Working Paper Series

The Lucas Critique and the Stability of Empirical Models

WP 06-05

Thomas A. Lubik
Federal Reserve Bank of Richmond

Paolo Surico
Bank of England and
University of Bari

THE FEDERAL RESERVE BANK OF RICHMOND
RICHMOND ▪ BALTIMORE ▪ CHARLOTTE

The Lucas Critique and the Stability of Empirical Models*

Thomas A. Lubik
Federal Reserve Bank of Richmond

Paolo Surico
Bank of England and University of Bari

July 2006

Working Paper No. 06-05

Abstract

This paper re-considers the empirical relevance of the Lucas critique using a DSGE sticky price model in which a *weak* central bank response to inflation generates equilibrium indeterminacy. The model is calibrated on the magnitude of the historical shift in the Federal Reserve's policy rule and is capable of generating the decline in the volatility of inflation and real activity observed in U.S. data. Using Monte Carlo simulations and a backward-looking model of aggregate supply and demand, we show that shifts in the policy rule induce breaks in both the reduced-form coefficients and the reduced-form error variances. The statistics of popular parameter stability tests are shown to have low power if such heteroskedasticity is neglected. In contrast, when the instability of the reduced-form error variances is accounted for, the Lucas critique is found to be empirically relevant for both artificial and actual data.

JEL Classification: C52, E38, E52.

Key Words: Lucas critique, heteroskedasticity, parameter stability tests, rational expectations, indeterminacy.

*We are grateful to Luca Benati, Andrew Blake, Jon Faust, Jan Groen, Haroon Mumtaz, Serena Ng, Christoph Schleicher, Frank Schorfheide, Shaun Vahey and seminar participants at the Reserve Bank of New Zealand, the Bank of England, the 2006 meeting of the Society of Computational Economics, the conference on "Macroeconometrics and Model Uncertainty" held on 27-28 June 2006 at the Reserve Bank of New Zealand, and our discussant Timothy Kam for valuable comments and suggestions. The views expressed in this paper are those of the authors, and do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System, nor the Bank of England or the Monetary Policy Committee.

Address for correspondence: Thomas A. Lubik: Research Department, Federal Reserve Bank of Richmond, P.O. Box 27622, Richmond, VA 23261. Thomas.Lubik@rich.frb.org. Paolo Surico: Bank of England, Threadneedle Street, London, EC2R 8AH. paolo.surico@bankofengland.co.uk.

1 Introduction

The Lucas critique comes as close to a natural law as seems possible in macroeconomics. While its theoretical validity is largely uncontested, there is, however, a surprising lack of empirical support in the literature. A host of studies have argued that the relevance of the Lucas critique is limited in practice. In particular, shifts in policy seem not to have any significant effects on the stability of backward representations of macroeconomic models for various historical episodes. We argue in this paper that the evident inapplicability of the Lucas critique is due to problems with the size and power of the econometric tests used. We use a simple structural model of the U.S. economy as a data-generating process (DGP) to illustrate these issues both conceptually and by means of a Monte Carlo analysis. Our empirical findings confirm that the Lucas critique is relevant for the shift in U.S. monetary policy behavior in the early 1980s.

It is difficult to underestimate the importance of Lucas (1976) in the development of modern macroeconomic thought. The introduction of rational expectations in macroeconomics in the middle of the 1970s represented an intellectual revolution for the profession and a serious challenge for large-scale, backward-looking econometric models that were used for policy analysis. Lucas argued that changes in policy have an immediate effect on agents' decision rules since they are inherently forward-looking and adapt to the effects of the new policy regime. An important corollary of this argument is that any policy evaluation based on backward-looking macroeconomic models is misleading whenever such policy shifts occur.

Empirical studies of the validity of the Lucas critique (e.g. Estrella and Fuhrer, 2003, Rudebusch, 2005) suggest that it is unimportant in practice. Tests for parameter stability in backward-looking specifications or reduced forms of macroeconomic relationships typically fail to reject the null of structural stability in the presence of well-documented policy shifts. This evidence would support the conclusion that policy changes are, in the terminology of Leeper and Zha (2003), 'modest' enough not to alter the behavior of private agents in a manner that is detectable by the econometrician. A further implication is that backward-looking monetary models of the type advocated by Rudebusch and Svensson (1999), which perform well empirically, are safe to use in policy experiments.

We argue in this paper that the failure of these studies to detect Lucas-critique effects rests on the use of parameter stability tests such as the Chow break-point test and the superexogeneity test. These tests implicitly assume equality of the variances of the reduced-form innovations across sub-samples. It is, however, well known, at least since Toyoda

(1974), that even moderate heteroskedasticity has a considerable impact on the significance levels of the null of parameter stability, which leads to an incorrect acceptance of parameter invariance. Moreover, and similar to Lindé (2001), we also find evidence of a small-sample bias which tends to hide the instability of backward-looking specifications.

At the heart of Lucas' critique lies an emphasis on the use of fully-specified, optimisation-based dynamic stochastic general equilibrium (DSGE) models for policy analysis. We follow this approach and work within the confines of a structural New Keynesian model that has been widely used for monetary policy analysis. We treat this model as our DGP from which we generate simulated time series for a Monte Carlo analysis. This allows us to control the environment in which the policy change occurs. Otherwise, it might be difficult to distinguish between actual policy shifts and changes in the economy's driving processes.

Our main experiment is to model the monetary policy change in the U.S. that occurred at the turn of the 1980s. We follow Clarida et al. (2000) and Lubik and Schorfheide (2004) in assuming that policy during the 1970s responded only weakly to inflation, whereas with the start of Volcker's tenure policy became much more aggressively anti-inflationary. In the context of our model, this implies that in the first sub-sample the rational expectations (RE) equilibrium was indeterminate, and determinate later on. On the basis of the approach developed by Lubik and Schorfheide (2003) to describe indeterminate equilibria, we show, first, that a policy shift of the magnitude observed in the Federal Reserve's monetary policy rule is capable of reproducing the historical decline in the volatility of real activity, inflation, and the interest rate. Furthermore, we demonstrate that such a decline invalidates the assumption of constant reduced-form innovation variances in a widely used backward-looking model of aggregate supply and aggregate demand estimated over the sub-samples associated with the regime shift.

Monte Carlo simulations show that the instability of the error variances severely affects the power of the Chow test and the superexogeneity test to the effect that it prevents the rejection of the *incorrect* null hypothesis of parameter stability. When the heteroskedasticity in the estimated backward-looking specifications is accounted for, we find robust evidence in favor of the empirical relevance of the Lucas critique. An application to U.S. data reveals that controlling for the decline in the variance of the estimated error terms matters also in practice. We show that this overturns the results from parameter stability tests which are based on the (incorrect) assumption of equal error variance between sub-samples. Hence, we conclude that the Lucas critique is alive and well.

To our knowledge, this is the first paper that investigates the effect of equilibrium indeterminacy on the stability of reduced-form models. The message of the paper is, however, not predicated on policy changes that induce switches between determinacy and indeterminacy, but this assumption helps to sharpen our argument. We also report results from simulations that preserve a specific regime.

The paper is organized as follows. Section 2 presents and discusses analytical results in the context of simple examples on how policy changes in structural models affect their reduced-form representations and how this relates to the debate about the empirical validity of the Lucas critique. Section 3 introduces the structural model that we use as DGP for the Monte Carlo study. This section also shows that the policy-induced shift from indeterminacy to determinacy is capable of explaining the U.S. Great Moderation. Section 4 describes the simulation strategy for testing the empirical relevance of the Lucas critique. In section 5, we report the Monte Carlo evidence on the structural stability of a backward-looking model of aggregate supply and demand, and provide a sensitivity analysis for variations in the values of some parameters of the structural model. The sixth section contains empirical results obtained on actual U.S. data, while the last section concludes.

2 Rational Expectations Equilibria and the Lucas Critique

We assume throughout our paper that the data are generated by a DSGE model. Analyzing the effects of parameter changes thus requires an understanding of the reduced-form properties of structural linear rational expectations models. In this sense, our paper is similar to Lindé (2001) and Rudebusch (2005). However, we go further than the earlier literature in analyzing the effects of a policy break that changes the equilibrium properties of the economy, specifically, a change from indeterminacy to determinacy. In this section, we therefore give a brief overview on indeterminacy in linear rational expectations models.

2.1 Determinate and Indeterminate Equilibria

Consider the simple expectational difference equation:

$$x_t = aE_t x_{t+1} + \varepsilon_t, \tag{1}$$

where a is a parameter, ε_t is a white noise process with mean zero and variance σ^2, and E_t is the rational expectations operator conditional on information at time t. It is well known that the type of solution depends on the value of the parameter a. If $|a| < 1$ there is a unique ('determinate') solution which is simply:

$$x_t = \varepsilon_t.$$

On the other hand, if $|a| > 1$, there are multiple solutions and the rational expectations equilibrium is indeterminate.

In order to derive the entire set of solutions we follow the approach developed by Lubik and Schorfheide (2003). For this purpose it is often convenient to rewrite the model by introducing endogenous forecast errors $\eta_t = x_t - E_{t-1}x_t$. Define $\xi_t = E_t x_{t+1}$ so that equation (1) can be rewritten as:

$$\xi_t = \frac{1}{a}\xi_{t-1} - \frac{1}{a}\varepsilon_t + \frac{1}{a}\eta_t. \tag{2}$$

Under indeterminacy this is a stable difference equation which imposes no further restrictions on the evolution of the endogenous forecast error η_t.[1] Hence, any covariance-stationary stochastic process for η_t is a solution for this linear rational expectations model. The forecast error can then be expressed as a linear combination of the model's fundamental disturbances and extraneous sources of uncertainty, typically labeled 'sunspots'. The coefficients on the shocks in this decomposition generally depend on parameters of the model and on parameters that index a specific sunspot equilibrium. To wit, the forecast error can be written as:

$$\eta_t = m\varepsilon_t + \zeta_t,$$

where the sunspot ζ_t is a martingale-difference sequence, and m is an unrestricted parameter.[2] Substituting this into Eq. (2) yields the full solution under indeterminacy:

$$x_t = \frac{1}{a}x_{t-1} + m\varepsilon_t - \frac{1}{a}\varepsilon_{t-1} + \zeta_t.$$

The evolution of x_t now depends on an additional (structural) parameter m which indexes specific rational expectations equilibria.

Indeterminacy affects the behavior of the model in three main ways. First, indeterminate solutions exhibit a much richer lag structure and more persistence than the corresponding determinate solution. This feature has been exploited by Lubik and Schorfheide (2004) in distinguishing between the two types of rational expectations equilibria in U.S. data. In the simple example, this is strikingly evident: under determinacy the solution for x_t is white noise, while under indeterminacy the solution is described by an ARMA(1,1) process. Specifically, the (composite) error term exhibits both serial correlation and a different variance when compared to the determinate solution. Second, under indeterminacy

[1] In the case of determinacy, the restriction imposed is that $\xi_t = 0$, $\forall\, t$, which implies $\eta_t = \varepsilon_t$.

[2] There is a technical subtlety in that ζ_t is actually, in the terminology of Lubik and Schorfheide (2003), a reduced-form sunspot shock, with $\zeta_t = m_\zeta \zeta_t^*$. Furthermore, in less simple models, there would be additional restrictions on the coefficients which depend on other structural parameters.

sunspot shocks can affect equilibrium dynamics. Other things being equal, data generated by sunspot equilibria are inherently more volatile than their determinate counterparts. The third implication, especially emphasized by Lubik and Schorfheide (2003), is that indeterminacy affects the response of the model to *fundamental* shocks, whereas the response to sunspot shocks is uniquely determined. In the example, innovations to ε_t could either increase or decrease x_t depending on the sign of m.

2.2 Indeterminacy and the Lucas Critique

We now try to provide some insight into the analytics of the Lucas critique by means of a simple example. We consider the following two-equation model which describes the evolution of an economic variable x_t and a policy variable y_t:

$$x_t = aE_t x_{t+1} + bE_t y_{t+1} + \varepsilon_{1t}, \tag{3}$$

$$y_t = cx_t + \varepsilon_{2t}. \tag{4}$$

ε_{1t}, ε_{2t} are white noise processes with variances σ_1^2, σ_2^2, respectively. a, b are structural parameters, and c is a policy parameter. We assume for simplicity that all parameters are positive. Eq. (4) is a feedback rule of the type often used in monetary policy models.

The model has a unique rational expectations equilibrium if $0 < c < \frac{1-a}{b}$, the solution of which is:

$$\begin{aligned} x_t &= \varepsilon_{1t}, \\ y_t &= c\varepsilon_{1t} + \varepsilon_{2t}. \end{aligned}$$

As in the simple example above, both variables are white noise processes. If $c > \frac{1-a}{b}$ the solution is indeterminate. The laws of motion for the two variables are as follows:

$$\begin{aligned} x_t &= \frac{1}{a+bc} x_{t-1} + m\varepsilon_{1t} - \frac{1}{a+bc}\varepsilon_{1t-1} + \zeta_t, \\ y_t &= \frac{1}{a+bc} y_{t-1} + mc\varepsilon_{1t} - \frac{c}{a+bc}\varepsilon_{1t-1} + \varepsilon_{2t} - \frac{1}{a+bc}\varepsilon_{2t-1} + c\zeta_t. \end{aligned}$$

Again, the change in the stochastic properties of the variables, when moving across the parameter space, is quite evident. If (3) - (4) is the DGP, then the reduced-form equations above are the representations upon which tests for the empirical relevance of the Lucas critique would have to be based.

The type of experiment we are interested in is an exogenous, unanticipated change in the policy parameter c. We can distinguish four different scenarios: a shift from determinacy to indeterminacy, from indeterminacy to determinacy, and changes that preserve previously determinate and indeterminate equilibria. If the break in c is such that the solution stays determinate, the behavior of x_t is unaffected while the variance of y_t changes. However, an econometrician could not distinguish between a change in σ_2^2 and the policy coefficient by observing y_t. A change in the variance of ε_{1t}, on the other hand, could be deduced from observations on both variables.

The more interesting case is a change in policy behavior that moves c across the boundary $\frac{1-a}{b}$ that separates the determinacy from the indeterminacy region. Suppose the economy is initially in an indeterminate equilibrium associated with a policy parameter c_0. An unexpected (and believed to be permanent) policy shift to $c_1 < \frac{1-a}{b} < c_0$ moves the economy to a determinate equilibrium. This implies a dramatic change in the nature of the stochastic processes for x_t and y_t, as they switch from persistence to white noise. Specifically, both their variance and the degree of auto-correlation decline. Furthermore, sunspot shocks, and the extra volatility associated with them, no longer affect the dynamics. It is this type of scenario we have in mind in our empirical analysis. The standard example is the change in monetary policy that occurred during Paul Volcker's tenure at the Board of Governors. Naturally, similar reasoning applies to the opposite case when a policy change moves the economy from a determinate to an indeterminate equilibrium.

Yet, even if equilibrium properties are unaffected by shifts in policy, the properties of the reduced-form model are not. Crucially, the variance of the error term in a regression of the endogenous variables on their first lag would (i) change with the shift in policy, and (ii) be correlated with the regressor. The latter issue is, of course, well known, and methods to deal with this, such as IV-estimation, are now widely used.[3] The former is not as well appreciated, however. Consequently, the main thrust of our paper is directed at this form of heteroskedasticity, which reduces the power of parameter stability tests based on reduced-form regressions.[4]

3 The Structural Model

In this section, we lay out a small-scale sticky-price monetary model which will serve as DGP for the Monte Carlo experiments. After describing the calibration, we use the structural

[3] See, however, Lubik and Schorfheide (2005) for a set of examples in the context of DSGE-models where IV-methods fail.

[4] Structural estimation methods are immune against this problem since the structure of the reduced-form error would be reflected in, say, the likelihood function. This is likely the main reason why the GMM-based results of Collard et al. (2002) are an outlier in the literature in that they find strong evidence for the empirical relevance of the Lucas critique.

model to assess the extent to which the 'good policy' and the 'good luck' explanations of the U.S. Great Moderation are capable of replicating the facts.

3.1 Specification

Our simulation analysis is based on a log-linearized, microfounded New-Keynesian sticky price model of the business cycle of the kind popularized by Clarida et al. (1999), King (2000) and Woodford (2003) among others. The model consists of three aggregate relationships which describe the dynamic behavior of output y_t, inflation π_t, and the nominal interest rate R_t:

$$y_t = E_t y_{t+1} - \tau(R_t - E_t \pi_{t+1}) + g_t, \tag{5}$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa (y_t - z_t), \tag{6}$$

$$R_t = \rho_R R_{t-1} + (1 - \rho_R) \left(\psi_\pi \pi_t + \psi_y (y_t - z_t) \right) + \varepsilon_{R,t}. \tag{7}$$

All variables are expressed in percentage deviations from the steady state.

Eq. (5) is a log-linearized IS-curve derived from a household's intertemporal optimization problem in which consumption and nominal bond holdings are the control variables. Since there is no physical capital in this economy, consumption is proportional to total resources up to an exogenous process g_t. The latter is typically interpreted as a government spending shock or a shock to preferences.[5] The parameter $\tau > 0$ represents the intertemporal elasticity of substitution.

The Phillips-curve relationship (6) describes inflation dynamics as a function of output. It captures the staggered feature of an economy with Calvo-type price setting in which firms adjust their optimal price with a constant probability in any period, independently of the time elapsed from the last adjustment. The discrete nature of price setting creates an incentive to adjust prices by more the higher expected future inflation is. The contemporaneous rate of inflation is thus related to the difference between output and the stochastic marginal cost of production z_t via the parameter $\kappa > 0$, which can be interpreted as the inverse of the sacrifice ratio. $0 < \beta < 1$ is the agents' discount factor.

The policy rule (7) characterizes the behavior of the monetary authorities according to which the central bank adjusts the policy variable in response to current inflation and the output gap $(y_t - z_t)$. ψ_π, $\psi_y \geq 0$ are the policy coefficients. These adjustments are implemented smoothly, with $0 < \rho_R < 1$ capturing the degree of interest rate inertia. The

[5] The IS curve can easily be reinterpreted as a schedule explaining the behavior of the 'output gap' defined as the difference between actual output and the hypothetical flexible price level of output (see Clarida et al. 1999). In this case, the shock g_t is also a source of potential output variations.

random variable $\varepsilon_{R,t}$ stands for the monetary policy shock, which can be interpreted either as unexpected deviations from the policy rule or as a policy implementation error. We assume it to be a white noise process with mean zero and variance σ_{ε}^2.

The model description is completed by specifying the stochastic properties of the exogenous shocks g_t and z_t. We assume they are first-order autoregressive processes with lag-coefficients $0 \leq \rho_g, \rho_z < 1$. Their innovations are assumed to be *i.i.d.* with variances σ_g^2 and σ_z^2, respectively. The equation system (5) - (7), together with the two processes for the exogenous shocks, describes a linear rational expectations model that can be solved using the methods described in Sims (2002). We solve the model for both determinacy and indeterminacy. In the latter case, we follow the approach developed by Lubik and Schorfheide (2003).

3.2 Calibration

We choose parameter values in line with the estimates in Lubik and Schorfheide (2004), who analyzed a model similar to ours. The values of the structural parameters are reported in Panel A of Table 1. We assign the same values in both subsamples to focus on changes in the policy coefficients. Unlike Lubik and Schorfheide (2004), however, we set the variance of sunspot shocks to zero in the case of indeterminacy. This is designed to facilitate the comparison between our simulations and the previous literature, which does not take the presence of sunspot fluctuations into account.

A large body of empirical literature has documented the dramatic change in the conduct of U.S. monetary policy at the turn of the 1980s. The established consensus is that the nominal interest rate response to inflation in estimated policy rules became substantially more aggressive in the early 1980s. As the number of available estimates is quite large, we set the coefficients in the policy rules for Periods I and II to values that are broadly in line with the evidence presented by Clarida et al. (2000). Within the indeterminacy and determinacy regions, the results are fairly robust to deviations from the baseline values reported in Panel B of Table 1.

The interest rate response to inflation over the indeterminacy sample does not guarantee a unique rational expectations equilibrium because $\psi_\pi = 0.4$ violates the Taylor principle.[6] On the other hand, the parameter constellation associated with Period II guarantees a determinate equilibrium. As detailed above, the solution of the model under indeterminacy

[6] The additional requirement is that the other policy coefficients in the rule are not 'too large' to compensate for the weak inflation response. The analytics and intuition of indeterminate equilibria in the New Keynesian monetary model are discussed extensively by Lubik and Marzo (2006).

is affected along various dimensions. Most importantly, equilibrium dynamics can be driven by sunspot fluctuations. Secondly, the transmission of fundamental shocks can be altered relative to the unique rational expectations solution. Having set the variance of sunspot disturbances to zero, we will thus focus on the transmission mechanism effect.

3.3 Replicating the Stylized Facts

Before evaluating the empirical relevance of the Lucas critique, we assess the extent to which the magnitude of historical policy shifts can account for the decline in volatility and persistence of inflation, output, and the interest rate observed in U.S. data at the beginning of the 1980s. A large number of contributions, including Kim and Nelson (1999), McConnell and Perez-Quiros (2000), and Blanchard and Simon (2001), document a sharp decline in the volatility of many U.S. macroeconomic variables. The variance of inflation declined by more than two thirds and the variance of output by just less than one half. Following Stock and Watson (2002), this set of facts has come to be known as the Great Moderation.

There is, however, a simmering debate whether this was due to good luck (a decline in the volatility of exogenous shocks) or good policy (a shift to a more stabilizing policy regime). Empirical evidence by Stock and Watson (2002) and Kim et al. (2004) shows that a sizeable part of such decline was driven by a reduction in the volatility of the reduced-form innovations of estimated backward-looking specifications for inflation and output. This has been invariably interpreted as *prima facie* evidence in favor of good luck.[7]

We thus ask the question whether the good luck and good policy scenarios are capable of explaining this decline in volatilities in the context of our model. We present results for two sets of model calibrations. In the first experiment, we change the policy parameters such that the equilibrium switches from indeterminacy to determinacy. Differences in the dynamics are therefore driven only by a shift in the monetary policy rule, as all non-policy parameters of the model are kept fixed across regimes. In the second calibration, we simulate a decline in the variance of the structural shocks such as to match the magnitude of the decline in the volatility of inflation and output observed in the data. All other parameters are fixed in the second simulation. In particular, the Taylor principle applies and the equilibrium is therefore always determinate.

Simulation results for the good policy hypothesis are reported in Table 2. Panel A displays the standard deviation of output, inflation and the interest rate. All variables are far

[7] In work in progress, Benati and Surico show, however, that this interpretation is unwarranted since a move from indeterminacy to determinacy is also consistent with a decline in the innovation variances of VAR models.

less volatile in the post-1982 simulated sample. The volatility of output and the interest rate declined by 42%, while the standard deviation of inflation was reduced by 67% when moving from indeterminacy to determinacy. Panel B reports a measure of persistence obtained as the sum of autoregressive coefficients from the estimates of an univariate $AR(p)$ process. In line with the evidence in Cogley and Sargent (2005), the most dramatic change occurs for inflation whose inertia drops by 58%. The impact on the interest rate is considerably smaller. This can be explained by the higher smoothing parameter under determinacy which outweighs the impact of the 'extra' persistence induced by monetary policy under indeterminacy.[8]

We contrast this with the statistics obtained under the good luck hypothesis, reported in Table 3. While a decline in the variance of the structural shocks can replicate the drop in volatility, it cannot match the observed decrease in persistence, in particular the well-documented decline in inflation persistence. This strongly points toward a change in the transmission mechanism. A similar conclusion has been reached by McConnell and Perez-Quiros (2000), Boivin and Giannoni (2005) and Canova, Gambetti and Pappa (2005), pinpointing monetary policy changes. Our specific explanation is the shift from an indeterminate to a determinate equilibrium induced by a change in the Federal Reserve's policy behavior. As we demonstrated in Section 2 such a shift changes the dynamic properties of the reduced form of the model.

We also want to emphasize that the Lucas critique is economically relevant in our setting. Two metrics that are typically considered for this assessment are the sum of the autoregressive coefficients of inflation and the value of a loss function that weighs the unconditional variances of inflation and output. The results of Table 2 indicate that the critique is, indeed, economically important in that both the persistence of inflation and the sum of the unconditional variances of inflation and output decline drastically in the move to a more anti-inflationary policy.

Overall, we find that both explanations are capable of explaining the Great Moderation. However, only the good policy hypothesis successfully reproduces the decline in inflation persistence. A shift from passive to active monetary policy can account *on its own* for the sharp decline in the volatility and persistence of inflation and output observed in the early 1980s. Furthermore, the decline in the volatility of the series is likely to contaminate the

[8] On the other hand, output persistence actually increases, which is likely due to the behavior of the ex-ante real interest rate. Across regimes, the persistence of the nominal interest rate is almost unchanged and close to 0.8. In contrast, the persistence of expected inflation, whose behavior is similar to actual inflation, declines remarkably, from 0.7 to 0.3. This implies that the persistence of the ex-ante real interest rate is higher in the second sub-sample, which makes output more persistent.

properties of the error terms in estimated reduced-form models. As many widely used break point tests like the Chow test are based on the assumption of constancy of the innovations variance, we show below that the neglected heteroskedasticity leads to incorrect inference on the empirical relevance of the Lucas critique.

4 Testing for the Lucas Critique

This section presents a simple algorithm to investigate the empirical relevance of the Lucas critique. To focus on the importance of a change in monetary policy, we simulate a shift in the coefficients of the policy rule while keeping the parameters describing the structure of the economy fixed across simulations. For each simulation, we obtain two sets of artificial data. The first set is generated from an indeterminate equilibrium and is associated with the pre-1979 estimates of the policy rule typically found in the literature. The second sample is generated under the assumption of determinacy and it corresponds to the post-1979 description of policy behavior. Any difference in the estimates on the two sets of artificial data is thus only attributable to changes in policy. As the Lucas critique is about (in)stability of reduced-form parameters, we specify a backward-looking representation of the economy and then assess the impact of a policy shift on the reduced-form representation.

4.1 A Backward-Looking Model

The backward-looking model of aggregate supply and demand is a quarterly version of the specification in Svensson (1997), which has been used by Rudebusch and Svensson (1999) and Rudebusch (2005) among many others:

$$\pi_t = \alpha_1 \pi_{t-1} + \alpha_2 \pi_{t-2} + \alpha_3 \pi_{t-3} + \alpha_4 \pi_{t-4} + \alpha_y y_{t-1} + u_t^{\pi}, \tag{8}$$

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \beta_4 y_{t-4} + \beta_r (R - \pi)_{t-1} + u_t^{y}. \tag{9}$$

Inflation depends on its own past values within a year and on the lagged value of real activity. Aggregate demand is characterized by an autoregressive structure with four lags augmented by the lagged value of the real interest rate. The model is closed with a standard Taylor-type rule with interest rate smoothing:

$$R_t = \gamma_R R_{t-1} + \gamma_{\pi} \pi_t + \gamma_y y_t + u_t^R. \tag{10}$$

The model (8)-(10) is estimated on simulated data under the assumption that the DGP is the structural model (5)-(7).

We want to emphasize that we do not restrict the backward-looking model to the reduced-form representation of the DSGE model. In particular, the backward-looking specification is allowed, but not required, to have a richer lag structure relative to the DSGE model. In other words, we let the data choose what lags of the dependent variables have explanatory power. Our reasoning is twofold. First, earlier contributions have used very similar backward-looking specifications and we wish to contrast our results with those of the previous literature. Second, and more importantly, we have shown in Section 2 that equilibrium indeterminacy is an additional source of serial correlation in the error terms.[9] Augmenting the backward-looking specification with additional lags of the dependent variable is a simple way to make the residuals white noise and therefore control for serial correlation.

4.2 Simulation Strategy

To assess the empirical relevance of the Lucas critique, we design a Monte Carlo experiment in which we postulate a break of the magnitude of the historical shift in the Federal Reserve's policy rule. All other parameters of the structural model are kept fixed across policy regimes, and we are thus isolating any Lucas-critique effects.

The procedure in the simulations is as follows:

1. Solve the structural model under both indeterminacy and determinacy, and generate two samples of 82 and 61 observations[10] for output, inflation and the interest rate.[11]

2. For each artificial sample, estimate the backward-looking equations for output, inflation, and the interest rate.

3. Perform tests for error variance constancy and parameter stability of the backward-looking equations across the two periods; compute the relevant statistics and probability values.

4. Repeat steps 1. to 3. $20,000$ times; for each sub-period select the median values of the statistics of interest.

[9] We are grateful to Adrian Pagan for raising our awareness of this issue.

[10] The number of observations has been chosen to match the quarterly data points which are typically used in sub-sample analyses of US data (see Lubik and Schorfheide, 2004, and Clarida et al., 2000). The first period ranges from 1960:1 to 1979:2 while the second corresponds to 1982:4 to 1997:4. In each simulated sample, 100 extra observations are produced to provide us with a stochastic vector of initial conditions, which are then discarded.

[11] The solution of the model under indeterminacy is computed using the continuity assumption in Lubik and Schorfheide (2004). Similar results, not reported but available upon request, are obtained by imposing that the structural shocks are orthogonal to the sunspot shocks.

5. Prior to the simulations, Steps 1 - 4 are carried out under the assumption of no policy regime shift, and small-sample critical values are computed from the actual distributions of the relative statistics under such scenario. This ensures a correct size of the tests for parameter instability.

If the probability of rejecting parameter stability in Steps 1 - 4 computed under the assumption of a policy break is higher than the probability in Step 5 based on the assumption of no change in the policy parameters, then the Lucas critique is judged empirically relevant.

5 Results on Simulated Data

This section presents the results of the Monte Carlo analysis. We present the relevant statistics of the test for equal innovation variances and parameter stability together with some robustness checks.

5.1 Testing for Equal Innovation Variances

Table 4 reports residual standard deviations from the estimated model (8) - (10) together with the statistics of the Goldfeld-Quandt test of innovation variance constancy.[12] The variances of reduced-form innovations in the output and inflation equation exhibit a dramatic decline from the indeterminacy to the determinacy period of more than 50%. The decline for the interest rate equation is less dramatic. The null hypothesis of stability across samples is overwhelmingly rejected for the output and inflation equations.

This leads us to conclude that heteroskedasticity is present in the reduced-form equations over the two sub-samples. From the point of view of a structural DSGE model this is not surprising as we demonstrated in Section 2: it is precisely what would be expected in the presence of a monetary policy shift. What makes this observation potentially relevant for the empirical literature on the Lucas critique is that, so far, inference had been based on parameter stability tests that neglected this feature of the data. More specifically, the Chow test and the superexogeneity test, used by Favero and Hendry (1992), Lindé (2001), and Rudebusch (2005) among many others, implicitly assume homoskedasticity of the error variances between sample regimes and are therefore subject to our criticism.

Toyoda (1974) has demonstrated that the size and power of the Chow test can be considerably affected by neglecting differences in reduced-form error variances. The problem

[12] Estimates of the reduced-form coefficients are available from the authors upon request. We chose not to report these since the actual estimates are immaterial to our discussion. What matters is the joint significance, since it is well known that two sets of parameters can be jointly different from each other even though the parameters are not statistically different individually.

is most serious when samples of similar size are used. When overall sample size is small even moderate degrees of heteroskedasticity reduce the power of the test significantly. Both scenarios apply to the historical episode of the monetary policy shift in the U.S. Perhaps not surprisingly, the conclusions drawn in the literature almost always go against the statistical importance of Lucas' critique.[13] We consequently ask the question whether the failure to detect any effects is due to a lack of power of these tests.

5.2 Parameter Stability Tests

To quantify the importance of heteroskedasticity for the results of conventional parameter stability tests, we follow the literature and use the Chow and superexogeneity tests on the two simulated samples. The superexogeneity test measures the probability of rejecting the null of parameter stability in the output and inflation equations conditional on having rejected the null hypothesis of parameter stability in the interest rate equation.[14]

We present results for three versions of the parameter stability and superexogeneity tests. The first version is based on the OLS residual sum of squares, RSS, of (8) and (9), and therefore implicitly (and incorrectly) assumes homoskedasticity. A prominent example of this class of tests is the Chow statistics used by Lindé (2001) and Rudebusch (2005):

$$Chow = \frac{\left[RSS_{full} - (RSS_1 + RSS_2)\right]/m}{(RSS_1 + RSS_2)\,/\,(T-k)}, \tag{11}$$

where m is the number of restrictions in an equation of k parameters and $i = 1, 2$ indexes the sub-samples.

The second and third versions of the tests, in contrast, account for possible heteroskedasticity across sub-samples. In particular, we use a GLS version based on a two-step procedure. In the first step, Eqs. (8) and (9) are estimated by OLS over the two sub-samples. In the second step, the variables are normalized by the square root of the estimated innovation variance-covariance matrix. Statistics of interest are then computed using the residuals of the OLS estimates of (8) and (9) whereby the original variables are replaced with the transformed variables.

The third version is based on the difference of the parameter estimates, $\hat{\theta}_i$. The test statistic is:

$$Wald = \left(\hat{\theta}_1 - \hat{\theta}_2\right)' \left(\hat{V}_1 + \hat{V}_2\right)^{-1} \left(\hat{\theta}_1 - \hat{\theta}_2\right), \tag{12}$$

[13] An exception is Lindé (2001) who argues in favor of its empirical relevance on the basis of an empirical money demand equation, showing that the superexogeneity tests suffers from a serious small-sample problem.

[14] See Favero and Hendry (1992) for the strategy behind this test.

where $\hat{V}_i$ is the estimated variance-covariance matrix of the parameters corrected for heteroskedasticity and autocorrelation in the error terms. Following Newey and West (1987), we set the truncation lag in the autocorrelation function of the residuals, q_i, to $floor(4 * (T_i/100)\hat{}(2/9))$, which corresponds to 3 in our sub-samples.[15]

Parameter stability tests based on the RSS such as (11) and on the Wald form such as (12) give the same algebraic results only in one special case: when the restrictions on the parameters are linear and the error covariance matrix is homoskedastic (see Hamilton, 1994, Ch. 8, and Hansen, 2001). The Newey-West correction for heteroskedasticity and autocorrelation, in fact, has no influence on the RSS, implying that only an expression such as (12) is suited for dealing with heteroskedasticity.

Table 5 reports the results of the three versions of the parameter stability and superexogeneity tests of 5% empirical size. The size-corrected power of the respective tests, i.e., the probability of rejecting the null hypothesis when it is false, is computed using the empirical 5% significance level and the simulation strategy described in Section 4.2. According to the standard Chow test (labeled 'OLS-based') there is little evidence of parameter instability in the output and inflation equations indicated by probabilities that are never larger than 0.10. Yet, the Chow test is capable of detecting instability in the interest rate equation.

The power of the test to detect parameter instability increases noticeably for our second and third versions which take the instability of the reduced-form innovations variance into account. When a GLS-based correction is used, the probability of (correctly) rejecting the null hypothesis increases by a factor of five for both the output and the inflation equations. While the power of the test for the inflation specification is good, it is still less than satisfactory in the case of output. This need not be surprising per se for two reasons. First, there may still be a small-sample issue. We investigate this possibility further below. Second, it may very well be the case that the effect of the policy parameter change on the behavior of the output specification is statistically small. In other words, the policy intervention might be modest in the Leeper-Zha sense. Additionally, the estimation of a similar model in Lubik and Schorfheide (2004) reveals that output over the sample period (and conditional on the structural model) is almost exclusively driven by technology shocks, and that the feedback from the policy equation is minor. Similar conclusions can be drawn in the case of the Newey-West correction.

The simulation evidence on the performance of the superexogeneity tests corroborates

[15] We obtain similar results by setting $q_i = 0$. Moreover, the residuals display no sign of serial correlation, suggesting that four lags of the dependent variable in the inflation and output equations provide a reasonably good approximation of the dynamics of the model.

the notion of a bias induced by neglecting heteroskedasticity. Incidentally, the probabilities of rejecting the null of superexogeneity reported in Panel B of Table 5 are virtually identical. We conclude that the test for parameter stability using either a GLS-based or a Newey-West correction for the presence of heteroskedasticity and serial correlation always have higher power than the tests based on OLS.

5.3 Sensitivity Analysis

We assess the robustness of our assessment along three dimensions. We first modify our baseline calibration with respect to a few key parameters, while maintaining the regime shift from indeterminacy to determinacy. Secondly, we keep the baseline specification for the structural parameters, but simulate a policy shift between determinate regimes. Finally, we address the small-sample problems emphasized by Lindé (2001).

The results for variations of our baseline calibration are reported in Table 6. We consider three cases. The simulation using a smaller Phillips-curve slope coefficient $\kappa = 0.2$ reinforces the neglected heteroskedasticity bias as the power of the tests on the stability of the output and inflation equations become 0.18 and 0.60, respectively. Similar conclusions are reached on the basis of the superexogeneity tests in the last two columns or using even smaller values of κ. When aggregate demand is less sensitive to interest rate movements, $\tau^{-1} = 3$, the GLS-based probabilities of rejecting the null for the output and inflation equations are three times as large as those neglecting the instability of the reduced-form innovations variance in the column OLS-based.

Lastly, to appreciate fully the effect of a shift from indeterminacy to determinacy, the bottom panel reports the results based on the value of 0.2 for σ_ζ, the standard deviation of sunspot shocks estimated by Lubik and Schorfheide (2004). This modification does not vary the baseline results much and confirms the conclusion reached by Castelnuovo and Surico (2006) that the passive monetary policy regime influenced U.S. aggregate fluctuations through a change in the transmission mechanism, rather than through sunspot shocks.

The second robustness check analyzes to what extent our results are driven by the change in the equilibrium properties of the model. As demonstrated in Section 2, the policy-induced shift from indeterminacy to determinacy reduces both persistence and volatility of the endogenous variables. Policy shifts that maintain a determinate equilibrium, on the other hand, have a less dramatic effect on the reduced-form error variance.[16] Our experiment is to

[16] Rudebusch (2005) rules out the possibility of a shift from indeterminacy to determinacy by adjusting the estimated inflation response coefficient upwards. This is done out of concern for avoiding instability in the equation system. However, the issue with indeterminacy is that there is not enough instability in the

change the inflation coefficient from 1.5 to 2.5. The other structural and policy parameters are as in Table 1. Two observations emerge: first, the policy shift within the determinacy region is less successful in replicating the stylized facts, in particular the decline in output volatility and persistence. While obviously far from conclusive, we regard this as a simple plausibility check on the driving forces behind the Great Moderation. Secondly, tests for error variance constancy, reported in Table 8, show that the null of homoskedasticity is not rejected at typical confidence levels with the exception of inflation. Differences between variances are in any case noticeably smaller than in the baseline case. Not surprisingly, this affects the improvement in power of the heteroskedasticity-adjusted test statistics in Table 9. The size-adjusted probabilities of rejecting parameter stability and superexogeneity are now more similar.

Small-sample problems as documented by Lindé (2001) in the case of the superexogeneity test are also a possible explanation. We return to our baseline calibration and simulate a shift from indeterminacy to determinacy, but this time impose a sub-sample size of 150 observations, twice as many as in the baseline case. The results can be found in Table 10. Small-sample problems are clearly present for all three versions of the test for parameter stability. Interestingly, a doubling of sample size has little effect on the probabilities of rejecting the null hypothesis for the output equation, while the power of heteroskedasticity-adjusted statistics improves by more than that of the simple OLS-based statistics. A similar conclusion can be drawn for the superexogeneity test.

6 Detecting the Lucas Critique in Practice: The Volcker Policy Shift

We conclude our analysis by an application to historical data. In fact, many contributions, for instance Stock and Watson (2002), Kim, Nelson and Piger (2004), and Cogley and Sargent (2005), have shown that post-war U.S. data are characterized by a substantial amount of heteroskedasticity. The conceptual background for our simulation analysis is the break in U.S. monetary policy behavior in the early 1980s, characterized by the tenure of Paul Volcker. We are interested in whether parameter stability tests adjusted for heteroskedasticity detect this shift in the context of a reduced-form model.

The full sample spans the period 1960:1 - 2005:3. Quarterly data were collected from the FRED database at the Federal Reserve Bank of St. Louis. We use data on output growth, measured as the quarter-to-quarter change in real GDP; inflation as the quarter-to-quarter

system since there are infinitely many *stable* adjustment paths.

change in the GDP deflator; and the federal funds rate as policy instrument.

Figure 1 depicts the statistics of two recursive tests for parameter stability of the reduced-form model (8)-(10). The minimum length of a sub-sample is eight years, implying that in the first recursion, Period I ends in 1968:4, whereas in the last recursion, Period II begins in 1998:1. The left column reports the results of the Chow test (11) which is based on the RSS and therefore implicitly assumes spherical disturbances. The right column refers to the Wald test (12) which, in contrast, is corrected for heteroskedasticity and autocorrelation in the error terms.

The top dashed horizontal line in each panel represents the 5% empirical critical value based on $10,000$ bootstrap repetitions in which we impose the null of parameter stability. The dashed horizontal line in the middle stands for the 5% asymptotic critical value computed by Andrews (1993) for tests of parameter instability with unknown change point within the middle 70% of the sample. For sake of comparison, the critical values of the same tests, but with known break date, are reported as dotted horizontal lines.[17] The null hypothesis of parameter stability is rejected if the maximum value of the statistics is above the critical value.

In analogy with the Monte Carlo simulations in Section 5, using the Wald test dramatically changes the inference on parameter stability. The *sup* of the Wald statistics in the right column are above the bootstrapped 5% critical values for all equations. In contrast, the recursive Chow tests in the left column uniformly fail to reject the null hypothesis of parameter stability. The only possible exception is the interest rate equation conditional on a priori knowledge that a break occurred in 1980:2. However, even if the econometrician were equipped with information, say, on the basis of the empirical literature on monetary policy rules (Clarida et al., 2000), the statistics of the Chow tests would not detect any structural change in the output and inflation equations, being far below the relevant 5% critical values.

While the focus of the paper is on the policy shift associated with the beginning of Volcker's tenure, we cannot exclude, in principle, that the backward-looking model (8)-(10) is in fact characterized by multiple structural breaks. To investigate this possibility, we use the tests for (pure) multiple changes at unknown dates proposed by Bai and Perron (1998). Following the practical recommendations in Bai and Perron (2003), we set the number of maximum structural breaks, M, to 4, and consider a trimming ε of 0.15. In analogy with

[17] Andrews and Fair (1988) show that under the null hypothesis of parameter stability, $\hat{\theta}_1 = \hat{\theta}_2$, the Wald statistic is asymptotically distributed as a χ^2 random variable with k degrees of freedom.

the testing procedure on simulated data, different variances of the residuals are allowed across segments. P-values are based on $10,000$ bootstrap repetitions obtained under the null of parameter stability. The significance level is 5%.

The first two columns of Table (11) report the $UDmax$ and $WDmax$ double maximum statistics for a test of no structural break against the alternative of an unknown number of breaks given the upper bound M. According to both statistics, there is at least one structural change in all three equations. To determine the number of breaks we then examine the Sup-F$(\ell + 1|\ell)$ statistics for testing the null of ℓ break(s) against the alternative that an additional break exists. The third column of Table (11) shows that the hypothesis that only one change occurred is never rejected against the alternative of at least two changes. The estimates of the break date obtained by minimizing the sum of the RSS over the sub-samples are 1983:2 for output, 1981:2 for inflation and 1980:3 for the policy rule.

7 Conclusions

We present two main arguments in this paper. First, we emphasize that within the framework of structural DSGE models, a change in policy parameters affects the stability of both reduced-form coefficients and reduced-form error variances. We show that tests for parameter instability based on the assumption of homoskedasticity have low power. We consequently argue that this is behind the tendency in the literature to reject the empirical relevance of the Lucas critique. We suggest adjustments for heteroskedasticity in commonly used parameter stability tests, and show by means of a Monte Carlo analysis that the power of these tests is improved and Lucas critique effects are detected in simulated data.

As an empirical example we test for the presence of Lucas critique effects on reduced-form specifications of output and inflation using post-war U.S. data. We do, indeed, find evidence of a break in the behavior of the Federal Reserve in 1980, and of parameter instability across the sub-samples using our heteroskedasticity-adjusted test statistic. We believe this is a new finding in the literature. Obviously, our results are model-dependent, and further investigation in a richer framework is certainly warranted.

We want to conclude on a somewhat critical note. Given a fully-specified, structural DSGE model as a DGP, the question arises why researchers should bother at all with reduced-form specifications that are subject to the Lucas critique. One answer is certainly ease of implementation. Furthermore, very little is known about structural break tests within the context of estimated DSGE-models.[18] However, a deeper issue is whether DSGE-

[18] A notable exception is Ireland (2001). Initial steps in this direction have also been made by Canova

models that are used for policy analysis are not themselves subject to the Lucas critique. Implicitly, Lucas' argument rests on the notion that the information set of economic agents and their decision problems were not fully specified in traditional macroeconometric models. Yet, with the use of *ad hoc* monetary policy rules that very issue surely comes up in DSGE models that do not include optimizing policy-makers.

(2005) and Justiniano and Primiceri (2006).

References

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

[2] Andrews, Donald.W.K. and Ray C. Fair (1988): "Inference in Nonlinear Econometric Models with Structural Change". *The Review of Economic Studies*, 55, 615-639.

[3] Bai, Jushan and Pierre Perron (1998): "Estimating and Testing Linear Models with Multiple Structural Changes". *Econometrica* 66, 47-78.

[4] Bai, Jushan and Pierre Perron (2003): "Computation and Analysis of Multiple Structural Change Models". *Journal of Applied Econometrics* 18, 1-22.

[5] Blanchard, Olivier J. and John A. Simon (2001): "The Long and Large Decline in U.S. Output Volatility". *Brookings Papers on Economic Activity*, 1, 135 – 164.

[6] Boivin, Jean, and Marc Giannoni (2005): "Has Monetary Policy Become More Effective?" Forthcoming, *The Review of Economics and Statistics*.

[7] Canova, Fabio (2005): "Policy and the Evolution of the US Economy: 1948-2002". *Universitat Pompeu Fabra Working Paper* No. 919.

[8] Canova, Fabio, Luca Gambetti, and Evi Pappa (2005): "The Structural Dynamics of US Output and Inflation: What Explains the Changes?" *Universitat Pompeu Fabra Working Paper* No. 921.

[9] Castelnuovo, Efrem, and Paolo Surico (2006): "The Price Puzzle: Fact or Artifact?". *Bank of England Working Paper* No. 288.

[10] Clarida, Richard, Jordi Galí, and Mark Gertler (1999): "The Science of Monetary Policy: A New Keynesian Perspective". *The Journal of Economic Literature,* XXVII, 1661-1707.

[11] Clarida, Richard, Jordi Galí, and Mark Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory". *The Quarterly Journal of Economics*, 115(1), 147-180.

[12] Cogley, Timothy W. and Thomas Sargent (2005): "Drifts and Volatilities: Monetary Policies and Outcomes in the Post World War II U.S.". *Review of Economic Dynamics*, 8(2), 262-302.

[13] Collard, Fabrice, Patrick Fève, and François Langot (2002): "Structural Inference and the Lucas Critique". *Annales d'Economie et de Statistiques*, 67/68, 183-2002.

[14] Estrella, Arturo and Jeffrey C. Fuhrer (2003): "Monetary Policy Shifts and the Stability of Monetary Policy Models". *The Review of Economics and Statistics,* 85(1), 94-104.

[15] Favero, Carlo and David Hendry (1992): "Testing the Lucas Critique: A Review". *Econometric Reviews,* 11(3), 265-306.

[16] Hamilton, James D. (1994): *Time Series Analysis*. Princeton University Press. Princeton, NJ.

[17] Hansen, Bruce, E. (2001): "The New Econometrics of Structural Change: Dating Breaks in U.S. Labor Productivity". *Journal of Economic Perspective*, 15, 117-128.

[18] Ireland, Peter N. (2001): "Sticky-Price Models of the Business Cycle: Specification and Stability". *Journal of Monetary Economics*, 47, 3-18.

[19] Justiniano, Alejandro, and Giorgio Primiceri (2006): "The Time Varying Volatility of Macroeconomic Fluctuations". Manuscript, Northwestern University

[20] Kim, Chang-Jin, and Charles R. Nelson (1999): "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle". *The Review of Economics and Statistics*, 81, 608 – 616.

[21] Kim, Chang-Jin, Charles R. Nelson, and Jeremy Piger (2004): "The Less-Volatile U.S. Economy: A Bayesian Investigation of Timing, Breadth, and Potential Explanations". *Journal of Economic and Business Statistics*, 22, 80-93.

[22] King, Robert G. (2000): "The New IS-LM Model: Language, Logic, and Limits". *Federal Reserve Bank of Richmond Economic Quarterly*, 86(3), 45-103.

[23] Leeper, Eric M. and Tao Zha (2003): "Modest Policy Interventions". *Journal of Monetary Economics*, 50, 1673-1700.

[24] Lindé, Jesper (2001): "Testing for the Lucas Critique: A Quantitative Investigation". *American Economic Review*, 91(4), 986-1005.

[25] Lubik, Thomas A., and Massimiliano Marzo (2006): "An Inventory of Simple Monetary Policy Rules in a New Keynesian Macroeconomic Model". Forthcoming, *International Review of Economics & Finance.*

[26] Lubik, Thomas A., and Frank Schorfheide (2003): "Computing Sunspot Equilibria in Linear Rational Expectations Models". *Journal of Economic Dynamics and Control*, 28(2), 273-285.

[27] Lubik, Thomas A., and Frank Schorfheide (2004): "Testing for Indeterminacy: An Application to U.S. Monetary Policy". *American Economic Review*, 94(1), 190-217.

[28] Lubik, Thomas A., and Frank Schorfheide (2005): "A Bayesian Look at New Open Economy Macroeconomics". *NBER Macroeconomics Annual*, pp. 313-366.

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

[30] McConnell, Margaret M., and Gabriel Perez-Quiros (2000): "Output Fluctuations in the United States: What has Changed Since the Early 1980's ?" *American Economic Review*, 90(5), 1464 – 1476.

[31] Newey, Whitney K. and Kenneth West (1987): "A Simple, Positive Semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix,". *Econometrica*, 55, 703-708.

[32] Rudebusch, Glenn D. (2005): "Assessing the Lucas Critique in Monetary Policy Models". *Journal of Money, Credit and Banking*, 37(2), 245-272.

[33] Rudebusch, Glenn D., and Lars E.O. Svensson (1999): "Policy Rules for Inflation Targeting". In: *Monetary Policy Rules*, ed. by John B. Taylor. Chicago: University of Chicago Press, 203-246.

[34] Sims, Christopher A. (2002): "Solving Linear Rational Expectations Models". *Computational Economics*, 20 (1-2), 1-20.

[35] Stock, James, and Mark Watson (2002): "Has the Business Cycle Changed and Why?" *NBER Macroeconomics Annual*.

[36] Svensson, Lars E.O. (1997): "Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets". *European Economic Review*, 41, 1111-1146.

[37] Toyoda, Toshihisa (1974): "Use of the Chow Test Under Heteroskedasticity". *Econometrica*, 42, 601-608.

[38] Woodford, Michael (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press. Princeton, NJ.

Table 1: Model Parameters

Panel A: Structural Parameters

Subsample	β	κ	τ^{-1}	ρ_g	ρ_z	σ_g	σ_z
I and II	0.99	0.7	1.4	0.7	0.8	0.3	1.1

Panel B: Monetary Policy Rule

Subsample	ψ_π	ψ_y	ρ_R	σ_R
I: Indeterminacy	0.4	0.2	0.6	0.3
II: Determinacy	1.5	0.5	0.8	0.3

Table 2: Good Policy Hypothesis: Descriptive Statistics

Panel A: Standard Deviations

Equation	I: Indeterminacy	II: Determinacy	% Change
Output	2.52	1.47	-41.58
Inflation	2.17	0.72	-66.91
Interest Rate	0.81	0.47	-41.85

Panel B: Persistence

Equation	I: Indeterminacy	II: Determinacy	% Change
Output	0.69	0.75	8.96
Inflation	0.69	0.29	-58.07
Interest Rate	0.85	0.77	- 9.33

Table 3: Good Luck Hypothesis: Descriptive Statistics

Panel A: Standard Deviations

Equation	I: High Variance	II: Low Variance	% Change
Output	1.50	0.80	-46.57
Inflation	0.72	0.26	-64.65
Interest Rate	0.48	0.21	-57.65

Panel B: Persistence

Equation	I: High Variance	II: Low Variance	% Change
Output	0.78	0.83	6.50
Inflation	0.30	0.33	9.27
Interest Rate	0.76	0.80	5.08

Table 4: Innovation Standard Errors: Descriptive Statistics and Stability Tests

Equation	I: Indeterminacy	II: Determinacy	Statistics (p-value)
Output	1.79	0.88	4.11 (0.00)
Inflation	1.55	0.65	5.64 (0.00)
Interest Rate	0.25	0.23	1.22 (0.23)

Table 5: Parameter Stability Tests

Panel A: Size-Adjusted Probability of Rejecting Parameter Stability

Equation	OLS-based	GLS-based	Newey-West corrected
Output	0.03	0.17	0.11
Inflation	0.09	0.45	0.38
Interest Rate	0.91	0.94	0.92

Panel B: Size-Adjusted Probability of Rejecting Super-Exogeneity

Equation	OLS-based	GLS-based	Newey-West corrected
Output	0.02	0.11	0.11
Inflation	0.09	0.46	0.37

Table 6: Parameter Stability Tests
-Sensitivity Analysis-

	Parameter Stability		Superexogeneity	
	OLS-based	GLS-based	OLS-based	GLS-based
$\kappa = 0.2$				
Output	0.03	0.18	0.03	0.13
Inflation	0.08	0.60	0.22	0.61
Interest Rate	0.30	0.44	-	-
$\tau^{-1} = 3$				
Output	0.05	0.20	0.04	0.14
Inflation	0.18	0.53	0.16	0.54
Interest Rate	0.97	0.99	-	-
$\sigma_\zeta = 0.2$				
Output	0.03	0.16	0.03	0.11
Inflation	0.08	0.45	0.08	0.45
Interest Rate	0.84	0.89	-	-

Table 7: Good Policy Hypothesis: Descriptive Statistics
- from Determinacy to Determinacy -

Panel A: Standard Deviations

Equation	I: Determinacy	II: Determinacy	% Change
Output	1.49	1.50	0.35
Inflation	0.69	0.52	-24.57
Interest Rate	0.73	0.50	-31.75

Panel B: Persistence

Equation	I: Determinacy	II: Determinacy	% Change
Output	0.77	0.75	1.98
Inflation	0.47	0.25	-45.66
Interest Rate	0.78	0.77	- 1.38

Table 8: Innovation Standard Errors: Descriptive Statistics and Stability Tests
- from Determinacy to Determinacy -

Equation	I: Determinacy	II: Determinacy	Statistics (p-value)
Output	0.89	0.90	0.97 (0.55)
Inflation	0.59	0.48	1.55 (0.05)
Interest Rate	0.26	0.23	1.28 (0.18)

Table 9: Parameter Stability Tests
- from Determinacy to Determinacy -

Panel A: Size-Adjusted Probability of Rejecting Parameter Stability

Equation	OLS-based	GLS-based	Newey-West corrected
Output	0.06	0.16	0.10
Inflation	0.10	0.23	0.20
Interest Rate	0.48	0.61	0.61

Panel B: Size-Adjusted Probability of Rejecting Super-Exogeneity

Equation	OLS-based	GLS-based	Newey-West corrected
Output	0.06	0.11	0.11
Inflation	0.21	0.23	0.20

Table 10: Parameter Stability Tests
-from Indeterminacy to Determinacy: sub-samples of 150 obs-

Panel A: Size-Adjusted Probability of Rejecting Parameter Stability

Equation	OLS-based	GLS-based	Newey-West corrected
Output	0.03	0.12	0.12
Inflation	0.19	0.78	0.75
Interest Rate	0.99	0.99	0.99

Panel B: Size-Adjusted Probability of Rejecting Super-Exogeneity

Equation	OLS-based	GLS-based	Newey-West corrected
Output	0.03	0.11	0.09
Inflation	0.15	0.76	0.71

Table 11: Tests for Multiple Structural Changes at Unknown Dates

Equation	UDmax (p-value)[a]	WDmax (p-value)[a]	SupF(2|1) (p-value)[a]	Break dates
Output	31.35 (0.010)	31.35 (0.015)	16.57 (0.117)	1983Q2
Inflation	35.35 (0.010)	35.35 (0.015)	11.35 (0.588)	1981Q2
Interest Rate	32.12 (0.035)	32.12 (0.063)	11.63 (0.443)	1980Q3

[a] bootstrapped p-values, trimming = 0.15, maximum number of breaks = 4.

Figure 1: Parameter Instability Tests at Unknown Break Date