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

Assessing The Effects of Policy Changes: Lessons from the European 1992 Experience

Pierre-Yves HÉNIN, Marie PODEVIN *

ABSTRACT. – The paper focuses on the experience of the EMS breakdown in September 1992 as a 'crucial experiment', allowing to identify subsequent changes in policy rules and to assess the invariance of private agents' behaviour to these changes. After a brief summary of the facts, the breakdown and the aftermath, we provide estimates of the change in monetary and budget policy rules for Germany (the benchmark), two 'stayers' (France and the Netherlands) and three 'leavers' (Italy, Spain and the UK). Finally, we perform tests of superexogeneity and invariance of reduced form equations for real activity, inflation and the term structure of interest rates.

Évaluer les effets de changements dans les politiques économiques : leçons de l'expérience de l'éclatement du SME en 1992

RÉSUMÉ. – L'article s'attache à l'épisode d'éclatement du SME en Septembre 1992 comme « expérience cruciale » permettant d'identifier des changements subséquents dans les politiques macroéconomiques et d'évaluer l'invariance du comportement des agents privés en présence de ces changements. Nous estimons les changements intervenus dans les politiques monétaires et budgétaires en Allemagne, France et Pays-Bas d'une part (les « stayers»), en Italie, Espagne et Royaume-Uni d'autre part (les « leavers »). Nous effectuons ensuite des tests d'invariance et de superexogénéité d'équations d'activité, d'inflation et de la structure par terme des taux d'intérêt.

---

* P.Y. HÉNIN: Université de Paris I – EUREQua et CEPREMAP; M. PODEVIN: Université de Paris I – EUREQua.
Useful remarks by two referees, one editor of this special issue and the participants at the ADRES-CEPREMAP-CREST Conference on "The Econometrics of Policy Evaluation" – especially our discussant K. WALLIS and S. GREGOIR – are gratefully acknowledged, without implication on the remaining shortcomings of this paper.

1 Introduction

The econometric evaluation of fiscal and monetary policies was first associated with the building and the use of large macroeconomic models. The LUCAS [1976] critique was in fact directed towards invalid uses of those models for policy evaluation. A second step in the literature turned to reduced form models and especially VAR specifications, thus focusing on the identification of policy innovations and the assessment of their long run persistence.[1] A third step emphasized the importance of the interventions rather than the 'message in the innovations' (SCHWERT [1979]), therefore focusing on issues of structural stability and invariance. This research program allowed for an evaluation of the empirical relevance of the *Lucas critique* (FAVERO and HENDRY [1992]) using, *inter-alia*, tests for superexogeneity and invariance (ENGLE and HENDRY [1993]). The evidence from these tests hardly plays in favour of that relevance. Nevertheless, the low power of invariance tests suggests to focus on major episodes of policy changes, which play the role of quasi 'natural experiments'. The change in the US monetary policy initiated by *P. Volcker* provides such an opportunity, exploited by many authors since BLANCHARD [1984] (see especially MILLER and ROBERDS [1991]). However, the designation of 'natural experiments' for episodes of macro policy changes may be disputed, as such episodes provide only poor substitutes to the distinction of a 'control group' and 'comparison groups' which are used in micro econometric policy evaluation.[2]

The European experience in the nineties may provide such a quasi-experimental situation. The break-up of the EMS in 1992 has the further advantage that we may observe different situations with respect to a common institutional shock. To some extent, the existence of 'stayers', *ie*, countries maintaining their participation to the EMS, provides the equivalent of a 'comparison group' as required in a natural experiment. This paper proposes to use this experience as an experiment to assess the effects of policy changes. We first document and test changes in monetary and fiscal policy rules in a sample group of six European countries –'leavers' or 'stayers' in the EMS. Second, we test for weak exogeneity and invariance in behavioural rules assessing the structural stability of reduced forms for output, inflation and the term structure equations. Our results weakly support the empirical relevance of the *Lucas critique*.

In order to introduce the issue, let us briefly review the main features of the 1992 EMS break-up. The problems in the EMS had arisen since 1990 from the pressures made by the unification on the German fiscal and monetary policies. Opportunities for resolving the resulting tensions by realignment,

1. VAR models basically reflecting reduced forms are not immune from the Lucas critique, contrary to small dynamic general equilibrium models specified in terms of deep parameters describing technology and preferences. See *eg*, LINDE [2000] for an evaluation of monetary rules free of the *Lucas critique* and COLLARD, FEVE and LANGOT [2002] for an empirical illustration of the issue using such a small dynamic general equilibrium model.
2. See the JBES symposium, and especially B. MEYER's [1995] contribution on the use of natural and quasi-experiments in program or policy evaluation.

including a DM appreciation, were missed and the Bundesbank resisted its partners' pressures to cut interest rates. The deteriorated perspectives of ratification of the Maastricht Treaty following the negative issue of the Danish referendum, resulted in an open crisis in September 1992, with the 'temporary' suspension of the participation of the sterling and the lira and the subsequent large devaluations of the central parity of the Spanish peseta.[3] After 10 months of turbulence, EMS members agreed on wider bands. This opened the way to stabilization and ultimately to the convergence towards the euro.

The aftermath of the 1992 break-up provides an interesting opportunity to compare the behaviour and the performances of the leavers, Italy, the UK and *de facto* Spain, *vis-à-vis* the stayers, like Germany, France or the Netherlands. GORDON's [1999] analysis does not support the 'conventional wisdom' that leavers enjoyed a free lunch.[4] Only the UK seems to have benefited from a significant windfall from the break-up. A deeper examination of the 1992 aftermath is beyond our scope and we will rather focus on the characterization of this episode in terms of policy changes, and their consequences on private behaviour and economic performance.

The paper is organized as follows. Section 2 is devoted to the estimation of policy rules for six countries, with a particular attention paid to their stability after the break-up. Section 3 introduces and implements tests for invariance of the output gap, inflation and the term structure equations and Section 4 concludes.

2 Leavers *vs* Stayers: Have their Monetary and Fiscal Policy Changed?

A priori, exiting from an exchange rate agreement reintroduces a room for autonomous national monetary policy. There is therefore a presumption that the rule describing monetary policies in the leaver countries could have changed after the break-up. We expect *a priori* more stability in the policy rule for the stayers. Standard open macroeconomics also predicts that, under high capital mobility, exchange rate flexibility would render fiscal policy less effective. In these circumstances, it may have been more efficient for the leavers to switch to more neutral, or passive, budget policies. Although the incentives for the stayers to change their budget policy were weaker, they were prone to look for fiscal consolidation, as a way to strenghten the credibility of their stayer status.

In the following formal and quantitative exercise, we are going to limit our attention to a sample group of EMS countries. As the leavers, we consider

3. On the EMS crisis, see EICHENGREEN and WYPLOSZ [1993], GROSS and THYGESEN [1998] or ROSE and SVENSSON [1994].

4. See also GROENEVELD *et alii.* [1998] for an analysis of the impact of the EMS crisis on the process of European convergence.

Italy, the UK and Spain.[5] France and the Netherlands are taken as representative stayers, and Germany itself is considered as a benchmark case. We first consider the specification and the results for monetary policy rules.

2.1 Specification and Stability of Monetary Rules

The estimation of monetary rules has met an amazing success in recent years. The usual framework is provided by a *Taylor* rule, generally augmented in order to cope for interest rate persistence (or smoothing). The changes of monetary rules according to the Fed chairman is now well documented (JUDD and RUDEBUSCH [1998]). Less evidence is available for the European countries, but the influential study by CLARIDA, GALÍ and GERTLER [1998] illustrates the interest of comparative estimates of (generalized) *Taylor* rules. Although relevant for the US, a *Taylor* rule equation is not adequate for the European case of open-medium to small-economies. An open economy version under the assumption of high but finite capital mobility (or imperfect currency substitutability) has to introduce the foreign rate as a major determinant of domestic interest rate. While this point is usually acknowledged, estimates of European interest rate equations neglect another major determinant, *ie*, the expected depreciation rate of the national currency.[6] Of course, this variable is not exogenous and we must account for this feature in the estimation.

Keeping the CLARIDA, GALÍ, GERTLER's [1998] notations, the typical backward-looking monetary policy rule[7] is:

$$(1) \quad \Delta r_t = (1-\rho)\left(\alpha + \beta\pi_t + \gamma y_{t-1} + \xi r^f_{t-1} + \varphi\Delta\widetilde{e}_t\right) + (\rho - 1)\, r_{t-1} + \varepsilon_t$$

where r_t is the 3-month money market rate, $\pi_t \equiv P_t - P_{t-4}$ denotes the annual inflation rate, y_{t-1} is an output gap measure and r^f_t is the short foreign rate (the US T -Bill for Germany, the German rate for all the other countries). $\Delta\widetilde{e}_t$ is a predictor of the national currency depreciation, conditional on information available in $t-1$, thus predetermined in the equation. [8] ρ is the persistence parameter, associated to an interest rate smoothing component in the central bank objective function.

CLARIDA *et alii.* [1998] emphasize the importance of the choice of the regression period which ought to be associated to a stable (or near-stable) institutional environment. They retain rather short sample periods with monthly data considering that the strong EMS prevailing from 1990 to the mid-1992 precludes any national rule other than a full alignment on German rates.

5. Remember that Spain was not a leaver in the institutional sense, as it did not quit the EMS. However, the strategy of large depreciation of the peseta resulted in equivalent practical effects.

6. An alternative solution for an open economy could be to depart from a *Taylor* rule by controlling a mci (monetary condition index) combining the interest and exchange rates (BALL [1998]).

7. As usual in the literature, we assume that such an interest rate equation identifies the policy rule, from a reduced form equation associated to the equilibrium in the money market.

8. Substituting the $\Delta\widetilde{e}_t$ predictor conditional to 'available information' is equivalent to regressing on realized Δe_t with a set of instuments characterizing this 'available information' at $t-1$.

Our perspective is different. We want to estimate the equation ruling the interest rate until the break-up, therefore under the early 90's. Indeed, our exploratory recursive regressions do not support the idea of a stronger alignment on the German rate during the 'hard EMS' period.[9] Furthermore, as we want to relate this study of monetary rules with equations for a budget rule, output gap and inflation, we use quarterly data, further limiting the degrees of freedom of our estimates.

Our inference strategy consists in considering the period preceding the break-up as the reference period. For every country, we first present a baseline estimate on a sample period spanning from 1979 (or later, up to 1983-2, previously presented as the beginning of a calmer period) to 1992-2. We then present estimates for longer sample periods, including the post-break-up experience until 1993-2, 1994-4 or 1996-4. This examination of different post-break horizons is performed in order to recover information on temporary changes in policy.[10] The choice of the first horizon (1993-2) is commanded by the agreement on wider bands, in August 1993, which ended the period of greatest uncertainty.

2.2 **Comparative Evidence on the Monetary Rules**

Results for baseline interest rate equations, and their stability after the 1992 break-up, are presented on table 1, a and b.

As it is widely documented,[11] we consider that changes in interest rates have only a negligible impact on the current inflation rate, which allows us to estimate the equation (1) using OLS.[12]

For *Germany*, the baseline period is limited to 1980-4 -1989-3, ending before the unification shock. The foreign rate is the US one and a second order lag is found significant.[13]

As reported on the first line of table one, the German monetary rule is found constant over the baseline period, both using *Chow* mid-sample test and HANSEN's [1992] Jit (joint invariance test).[14] When examining the stability of this rule, after the break-up, we notice that the null of parameter constancy is rejected for the 1993-2 and the 1996-4 horizon, either by the *Chow* predictive test, or by the Jit one. A closer examination of the parameters reveals an increase in the persistence parameter ρ and a temporary decrease in the impact of the inflation rate.[15]

---

9. As shown by SVENSSON [1994], the system of exchange rate bands allows for a margin of independence of national monetary policies even with free international capital mobility.

10. Due to the *ex post* evidence that policies have ultimately converged, it would be irrelevant to test for permanent policy changes. For the same reason, we disregard the data posterior to 1996.

11. See *eg*, CHRISTIANO, EICHENBAUM and EVANS [1999]. If a change in the interest rate impacts inflation only with one lag, π_t is predetermined and therefore an eligible regressor in equation (1).

12. *De facto*, we use non linear OLS in order to recover directly the (just identified) parameters of equation (1).

13. We cope with this second order lag by introducing the lagged rate in difference. The coefficient of R_{t-1}, in levels, is still the only one relevant for assessing the long run effects.

14. Hansen's Jit test is performed on the L_c statistic computed from the vector of the scores and their sample cumulative sums (HANSEN [1992], equation 9, see also appendix II of this paper). Critical values are taken from table 1 in HANSEN [1992, p. 534].

15. We comment here changes in parameters from estimations performed on longer samples. A more specific inference on the post 'break-down' values would have required a new estimation bearing only on the additional observations. However, this solution is precluded as we lack sufficient degrees of freedom at the horizons considered.

TABLE 1

Estimated Monetary Rule (baseline and 'post-break')

Estimated equation:

$$\Delta r_t = (1-\rho)(\alpha + \beta\pi_t + \gamma y_{t-1} + \xi r^f_{t-1} + \varphi\Delta\tilde{e}_t) + (\rho - 1)r_{t-1} + \varepsilon_t$$

Germany[1]	α	β	γ	ξ	φ	δ	$\overline{R}^2$	Chow/Jit	AR1-4
80.4-89.3	1.54	0.200	0.625	0.466	–	0.507	0.57	0.42/1.0	1.62
	(1.9)	(1.36)	(3.8)	(4.0)		(4.8)			
80.4-93.2	3.72	0.285	1.00	0.268	–	0.751	0.44	2.70*/1.34	0.27
	(4.1)	(2.0)	(5.5)	(2.2)		(10.2)			
80.4-94.4	3.82	0.325	0.969	0.245	–	0.756	0.48	0.90/1.87*	1.75
	(5.0)	(2.4)	(5.9)	(2.5)		(11.4)			
80.4-96.4	3.19	0.217	1.08	0.328	–	0.826	0.43	0.86/2.05*	1.75
	(3.2)	(1.1)	(4.6)	(2.4)		(14.6)			
France									
83.1-92.2	4.69	0.674	0.223	0.285	0.070	0.358	0.43	0.42/1.01	1.49
	(11.6)	(15.2)	(3.9)	(6.2)	(2.1)	(2.9)			
83.1-93.2	4.06	0.623	0.133	0.442	0.084	0.420	3.35	15.4**/1.89+	1.10
	(6.3)	(9.1)	(1.3)	(3.4)	(1.4)	(2.2)			
83.1-94.4	4.3	0.697	0.600	0.300	0.084	0.660	0.27	8.6**/1.88+	1.75
	(3.7)	(8.0)	(2.3)	(1.65)	(0.8)	(3.7)			
83.1-96.4	2.57	0.785	0.745	0.469	0.274	0.737	0.24	7.3**/2.05*	1.75
	(1.7)	(7.2)	(2.3)	(1.85)	(1.6)	(5.9)			
Netherlands									
80.1-91.2	2.00	–	0.242	0.775	–	0.282	0.76	0.40/1.26	1.21
	(6.9)		(3.0)	(17.0)		(4.4)			
80.1-93.2	2.06	–	0.234	0.759	–	0.283	0.75	0.56/1.39	1.02
	(7.2)		(2.8)	(17.7)		(4.5)			
80.1-94.4	1.91	–	0.242	0.770	–	0.308	0.72	0.92/2.10**	1.88
	(6.6)		(2.7)	(17.8)		(4.8)			
80.1-96.4	1.24	–	0.203	0.847	–	0.397	0.60	1.78*/3.58**	7.52**
	(3.6)		(2.0)	(17.4)		(5.3)			

*Note 1: Under the coefficients are given the t-statistics. The Chow statistics refers to mid-sample for the baseline, pre-break sample, and to its extensions up to the successive horizons for the other lines. The Jit is the HANSEN [1992] statistics over respectively the baseline and the extended samples. AR1-4 is a LM statistics for joint autocorrelation of order to 4. Superscripts for the Chow, Jit and AR statistics are respectively a + for significance at 10%, a * for significance at 5%, ** for significance at 1%.*

TABLE 1

Estimated Monetary Rule (following)

Estimated equation:

$$\Delta r_t = (1-\rho)(\alpha + \beta\pi_t + \gamma y_{t-1} + \xi r^f_{t-1} + \varphi\Delta\tilde{e}_t) + (\rho - 1)r_{t-1} + \varepsilon_t$$

United Kingdom	α	β	γ	ξ	φ	δ	$\overline{R}^2$	Chow/Jit	AR1-4
79.2-92.2	7.24	0.243	0.488	0.391	0.041	0.474	0.33	0.49/0.76	0.24
	(6.2)	(3.2)	(3.1)	(2.2)	(1.1)	(3.9)			
79.2-93.2	7.44	0.402	0.634	0.150	0.066	0.590	0.27	2.14/1.16	0.30
	(5.0)	(3.5)	(2.9)	(0.7)	(1.4)	(5.2)			
79.2-94.4	6.49	0.469	0.877	0.173	0.096	0.711	0.22	1.20/1.50	0.35
	(3.3)	(3.1)	(3.2)	(0.5)	(1.6)	(9.0)			
79.2-96.4	5.38	0.464	1.015	0.307	0.106	0.751	0.21	0.78/1.56	0.41
	(3.0)	(2.7)	(3.4)	(1.0)	(1.74)	(12.0)			
Italy									
80.1-92.2	7.45	0.567	– 0.42	0.284	–	0.558	0.45	0.98/1.34	0.46
	(9.4)	(5.0)	(– 1.6)	(1.76)		(5.0)			
80.1-93.2	7.46	0.573	– 0.24	0.266	–	0.586	0.36	8.99**/1.02	1.36
	(8.4)	(6.2)	(– 0.7)	(1.1)		(5.9)			
80.1-94.4	7.03	0.691	0.595	0.082	–	0.760	0.31	4.56**/1.26	0.90
	(4.2)	(5.6)	(1.4)	(0.2)		(12.1)			
80.1-96.4	5.64	0.683	0.492	0.276	–	0.771	0.29	3.05**/1.18	1.05
	(3.6)	(5.2)	(1.2)	(0.8)		(12.5)			
Spain									
80.1-92.2	6.57	0.577	0.564	0.340	–	0.436	0.39	0.34/1.82$^+$	1.23
	(4.3)	(3.2)	(2.7)	(1.4)		(4.0)			
80.1-93.2	6.74	0.515	0.507	0.409	–	0.435	0.40	2.76*/1.79$^+$	1.22
	(4.3)	(3.5)	(2.9)	(2.1)		(4.2)			
80.1-94.4	5.37	0.727	0.818	0.303	–	0.503	0.38	2.46*/1.92*	1.15
	(3.0)	(4.8)	(5.5)	(1.5)		(5.5)			
80.1-96.4	4.96	0.754	0.850	0.321	–	0.510	0.38	0.31/2.22*	1.26
	(3.3)	(5.6)	(6.3)	(1.6)		(6.1)			

The baseline sample period for *France* spans from 1983-1 to 1992-2, the French economic policy being too unstable between 1979 and 1982. Both the *Chow* test and the *Jit* lead to accept the hypothesis of a constant rule over this period, while both tests reject constancy on the longer samples. As in *Germany*, we get a larger persistence in French interest rates after 1992.

We may remark that, except on the full sample, this French rule is almost inflation neutral, the French real rate being decreasing in response to the French inflation and roughly constant in response to a parallel increase in the French inflation rate and the German nominal interest rate. With such a rule, the French monetary policy would fight inflation only through importing the Bundesbank policy.

Looking at the long run coefficients, it appears that, after 1992, the French monetary policy becomes more reactive to the domestic inflation, the domestic GDP gap and the German rate.

From the inception of the snake, the *Dutch* guilder stayed aligned on the DM and the Dutch monetary policy was strongly committed to follow the German one. The *Jit* test rejects the null of constant coefficients after the break-up of 1992.

A direct examination of the 1992 episode reveals that the *Bank of England* took the opportunity of the break-up to lower its interest rates with respect to the German ones. Surprisingly, this feature is not detected as a case of structural instability by either the *Chow* or the *Jit* tests. However, point estimates, like the recursive ones, are informative on parameter changes. British rates therefore become more persistent. The influence of the German rates weakens. Stable short run response to domestic inflation and GDP gap, together with an increasing autoregressive coefficient, mean that the UK monetary policy actually became more responsive to domestic conditions. Thus, leaving the EMS allowed the UK to recover an autonomous monetary policy in accordance with the *a priori* prediction.

Italy has a tradition of accommodative monetary policy, with a significant contribution of seignorage to budget financing. Our estimates of the monetary rule actually imply an underindexation of nominal rates to parallel increases in both the domestic inflation rate and the German monetary rate. No structural change is detected by the *Jit* test, although the *Chow* and recursive tests are very supportive of changes between 1991 and 1993, especially in 1992.

As in the UK, the exit from EMS allows for a policy temporarily more independent from the German rates. As the rule becomes more reactive to domestic inflation, it is not more permissive than during the EMS period.

Spain was a late comer in the European Union and especially in its nominal convergence process. The eighties were still a period of monetary instability and of structural reforms in financial markets and institutions. A reasonable interest rule is estimated for the period 1979.2 to 1990.4. With a dummy covering the episode 1987.1-1987.2, the mid-sample *Chow* test does not reject the stability of the rule over this decade. The *Jit* rejects stability over the 1979-1994 or -1996 sample periods while the predictive *Chow* test does not.

The study of the monetary rules in our sample group of European countries is mildly supportive of *a priori* predictions. In accordance with the theory, the leavers *de jure*, the *UK* and *Italy*, or *de facto*, *Spain*) is found to have gained more autonomy with respect to the German monetary policy. Paradoxically, structural changes in the interest rate equation affect the stayers as much as the leavers. This may result from a change in the German rule which has essentially preceded the monetary unification and implied more persistence, a stronger commitment to domestic objectives and less dependence with respect to the US rates. While keeping a narrow dependency to the German rate, the French rule became also more sensitive to domestic objectives. Only in the *Netherlands* did the stronger alignment on German rates reduce the sensitivity to the domestic GDP gap.

From these estimates of changing monetary rules, we keep the various fitted regressors and dummy variables for a later use in weak exogeneity and invariance tests.

2.3 Comparative Evidence on Changing Fiscal Policies

There is no consensus on the proper specification of budget rules such as the one prevailing on monetary rules. First, it is unclear that focusing on the budget balance is sufficient enough, as expenditure cuts (resp. increases) and tax increases (resp. cuts) are far to be equivalent; secondly, the behavior of fiscal authorities, from the elected political majorities to the administrative bureaucraties, may be less systematic and predictable than the one of central banks. This suggests that estimated fiscal rules will not perform so well as monetary policy rules.[16] However, it is interesting to attempt to characterize the shifts in fiscal policies in terms of a changing rule.

Due to difficulties of building a structural primary surplus (on a quarterly basis), we consider an equation determining the actual global surplus.

The first mechanism to be embedded in a fiscal rule is related to the sustainability criterion:[17] a sufficient condition is the stationarity of the global deficit ratio (to GDP or trend GDP) (KREMERS [1989]) or a positive effect of the inherited debt ratio on the primary surplus (BOHN [1995]).[18] Consequently, we *a priori* introduce both the lagged surplus ratio and the inherited debt ratio as regressors in the fiscal policy equation. The second mechanism is related to the stabilization purpose, leading us to expect a negative coefficient for the output gap and a positive one for inflation. However, these expected effects may be counteracted by the automatic effects which entail a positive impact of the gap on the actual global surplus and a (possibly weak) positive effect of inflation. The lagged interest rate is also introduced with two possibly conflicting effects: A negative effect as a determinant of current interest payments but a positive impact associated with sustainability requirements.

Our expression for the budget rule, expressed as response of the actual global surplus ratio, takes the following form,

$$\Delta s_t = a + b\Delta s_{t-1} + cs_{t-2} + d\pi_t + er_{t-1} + fy_{t-1} + u_t \quad (2)$$

16. The impact of *a priori* specified fiscal rules was studied by BRYANT and ZHANG [1996] in various macroeconomic models. Empirical fiscal rules are estimated by TAYLOR [2000] as response of surplus and by FAIR [1999] as response of tax rates. See PEREZ and HIEBERT [2001] for further elaboration on tax rules in macro models.

17. Sustainability of fiscal policy requires that the discounted government debt does not diverge or, in a more practical definition (KREMERS [1989]), that the debt-to-gdp ratio remains bounded.

18. As cyclical adjustments only involve stationary transformations, they do not affect these sustainability conditions.

TABLE 2a

Estimated Fiscal Rules (baseline and 'post-break')1

Germany	Cst	s_{t-2}	dS_{t-1}	dS_{t-4}	y_{t-1}	r_{t-1}	π_{t-1}	b_{t-1}	R^2	Chow/Jit	AR1-4
79.3-90.4	-1.91	0.40	0.60	0.46	0.12	-0.03	-0.07	0.04	0.90	0.71/ 2.29	1.52
	(-3.98)	(-9.12)	(8.61)	(5.14)	(4.44)	(-1.52)	(-2.33)	(3.46)			
79.3-93.2	-0.35	-0.24	0.45	0.16	0.01	-0.06	-0.007	0.007	0.80	2.72**/ 1.72	0.83
	(-0.97)	(-6.30)	(5.67)	(1.87)	(0.66)	(-2.74)	(-0.26)	(0.76)			
79.3-94.4	-0.03	-0.21	0.46	0.12	-0.001	-0.05	-0.02	-0.002	0.77	2.78**/ 1.81	0.97
	(-0.11)	(-5.72)	(5.73)	(1.28)	(-0.08)	(-2.15)	(-0.79)	(-0.33)			
79.3-96.4	0.33	-0.18	0.52	0.08	-0.01	-0.02	-0.04	-0.01	0.76	2.46**/2.73*	1.52
	(1.40)	(-5.3)	(7.01)	(0.93)	(-0.86)	(-1.26)	(-1.58)	(-2.63)			
France											
83.1-91.4	0.16	-0.14	0.62	–	0.04	-0.06	0.02	–	0.87	0.68/ 1.1	2.14
	(0.92)	(-3.37)	(6.44)		(2.23)	(-2.20)	(1.03)				
83.1-93.2	0.30	-0.12	0.51	–	0.03	-0.07	0.04	–	0.90	2.61*/ 0.64	1.69
	(2.2)	(-3.14)	(5.23)		(1.65)	(-2.9)	(2.05)				
83.1-94.4	0.25	-0.08	0.43	–	0.01	-0.05	0.02	–	0.87	2.60*/ 0.79	1.20
	(1.96)	(-4.03)	(4.54)		(0.83)	(-3.14)	(1.70)				
83.1-96.4	0.11	-0.09	0.51	–	0.01	-0.03	0.03	–	0.87	2.09*/ 0.69	1.35
	(1.2)	(-3.15)	(6.28)		(1.04)	(-2.85)	(0.92)				
Netherlands											
80.4-90.4	-2.46	-0.26	0.55	-0.08	0.05	-0.06	0.10	0.02	0.82	0.98/ 1.36	1.11
	(-2.80)	(-3.31)	(5.76)	(0.50)	(1.33)	(-1.69)	(1.94)	(1.90)			
80.4-93.2	-2.99	-0.21	0.46	-0.26	0.03	-0.07	0.14	0.03	0.81	0.82/ 1.58	0.77
	(-4.13)	(-3.50)	(5.77)	(-2.42)	(0.98)	(-2.18)	(3.52)	(4.28)			
80.4-94.4	-3.05	-0.23	0.49	-0.18	0.02	-0.04	0.11	0.03	0.81	0.72/ 1.55	0.12
	(-4.38)	(-4.38)	(6.81)	(-2.14)	(0.88)	(-1.73)	(3.52)	(4.23)			
80.4-96.4	-2.47	-0.18	0.52	-0.24	0.03	-0.07	0.12	0.02	0.80	0.68/ 0.1	1.56
	(-3.92)	(-3.92)	(7.64)	(-2.97)	(1.13)	(-3.47)	(4.10)	(4.16)			

Note: with r_t the money market rate, π_t the yearly inflation rate, y_{t-1} the output gap, s_t the budget ratio and b_t the public debt.

TABLE 2b

Baseline and 'post-break' Fiscal Rules (following)

United Kingdom	Cst	s_{t-2}	dS_{t-1}	dS_{t-4}	y_{t-1}	r_{t-1}	π_{t-1}	b_{t-1}	R^2	Chow/Jit	AR1-4
81.1-91.4	0.22	-0.77	-0.81	–	0.29	-0.22	0.25	-0.03	0.59	0.40/ 2.40	1.35
	(0.07)	(-4.33)	(-5.35)		(1.94)	(-1.71)	(2.51)	(-0.52)			
81.1-93.2	-3.99	-0.54	-0.68	–	0.31	-0.09	0.21	0.04	0.53	0.95/ 1.85	0.74
	(-1.42)	(-3.63)	(-4.79)		(2.11)	(-0.72)	(2.07)	(0.93)			
81.1-94.4	-3.34	-0.52	-0.67	–	0.27	-0.05	0.19	0.03	0.52	0.78/ 1.81	0.60
	(-1.32)	(-3.86)	(-5.01)		(2.21)	(-0.51)	(2.06)	(0.65)			
81.1-96.4	-3.04	-0.54	-0.71	–	0.27	-0.07	0.20	0.02	0.53	0.71/ 1.81	0.63
	(-1.51)	(-4.1)	(-5.67)		(2.41)	(-0.65)	(2.26)	(0.74)			
Italy	Cst	s_{t-2}	dS_{t-1}	dS_{t-2}	y_{t-1}	r_{t-1}	π_{t-1}	b_{t-1}	R^2	Chow/Jit	AR1-4
80.1-92.2	-5.83	-0.49	0.04	0.53	0.05	-0.07	0.06	0.02	0.86	0.55/2.32	2.22+
	(-7.39)	(-9.37)	(0.31)	(4.69)	(1.75)	(-2.97)	(2.66)	(3.78)			
80.1-93.2	-5.88	-0.41	0.11	0.47	0.03	-0.04	0.06	0.02	0.82	0.95/3.32**	1.26
	(-6.88)	(-8.07)	(0.91)	(3.89)	(1.04)	(-1.95)	(2.58)	(5.06)			
80.1-94.4	-5.87	-0.42	0.16	0.43	0.02	-0.05	0.06	002	0.82	0.94/3.19*	1.64
	(-6.94)	(-8.21)	(1.30)	(3.69)	(1.07)	(-3.44)	(3.45)	(6.06)			
80.1-96.4	-3.22	-0.21	0.49	0.15	-0.004	-0.04	0.04	0.02	0.74	1.68+/4.47**	3.52*
	(-4.02)	(-5.38)	(3.95)	(1.21)	(-0.17)	(-2.22)	(2.08)	(3.94)			
Spain	Cst	s_{t-2}	dS_{t-1}	dS_{t-4}	y_{t-1}	r_{t-1}	π_{t-1}	b_{t-1}	R^2	Chow/Jit	AR1-4
79.2-91.2	0.51	-0.35	0.60	0.36	0.21	-0.03	-0.01	-0.03	0.85	0.54/1.46	1.94
	(0.77)	(-5.97)	(7.53)	(2.92)	(4.18)	(-2.91)	(-0.41)	(-2.90)			
79.2-93.2	0.77	-0.35	0.53	0.18	0.17	-0.02	-0.04	-0.04	0.80	2.46*/2.17	2.65*
	(0.97)	(-5.33)	(6.58)	(1.40)	(3.57)	(-1.20)	(-1.08)	(-3.12)			
79.2-94.4	0.81	-0.27	0.54	-0.04	0.08	-0.006	-0.06	-0.04	0.75	2.80**/2.10	1.50
	(0.97)	(-4.32)	(6.78)	(-0.35)	(2.40)	(-0.46)	(-1.27)	(-2.68)			
79.2-96.4	0.006	-0.19	0.62	-0.04	0.07	-0.02	-0.07	-0.01	0.72	3.09**/2.82*	1.96
	(0.007)	(-3.12)	(7.76)	(-0.37)	(1.93)	(-1.29)	(-0.16)	(-1.02)			

where s_t is the actual surplus to gdp ratio, π_t the inflation rate, r_t the interest rate, y_t a measure of the output gap and b_t a debt-to-gdp ratio.[19] As the effect of the surplus on current inflation is assumed negligible, the equation (2) is estimated using OLS, with π_t a valid regressor.

The evidence from estimated budget rule for *Germany* is that instability occurred before the EMS break-up due to the unification effect. We thus retain for the baseline estimate a sample period ending in 1990-4. A higher persistence in the deficit ratio after the unification, together with a smaller response to past debt, means that the German fiscal policy has become less likely to satisfy sustainability requirements.

The *French* fiscal policy is also characterized by deeper deficits in the nineties and a build-up of government debt. This worsening, detected as a significant structural change by the *Chow* test but not by the *Jit* tests, is clearly shown up by the recursive statistics, as arising during the fiscal year 1991.

The *Netherlands* are of the few European countries which have not experienced a worsening of their budget balance in the early nineties. The main structural shift[20] detected by the recursive statistics in 1992.1 is associated with an underprediction of the surplus, therefore indicative of a stronger budget policy.[21] The Dutch budget surplus especially becomes more reactive to the inherited debt level, as required by BOHN [1995] sustainability criterion.

The *UK* budget balance deteriorated, from an historical global surplus in 1988 to a large deficit in 1993. According to recursive statistics, the interventions on the surplus equation occurred in 1989.1 and 1990.3. We introduce dummies to account for these fiscal shocks and stabilize the rule estimated over the baseline period. As in other countries, some fiscal shocks have therefore preceded the EMS break-up, possibly contributing to weaken the position of the pound within the EMS. The main change affecting the budget equation after 1991 is the drastic increase in deficit persistence, violating the sustainability condition.

With a tradition of large budget deficits, *Italy* performed a consolidation effort in 1989, a setback in 1990-91 and began its march towards satisfying the Maastricht deficit criteria from 1995. These policy changes lead to a significant coefficient instability, after the EMS breakup, according to the predictive *Chow* tests and the *Jit*.

After a dramatic budget consolidation initiated in 1985, *Spain* experimented the deterioration common to many European countries from 1990. While not significant prior to 1994-4 according to the *Jit* tests, interventions to the budget process are detected by the *Chow* tests and the recursive statistics in late 1991 and 1992.

Cross country evidence supports the view that the deterioration of budget balances in Europe preceded the EMS problems, and were followed by more deficit persistence, generally violating the sustainability requirements. Our sample termination date is too early for the final effort aiming at the satisfaction of the convergence criteria to show up. On the basis of the estimated rule, we are left with the paradoxical evidence that the Maastricht Treaty may well have postponed the unavoidable budget consolidation.[22]

19. See the appendix, for further information on the data.
20. However, the structural stability tests are also suggesting that some changes occurred previously, within the baseline period.
21. The structural instability also results in a strong residual autocorrelation on the longer samples.
22. See M. CARRÉ [2000] for a theoretical analysis of the impact of the EMU deadline on a possible postponement of fiscal adjustments in Europe.

3 Reduced Form Behavioral Equations: Structural Stability and Invariance

The early nineties were a time of significant policy changes for a good many of the European countries, as documented on our sample group. Let us now examine to which extent these policy changes have resulted in structural changes in behavioral relationships, by examining reduced form equations for the GDP gap, the (CPI) inflation rate and the term structure. Of course, this choice will leave unexplored the stability of more structural behavioral equations considered in the literature, like the consumption function (FAVERO [1993]), money demand (FISHER and PEYTRIGNET [1991], and many references by D. Hendry and his coauthors, see FAVERO and HENDRY [1992]), investment (OLINER, RUDEBUSCH and SICHEL [1996]). As BLANCHARD [1984], MILLER and ROBERDS [1991], ESTRELLA and FUHRER [1999], we choose to focus on relationships more directly related to the ultimate success of policies. Before proceeding with the specification and estimation of these equations, we first introduce our approach to the problems of exogeneity and invariance.

3.1 A Framework for Evaluating Exogeneity and Invariance

In order to assess the exogeneity and invariance properties in our empirical model, we use the following testing strategy, adapted from ENGLE and HENDRY [1993]. Let y_t be a variable of interest, resulting from agents behavior, and z_t an instrument or policy variable. For sake of parsimony, we do not introduce exogenous conditioning variables which would leave the argument unchanged.

The assumption is the following: agents structurally react to expected values $\widetilde{z}_t$ of the policy instrument z_t, resulting from the rule.[23]

(3) $$z_t = b_{0,t} + b_{1,t} y_{t-1} + b_{2,t} z_{t-1} + \sigma_t \varepsilon_t$$

where the $b_{i,t}$ and σ_t are potentially time varying parameters and ε_t an i.i.d. Gaussian disturbance.

The behavioral equation is assumed to be:

(4) $$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 \widetilde{z}_t + \alpha_3 \widetilde{z}_{t-1} + v_t$$

with v_t an i.i.d. Gaussian disturbance. For $\alpha_3 \neq 0$, equation (4) allows for a delayed effect of the policy instrument, as documented in the literature.

23. We assume here parsimonious lag specifications. Longer lag distributions may be introduced without changing the argument. It follows from (3) that the relevant information for agents is the set of part realizations of the process $(y_{t-i}, z_{t-i}, i > 0)$.

A conditional empirical model of y_t, obtained as the mere linear regression of y_t on the r.h.s. variables in equation 4, is:

$$(5) \qquad y_t = a_0 + a_1 y_{t-1} + a_2 z_t + a_3 z_{t-1} + u_t$$

We are concerned with the following null hypotheses, respectively of weak exogeneity and invariance:

H_0^{WE} : *z_t and z_{t-1} are weakly exogeneous w.r.t. y_t in the conditional model (5) for α_2 and α_3.*[24]

H_0^{INV}: *a_2 and a_3 in (5) are invariant for parameter changes affecting the rule (3), ie, b_{it} or σ_t.*

As usual, weak exogeneity means that we may recover an evaluation of the impact of the policy parameters from the regression (5) on realized values of z, disregarding information from the marginal equation (3). Further, invariance would entail that this inference is not affected by changes in parameters of equation (3) in our study.

These null hypotheses are straightforward extensions of the usual ones when we allow for a lagged conditioning and include α_3 within the set of parameters of interest. A model representation of the underlying factorization of the joint distribution as the product of a conditional distribution and a marginal one.

ie, $D_j\left(y_t, y_{1t}, z_t, z_{1t} \,\middle|\, y_{t-1}, y_{1t-1}, z_{t-1}, z_{1t-1}\right) = D_c\left(y_t / y_{1t}, z_t, z_{1t}\right).$
$D_m\left(y_{1t}, z_t, z_{1t} / y_{t-1}, y_{1t-1}, z_{t-1}, z_{1t-1}\right)$

is provided in appendix I.[25]

As noticed by the editor, such a regression model representation would coincide with a conditional distribution only when the variables of interest follow a joint normal distribution.

In order to render these null hypotheses operational, we introduce the two following augmented equations: First, equation (6):

$$(6) \; y_t = a'_0 + a'_1 y_{t-1} + a'_2 z_t + a'_3 z_{t-1} - a'_4(z_t - \widetilde{z}_t) - a'_5(z_{t-1} - \widetilde{z}_{t-1}) + w_t$$

where expected values $\widetilde{z}_t$ are computed from an invariant policy rule (3)

and, second,

$$(7) \quad \begin{aligned} y_t = {} & a''_0 + a''_1 y_{t-1} + a''_2 z_t + a''_3 z_{t-1} - a''_4(z_t - \widetilde{z}_t) - \alpha''_5(z_{t-1} - \widetilde{z}_{t-1}) \\ & - a''_6(\tilde{z}_t - \widetilde{\widetilde{z}}_t) - a''_7(\tilde{z}_{t-1} - \widetilde{\widetilde{z}}_{t-1}) + \eta_t \end{aligned}$$

where $\widetilde{\widetilde{z}}_t$ stands for the expected value of the instrument accounting for the possible changes in the rule from the time-varying parameters equation (3), with,

24. Although the parameters α_2 and α_3 do not formally appear in equation 5, notice that a_2 and a_3 precisely coincide with α_2 and α_3 under H_0^{WE}.

25. In the previous decomposition, the notation y_{1t} is introduced for y_{t-1} (resp. $z_{1t} \equiv z_{t-1}$), following the notational convention used for expressing a VAR model in its companion form, in order to compute conditional expectations.

$$\widetilde{\widetilde{z}}_t = \widehat{b}_{0,T(t)} + \widehat{b}_{1,T(t)} y_{t-1} + \widehat{b}_{2,T(t)} z_{t-1} \quad (8)$$

Where $T(t)$ indicates the information set –practically the sample period– on which rely the estimates $\widehat{b}_{iT(t)}$ of the parameters used for computing expectations where agents are aware of the policy changes.

In this framework, a test for the weak exogeneity hypothesis is obtained by taking the joint restriction $a'_4 = a'_5 = 0$ in equation (6) as the null hypothesis against the two sided alternative $a'_4 \neq 0$, $a'_5 \neq 0$.[26] Similarly an invariance test derives from assessing the restriction $a''_6 = a''_7 = 0$ against the unconstrained alternative in equation (7).

Superexogeneity, as the combination of weak exogeneity and invariance, can further be tested as the restriction $a''_4 = a''_5 = a''_6 = a''_7 = 0$ in (7). Under the expectational model (4), both weak exogeneity and invariance would fail. When this model is true, we expect both $a'_i \neq 0, i = 4{,}5$, in (6) and $a''_i \neq 0$, i = 4,..., 7, in (7). A failure to reject the nulls would therefore provide evidence that the *Lucas critique* is not empirically relevant.

Our test is of course related to alternative superexogeneity tests proposed in the literature. ENGLE and HENDRY [1993] use cross products from the limited development of $(a_{2t}\widetilde{z}_t)$ instead of our $\widehat{\widetilde{z}}_t \equiv (\widetilde{z}_t - \widetilde{z})$ terms. ERICSSON, HENDRY and MIZON's [1998] so called 'test II' may be viewed as the approximation of these terms by a set of dummies sufficient to account for any source of non constancy in the policy equation (3).[27] Our test is more parsimonious than this 'test II' as it does not require to introduce as many dummy variables as there are observations affected by policy change. With respect to the general test proposed by ENGLE and HENDRY [1993], we think that our test may have more power for a defined policy change.[28] Indeed, the extra regressors we introduce are likely to track more closely the information relevant under the alternative. However, ENGLE and HENDRY's [1993] test may be more robust in the case of undefined policy changes. Indeed, their test rests on estimated changes of moments, which are not conditional on a particular specification of the policy change. Estimation of our extra parameters in equations (6) and (7) is conditional on a specification of the policy changes and therefore may lack of robustness in case of misspecification of these changes. Our approach may also be related to MILLER and ROBERDS' [1991] evaluation of the effect of policy changes by assessing the different impacts on respectively conditional and unconditional forecasts of policy variables.

3.2 Reduced Form Equations: Specification and Structural Stability Tests

At this point of the study, we consider backward reduced form specifications of equations for output, inflation and the long term interest rate, as in

26. This test of weak exogeneity was introduced by REVANKAR and YOSHINO [1990]. Its equivalence with a *Haussman* test follows from HOLLY [1982].

27. See FISHER and PEYTRIGNET [1991] for an early application.

28. Of course, the empirical implementation of this potential gain of power rests on the possibility of building efficient measures of the expectational variables conditional on estimated dates of policy change.

MILLER and ROBERDS [1991]. Invariance studies of the inflation rate include ALOGOSKOUFIS and SMITH [1991], while ESTRELLA and FUHRER [1999] consider both backward and forward equations for inflation and output[29].

Output equations take the general form:[30]

$$(9)\ y_t = a_0 + a_1 y_{t-1} + a_1 x_t + a_2 \pi_{t-1} + \sum_i a_{3i} r_{t-i} + \sum_i a_{4i} s_{t-i} + u_{1t}$$

where y_t stands for the output gap, x_t for the foreign market growth, π_{t-1} for the yearly inflation rate, r_{t-i} and s_{t-i} for values of the (short) interest rate and the budget balance ratio.

Inflation equations are specified as:

$$\begin{aligned} \Delta\pi_t =& c_0 + c_1\pi_{t-1} + \sum_{i=1}^{k} c_{2i}\Delta\pi_{t-i} + c_{3i}\omega_t \\ &+ \sum_{i=1}^{k} c_{4i} r_{t-i} + \sum_{i=1}^{k} c_{5i} s_{t-i} + u_{2,t} \end{aligned} \quad (10)$$

and the long term interest rate equations are variants of the following:

$$\begin{aligned} \Delta R_t =& d_0 + d_1 S_{t-1} + d_2\Delta R_{t-1} + d_3\Delta r_t + d_4\Delta r_{t-1} \\ &+ d_5\pi_{t-1} + \sum_{i=0}^{k} d_{6i} s_{t-i} + u_{3t} \end{aligned} \quad (11)$$

with ω_t the (growth rate) in oil and raw materials price, ΔR_t the change in the long rate, S_t the difference (spread) between the long run (ten years) and the 3-month interest rate.

Given these conditional behavioral equations, our empirical strategy will include the following steps:

i) Estimate these equations on the baseline sample period, prior to 1992-3, selecting a proper specification, through the choice of lags and variable deletion from the general specifications (9, 10, 11).

ii) Reestimate the same equations on longer sample periods in order to check for constancy at the three selected horizons (1993-2, 94-4, 96-4).

iii) Estimate these equations, augmented by extra regressors conformably to (6) and (7) in order to perform our weak exogeneity, constancy and invariance tests.

---

29. BLANCHARD [1984] found the term structure equation structurally more sensitive than the *Phillips* curve to the VOLCKER [1979] policy break. FUHRER [1996] provides a nice approach to the invariance in long term interest rate equation.

30. We do not *a priori* impose neutrality by constraining the coefficients of the nominal interest and inflation rate to sum up to zero this restriction being generally rejected by the data. Notice that besides the policy instruments, we introduce a minimal set of exogenous conditioning variables, *ie*, the foreign market growth in the output equation and the oil and raw material prices in the inflation equation.

TABLE 3

Constancy Tests for the 'post-break' Behavioural Reduced Form Equations

Output gap equation[1]

Country	Predetermination test[2]	Constancy (Chow/Jit)		
		baseline to 93.2	baseline to 94.4	baseline to 96.4
Germany	0.17	5.42** / 1.86	3.96** / 1.97	2.86** / 1.59
France	0.08	0.86 / 1.59	0.74 / 1.22	1.17 / 1.82
Netherlands	0.10	0.06+ / 2.14*	1.40 / 2.07	1.25 / 2.07
United Kingdom	0.00	1.83 / 1.45	1.24 / 1.56	1.01 / 1.29
Italy	0.84	2.14+ / 0.99	1.97+ / 1.38	2.56** / 1.55
Spain	2.13	35.8** / 3.42**	26.2** / 3.10**	19.6** / 2.94**

Inflation equation

Country	Constancy (Chow/Jit)		
	baseline to 93.2	baseline to 94.4	baseline to 96.4
Germany	15.8** / 1.65	10.2** / 1.26	7.01** / 1.15
France	0.38 / 1.65	0.76 / 1.58	0.62 / 1.85
Netherlands	1.99+ / 1.83	1.54 / 2.05+	2.13* / 1.22
United Kingdom	0.36 / 2.06	0.43 / 2.26+	0.35 / 2.53*
Italy	1.43 / 2.91**	1.06 / 2.85**	0.89 / 2.97**
Spain	1.86 / 1.08	1.34 / 1.45	0.99 / 1.46

Long rate equation

Country	Constancy (Chow/Jit)		
	baseline to 93.2	baseline to 94.4	baseline to 96.4
Germany	2.42* / 1.55+	2.82** / 2.11	2.19* / 1.58+
France	0.90 / 1.10	2.87** / 1.55+	1.80+ / 1.27
Netherlands	1.44 / 1.56+	1.83+ / 1.50+	1.26 / 1.72*
United Kingdom	0.50 / 1.79	1.15 / 1.28	0.71 / 1.42
Italy	1.43 / 0.84	4.48 / 0.75	1.26 / 1.23
Spain	1.58 / 1.01	2.56* / 1.58*	1.87* / 1.85*

Note 1: For constancy test at every horizon, we provide first the predictive Chow statistics, second the HANSEN *[1992]'s Jit statistics are respectively a^{+} for significance at 10%, a* for significance at 5%, ** for significance at 1%.*

Note 2: The instrument set used in the auxiliary equation for predetermination test is the following: $x_t, x_{t-1}, \omega_t, \omega_{t-1}$ and g_t (for the world demand, the averaged oil and raw material inflation rate and the fiscal expenditure ratio). The predetermination of the budget ratio s_t in the output gap equation implies a zero restriction on the residual of the restriction is accepted), then the null hypothesis of predetermination is accepted and the balance ratio s_t is predetermined. Wald statistics reported in the table are to compared to a 5% critical value of 4.

Detailed estimates of the behavioral equations for the different countries are not reported, but are available from the authors. On table 3, we summarize results on the constancy of these equations over the three extended, post EMS breakup, samples. A first glance at these results shows that the estimates of behavioral equations are rather more constant than the estimates of policy rules, a feature non supportive of the *Lucas critique*.

Overall evidence of non constancy, mainly from the *Chow* test, is the strongest for Germany, where the three equations are affected.[31] On the opposite, the output and inflation equations are found constant for France, with the interest rate equation mainly affected at the 1994 horizon. The Netherlands present an intermediate pattern, with non constancy affecting the output and inflation equation at the latest horizon. There are interesting similarities between the constancy results for the UK and Italy. Output and inflation equations for those countries suffer from non constancy mainly at the 1994 and 1996 horizons. However, these two main 'leaver' countries have their long rate equations apparently unaffected by the large 1994 shock. Alone with France, Spain have a stable inflation equation, together with a late non constancy in the output equation and in the long rate equation.

Under the testing strategy advocated by ERICSSON and HENDRY [1999], evidence against the relevance of the *Lucas critique* is obtained when, for a given sample period:

i) constancy of the policy rule(s) is rejected,

ii) the conditional behavioral equation is however found invariant.

According to this protocol, results for France are typical of an evidence for the unrelevance of *Lucas critique*: While both monetary and fiscal policies are found non constant at the three horizons of interest, we cannot reject the hypotheses of parameter stability in the output and inflation equations. Symmetrically, results for Germany are supportive of the *Lucas critique*. Other countries deliver less clear-cut messages with Italy closer to the German case, and Spain closer to the French one.

However, following this strategy, we do not use formal invariance tests, but only unspecific ones, derived from the comparison of constancy tests for the behavioral equations and for the policy equations. Let us turn now to our preferred strategy, introduced in the previous section, which rests on more specific tests of weak exogeneity and invariance.

3.3 Exogeneity and Invariance Tests

According to the strategy presented in paragraph 4.1, the equations (9, 10, 11) are augmented, as equation (7), by measures of $(z_{t-1} - \widetilde{z}_{t-i}) = \widehat{z}_{t-i}$, the difference between the realized values of the policy instrument and their preset values according to the rules, and by measures of $\widehat{\widehat{z}}_{t-i} = \widetilde{z}_{t-i} - \overset{\approx}{z}_{t-i}$, the deviations of instrument values expected according to the full sample rule

31. See WEBER [1996] for an empirical analysis of the German economy before and after the unification.

TABLE 4
Augmented Regression Tests of Exogeneity and Invariance

Germany[1]

Equation	Sample	Weak exogeneity			Invariance			Super exogeneity		
		r_t^1	s_t^1	mix[2]	r_t^1	s_t^1	mix[2]	r_t^2	s_t^2	mix[3]
Output gap	93.2	a	a	a	0.017	0.051	0.0000	0.013	a	0.0002
	94.4	a	a	a	0.035	0.037	0.0001	0.039	0.009	0.0003
	96.4	a	a	a	a	0.017	0.026	a	0.046	0.083
Inflation	93.2	a	a	a	0.083	a	a	a	0.075	a
	94.4	a	a	a	0.064	a	a	a	0.062	a
	96.4	a	a	a	a	a	a	a	a	a
Spread	93.2	a	a	a	0.056	a	a	a	0.026	a
	94.4	a	a	a	a	a	a	a	a	a
	96.4	a	a	a	a	a	a	a	a	a

France

Equation	Sample	Weak exogeneity			Invariance			Super exogeneity		
		r_t^1	s_t^1	mix[2]	r_t^1	s_t^1	mix[2]	r_t^2	s_t^2	mix[3]
Output gap	93.2	a	a	a	a	0.07	a	a	0.08	a
	94.4	a	a	a	a	a	a	a	a	a
	96.4	a	a	a	a	a	a	a	a	a
Inflation	93.2	a	a	a	a	a	a	a	a	a
	94.4	a	a	a	a	a	a	a	a	a
	96.4	a	a	a	a	a	a	a	a	a
Spread	93.2	a	a	a	a	a	a	a	a	a
	94.4	0.031	a	0.011	0.002	a	0.005	0.004	0.065	0.002
	96.4	a	0.033	a	0.005	0.073	0.006	0.064	0.046	0.014

Netherlands

Equation	Sample	Weak exogeneity			Invariance			Super exogeneity		
		r_t^1	s_t^1	mix[2]	r_t^1	s_t^1	mix[2]	r_t^2	s_t^2	mix[3]
Output gap	93.2	a	a	a	a	a	0.039	a	a	a
	94.4	a	a	a	a	a	0.07	a	a	0.036
	96.4	a	a	a	a	0.011	0.003	a	0.064	0.005
Inflation	93.2	a	a	a	0.057	0.008	0.025	0.026	0.008	0.021
	94.4	a	a	a	0.008	0.007	0.021	0.003	0.005	0.013
	96.4	a	a	a	0.022	0.049	0.067	0.022	0.076	0.097
Spread	93.2	a	a	a	a	a	a	a	a	a
	94.4	a	a	a	0.03	a	0.049	a	a	a
	96.4	a	a	a	0.01	a	0.005	0.036	a	0.029

Note 1: The colums headed r_t and s_t refer to test relative respectively to the monetary (r_t) and fiscal (s_t) instruments. The colum headed ‘mix’ refers to test relative to both instruments.

1. *The reported numbers are p-values for a rejection of the null hypothesis in a Wald test for one restriction (then, “a” means an acceptance fo the null hypothesis)*
2. *The reported numbers are p-values for a rejection of the null hypothesis in a Wald test with two restrictions.*
3. *The reported numbers are p-values for a rejection of the null hypothesis in a Wald test with four restrictions.*

United Kingdom

Equation	Sample	Weak exogeneity			Invariance			Super exogeneity		
		r_t^1	s_t^1	mix^2	r_t^1	s_t^1	mix^2	r_t^2	s_t^2	mix^3
Output gap	93.2	a	a	a	a	0.078	a	a	a	a
	94.4	a	a	a	a	0.082	0.079	a	a	a
	96.4	a	a	a	a	0.026	0.026	a	0.063	0.043
Inflation	93.2	a	a	a	a	a	a	a	a	a
	94.4	a	a	a	a	a	a	a	a	a
	96.4	a	a	a	a	a	a	a	a	a
Spread	93.2	a	a	a	a	a	a	a	a	a
	94.4	a	a	a	a	0.058	a	0.078	a	a
	96.4	a	a	a	a	a	a	a	a	a

Italy

Equation	Sample	Weak exogeneity			Invariance			Super exogeneity		
		r_t^1	s_t^1	mix^2	r_t^1	s_t^1	mix^2	r_t^2	s_t^2	mix^3
Output gap	93.2	a	a	a	a	a	a	a	a	a
	94.4	a	a	a	a	a	0.015	a	a	0.035
	96.4	a	a	a	0.056	0.061	0.068	0.044	a	0.086
Inflation	93.2	a	a	a	a	a	a	a	a	a
	94.4	a	a	a	0.047	0.026	0.051	a	0.088	a
	96.4	a	a	a	a	a	a	a	a	a
Spread	93.2	a	a	a	a	a	a	a	a	a
	94.4	a	a	a	a	a	a	a	a	a
	96.4	a	a	a	a	a	a	0.099	a	a

Spain

Equation	Sample	Weak exogeneity			Invariance			Super exogeneity		
		r_t^1	s_t^1	mix^2	r_t^1	s_t^1	mix^2	r_t^2	s_t^2	mix^3
Output gap	93.2	a	a	a	0.081	0.0000	0.0000	0.091	0.0000	0.0000
	94.4	a	0.044	0.06	a	0.0022	0.0003	a	0.008	0.0012
	96.4	a	a	a	a	0.0001	0.0002	a	0.0009	0.0017
Inflation	93.2	a	a	a	0.08	a	a	0.062	a	a
	94.4	a	a	a	a	a	0.022	a	a	a
	96.4	a	a	a	0.05	a	0.045			
Spread	93.2	a	a	a	a	a	a	a	a	a
	94.4	a	a	a	0.004	0.002	0.002	0.014	0.02	0.013
	96.4	a	a	0.07	0.003	0.002	0.006	0.008	0.005	0.006

from the ones expected according to the baseline sample estimate.[32] We report in table 4, for the three horizons, tests of weak exogeneity (exclusion of the $\widehat{z}_{t-i}$) of invariance (exclusion of the $\widehat{\hat{z}}_{t-i}$) and superexogeneity (joint restrictions on both the $\widehat{z}_{t-i}$ and the $\widehat{\hat{z}}_{t-i}$). For $z_t = r_t$, the interest rate, and $z_t = s_t$, the surplus ratio, we consider respectively the impact of monetary policy and the one of fiscal policy. The 'policy mix' columns refer to restrictions bearing jointly on the interest rate and the budget surplus variables. The restrictions are evaluated through *Wald* tests. For sake of comparability, only the p-values are reported on the various panels of tables 4, a to f, in cases of rejection of the null.

3.3.1 **The Stayers**

In the *German* case, both the policy rules and the behavioral equations are found non constant.

The output growth equation is found weakly exogenous but not invariant to both monetary and fiscal policy changes, especially at the two longest horizons.

The non constancy of the inflation equation is explained by a lack of invariance with respect to budget policy changes only at the short (1993-2) and intermediate (1994-4) horizons. While the long term rate equation is not constant at all the horizons, it suffers from a lack of invariance only at the 1993-2 horizon due to the monetary policy changes.

In opposition to the German case, the behavioral equations in *France* were found basically constant (except for the long rate equation after 1993). However, tests reject the invariance of the output growth equation with respect to the budget policy changes at the intermediate horizon. This result of a lack of invariance, despite constancy, may look contradictory.

However, the invariance tests are designed to detect deviations from constancy towards a very specific alternative, with respect to which the generic constancy tests may severely lack power.

Evidence on policy changes is milder for the *Dutch* monetary and budget policies, and the evidence against constancy of the behavioral equations is also weak.

However, invariance is generally rejected for the inflation equation and, to a lesser extent, in the output gap and long rate equations. This invariance, while the evidence of domestic policy shift is limited, may be due to a direct contagion of the instability of the German equations.

3.3.2 **The Leavers**

In the *UK*, there is no strong evidence of changes in policy rules, except for the budget in 1993.

However, the invariance tests on the output equation reject the null, mainly due to a lack of invariance w.r.t. changes in budget policy. We also reject invariance of the long rate equation w.r.t. the budget changes at the horizon 1994-4. The inflation equation passes all the tests of exogeneity and invariance.

---

32. Practically, the $\widehat{b}_{iT(t)}$ parameters accounting for policy changes in equation (8) are estimated over the full sample period, and the "constant rule" parameters $\widehat{b_{iT_o}}$ are estimated over the baseline sample period.

The lack of invariance of the output process w.r.t. to budget policy is thus the main result consistent with the *Lucas critique* in the *UK*.

Italy experienced changes in both its monetary and budgetary policies intervening between 1992 and 1995. Evidence on behavioral equations constancy is mitigated.

At the same intermediate horizon, the tests also reject invariance of the inflation equation to both policy shocks.

Super exogeneity of the policy mix in the output equation is rejected for the two longest horizons.

Thus, the Italian experience brings partial evidence in favor of the importance of the *Lucas critique*.

There is evidence of change in the budget rule in *Spain*, and, to a lesser extent, in the monetary rule. The rejection of invariance (hence, of super exogeneity) of the output gap equation w.r.t. budgetary changes is consistent with the predictions from the *Lucas critique*, as well as the changes in inflation equation in response to monetary policy shifts.

The non constancy of the long rate equation is associated to non invariance w.r.t. budget changes at the intermediate and longest horizons.

On the overall, we observe contrasted results, with only a few cases of rejection of the weak exogeneity hypothesis, but a lesser evidence of invariance of the output and inflation equations to policy shifts.

4 Concluding Remarks

The German unification shock, the EMS crisis in 1992 and its aftermath have led to significant changes in policy rules. From the examination of a 6-country sample, it does not seem that these changes were systematically larger in those countries leaving the EMS than in the 'stayers'. Furthermore, evidence of non constancy of three conditional equations, for output gap, inflation and the long term interest rate, is mild, except in Germany.

The proposed strategy of directly testing for invariance using variable addition tests, allowing for lagged conditioning, appears more powerful than the indirect one, which rests only on unspecific constancy tests.

In several cases, invariance is rejected although constancy was not, as the invariance test has more power to detect specific deviations from structural stability.

With respect to the general issue of evaluating the empirical relevance of the *Lucas critique*, our results do not lead to a clear cut conclusion. In nearly half of the cases, significant changes in policy rules fail to induce non constancy and non invariance. Taking ERICSSON and IRONS [1995] as a benchmark, this study also points toward a message of skepticism *vis-à-vis* the empirical relevance of the *Lucas critique*. However, our results using specific invariance tests provide more cases where the *Lucas* argument works. Therefore, we cannot conclude that the *Lucas critique* is a theoretical statement that can be ignored by the empirical model builders. Further work is to be done in order to check if tests using structural models would lead to more decisive conclusions. ▼

• References

ALOGOSKOUFIS G.S., SMITH R.P. (1991). – « The Phillips Curve, the Persistence of Inflation and the *Lucas Critique* », *American Economic Review*, 81, p. 1254-1275.

BALL L. (1998). – « Policy Rules for Open Economies », NBER *Working Paper* 6760.

BLANCHARD O.J. (1984). – « The *Lucas Critique* and the Volcker Deflation », *American Economic Review*, AEA Papers and Proceedings, 74, p. 211-215.

BOHN H. (1995). – « On Testing the Sustainability of Budget Deficits in a Stochastic Environment, CES *Working Paper*, Munchen 91.3.

BRYANT R.C., ZHANG L. (1996). – « Intertemporal Fiscal Policy in Macro-Economic Models: Introduction and Major Alternatives », Brookings *Discussion Papers* in International Economics, n°123.

CARRÉ M. (2000). – « Stabilization with a Deadline », *European Economic Review*, 44, p. 71-90

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

CHRISTIANO L.C., EICHENBAUM M., EVANS C. (1999). – Monetary Policy Shocks: What Have we Learned and to What End?, in J.B. Taylor and M. Woodford, edit, *Handbook of Macroeconomics*, vol. la, chap. II, p. 65-148.

COLLARD F., FEVE P., LANGOT F. (2002). – « Structural Inference and the *Lucas Critique* », *Annales d'Économie et de Statistique*, this issue.

EICHENGREEN B., WYPLOSZ C. (1993). – « The Unstable EMS », *Brookings Papers on Economic Activity*, &, p. 51-143.

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

ERICSSON N., HENDRY D.F., MIZON G. (1998). – « Exogeneity, Cointegration and Economic Policy Analysis », *Journal of Business and Economic Statistics*, 16, p. 370-387.

ERICSSON N., HENDRY D.F. (1999). – « Encompassing and Rational Expectations: How Sequential Corroboration Can Imply Refutation », *Empirical Economics*, 24, p. 1-21.

ERICSSON N., IRONS D.J.S. (1995). – « The *Lucas Critique* in Practice: Theory without Measurement », in K.D. Hoover (Ed.), *Macroeconometrics: Developments, tensions and prospects*, Kluwer, chap. 8, p. 263-312.

ESTRELLA A., FUHRER J.C. (1999). – « Are 'Deep' Parameters Stable? The *Lucas Critique* as an Empirical Hypothesis », *Mimeo*, Federal Reserve Banks of Boston and New York.

FAIR R.C. (1999). – « A Fiscal Policy Rule for Stabilization », *Mimeo* Cowles Foundation, Yale University.

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

FAVERO C. (1993). – « Error Correction and Forward Looking Models for UK Consumption Expenditures », *Oxford Bulletin of Economics and Statistics*, 55 (4), p. 453-472.

FISHER A.M., PEYTRIGNET M. (1991). – « The *Lucas Critique* in Light of Swiss Monetary Policy », *Oxford Bulletin of Economics and Statistics*, 53(4).

FOUNTAS S., PAPAGAPITOS A. (1997). – « Policy Effectiveness in the Post-ERM Era: Evidence from Six Countries, *Open Economies Review*, 8, p. 189-201.

FUHRER J.C. (1996). – « Monetary Policy Shifts and Long Term Interest Rates », *Quarterly Journal of Economics*, p. 1199-1209.

GORDON R.J. (1999). – « The Aftermath of the 1992 Breakup: Was There a Macroeconomic Free Lunch? CEPR *Discussion Paper* n°2073.

GROENEVELD J.M., KOEDIJK K.G., Kool C.J. (1998). – « Credibility of European Economic Convergence », *Weltwirtschaffiches Archiv*, 134, p.1-24.

GROSS D., THYGESEN C.J. (1998). – *European Monetary Integration*, 2nd edition Longman.

HANSEN B.E. (1992). – « Tests for Parameter Instability in Linear Models », *Journal of Policy Modelling*, 14(4), p.517-533.

HENDRY D.F. (1995). – *Dynamic Econometrics*, Oxford University Press.

HOLLY A. (1982). – « A Remark on Haussman's Specification Test », *Econometrica*, 50, p. 749-759.

JUDD J.P., RUDEBUSCH G.D. (1998). – « Taylor's Rule and the Fed », *Economic Review*, Federal Reserve Bank of San Francisco n°3, p. 3-14.

JUDGE G.G., HILL R.C., GRIFFITHS W.E., LÜTKEPOHL H., LEE T.C. (1988). – *Introduction to the Theory and Practice of Econometrics*, Wiley.

KREMERS J. (1989). – « US Federal Indebtness and the Conduct of Fiscal Policy », *Journal of Monetary Economics*, 23, p.219-238.

LINDÉ J. (2000). – « Monetary Policy Analysis in Backward Looking Models », *Discussion paper* presented at the *ADRES-CEPREMAP-CREST conference* on the "Econometrics of Policy Evaluation".

LUCAS R.E. (1976). – « Econometric Policy Evaluation : a Critique », in "The Phillips Curve and Labour Market", *Carnegie-Rochester Conference Series*, 1, p. 19-46.

MEYER B.D. (1985). – « Natural and Quasi-Experiments in Economics », *Journal of Business and Economic Statistics*, 13(2), p. 151-161.

MILLER P.J., ROBERDS W.T. (1991). – « The Quantitative Significance of the *Lucas Critique* », *Journal of Business Economics and Statistics*, 9, p. 361-387.

OLINER S.D., RUDEBUSCH G.D., SICHEL D. (1996). – « The *Lucas Critique* Revisited: Assessing the Stability of Empirical Euler Equations for Investment », *Journal of Econometrics*, 70, p. 291-316.

PEREZ J.J., HIEBERT P. (2001). – « Forecast-error-based Fiscal Policy Rules for Macroeconomic Models », *Mimeo*, European Central Bank.

REVANKAR N.S., YOSHINO N. (1990). – « An Expanded Equation Approach to Weak Exogeneity Tests in Structural Systems and a Monetary Application », *Review of Economics and Statistics*, 72, p. 173-177.

ROSE A., SVENSSON L. (1994). – « European Exchange Rate Credibility before the Fall », *European Economic Review*, 38, p. 1185-1216.

SCHWERT G.W. (1979). – « Tests of Causality: The Message in the Innovations », *Carnegie Rochester Conference Series on Public Policy*, 10, p. 55-96.

SVENSSON L.E.O. (1994). – « Why Exchange Rate Bands? Monetary Independance in Spite of Fixed Exchange Rates », *Journal of Monetary Economics*, 33, p.157-199.

TAYLOR J.B. (1999). – « The Robustness and Efficiency of Monetary Policy Rules as Guidelines for Interest Rate Setting by the European Central Bank, *Journal of Monetary Economics*, 43, p. 655-679.

TAYLOR J.B. (2000). – « Reassessing Discretionary Fiscal Policy », *Mimeo*.

WEBER A.A. (1996). – « Germany before and after Unification: A structural VAR Analysis », *Economic Modelling*, 13, p. 575-601.

APPENDIX

Data Appendix

All data, but a few exceptions indicated below, are taken from the IMF-IFS data base, 1999 edition.

r_t	Money market rate	IFS code 60b
π_t	Consumer price index	IFS code 64
y_t	GDP (HP detrended)	IFS code 99br
e_t	Exchange rate, local money per US	IFS code ee
R_t	Government bonds yield	IFS code 61%
x_t	World demand (weighted indicator)	Source: Banque de France
ω_t	Average of oil and raw material inflation rate	IFS code 00176 axd and aa'z
$s_t^{(a)}$	Global balance, public sector (deflated by GDP)	FERI database
$b_t^{(a)}$	Public debt (deflated by GDP)	FERI database

(a) The FERI indicators of public debt and surplus have be retained, preferably to their IFS counterpart, due to their broader coverage.

Appendix I: On Weak Exogeneity with Lagged Conditioning

Due to assumed -contractual or technological- rigidities, the effect of $\widetilde{z}_{t-1}$ on y_t in equation (4) cannot be revised in $t-1$, after the observation of the actual realization of z_{t-1}. Therefore, we are led to extend the set of parameters of interest beyond α_2 – *ie,* the contemporanous conditioning – to α_3 – *ie,* the parameter of lagged conditioning. In order to properly redefine the weak exogeneity condition, we need to reformulate the underlying factorization of the relevant joint distribution as the product of a conditional process and a marginal one.

From equations (4) and (3), in the main text, we derive the following reduced form equation for y_t:

$$\text{(A1)} \quad \begin{aligned} y_t = {} & [\alpha_o + (a_2 + \alpha_3)\, b_o] + (\alpha_1 + \alpha_2 b_1)\, y_{t-1} \\ & + \alpha_3 b_1 y_{t-2} + \alpha_2 b_2 z_{t-1} + \alpha_3 b_2 z_{t-2} + \nu_t \end{aligned}$$

Equation (A1) fully characterizes the process of y_t given the process of z_t and the law of ν_t. However, considered isolately, it does not allow for a proper inference on the value of the parameters of interest α_2 and α_3, without knowledge of the marginal process of z_t and z_{t-1}. Therefore, the z_{t-i} are not weakly exogenous for α_2 and α_3 in equation (A1).

Introducing $y_{1t} \equiv y_{t-1}$, $z_{1t} \equiv z_{t-1}$, as indicated in footnote 25, we get the matrix representation of the joint distribution:

$$
\text{(A2)}\quad \begin{pmatrix} y_t \\ y_{1t} \\ z_t \\ z_{1t} \end{pmatrix} = \begin{pmatrix} \alpha_o + b_o(\alpha_2+\alpha_3) \\ 0 \\ b_o \\ b_o \end{pmatrix} + \begin{pmatrix} \alpha_1+\alpha_2 b_1 & \alpha_3 b_1 & \alpha_2 b_2 & \alpha_3 b_3 \\ 1 & 0 & 0 & 0 \\ b_1 & 0 & b_1 & 0 \\ 0 & b_1 & 0 & b_2 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ y_{1,t-1} \\ z_{t-1} \\ z_{1,t-1} \end{pmatrix} + \begin{pmatrix} \nu_t \\ 0 \\ \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix}
$$

where the (ν_t, ε_t) Gaussian disturbances are assumed both contemporeanously and serially uncorrelated. Conditionning y_t w.r.t. z_t and z_{t-1}, one get the following structural representation:

$$
\text{(A3)}\quad \begin{pmatrix} 1 & 0 & -\alpha_2 & -\alpha_3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_t \\ y_{1t} \\ z_t \\ z_{1t} \end{pmatrix} = \begin{pmatrix} \alpha_o \\ 0 \\ b_o \\ b_o \end{pmatrix} + \begin{pmatrix} \alpha_1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ b_1 & 0 & b_2 & 0 \\ 0 & b_1 & 0 & b_2 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ y_{1,t-1} \\ z_{t-1} \\ z_{1,t-1} \end{pmatrix} + \begin{pmatrix} \nu_t - \alpha_2\varepsilon_t - \alpha_3\varepsilon_{t-1} \\ 0 \\ \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix}
$$

where the first equation defines the linear projection of y_t on z_t and z_{t-1}, given y_{t-1} and z_{1t-1} and the other equations define the marginal model for the regressors.

The disturbance of the first equation depends upon the marginal disturbances ε_t and ε_{t-1}, and this equation thus does not provide a proper conditional model. Inference on (α_2, α_3) from the regression equation (5) will be invalid due to the correlation between regressors and the disturbance terms.

To show how this failure of predetermination of z_t and z_{t-1} involves a failure in weak exogeneity, let us consider the textbook expression (*eg*, JUDGE *et alii* [1988, p. 609] of the simultaneity bias between the parameters of interest and their least square estimates $(\widehat{a}_2, \widehat{a}_3)$ on the first equation in the system (A3):

$$
\text{(A4)}\quad \begin{pmatrix} \alpha_2 \\ \alpha_3 \end{pmatrix} = \operatorname{plim} \begin{pmatrix} \widehat{a}_2 \\ \widehat{a}_3 \end{pmatrix} - \begin{pmatrix} \upsilon_{zz} & \upsilon_{zz_1} \\ \upsilon_{zz_1} & \upsilon_{zz} \end{pmatrix} \cdot \begin{pmatrix} -(\alpha_2 + b_2\alpha_3)\,\sigma_\varepsilon^2 \\ -\alpha_3\,\sigma_\varepsilon^2 \end{pmatrix}
$$

where the υ_{ij} are the (ij) terms of the inverse of the variance-covariance matrix of the regressors present in equation (A3-eq1), *ie*, z_t, z_{t-1} and y_{t-1}, and $\left[-(\alpha_2 + b_2\alpha_3)\,\sigma_\varepsilon^2, -\alpha_3\,\sigma_\varepsilon^2\right]$ are the asymptotic covariances of the disturbance with respectively z_t and z_{t-1}. This bias depends upon the marginal process of z_t, not only through σ_ε^2 but also from the whole unconditional variance-covariance structure of z_t with z_{t-1}, and y_{t-1}. Thus, expression A4 explicitly shows that (z_t, z_{t-1}) are not weakly exogenous for (α_2, α_3) in (A3, eq.1).

In case of a non constant marginal model, (equation 3 in the main text), and using the following decomposition for time varying parameters $b_{it} = \overline{b}_i + \widehat{b}_{it}, \sigma_t = \overline{\sigma} + \widehat{\sigma}_t,$ we get:

$$\begin{array}{rcl} \widetilde{z}_t - z_t & = & \varepsilon_t \\ \underset{t}{\overset{\approx}{z}} - z_t & = & \widehat{b}_{ot} + \widehat{b}_{1,t} y_{t-1} + \widehat{b}_{2,t} z_{t-1} + \widehat{\sigma}_t \varepsilon_t \end{array}$$

and the vector representation stacking the conditional equation for y_t and the marginal model becomes:

$$\begin{pmatrix} 1 & 0 & -\alpha_2 & -\alpha_3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_t \\ y_{1t} \\ z_t \\ z_{1t} \end{pmatrix} = \begin{pmatrix} \alpha_o + (\alpha_o + \alpha_3) \widehat{b}_{ot} \\ 0 \\ b_{ot} \\ b_{o,t-1} \end{pmatrix}$$

$$\text{(A5)} \qquad + \begin{pmatrix} \alpha_1 + \alpha_2 \widehat{b}_{1t} & \alpha_3 \widehat{b}_{1t} & \alpha_2 \widehat{b}_{1t} & \alpha_3 \widehat{b}_{2t} \\ 1 & 0 & 0 & 0 \\ b_{1t} & 0 & b_{2t} & 0 \\ 0 & b_{1,t-1} & 0 & b_{2,t-1} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ y_{1,t-1} \\ z_{t-1} \\ z_{1,t-1} \end{pmatrix}$$

$$+ \begin{pmatrix} \nu_t - \alpha_2 \widehat{\sigma}_t \varepsilon_t + \alpha_3 \widehat{\sigma}_{t1} \varepsilon_{t-1} \\ 0 \\ \sigma_t \varepsilon_t \\ \sigma_{t-1} \varepsilon_{t-1} \end{pmatrix}$$

which illustrates the lack of invariance of the conditional regression equation (A3-eq1) for interventions on $\widehat{b}_{it}$ and $\widehat{\sigma}_t$.

Appendix II: The Jit Test

The *Jit* (Joint (in)stability test), introduced by HANSEN [1992] is based upon the cumulative first order conditions (the scores in a maximum likelihood context):

$$S_{it} = \sum_{j=1}^{t} f_{ij}, \qquad \forall\, t = 1 \ldots T$$

where $i = 1 \ldots m$ are the scores associated to the m stochastic regressors and $i = m+1$ the condition associated to the variance of disturbances.

The *Jit* statistic is $L_c = T^{-1} \sum^{T} \widehat{S}_t' \widehat{V}^{-1} \widehat{S}_t$ where $\widehat{S}_t$ is the matrix of $\widehat{S}_{it}$ vectors, the empirical counterpart of the S_{it}, and $V = \sum_{t=1}^{T} \widehat{f}_t \widehat{f}_t'$.

The *Jit* test exploits the property that the alternative of non constancy entails a failure of the orthogonality between regressors and disturbances over certain subsamples.