RESEARCH ARTICLE

Revisiting the Great Moderation: Policy or Luck?

Patrick Minford · Zhirong Ou · Michael Wickens

Abstract We investigate the relative roles of monetary policy and shocks in causing the Great Moderation, using indirect inference where a DSGE model is tested for its ability to mimic a VAR describing the data. A New Keynesian model with a Taylor Rule and one with the Optimal Timeless Rule are both tested. The latter easily dominates, whether calibrated or estimated, implying that the Fed's policy in the 1970s was neither inadequate nor a cause of indeterminacy; it was both optimal and essentially unchanged during the 1980s. By implication it was largely the reduced shocks that caused the Great Moderation—among them monetary policy shocks the Fed injected into inflation.

Keywords Great moderation · Causes · Indirect inference · Test · Wald statistics

JEL Classification E42 · E52 · E58

P. Minford
Cardiff Business School, Cardiff University, E26, Aberconway building, Colum Drive, Cardiff, UK CF10 3EU
e-mail: MinfordP@cardiff.ac.uk

P. Minford · M. Wickens
CEPR, London, UK

M. Wickens
e-mail: WickensM1@cardiff.ac.uk

Z. Ou (✉)
Cardiff Business School, Cardiff University, B14, Aberconway Building, Colum Drive, Cardiff, UK CF10 3EU
e-mail: ouz@cardiff.ac.uk

M. Wickens
Cardiff Business School, Cardiff University, S36, Aberconway Building, Colum Drive, Cardiff CF10 3EU, UK

M. Wickens
University of York, York, UK

1 Introduction

John Taylor suggested in Taylor (1993) that an interest rate rule for well-conducted monetary policy fitted the Fed's behaviour since 1987 rather well in a single equation regression. Since then a variety of similar studies have confirmed his finding—most of these have focused on a data sample beginning in the early-to-mid 1980s. For the period from the late 1960s to the early 1980s the results have been more mixed. Thus Clarida et al. (2000) reported that the Taylor Rule fitted but with a coefficient on inflation of less than unity; in a full New Keynesian model this fails under the usual criteria to create determinacy in inflation and they argue that this could be the reason for high inflation and output volatility in this earlier post-war period. They concluded that the reduction in macro volatility between these two periods (the 'Great Moderation') was due to the improvement in monetary policy as captured by this change in the operative Taylor Rule.

This view of the Great Moderation has been widely challenged in econometric studies of the time series. These have attempted to decompose the reduction in macro variance into the effect of parameter changes and the effect of shock variances. Virtually all have found that the shock variances have dominated the change and that the monetary policy rule operating therefore did not change very much.

A further questioning of the Taylor Rule account of the post-war monetary policy has come from Cochrane (2011) and others (Minford et al. 2002) who argue that the Taylor Rule is not identified as a single equation because a DSGE model with a different monetary policy rule (such as a money supply rule) could equally well generate an equation of the Taylor Rule form. Therefore much of the work that estimates the Taylor Rule could be spurious. However, identification can be achieved when the Taylor Rule is embedded in a full DSGE model because of the over-identifying restrictions implied by rational expectations. There then remains the question of whether such a model is as good a representation as one that is in general the same but has an alternative monetary policy rule[1].

This points the way to a possible way forward for testing models of monetary policy. One may specify a complete DSGE model with alternative monetary rules and use the over-identifying restrictions to determine their differing behaviours. One may then test which of these is more acceptable from the data's viewpoint and hence comes closest to the true model. This is the approach taken here.

We look at a particular rival to the Taylor Rule, the Optimal Timeless Rule. This is of interest because in it the Fed is playing a more precisely optimising role than it does in the Taylor Rule which is a simple rule that can be operated with limited current information, namely for output and inflation. The Optimal Timeless Rule assumes that

---

[1] Rules of the Taylor type are generally found to fit the data well, either as a stand-alone equation in regression analysis, or as part of a full model in DSGE analysis. Giannoni and Woodford (2005) is a recent example of the former, whereas Smets and Wouters (2007) and Ireland (2007) are examples of the latter. However, besides the usual difficulties encountered in applied work (e.g., Castelnuovo 2003 and Carare and Tchaidze 2005), these estimates face an identification problem pointed out by Minford et al. (2002) and Cochrane (2011).

Lack of identification occurs when an equation could be confused with a linear combination of other equations in the model. In the case of the Taylor Rule, DSGE models give rise to the same correlations between interest rate and inflation as the Taylor Rule, even if the Fed is doing something quite different, such as targeting the money supply. For example, Minford (2008) shows this in a DSGE model with Fischer wage contracts. For details see Minford and Ou (2013).

the Fed can solve the DSGE model for all the shocks and so choose in a discriminating way its reaction to each shock. Other than this Optimal Timeless Rule we also look at variants of the Taylor Rule, including one that closely mimics the Optimal Timeless Rule.

To make our testing bounded and tractable we use the monetary rule in conjunction with the most widely-accepted parsimonious DSGE model representation—where the model is reduced to two equations, a forward-looking 'IS' curve and a New Keynesian Phillips curve, plus the monetary rule. We allow each rule/model combination to be calibrated with the best chance of matching the data and then test on that best calibration, using the method of Indirect Inference under which the model's simulated behaviour is formally tested for congruence with the behaviour of the data. Our efforts here join others that have brought DSGE models to bear on this issue—notably, Ireland (2007); Smets and Wouters (2007) and the related Le et al. (2011) and Fernandez-Villaverde et al. (2009, 2010). These authors have all used much larger DSGE models, in some cases data that was non-stationary, and in most cases Bayesian estimation methods. Their work is largely complementary to ours and we discuss their findings in relation to ours below. Bayesian estimation is a method for improving on calibrated parameters but our method of indirect inference takes matters further and asks if the finally estimated model is consistent overall with the data behaviour; if not it searches for some set permissible within the theory that is consistent, getting as close to consistency as possible given the model and the data. This method is the major innovation we introduce for the treatment of the topic here; it is a method based on classical statistical inference which we explain and defend carefully below.

In using indirect inference to test the models we deviate from the popular use of Bayesian methods to evaluate models. This is because Bayesian evaluation of a model (by likelihood and odds ratio tests) does not test the model as a whole against the data; indeed Bayesians dismiss the idea of 'testing models'. Bayesian estimates depend on the choice of prior distributions which are designed to restrict the estimates. By testing a model as a whole, we mean testing jointly both the effect of imposing the priors and the usual structural restrictions on the model. In contrast, Bayesian evaluation examines which of two specifications of a structural model is more probable given the priors assumed; any ranking of the two models is thus dependent on the priors chosen[2] . Thus Bayesian methods cannot be used to test models against the data in the sense in which we wish to test (and rank) a model.

It may, of course, still be argued that it is wrong to test models as a whole against the data and that one should only check improvements conditional on prior assumptions that should not be challenged. It is, however, hard to argue that any set of prior assumptions can be taken for granted as true and beyond challenge as can be seen from the number of 'schools of thought' still in existence in macroeconomics. This wide divergence of beliefs has, if anything, been exacerbated by the financial crisis of the late 2000s. Whether one likes it or not, as a macroeconomist, one must recognise that to establish a model scientifically to the satisfaction of other economists and

---

[2] With the use of a flat prior Bayesian ranking is equivalent to assessing the likelihood ratio of two models. This ranking is similar to the ranking we obtain under indirect inference except that the criterion is the likelihood of the data rather than the likelihood of the data representation. However we must emphasise that our first concern is testing against the data for rejection; only for models that are not rejected are we concerned with rank.

policymakers, it needs to be shown that a model being proposed for a policy use is consistent with the data in a manner that enables it to be used for that purpose. We show below that indirect inference fulfils that need.

Direct inference, as in the Likelihood Ratio test, is an alternative to indirect inference for testing models against the data. However, as we elaborate below, Likelihood Ratio tests appear to have considerably less power in small samples than the indirect inference test we use here. By implication, indirect inference will therefore provide a more powerful discrimination between the models.

In Section 2 we review the work on the Great Moderation. In Section 3 we set out the model and the rules to be tested, and in Section 4 our test procedure, together with a discussion of the alternatives. Section 5 shows the results, while Section 6 draws out the implications for the Great Moderation and Section 7 concludes.

2 Causes of the Great Moderation

The Great Moderation refers to the period during which the volatility of the main economic variables was relatively modest. This began in the US around the early 1980s although there is no consensus on the exact date. Figure 1 below shows the time paths of three main US macro variables from 1972 to 2007: the nominal Fed interest rate, output gap and CPI inflation. It shows the massive fluctuation of the 1970s ceased after the early 1980s, indicating the economy's transition from the Great Acceleration to the Great Moderation.

Changes in the monetary policy regime could have produced the Great Moderation. This is typically illustrated with the three-equation New Keynesian framework, consisting of the IS curve derived from the household's optimization problem, the Phillips curve derived from the firm's optimal price-setting behaviour, and a Taylor Rule approximating the Fed's monetary policy. Using simulated behaviour from models of this sort, a number of authors suggest that the US economy's improved stability was largely due to stronger monetary policy responses to inflation (Clarida et al. 2000; Lubik and Schorfheide 2004; Boivin and Giannoni 2006 and Benati and Surico 2009). The contrast is between the 'passive' monetary policy of the 1970s, with low Taylor Rule responses, and the 'active' policy of the later period in which the conditions for a unique stable equilibrium (the 'Taylor Principle') are met, these normally being that the inflation response in the Taylor Rule be greater than unity. Thus it was argued that the indeterminacy caused by the passive 1970s policy generated
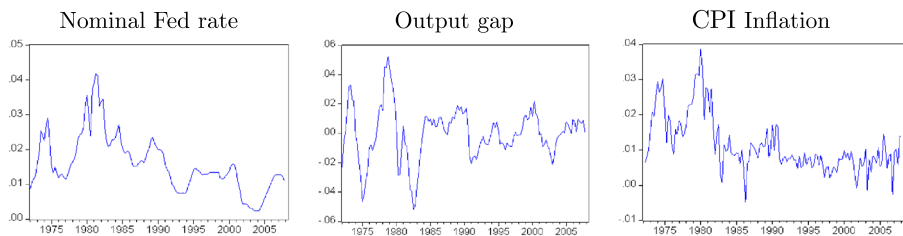
Data source: the Federal Reserve Bank of St. Louis (http://research.stlouisfed.org/fred2/, accessed Nov. 2009).

Fed rate and inflation unfiltered; the output gap is the log deviation of real GDP from its HP trend.

Fig. 1 Time paths of main macro variables of the US economy

sunspots and so the Great Acceleration; with the Fed's switch this was eliminated, hence the Great Moderation.

By contrast other authors, mainly using structural VAR analysis, have suggested that the Great Moderation was caused not by policy regime change but by a reduction in the variance of shocks. Thus Stock and Watson (2002) claimed that over 70 % of the reduction in GDP volatility was due to lower shocks to productivity, commodity prices and forecast errors. Primiceri (2005) argued that the stagflation in the 1970s was mostly due to non-policy shocks. A similar conclusion was drawn by Gambetti et al. (2008), while Sims and Zha (2006) found in much the same vein that an empirical model with variation only in the variance of the structural errors fitted the data best and that alteration in the monetary regime—even if assumed to occur—would not much influence the observed inflation dynamics.

The logic underlying the structural VAR approach is that, when actual data are modelled with a structural VAR, their dynamics will be determined both by the VAR coefficient matrix that represents the propagation mechanism (including the monetary regime) and by the variance-covariance matrix of prediction errors which takes into account the impact of exogenous disturbances. Hence by analysing the variation of these two matrices across different subsamples it is possible to work out whether it is the change in the propagation mechanism or in the error variability that has caused the change in the data variability. It is the second that these studies have identified as the dominant cause. Hence almost all structural VAR analyses have suggested 'good shocks' (or 'good luck') as the main cause of the Great Moderation, with the change of policy regime in a negligible role.

Nevertheless, since this structural VAR approach relies critically on supposed model restrictions to decompose the variations in the VAR between its coefficient matrix and the variance-covariance matrix of its prediction errors, there is a pervasive identification problem. As Benati and Surico (2009) have pointed out, the problem that 'lies at the very heart' is the difficulty in connecting the structure of a DSGE model to the structure of a VAR. In other words one cannot retrieve from the parameters of an SVAR the underlying structural parameters of the DSGE model generating it, unless one is willing to specify the DSGE model in detail. None of these authors have done this. Hence one cannot know from their studies whether in fact the DSGE model that produced the SVAR for the Great Acceleration period differed or did not differ from the DSGE model producing the SVAR for the Great Moderation period. It is not enough to say that the SVAR parameters 'changed little' since we do not know what changes would have been produced by the relevant changes in the structural DSGE models. Different DSGE models with similar shock distributions could have produced these SVARs with similar coefficients and different shock distributions.

Essentially it is this problem that we attempt to solve in the work we present below. We estimate a VAR for each period and we then ask what candidate DSGE models could have generated each VAR. Having established which model comes closest to doing so, we then examine how the difference between them accounts for the Great Moderation. Since these models embrace the ones put forward by the authors who argue that policy regime change accounts for it, we are also able to evaluate these authors' claims statistically. Thus we bring evaluative statistics to bear on the authors who claim policy regime change, while we bring identification to bear on the authors who use SVARs.

3 A Simple New Keynesian Model for Interest Rate, Output Gap and Inflation Determination

We follow a common practice among New Keynesian authors of setting up a full DSGE model with representative agents and reducing it to a three-equation framework consisting of an IS curve, a Phillips curve and a monetary policy rule.

Under rational expectations the IS curve derived from the household's problem and the Phillips curve derived from the firm's problem under Calvo (1983) contracts can be shown to be:

$$x_t = E_t x_{t+1} - \left(\frac{1}{\sigma}\right)\left(\widetilde{\iota}_t - E_t \pi_{t+1}\right) + v_t \tag{1}$$

$$\pi_t = \chi E_t \pi_{t+1} + \gamma x_t + \kappa u_t^w \tag{2}$$

where x_t is the output gap, $\widetilde{\iota}_t$ is the deviation of interest rate from its steady-state value, π_t is the price inflation, and v_t and u_t^w are the 'demand shock' and 'supply shock', respectively.[3]

We consider three monetary regime versions widely suggested for the US economy. These are the Optimal Timeless Rule when the Fed commits itself to minimizing a typical quadratic social welfare loss function; the original Taylor Rule; and an interest-rate-smoothed Taylor Rule.

In particular, the Optimal Timeless Rule is derived following Woodford (1999)'s idea of ignoring the initial conditions confronting the Fed at the regime's inception. It implies that, if the Fed was a strict, consistent optimizer, it would have kept inflation always equal to a fixed fraction of the first difference of the output gap, ensuring

$$\pi_t = -\frac{\alpha}{\gamma}\left(x_t - x_{t-1}\right) \tag{3}$$

where α is the relative weight it puts on the loss from output variations against inflation variations and γ is the Phillips curve constraint (regarding stickiness) it faces.[4]

Unlike Taylor Rules that specify a systematic policy instrument response to economic changes, this Timeless Rule sets an optimal trade-off between economic outcomes—here, it punishes excess inflation with a fall in the output growth rate. It then chooses the policy instrument setting to achieve these outcomes; thus the policy response is implicit. Svensson and Woodford (2004) categorized such a rule as 'high-level monetary policy'; they argued that by connecting the central bank's monetary actions to its ultimate policy objectives this rule has the advantage of being more transparent and robust.[5]

---

[3] Our Supporting Annex shows the full derivation.

[4] See also Clarida et al. (1999) and McCallum and Nelson (2004). This is based on defining social welfare loss as 'the loss in units of consumption as a percentage of steady-state output' as in Rotemberg and Woodford (1998)—also Nistico (2007); it is conditional on assuming a particular utility function and zero-inflation steady state—more details can be found in our Annex.

[5] Svensson and Woodford (2004) also comment that such a rule may produce indeterminacy; however this does not occur in the model here.

Thus, in order to implement the Optimal Timeless Rule the Fed must fully understand the model (including the shocks hitting the economy) and set its policy instrument (here the Fed rate) to whatever supports the Rule. Nevertheless, the Fed may make errors of implementation that cause the rule not to be met exactly—'trembling hand' errors, ξ_t. Here, since (3) is a strict optimality condition, we think of such policy mistakes as due either to an imperfect understanding of the model or to an inability to identify and react to the demand and supply shocks correctly. This differs from the error in typical Taylor Rules, which consist of the Fed's discretionary departures from the rule.

Thus the three model economies with differing monetary policy settings are readily comparable. These are summarised in Table 1.[6]

Since these models differ only in the monetary policies being implemented, by comparing their capacity to fit the data one should be able to tell which rule, when included in a simple New Keynesian model, provides the best explanation of the facts and therefore the most appropriate description of the underlying policy. We go on to investigate this in what follows.

4 Model Estimation and Evaluation by Indirect Inference

Indirect inference has been widely used in the estimation of structural models (e.g., Smith 1993; Gregory and Smith 1991, 1993; Gourieroux et al. 1993; Gourieroux and Monfort 1996 and Canova 2005). Here we make a further use of indirect inference to evaluate an already estimated or calibrated (DSGE) macroeconomic model using classical statistical inference. This is related to, but is different from, estimating a macroeconomic model by indirect inference. The common feature is the use of an auxiliary model in addition to the structural macroeconomic model. For a full description of the method of indirect inference see also Le et al. (2011). In addition to testing a particular prior numerical specification of the DSGE model, we examine how we might compare and test alternative numerical specifications of the model. Next we set out the main features of indirect inference.

4.1 Estimation

Estimation by indirect inference chooses the parameters of the macroeconomic model so that when this model is simulated it generates estimates of the auxiliary model similar to those obtained from the observed data. The optimal choice of parameters for the macroeconomic model are those that minimize the distance between a given function of the two sets of estimated coefficients of the auxiliary model. Common choices of this function are (i) the actual coefficients, (ii) the scores, and (iii) the impulse response functions. In effect, estimation by indirect inference provides an optimal calibration.

---

[6] Note all equation errors are allowed to follow an AR(1) process when the models are tested against the data so omitted variables are allowed for. We also transform the Taylor Rules to quarterly versions so the frequency of interest rate and inflation is consistent with other variables in the model. All constant terms are dropped as demeaned, detrended data will be used- see section 'Data and Results' below.

Table 1 Competing rival models

Baseline framework	
IS curve	$x_t = E_t x_{t+1} - \left(\frac{1}{\sigma}\right)\left(\tilde{\imath}_t - E_t \pi_{t+1}\right) + v_t$
Phillips curve	$\pi_t = \beta E_t \pi_{t+1} + \gamma x_t + \kappa u_t^w$
Monetary policy versions	
Optimal Timeless Rule (model one)	$\pi_t = -\frac{\alpha}{\gamma}\left(x_t - x_{t-1}\right) + \xi_t$
Original Taylor Rule (model two) & transformed equation	$i_t^A = \pi_t^A + 0.5x_t + 0.5\left(\pi_t^A - 0.02\right) + 0.02 + \xi_t \tilde{\imath}_t = 1.5\pi_t + 0.125x_t + \xi_t'$
'IRS' Taylor Rule (model three) & transformed equation	$i_t^A = (1-\rho)[\alpha + \gamma_\pi(\pi - \pi^-) + \gamma_x x_t] + \rho i_{t-1}^A + \xi_t \tilde{\imath}_t = (1-\rho)[\gamma_\pi \pi_t + \gamma_x' x_t] + \tilde{\rho}\tilde{\imath}_{t-1} + \xi_t'$

Suppose that y_t is an $m \times 1$ vector of observed data, $t=1,\ldots,T$, $x_t(\theta)$ is an $m \times 1$ vector of simulated time series generated from the structural macroeconomic model, θ is a $k \times 1$ vector of the parameters of the macroeconomic model and $x_t(\theta)$ and y_t are assumed to be stationary and ergodic. The auxiliary model is $f[y_t, \alpha]$. We assume that there exists a particular value of θ given by θ_0 such that $\{x_t(\theta_0)\}_{s=1}^S$ and $\{y_t\}_{t=1}^T$ share the same distribution, i.e.

$$f[x_t(\theta_0), a] = f[y_t, \alpha]$$

where α is the vector of parameters of the auxiliary model, and the existence of a binding function relating θ to α.

The likelihood function for the auxiliary model defined for the observed data $\{y_t\}_{t=1}^T$ is

$$\mathscr{L}_T(y_t; \alpha) = \sum_{t=1}^T \log f[y_t, \alpha]$$

The maximum likelihood estimator of α is then

$$a_T = \arg \max_\alpha \mathscr{L}_T(y_t; \alpha)$$

The corresponding likelihood function based on the simulated data $\{x_t(\theta)\}_{s=1}^S$ is

$$\mathscr{L}_S[x_t(\theta); \alpha] = \sum_{t=1}^S \log f[x_t(\theta), \alpha]$$

with

$$a_S(\theta) = \underset{a}{\operatorname{argmax}} \, \mathscr{L}_S[x_t(\theta); \alpha]$$

The simulated quasi maximum likelihood estimator (SQMLE) of θ is

$$\theta_{T,S} = \underset{\theta}{\operatorname{argmax}} \, \mathscr{L}_T[y_t; \alpha_S(\theta)]$$

This value of θ corresponds to the value of α that maximises the likelihood function using the observed data. Further, as $x_t(\theta)$ and y_t are assumed to be stationary and ergodic, from Canova (2005),

$$plim \ a_T = plim \ a_S(\theta) = \alpha.$$

it can then be shown that

$$T^{1/2}(a_S(\theta) - \alpha) \rightarrow N[0, \Omega(\theta)]$$

$$\Omega(\theta) = E\left[-\frac{\partial^2 \mathscr{L}[\alpha(\theta)]}{\partial \alpha^2}\right]^{-1} E\left[\frac{\partial \mathscr{L}[\alpha(\theta)]}{\partial \alpha}\frac{\partial \mathscr{L}[\alpha(\theta)]'}{\partial \alpha}\right] E\left[-\frac{\partial^2 \mathscr{L}[\alpha(\theta)]}{\partial \alpha^2}\right]^{-1}$$

The covariance matrix can be obtained either analytically or by bootstrapping the simulations.

The method of simulated moments estimator (EMSME) may be extended to estimating a function $g(\theta)$ of θ. Let $g(a_T)$ and $g(\alpha_S(\theta))$ denote a continuous $p \times 1$ vector of functions which could, for example, be moments or scores, and let the mean functions be $G_T(a_T) = \frac{1}{T}\sum_{t=1}^{T}g(a_T)$ and $G_S(\alpha_S(\theta)) = \frac{1}{S}\sum_{s=1}^{S}g(\alpha_S(\theta))$. We require that $a_T \rightarrow \alpha_S$ in probability and that $G_T(a_T) \rightarrow G_S(\alpha_S(\theta))$ in probability for each θ. The EMSME is

$$\theta_{T,S} = \underset{\theta}{\text{argmin}} \ [G_T(a_T) - G_S(\alpha_S(\theta))]' W(\theta)[G(a_T) - G_S(\alpha_S(\theta))]$$

where $W(\theta_0)$ is the inverse of the variance-covariance matrix of the distribution of $G_S(a_S) - \overline{G[a_S(\theta_0)]}$. The estimator is consistent and asymptotically normal- Smith (1993), Gourieroux et al. (1993) and Canova (2005).

4.2 Model Evaluation

In model evaluation indirect inference is used in a different way. The aim here is to compare the performance of an auxiliary model based on observed data with its performance based on data simulated from a calibrated or previously estimated macroeconomic model. We choose the auxiliary model to be a VAR and base our test on a function of the VAR coefficients. The test statistic is formed from the minimand of the EMSME evaluated using estimates of α derived from observed data and data simulated from the given numerically specified DSGE model. The distribution of this Wald-type of test statistic is obtained numerically through bootstrapping.

Non-rejection of the null hypothesis is taken to indicate that dynamic behaviour of the macroeconomic model is not significantly different from that of the observed data. Rejection is taken to imply that the macroeconomic model is incorrectly specified. Comparison of the impulse response functions of the observed and simulated data should reveal in what respects the macroeconomic model fails to capture the auxiliary model.

A formal statement of the inferential problem is as follows. Using the same notation as before, we define y_t an $m \times 1$ vector of observed data ($t = 1, \ldots, T$), $x_t(\theta)$ an $m \times 1$ vector of simulated time series of S observations generated from the structural macroeconomic model, θ a $k \times 1$ vector of the parameters of the macroeconomic model. $x_t(\theta)$ and y_t are assumed to be stationary and ergodic. We set $S = T$ since we require that the actual data

sample be regarded as a potential replication from the population of bootstrapped samples. The auxiliary model is $f[y_t, \alpha]$; an example is the $VAR(p) y_t = \sum_{i=1}^{p} A_i y_{t-1} + \eta_t$ where α is a vector comprising elements of the A_t and of the covariance matrix of y_t. Under the null hypothesis $H_0 : \theta = \theta_0$, the stated values of θ whether obtained by calibration or estimation; the auxiliary model is then $f[x_t(\theta_0), \alpha(\theta_0)] = f[y_t, \alpha]$. We wish to test the null hypothesis through the $q \times 1$ vector of continuous functions $g(\alpha)$. Such a formulation includes impulse response functions. Under $H_0 : g(\alpha) = g[\alpha(\theta_0)]$.

If a_T denotes the estimator of α using actual data and $a_S(\theta_0)$ is the estimator of α based on simulated data for θ_0, we may obtain $g(a_T)$ and $g[a_S(\theta_0)]$. Using N independent sets of simulated data obtained using the bootstrap we can also define the bootstrap mean of the $g[a_S(\theta)]$, $\overline{g[a_S(\theta_0)]} = \frac{1}{N} \sum_{k=1}^{N} g_k[a_S(\theta_0)]$. The Wald test statistic is based on the distribution of $g(a_T) - \overline{g[a_S(\theta_0)]}$ where we assume that $g(a_T) - \overline{g[a_S(\theta_0)]} \xrightarrow{p} 0$. The resulting Wald statistic (WS) may be written as

$$WS = \left(g(a_T) - \overline{g[a_S(\theta_0)]} \right)' W(\theta_0) \left(g(a_T) - \overline{g[a_S(\theta_0)]} \right)$$

where $W(\theta_0)$ is the inverse of the variance-covariance matrix of the distribution of $g(a_T) - \overline{g[a_S(\theta_0)]}$. $W(\theta_0)^{-1}$ can be obtained from the asymptotic distribution of $g(a_T) - \overline{g[a_S(\theta_0)]}$ and the asymptotic distribution of the Wald statistic would then be chi-squared. The empirical distribution of the Wald statistic is derived using bootstrap methods as follows.

> *Step 1: Determine the errors of the economic model conditional on the observed data and $\widehat{\theta}$.*
>
> Solve the DSGE macroeconomic model for the structural the errors ε_t given $\widehat{\theta}$ and the observed data.[7] The number of independent structural errors is taken to be less than or equal to the number of endogenous variables. The errors are not assumed to be normal.
>
> *Step 2: Construct the empirical distribution of the structural errors*
>
> On the null hypothesis the $\{\varepsilon_t\}_{t=1}^{T}$ errors are omitted variables. Their empirical distribution is assumed to be given by these structural errors. The simulated disturbances are drawn from these errors. In some DSGE models the structural errors are assumed to be generated by autoregressive processes. This is the case with the New Keynesian model here; we discuss below the precise assumptions made.

As our test is based on a comparison of the VAR coefficient vector itself rather than a multi-valued function of it such as the IRFs

$$g(a_T) - g(\alpha_S(\theta)) = a_T - \alpha_S(\theta)$$

and

---

[7] Some equations may involve calculation of expectations. The method we use here is the robust instrumental variables estimation suggested by McCallum (1976) and Wickens (1982): we set the lagged endogenous data as instruments and calculate the fitted values from a VAR(1)—this also being the auxiliary model chosen in what follows.

$$G_T(a_T) - G_S\left(\alpha_S\left(\widehat{\theta}\right)\right) = a_T - \alpha_S\left(\widehat{\theta}\right)$$

The distribution of $a_T - \alpha_S\left(\widehat{\theta}\right)$ and its covariance matrix $W\left(\widehat{\theta}\right)^{-1}$ are estimated by bootstrapping $\alpha_S\left(\widehat{\theta}\right)$. We draw N bootstrap samples of the structural model and estimate the auxiliary VAR on each.[8] N is generally set to 1,000. From these samples we compute the sample mean and covariance matrix. We also obtain the bootstrap distribution of the Wald statistic $\left[a_T - \alpha_S\left(\widehat{\theta}\right)\right]' W\left(\widehat{\theta}\right)\left[a_T - \alpha_S\left(\widehat{\theta}\right)\right]$ and a confidence interval. Such a distribution is generally more accurate for small samples than the asymptotic distribution. It is also shown by Le et al. (2011) to be consistent as the Wald statistic is asymptotically pivotal, and to have good accuracy in small sample Monte Carlo experiments.[9]

4.3 An Extension

A potential problem is that the given numerical values of θ that are being tested are not the 'true' values of θ. We therefore extend our test procedure by searching for alternative values of θ that might perform better in the test. This involves calculating a minimum-value Wald statistic for each period using a powerful algorithm based on Simulated Annealing (SA) in which search takes place in wide neighbourhood of the initial values of θ where the optimising search is accompanied by random jumps around the parameter space[10]. The merit of this extended procedure is that we are then testing the best possible numerical specifications of each model against actual data.

Several outcomes are possible.

a. One model is rejected, but the other is not. In this case only one model is compatible with actual data and the other can therefore be disregarded.
b. Both models are rejected, but the Wald statistic of one is lower than that of the other.
c. Neither model is rejected, but the Wald statistic for one is lower than for the other.

---

[8] The bootstraps in our tests are all drawn as time vectors so contemporaneous correlations between the innovations are preserved.

[9] Specifically, they found that the bias due to bootstrapping was just over 2 % at the 95 % confidence level and 0.6 % at the 99 % level. They suggested possible further refinements in the bootstrapping procedure which could increase the accuracy further; however, we do not feel it necessary to pursue these here.

[10] We use a Simulated Annealing algorithm due to Ingber (1996). This mimics the behaviour of the steel cooling process in which steel is cooled, with a degree of reheating at randomly chosen moments in the cooling process—this ensuring that the defects are minimised globally. Similarly the algorithm searches in the chosen range and as points that improve the objective are found it also accepts points that do not improve the objective. This helps to stop the algorithm being caught in local minima. We find this algorithm improves substantially here on a standard optimisation algorithm- Chib et al (2010) report that in their experience the SA algorithm deals well with distributions that may be highly irregular in shape, and much better than the Newton–Raphson method.

Our method used our standard testing method: we take a set of model parameters (excluding error processes), extract the resulting residuals from the data using the LIML method, find their implied autoregressive coefficients (AR(1) here) and then bootstrap the implied innovations with this full set of parameters to find the implied Wald value. This is then minimised by the SA algorithm.

The *p*-value of the Wald statistic can loosely be described as the probability of the model being true given the data: it is the percent of the distribution to the right of the data Wald value, or (100 minus the Wald percentile)/100. Hence, ranking the models by their *p*-value, suggests in case b) that this ranking is merely information about possible misspecifications. In case c) it suggests that one can regard the model with the higher *p*-value as the better approximation to the 'true' model.

4.4 Comment

We explained earlier that we have chosen to use indirect inference to test a model, whether or not it has been estimated by Bayesian methods, rather than standard Bayesian evaluation methods because we wish to test the whole model against the data, including the assumptions embodied in the priors in the case of Bayesian estimates. Even a major model like the Smets and Wouters (2007) model of the U.S., that has been carefully estimated by Bayesian methods, is rejected by our indirect inference test, see Le et al. (2011). We also claimed that indirect inference has greater power than another direct test, the Likelihood Ratio (LR) test. This claim is based on Le et al. (2012) who find that LR is much less powerful in small samples as a test of specification than a Wald test based on indirect inference. Presumably this result is related to the nature of the two tests. The LR test is based on a model's in-sample current forecasting ability, whereas the Wald is an in-sample test based on the model's ability to replicate data behaviour as represented by the VAR coefficients and the data variances which reflect the causal processes at work in the data. Models that are somewhat mis-specified may still be able to forecast well in sample as the error processes will capture some of the effects of mis-specification, but mis-specified models imply a reduced form that differs materially from the true one. Similarly, a VAR approximation to a mis-specified reduced form will deviate from the VAR associated with the true model.

Table 2 reproduces the findings by Le et al. (2012) comparing the two tests (on the Smets-Wouters model, for a 3-variable VAR(1)).

In sum, we could use LR instead of indirect inference as a test of our competing models. But it would be a much weaker test and hence we would get much less discrimination between the models. As will be seen below, the indirect inference Wald test discriminates powerfully between these models.

5 Data and Results

We evaluate the models against the US experience since the breakdown of the Bretton Woods system using quarterly data published by the Federal Reserve Bank of St. Louis from 1972 to 2007.[11] This covers both the Great Acceleration and the Great Moderation episodes of the US history.

The time series involved for the given baseline model include $\widetilde{\iota}_t$, measured as the deviation of the current Fed rates from its steady-state value, the output gap x_t,

---

[11] http://research.stlouisfed.org/fred2/.

Table 2 Rejection rates for Wald and likelihood ratio for 3-Variable VAR(1)

Model falseness (%)	Rejection rate	
	Wald	LR
0	5	5
1	19.8	6.3
3	52.1	8.8
5	87.3	13.1
7	99.4	21.6
10	100	53.4
15	100	99.3
20	100	99.7

Source: Le et al. (2012)

approximated by the percentage deviation of real GDP from its HP trend, and the quarterly rate of inflation π_t, defined as the quarterly log difference of the CPI[12]

We should find a break in the VAR process reflecting the start of the Great Moderation. Accordingly we split the time series into two subsamples and estimate the VAR representation before and after the break; the baseline model is then evaluated against the VAR of each subsample separately. We set the break at 1982Q3. Most discussions of the Fed's behaviour (especially those based on Taylor Rules) are concerned with periods that begin sometime around the mid-1980s but we chose 1982 as the break point here because many (including Bernanke and Mihov 1998, and Clarida et al. 2000) have argued that it was around then that the Fed switched from using non-borrowed reserves to setting the Fed Funds rate as the instrument of monetary policy. Such a choice is consistent with the Qu and Perron (2007) test which gives a 95 % confidence interval between 1980Q1 and 1984Q4 .[13]

For simplicity, the data we use are demeaned so that a VAR(1) representation of them contains no constants but only nine autoregressive parameters in the coefficient

---

[12] Note by defining the output gap as the HP-filtered log output we have effectively assumed that the HP trend approximates the flexible-price output in line with the bulk of other empirical work. To estimate the flexible-price output from the full DSGE model that underlies our three-equation representation, we would need to specify that model in detail, estimate the structural shocks within it and fit the model to the unfiltered data, in order to estimate the output that would have resulted from these shocks under flexible prices. This is a substantial undertaking well beyond the scope of this paper, though something worth pursuing in future work.

Le et al. (2011) test the Smets and Wouters (2007) US model by the same methods as we use here. This has a Taylor Rule that responds to flexible-price output. It is also close to the timeless optimum since, besides inflation, it responds mainly not to the level of the output gap but to its rate of change and also has strong persistence so that these responses cumulate strongly. Le et al. find that the best empirical representation of the output gap treats the output trend as a linear or HP trend instead of the flexible-price output—this Taylor Rule is used in the best-fitting 'weighted' models for both the full sample and the sample from 1984. Thus while in principle the output trend should be the flexible-price output solution, it may be that in practice these models capture this rather badly so that it performs less well than the linear or HP trends.

We have also purposely adjusted the annual Fed rates from the Fred to quarterly rates so the frequencies of all time series kept consistent on quarterly basis. The quarterly interest rate instead state is given by $i_{ss} = \frac{1}{\beta} - 1$.

[13] The Qu-Perron test suggests 1984Q3 as the most likely within the range. We show in the Supporting Annex that our tests are robust to this later choice of switch date.

matrix; a linear trend is also taken out of the interest rate series for the post-1982 sample to ensure stationarity.[14]

The model is calibrated by choosing the parameters commonly accepted for the US economy in the literature. The error processes these imply for each structural equation are then backed out and estimated as explained above. As we will go on to re-estimate all these parameters in a second stage of evaluation, we comment further on them at that point. We now go on to review the performance of each model with these calibrated parameters; since these are widely used in other papers, this allows us to relate our findings more easily to existing work, as well as illustrating the essential elements of our methods.

5.1 Results With Calibrated Parameters

The test results for the models considered are presented in what follows; these are based on the nine autoregressive coefficients of a VAR(1) representation and three variances of the model variables, the chosen descriptors of the dynamics and volatility of the data as discussed above. Our evaluation is based on the Wald test, and we calculate two kinds of Wald statistic, namely, a 'directed Wald' that accounts either only for dynamics (the VAR coefficients) or only for the volatility (the variances) of the data, and a 'full Wald' where these features are jointly evaluated. In both cases we report the Wald statistic as a percentile, i.e. the percentage point where the data value comes in the bootstrap distribution. The models' performance in each subsample follows.

5.1.1 Model Performance in the Great Moderation

We start with the post-1982 period, the Great Moderation subsample, as this has been the main focus of econometric work to date. Table 3 summarises the performance of the different models. The Optimal Timeless Rule model passes the tests by a comfortable margin, both overall, with a Wald percentile of 77.1 (implying a *p*-value of 0.229), and specifically on the dynamics alone (a *p*-value of 0.136) and the volatilities alone (0.104). The conclusion is that the US facts do not reject the Timeless Rule model as the data-generating process post-1982.

This is not the case, however, when Taylor Rules of the standard sort are substituted for it. The Table 3 suggests when the original Taylor Rule or the interest-rate-smoothed Taylor Rule is combined with the same IS-Phillips curve framework on these

---

[14] We show the plots and unit root test results of these in the Annex.

T-value normalization of the Wald percentiles is calculated based on Wilson and Hilferty (1931)'s method of transforming chi-squared distribution into the standard normal distribution. The formula used here is:

$$Z = \left\{ \left[(2M^{squ})^{1/2} - (2n)^{1/2} \right] \middle/ \left[\left(2M^{squ^{95th}} \right)^{1/2} - (2n)^{1/2} \right] \right\} \times 1.645$$

where M^{squ} is the square of the Mahalanobis distance calculated from the Wald statistic equation with the real data, $M^{squ^{95th}}$ is its corresponding 95 % critical value on the simulated (chi-squared) distribution, n is the degrees of freedom of the variate, and Z is the normalized *t* value; it can be derived by employing a square root and assuming n tends to infinity when the Wilson and Hilferty (1931)'s transformation is performed.

Table 3 Wald statistics for calibrated models in the great moderation

Tests for chosen data features	Baseline model with		
	Opt. Timeless Rule (model one)	Original Taylor Rule (model two)	'IRS' Taylor Rule (model three)
Directed Wald for dynamics	86.4	100	99.8
Directed Wald for volatilities	89.6	99.2	99
Full Wald for dynamics & volatilities	77.1	100	99.7

commonly accepted calibrations, from all perspectives the post-1982 data strongly reject the model at 99 %.

5.1.2 Model Performance in the Great Acceleration

We now proceed to evaluate how the models behave before 1982, the Great Acceleration period. Table 4 reveals the performance of the Optimal Timeless Rule model.

We can see that although the model does not behave as well here as it did in the Moderation subsample in explaining the data dynamics, with a directed Wald of 98.2, the directed Wald for data volatilities at 89.6 lies within the 90 % confidence bound. Overall, the full Wald percentile of 97.3 falls between the 95 % and the 99 % confidence bounds. So while the model fits the facts less well than in the case of the Great Moderation, it just about fits those of the turbulent Great Acceleration episode if we are willing to reject at a higher threshold. As we will see next, it also fits them better than its rival Taylor Rule models.

Unfortunately we are unable to test the DSGE model with the generally proposed pre-1982 Taylor Rules because the solution is indeterminate, the model not satisfying the Taylor Principle. Such models have a sunspot solution and therefore any outcome is possible and also consistent formally with the theory. The assertion of those supporting such models is that the solutions, being sunspots, accounted for the volatility of inflation. Unfortunately there is no way of testing such an assertion. Since a sunspot can be anything, any solution for inflation that occurred implies such a sunspot—equally of course it might not be due to a sunspot, rather it could be due to some other unspecified model. There is no way of telling. To put the matter technically in terms of indirect inference testing using the bootstrap, we can solve the model for the sunspots that must have occurred to generate the outcomes; however, the sunspots that occurred cannot be meaningfully bootstrapped because by definition the sunspot variance is infinite. Values drawn from an infinite-variance distribution cannot give a valid estimate of the distribution, as they will represent it with a finite-variance distribution. To draw representative random values we would have to impose an infinite variance; by implication all possible outcomes would be embraced by the simulations of the model

Table 4 Wald statistics for calibrated models in the great acceleration

| Baseline model with | Wald percentiles for chosen features[a] | | |
	Directed Wald for dynamics	Directed Wald for volatilities	Full Wald for dyn. & vol.
Opt. Timeless Rule	98.2	89.6	97.3
'Weak Taylor Rule' variants ($\widetilde{\iota}_t = \rho\widetilde{\iota}_{t-1} + \gamma_\pi \pi_t \xi_t$)			
$\rho = 0$, $\gamma_\pi = 1.001$ $\widetilde{\xi}_t$ AR(1)	100 (39.81)	78.9 (0.22)	100 (40.24)
$\rho = 0.3$, $\gamma_\pi = 0.7007$ $\widetilde{\xi}_t$ AR(1)	100 (30.26)	92 (1.08)	100 (28.01)
$\rho = 0.5$, $\gamma_\pi = 0.5005$ $\widetilde{\xi}_t$ AR(1)	100 (22.69)	95.9 (1.77)	100 (21.98)
$\rho = 0.7$, $\gamma_\pi = 0.3003$ $\widetilde{\xi}_t$ iid	100 (19.26)	98.2 (2.73)	100 (18.24)
$\rho = 0.9$, $\gamma_\pi = 0.1001$ $\widetilde{\xi}_t$ iid	100 (9.09)	99 (3.56)	100 (9.03)

[a] Normalized *t*-values in parenthesis

and hence the model cannot be falsified. Thus the pre-1982 Taylor Rule DSGE model proposed is not a testable theory of this period.[15]

However, it is open to us to test the model with a pre-1982 Taylor Rule that gives a determinate solution; we do this by making the Taylor Rule as unresponsive to inflation as is consistent with determinacy, implying a long-run inflation response of just above unity. Such a rule shows considerably more monetary 'weakness' than the rule typically used for the post-1982 period, as calibrated here with a long-run response of interest rate to inflation of 1.5.

We implement this weak Taylor Rule across a spectrum of combinations of smoothing coefficient and short-run response to inflation, with in all cases the long-run coefficient equalling 1.001. The Wald test results are shown in Table 4. What we see here is that with a low smoothing coefficient the model encompasses the variance of the data well, in other words picking up the Great Acceleration. However, when it does so, the data dynamics reject the model very strongly. If one increases the smoothing coefficient, the model is rejected less strongly by the data dynamics and also overall but it is then increasingly at odds with the data variance. In all cases the model is rejected strongly overall by the data, though least badly with the highest smoothing coefficient. Thus the testable model that gets nearest to the position that the shift in US post-war behaviour was due to the shift in monetary regime (reflected in Taylor Rule coefficients) is rejected most conclusively.

---

[15] We could use the approach suggested in Minford et al. (2011) in which the monetary authority embraces a terminal condition designed to eliminate imploding (as well as exploding) sunspots. In this case the model is forced to a determinate solution even when the Taylor Principle does not hold. However in our sample here we find that the model only fails to be rejected with inflation response parameters well in excess of unity—see below—while as we see from Table 4 being consistently rejected for parameters that get close to unity. So parameter values below unity, where the Taylor Principle does not apply, seem unlikely to fit the facts and we have not therefore pursued them here using this terminal condition approach.

5.2 Simulated Annealing and Model Tests With Final Parameter Selection

The above results based on calibration thus suggest that the Optimal Timeless Rule, when embedded in our IS-Phillips curves model, outperforms testable Taylor Rules of the standard sort in representing the Fed's monetary behaviour since 1972. In both the Great Acceleration and the Great Moderation the only model version that fails to be strongly rejected is the one in which the optimal timeless policy was effectively operating[16].

However, fixing model parameters in such a way is an excessively strong assumption in terms of testing and comparing DSGE models. This is because the numerical values of a model's parameters could in principle be calibrated anywhere within a range permitted by the model's theoretical structure, so that a model rejected with one set of assumed parameters may not be rejected with another. Going back to what we have just tested, this could mean that the Taylor Rule models were rejected not because the policy specified was incorrect but because the calibrated IS and Phillips curves had failed to reflect the true structure of the economy. Thus, to compare the Timeless Rule model and Taylor Rule models thoroughly one cannot assume the models' parameters are fixed always at particular values; rather one is compelled to search over the full range of potential values the models can take and test if these models, with the best set of parameters from their viewpoints, can be accepted by the data.

Accordingly we now allow the model parameters to be altered to achieve for each model the lowest Wald possible, subject to the theoretical ranges permitted by the model theory.[17] This estimation method is that of Indirect Inference; we use the Simulated Annealing (SA) algorithm for the parameter search, as discussed above in Section 4. In this process we allow each model to be estimated with different

---

[16] In a recent paper Ireland ([2007](#)) estimates a model in which there is a non-standard 'Taylor Rule' that is held constant across both post-war episodes. His policy rule always satis.es the Taylor Principle because unusually it is the change in the interest rate that is set in response to inflation and the output gap so that the long-run response to inflation is infinite. He distinguishes the policy actions of the Fed between the two subperiods not by any change in the rule's coefficients but by a time-varying inflation target which he treats under the assumptions of 'opportunism' largely as a function of the shocks to the economy. Ireland's model implies that the cause of the Great Moderation is the fall in shock variances. However, since these also cause a fall in the variance of the inflation target, which in turn lowers the variance of inflation, part of this fall in shock variance can be attributed to monetary policy.

    It turns out that Ireland's model is hardly distinguishable from our Optimal Timeless Rule model. His rule changes the interest rate until the Optimal Timeless Rule is satisfied, in effect forcing it on the economy. Since the Ireland rule is so similar to the Optimal Timeless Rule, it is not surprising that its empirical performance is also similar- thus the Wald percentiles for it are virtually the same.

    Ireland's rule can in principle be distinguished from the Optimal Timeless Rule via his restriction on the rule's error. However we cannot apply this restriction within our framework here so that Ireland's rule in its unrestricted form here only differs materially from the Optimal Timeless Rule in the interpretation of the error. However from a welfare viewpoint it makes little difference whether the cause of the policy error is excessive target variation or excessively variable mistakes in policy setting; the former can be seen as a type of policy mistake. Thus both versions of the rule imply that what changed in it between the two subperiods was the policy error. In effect, we can treat Ireland's rule as essentially the same in our model context as the Optimal Timeless Rule, and while he calls it a Taylor Rule, it is quite distinct from such rules as defined here.

[17] We fix the time discount factor β and the steady-state consumption-output ratio $\frac{C}{Y}$ as calibrated; other parameters are allowed to vary within ±50% of the calibrated values—which are set as initial values here— unless stated otherwise.

parameters for each episode. Thus we are permitting changes between the episodes in both structural parameters and the parameters of monetary policy; in so doing we are investigating whether either structural or policy rule changes were occurring and so contributing to the Great Moderation.[18]

5.2.1 The Estimated Optimal Timeless Rule Model

The SA estimates of the Timeless Rule model in both the post-war subperiods are reported in Table 5 - for the main parameters and Table 6 - for the shocks persistence. We can see that this estimated model is not very different from its calibrated version in the Great Moderation. However for the Great Acceleration period the estimation now suggests substantially lower elasticities of intertemporal consumption (the inverse of σ) and labour supply (the inverse of η), and a much higher Calvo contract non-adjusting probability (ω); with lower γ the latter implies a much flatter Phillips curve. The estimation also suggests the Fed had a low relative weight on output variations (α) pre-1982 but that high nominal rigidity forced it to reduce inflation more strongly in response to output growth (due to higher α/γ). The shocks' persistence is not much altered in either period from that in the calibrated model.

Table 7 shows that estimation brings the model substantially closer to the data. This is particularly so for the pre-1982 period where the calibrated model was rejected at 95 % confidence; here the necessary parameter changes were substantial to get the model to fit, as we have just seen. The Full Wald percentile in both episodes is now around 70 %, so that the model easily fails to be rejected at 95 %.

5.2.2 Taylor Rule Model Under Estimation

In estimating the Taylor Rule model alternative we substitute the smoothed version (Table 1) for the Optimal Timeless Rule in the identical IS-Phillips curves framework. This specification covers all Taylor Rule versions we considered in the earlier evaluation, as when ρ is zero it reduces to the original Taylor Rule while when γ_π is just above unity it turns to be a weak Taylor Rule variant.

As with the Optimal Timeless Rule model the estimation process achieves a substantial improvement in the closeness of the Taylor Rule model to the data in both episodes. Pre-1982 the best weak Taylor Rule version was strongly rejected; after estimation it is still rejected at the 95 % level but not at the 99 % level. Most importantly, the estimates include a much stronger Taylor Rule response to inflation than the calibrated version for this early episode; hence the evidence supports the view that the Taylor Rule principle was easily satisfied in this period. The response is essentially the same as that found in the later period by this estimation process: the weaker the response, the further the model is from fitting the data. Tables 5 and 6 show the details. The elasticity of intertemporal consumption and that of labour are found to

---

[18] It could be argued that deep parameters such as the elasticity of intertemporal substitution and Calvo price-change probabilities should remain fixed across the two periods. However, with such radically different environments these parameters could have differed; for example Le et al. (2011) find evidence that the degree of nominal rigidity varied across periods and interpret this as a response to changing variability. Here therefore we allow the data to determine the extent of change.

The results for the best testable weak Taylor Rule version as in Table 4.

Table 5 SA estimates of the competing models

Main model parameters	Calibrated values	Opt. Timeless Rule model SA estimates		Taylor Rule models SA estimates	
		pre-[a]	post-	pre-	post-
β	0.99	fixed	fixed	fixed	fixed
σ	2	1.01	1.46	1.15	1.16
η	3	2.04	3.23	2.66	3.85
ω	0.53	0.79	0.54	0.79	0.61
$\frac{G}{Y}$	0.23	fixed	fixed	fixed	fixed
$\frac{Y}{C}$	$\frac{1}{0.77}$	fixed	fixed	fixed	fixed
κ	0.42	0.06	0.40	0.06	0.25
γ	2.36	0.19	2.06	0.23	1.33
α	0.39	0.20	0.58	n.a	n.a
$\frac{\alpha}{\gamma} \equiv \frac{1}{\theta}$	$\frac{1}{6}$	$\frac{1}{0.95}$	$\frac{1}{3.6}$	n.a	n.a
θ	6	0.95	3.6	n.a	n.a
γ_π	1.44	n.a	n.a	2.03	2.06
γ'_x	0.14	n.a	n.a	.001	0.06
ρ	0.76	n.a	n.a	0.42	0.89

[a] Break point at 1982

be fairly similar to those estimated with the Optimal Timeless Rule, as is the Calvo rigidity parameter which is again higher in the first episode. For the model to get close to the data there needs to be interest rate smoothing in both episodes.

The resulting Wald statistics in Table 8 thus show that the Taylor Rule model is now close to passing at 95 % pre-1982 and passes comfortably post-1982. However relative to the Timeless Rule it is substantially further from the data, as summarized in Table 9 where the *p*-values are also reported. This suggests that, although it is possible to fit the post-1982 period with a Taylor Rule model, policy is better understood in terms of the Timeless Rule model.

Table 6 SA estimates of shock persistence

Persistence of shocks	Opt. Timeless Rule model Calibrated val.		SA estimates		Taylor Rule model Calibrated val.		SA estimates	
	pre-[a]	post-	pre-	post-	pre-	post-	pre-	post-
ρ_v	0.88	0.93	0.92	0.94	n.a	0.93	0.91	0.95
ρ_{u^w}	0.91	0.80	0.86	0.79	n.a	0.80	0.87	0.77
ρ_ε	0.59	0.38	0.14	0.42	n.a	0.39	0.58	0.40

[a] Break point at 1982

Table 7 Performance of the Timeless Rule Model under Calibration and Estimation

Tests for chosen features	Pre-1982 under		Post-1982 under	
	Calibration	Estimation	Calibration	Estimation
Directed Wald for dynamics	98.2	81.9	86.4	77.7
Directed Wald for volatilities	89.6	32.5	89.6	90.3
Full Wald for dynamics & volatilities	97.3	71.7	77.1	68.6

5.2.3 The Identification Problem Revisited in the Light of our Results

Having established that the Optimal Timeless Rule model gives the best representation of the key features of the US post-war data, we can now ask whether this model can also account for the single-equation findings for the Taylor Rule.

The above suggests that the widespread success reported in single-equation Taylor Rule regressions on US data could simply represent some sort of statistical relation emerging from the model with the Optimal Timeless Rule. To examine this possibility, we treat the Optimal Timeless Rule model as the true model and ask whether the existence of empirical Taylor Rules would be consistent with that. Technically this is again a process of model evaluation basing on indirect inference; but instead of a VAR here Taylor Rule regression coefficients are used as the data descriptors for the model to fit.

Table 10 shows the OLS estimates of several popular Taylor Rule variants when these are fitted, respectively, to data for both the post-war episodes. To compare the regression results here with those commonly found in the US Taylor Rule literature where unfiltered interest rate data is normally used we must emphasize that here for the post-1982 subsample a linear trend is taken out of the interest rate series so that stationarity is ensured. These Taylor Rules, when estimated on the stationary data we have used here, generally fail to satisfy the Taylor Principle, in much the same way as in pre-1982. Thus econometrically the standard estimates of the long-run Taylor Rule response to inflation post-1982 are biased by the non-stationarity of the interest rate. There is little statistical difference in the estimates across the two periods. The reported Wald percentiles indicate that these empirical 'Taylor Rules' are indeed consistent with what the Timeless Rule implies: in both panels the Taylor Rule regressions estimated

Table 8 Performance of Taylor Rule model under calibration and estimation

Tests for chosen features	Pre-1982 under		Post-1982 under	
	Calibration[a]	Estimation	Calibration	Estimation
Directed Wald for dynamics	100	98	99.8	89.6
Directed Wald for volatilities	99	40.6	99	94.9
Full Wald for dynamics & volatilities	100	96.1	99.7	87.6

[a] The results for the best testable weak Taylor Rule version as in Table 4

Table 9 Summary of model performance with estimated parameters

Tests for chosen features	Pre-1982 with		Post-1982 with	
	Timeless rule	Taylor rule	Timeless rule	Taylor rule
Directed Wald for dynamics (and p-value)	81.9 (0.181)	98 (0.020)	77.7 (0.223)	89.6 (0.104)
Directed Wald for volatilities (and p-value)	32.5 (0.675)	40.6 (0.604)	90.3 (0.097)	94.9 (0.051)
Full Wald for dyn. & vol. (and p-value)	71.7 (0.283)	96.1 (0.039)	68.6 (0.314)	87.6 (0.124)

are all within or on the 95 % confidence bounds implied by the estimated Timeless Rule Model.

This illustrates the identification problem with which we began this paper: a Taylor Rule regression having a good fit to the data may well be generated by a model where there is no structural Taylor Rule at all. Here we suggest that the Timeless Rule model we have found gets closest to fitting US data in each episode is also generating these Taylor Rule single-equation relationships.

5.2.4 The 'Interest Rate Smoothing' Illusion: A Further Implication

Another issue on which the above sheds light is the phenomenon of 'interest rate smoothing'. Clarida et al. (1999) noted that the Optimal Timeless Rule required nominal interest rate to be adjusted in a once-and-for-all manner, but that empirical evidence from Taylor Rule regressions usually displayed clear interest rate smoothing. This they argued created a 'puzzle': that sluggish interest rate movements could not be justified as optimal.

While various authors have tried to explain such a discrepancy either from an economic (e.g., Rotemberg and Woodford 1997, 1998; Woodford 1999, 2003a, b) or from an econometric (e.g., Sack and Wieland 2000; Rudebusch 2002) viewpoint, the

Table 10 'Taylor Rules' in the Data (with OLS): consistency with the estimated Timeless Rule Model

'Taylor Rule' versions	γ_π	γ_x	ρ	Adj.R^2	Wald percentiles
Panel A: 'Taylor Rules' in the Great Acceleration					
$\widetilde{\iota}_t = \gamma_\pi \pi_t + \gamma_x x_t + \rho \widetilde{\iota}_{t-1} + \xi_t$	0.09	0.06	0.90	0.84	31.9
$\widetilde{\iota} = \gamma_\pi \pi_t + \gamma_x x_t + \xi_t \quad \xi_t = \rho_\xi \xi_{t-1} + \varepsilon_t$	0.30	0.07	0.92	0.85	69.1
$\widetilde{\iota}_t = \gamma_\pi \pi_{t-1} + \gamma_x x_{t-1} + \xi_t$	0.60	−0.01	n/a	0.24	36.9
$\widetilde{\iota}_t = \gamma_\pi \pi_{t-1} + \gamma_x x_{t-1} + \rho \widetilde{\iota}_{t-1} + \xi_t$	−0.11	0.06	0.82	0.83	68.5
Panel B: 'Taylor Rules' in the Great Moderation					
$\widetilde{\iota}_t = \gamma_\pi \pi_t + \gamma_x x_t + \rho \widetilde{\iota}_{t-1} + \xi_t$	0.08	0.05	0.89	0.92	11.5
$\widetilde{\iota} = \gamma_\pi \pi_t + \gamma_x x_t + \xi_t \quad \xi_t = \rho_\xi \xi_{t-1} + \varepsilon_t$	0.07	0.06	0.93	0.90	95.6
$\widetilde{\iota}_t = \gamma_\pi \pi_{t-1} + \gamma_x x_{t-1} + \xi_t$	0.26	0.13	n/a	0.24	17.1
$\widetilde{\iota}_t = \gamma_\pi \pi_{t-1} + \gamma_x x_{t-1} + \rho \widetilde{\iota}_{t-1} + \xi_t$	0.03	0.04	0.89	0.91	89.7

Taylor Rule regressions above show that 'smoothing' is a regression result that is generated by the Optimal Timeless Rule model, in which there is no smoothing present. The source of inertia in the model is the persistence in the shocks themselves.

6 What Caused the Great Moderation?

We have found that the Optimal Timeless Rule is the best guide to US monetary policy since the Bretton Woods; we have also obtained estimates of the model under a Taylor Rule, which though fitting the data considerably less well nevertheless fail to be rejected in absolute terms by the data. These models enable us finally to examine the causes of the Great Moderation. We have made a number of empirical findings about changes in the structural parameters, the parameters of the monetary rule trade-off, and the behaviour of the shocks. We now examine the contribution of each of these changes to the Great Moderation.

Table 11 shows that under our preferred model with the Optimal Timeless Rule the Great Moderation is almost entirely the result of reduced volatility in the shocks. There is a small contribution to lowered inflation variance from the policy parameters; but otherwise the contribution from both structural and policy parameters is slightly to increase macro variance in the later period. If one then examines which shocks' volatility fell, the Table 12 following shows that it did so for all three of our shocks, with a fall in standard deviation of 60–70 %.

If we look at the Taylor Rule model the story is essentially the same. As we saw above the inflation response of the Taylor Rule hardly changes across the two periods. The main change is a doubling of the smoothing parameter which accordingly contributes about a third of the reduction in interest rate variance. Otherwise structural and policy parameter changes contribute negligibly to the variance reduction. Thus again the reduction in shock variability dominates as the cause of the Great Moderation. Here too all the shocks have large falls in standard deviation; the largest at 86 % is the monetary shock (Tables 13 and 14).

Thus what we find is that the Great Moderation is essentially a story of 'good shocks' as proposed in the time-series studies we cited earlier. Also we have found no evidence of the weak monetary regime regarded by an earlier DSGE model literature as responsible for the Great Acceleration and in the same vein no evidence of much change in the monetary regime during the Great Moderation. However, what we do find about monetary policy is that the 'trembling hand' trembled enormously more in the earlier period than in the later; thus monetary error is a large source of the Great Acceleration and its reduction an important reason for the Moderation. For those that

Table 11 Accountability of factor variations for reduced data volatility (Timeless Rule model)

Reduced data volatility caused by	Interest rate	Output gap	Inflation
Reduced shocks	115.3 %	106.9 %	90 %
Chg in policy paras	−4.3 %	−2.5 %	12.7 %
Chg in structural para	−11 %	−4.5 %	−2.7 %

Table 12 Reduced size of shocks (Timeless Rule model)

Standard deviation of	Pre-1982	Post-1982	Reduction
Demand shock	0.0625	0.02	60 %
	(0.0050)	(0.0012)	
Supply shock	0.4767	0.1419	70 %
	(0.0667)	(0.0298)	
Policy shock	0.0148	0.0055	63 %
	(0.0127)	(0.0032)	

[a] Values in parentheses are sample estimates of standard deviation of innovations

[b] The standard deviation of the shocks is calculated using sd(err.) = sd(innov)/(1-rho); rho is the sample estimate of shock persistence reported in Table 6

embrace a Taylor Rule model in spite of its poorer data fit the story is the same—in this case monetary 'judgement' was substantially more erratic in its effect in the earlier period.

6.1 A Comparison With Other Recent DSGE Models

As we noted earlier, Ireland (2007), Smets and Wouters (2007), Le et al. (2011) and Fernandez-Villaverde et al. (2009, 2010) have also estimated models of these periods and we can compare their results in a general way with ours. Other than Ireland, these models follow the model of Christiano et al. (2005). Smets and Wouters use this model with some small modifications; they estimate it by Bayesian methods. Le et al. add a competitive sector and reestimate the model using Indirect Inference, since they found the model was rejected quite badly overall by the data with the previously estimated ones. When reestimated in this way they found that the model was accepted, at 99 % for the full post-war period and at 95 % for the Great Moderation period, for the key subset of variables, output, inflation and interest rate when represented by a VAR(1). Fernandez-Villaverde, Guerron-Quintana and Rubio-Ramirez add moving volatility in the errors and drift in the parameters of the Taylor Rule; like Smets and Wouters they estimate the model by Bayesian methods.

What is striking about all these studies is that none of them find evidence of much difference in monetary regime between the two periods—interestingly, Fernandez-Villaverde, Guerron-Quintana and Rubio-Ramirez find variations of 'monetary toughness' *within* both periods, while not finding much difference on average *across* the two. Both Smets and Wouters and Le et al. in their reworking of them find little change in the inflation response coefficient of the Taylor Rule. In this these models echo Ireland,

Table 13 Accountability of factor variations for reduced data volatility (II) (Taylor Rule model)

Reduced data volatility caused by	Interest rate	Output gap	Inflation
Reduced shocks	66.7 %	99 %	99.9 %
Chg in policy paras	34.2 %	1.8 %	3.9 %
Chg in structural para	−0.9 %	−0.8 %	−3.8 %

Table 14 Reduced size of shocks (II) (Taylor Rule model)

Standard deviation of	Pre-1982	Post-1982	Reduction
Demand shock	0.0533 (0.0050)	0.0280 (0.0012)	48 %
Supply shock	0.5777 (0.0751)	0.1474 (0.0339)	75 %
Policy shock	0.0145 (0.0061)	0.0020 (0.0012)	86 %

[a] Values in parentheses are sample estimates of standard deviation of innovations

[b] The standard deviation of the shocks is calculated using sd(err.) = sd(innov)/(1-rho); rho is the sample estimate of shock persistence reported in Table 6

even though their Taylor Rule representations differ from his. Thus these studies agree with ours in finding that it is the shocks that account for the difference in volatility.

Nevertheless, all also agree with us that the scale of the monetary shock has declined between the two periods. Thus a pattern is visible in ours, Ireland's and these other studies: while the monetary regime did not apparently change much, the scale of the monetary 'error' fell between the two periods. Ireland interprets this, based on his connection of it with other shocks, as 'opportunism', where the Fed was allowing the inflation target to drift with events, pushing it downwards when events allowed this to be done with less perceived cost. Other studies, like ours, do not model it other than as a pure error.

An important implication of the lack of regime change is that there is no evidence of indeterminacy in the earlier period according to any of these studies including ours. Thus all these studies that are based on full information system estimates cannot find the evidence that appears to come out of single-equation studies that the earlier period's Taylor Rule responded weakly to inflation. As we have seen this is consistent with the lack of identification of the Taylor Rule as a single equation; indeed as we have seen the models that fit the data overall could easily have 'generated' single-equation Taylor Rules of this 'weak' type.

7 Conclusion

In this study we have used the method of indirect inference to estimate and test a three-equation DSGE model against the data for the Great Acceleration and the Great Moderation. The method has the advantage over alternatives that it tests the model overall in its ability to fit the data's behaviour. Nevertheless, in spite of differences in method, our results echo those of other recent work where DSGE models of greater complexity than ours have been estimated by a variety of methods. We have found that the monetary regimes being followed in the two periods are rather similar. We have also found that, while these regimes can be represented by Taylor Rules of the usual sort, they more closely fit the facts if represented by an Optimal Timeless Rule, essentially the same as the Taylor Rule form suggested by Ireland, which he also finds fits the facts best.

A corollary of this finding is that there is no evidence of indeterminacy due to the 'weakness' of the monetary regime during the Great Acceleration. Previous findings to

this effect seem to have arisen from single-equation estimates that suffered from a lack of identification and are quite consistent with the DSGE models estimated here.

By implication we also find, in common with these other studies using full DSGE models, that the Great Moderation was mainly the result of 'good shocks'—a fall in the variance of the errors in the model. This reinforces the results of a large number of time-series studies using Structural VARs, but it does so through finding structural DSGE parameters that can replicate these VARs and so allows them to be interpreted structurally.

Nevertheless, the falling variance of shocks includes that of monetary shocks. Within this fall lies the remedying of a failure of monetary policy. Whether this failure was due to an 'opportunistic' pursuit of varying inflation targets as in Ireland (2007), to sheer inefficiency, or to some other reason, our work cannot say; this remains a fruitful avenue for future work. Clearly and perhaps not surprisingly given the size and novelty of the shocks bombarding the 1970s economy, monetary policy was far from perfect in this early period. But at least we and other recent DSGE modellers are clear that it was not just plain stupid.

References

Benati L, Surico P (2009) VAR analysis and the great moderation. Am Econ Rev Am Econ Assoc 99((4)): 1636–1652

Bernanke BS, Mihov I (1998) Measuring monetary policy. Q J Econ 113(3):869–902, MIT Press

Boivin J, Giannoni MP (2006) Has monetary policy become more effective? Rev Econ Stat 88((3)):445–462, MIT Press

Calvo GA (1983) Staggered prices in a utility maximising framework. J Monet Econ 12:383–398

Canova F (2005) Methods for applied macroeconomic research. Princeton University Press, Princeton

Carare A, Tchaidze R (2005) The use and abuse of Taylor Rules: how precisely can we estimate them?. IMF Working Paper, No. 05/148, July

Castelnuovo E (2003) Taylor Rules, omitted variables, and interest rate smoothing in the US. Econ Lett 81(2003):55–59

Chib S, Kang KH, Ramamurthy S (2010) Term structure of interest rates in a DSGE model with regime changes. Olin Business School working paper, Washington University, http://apps.olin.wustl.edu/MEGConference/Files/pdf/2010/59.pdf, especially appendix D, p 46

Christiano LJ, Eichenbaum M, Evans CL (2005) Nominal rigidities and the dynamic effects of a shock to monetary policy. J Polit Econ 113(1):1–45, University of Chicago Press

Clarida R, Gali J, Gertler ML (1999) The science of monetary policy: a new keynesian perspective. J Econ Lit 37(4):1661–1707

Clarida R, Gali J, Gertler ML (2000) Monetary policy rules and macroeconomic stability: evidence and some theory. Q J Econ 115((1)):147–180, MIT Press

Cochrane JH (2011) Determinacy and Identification with Taylor Rules. J Polit Econ 119(3):565–615, University of Chicago Press

Fernandez-Villaverde J, Guerron-Quintana P, Rubio-Ramirez JF (2009) The new macroeconometrics: a Bayesian approach. In: O'Hagan A, West M (eds) Handbook of applied Bayesian analysis. Oxford University Press, Oxford

Fernandez-Villaverde J, Guerron-Quintana P, Rubio-Ramirez JF (2010) Reading the recent monetary history of the U.S., 1959–2007. Federal Reserve Bank of Philadelphia working paper, No. 10–15

Gambetti L, Pappa E, Canova F (2008) The structural dynamics of U.S. output and inflation: what explains the changes? J Money Credit Bank 40(2–3):369–388, Blackwell Publishing

Giannoni MP, Woodford M (2005) Optimal interest rate targeting rules. In: Bernanke BS, Woodford M (eds) Inflation targeting. University of Chicago Press, Chicago

Gourieroux C, Monfort A (1996) Simulation based econometric methods, CORE Lectures Series. Oxford University Press, Oxford

Gourieroux C, Monfort A, Renault E (1993) Indirect inference. J Appl Econ 8:S85–S118, John Wiley & Sons, Ltd

Gregory AW, Smith GW (1991) Calibration as testing: inference in simulated macroeconomic models. J Bus Econ Stat Am Stat Assoc 9(3):297–303

Gregory AW, Smith GW (1993) Calibration in macroeconomics. In: Maddla GS (ed) Handbook of statistics, 11th edn. Elsevier, St. Louis, pp 703–719

Ingber L (1996) Adaptive simulated annealing (ASA): lessons learned. Control Cybern 25:33

Ireland PN (2007) Changes in the Federal Reserve's Inflation Target: causes and consequences. J Money Credit Bank 39(8):1851–1882, Blackwell Publishing

Le VPM, Meenagh D, Minford APL, Wickens MR (2011) How much nominal rigidity is there in the US economy? Testing a new Keynesian DSGE model using indirect inference. J Econ Dyn Control 35(12): 2078–2104, Elsevier

Le VPM, Meenagh D, Minford APL, Wickens MR (2012) Testing DSGE models by Indirect inference and other methods: some Monte Carlo experiments. Cadiff University Working Paper Series, E2012/15, June

Lubik TA, Schorfheide F (2004) Testing for indeterminacy: an application to U.S. monetary policy. Am Econ Rev Am Econ Assoc 94(1):190–217

McCallum BT (1976) Rational Expectations and the Natural Rate Hypothesis: some consistent estimates. Econometrica Econ Soc 44(1):43–52

McCallum BT, Nelson E (2004) 'Timeless perspective vs. discretionary monetary policy'. Review, Federal Reserve Bank of St. Louis 86(2):43

Minford P (2008) Commentary on 'Economic projections and rules of thumb for monetary policy'. Review, Federal Reserve Bank of St. Louis 90(4):331

Minford P, Ou Z (2013) Taylor rule or optimal timeless policy? Reconsidering the Fed's behaviour since 1982. Econ Model 32:113–123

Minford P, Perugini F, Srinivasan N (2002) Are interest rate regressions evidence for a Taylor rule? Econ Lett 76(1):145–150

Minford P, Perugini F, Srinivasan N (2011) Determinacy in New Keynesian Models: a role for money after all? Int Financ 14(2):211–229

Nistico S (2007) The welfare loss from unstable inflation. Econ Lett 96(1):51–57, Elsevier

Primiceri GE (2005) Time varying structural vector autoregressions and monetary policy. Rev Econ Stud 72: 821–825, The Review of Economic Studies Limited

Qu Z, Perron P (2007) Estimating and testing structural changes in multivariate regressions. Econometrica 75(2):459–502

Rotemberg JJ, Woodford M (1997) An optimization-based econometric model for the evaluation of monetary policy. NBER Macroecon Annu 12:297–346

Rotemberg JJ, Woodford M (1998) An optimization-based econometric model for the evaluation of monetary policy. NBER Technical Working Paper, No. 233

Rudebusch GD (2002) Term structure evidence on interest rate smoothing and monetary policy inertia. J Monet Econ 49(6):1161–1187, Elsevier

Sack B, Wieland V (2000) Interest-rate smoothing and optimal monetary policy: a review of recent empirical evidence. J Econ Bus 52:205–228

Sims CA, Zha T (2006) Were there regime switches in U.S. monetary policy? Am Econ Rev Am Econ Assoc 96(1):54–81

Smets F, Wouters R (2007) Shocks and frictions in US business cycles: a Bayesian DSGE approach. Am Econ Rev Am Econ Assoc 97(3):586–606

Smith A (1993) Estimating nonlinear time-series models using simulated vector autoregressions. J Appl Econ 8:S63–S84

Stock JH, Watson MW (2002) Has the Business Cycle Changed and Why? NBER Working Papers 9127, National Bureau of Economic Research, Inc

Svensson LEO, Woodford M (2004) Implementing optimal policy through inflation-forecast targeting. In: Bernanke BS, Woodford M (eds) Inflation targeting. University of Chicago Press, Chicago

Taylor JB (1993) Discretion versus policy rules in practice. Carn-Roch Conf Ser Public Policy 39:195–214

Wickens MR (1982) The efficient estimation of econometric models with rational expectations. Rev Econ Stud 49(1):55–67, Blackwell Publishing

Wilson EB, Hilferty MM (1931) The distribution of chi-square. Proc Natl Acad Sci 17(12):684–688

Woodford M (1999) Optimal monetary policy inertia. Manch Sch 67:1–35, University of Manchester

Woodford M (2003a) Optimal interest rate smoothing. Rev Econ Stud 70(4):861–886

Woodford M (2003b) Interest and prices: foundations of a theory of monetary policy. Princeton University Press, Princeton