Testing for the Lucas Critique: A Quantitative Investigation

Jesper Lindé*

May 25, 2000

Abstract

In this paper, I try to shed some new light on the "puzzle" why the Lucas critique, believed to be important by most economists, seems to have received very little empirical support. I use a real business cycle model to verify that the Lucas critique is quantitatively important in theory, and to examine the properties of the super exogeneity test, which is used to detect the applicability of the Lucas critique in practice. The results suggest that the super exogeneity test is not capable of detecting the relevance of Lucas critique in practice in small samples.

Keywords: Lucas critique; real business cycle model; super exogeneity test; Taylor rules; money demand; consumption function.

JEL Classification Numbers: C52, C22, E41.

*Research Department, Sveriges Riksbank (the Swedish cental bank). *Address*: SE-103 37 Stockholm, Sweden. *E-mail*: jesper.linde@riksbank.se. I would like to thank Lawrence Christiano, Tor Jacobson, Søren Johansen, David Hendry, Eric Leeper, Lars Ljungqvist, Paul Söderlind and Anders Vredin for very helpful comments and discussions. I have also benefitted from comments by seminar participants at the Stockholm School of Economics, the Trade Union Institute for Economic Research (FIEF), the conference on Macroeconomic Transmission Mechanisms at NTNU in Trondheim, and the Econometric Society European Meeting 1999 in Santiago de Compostela. The views expressed in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Executive Board of Sveriges Riksbank. Financial support from the *Tore Browaldh Foundation* and the *Jan Wallander and Tom Hedelius Foundation* is gratefully acknowledged.

1 Introduction

In a very influential article, Robert E. Lucas, Jr., (1976) raised a serious critique against econometric models that were used for policy evaluation. Lucas's argument was that shifts in economic policy change how policy affects the economy since agents in the economy are forward rather than backward-looking and adapt their expectations and behavior to the new policy stance. Thus, past behavior can be a poor guide for assessing the effects of policy actions. For this reason, Lucas concluded that reduced-form econometric models cannot provide useful information about the actual consequences of alternative policies because the structure of the economy will change when policy changes, thereby rendering the estimated parameters in reduced-form econometric models nonconstant.

Instead, Lucas (1975, 1977), Finn E. Kydland and Edward C. Prescott (1982) and others initiated a new research program, often termed the real business cycle (or equilibrium business cycle) approach, where the models used for policy analysis are immune to the Lucas critique in that they are equilibrium models with forward-looking behavior. Other researchers were concerned about the applicability of the Lucas critique in practice; see e.g. the discussion in Christopher A. Sims (1982). Robert F. Engle, David F. Hendry and Jean-Francois Richard (1983) introduced the concept of super exogeneity, and argued that it could be used to test for the empirical relevance of the Lucas critique. Subsequent papers (e.g. Engle and Hendry, 1993) have shown how this concept can be applied.

Recently, the empirical relevance of the Lucas critique has received increased attention. A possible explanation for this is the extensive use of backward-looking models in monetary policy analysis; cf. Laurence Ball (1997), Lars E.O. Svensson (1997), Glenn D. Rudebusch and Svensson (1998) and John B. Taylor (1999). Jeffrey C. Fuhrer (1997) maintains that backward-looking behavior seems to be a better approximation of reality in macroeconomic models than forward-looking behavior. Arturo Estrella and Fuhrer (1999) argue that the Lucas critique is an empirically testable hypothesis. They provide evidence that when there is a change in monetary policy regime, some forward-looking models may be less stable than their better fitting backward-looking counterparts, which they contend is an observation inconsistent with the Lucas critique. In addition, most - if not all - of the many papers which have used the concept of super exogeneity to examine the Lucas critique empirically have found no evidence in favor of the proposition; see the survey by Neil R. Ericsson and John S. Irons (1995).

Two natural questions then arise. Is it obvious that the Lucas critique is quantitatively important in theory in a statistical sense? In other words, even if the Lucas critique is theoretically important qualitatively, it might not be so quantitatively in a statistical context.[1] Second, given that the answer to the first question is yes (or no), does the test for super exogeneity work in the sense that it really reveals the presence (or absence) of the Lucas critique in observed data?

In this paper, I examine these two questions in an attempt to shed some new light on the "puzzle": why is it that the Lucas critique, although regarded as highly important, does not seem to be important in studies of real-world data? My approach is to set up a version of Thomas F. Cooley and Gary D. Hansen's (1995) real business cycle model with money, modified to include government expenditures and a Taylor-type policy rule (see Taylor, 1993) for nominal money growth similar to that analyzed by Benneth T. McCallum (1984, 1988). This policy rule is then estimated on U.S. data for the recent periods in office of Federal Reserve's chairmen Burns, Volcker, and Greenspan. According to John P. Judd and Rudebusch (1998), the conduct of monetary policy has varied systematically between these periods.[2]

I study the dependence of two relationships on the monetary policy rule. The first relationship is a traditional money demand function (similar to those analyzed in Stephen M. Goldfeld and Daniel E. Sichel, 1990) and the second is a "Keynesian" consumption function. An important reason for focusing on money demand and consumption is that most empirical studies in the field have applied the super exogeneity test to these relations; see Ericsson and Irons (1995).

To investigate the theoretical applicability of the Lucas critique for these relationships, I begin by deriving analytical solutions for the money demand and consumption

[1] For a discussion along this line, see Eric M. Leeper's (1995) comments on the paper by Ericsson and Irons (1995) .

[2] Judd and Rudebusch (1998) start out by noting that there is instability in the Fed reaction function. They then find support for the hypothesis that the Fed monetary policy rule has varied systematically with the different periods in office of Fed chairmen Burns, Volcker, and Greenspan. As in their analysis, chairman Miller is omitted here because of his very short tenure.

functions to see whether the parameters are dependent on the monetary policy rule. The equilibrium model is then calibrated with different estimated monetary policy regimes to study whether the properties of the estimated money demand and consumption functions change significantly when there is a monetary regime shift. This is done using a simple Gregory C. Chow (1960) test.

Finally, I examine the small sample properties of the super exogeneity test. A test of super exogeneity is applied on money demand/consumption along with the monetary policy rule by means of a Monte-Carlo simulation. The purpose is thus to check whether the test is actually able to identify the relevance/nonrelevance of the Lucas critique in the model economy.

One possible reason why the test may have failed to detect parameter instability in behavioral relationships, despite presence of parameter instability in policy rules, is that the effects of changes in policy rules are very difficult to distinguish from the other shocks that hit the economy. However, in the general equilibrium framework used here, it is in fact possible to control for these effects by "going back in time" and performing the super exogeneity test conditional on all other shocks except for the change in the monetary policy rule.

The results in the paper are as follows. First, it is demonstrated that in the equilibrium model, the parameters in the money demand and consumption functions are functions of the parameters in the monetary policy rule.[3] Consequently, the Lucas critique is, as expected, at least qualitatively important in the model. Second, with a standard parameterization of the model and by considering the estimated Federal Reserve monetary policy rules for nominal money growth during Burns, Volcker and Greenspan office periods, it is found that the Lucas critique is theoretically important in a statistically significant way. When the parameters in the estimated Taylor-type monetary policy rule change, the estimated money demand and consumption functions display considerable parameter

[3] As regards money demand, the results both contradict and support the proposition in Lucas (1988) that money demand is a structural relation (invariant to policy parameters). The "true" money demand function in the general equilibrium model is independent of parameters in the monetary policy rule as suggested by Lucas. However, the "true" money demand function derived in the model is not in the form typically estimated by economists. Moreover, when it is rewritten in a traditional form (as in Goldfeld and Sichel, 1990), the reduced form parameters turns out to be dependent on the monetary policy rule. Robert G. King (1988) discusses this issue.

instability from both a statistical and an economic point of view. I thus have a model where the super exogeneity test should be able to identify the empirical relevance of the Lucas critique.

Despite this, it is found that the super exogeneity test far too often fails to reject the false null hypothesis that the Lucas critique does not apply when there is a change in the conduct of monetary policy. This lack of power for the super exogeneity test is then, quite naturally, a possible explanation as to why the Lucas critique has not been found in the data.

The findings of this paper also have some general implications for empirical testing of the relevance of backward versus forward-looking models. First, only the truly forward-looking model will have parameters invariant to the monetary regime. Thus, the preliminary results in Estrella and Fuhrer (1999) suggesting that the Lucas critique is not important in practice may be due to model misspecification. Second, these types of tests will presumably have very weak power in small samples, as is the case for the super exogeneity test.

The structure of the paper is as follows. In the next section, I introduce the theoretical model, and indicate how to compute the equilibrium. Estimation and calibration issues are addressed in Section 3. In Section 4, I derive and discuss the theoretical properties of the money demand and consumption functions used throughout the paper. Next, in Section 5, I examine whether the Lucas critique is significantly important in the equilibrium business cycle model by testing for parameter stability between the different estimated monetary policy regimes on simulated data from the model. The super exogeneity test is briefly presented in Section 6, along with the results of the Monte-Carlo simulations. Section 7 concludes.

2 The equilibrium model

In this section, I describe and solve a slightly modified version of Cooley and Hansen's (1989, 1995) monetary equilibrium business cycle model. The model is a standard real business cycle model with some additional features. A stochastic nominal money supply interacts with a cash-in-advance technology and one-period nominal wage contracts, which creates short run real effects of nominal money supply shocks. Originally, Cooley and Hansen (1995) calibrated the model to quarterly data. Here, one period is one year. The reason for this is that I think it is reasonable that real effects on monetary policy within the model last for a year rather than a quarter, which is an implication of the one-period nominal wage contract setting.[4]

The difference between the model in this paper and the one in Cooley and Hansen (1995) is that the central bank is here assumed to use a policy rule when it decides on the nominal money supply growth in each period similar to that suggested by McCallum (1984, 1988). More specifically, the growth rate in nominal money supply in period t is assumed to follow a Taylor-type policy rule and depend on the output gap, the difference between actual and targeted inflation rate (hereafter named inflation gap), an uncontrollable shock, and the growth rate in nominal money in period $t-1$. This specification is supposed to capture the real world phenomenon that central banks use money supply to affect inflation and output gaps, although they act gradually and do not have perfect control of the process. It is shown that this monetary policy rule for nominal money growth has can be rewritten as a Taylor rule for the nominal interest rate.

In the model I abstract from population and technological growth and represent all variables in per capita terms.

Finally, a notational comment; in the following, capital letters denote economy wide averages which the agent takes as given and small letters individual specific values which the agent internalizes.

2.1 An equilibrium monetary business cycle model

Infinitely many identical infinitely lived agents maximize expected utility with preferences summarized by

$$\mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t u\left(c_{1t}, c_{2t}, h_t\right), \tag{1}$$

$$u\left(c_{1t}, c_{2t}, h_t\right) \equiv \alpha \ln(c_{1t}) + (1-\alpha)\ln\left(c_{2t}\right) - \gamma h_t$$

[4] I would like to emphasize that the qualitative aspects of the results in the paper are not at all dependent on whether I calibrate the model to match quarterly or yearly data.

where c_{1t} is consumption of the "cash good" in period t, c_{2t} is consumption of the "credit good," and h_t is the share of available time spent in employment which enters linearly in (1) because of the "indivisible labor" assumption (see Hansen, 1985). In (1), β is the subjective discount factor, γ the disutility the agent gets from working, while α reflects the trade-off between consumption of the cash and credit goods.

The flow budget constraint facing the agent is

$$c_{1t} + c_{2t} + i_t + \frac{m_{t+1}}{P_t} + \frac{b_{t+1}}{P_t} = \left(\frac{W_t^c}{P_t}\right) h_t + R_t^K k_t + \frac{m_t}{P_t} + (1 + R_{t-1}) \frac{b_t}{P_t} + \frac{TR_t}{P_t} \tag{2}$$

where i_t denotes the agent's investment, m_{t+1} and b_{t+1} the agent's holdings of nominal money and government bonds in the end of period t, P_t the aggregate price level, W_t^c the contracted nominal wage, R_t^K the gross real return on the capital stock k_t, R_{t-1} the nominal interest rate on government bonds in between period $t-1$ and t, and TR_t nominal lump-sum transfers (or taxes if negative) from the government.

The agent has the following cash-in-advance constraint for the cash-good c_{1t},

$$P_t c_{1t} = m_t + (1 + R_{t-1})\, b_t + TR_t - b_{t+1} \tag{3}$$

which always holds with equality since the nominal interest rate will always be positive in this model.

The government's budget constraint is

$$P_t G_t + TR_t = M_{t+1} - M_t + B_{t+1} - (1 + R_{t-1}) B_t \tag{4}$$

where G is exogenous public consumption expenditures, and M and B aggregate nominal money supply and government bonds. As in Cooley and Hansen (1995), I will assume that $B_t = 0$ for $t \geq 0$ and only use it to compute the nominal interest rate in the economy. It can be shown that the nominal interest rate in equilibrium is given by

$$R_t = \frac{\alpha}{1-\alpha} \frac{C_{2t}}{C_{1t}} - 1 \tag{5}$$

where C_{1t} and C_{2t} are aggregate consumption of the cash and credit goods, respectively.

Government consumption, G, in (4) is assumed to be generated by the following stationary AR(1)-process,

$$\ln G_{t+1} = (1 - \rho_{\ln G}) \ln G + \rho_{\ln G} \ln G_t + \varepsilon_{\ln G, t+1} \tag{6}$$

where $0 < \rho_{\ln G} < 1$ and $\varepsilon_{\ln G} \sim i.i.d.\ N\left(0, \sigma^2_{\ln G}\right)$.

Aggregate nominal money supply is assumed to evolve according to

$$M_{t+1} = e^{\mu_t} M_t \tag{7}$$

where the growth rate in nominal money supply in period t, defined as $\Delta \ln M_{t+1}$ and denoted μ_t, is assumed to be determined by

$$\begin{aligned} \mu_t &= \eta \mu_{t-1} - \lambda_\pi \left(\pi_t - \pi^*\right) - \lambda_Y \left(\ln Y_t - \ln Y^*\right) + \xi_t,\ 0 < \eta < 1, \\ \xi &\sim i.i.d.\ \textit{Log Normal},\ \mathrm{E}[\xi] = (1-\eta)\mu,\ \mathrm{Var}(\xi) = \sigma^2_\xi \end{aligned} \tag{8}$$

where π_t is defined as $\ln P_t - \ln P_{t-1}$, and λ_π and λ_Y measure how the central bank reacts to deviations in the inflation $(\pi_t - \pi^*)$ and the output gap $(\ln Y_t - \ln Y^*)$.[5] The implicit assumption underlying the specification in (8) is that the central bank tries to stabilize inflation and/or output, and one might think of (8) as an implied monetary policy rule for a central bank which has been attached a conventional quadratic loss function in the inflation and output gaps. For simplicity, we set π^* and $\ln Y^*$ in (8) equal to steady state (dropping time subscripts) nominal money supply growth (μ) and log of output ($\ln Y$), respectively. I think of the error term, ξ, as policy shocks from the perspective of the private sector. By introducing the persistence component $\eta\mu_{t-1}$, I assume that the central bank reacts gradually to shocks which hits the economy.[6]

The production function is assumed to have constant returns to scale and be of Cobb-Douglas type

$$Y_t = e^{\ln Z_t} K_t^\theta H_t^{1-\theta} \tag{9}$$

[5] Although we assume that ξ is log normally distributed, we require that ξ has mean $(1-\eta)\mu$, and variance σ^2_ξ as seen in (8). By using that $\mathrm{E}[\xi] = e^{\mathrm{E}[\ln \xi] + \frac{1}{2}\mathrm{Var}(\ln \xi)}$ and that $\mathrm{Var}(\xi) = \mathrm{E}\left\{(\xi - \mathrm{E}[\xi])^2\right\} = \mathrm{E}\left[\xi^2\right] - \left[(1-\eta)\mu\right]^2 = e^{2\mathrm{E}[\ln \xi] + \mathrm{Var}(\ln \xi)} - \left[(1-\eta)\mu\right]^2$ since ξ is log-normally distributed, one can pin down the mean and the variance for $\ln \xi$ as $-\frac{1}{2}\ln\left(\sigma^2_\xi + \left[(1-\eta)\mu\right]^2\right) + 2\ln\left((1-\eta)\mu\right)$ and $\ln\left(\sigma^2_\xi + \left[(1-\eta)\mu\right]^2\right) - 2\ln\left((1-\eta)\mu\right)$ respectively.

[6] The Taylor-type rule is normally specified in terms of the nominal interest rate, see for instance Taylor (1998). Ignoring government expenditures ($G_t = 0$), and log-linearizing (5) and (18), it is possible to derive the following Taylor-type rule within the model for the nominal interest rate $R_t = -\frac{\lambda_\pi \pi^* + \lambda_Y \ln Y^*}{\kappa_3} + (1 + \eta - \eta \mathrm{L}) R_{t-1} + \frac{1 + \lambda_\pi - \eta \mathrm{L}}{\kappa_3}\pi_t + \frac{\lambda_Y + (1+\eta \mathrm{L})(1-\mathrm{L})\left(1 + \delta \frac{K}{Y}\right)}{\kappa_3} \ln Y_t + \varepsilon^R_t$ where $\varepsilon^R_t \equiv \frac{\xi_t - (1-\eta \mathrm{L})(1-\mathrm{L})\delta \frac{K}{Y} \ln I_t}{\kappa_3}$, $\kappa_3 = \frac{C - C_1}{C} > 0$ since $C \equiv C_1 + C_2$ and L is the lag operator. Thus, it is possible to transform the Taylor inspired rule for μ to a standard rule for R in the model, although we can no longer treat the error term as exogenous. But here it only possible to use the rule for μ_t, since R_t is an endogenous equilibrium relative price.

where K_t and H_t are aggregate (average) capital stock and hours worked, respectively, and Z_t is the technology level. The perfect competition zero profit maximizing conditions for the representative firm are

$$W_t^c = (1-\theta)\, e^{\ln Z_t} \left(\frac{K_t}{H_t}\right)^{\theta} P_t \tag{10}$$

and

$$R_t^K = \theta e^{\ln Z_t} \left(\frac{K_t}{H_t}\right)^{\theta-1}. \tag{11}$$

The nominal wage W_t^c is assumed to be set in period $t-1$ (see Cooley and Hansen (1995) for further details on the nominal wage arrangement) as

$$W_t^c = (1-\theta)\, e^{\mathrm{E}_{t-1} \ln Z_t} \left(\frac{K_t}{\mathrm{E}_{t-1} H_t}\right)^{\theta} \mathrm{E}_{t-1} P_t \tag{12}$$

where the capital stock in period t is known in $t-1$. If we combine (10) and (12) in natural logarithms, using (15) below, we obtain

$$\ln H_t = \mathrm{E}_{t-1} \ln H_t + \frac{1}{\theta} \left(\ln P_t - \mathrm{E}_{t-1} \ln P_t\right) + \frac{1}{\theta} \varepsilon_t^{\ln Z}. \tag{13}$$

Similarly, one realizes that the natural logarithm of h_t for an agent in equilibrium is given by

$$\ln h_t = \mathrm{E}_{t-1} \ln H_t + \frac{1}{\theta} \left(\ln P_t - \mathrm{E}_{t-1} \ln P_t\right) + \frac{1}{\theta} \varepsilon_t^{\ln Z}. \tag{14}$$

In (9), the technology level is assumed to be exogenous and the natural log of it to follow a stationary AR(1)-process

$$\ln Z_{t+1} = \rho_{\ln Z} \ln Z_t + \varepsilon_{\ln Z, t+1},\ \varepsilon_{\ln Z} \sim i.i.d.\ N\left(0, \sigma^2_{\ln Z}\right). \tag{15}$$

Individual and aggregate investment in period t produces productive capital in period $t+1$ according to

$$k_{t+1} = (1-\delta)\, k_t + i_t \tag{16}$$

and

$$K_{t+1} = (1-\delta)\, K_t + I_t \tag{17}$$

where δ is the rate of capital depreciation.

The aggregate resource constraint

$$Y_t = C_{1t} + C_{2t} + I_t + G_t \equiv C_t + I_t + G_t \tag{18}$$

also holds in every period where C_t is total consumption.

2.2 Equilibrium in the model

The equilibrium in the model consists of a set of decision rules for the agents $\ln k_{t+1} = k(\mathbf{S}_t,\ \ln k_t,\ \ln \hat{m}_t)$, $\ln \hat{m}_{t+1} = \hat{m}(\mathbf{S}_t,\ \ln k_t,\ \ln \hat{m}_t)$ and $\ln h_t = h(\mathbf{S}_t,\ \ln k_t,\ \ln \hat{m}_t)$, and a set of aggregate decision rules $\ln K_{t+1} = K(\ \mathbf{S}_t)$, $\ln H_t = H(\mathbf{S}_t)$, $\ln \hat{P}_t = \hat{P}(\mathbf{S}_t)$ where $\mathbf{S}_t = \left[\ln Z_{t-1}, \varepsilon_t^{\ln Z}, \mu_{t-1}, \xi_t, \ln G_t, \ln K_t, \ln \hat{P}_{t-1}\right]'$ such that; (i) agents maximize utility, (ii) firms maximize profits, and (iii), individual decision rules are consistent with aggregate outcomes. Equilibrium condition (iii) implies that $k(\mathbf{S}_t,\ \ln K_t,\ 1) = K(\ \mathbf{S}_t)$, $\hat{m}(\mathbf{S}_t,\ \ln K_t,\ 1) = 1$, and $h(\mathbf{S}_t,\ \ln K_t,\ 1) = H(\mathbf{S}_t)$ for all $\mathbf{S}_t$.

I use the same method as Cooley and Hansen (1989, 1995) (described in detail by Hansen and Prescott, 1995) to compute the equilibrium in this model.[7]

3 Estimation and calibration

The parameters in the model are determined in two ways. About half of the parameters (η, μ, σ_ξ^2, λ_π, λ_Y, $\rho_{\ln G}$, $\sigma_{\ln G}^2$ and $g \equiv \frac{\bar{G}}{\bar{Y}}$) are estimated on U.S. data 1960-1997 with Instrumental Variables method (IV) and Maximum Likelihood (ML). The other half of the parameters (α, β, δ, γ, θ, $\rho_{\ln Z}$ and $\sigma_{\ln Z}^2$) are adapted from Cooley and Hansen (1995), and chosen so that the model's steady state properties are consistent with U.S. growth facts.

To estimate the parameters η, μ, σ_ξ^2, λ_π, and λ_Y in the monetary policy rule (8) for different Fed chairmen periods, I collected quarterly data on real gross national product per capita in natural logarithms ($\ln Y_t$), growth rate in nominal money supply (μ_t) and the inflation rate in the consumer price index (π_t). The reason for collecting quarterly rather than yearly data is to get more observations in each subsample. To compute measures of $\ln Y_t - \ln Y^*$ and $\pi_t - \pi^*$, I simply filtered the series for output and inflation rate with the Hodrick-Prescott (H-P) filter (see Hodrick and Prescott, 1997).[8] It is standard to use H-P filtered output as a measure of the output gap, but is less clear how to compute an appropriate measure of π^* from historical data as discussed by Judd and Rudebusch

[7] More details on how to compute the equilibrium is provided in Appendix B which is available on request from the author.

[8] I use the common value 1600 (quarterly data) for the smoothness coefficient λ in the H-P filter. See Appendix A for a detailed description of the raw data and data transformations.

(1998).[9] Since the model does not distinguish between money controlled by the Fed (the monetary base, M0) and money used in the private transactions (M2), I compromise between them and use M1 as a measure of money as in Cooley and Hansen (1989, 1995). The reason for estimating with IV rather than Ordinary Least Squares (OLS), is that OLS is likely to be a biased and inconsistent estimator due to the fact that we may have contemporaneous correlation between the error term and the regressors in (8). In terms of the theoretical model used in this paper, there will, via the equilibrium decision rules, be a positive correlation between the error term ξ_t and the regressors π_t and $\ln Y_t$ in (8). As instruments in the estimation, I therefore use $(\ln Y - \ln Y^*)_{t-1}$, μ_{t-1} and $(\pi - \pi^*)_{t-1}$ which are uncorrelated with the error term ξ_t in (8). In addition to that, the estimated λ_π and λ_Y will be correlated in general, which also makes inference about individual parameters tricky.

I estimate the monetary policy rule (8) with IV for the whole sample period ($1960Q2 - 1997Q4$), for chairman Burn's office period ($1970Q1 - 1978Q1$), chairman Volcker's office period ($1979Q3 - 1987Q2$), chairman Greenspan's office period ($1987Q3 - 1997Q4$), and omit chairman Miller as in Judd and Rudebusch (1998) because of his short tenure. The results of the estimations are reported in Table 1 (a constant is included in the regressions but is not spelled out).

Table 1: IV estimation results for the monetary policy rule (8).

Estimation	Estimation output								
period	$\hat{\eta}$	$\hat{\lambda}_\pi$	$\hat{\lambda}_Y$	$\hat{\sigma}_\xi$	$\bar{R}^2$	D-W	B-G χ^2 (4)	J-B	T
Whole	0.930 (0.028)	0.204 (0.083)	0.068 (0.081)	0.0127	0.89	1.34	42.19 (0.000)	0.596 (0.742)	151
Burns	0.514 (0.151)	0.182 (0.087)	−0.166 (0.148)	0.0073	0.73	1.96	14.17 (0.007)	0.715 (0.699)	33
Volcker	0.717 (0.116)	0.377 (0.203)	−0.137 (0.249)	0.0158	0.73	1.59	11.86 (0.019)	1.289 (0.525)	32
Greenspan	0.919 (0.054)	0.532 (0.540)	0.013 (0.262)	0.0153	0.91	0.98	17.37 (0.002)	1.249 (0.536)	42

Note: Standard errors in parenthesis for $\hat{\eta}$, $\hat{\lambda}_\pi$ and $\hat{\lambda}_Y$, and small sample adjusted p-values in parenthesis for the Breusch-Godfrey autocorrelation test (null hypothesis no autocorrelation up to 4 lags) and the Jarque-Bera normality test (null hypothesis normally distributed residuals). A constant, $(\ln Y - \ln Y^*)_{t-1}$, μ_{t-1} and $(\pi - \pi^*)_{t-1}$ have been used as instruments. T denote the number of observations in the regressions.

[9] Although my approach regarding $\pi - \pi^*$ appears to be as good as any other considerable alternative (see Judd and Rudebusch), I have nevertheless experimented with other measures (such as the average inflation rate during a given chairmen's term), but it did not have any impact on the conclusions drawn in the paper.

From Table 1, the implied yearly estimates of η for the different periods are approximately 0.751 (0.930^4), 0.070, 0.264 and 0.713 respectively. For σ_ξ, the implied yearly estimates are $\sqrt{\hat{\sigma}_\xi^2 \left(1+\hat{\eta}^2+\hat{\eta}^4+\hat{\eta}^6\right)}$.

The D-W and Breusch-Godfrey statistics indicates presence of positive autocorrelation in the regressions, suggesting difficulties to interpret the significance levels of the estimates of η, λ_π and λ_Y. However, use of the asymptotic χ^2-distribution for the Breusch-Godfrey test is very likely to yield an oversized test (i.e. an exaggerated probability of rejecting a true null hypothesis of no autocorrelation) for sample sizes as small as the present ones. Simulated small sample adjusted p-values for the Breusch-Godfrey test confirm the size problem, and result in a non-significant autocorrelation effect.[10] Not surprisingly, we get the highest estimate of λ_π during chairman Greenspan's office period, and the lowest for chairman Burns.

All the regressions show signs of positive autocorrelation, so it is therefore difficult to say something about the significance levels of these estimates. For the whole sample period, the presence of autocorrelation is what one would expect as a result of the regime shifts that have occured. For the different chairmens' office periods, the autocorrelation is much less pronounced in accordance with Judd and Rudebusch (1998). Despite the autocorrelation, the estimates in Table 1 are still consistent and thus suffice for the purpose of this paper.[11] Not surprisingly, we get the highest estimate of λ_π during chairman Greenspan's office period, and the lowest for chairman Burns.

To estimate $\rho_{\ln G}$ and $\sigma^2_{\ln G}$ in (6), I collected yearly data series on real government expenditures on consumption and investment per capita in natural logarithms, and filtered the series with the Hodrick-Prescott (H-P) filter (see Robert J. Hodrick and Prescott,

[10] The small sample adjusted B-G test statistics have been computed by: (i) estimating a VAR-model with 6 lags including the variables $\ln Y_t - \ln Y^*$, $\pi_t - \pi^*$ and μ_t (using likelihood ratio, autocorrelation and normality tests to determine the lag order) on data for the different periods; (ii) using the estimated VAR-model as a data generating process to simulate artificial samples of data; (iii) estimating the regression (8) on the simulated data with IV and then computing the associated B-G statistics. From the resulting distributions of B-G statistics, the small sample adjusted p-values are computed as the fraction of simulated B-G statistics that are larger than the estimated ones. The resulting p-values by this procedure are 0.683, 0.613, 0.663 and 0.458 for Whole sample, Burns, Volcker and Greenspan regimes respectively.

[11] It can also be argued that the asymptotic χ^2-distribution for the Breusch-Godfrey test is very likely to be oversized (i.e. an exaggerated probability of rejecting a true null hypothesis of no autocorrelation) for sample sizes as small as the present ones ($T \approx 30$).

1997) to get a measure of $\ln G_t$.[12] I then estimated (6) with ML with the result (standard error in parenthesis)

$$\ln G_t = \underset{(0.1040)}{0.7706} \ln G_{t-1} + \hat{\varepsilon}_{\ln G,t},\ \hat{\sigma}_{\ln G} = 0.01747,\ \text{D-W} = 0.94,\ T = 37,\ \bar{R}^2 = 0.60. \quad (19)$$

(19) shows tendencies of positive autocorrelation, but when I augmented the estimation with more lags on the dependent variable to remove this autocorrelation, I found that the estimated parameters were not much affected.

To compute values for μ and g, I took averages of nominal money growth and the ratio of government expenditures to gross national product to get 0.05345 and 0.21038 respectively.

Since Cooley and Hansen (1995) calibrated their model on quarterly data, I have mapped some of their parameters to fit yearly growth facts. More specifically; β is set to 0.9567 (0.989^4) instead of 0.989; δ is set to 0.07386 $\left(1-(1-0.019)^4\right)$ instead of 0.019; $\rho_{\ln Z}$ is set to 0.8145 (0.95^4) instead of 0.95; $\sigma_{\ln Z}$ is set to 0.013 $\left(\sqrt{0.00721\,(1+0.95^2+0.95^4+0.95^6)}\right)$ instead of 0.00721. γ is set so that hours worked as share of available time in steady state, H, equal 0.30.[13] The parameter values for α and θ are set to 0.84 and 0.40 as in Cooley and Hansen (1995).

4 Investigated relationships

In this section, I will derive the relationships that will be used later in the paper to investigate the small sample properties of the super exogeneity test.

4.1 Application 1: Money demand

By using the cash-in-advance constraint (3) in equilibrium, $C_{1t} = M_{t+1}/P_t - G_t$ (log linearized around steady state), (5) (log linearized around steady state, where I have used the approximation $\ln(1+R_t) \approx R_t$ since R_t is a small number), and the resource

[12] I use the common value 100 (yearly data) for the smoothness coefficient λ in the H-P filter. See also Appendix B for a detailed description of the raw data and data transformations.

[13] Formally, we have that $\gamma = \frac{(1-\theta)\left(\alpha\beta\frac{1}{e^{\mu}}+1-\alpha\right)\left(\frac{1-\beta(1-\delta)}{\beta\theta}\right)}{H\left((1-g)\left(\frac{1-\beta(1-\delta)}{\beta\theta}\right)-\delta\right)}$, which can be used to compute $\gamma = 3.404$ given the values for the other parameters. This value is higher than Cooley and Hansen's value (2.53) since I have government expenditures in the model.

constraint (log linearized around steady state) (18), it is possible to derive the following "true" money demand equation

$$\ln\left(\frac{M_{t+1}}{P_tY_t}\right)+\frac{\hat{P}G}{1-\hat{P}G}\ln\left(\frac{M_{t+1}}{P_tG_t}\right)=\kappa_0+\delta\kappa_1\left(\ln Y_t-\ln I_t\right)+\kappa_2\left(\ln Y_t-\ln G_t\right)-\kappa_3R_t \quad (20)$$

where steady state is denoted by dropping time subscripts, κ_1 denotes the capital-consumption ratio in steady state, κ_2 the government expenditure-consumption ratio in steady state, and κ_3 is shorthand notation for $\left(C-C_1\right)/C = C_2/C > 0$. Note that M_{t+1} in (20) corresponds to M_t in the data (M_t in the data is the money stock at the end of period t) since M_{t+1} is determined at the end of period t.

Since $\ln I_t$ and $\ln G_t$ normally do not enter in the estimation of a money demand equation, (20) is approximated with the following traditional money demand equation

$$\ln\left(\frac{M_{t+1}}{P_t}\right)=\beta_0+\beta_1\ln Y_t+\beta_2R_t+\beta_3\ln\left(\frac{M_t}{P_{t-1}}\right)+\varepsilon_{MD,\,t}. \quad (21)$$

The purpose of replacing the true money demand equation in (20) with (21), is that the latter is of the form used in practical work, see Goldfeld and Sichel (1990), whereas the former is not. Another obvious reason is that the Lucas critique is not relevant in (20), since it is a structural relationship as suggested by Lucas (1988). By using the monetary policy rule, (8), and the relations used to derive (20), it is possible to establish that

$$\begin{aligned}
\beta_0 &\equiv 0,\\
\beta_1 &\equiv (1-\eta\mathrm{L})(1-\mathrm{L})\left(1-\hat{P}G\right)(1+\delta\kappa_1+\kappa_2),\\
\beta_2 &\equiv (1-\eta\mathrm{L})(1-\mathrm{L})\left(1-\hat{P}G\right)\kappa_3,\\
\beta_3 &\equiv (1+\eta-\eta\mathrm{L}),\\
\varepsilon_{MD,\,t} &\equiv -(1-\eta\mathrm{L})(1-\mathrm{L})\left[\left(1-\hat{P}G\right)\delta\kappa_1\ln I_t+\left(\left(1-\hat{P}G\right)\kappa_2+\hat{P}G\right)\ln G_t\right],
\end{aligned}$$

where the lag-operator L is defined by $\mathrm{L}X_t \equiv X_{t-1}$. This shows that, contrary to the findings in Lucas (1988), that the parameters in a traditional money demand equation derived in this general equilibrium framework are dependent on the monetary policy rule (8) explicitly via the parameter η and implicitly via the error term (when the monetary policy rule changes, the decision rule for the capital stock will change as well which will affect the behavior of private investment).

4.2 Application 2: A "Keynesian" consumption function

By using the same equations and approximations as in the previous subsection, it is possible to derive the following "true" consumption function

$$\ln C_t = \kappa_4 - \tfrac{\hat{P}}{1-\hat{P}G-\hat{P}C}\left(Y \ln Y_t - \delta K \ln I_t\right) + \tfrac{1-\hat{P}G}{1-\hat{P}G-\hat{P}C}\kappa_3 R_t - \tfrac{1}{1-\hat{P}G-\hat{P}C}\ln\left(\frac{M_{t+1}}{P_t}\right) \quad (22)$$

where steady state is denoted by dropping time subscripts, and $Y = C + \delta K + G$. It can be shown that $1 - \hat{P}G - \hat{P}C < 0$, so that an increase in output/nominal interest rate increases/decreases consumption since $\hat{P} > 0$ and $1 - \hat{P}G > 0$.

Since $\ln I_t$ and $\ln\left(M_{t+1}/P_t\right)$ normally do not enter in the estimation of a traditional consumption function, (22) is approximated with the following "Keynesian flavored" consumption function

$$\ln C_t = \gamma_0 + \gamma_1 \ln Y_t + \gamma_2 R_t + \gamma_3 \ln C_{t-1} + \varepsilon_{CF,\,t}. \quad (23)$$

Again, the purpose of replacing the true consumption function in (22) with (23), is that the latter is of the form used in practical work whereas the former is not. By inserting (22) in (21), we have that

$$\begin{aligned}
\gamma_0 &\equiv \kappa_4, \\
\gamma_1 &\equiv -\frac{(1-\eta\mathrm{L})\left[\hat{P}Y + \left(1-\hat{P}G\right)(1+\delta\kappa_1+\kappa_2)\right]}{1-\hat{P}G-\hat{P}C}, \\
\gamma_2 &\equiv \frac{2\,(1-\eta\mathrm{L})\left(1-\hat{P}G\right)\kappa_3}{1-\hat{P}G-\hat{P}C}, \\
\gamma_3 &\equiv \eta, \\
\varepsilon_{CF,\,t} &\equiv \frac{(1-\eta\mathrm{L})}{1-\hat{P}G-\hat{P}C}\left[\left(\left(1-\hat{P}G\right)+\hat{P}Y\right)\delta\kappa_1 \ln I_t + \left(\left(1-\hat{P}G\right)\kappa_2 + \hat{P}G\right)\ln G_t\right].
\end{aligned}$$

As was the case with the money demand function, the parameters in the consumption function in (23) are dependent, explicitly and implicitly, on the monetary policy rule.

5 The quantitative importance of the Lucas critique in the model

In this section, I will investigate whether the Lucas critique seems to be significant in a statistical sense in the equilibrium model by simulating the model for the estimated Taylor rules for nominal money growth, and estimate the relationships (21), (23) and (8) on simulated data.

The procedure in the simulations has been as follows:

1. Simulate the model for T periods under the assumption that the monetary policy rule changes completely unexpectedly after $T/2$ periods from one regime to another (for example, from Burns to Volcker and Burns to Greenspan).[14]

2. Estimate the money demand equation (21), consumption function (23) and monetary policy rule (8) with OLS on the first $1, ..., T/2$ observations in the simulated sample. Denote the estimated parameter vectors $\hat{\boldsymbol{\beta}}_{MD}$, $\hat{\boldsymbol{\beta}}_{CF}$ and $\hat{\boldsymbol{\beta}}_{TR}$ respectively.

3. Estimate (21), (23) and (8) with OLS on the last $T/2+1, ..., T$ observations in the simulated sample. Denote the estimated parameter vectors $\hat{\boldsymbol{\alpha}}_{MD}$, $\hat{\boldsymbol{\alpha}}_{CF}$ and $\hat{\boldsymbol{\alpha}}_{TR}$ respectively.

4. Use a version of the F-test, often called the Chow breakpoint test, to examine if the null hypotheses $\boldsymbol{\alpha}_{MD} = \boldsymbol{\beta}_{MD}$, $\boldsymbol{\alpha}_{CF} = \boldsymbol{\beta}_{CF}$ and $\boldsymbol{\alpha}_{TR} = \boldsymbol{\beta}_{TR}$ are maintained at the 5 percent significance level.

5. Repeat Step 1 to Step 4 many (N) times to compute probabilities for how often the null hypotheses are maintained for the given significance level.

6. To get correct significance levels, steps 1 - 5 above are carried out twice. In the first round, small sample critical values are computed under the (true) null hypotheses $H_0 : \boldsymbol{\alpha}_{MD} = \boldsymbol{\beta}_{MD}$, $H_0 : \boldsymbol{\alpha}_{CF} = \boldsymbol{\beta}_{CF}$ and $H_0 : \boldsymbol{\alpha}_{TR} = \boldsymbol{\beta}_{TR}$ (that is, compute

[14] The simulations are made in the GAUSS programming language, using the random number generator RDND with RDNDSEED set to $159425 + iter$ for $iter = 1, 2, ..., N$. To get a stochastic initial state in each simulation, the model is simulated for $T + 100$ periods, where the first 100 are then discarded in all the estimations.

the distribution of F-statistics when there has been no regime shift). In the second round, these adjusted critical values ensures a correct size in the F-testing for regime shifts.

If the computed probabilities in Step 5 (in the second round) of rejecting parameter stability are lower/higher than the given significance levels, the Lucas critique is/is not relevant in this model in a statistical sense.[15]

The critical assumptions in the above scheme are clearly made in Step 1, and I would like to briefly comment on them. First, I have chosen to change monetary policy regime in the middle of the sample. The reason for this choice is that it implies the highest possible power in the testing. Secondly, I have chosen to model the once and for all change in monetary policy regime as a completely unexpected shift in the estimated monetary policy rule where I let the economy bring the state vector from the last period in the previous regime (period $T/2$) to the first period in the new regime (period $T/2+1$). By this procedure, I implicitly assume a first order Markov chain for the different monetary policy regimes where I let the diagonal elements in the transition matrix approach unity. This assumption is very convenient since it allows me to use the same decision rules for the first $T/2$ periods and then change to new decision rules only once in the beginning of period $T/2+1$ for the remaining $T/2$ periods.

The results of this exercise for the different estimated monetary policy rules for sample sizes $T = 100$ and $T = 200$, are provided in Table 2.[16]

As seen in Table 2, the probabilities of rejecting the null hypothesis of parameter stability between regimes are in almost every case clearly higher than the given significance level both for $T = 100$ and $T = 200$. It is only the regime shift from "Whole sample" to Greenspan that is not statistically significant (the probabilities for the vice versa case

[15] Since there are lag operators in the parameters and endogeneity issues in the error terms in the approximated money demand (21) and consumption function (23) relationships, I have checked the robustness of the results with respect to: (i) introducing one more lag on the regressors in (21) and (23) although they are extraneous; (ii) estimating the equations with instrument variables (IV) instead. The results of this exercise, which do not change the qualitative aspects of the results in the paper, are reported in Appendix B which is available on request from the author.

[16] I do not use more than $T = 200$ in the simulations since $T \approx 100$ seems to be an upper bound on observations in the studies which have applied the super exogeneity test to assess the practical importance of the Lucas critique on yearly data. To get reliable probabilities, I have simulated the model $N = 100000$ times.

Table 2: Chow test probabilities of rejecting the null hypothesis of parameter stability in money demand, consumption and the monetary policy rule at the 5 percent significance level.

	$T = 100$				$T = 200$			
	Comparison regime				Comparison regime			
	WS	B	V	G	WS	B	V	G
Benchmark regime	*The money demand function* (21)							
Whole sample (WS)	0.050	0.219	0.083	0.046	0.050	0.416	0.121	0.044
Burns (B)	0.654	0.050	0.328	0.463	0.916	0.050	0.711	0.785
Volcker (V)	0.422	0.319	0.050	0.212	0.607	0.671	0.050	0.330
Greenspan (G)	0.092	0.208	0.062	0.050	0.101	0.440	0.088	0.050
Benchmark regime	*The consumption function* (23)							
Whole sample	0.050	0.277	0.123	0.044	0.050	0.258	0.116	0.041
Burns	0.462	0.050	0.164	0.331	0.708	0.050	0.438	0.596
Volcker	0.502	0.354	0.050	0.312	0.708	0.673	0.050	0.513
Greenspan	0.080	0.283	0.096	0.050	0.084	0.352	0.114	0.050
Benchmark regime	*The Taylor rule for monetary policy* (8)							
Whole sample	0.050	0.863	0.606	0.078	0.050	0.999	0.941	0.093
Burns	0.943	0.050	0.258	0.687	1.000	0.050	0.517	0.975
Volcker	0.584	0.154	0.050	0.209	0.944	0.365	0.050	0.469
Greenspan	0.073	0.481	0.209	0.050	0.100	0.926	0.465	0.050

Note: The diagonals equals 0.050 since the critical values used in the testing have been simulated under the null hypothesis of no regime shift (that is, $H_0 : \boldsymbol{\alpha}_{MD} = \boldsymbol{\beta}_{MD}$, $H_0 : \boldsymbol{\alpha}_{CF} = \boldsymbol{\beta}_{CF}$ and $H_0 : \boldsymbol{\alpha}_{TR} = \boldsymbol{\beta}_{TR}$ are true). The Chow (1960) statistic underlying the computation of the probabilities is defined as $\frac{\left(\hat{\sigma}_T^2 - \frac{T_1}{T}\hat{\sigma}_{T_1}^2 - \frac{T_2}{T}\hat{\sigma}_{T_2}^2\right)/k}{\left(\frac{T_1}{T}\hat{\sigma}_{T_1}^2 + \frac{T_2}{T}\hat{\sigma}_{T_2}^2\right)/(T-2k)}$ and it follows the F-distribution with $k, T-2k$ degrees of freedom where k is the number of parameter restrictions that are being tested (here, $k = 4$), T the total number of observations ($T \equiv T_1 + T_2$) and $\hat{\sigma}_T^2$, $\hat{\sigma}_{T_1}^2$, and $\hat{\sigma}_{T_2}^2$denote the estimated standard error of the regression during both monetary regimes, the first monetary regime, and the second monetary regime respectively. The probabilities reported are computed from $N = 100000$ simulations.

are also low, but clearly significant) for the money demand and consumption function.[17] Not surprisingly, the probabilities for the monetary policy rule are also very low for this regime shift. The reason being that the estimated yearly residual standard error in the monetary policy rule, $\sqrt{\hat{\sigma}_{\xi}^{2}\left(1+\hat{\eta}^{2}+\hat{\eta}^{4}+\hat{\eta}^{6}\right)}$, is relatively high for these regimes as can be seen from Table 1, which makes it hard to detect this particular monetary regime shift.

Some other general comments on the content in Table 2 are also warranted. First, when T increases from 100 to 200, the Chow test discovers the shift in monetary regime more frequently as expected. Although Table 2 only report results for the 5 percent significance level, it is immediately recognized that the choice of significance level is of little importance here, since the size of the probabilities in almost every case are so large. Third, we see that the probabilities are on average higher for the monetary policy rule than the behavioral equations, which is quite natural since the effects of changes in the monetary rules on the money demand and consumption are somestimes offsetted by other shocks in the model. Finally, note that the probabilities in the diagonal (for the case of no regime shift) equals 0.05 since small sample critical values have been used.

The conclusion is that the Lucas critique is quantitatively important in this model in a statistical sense for every regime shift except one, and the super exogeneity test should recognize that.

It is also interesting to examine if the changes are economically meaningful, and whether the estimated parameters in the approximated relationships (21) and (23) are reasonable, and if the equations pass econometric specification tests for autocorrelation and normality of residuals underlying the Chow test. In Table 3a and 3b, I provide the OLS estimation results for the approximated money demand and consumption equations during the different monetary policy regimes.

From Tables 3a and 3b, we see that in most cases, the parameter estimates are reasonable from an economic point of view. As indicated by the p-values, both the specifications also pass all the statistical tests. So, both from an economic and econometric point of view, the estimated equations are perfectly acceptable. We also see that the point esti-

[17] Note that the standard errors for the probabilities in Table 2, $\hat{P}\left(1-\hat{P}\right)/\sqrt{N}$, are very small (around 0.001) and every regime shift except "Whole sample" to Greenspan are consequently statistically significant on the 1 percent level.

Table 3a: Estimation results of (21) for different monetary regimes.

	Estimation output							
Monetary regime	β_0	β_1	β_2	β_3	R^2	D-W	Durbin's h	J-B
Whole sample	-0.122 (0.022)	0.139 (0.027)	-0.137 (0.030)	0.764 (0.038)	0.93	1.93	0.55 (0.881)	16.83 (0.385)
Burns	-0.114 (0.010)	0.151 (0.007)	0.027 (0.023)	0.790 (0.012)	0.97	2.18	-1.27 (0.767)	1.90 (0.956)
Volcker	-0.080 (0.009)	0.115 (0.011)	-0.038 (0.025)	0.841 (0.014)	0.95	2.32	-2.33 (0.337)	7.38 (0.787)
Greenspan	-0.097 (0.015)	0.139 (0.019)	-0.115 (0.030)	0.798 (0.026)	0.94	2.09	-0.65 (0.884)	23.32 (0.267)

Note: Standard errors for regression coefficients in parentheses. The values within parentheses for Durbin's h test for autocorrelation (null hypothesis no first-order autocorrelation) and the Jarque-Bera (J-B) normality test (null hypothesis normally distributed residuals) measure the likelihood that the computed test statistics are insignificant (significance level 5 percent). I use $T = 200$ observations in each estimation. All the statistics reported are averages of $N = 100000$ estimations.

mates ("true" parameters) in many cases change substantially from an economic point of view when the monetary regime change. In particular this is the case for Burns and Greenspan/"Whole sample" regimes with respect to the monetary policy channel, i.e. the parameters β_2 and γ_2. In view of the estimation results of the monetary policy rules provided in Table 1 in Section 3, this is not surprising since the estimated monetary policy rules for these periods are very different.

Table 3b: Estimation results of (23) for different monetary regimes.

	Estimation output							
Monetary regime	γ_0	γ_1	γ_2	γ_3	R^2	D-W	Durbin's h	J-B
Whole sample	-0.123 (0.021)	0.209 (0.031)	-0.105 (0.044)	0.785 (0.032)	0.96	2.09	-0.65 (0.870)	21.85 (0.295)
Burns	-0.133 (0.012)	0.217 (0.008)	0.093 (0.026)	0.794 (0.012)	0.99	1.98	0.16 (0.957)	7.63 (0.717)
Volcker	-0.104 (0.011)	0.165 (0.013)	0.058 (0.030)	0.839 (0.013)	0.97	2.40	-2.87 (0.145)	19.25 (0.697)
Greenspan	-0.106 (0.015)	0.207 (0.025)	-0.080 (0.042)	0.805 (0.024)	0.96	2.26	-1.88 (0.529)	23.18 (0.252)

Note: See Table 3a.

To sum up, the results in this section suggest that the Lucas critique is quantitatively important in this typical equilibrium model with a standard parameterization in the sense that the structure of the economy, and thus how the economy reacts to different policies, changes in a statistically significant way when the monetary policy rule changes.

6 Does the super exogeneity test detect the Lucas critique in practice?

In the previous section, we verified that the Lucas critique is quantitatively important in a statistically significant way in a typical equilibrium model with a standard parameterization when we moved between different monetary policy rule regimes. A natural question then arises: Why has, in principle, every paper that has tried to test whether the Lucas critique is significantly important in practice, using tests for super exogeneity, rejected the Lucas critique in practice? One possible explanation, investigated in this section, is that the test for super exogeneity is not able to measure the relevance of the Lucas critique in data sufficiently well.

The structure of this section is the following. First, following Engle and Hendry (1993), and Ericsson and Irons (1995) closely, I give a short introduction to the concept of super exogeneity, discuss how one can test for it, and for the sake of clarity also provide a formal definition of it. Then I use the theoretical model, where the Lucas critique is applicable, to investigate whether the super exogeneity test has enough power in small samples on simulated data. I present results for both the money demand equation and the consumption function in the model.

6.1 Super exogeneity: concept, testing and a formal definition

Consider the following simple (presented in reduced form) forward looking model in the spirit of Lucas (1976):

$$\begin{aligned} x_t &= \theta \mathrm{E}_t \sum_{j=0}^{\infty} z_{t+j} + \varepsilon_t,\ \mathrm{E}_t \varepsilon_{t+j} = 0, \\ z_t &= \phi z_{t-1} + \upsilon_t,\ -1 < \phi < 1,\ \mathrm{E}_t \upsilon_{t+j} = 0 \end{aligned} \tag{24}$$

In (24), z_t is the policy variable and x_t the target variable. Solving (24) for x_t as function of z_t, we obtain

$$x_t = \beta_z z_t + \varepsilon_t \tag{25}$$

where $\beta_z \equiv \theta/(1-\phi)$. If the econometrician estimates (25), ignoring the dependency of β_z on ϕ, then policy simulations based on the estimate $\hat{\beta}_z$ for alternative paths of $\{z_{t+j}\}_{j=0}^{\infty}$

(treating v_{t+j} as a fixed exogenous shock), and thus for alternative paths of ϕ, will give misleading results.

Testing for the constancy/nonconstancy of ϕ and β_z in (24) and (25) by estimating these equations then provides a simple way of testing for the Lucas critique; if β_z is constant but ϕ is not, then the Lucas critique cannot apply. z_t is then, loosely speaking, said to be super exogenous to β_z. The testing procedure can generate three other combinations of constancy/nonconstancy for ϕ and β_z, but those combinations can arise from other sources (that is, changes in other policy variables) and thus not constitute evidence for or against the Lucas critique in practice.[18]

Now, let us consider a general case and give a formal definition of super exogeneity. Formally, the joint distribution of x_t and z_t conditional on the sigma field, denoted $\mathcal{F}_t$, consisting of $x_{t-1}, x_{t-2}, ...$ and $z_{t-1}, z_{t-2}, ...$, and the current and past observations on all other valid conditioning variables, can be written

$$D\left(x_t, z_t \mid \mathcal{F}_t, \lambda_t\right) = D_{x|z}\left(x_t \mid \; z_t, \mathcal{F}_t, \lambda_{1t}\right) D_z\left(z_t \mid \mathcal{F}_t, \lambda_{2t}\right)$$

where D, $D_{x|z}$ and D_z denote the joint density, the conditional density of x_t given z_t, and the marginal density of z_t, respectively, and λ_t, λ_{1t} and λ_{2t} the corresponding parameters. Engle et al. (1983) define z_t as weakly exogenous for a set of parameters θ if: (i) θ is a function of the parameters λ_{1t} alone, and (ii), λ_{1t} and the parameters λ_{2t} of the marginal model for z_t are variation free.[19] Finally, Engle et al. (1983) define z_t as super exogenous for if z_t is weakly exogenous for θ and λ_1 is invariant to changes in λ_2 (that is, changes in λ_2 do not imply changes in λ_1). The bottom line is that weak exogeneity of z_t for θ is sufficient for the conditional model to be used in forecasting analysis, whereas super exogeneity is required for policy analysis.

[18] In the context of the theoretical equilibrium model, it may be that the processes for the other exogenous shocks $\ln Z_t$ and/or $\ln G_t$ have changed.

[19] With the term "variation free" Engle et al (1983) mean that over periods of constant λ_{2t}, there is no information in λ_2 that would help estimating λ_1. Note that variation free and invariance are different concepts since if $\lambda_{1t} = \phi\lambda_{2t}$ then λ_1 is said to be variation free but not invariant to λ_2, but if the relation instead is $\lambda_1 = \phi_t\lambda_{2t}$ $\forall t$, then λ_1 is said to be both variation free and invariant to λ_2.

6.2 Application 1: The money demand equation

In this subsection, I will investigate if the test for super exogeneity is able to judge the relevance of the Lucas critique in small samples by using simulated data from the equilibrium model for the money demand equation and the Taylor rule for nominal money growth. We might think of money demand as x_t and monetary policy as z_t in (24).

The procedure in applying the test of super exogeneity on the money demand (21) together with the nominal money growth policy rule (8), is similar to that in Section 5. It involves the following steps:

1. Simulate the model for T periods under the assumption that the monetary policy rule changes completely unexpectedly after $T/2$ periods from one regime to another (for example, from Burns to Volcker and Burns to Greenspan).[20]

2. Estimate the money demand equation (21) and monetary policy rule (8) with OLS on the first $1, ..., T/2$ observations in the simulated sample. Denote the estimated parameter vectors $\hat{\boldsymbol{\beta}}_{MD}$ and $\hat{\boldsymbol{\beta}}_{TR}$ respectively.

3. Estimate (21) and (8) with OLS on the last $T/2+1, ..., T$ observations in the simulated sample. Denote the estimated parameter vectors $\hat{\boldsymbol{\alpha}}_{MD}$, $\hat{\boldsymbol{\alpha}}_{CF}$ and $\hat{\boldsymbol{\alpha}}_{TR}$ respectively.

4. Use a version of the F-test, the Chow breakpoint test, to examine if the null hypothesis $\boldsymbol{\alpha}_{MD} = \boldsymbol{\beta}_{MD}$ is rejected when the null $\boldsymbol{\alpha}_{TR} = \boldsymbol{\beta}_{TR}$ can be rejected on the ten, five and one percent significance level.

5. Repeat Step 1 to Step 4 many (N) times to compute the power of the super exogeneity test, that is, probabilities for how often the null hypothesis $\boldsymbol{\alpha}_{TR} = \boldsymbol{\beta}_{TR}$ is rejected and the null hypothesis $\boldsymbol{\alpha}_{MD} = \boldsymbol{\beta}_{MD}$ is rejected simultaneously for the different significance levels.

[20] As in Section 5, the random number generator RDND with RDNDSEED in GAUSS is set to $159425 + iter$ for $iter = 1, 2, ..., N$. To get a stochastic initial state in each simulation, the model is simulated for $T + 100$ periods, where the first 100 are then discarded in all the estimations discussed below.

According to the Lucas critique, which we verified in Section 5 was relevant in this model, the null hypothesis

$$H_0\colon \boldsymbol{\alpha}_{MD} = \boldsymbol{\beta}_{MD} \mid \boldsymbol{\alpha}_{TR} \neq \boldsymbol{\beta}_{TR}$$

is false so if the computed probabilities in Step 5 of rejecting parameter stability in (8) and (21) at the same time are low, the power of the super exogeneity test is low in small samples. On the other hand, if the computed probabilities are found to be high, then the power of the super exogeneity test in small samples is satisfactory.[21]

For the Chow test in Step 4, I use small sample adjusted critical values generated in Section 4 rather than asymptotic values. The reason for doing so is to get the correct nominal significance level for each equation involved in the testing.[22]

As in Section 5, the critical assumptions in step 1 to 5 are clearly made in step 1, and since I already discussed them there I will not spend time on discussing them here again.

The results of this exercise for the different estimated monetary policy rules, and sample sizes $T = 100$ and $T = 200$, are provided in Table 4.

The general message from Table 4 is striking. For $T = 100$, which is an upper bound on yearly observations that has been used in the literature so far, the power of the super exogeneity test is very low in most cases although we have given the test the best possible environment for detecting a regime shift. The average power probability is as low as 0.30 for $T = 100$ (the regime shift "Whole sample to Greenspan" excluded since it was not statistically significant for this sample size, see Table 2). The test has highest power for the regime shifts from Burns to "Whole sample"/Greenspan which is not surprising given the results in Tables 3a and 3b. For $T = 200$, which is a considerable amount of data with as much as 100 yearly observations in each regime, the super exogeneity test has satisfactory power properties in only a few cases, especially for the regime shifts from Burns to "Whole sample"/Greenspan as was the case for $T = 100$. On the whole, however, the power probability is not higher than 0.45 on average, which is not satisfactory. The results imply that the probability of rejecting the (false) null hypothesis that the Lucas

[21] Note that the power of a test is formally defined as $1 - \beta$, where β, the type II error, is defined as Pr (do not reject $H_0|H_0$ is false). In our case, β is the probability of not rejecting stability in money demand, while rejecting stability in the monetary policy rule simultaneously.

[22] It should be noted, however, that the qualitative aspects of the results in Table 2, Table 4 and 5 are robust with respect to this adjustment of the critical values.

Table 4: Power of the super exogeneity test in small samples: Chow test probabilites of rejecting stability in money demand when stability in the monetary policy rule is rejected.

	$T = 100$				$T = 200$			
	Comparison regime				Comparison regime			
	WS	B	V	G	WS	B	V	G
Benchmark regime	10 percent significance level							
Whole sample (WS)	NC	0.380	0.183	0.163	NC	0.614	0.239	0.162
Burns (B)	0.759	NC	0.535	0.602	0.953	NC	0.850	0.854
Volcker (V)	0.569	0.541	NC	0.371	0.722	0.839	NC	0.477
Greenspan (G)	0.276	0.387	0.163	NC	0.296	0.618	0.186	NC
Benchmark regime	5 percent significance level							
Whole sample	NC	0.229	0.103	0.102	NC	0.417	0.124	0.097
Burns	0.658	NC	0.412	0.507	0.911	NC	0.776	0.791
Volcker	0.472	0.406	NC	0.272	0.618	0.755	NC	0.374
Greenspan	0.188	0.259	0.095	NC	0.206	0.451	0.100	NC
Benchmark regime	1 percent significance level							
Whole sample	NC	0.066	0.030	0.061	NC	0.112	0.023	0.043
Burns	0.461	NC	0.217	0.340	0.798	NC	0.600	0.643
Volcker	0.314	0.210	NC	0.158	0.402	0.548	NC	0.210
Greenspan	0.110	0.091	0.034	NC	0.123	0.165	0.027	NC

Note: NC is shorthand notation for not computed. Small sample critical values generated in Section 4 are used. The probability in each entry is formally defined as $\Pr\left(F_{MD} > F^{MD}_{0.\text{XX}} \mid F_{TR} > F^{TR}_{0.\text{XX}}\right)$where F_{TR} denotes the computed Chow statistic for the monetary policy rule, F_{MD} the corresponding statistic for money demand, and XX the i'th percentile in the simulated distributions ($90, 95$ or 99'th percentile) under the null hypothesis. Under the null hypothesis, the Chow statistic follows the F-distribution with $k, T-2k$ degrees of freedom where k is the number of parameter restrictions (here, $k = 4$ for both money demand and the Taylor rule) and T is the number of observations.

critique does not apply, and thus that the parameters in the money demand function are super exogenous to the monetary policy rule, is too high regardless of chosen significance level and sample size in most cases. Of course, one may argue that the low probabilities in Table 4 are due to that the changes in the monetary policy regime are not large enough, and every test has low power when the alternative is close to the null. The evidence against this type of argument is provided in Table 1, 3a and 3b, where we see that there actually are big changes in parameters which are manifested in the single equation stability testing results provided in Table 2. So the low level of the probabilities reported in Table 4, are not due to that the alternative hypotheses are close to the null, but rather due to inherently low power for the super exogeneity test in this setting.

A couple of further comments on specific features in Table 4 are warranted. First, when the sample size increased from $T = 100$ to $T = 200$, the power of the super exogeneity test only increased from about 0.30 to 0.45 on average (the regime shift from "Whole sample to Greenspan" excluded since it is not statistically significant for these sample sizes). This is an indication of that considerably larger sample sizes are needed to get an acceptable power level. Some experiments indicate that we need around $T = 1000$ yearly observations to get a power of 0.90 on average. Finally, Table 4 confirms the classical conflict between the choice of significance level (probability of committing a type I error) and the probability of not committing an type II error (that is, not reject a false null hypothesis): as the nominal significance level decreases, we see that the power also decreases for a given sample size T.

6.3 Application 2: The "Keynesian" consumption function

In this subsection, I will investigate the properties of the test for super exogeneity by using simulated data from the equilibrium model to estimate the consumption function together with the Taylor rule for nominal money growth. We might think of consumption as x_t and monetary policy as z_t in (24).

The procedure in the testing for super exogeneity here is exactly the same as in the previous subsection, except for that the consumption function in (23) is used instead of money demand in (21), and thus not repeated here. The results are reported in Table 5.

Table 5: Power of the super exogeneity test in small samples: Chow test probabilites of rejecting stability in consumption when stability in the monetary policy rule is rejected.

	$T = 100$				$T = 200$			
	Comparison regime				Comparison regime			
	WS	B	V	G	WS	B	V	G
Benchmark regime	10 percent significance level							
Whole sample (WS)	NC	0.455	0.258	0.165	NC	0.509	0.291	0.142
Burns (B)	0.586	NC	0.307	0.465	0.795	NC	0.599	0.701
Volcker (V)	0.635	0.534	NC	0.469	0.792	0.798	NC	0.628
Greenspan (G)	0.268	0.470	0.229	NC	0.250	0.572	0.252	NC
Benchmark regime	5 percent significance level							
Whole sample	NC	0.288	0.140	0.101	NC	0.263	0.117	0.069
Burns	0.468	NC	0.204	0.359	0.707	NC	0.462	0.594
Volcker	0.534	0.427	NC	0.368	0.714	0.696	NC	0.531
Greenspan	0.168	0.326	0.136	NC	0.152	0.361	0.130	NC
Benchmark regime	1 percent significance level							
Whole sample	NC	0.032	0.018	0.041	NC	0.008	0.007	0.025
Burns	0.272	NC	0.079	0.205	0.509	NC	0.239	0.393
Volcker	0.362	0.220	NC	0.231	0.547	0.480	NC	0.366
Greenspan	0.095	0.096	0.038	NC	0.042	0.035	0.015	NC

Note: NC is shorthand notation for not computed. Small sample critical values generated in Section 4 are used. The probability in each entry is formally defined as $\Pr\left(F_{CF} > F_{0.\text{XX}}^{CF} \mid F_{TR} > F_{0.\text{XX}}^{TR}\right)$where F_{TR} denotes the computed Chow statistic for the monetary policy rule, F_{CF} the corresponding statistic for consumption, and XX the i'th percentile in the simulated distributions ($90, 95$ or 99'th percentile) under the null hypothesis. Under the null hypothesis, the Chow statistic follows the F-distribution with $k, T - 2k$ degrees of freedom where k is the number of parameter restrictions (here, $k = 4$ for both consumption and the Taylor rule) and T is the number of observations.

Table 5 confirms the findings in Table 4: we see again that the power of the super exogeneity test is unacceptably low in small samples, and the implied probabilities of failing to reject the false null hypothesis that the Lucas critique is not relevant, are far too high (in fact, they are actually slightly higher than those in Table 4 on average).

To sum up this section, it seems to be a robust finding in this model that the test for super exogeneity have very low power in small samples - although we have given the test the best possible environment for detecting the regime shifts - and hence is not able to shed light on the practical importance of the Lucas critique in small samples.[23]

7 Conclusions

In this paper, I have tried to shed new light on why the Lucas critique has received very little empirical support (see Ericsson and Irons, 1995). A standard real business cycle model with a Taylor-type rule for nominal money (M1) growth was calibrated to different Federal Reserve regimes (the periods in office of chairmen Burns, Volcker and Greenspan). It was found that the super exogeneity test, which is used to identify the relevance/nonrelevance of the Lucas critique in practice, does not have enough power in small samples. This is a possible explanation why the Lucas critique has received very little empirical support.

Given the results, why does the test fail? The most important reason seems to be that the testing of super exogeneity involves the use of "conditional" models. In such models, we do not condition on all shocks that hit the economy and affect parameter stability. In the context of the analysis here, there are shocks to government expenditures and technology that are not explicitly taken into account in the testing. The fact that the test too often fail to rejects the false null hypothesis of parameter stability in money demand/consumption given parameter instability in the monetary policy rule, might be explained by shocks in government expenditures and technology which "neutralize" the

[23] Since monetary policy rules nowadays normally are estimated on nominal interest rates, I have also investigated this combination by estimating a standard Taylor rule $R_t = \beta_0 + \beta_1 (\pi_t - \pi^*) + \beta_2 (\ln Y_t - \ln Y^*) + \beta_3 R_{t-1} + \varepsilon_t^R$ on simulated data and applied the super exogeneity test on this relation together with (21) / (23). But since the results were qualitatively the same as those reported in Table 4 and 5 (in fact, the properties of the super exogeneity test were slightly worsened in this setting), they are not reported.

effects of the changes in monetary policy. This is one of the classical problems in time series analysis that cannot be controlled for in real-world data. However, in the general equilibrium framework used here, it is in fact possible to control for these effects by "going back in time" and performing the super exogeneity test conditional on these shocks except for the change in monetary policy rule.[24] Although not reported in the paper, I carried out this kind of experiment, and when controlling for the influences of other shocks on money demand/consumption, the power of the super exogeneity test is substantially higher.[25]

The tests reported here were based on samples of 100 and 200 observations (one observation is one year). The test of super exogeneity only does marginally better with the larger sample size, which is troublesome since 200 yearly observations comprise a substantially greater amount of data than has been used so far in the empirical literature. This is a clear indication that the sample size has to be increased considerably before the test may work accurately.

To simplify the experiments conducted throughout the paper, it was assumed that the regime shift in monetary policy, which takes place in the middle of the artificial samples, is completely unanticipated and that the agents in the economy anticipate the new regime to last forever. In real-world data, this assumption is likely to be invalid. In reality, it is more likely that people gradually learn about the workings of a new regime and therefore adjust their behavior gradually over time. This is problematical for a researcher who wants to test for the empirical applicability of the Lucas critique on real-world data, because it will be much harder for the super exogeneity test to judge the relevance/nonrelevance of the Lucas critique in such a situation. Here, even though my simplifying assumption here gives the super exogeneity test the best possible environment for revealing the empirical importance of the Lucas critique, it nevertheless fails to do so accurately.

Another limitation of the paper is that I have implicitly assumed that regime shifts in monetary policy do not affect the institutional design of the economy (that is, the

[24] More precisely, "going back in time" means the following. Let $\{\varepsilon_t\}_{t=1}^{T}$ denote a vector valued sequence of exogenous shocks (here government expenditures $\varepsilon_{\ln G}$, technology $\varepsilon_{\ln Z}$ and monetary policy ξ) for a particular sample $1, 2, ..., T$ and $\{\mu_t = \mu(\vartheta, \xi_t)\}_{t=1}^{T}$ a realization of monetary policy given a monetary policy rule (ϑ) . By going back in time, the money demand/consumption function and the monetary policy rule can be estimated on the same realization of shocks ε by generating (different) sequences of μ for different monetary policy regimes (ϑ).

[25] The results of this exercise are reported in Appendix B which is available on request from the author.

one-period nominal wage contacting assumption in the model) and the cash-in-advance constraint "as a social convention for using money". The former assumption can be motivated by the real-world observation that institutions, for example labor market arrangements, change very slowly over time. Consequently, when testing for the Lucas critique in practice, the effects of institutional changes should be of "second order", while the effects of monetary policy regime shifts examined in this paper should be of "first order".[26] The latter assumption can be motivated by the fact that the considered changes in monetary policy are relatively small. Cooley and Hansen (1989) also maintain this assumption throughout their analysis, although they consider much larger variations in monetary policy (e.g. they vary the inflation rate between -4 and 400 percent).

I have chosen to follow Judd and Rudebusch's (1998) empirical classification of monetary policy regimes in the U.S. An alternative would have been to adopt a more theoretical approach, as in e.g. Rudebusch and Svensson (1998), by considering monetary policy rules that are consistent with some kind of optimizing behavior of a central bank.[27] I would be very surprised, however, if this alternative approach would change the qualitative conclusions regarding the properties of the super exogeneity test on small samples.[28]

To sum up, the results in this paper suggest that more research on the properties of the super exogeneity test is warranted before drawing the conclusion that the Lucas critique is not empirically relevant for various economic relationships.

[26] If a sufficiently long period is covered (50 years or so), the effects of institutional changes may be larger; see for instance Klas Fregert and Lars Jonung (1998). When considering the effects of the Lucas critique here, however, I have in mind a shorter period (5 to 10 years or so).

[27] That is, in this setting, by examining the coefficients η, λ_π and λ_Y in the monetary policy rule (8) that are consistent with a given loss function of the central bank.

[28] I have experimented with a quadratic loss function of the form $L_t = \frac{1}{2}\left[(\pi_t - \pi^*)^2 + \lambda(\ln Y_t - \ln Y^*)^2\right]$ for the central bank for values of λ equal to 0, 0.3, 1 and 5 as in Svensson and Rudebusch (1998) and found that the qualitative conclusions were unaffected.

Appendix A Data sources and definitions

In this appendix, I provide the sources of the data collected in Table A.1 below.

Table A.1: The data set.

Variables	Sample period	Source
GNP	1960Q1-1997Q4	FRED database, Federal Reserve Bank of St. Louis
GEC	1960Q1-1997Q4	FRED database, Federal Reserve Bank of St. Louis
M1	1959Q1-1997Q4	FRED database, Federal Reserve Bank of St. Louis
POP	1960-1996	OECD Main Economic Indicators
CPI	1959Q1-1997Q4	FRED database, Federal Reserve Bank of St. Louis

Note: All real macroeconomic variables are measured in 1992 billion U.S. dollars. Abbreviations; GNP denotes real (fixed, seasonally adjusted) gross national product; GEC real (chained, seasonally adjusted) government consumption and investment; M1 (not seasonally adjusted) nominal money supply 1; CPI (not seasonally adjusted) consumer price index; POP average U.S. population (for 1997, POP is set equal to average gross growth rate times the value for 1996).

The transformations made to generate the variables used in Table 1 are displayed in Table A.2.

Table A.2: Generation of composite quaterly data series.

Variable	Sample period	Calculation formula
$\ln Y$	1960Q1-1997Q1	$\ln(\text{GNP}/\text{POP})$
μ	1960Q1-1997Q4	$\ln(\text{M1}_t/\text{M1}_{t-4})$
π	1960Q1-1997Q4	$\ln(\text{CPI}_t/\text{CPI}_{t-4})$
$\ln G$	1960Q1-1997Q4	$\ln(\text{GEC}/\text{POP})$

Note: To get measures of $\ln Y - \ln Y^*$ and $\pi - \pi^*$, $\ln Y$ and π are then subject to Hodrick-Prescott filtering with the smoothness coefficient λ set to 1600.

To transform the quarterly data for $\ln G$ in Table A.2 to yearly data, which are used in (19), I added up all the quarterly observations within a year to get a yearly observation and Hodrick-Prescott filtered the resulted series with the smoothness coefficient λ set to 100. To compute measures of the ratio of government expenditures to output and the growth rate in nominal money supply in steady state, g and μ respectively, I computed the sums $\frac{1}{152}\sum_{t=1960Q1}^{1997Q4}(\text{GEC}_t/\text{GNP}_t)$ and $\frac{1}{152}\sum_{t=1960Q1}^{1997Q4}\mu_t$.

References

Ball, Laurence, (1997), "Efficient Rules for Monetary Policy", NBER Working Paper Series, No. 5952.

Chow, Gregory C., (1960), "Tests of Equality between Sets of Coefficients in Two Linear Regressions", *Econometrica*, Vol. 28, No. 3, pp. 591-605.

Cooley, Thomas F. and Hansen, Gary D., (1989), "The Inflation Tax in a Real Business Cycle model", *American Economic Review*, Vol. 79, No. 4, pp. 733-748.

Cooley, Thomas F. and Hansen, Gary D., (1995), "Money and the Business Cycle", Chapter 7 in Cooley, Thomas F. (Ed.), *Frontiers of Business Cycle Research*, Princeton University Press, Princeton.

Engle, Robert F., Hendry, David F. and Richard, Jean-Francois, (1983), "Exogeneity", *Econometrica*, Vol. 51, No. 2, pp. 277-304.

Engle, Robert F. and Hendry, David F., (1993), "Testing Super Exogeneity and Invariance in Regression Models", *Journal of Econometrics*, Vol. 56, No. 2, pp. 119-139.

Ericsson, Neil R. and Irons, John S., (1995), "The Lucas Critique in Practice: Theory without Measurement", Chapter 8 in Hoover, Kevin D. (Ed.), *Macroeconometrics: Developments, Tensions, and Prospects*, Kluwer Academic Publishers, Boston.

Estrella, Arturo and Fuhrer, Jeffrey C., (1999), "Are 'Deep' Parameters Stable? The Lucas Critique as an Empirical Testable Hypothesis", Federal Reserve Bank of Boston Working Papers, No. 99-4.

Fregert, Klas and Jonung, Lars, (1998), "Monetary Regimes and Endogenous Wage Contracts: Sweden 1908-1995", School of Economics and Management Working Paper Series 1998:3, Department of Economics, University of Lund.

Fuhrer, Jeffrey C., (1997), "The (Un)Importance of Forward-Looking Behavior in Price Specification", *Journal of Money, Credit and Banking*, Vol. 29, No. 3, pp. 338-350.

Goldfeld, Stephen M. and Sichel, Daniel E., (1990), "The Demand for Money", in Friedman B. M. and Hahn F. H. (Eds.), *Handbook of Monetary Economics*, Vol. 1, pp. 299-356, North-Holland, Amsterdam.

Hansen, Gary D., (1985), "Indivisible Labor and the Business Cycle", *Journal of Monetary Economics*, Vol. 16, No. 3, pp. 309-327.

Hansen, Gary D. and Prescott, Edward C., (1995), "Recursive Methods for Computing Equilibria of Business Cycle Models", Chapter 2 in Cooley, Thomas F. (Ed.), *Frontiers of Business Cycle Research*, Princeton University Press, Princeton.

Hodrick, Robert J. and Prescott, Edward C., (1997), "Postwar U.S. Business Cycles: An Empirical Investigation", *Journal of Money, Credit and Banking*, Vol. 29, No. 1, pp. 1-16.

Judd, John P. and Rudebusch, Glenn D., (1998), "Taylor's rule and the Fed: 1970-1997", *Federal Reserve Bank of San Francisco Economic Review*, No. 3, pp. 3-16.

King, Robert G., (1988), Comment on "Money Demand in the United States: A Quantitative Review" by Lucas (1988), *Carnegie-Rochester Conference Series on Public Policy*, Vol. 29, pp. 169-172.

Kydland, Finn E. and Prescott, Edward C., (1982),"Time to Build and Aggregate Fluctuations", *Econometrica*, Vol. 50, No. 6, pp. 1345-1370.

Leeper, Eric M., (1995), Commentary on "The Lucas Critique in Practice: Theory without Measurement" by Ericsson and Irons (1995), in Hoover, Kevin D. (Ed.), *Macroeconometrics: Developments, Tensions, and Prospects*, pp. 313-324, Kluwer Academic Publishers, Boston.

Lucas, Robert E. Jr., (1975), "An Equilibrium Model of the Business Cycle", *Journal of Political Economy*, Vol. 83, No. 6, pp. 1113-1144.

Lucas, Robert E. Jr, (1976), "Econometric Policy Evaluation: A Critique", *Carnegie-Rochester Conference Series on Public Policy*, Vol. 1, pp. 19-46.

Lucas, Robert E. Jr, (1977), "Understanding Business Cycles", *Carnegie-Rochester Conference Series on Public Policy*, Vol. 5, pp. 7-29.

Lucas, Robert E. Jr, (1988), "Money Demand in the United States: A Quantitative Review", *Carnegie-Rochester Conference Series on Public Policy*, Vol. 29, pp. 137-167.

McCallum, Bennett T., (1984), "Monetarist Rules in the Light of Recent Experience", *American Economic Review*, Vol. 74, No. 2, pp. 388-391.

McCallum, Bennett T., (1988), "Robustness Properties of a Rule for Monetary Policy", *Carnegie-Rochester Conference Series on Public Policy*, Vol. 29, pp. 173-204.

Rudebusch, Glenn D. and Svensson, Lars E.O., (1998), "Policy Rules for Inflation Targeting", in Taylor, John B. (Ed.), *Monetary Policy Rules*, University of Chicago Press, forthcoming.

Sims, Christopher A., (1982), "Policy Analysis with Econometric Models", *Brookings Papers on Economic Activitity*, No. 1, pp. 107-152.

Svensson, Lars E.O., (1997), "Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets", *European Economic Review*, Vol. 41, No. 6, pp. 1111-1146.

Taylor, John B., (1993), "Discretion Versus Policy Rules in Practice", *Carnegie-Rochester Conference Series on Public Policy*, Vol. 39, pp. 195-214.

Taylor, John 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*, Vol. 43, No. 3, pp. 655-679.

Appendix B Additional results for "Testing for the Lucas Critique: A Quantitative Investigation"

B.1 Computation of equilibrium

In this subsection, I give more details on how to compute the equilibrium in this model.

In order to make all variables in the deterministic version of the model above converge to a steady state, I transform the nominal variables by dividing m_{t+1} and P_t with M_{t+1}, and m_t with M_t. If we introduce the notation

$$\hat{m}_{t+s} \equiv \frac{m_{t+s}}{M_{t+s}} \text{ and } \hat{P}_{t+s} \equiv \frac{P_{t+s}}{M_{t+s+1}}$$

and use the transformations to rewrite the equations (2), (3), (4), (8), (13) and (14), the representative agent's optimization problem can, following Hansen and Prescott (1995), be expressed as the recursive dynamic programming problem:

$$V(\mathbf{S}_t, \hat{m}_t, k_t) \equiv \max_{\{\hat{m}_{t+1}, h_t, k_{t+1}\}} \left[\alpha \ln(c_{1t}) + (1-\alpha)\ln(c_{2t}) - \gamma h_t + \beta \mathrm{E}_t V(\mathbf{S}_{t+1}, \hat{m}_{t+1}, k_{t+1})\right]$$

$$s.t. \tag{B.1}$$

$$\begin{aligned}
c_{1t} &= \frac{\hat{m}_t + e^{\mu_t} - 1}{e^{\mu_t}\hat{P}_t} - G_t, \\
c_{2t} &= (1-\theta)\, e^{\ln Z_t} \left(\frac{K_t}{H_t}\right)^{\theta} h_t + \left(1 + R_t^K - \delta\right) k_t - k_{t+1} - \frac{\hat{m}_{t+1}}{\hat{P}_t}, \\
\mu_t &= \tfrac{\eta}{1+\lambda_\pi}\mu_{t-1} - \tfrac{\lambda_\pi}{1+\lambda_\pi}\left(\ln \hat{P}_t - \ln \hat{P}_{t-1} - \pi^*\right) - \tfrac{\lambda_Y}{1+\lambda_\pi}\left(\ln Y_t - \ln Y^*\right) + \tfrac{1}{1+\lambda_\pi}\xi_t, \\
H_t - \mathrm{E}_{t-1}\ln H_t &= \tfrac{1}{\theta(1+\lambda_\pi)+(1-\theta)\lambda_Y}\left(\ln \hat{P}_t - \mathrm{E}_{t-1}\ln \hat{P}_t\right) + \tfrac{1+\lambda_\pi-\lambda_Y}{\theta(1+\lambda_\pi)+(1-\theta)\lambda_Y}\varepsilon_t^{\ln Z} + \tfrac{\xi_t - (1-\eta)\mu}{\theta(1+\lambda_\pi)+(1-\theta)\lambda_Y}, \\
h_t - \mathrm{E}_{t-1}\ln H_t &= \tfrac{1}{\theta(1+\lambda_\pi)+(1-\theta)\lambda_Y}\left(\ln \hat{P}_t - \mathrm{E}_{t-1}\ln \hat{P}_t\right) + \tfrac{1+\lambda_\pi-\lambda_Y}{\theta(1+\lambda_\pi)+(1-\theta)\lambda_Y}\varepsilon_t^{\ln Z} + \tfrac{\xi_t - (1-\eta)\mu}{\theta(1+\lambda_\pi)+(1-\theta)\lambda_Y}, \\
\ln K_{t+1} &= K(\mathbf{S}_t),\ \ln H_t = H(\mathbf{S}_t),\ \ln \hat{P}_t = \hat{P}(\mathbf{S}_t).
\end{aligned}$$

In (B.1), $\mathbf{S}_t$ is a 1×8 row vector which contains all the aggregate state variables $\ln Z_{t-1}$, $\varepsilon_t^{\ln Z}$, μ_{t-1}, ξ_t, $\ln G_t$, $\ln K_t$, $\ln \hat{P}_{t-1}$ and a constant term. If $\lambda_\pi = 0$, then $\ln \hat{P}_{t-1}$ vanishes in

$\mathbf{S}_t$.[1] In maximization of (B.1), the agent takes the economy-wide aggregate (average) variables as given. The functions K, $\hat{P}$ and H describe the relationship perceived by agents between the aggregate decision variables and the state of the economy. As the solution to the problem in (B.1), we have the agent's decision rules $\ln k_{t+1} = k(\mathbf{S}_t,\ \ln k_t,\ \ln \hat{m}_t)$, $\ln \hat{m}_{t+1} = \hat{m}(\mathbf{S}_t,\ \ln k_t,\ \ln \hat{m}_t)$ and $\ln h_t = h(\mathbf{S}_t,\ \ln k_t,\ \ln \hat{m}_t)$. The competitive equilibrium is obtained when the individual and average decision rules coincide for $\ln k_t = \ln K_t$ and $\ln \hat{m}_{t+1} = \ln \hat{m}_t = 0$.

Since it is impossible to derive the decision rules analytically, I have used the same method as Cooley and Hansen (1995) and computed the decision rules numerically by approximating the original problem with a second order Taylor expansion around the constant steady state values in the nominal-growth adjusted economy. As a consequence of this approximation, the method produces linear decision rules (in natural logarithms for K_{t+1}, H_t and $\hat{P}_t$). The algorithm utilized is described in detail in Hansen and Prescott (1995).

B.2 Power properties of the super exogeneity test when controlling for other shocks

Results are presented for the case where I control for the effects of other shocks, as discussed in the Conclusions.

The procedure for applying the test of super exogeneity to the money demand relation (21) together with the nominal money growth policy rule (8), involves the following steps in this case:

1. Simulate the model for T periods for a given monetary policy regime (Whole sample, Burns, Volcker or Greenspan).[2]

2. Estimate (21) and (8) with OLS on the simulated data. Denote the estimated

[1] Note that the household budget constraint on line 4 in (B.1) incorporates the fact that the contracted nominal wage divided by the price level equals the equilibrium marginal product of labor since firms unilaterally determine hours worked in period t.

[2] The simulations are made in the GAUSS programming language, using the random number generator RDND with RDNDSEED set to $159425 + iter$ for $iter = 1, 2, ..., N$. To get a stochastic initial state in each simulation, the model is simulated for $T + 100$ periods, where the first 100 are then discarded in the estimations.

parameter vectors $\hat{\boldsymbol{\beta}}_{MD}$ and $\hat{\boldsymbol{\beta}}_{TR}$ respectively.

3. Simulate the model for T periods again, using the same stochastic shock realizations for the exogenous processes as in step 1, under the assumption that the monetary policy rule has changed from one regime to another (for example, from Burns to Volcker and Burns to Greenspan).

4. Estimate (21) and (8) with OLS again on the new data. Denote these estimated parameter vectors $\hat{\boldsymbol{\alpha}}_{MD}$ and $\hat{\boldsymbol{\alpha}}_{TR}$ respectively.

5. Use the F-test to examine if the null hypothesis $\hat{\boldsymbol{\alpha}}_{MD} = \hat{\boldsymbol{\beta}}_{MD}$ is maintained while the null $\hat{\boldsymbol{\alpha}}_{TR} = \hat{\boldsymbol{\beta}}_{TR}$ can be rejected at the 10, 5 and 1 percent significance level.

6. Repeat step 1 to step 5 many (N) times to compute probabilities for rejection of the null hypothesis $\hat{\boldsymbol{\alpha}}_{TR} = \hat{\boldsymbol{\beta}}_{TR}$ while the null hypothesis $\hat{\boldsymbol{\alpha}}_{MD} = \hat{\boldsymbol{\beta}}_{MD}$ is maintained for different significance levels.

By this procedure, we control for the effects of the realizations of the two exogenous shocks in the model (technology shocks and government consumption) when evaluating the small sample properties of the super exogeneity test. All other aspects of the set up remain unchanged. For the consumption function (23), the same procedure is adopted.

The results for these equations along with the monetary policy rule (8), corresponding to Tables 4 and 5 in the paper, are presented in Tables B.1 and B.2.

As we can from Tables B.1 and B.2, the results are now very different from those in Tables 4 and 5. The super exogeneity test has much higher power in this setup in almost every case. Since the Lucas critique is still not statistically significant (for these sample sizes) when there is a regime shift from "Whole sample" to Greenspan, the power of the super exogeneity test is consequently somewhat higher in that case. The average power probabilities for $T = 100$ now equals 0.75 and 0.85 (the regime shift from "Whole sample" to Greenspan excluded since it is not statistically significant for the given sample size) for money demand and consumption whereas the corresponding numbers in Tables 4 and 5 are 0.30 and 0.28. For $T = 200$ they are about 0.85 and 0.90, corresponding to 0.45 and 0.39 in Tables 4 and 5.

Table B.1: Power of the super exogeneity test in small samples: Chow test probabilites of rejecting stability in money demand when stability in the monetary policy rule is rejected.

	$T = 100$				$T = 200$			
	Comparison regime				Comparison regime			
	WS	B	V	G	WS	B	V	G
Benchmark regime	10 percent significance level							
Whole sample (WS)	NC	0.998	0.954	0.088	NC	1.000	0.982	0.181
Burns (B)	0.991	NC	0.985	0.923	1.000	NC	1.000	0.996
Volcker (V)	0.932	0.996	NC	0.513	0.983	1.000	NC	0.906
Greenspan (G)	0.023	0.968	0.644	NC	0.101	0.999	0.899	NC
Benchmark regime	5 percent significance level							
Whole sample	NC	0.997	0.922	0.048	NC	1.000	0.971	0.101
Burns	0.985	NC	0.963	0.896	1.000	NC	1.000	0.993
Volcker	0.880	0.987	NC	0.374	0.974	1.000	NC	0.845
Greenspan	0.011	0.955	0.483	NC	0.043	0.998	0.844	NC
Benchmark regime	1 percent significance level							
Whole sample	NC	0.994	0.821	0.022	NC	1.000	0.937	0.032
Burns	0.962	NC	0.860	0.833	0.999	NC	1.000	0.984
Volcker	0.695	0.936	NC	0.175	0.948	1.000	NC	0.655
Greenspan	0.007	0.913	0.215	NC	0.013	0.996	0.680	NC

Note: NC is shorthand notation for not computed. The probability in each entry is formally defined as $\Pr\left(F_{MD} > F_{0.\mathrm{XX}}^{MD} \mid F_{TR} > F_{0.\mathrm{XX}}^{TR}\right)$where F_{TR} denotes the computed Chow statistic for the monetary policy rule, F_{MD} the corresponding statistic for money demand, and XX the i'th percentile in the simulated distributions (90, 95 or 99'th percentile) under the null hypothesis. Under the null hypothesis, the Chow statistic follows the F-distribution with $k, T-2k$ degrees of freedom where k is the number of parameter restrictions (here, $k = 4$ for both money demand and the Taylor rule) and T is the number of observations.

Table B.2: Power of the super exogeneity test in small samples: Chow test probabilites of rejecting stability in consumption when stability in the monetary policy rule is rejected.

	$T = 100$				$T = 200$			
	Comparison regime				Comparison regime			
	WS	B	V	G	WS	B	V	G
Benchmark regime	10 percent significance level							
Whole sample (WS)	NC	0.997	0.962	0.127	NC	1.000	0.992	0.149
Burns (B)	0.996	NC	0.984	0.973	1.000	NC	1.000	0.999
Volcker (V)	0.979	0.998	NC	0.922	0.995	1.000	NC	0.987
Greenspan (G)	0.051	0.990	0.897	NC	0.094	0.999	0.968	NC
Benchmark regime	5 percent significance level							
Whole sample	NC	0.995	0.940	0.120	NC	1.000	0.985	0.108
Burns	0.994	NC	0.959	0.955	1.000	NC	1.000	0.999
Volcker	0.968	0.995	NC	0.866	0.992	1.000	NC	0.980
Greenspan	0.040	0.982	0.840	NC	0.061	0.998	0.951	NC
Benchmark regime	1 percent significance level							
Whole sample	NC	0.987	0.877	0.072	NC	0.998	0.965	0.088
Burns	0.985	NC	0.840	0.907	0.999	NC	0.999	0.996
Volcker	0.935	0.975	NC	0.663	0.984	1.000	NC	0.958
Greenspan	0.035	0.957	0.671	NC	0.038	0.996	0.897	NC

Note: NC is shorthand notation for not computed. The probability in each entry is formally defined as $\Pr\left(F_{CF} > F^{CF}_{0.\mathrm{XX}} \mid F_{TR} > F^{TR}_{0.\mathrm{XX}}\right)$ where F_{TR} denotes the computed Chow statistic for the monetary policy rule, F_{CF} the corresponding statistic for consumption, and XX the i'th percentile in the simulated distributions ($90, 95$ or 99'th percentile) under the null hypothesis. Under the null hypothesis, the Chow statistic follows the F-distribution with $k, T-2k$ degrees of freedom where k is the number of parameter restrictions (here, $k = 4$ for both consumption and the Taylor rule) and T is the number of observations.

B.3 Robustness of results with one more lag included in money demand and consumption

Since there are lag operators in the parameters in (21) and (23), I have checked the robustness of the results presented in the paper by replacing the regression equations (21) and (23) with

$$\begin{aligned} \ln\left(\frac{M_{t+1}}{P_t}\right) &= \beta_0 + \beta_1 \ln Y_t + \beta_2 \ln Y_{t-1} + \beta_3 R_t + \beta_4 R_{t-1} \\ &\quad + \beta_5 \ln\left(\frac{M_t}{P_{t-1}}\right) + \beta_6 \ln\left(\frac{M_{t-1}}{P_{t-2}}\right) + \varepsilon_{MD,\ t} \end{aligned} \tag{B.2}$$

and

$$\begin{aligned} \ln C_t &= \gamma_0 + \gamma_1 \ln Y_t + \gamma_2 \ln Y_{t-1} + \gamma_3 R_t + \gamma_4 R_{t-1} \\ &\quad + \gamma_5 \ln C_{t-1} + \gamma_6 \ln C_{t-2} + \varepsilon_{CF,\ t}. \end{aligned} \tag{B.3}$$

The results for these equations along with the monetary policy rule (8), corresponding to Tables 4 and 5 in the paper, are presented in Tables B.3 and B.4.

As we can from Tables B.3 and B.4, the results are not at all much different from those in Tables 4 and 5. The super exogeneity test still has unsatisfactory power in this case in most cases, although the average power probabilities for $T = 100$ have increased from about 0.30 and 0.28 to 0.36 and 0.37 (the regime shift from "Whole sample" to Greenspan is excluded since it is not statistically significant for the given sample size) for money demand and consumption respectively.

B.4 Robustness of results when estimating with IV instead of OLS

Since there are endogeneity issues in the error terms in (21) and (23), and also in the monetary policy rule (8), I have checked the robustness of the results presented in the paper by estimating with instrumental variables (IV) instead of OLS. As instruments, I use the regressors lagged one period.

Table B.3: Power of the super exogeneity test in small samples: Chow test probabilites of rejecting stability in (B.2) when stability in the monetary policy rule is rejected.

	$T = 100$				$T = 200$			
	Comparison regime				Comparison regime			
	WS	B	V	G	WS	B	V	G
Benchmark regime	10 percent significance level							
Whole sample (WS)	NC	0.425	0.250	0.151	NC	0.535	0.252	0.138
Burns (B)	0.828	NC	0.154	0.684	0.977	NC	0.935	0.922
Volcker (V)	0.663	0.624	NC	0.444	0.847	0.915	NC	0.618
Greenspan (G)	0.249	0.433	0.199	NC	0.295	0.632	0.222	NC
Benchmark regime	5 percent significance level							
Whole sample	NC	0.269	0.143	0.094	NC	0.319	0.120	0.081
Burns	0.746	NC	0.553	0.607	0.958	NC	0.904	0.881
Volcker	0.576	0.503	NC	0.344	0.780	0.872	NC	0.522
Greenspan	0.154	0.316	0.124	NC	0.193	0.438	0.119	NC
Benchmark regime	1 percent significance level							
Whole sample	NC	0.082	0.038	0.038	NC	0.056	0.014	0.024
Burns	0.583	NC	0.386	0.478	0.900	NC	0.804	0.780
Volcker	0.406	0.312	NC	0.214	0.633	0.750	NC	0.359
Greenspan	0.088	0.130	0.045	NC	0.060	0.115	0.020	NC

Note: NC is shorthand notation for not computed. Small sample critical values generated in Section 4 are used. The probability in each entry is formally defined as $\Pr\left(F_{MD} > F^{MD}_{0.\mathrm{XX}} \mid F_{TR} > F^{TR}_{0.\mathrm{XX}}\right)$where F_{TR} denotes the computed Chow statistic for the monetary policy rule, F_{MD} the corresponding statistic for money demand, and XX the i'th percentile in the simulated distributions (90, 95 or 99'th percentile) under the null hypothesis. Under the null hypothesis, the Chow statistic follows the F-distribution with $k, T-2k$ degrees of freedom where k is the number of parameter restrictions (here, $k = 8$ for both money demand and the Taylor rule) and T is the number of observations.

Table B.4: Power of the super exogeneity test in small samples: Chow test probabilites of rejecting stability in (B.3) when stability in the monetary policy rule is rejected.

	$T = 100$				$T = 200$			
	Comparison regime				Comparison regime			
	WS	B	V	G	WS	B	V	G
Benchmark regime	10 percent significance level							
Whole sample (WS)	NC	0.469	0.314	0.150	NC	0.497	0.320	0.141
Burns (B)	0.781	NC	0.616	0.671	0.914	NC	0.922	0.891
Volcker (V)	0.705	0.613	NC	0.490	0.820	0.918	NC	0.642
Greenspan (G)	0.248	0.451	0.233	NC	0.293	0.608	0.260	NC
Benchmark regime	5 percent significance level							
Whole sample	NC	0.333	0.193	0.086	NC	0.343	0.194	0.074
Burns	0.713	NC	0.509	0.587	0.873	NC	0.873	0.836
Volcker	0.628	0.516	NC	0.400	0.763	0.859	NC	0.546
Greenspan	0.165	0.340	0.144	NC	0.201	0.438	0.153	NC
Benchmark regime	1 percent significance level							
Whole sample	NC	0.141	0.058	0.030	NC	0.146	0.060	0.032
Burns	0.585	NC	0.338	0.446	0.774	NC	0.724	0.698
Volcker	0.455	0.300	NC	0.231	0.646	0.689	NC	0.360
Greenspan	0.073	0.157	0.048	NC	0.088	0.182	0.043	NC

Note: NC is shorthand notation for not computed. Small sample critical values generated in Section 4 are used. The probability in each entry is formally defined as $\Pr\left(F_{CF} > F^{CF}_{0.\mathrm{XX}} \mid F_{TR} > F^{TR}_{0.\mathrm{XX}}\right)$where F_{TR} denotes the computed Chow statistic for the monetary policy rule, F_{CF} the corresponding statistic for consumption, and XX the i'th percentile in the simulated distributions $(90, 95$ or 99'th percentile) under the null hypothesis. Under the null hypothesis, the Chow statistic follows the F-distribution with $k, T - 2k$ degrees of freedom where k is the number of parameter restrictions (here, $k = 8$ for both consumption and the Taylor rule) and T is the number of observations.

The results of this exercise, corresponding to Tables 4 and 5 in the paper, are presented in Tables B.5 and B.6.[3]

Table B.5: Power of the super exogeneity test in small samples: Chow test probabilites of rejecting stability in (21) when stability in the monetary policy rule is rejected.

	$T = 100$				$T = 200$			
	Comparison regime				Comparison regime			
	WS	B	V	G	WS	B	V	G
Benchmark regime	10 percent significance level							
Whole sample (WS)	NC	0.	0.	0.	NC	0.	0.	0.
Burns (B)	0.	NC	0.	0.	0.	NC	0.	0.
Volcker (V)	0.	0.	NC	0.	0.	0.	NC	0.
Greenspan (G)	0.	0.	0.	NC	0.	0.	0.	NC
Benchmark regime	5 percent significance level							
Whole sample	NC	0.	0.	0.	NC	0.	0.	0.
Burns	0.	NC	0.	0.	0.	NC	0.	0.
Volcker	0.	0.	NC	0.	0.	0.	NC	0.
Greenspan	0.	0.	0.	NC	0.	0.	0.	NC
Benchmark regime	1 percent significance level							
Whole sample	NC	0.	0.	0.	NC	0.	0.	0.
Burns	0.	NC	0.	0.	0.	NC	0.	0.
Volcker	0.	0.	NC	0.	0.	0.	NC	0.
Greenspan	0.	0.	0.	NC	0.	0.	0.	NC

Note: NC is shorthand notation for not computed. Small sample critical values generated in Section 4 are used. The probability in each entry is formally defined as $\Pr\left(F_{MD} > F^{MD}_{0.\mathrm{XX}} \mid F_{TR} > F^{TR}_{0.\mathrm{XX}}\right)$where F_{TR} denotes the computed Chow statistic for the monetary policy rule, F_{MD} the corresponding statistic for money demand, and XX the i'th percentile in the simulated distributions (90, 95 or 99'th percentile) under the null hypothesis. Under the null hypothesis, the Chow statistic follows the F-distribution with $k, T-2k$ degrees of freedom where k is the number of parameter restrictions (here, $k = 4$ for both money demand and the Taylor rule) and T is the number of observations.

As we can from Tables B.1 and B.2, the results are not at all much different from those in Tables 4 and 5. The super exogeneity test have low power in this case as well in almost every case.

[3] It should be mentioned, however, that the IV estimation results are more unstable than the OLS results, since all variables in the model are stationary. This makes it hard to find instrumental variables that are highly correlated with the regressors (but not with the error term).

Table B.6: Power of the super exogeneity test in small samples: Chow test probabilites of rejecting stability in (23) when stability in the monetary policy rule is rejected.

	$T = 100$				$T = 200$			
	Comparison regime				Comparison regime			
	WS	B	V	G	WS	B	V	G
Benchmark regime	10 percent significance level							
Whole sample (WS)	NC	0.	0.	0.	NC	0.	0.	0.
Burns (B)	0.	NC	0.	0.	0.	NC	0.	0.
Volcker (V)	0.	0.	NC	0.	0.	0.	NC	0.
Greenspan (G)	0.	0.	0.	NC	0.	0.	0.	NC
Benchmark regime	5 percent significance level							
Whole sample	NC	0.	0.	0.	NC	0.	0.	0.
Burns	0.	NC	0.	0.	0.	NC	0.	0.
Volcker	0.	0.	NC	0.	0.	0.	NC	0.
Greenspan	0.	0.	0.	NC	0.	0.	0.	NC
Benchmark regime	1 percent significance level							
Whole sample	NC	0.	0.	0.	NC	0.	0.	0.
Burns	0.	NC	0.	0.	0.	NC	0.	0.
Volcker	0.	0.	NC	0.	0.	0.	NC	0.
Greenspan	0.	0.	0.	NC	0.	0.	0.	NC

Note: NC is shorthand notation for not computed. Small sample critical values generated in Section 4 are used. The probability in each entry is formally defined as $\Pr\left(F_{CF} > F^{CF}_{0.\mathrm{XX}} \mid F_{TR} > F^{TR}_{0.\mathrm{XX}}\right)$where F_{TR} denotes the computed Chow statistic for the monetary policy rule, F_{CF} the corresponding statistic for consumption, and XX the i'th percentile in the simulated distributions ($90, 95$ or 99'th percentile) under the null hypothesis. Under the null hypothesis, the Chow statistic follows the F-distribution with $k, T - 2k$ degrees of freedom where k is the number of parameter restrictions (here, $k = 4$ for both consumption and the Taylor rule) and T is the number of observations.