test casts some doubt on the adequacy of the model for $p = 2$ as the null is accepted at a 1% level up to 20 lags, (it accepts the null hypothesis of the residual whiteness for $p = 4$ at any reasonable level). The conclusions of these different tests are mixed. As we prefer a parsimonious model, we select $p = 2$. On the whole, the number of parameters in the VARX model is equal to that of the RCVAR model. This allows for a direct comparison of their relative performances. We plot in-the-sample forecasts for three variables : interest rate, household consumption and inflation. For both models, we only used past available information.[7] These plots are reported in Figures 2, 3 and 4. They show that heteroskedasticity effects captured by the random coefficient model for the various equations describe appropriately the fluctuations in its variance and allow for the computation of relevant confidence interval that are not provided by the VAR model. Forecasts appear more accurate into RCVAR than in the VAR.

We plot, in Figure 5, the conditional standard errors for four variables under study. A computation of the mean conditional standard errors show that the Random Coefficient VARX model dominates in terms of forecast accuracy the standard VARX (for a similar number of degrees of freedom). Indeed, we get that average conditional standard errors are respectively 2%, 2% and 3% smaller in the case of the RCVARX than in the case of the VARX for the interest rate, inflation and household consumption. Moreover, for unemployment, using RCVAR represents a substantial benefit (10%). Note that only interest rate and consumption have random coefficients in their equations, this explains why standard errors are much more variable for these two series.

Finally, we address the question of a possible break in the household consumption equation in the late eighties. Figure 6 presents the value of the interest rate coefficient $\phi_{ci,t}$ in the household consumption equation between 1988 and 1994. From 1988 on, $\phi_{ci,t}$ decreases dramatically. From Model (10), we know that this implies a decline in the growth rate of consumption for a given value of the interest rate. This is consistent with the data. Despite the fact that the growth rate of income was not yet decreasing, savings – that is $-(c_t - y_t)$ – increase during the late eighties and the beginning of the nineties (see Figure 6). It certainly results from a complex mechanism. One reason, but not the only one, is that households may have anticipated the changes in the monetary authorities' preferences that would result from the independence of the Central Bank as controlling inflation was to become its main goal. Note that from Figure 1, $\phi_{ci,t}$ negative corresponds to $\phi_{iu,t}$ positive in the interest rate setting rule. Hence, in the nineties, increasing unemployment is no longer associated with smaller interest rate as it would be if fighting unemployment were the priority.

7. From the *Kalman* filter, we use only filtered quantities and no smoothed ones.

5 Conclusion

In this article, we propose to use random coefficient autoregressive models to test for the immunity of a descriptive behavioral equation to the *Lucas Critique*. We assume that coefficients in the equations of variables under the (partial) control of the authorities and those in the equation of the variables chosen by the private agents may co-move. We introduce test procedures to analyze the presence and features of this co-movement and hence to assess the empirical relevance of the critique. We applied these tests to study the immunity of the French household consumption equation to changes in the monetary policy. Invariance is rejected and *Lucas Critique* appears to be relevant. Our approach has several advantages. It permits to quantify the consequences of a change in the monetary policy equation on the coefficients of the behavioral equation. It thus gives a framework to do simulations of policy scenarios as well as forecasting as long as policy shocks are homogeneous over time. Our approach is essentially descriptive. It is therefore important to be able to identify and interpret the past changes in policy in order to produce informative simulations.

From a statistical point of view, testing the constancy of parameters in a Random Coefficient model is not straightforward because some of the standard assumptions are not satisfied. We use recent results by ANDREWS [1999, 2001] to show that the LR statistic is asymptotically nuisance parameter free and to provide its critical values. A small Monte Carlo study shows that this statistic tends to over-reject in small samples and therefore that empirical critical values are preferable to asymptotic ones. ▼

• References

ANDREWS D. (1999). – « Estimation when a Parameter Is on the Boundary », *Econometrica*, 67, p. 1341-1383.

ANDREWS D. (2001) – « Testing when a Parameter Is on the Boundary of the Maintained Hypothesis », *Econometrica*, 69, p. 683-734.

ATTANASIO O.R., WEBER G. (1993). – « Consumption Growth, the Interest Rate and Aggregation », *Review of Economic Studies*, 60, p. 631-649.

ATTANASIO O.R., BANKS J., MEGHIR C., WEBER G. (1995). – « Humps and Bumps in Lifetime Consumption », *NBER Working Paper* n° 5350.

BLOCH S., MAUREL F. (1991). – « Consommation-revenu permanent : un regard d'économètre », *Économie et Prévision*, 99, p. 113-144.

BONNET X., DUBOIS E. (1995). – « Peut-on comprendre la hausse imprévue du taux d'épargne des ménages depuis 1990 ? », *Économie et Prévision*, 121, p. 39-58.

CARROLL C. (1994). – « How does Future Income Affect Current Consumption », *Quarterly Journal of Economic*, vol. 109, n° 1, p. 111-147.

CARROLL C. (1997). – « Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis », *Quarterly Journal of Economic*, vol. 112, n° 1, p. 1-55

CARROLL C., KIMBALL M. (1996). – « On the Concavity of the Consumption Function », *Econometrica*, 64, p. 981-992.

CHOW G. (1984). – « Random and Changing Coefficient Models », in *Handbook of Econometrics*, Eds Z. Griliches and M. Intriligator, Elsevier Science.

CLARIDA R., GALI J., GERTLER M. (1998). – « Monetary Policy Rules in Practice: Some International Evidence », *European Economic Review*, 42, p. 1033-1067.

CLARIDA R., GALI J., GERTLER M. (1999). – « The Science of Monetary Policy: A New Keynesian Perspective », *Journal of Economic Literature*, 37, p. 1661-1707.

COOLEY T.F., LEROY S.F. (1985). – « A theoretical Macroeconometrics: A Critique », *Journal of Monetary Economics*, 16, p. 283-308.

COOLEY T.F., LEROY S.F., RAYMON N. (1984). – « Econometric Policy Evaluation: Note », *American Economic Review*, vol 74, (3), p. 467-470.

COOLEY T.F., PRESCOTT E. (1973). – « An Adaptive Regression Model », *Review of Economics and Statistics*, 73, p. 248-256.

DEATON A. (1991). – « Saving and Liquidity Constraints », *Econometrica*, 59, p. 1221-1248.

DOAN T., LITTERMAN R., SIMS C. (1984). – « Forecasting and Conditional Projection using Realistic Prior Distribution », *Econometric Reviews*, 3, p. 1-100.

ENGLE R.F., HENDRY D.F. (1993). – « Testing Superexogeneity and Invariance in Regression Models », *Journal of Econometrics*, 56 (1/2), p. 119-139.

ENGLE R.F., HENDRY D.F., RICHARD J.F. (1983). – « Exogeneity », *Econometrica*, 51, p. 277-304.

ERICSSON N.R., IRONS J.S. (1995). – « The *Lucas Critique* in Practice: Theory Without Measurement », in *Macroeconometrics: Developments, Tensions and Prospects*, K.D. Hoover ed., Kluwer Academic, p. 263-312.

FAVERO C., HENDRY D.F. (1992). – « Testing the *Lucas Critique*: A Review », *Econometric Reviews*, 11(3), p. 265-306.

FLORENS, J.-P., MOUCHART M. (1985). – « Conditioning in Dynamic Models », *Journal of Time Series Analysis*, Vol. 6, p. 15-34.

GOODFRIEND M., KING M.G. (1997).– « The New Neoclassical Synthesis and the Role of Monetary Policy », 1997 *NBER Macroeconomics Annual*.

GOURIEROUX C., HOLLY A., MONFORT A. (1982). – « The Likelihood Ratio Test, Wald Test and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters », *Econometrica*, Vol. 50, p. 63-80.

GOURIEROUX C., MONFORT A. (1989). – « Statistique et Modèles Économétriques », Vol. 2, *Economica*.

HAMILTON J. (1989). – « A New Approach to Economic Analysis of Nonstationary Time Series and the Business Cycle », *Econometrica*, Vol. 57, p. 357-384.

HARVEY A. (1989). – *Forecasting Structural Time Series Models and the Kalman Filter*, Cambridge University Press.

HARVILLE D. (1989). – *Matrix Algebra from a Statistician's Perspective*, Springer Verlag, New York.

HENDRY D.F. (1988). – « The Encompassing Implications of Feedback versus Feedforward Mechanisms in Econometrics », *Oxford Economic Papers*, 40, p. 132-149.

LEAMER E.E. (1985). – « Vector Autoregressions for Causal Inference? », in K. Brunner and A.H.Meltzer eds, *Understanding Monetary Regimes*, Carnegie-Rochester Conference Series on Public Policy, 22, Amsterdam North-Holland.

LUCAS R.E. Jr. (1976). – « Econometric Policy Evaluation: A Critique », in K.Brunner and A.Meltzer, eds The Phillips Curve and Labor Markets, vol 1, *Carnegie-Rochester Series on Public Policy*, *Journal of Monetary Economics*, suppl. 19-46.

MACNELIS P.D., NEFTÇI S.N. (1982). – « Policy-Dependent Parameters in the Presence of Optimal Learning: An application of Kalman Filtering to the Fair and Sargent Supply-side Equations », *Review of Economics and Statistics*, 64 (2), p. 296-306

MARCET A., SARGENT T. (1989). – « Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models », *Journal of Economic Theory*, 48, p. 337-368.

MEYN S.P., GUO L. (1993). – « Geometric Ergodicity of a Doubly Stochastic Time Series Model », *Journal of Time Series Analysis*, 14 (1), p. 93-108.

POURAHMADI M. (1986). – « On Stationarity of the Solution of a Doubly Stochastic Model », *Journal of Time Series Analysis*, 7 (2), p. 123-131

POURAHMADI M. (1988). – « Stationarity of the Solution of $X_t = A_t X_{t-1} + \varepsilon_t$ and Analysis of Non-Gaussian Dependent Random Variables », *Journal of Time Series Analysis*, 9 (3), p. 225-239

PSARADAKIS Z., SOLA M. (1996). – « On the Power of Tests for Superexogeneity and Structural Invariance », *Journal of Econometrics*, 72, p. 151-175.

SARGENT T.J. (1984). – « Autoregressions, Expectations and Advice », *Journal of Political Economy*, 74, p. 408-415.

SIMS C.A. (1982). – « Policy Analysis with Economic Models », *Brookings Papers on Economic Activity*, p. 107-164.

SIMS C.A. (1987). – « A Rational Expectations Framework for Short-Run Policy Analysis », in W.A. Barnett and K.J. Singleton eds. *New Approaches to Monetary Economics*, Cambridge University Press, p. 293-308.

TERASVIRTA T. (1994). – « Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models », *Journal of the American Statistical Association*, 89, p. 208-218.

TJOSTHEIM D. (1986). – « Some Doubly Stochastic Time Series Models », *Journal of Time Series Analysis*, 7 (1), p. 51-72.

TONG, H. (1990). – *Nonlinear Time Series*, Oxford University Press, Oxford.

WALSH C. (1998). – *Monetary Theory and policy*, MIT Press, Cambridge.

WIELAND V. (1998). – « Monetary Policy and Uncertainty about the Natural Unemployment Rate », Board of Governors of the Federal Reserve System, Finance and Economics *Discussion Paper* Series 98/22.

WICKENS M.R., MOLANA H. (1984). – « Stochastic Life Cycle Theory with Varying Interest Rates and Prices », *The Economic Journal*, 94, p. 133-147

APPENDIX A

Proof of Propositions 1 and 2

Proof of Proposition 1. We follow closely the approach proposed by ANDREWS [1999, 2001]. We first study the properties of the score and information matrix and then check a set of assumptions necessary to derive an asymptotic result. In our derivation of the time properties for the score function, we consider its strictly stationary counterpart. This partly results from the initial values we take for the expectation and variance of the random coefficients. They are equal to the unconditional mean and variance of the random process under study:

$$E\alpha_t = 0$$
$$vecV\alpha_t = \rho^2 \left[\left(I_{q^2} - T_v \otimes T_v \right)^{-1} vec\Lambda_v \Lambda_v' \right] = \rho^2 vecH_v$$

Under H_0, the score for the observation at date t is given by:

$$\left. \frac{\partial \ln l_t(\theta, v)}{\partial \theta'} \right|_{\rho^2=0} =$$

$$\begin{pmatrix} \left(Y_{t-1} \otimes I_n\right) \Sigma^{-1} \varepsilon_t \\ -\frac{1}{2} G_n' vec\left(\Sigma^{-1} - \Sigma^{-1} \varepsilon_t \varepsilon_t' \Sigma^{-1}\right) \\ -\frac{1}{2} Tr\left(K Z_t H_v Z_t' K' \Sigma^{-1} \left(I_n - \varepsilon_t \varepsilon_t' \Sigma^{-1}\right)\right) + \varepsilon_t' \Sigma^{-1} K Z_t T_v U_{v,t-1} \end{pmatrix}'$$

where $U_{v,t}$ is a process that satisfies the following autoregressive equation:

$$U_{v,t} = T_v U_{v,t-1} + H_v Z_t' K' \Sigma^{-1} \varepsilon_t = T_v U_{v,t-1} + H_v \zeta_t$$

and G_n is the duplication matrix (see HARVILLE [1997], page 352). We notice that in fact only the last component of the score is affected by the nuisance parameters, the first components are the ones we get in the model under the null. Similar computations lead to the following information matrix:

$$-E \frac{\partial^2 \ln l(\theta, v)}{\partial \theta \partial \theta'} = \begin{pmatrix} I_{00} & I_{0\rho^2}(v) \\ I_{\rho^2 0}(v) & I_{\rho^2}(v) \end{pmatrix} = \mathcal{I}(v)$$

where I_{00} is the information matrix of the model under the null that is block diagonal,

$$I_{\rho^2}(v) = \frac{1}{2} E Tr\left(K Z_t H_v Z_t' K' \Sigma^{-1} K Z_t H_v Z_t' K' \Sigma^{-1}\right) + E\left(U_{v,t-1} T_v Z_t' K' \Sigma^{-1} K Z_t T_v U_{v,t-1}\right)$$

and,

$$I_{\rho^2 0}(v) =$$
$$\left(EU_{v,t-1}T_v Z_t' K' \Sigma^{-1} \left(Y_{t-1}' \otimes I_n \right) \quad E vec\left(\Sigma^{-1} K Z_t H_v Z_t' K' \Sigma^{-1} \right)' G_n \right)$$

We now list a set of properties of the problem under study necessary to use Theorem 2 and 3 in ANDREWS [2001].

Constraints on the values of the parameters (Φ, v) imply absence of trending variables in the model under the null, so that Assumption 5 (ANDREWS [2001]) is satisfied. Under H_0, the problem we consider is a standard one and the likelihood function $l_0(\underline{\theta})$ is strictly concave. Let θ_0 be the value of $\underline{\theta}$ at this maximum, Assumption 4_0, 9 and 10 in ANDREWS [2001] are satisfied. On the other hand the likelihood $l(\theta, v)$ is continuously differentiable with respect to all its arguments and $l\left(\underline{\theta}, 0, v\right) = l_0\left(\underline{\theta}\right)$.

The most important property to establish is the functional convergence of the joint process of the score and information matrix as a process indexed by $v \in N$. We add the assumption on N that on this set, the eigenvalues of the information matrix $\mathcal{I}(v)$ are positive, strictly larger than 0 and bounded. Assumption 3* in ANDREWS [2001] then allows us to limit the study to the score function. Indeed, the information matrix is non random, does not depend on T and is symmetric. The two first components of the score are those of standard maximum likelihood model under gaussian assumption which are independent of v. The result holds. Only the last component of the score depends on v and we need to first prove a stochastic equicontinuity to get the required convergence. In our framework with strong distributional assumptions we have the following result:

LEMMA: $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left. \frac{\partial \ln l_t(\theta, v)}{\partial \rho^2} \right|_{\rho^2=0}$ is stochastically equicontinuous in $v \in N$, where N is a compact set.

PROOF: To obtain this result, we prove a *Lipschitz* property of the score in v. We compute:

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left. \frac{\partial \ln l_t(\theta, v_1)}{\partial \rho^2} \right|_{\rho^2=0} - \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left. \frac{\partial \ln l_t(\theta, v_2)}{\partial \rho^2} \right|_{\rho^2=0}$$

and decompose this expression into two quantities:

$$(I) = -\frac{1}{2\sqrt{T}} \sum_{t=1}^{T} Tr\left(K Z_t \left(H_{v_1} - H_{v_2} \right) Z_t' K \Sigma^{-1} \left(I_n - \varepsilon_t \varepsilon_t' \Sigma^{-1} \right) \right)$$

and

$$(II) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_t' \Sigma^{-1} K Z_t \left(T_{v_1} U_{v_1,t-1} - T_{v_2} U_{v_2,t-1} \right)$$

(I) can be rewritten under the form:

$$(I) = \frac{1}{2} Tr\left(\left(H_{v_1} - H_{v_2} \right) \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(\zeta_t \zeta_t' - E_t \left(\zeta_t \zeta_t' \right) \right) \right)$$

with $\zeta_t = Z_t' K' \Sigma^{-1} \varepsilon_t$ that is a gaussian heteroskedastic independent process. Let $\zeta_{2,t}$ denote $\zeta_t \zeta_t' - E_t \left(\zeta_t \zeta_t' \right)$, then ζ_{2t} is a martingale difference sequence, heteroskedastic but such that $E \max_i \left| \zeta_{2,i,t}^4 \right| < +\infty$ (under our distributional assumptions and the absence of trending variables). A central limit theorem applies. We get using *Cauchy-Schwartz* inequality:

$$(I) \leqslant \frac{1}{2} \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(\zeta_t \zeta_t' - E_t \left(\zeta_t \zeta_t' \right) \right) \right\|_2 \left\| H_{v_1} - H_{v_2} \right\|_2$$

where,

$$\left\| H_{v_1} - H_{v_2} \right\|_2^2 = vec \left(H_{v_1} - H_{v_2} \right)' vec \left(H_{v_1} - H_{v_2} \right)$$

but this last expression as by construction $vec H_{v_1} = \left[\left(I_{q^2} - T_{v_1} \otimes T_{v_1} \right)^{-1} vec \Lambda_{v_1} \Lambda_{v_1}' \right]$, is a continuous differentiable function in v_1 whose partial derivatives are continuous and therefore absolutely bounded on the compact set N, so that $\exists \xi_1 > 0$ such that:

$$\left\| H_{v_1} - H_{v_2} \right\|_2 \leqslant \xi_1 \sup_i \left| v_{1,i} - v_{2,i} \right|$$

and thus,

$$(I) \leqslant O_p(1) \sup_i \left| v_{1,i} - v_{2,i} \right|$$

We turn now to (II) we rewrite as follows:

$$(II) = \sum_{h=1}^{T} Tr\left(\widehat{\Gamma}_{\zeta,T}(h) \sqrt{T} \left(T_{v_1}^h H_{v_1} - T_{v_2}^h H_{v_2} \right) \right)$$

where $\widehat{\Gamma}_{\zeta,T}(h) = \frac{1}{T} \sum_{t=h}^{T} \zeta_{t-h} \zeta_t'$. The arguments we propose to use are similar to those used in the previous step. We consider the expressions:

$$\left\| T_{v_1}^h H_{v_1} - T_{v_2}^h H_{v_2} \right\|_2 \leqslant \left\| \left(T_{v_1}^h - T_{v_2}^h \right) H_{v_1} \right\|_2 + \left\| T_{v_2}^h \left(H_{v_1} - H_{v_2} \right) \right\|_2$$

we have:

$$\left\| T_{v_2}^h \left(H_{v_1} - H_{v_2} \right) \right\|_2 \leqslant \left\| T_{v_2}^h \right\|_1 \left\| H_{v_1} - H_{v_2} \right\|_2$$

where,

$$\left\| T_{v_2}^h \right\|_1 = \max \left\{ \sqrt{\lambda} \text{ such that } \lambda \text{ is an eigenvalue of } T_{v_2}^{h\prime} T_{v_2}^h \right\}$$

so that by assumption:

$$\left\| T_{v_2}^h \left(H_{v_1} - H_{v_2} \right) \right\|_2 \leqslant \frac{\xi_1}{(1+\delta)^h} \sup_i \left| v_{1,i} - v_{2,i} \right|$$

On the other hand,

$$T_{v_1}^h - T_{v_2}^h = \sum_{j=1}^{h-1} T_{v_2}^j \left(T_{v_1} - T_{v_2} \right) T_{v_1}^{h-j-1}$$

so that:

$$\begin{aligned}
\left\| \left(T_{v_1}^h - T_{v_2}^h \right) H_{v_1} \right\|_2 &\leqslant \sum_{j=1}^{h-1} \left\| T_{v_2}^j \left(T_{v_1} - T_{v_2} \right) T_{v_1}^{h-j-1} H_{v_1} \right\|_2 \\
&\leqslant \sum_{j=1}^{h-1} \left\| T_{v_2}^j \left(T_{v_1} - T_{v_2} \right) \right\|_2 \left\| T_{v_1}^{h-j-1} H_{v_1} \right\|_2 \\
&\leqslant \sum_{j=1}^{h-1} \left\| T_{v_2}^j \right\|_1 \left\| T_{v_1} - T_{v_2} \right\|_2 \left\| T_{v_1}^{h-j-1} \right\|_1 \left\| H_{v_1} \right\|_2 \\
&\leqslant \frac{(h-1) \left\| H_{v_1} \right\|_2}{(1+\delta)^{h-1}} \left\| T_{v_1} - T_{v_2} \right\|_2
\end{aligned}$$

By continuity of the partial derivatives of the function $\left\| T_{v_1} - T_{v_2} \right\|_2$ with respect to υ and the continuity of $\left\| H_{v_1} \right\|_2$ on the compact set N, we conclude that $\exists \xi_2 > 0$ such that:

$$\left\| \left(T_{v_1}^h - T_{v_2}^h \right) H_{v_1} \right\|_2 \leqslant \frac{(h-1)\,\xi_2}{(1+\delta)^{h-1}} \sup_i \left| v_{1,i} - v_{2,i} \right|$$

and $\exists \left(\xi_3, \xi_4 \right)$ such that:

$$\left\| T_{v_1}^h H_{v_1} - T_{v_2}^h H_{v_2} \right\|_2 \leqslant \frac{(\xi_3 h + \xi_4)}{(1+\delta)^{h-1}} \sup_i \left| v_{1,i} - v_{2,i} \right|$$

We now need to establish that:

$$\sum_{h=1}^{T} \left\| \widehat{\Gamma}_{\zeta,T}(h) \sqrt{T} \right\|_2 \frac{(\xi_3 h + \xi_4)}{(1+\delta)^{h-1}} = O_p(1)$$

but ζ_t is a gaussian heteroskedastic independent process. We split the set of indices $\{1,...T\}$ into two subsets $\left\{1,... \left[cT^a \right] \right\}$ and $\left\{ \left[cT^a \right] + 1,...T \right\}$ with $a < 1$. Under our assumptions, we know that a joint convergence of the esti-

mates $\widehat{\Gamma}_{\zeta,T}(h)$ for h in the first subset holds, this allows us to state that the first partial sum is $\sum_{h=1}^{[cT^a]} \left\| \widehat{\Gamma}_{\zeta,T}(h)\sqrt{T} \right\|_2 \frac{(\xi_3 h+\xi_4)}{(1+\delta)^{h-1}} = O_p(1)$. On the second subset, we use the fact that ζ_t are $O_p(1)$, their product is $O_p(1)$, we consider sum of $O_p(1)$ random variables weighted by absolutely summable constants due to the fact that $\frac{1}{1+\delta} < 1$, whence the result.

Under the regularity conditions listed above, Theorem 4(c) in ANDREWS [2001] says that:

$$\sup_{v\in N}\left(\text{lnl}\left(\widehat{\theta}_v, v\right) - \text{lnl}\left(\hat{\theta}_0, 0, v\right)\right) \Longrightarrow \frac{1}{2}\sup_{v\in N}\left(\left(\max\left\{Z_{\rho^2,v}, 0\right\}\right)^2\left(e'\mathcal{I}_v^{-1}e\right)^{-1}\right)$$

where $Z_{\rho^2,v}$ is the asymptotic distribution of the last component of the score and e is a vector whose all components are equal to 0 except for the last one which is equal to 1. Because the distribution in parentheses does not depend on v under the null, we have the standard result that:

$$2\left\{\text{lnl}\left(\widehat{\theta}\right) - \text{lnl}\left(\hat{\theta}_0, 0, v\right)\right\} \Longrightarrow \frac{1}{2}\chi^2(0) + \frac{1}{2}\chi^2(1)$$

so that the likelihood ratio test follows the asymptotic distribution stated in Proposition 1.

Proof of Proposition 2. We now turn to the situation in which dummy variables are present in the Coefficient generating process (Model (B)). For sake of simplicity, we limit our attention to the case of a unique dummy variable: $\mu_t = \mu \mathbf{1}_{t>[\tau_0 T]}$ but as said above the additional parameter τ_0 is given a priori and therefore is not included into the nuisance parameter set.

Under H_0, the score for the observation at date t is given by:

$$\left.\frac{\partial \text{lnl}_t(\theta, v)}{\partial \theta'}\right|_{\rho^2=0,\mu=0}$$
$$= \begin{pmatrix} \left(Y_{t-1}\otimes I_n\right)\Sigma^{-1}\varepsilon_t \\ -\frac{1}{2}G_n' vec\left(\Sigma^{-1} - \Sigma^{-1}\varepsilon_t\varepsilon_t'\Sigma^{-1}\right) \\ \Lambda_v' T_v^{t-[\tau_0 T]} Z_t' K'\Sigma^{-1}\varepsilon_t \mathbf{1}_{t>[\tau_0 T]} \\ -\frac{1}{2}Tr\left(K Z_t H_v Z_t' K'\Sigma^{-1}\left(I_n - \varepsilon_t\varepsilon_t'\Sigma^{-1}\right)\right) + \varepsilon_t'\Sigma^{-1}K Z_t T_v U_{v,t-1} \end{pmatrix}'$$

and the information matrix by:

$$-E\frac{\partial^2 \text{lnl}(\theta, v)}{\partial\theta\partial\theta'} = \begin{pmatrix} I_{00} & I_0(v)' \\ I_0(v) & I(v) \end{pmatrix} = \mathcal{I}(v)$$

where I_{00} is the information matrix of the model under the null that is block diagonal,

$$I_0(v) =$$
$$\begin{pmatrix} E\sum_{t>[\tau_0 T]} \Lambda_v' T_v^{t-[\tau_0 T]} Z_t' K' \Sigma^{-1} \left(Y_{t-1}' \otimes I_n\right) & 0 \\ E U_{v,t-1}' T_v Z_t' K' \Sigma^{-1} \left(Y_{t-1}' \otimes I_n\right) & E vec\left(\Sigma^{-1} K Z_t H_v Z_t' K' \Sigma^{-1}\right)' G_n \end{pmatrix}$$

and,

$$I(v) = \begin{pmatrix} I_\mu(v) & I_{\mu,\rho^2}(v) \\ I_{\rho^2,\mu}(v) & I_{\rho^2(v)} \end{pmatrix}$$

where:

$$I_\mu(v) = E \sum_{t>[\tau_0 T]} \Lambda_v' T_v^{t-[\tau_0 T]} Z_t' K' \Sigma^{-1} K Z_t T_v^{t-[\tau_0 T]} \Lambda_v$$

and,

$$I_{\mu,\rho^2}(v) = E \sum_{t>[\tau_0 T]} \Lambda_v' T_v^{t-[\tau_0 T]} Z_t' K' \Sigma^{-1} K Z_t T_v U_{v,t-1},$$

We again have to establish the stochastic equicontinuity of the score function. Only the component related to μ must be studied. We simply have to notice that it is equal to $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \Lambda_v' T_v^{t-[\tau_0 T]} \mathbf{1}_{t>[\tau_0 T]} \zeta_t$ with the above notations so that the difference between the score for the nuisance parameter v_1 and v_2 is equal to:

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(\Lambda_{v_1}' T_{v_1}^{t-[\tau_0 T]} \mathbf{1}_{t>[\tau_0 T]} - \Lambda_{v_2}' T_{v_2}^{t-[\tau_0 T]} \mathbf{1}_{t>[\tau_0 T]} \right) \zeta_t$$

and the stochastic equicontinuity follows with similar arguments to those used in the preceding proof once we set $v \in N$ a compact set.

Let $\delta = \left(\underline{\theta}, \mu\right)$ be the set of all parameters except for ρ^2. Let $G_{\delta v}$ be the limit in distribution (under H_0) of the score with respect to (wrt) δ in Model B. Let G_δ be the limit in distribution (under H_0) of the score wrt δ in the restricted model where $\rho^2 = \mu = 0$. Similarly, let $\mathcal{I}_{\delta v}$ and $\mathcal{I}_\delta$ be the block of the information matrix corresponding to δ. From Theorems 2 and 3 of ANDREWS [2001], the LR statistic converges to:

$$\begin{aligned} 2 \sup_{v \in N} \left(\ln l\left(\widehat{\theta_v}, v\right) - \ln l\left(\hat{\theta}_0, 0, v\right) \right) & \\ \Longrightarrow \sup_{v \in N} & \left(\left(\max\left\{ Z_{\rho^2,v}, 0 \right\} \right)^2 \left(e' \mathcal{I}_v^{-1} e \right)^{-1} \right) \\ + \sup_{v \in N} & G_{\delta v} \mathcal{I}_{\delta v}^{-1} G_{\delta v} - G_\delta \mathcal{I}_\delta^{-1} G_\delta \\ = & \frac{1}{2}\chi^2(0) + \frac{1}{2}\chi^2(1) + \chi^2(\dim \mu) \end{aligned}$$

To prove that $G_{\delta v} \mathcal{I}_{\delta v}^{-1} G_{\delta v}$ converges to a chi-square and is therefore nuisance parameter free, one needs to use the CLT of martingale difference sequences. The result of Proposition 2 follows.