US Post-war Monetary Policy and the Great Moderation: was the Fed doing a good job?

by

Zhirong Ou

A Thesis Submitted in Fulfilment of the Requirements for the Degree of Doctor of Philosophy of Cardiff University

Economics Section of Cardiff Business School, Cardiff University

September 2011

CARDIFF
UNIVERSITY

CARDIFF
BUSINESS
SCHOOL

DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed ....[illegible].......... (candidate)

Date ...05/09/2011

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of ....PhD...........(insert MCh, Md, MPhil, PhD etc, as appropriate)

Signed ....[illegible].......... (candidate)

Date 05/09/2011

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.
Other sources are acknowledged by footnotes giving explicit references.

Signed ....[illegible].......... (candidate)

Date 05/09/2011

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ....[illegible].......... (candidate)

Date 05/09/2011

To Patrick

Acknowledgements

I am heartily thankful to my supervisor, Patrick Minford, who has been giving me non-stop care, support and encouragement from the initial to the final stage of my PhD studies, in my good time and bad.

I also owe David Meenagh a lot for his generous help and comments during my completion of this thesis.

Special thanks also goes to Kent Matthews, who has fostered me by offering with trust a variety of valuable research-related opportunities, and Peter Smith, for his kind support and effort in my earlier attempt of application to the ESRC.

Jingwen Fan, Michael Hatcher, Jing Jiao, Mai Le, Chunping Liu, Kateryna Onishchenko, Tiantian Zhang and my other colleagues and the members of the Economics section have all played an important role both in my research life and in my daily life in Cardiff, for which I would like to express my sincere thanks and that I treasure the days I spent with them.

Lastly, I would like to send my deepest gratitude to my parents and my wife, for their unconditional support and tolerance, care and love. I understand there is no way to pay you back, but I will keep working hard to make you proud.

Abstract

Using indirect inference based on a VAR this thesis confronts the US data from 1972 to 2007 with a standard New Keynesian model in which either an optimal timeless policy or a Taylor rule is assumed prevails. By comparing the models' performance in fitting the dynamics and size of the data, it finds in both the episodes of the Great Acceleration and the Great Moderation that the Fed's underlying behaviour was better understood as the timeless optimum either under standard calibration or under estimation. The implication is that to the final analysis the Great Moderation is a result of improved environment as the volatility of shocks has fallen, rather than one occurred as policy improved. Smaller Fed managerial errors caused the moderation in inflation. Smaller supply shocks caused the moderation in output and smaller demand shocks the moderation in interest rates. In either episode the same model with differing Taylor rules of the standard sorts generally fails to fit the data well. But the optimal timeless rule model could have generated data in which Taylor rule regressions could have been discovered, creating an illusion that the Fed was following such policies and that the improved economy was caused by changed pattern of these.

Table of Contents

Chapter 4

List of Tables

List of Figures

Introduction

Since the breakdown of Bretton Woods in 1972 the US economy behaved first rather badly (the 'Great Acceleration' or 'Great Inflation') and then from sometime in the early 1980s until around 2007 rather well (the 'Great Moderation'). While economists have attempted to understand why these episodes differed so much, improved policy and improved environment are the two factors mostly focused on. In general, studies based on DSGE models have pointed to policy as the main cause of the improvement; those based on time-series models have typically pointed to environment. These methods are similar in a way that they both build on one's understanding about the Fed's monetary behaviour and decompose the improvement of economy between that and the environment. The difference is that the former is founded in theory, whereas the latter is founded in facts. However, while DSGE analysis has the advantage of establishing causality, it cannot ensure that the data are well accounted for; time-series analysis accounts for the data, but connecting the result of this to theory is difficult. Yet, since causality can only be established by theory, which is valid only if it fits the facts, a thorough investigation would require a method of evaluation that is founded both in theory and in facts. Thus a DSGE model that is not strongly rejected by the data is needed.

This thesis applies such principle to a reconsideration of the causes of the Great Moderation. It follows the existing literature by first identifying the underlying monetary policy being followed and then decomposing the change in data variability based on that. But it adopts a novel way of evaluation that builds the argument on a DSGE model selected according to its capacity in fitting the data using the method of indirect inference. The latter combines the DSGE causal framework and a time-series description of data—the two elements used to be separately adopted in episode comparison—to ensure that both theory and data are used in harness. The exercise yields quite different implications compared to the conventional wisdom in the literature: on the operation of monetary policy, it suggests the Fed's post-war behaviour was better understood as the timeless optimum rather than the widespread consensus of Taylor rule with shift; on the causes of the Moderation, it suggests the improved economy was a result of improved environment and monetary management but not one of improved policy. It also shows that illusion of regime switch could have been generated under the perspective of unidentified Taylor rule regressions when the Fed was pursuing the optimal timeless policy.

In the rest of this thesis I first survey the recent literature in modelling the Fed's post-war monetary behaviour and the causes of the US improvement (chapter 1); the two issues are integrated in a way that understanding of the former grounds the analytical basis of the latter. I review the progress and problems in the development of these and suggest a way forward for a more rigorous evaluation. Chapter 2 then takes the Moderation episode as an example to demonstrate how the Fed's true policy can be identified using the new method and how this could be confused with other regime versions usually suggested in the literature. This is followed by chapter 3 that extends the exercise to the Acceleration episode and re-decomposes the causes of the Moderation based on the extended results; it also illustrates how conventional understanding of the improvement could have been biased due to biased understanding of the prevailing policy as chapter 2 considers. The thesis is then closed by chapter 4 that extends the earlier discussion based on standard calibration to one based on model estimation with some concluding remarks regarding robustness and the models' working in detail.

Chapter 1

Approaching the Fed's Post-war Monetary Policy and the Causes of the Great Moderation: progress, problems and the way forward

Introduction

This chapter surveys the recent literature in approaching the Fed's post-war monetary policy and its impact on causing the Great Moderation. Existing studies on the former have mostly focused on estimation of simple rules, much in the spirit of Taylor (1993); others have built on these efforts to decompose the US improvement into policy and environment. However, simple-rule estimation of monetary policy is often intruded by the problem of identification, which when smuggled into the decomposition process undermines the implication for the latter. This chapter looks into such problem of identification failure. The aim of it is to summarise the difficulties (in terms of identification) encountered the existing literature and point the way to a possible way forward.

This chapter is organised as follows: section 1.1 reviews the earlier efforts in modelling the Fed's behaviour basing on Taylor rules and deduces from that the inherent problem of identification with all single-equation estimations; section 1.2 moves on to the debate over the causes of the Great Moderation and reviews how conventional methods of decomposition have suggested the essence of the improvement; it also reviews the strength and limitations of these methods and explains how identification failure of the Fed's behaviour could have distorted the implications; this is followed by section 1.3 that points to the way forward; section 1.4 concludes.

1.1 Modelling the Monetary Behaviour of the Fed: a Method Basing on Simple Rules

The conduct of monetary policy is in general an intricate practice of central banks in reality. For example, the Federal Open Market Committee (FOMC) of the Fed runs regular meetings each year and forms consensus regarding the direction and course of monetary policy basing on the member's reports on the recent status and prospects of the economy; the Monetary Policy Committee (MPC) of the Bank of England has a similar procedure in setting the official Bank Rate although its decision is not made upon consensus but voting. This complex nature of policy making has made the banks' behaviour rather difficult to understand. Thus in the common exercise of monetary analysis the principle governing such behaviour is usually modelled by economists parsimoniously with sim ple rules; these could be explicit (such as a money supply rule or an interest-rate setting rule) or implicit (such as a fixed exchange rate or other economic relationships)[1].

The adoption of simple rules for policy analysis can be dated back to Friedman (1960), who argued the growth of monetary base should be kept at a constant rate (the k-percent rule) irrespective of business cycles. Later examples are the McCallum (1988) rule and Taylor (1993) rule that suggested policy instruments (base money in the former and short-run nominal interest rates in the latter) should be adjusted in a way that it 'leans against the wind' according to the feedback from the economy. Currie and Levine (1985) compared the impact of simple rules in single open economy and in independent aggregate economy and pointed that a price rule using exchange rates as policy instrument is best for individual economies but would be disastrously destabilizing if all countries do the same. These monetary proposals have prescribed what the central bank should do in pursuit of stabilization policy. Of course as remedies they do not necessarily coincide with any underlying policy truly being pursued. However the work by Taylor has shown the interest rates recommended by his rule were surprisingly close to the actual Federal Fund rates between 1987 and 1992, for which he claimed the Fed's behaviour over this period was also well described by the rule. Since then the use of Taylor-type formulae has become rather popular in modelling of the prevailing monetary policy, and most efforts along this line have focused on the case of US.

1.1. 1 Differing versions of Taylor rules

[1] Of course the nature of policy making has made it hardly convincible that any central bank in reality would ever bank its monetary decision solely upon a simple numerical formula. But these rules, by summarizing much of the relevant information (such as the instruments or feedbacks or goals) regarding policy making in a compact manner, are widely accepted at least as a proxy that approximates the systematic part of the actual behaviour of major central banks—see Stuart (1996) for example.

The original Taylor rule states the Federal Fund rate should react directly to two 'gaps', one between inflation and its target, and the other between output and its natural-rate level. Numerically it takes the form:

$$i_t^A = \pi_t^A + 0.5x_t + 0.5(\pi_t^A - \pi^*) + g \qquad [1.1]$$

where i_t^A is annual nominal rate of interest, π_t^A is the annual inflation averaged over the last four quarters and x_t is the percentage deviation of real GDP from trend (the 'output gap'); it also assumes the US has an inflation target π^* and real GDP growth rate g both equalling 2 percent.

A number of variants have then been built upon this basic specification according to different assumptions. For example, Clarida, Gali and Gertler (1999) suggested the central bank may for some reasons be quite cautious in reacting to change in the fundamentals so that the adjustment of nominal interest rates would be completed in several successive small steps. They then proposed a Taylor rule version where interest rates are 'smoothed' as:

$$i_t^A = (1-\rho)[\alpha + \gamma_\pi(\pi_t^A - \pi^*) + \gamma_x x_t] + \rho i_{t-1}^A \qquad [1.2]$$

with ρ representing the degree of policy inertia that could arise from the central banks' cautiousness due to limited knowledge about the true model of the economy. Other rationales in support of such smoothing behaviour also include policy-makers' fear of financial instability (Goodfriend, 1991; Campbell, 1995), their compromise over a policy change (Goodfriend, 1991), their exploitation of private agents' forward-looking behaviour (Rotemberg and Woodford, 1997, 1998) and the existence of possible non-rational expectation-generating mechanism (Brayton, Levin, Tryon and Williams, 1997; Woodford, 2010). This could also reflect the optimal response for the policy-makers when they are concerned with the volatility of nominal interest rates according to Woodford (2003).

There are also other Taylor rule versions involving lagging or leading the inflation and output gap terms to reflect possible backward-looking or forward-looking behaviours. Examples of these are Rotemberg and Woodford (just cited) where the setting of interest rates is subject to a 'decision lag' so that

$$i_t^A = i^* + \sum_{k=1}^{n_i} \gamma_i (i_{t-k}^A - i^*) + \sum_{k=0}^{n_\pi} \gamma_\pi (\pi_{t-k}^A - \pi^*) + \sum_{k=0}^{n_x} \gamma_x x_{t-k} \qquad [1.3]$$

and Clarida, Gali and Gertler (2000) who assumed monetary decisions are forecast-based, thus

$$i_t^A = i^* + \gamma_\pi (E_t \pi_{t+k} - \pi^*) + \gamma_x E_t x_{t+k} \qquad [1.4]^2$$

Other suggested modifications would involve the use of unemployment gap instead of output gap or the use of unemployment growth or output growth as feedback from the economy as Carare and Tchaidze (2005) have summarised.

Yet in a more recent investigation Ireland (2007) has also proposed a Taylor rule version that is substantially different from the standard sort as above. Ireland's alternative assumes the Fed is effectively concerned with the change (rather than the level) of nominal interest rates in reacting to business cycles and it would stabilize inflation around an implicit, time-varying target that is endogenously determined by the shocks to the economy under the assumption of 'opportunism'. The rule specified by Ireland takes the form:

$$i_t^A = i_{t-1}^A + \gamma_\pi (\pi_t - \pi^*) + \gamma_g (g_t - g) - \Delta \pi_t^* + \xi_t \qquad [1.5]$$

where π_t^* is now the implicit inflation target, $(g_t - g)$ is the deviation of output growth rate from its steady-state level and other variables have their usual meanings. Note equation [1.5] has incorporated an error term ξ_t to denote possible 'policy mistakes'. In reality these could reflect the Fed governors' occasional discretionary response to something else (such as exchange rates or other macro-conditions including frictions in the financial market) unaccounted for by the announced rule, or perhaps more commonly, pure managerial mistake in executing the proposed policy due to 'trembling hand'. Yet whatever the interpretation is, this error term capturing

[2] Note $E_t x_{t+k}$ in equation [1.4] is defined as the expected output gap average over the periods between t and t+k.

the stochastic component of policy-making process would in principle be involved in all econometric modelling of the prevailing monetary behaviour basing on simple rules, Taylor-type or others, standard or non-standard.

1.1.2 Estimation of Taylor rules and the problem of identification

Rules like the above are widely estimated in literature in searching for the existing monetary policy, either as single-equations via regressions or as part of a complete model via full-model estimation. For the case of US, it seems most authors have agreed the Fed's post-war policy could be well characterized by some sort of Taylor rules with 'interest rates smoothing' and that it has been more active in stabilization since the early 1980s. Of course consensus has never been made upon its behaviour in detail—as it never will—as what it implies would be dependent heavily on the assumed 'reaction function', the measurement of involved variables, the time span, as well as the method of estimation—Judd and Rudebusch (1998), Fair (2000), Orphanides (2001, 2002) and Rudebusch (2002) are examples. Yet even econometricians might be able to 'solve' these usual difficulties in applied work, estimation of Taylor rules would still face an identification problem that baffles our understanding of the true policy, especially when the estimates are regression-based (Minford, Perugini and Srinivasan, 2001, 2002; Cochrane, 2007).

The problem of (non-) identification occurs when an equation could be confused with others so it does not by itself identify what it is. In the context of Taylor rule estimation it refers to the situation in which the rule being estimated has the same functional form as other economic relations such that the meaning of the rule is unclear. These 'other relations' could in principle be anything implied by a DSGE model. For example, Minford and his colleagues showed this could be a reduced-form equation of their money-supply rule model where there is no structural Taylor rule at all (see also Gillman, Le and Minford (2007), and Minford (2008) as illustrated in what follows), whereas basing on a model which did include an interest-rate setting rule Cochrane suggested it could just be the interest-rate solution. It could also be some pure statistical relationship between interest rates and inflation and others. In general there is no way to tell.

One may consider the following simple model as in Minford (2008) with a money-supply rule instead of a Taylor rule to see the point more clearly:

('IS' curve): $y_t = \gamma E_{t-1} y_{t+1} - \phi r_t + v_t$

(Phillips curve): $\pi_t = \zeta(y_t - y^*) + \nu E_{t-1}\pi_{t+1} + (1-\nu)\pi_{t-1} + u_t$

(Money supply target): $\Delta m_t = m + \mu_t$

(Money demand): $m_t - p_t = \psi_1 E_{t-1} y_{t+1} - \psi_2 R_t + \varepsilon_t$

(Fisher identity): $R_t = r_t + E_{t-1}\pi_{t+1}$

This model implies a Taylor-type relationship that looks like:

$$R_t = r^* + \pi^* + \gamma\chi^{-1}(\pi_t - \pi^*) + \psi_1\chi^{-1}(y_t - y^*) + w_t,$$

where $\chi = \psi_2\gamma - \psi_1\phi$, and the error term, w_t, is both correlated with inflation and output and autocorrelated; it contains the current money supply/demand and aggregate demand shocks and also various lagged values (the change in lagged expected future inflation, interest rates, the output gap, the money demand shock, and the aggregate demand shock). This particular Taylor-type relation was created with a combination of equations—the solution of the money demand and supply curves for interest rates, the Fisher identity, and the 'IS' curve for expected future output[3]. But other Taylor-type relationships could be created with combinations of other equations, including the solution equations, generated by the model. They will all exhibit autocorrelation

[3] From the money demand and money supply equation, $\psi_2 \Delta R_t = \pi_t - m + \psi_1 \Delta E_{t-1} y_{t+1} + \Delta\varepsilon_t - \mu_t$. Substitute for $E_{t-1} y_{t+1}$ from the IS curve and then inside that for real interest rates from the Fisher identity giving $\psi_2 \Delta R_t = \pi_t - m + \psi_1 (\frac{1}{\gamma})\{\varphi(\Delta R_t - \Delta E_{t-1}\pi_{t+1}) + \Delta y_t - \Delta v_t\} + \Delta\varepsilon_t - \mu_t$; then, rearrange this as $(\psi_2 - \frac{\psi_1\varphi}{\gamma})\Delta(R_t - R^*) = (\pi_t - m) - \frac{\psi_1\varphi}{\gamma}\Delta E_{t-1}\pi_{t+1} + \frac{\psi_1}{\gamma}\Delta(y_t - y^*) - \frac{\psi_1}{\gamma}\Delta v_t + \Delta\varepsilon_t - \mu_t$, where the constants R^* and y^* have been subtracted from R_t and y_t respectively, exploiting the fact that when differenced they disappear. Finally,

$$R_t = r^* + \pi^* + \gamma\chi^{-1}(\pi_t - \pi^*) + \psi_1\chi^{-1}(y_t - y^*) + \{(R_{t-1} - R^*) - \psi_1\varphi\chi^{-1}\Delta E_{t-1}\pi_{t+1} - \psi_1\chi^{-1}(y_{t-1} - y^*) - \psi_1\chi^{-1}\Delta v_t + \gamma\chi^{-1}\Delta\varepsilon_t - \gamma\chi^{-1}\mu_t\},$$

where we have used the steady state property that $R^* = r^* + \pi^*$ and $m = \pi^*$.

and contemporaneous correlation with output and inflation, clearly of different sorts depending on the combinations used.

It follows when econometricians estimate a Taylor-type equation as regression they would never know what they are really estimating although they might believe or assume they do. They would, of course, be able to retrieve the parameters of the proposed equation. But the estimates they obtain would hardly provide any useful information for understanding the Fed's behaviour simply because the equation itself—and henceforth its estimates—could be anything, thus unidentified. It follows that existing views that are backed by 'evidence' from Taylor rule regressions on the US post-war monetary policy are in general unreliable; unless the modeller knows from other sources that the Fed is indeed pursuing a Taylor rule or he is lucky enough to have specified a Taylor-type equation that is functionally the same as what is being followed, he is doomed to be fooled by a bunch of uninformative estimates, falling a victim of identification failure[4].

All the above applies to any simple-rule modelling of monetary behaviour basing on the regression method, regardless of the policy function chosen. However, if the modeller takes the alternative of including the specified 'rule' into a complete DSGE model with rational expectations and estimates it as part of the latter using full information methods as some recent examples—Smets and Wouters (2003), Lubik and Schorfheide (2004), Onatski and Williams (2004) and Ireland (2007)—have done, he would be able to circumvent the confusion and ensure that the estimates are correctly interpreted.

This is because the rational-expectations mechanism would impose a set of over-identifying restrictions through the expectations terms which involve in principle all the model parameters, including those of the 'monetary rule'. Thus when the expectations terms are substituted for, there would in general be more reduced-form

[4] While one may argue that various announcements, proposals and reports published by the central bank directly reveal to econometricians the bank's reaction function. However, what the Fed actually does is not necessarily the same thing as what its officials and governors say it does. So these documents while illuminating can complement but cannot substitute for econometric evidence.

parameters than structural parameters to ensure that the model is over-identified. In other words if the econometrician posits a monetary rule—say, a Taylor rule—and estimates it together with the rest of the model basing on full information methods, he would retrieve its coefficients as parts of the structural parameters without any confusion. Of course he would never confuse the rule he assumes with other model equations or the linear combinations of these as the latter are effectively restricted; nor would he ever confuse it with other policy alternatives as these would change the over-identifying restrictions and therefore the appearance of the reduced-form model—a point first made by Lucas (1976) in his 'critique' against conventional evaluation of optimal policy at the time. Thus modelling the existing monetary behaviour with differing rules and comparing one to another via full model estimation is completely possible.

However, the identification problem does not go away, even when a model is over-identified in this way. This is because the decision of including a Taylor rule or other exogenous alternatives into a DSGE model is in general induced by the observation that they fit the data well in single-equation estimation. Yet since these 'monetary rules' are unidentified as regressions and they may represent some other joint behaviours of the true model and the true policy, including these into a complete model and estimating them as part of it may result in quite misleading implications about the true policy being followed due to model misspecifications. Of course the more precise the included 'rule' is, the less the model will be mis-specified and the better it will fit the data. Thus to detect such misspecification and also to find the better description of policy being followed one would need not only to estimate possible model alternatives but also to compare their fit to data by formal testing; then by ruling out those rejected by the data and picking up the best performer amongst the survivors it is possible to identify the less mis-specified monetary rule as part of the less mis-specified model. This method of identifying monetary policy will be extensively used in the following chapters. For now, we move on to the debate over the causes of the Great Moderation and see how understanding of the Fed's behaviour could have directed the implications.

1.2 The Success of Great Moderation: a matter of policy or shocks?

One practical issue related to the study of prevailing monetary policy is how it would imply about the essence of Great Moderation which refers to the general improvement of economy in terms of stability. This occurred in the US in the early 1980s when the turbulence period since the breakdown of Bretton Woods ceased to prevail; the earlier episode is known as Great Acceleration or Great Inflation.

While a number of factors might have helped to explain the Great Moderation, improved policy and improved environment (in the form of milder shocks) are the two mostly focused on. Some studies have based the investigation on the real data with the aid of time-series models and pointed to environment; others have chosen the alternative of exploiting differing theories using the DSGE framework and pointed to policy. Yet, although these approaches are quite distinct on the analytical basis and have suggested rather conflicting views on the causes, the underlying principle governing these analyses is always to first identify the policy being followed and then decompose the data variability into policy and environment based on that identification. It follows one's view on the Great Moderation is not only a matter of employed methodology but also one of believed policy.

1.2. 1 The Time-series approach to Great Moderation: a method basing on 'facts'

The time-series approach to Great Moderation emphasizes the importance of decomposing the causes of the improvement basing on the real data. Most authors following this line have employed a structural VAR to reproduce the 'facts' and concluded with the help of that that the US Moderation was induced by luck when milder shocks caused the environment to improve. For example, Stock and Watson (2002) suggested over 70% of the reduced GDP volatility since the mid-80s was moderated by the decline of shocks to productivity, commodity prices and forecast errors; Primiceri (2005) focused on the rates of inflation and unemployment and argued that the Fed was innocent to the stagflation in the 70s; instead he attributed the predicament at the time mostly to non-policy shocks. While a similar conclusion was drawn by Gambetti, Pappa and Canova (2008), Sims and Zha (2006) found in the same vein that the best data-fitting model they had was one in which variation was only assumed to the variance of structural disturbance and that the observed inflation dynamics would not be much influenced even if alteration in monetary regime was allowed for.

The structural VAR analysis of this sort have used the VAR's capacity to capture the facts and exploited its estimates across different subsamples to trace out the causes of change in the former. The logic underlying this approach is that, when actual data are fitted into a VAR, the factors that determine their variability will be fully reflected on the estimates of the coefficient matrix and the variance-covariance matrix of prediction errors; the former condenses the propagation mechanism (including the structure and monetary policy) of the economy while the latter takes into account the impact caused by exogenous disturbances. Thus, by analysing how these two matrices vary when data of comparable episodes are fitted, 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 data variability. Of course to determine the exact causes these estimates need to be identified so that the message carried by the reduced-form models can be interpreted with economic meaning. While differing implications could be drawn as a result of open choice of identification schemes, most authors following this approach (including those just cited) have found it was the change in error variability that had dominated the change in US' data variability. Thus, almost all structural VAR studies of US Moderation have pointed to good shocks as the main cause of the improvement and suggested the impact of change in monetary policy was trivial.

Clearly, by working from data to theory the structural VAR approach has the merit of ensuring that the analysis can well account for the facts. However, since identification is required for decomposing the causes of change in data basing on estimates of reduced-form models, the weakness of this method is also apparent. The problem 'lies at the very heart', as Benati and Surico (2009) have argued, is that, the supposed theoretical restrictions on which identification relies are in general incompetent in connecting the structure of a DSGE model to the structure of a VAR. This is because the choice of such restrictions is ultimately an issue of believed structure and dynamics of the economy, whose 'reasonableness' is justified at the modeller's discretion[5]. This vague and indeterminate linkage between the facts and theory has then resulted in a hole of the method from which suspicions against its implications are often aroused—a problem of identification failure analogous to what would happen to Taylor rule regressions as the last section has considered. Thus, to circumvent confusion many authors have taken the alternative of basing their analyses directly on theories abstracted by DSGE models. This method is reviewed in what follows.

1.2. 2 The DSGE approach to Great Moderation: a method basing on theories

The DSGE approach to Great Moderation is one in which argument is built directly on theory. By exploiting differing model simulations, this method aims at finding some sort of theory whose implied dynamics could best mimic what is observed in reality. The focus is on how changes in the propagation mechanism, especially the monetary regime, of the model would affect the dynamic behaviour of the economy. Most authors following this approach have adopted the New Keynesian three-equation framework that consists of an 'IS' curve derived from the household's optimization problem, a Phillips curve derived from the firm's optimal price-setting decision and a monetary rule assumed pursued by the Fed—this last normally being a Taylor rule. Some have used a full DSGE model in which all micro-foundations are reserved for greater accuracy.

Basing on the mimicry of counterfactual experiments, this approach generally suggests that the Great Moderation in US was largely a result of improved policy instead of luck; examples of this are Clarida, Gali and Gertler (2000), Lubik and Schorfheide (2004), Boivin and Giannoni (2006) and Benati and Surico (2009)[6]. The contrast was made between the differing dynamics driven by a 'passive' policy

[5] This structure, of course, includes the monetary regime in place, and this is why a correct understanding of the existing monetary policy is so important in understanding the true causes of Great Moderation in structural VAR analyses.

[6] Ireland (2007) is one exception. While arguing that the US economy improved as a result of improved policy, Ireland has defined such an improvement as having a more conservative (implicit) inflation target that is driven by several structural shocks. His argument has then effectively pointed to the environment as the determinant of the Moderation.

assumed prevailed in the 1970s, and an 'active' policy of the later period when the Fed's response to inflation was supposed stronger—this usually refers to the satisfaction of the Taylor Principle (that ensures the existence of a unique stable equilibrium); in general it requires that the interest rates' response to inflation in the Taylor rule be greater than unity.

Thus a consensus theory advocated by these authors claims that the passive interest-rates response the Fed adopted in the 1970s had led the US economy into a region of indeterminacy, within which 'sunspot fluctuations' arose and appeared as the Great Acceleration as a result of (the private agents') self-fulfilling behaviour; these sunspots then ceased in the early 80s when the Fed became more active in combating inflation, for which self-fulfillingness was effectively suppressed and the economy moved into a region of determinacy, terminating the turbulence and giving the rise of the Great Moderation.

The cornerstone of the above explanation is the assumption that the Fed was pursuing a passive Taylor-type policy in the 1970s and then switched to another qualitatively more active thereafter. One recent empirical justification of this is Lubik and Schorfheide (2004), who, building on the Bayesian approach, compared the conditional probabilities of the US economy being respectively in a determinacy region and in an indeterminacy region with differing 'structural' priors. For all cases assumed they found the posterior probability suggested that the US economy was indeterminate before 1980 but was determinate afterwards. While they have allowed all parameters to change, it was rather consistent that their posterior estimates implied a passive pre-1980 Taylor rule and an active post-1980 rule. Clearly, another more conventional way of justifying this is to use 'evidence' from Taylor rule regressions directly estimated over the two subsamples.

However, while the New Keynesian-Taylor rule approach has presented a possible channel through which the transition from Great Acceleration to Great Moderation could have occurred, a crucial challenge to this explanation is the extent to which the proposed model has resembled the truth. Especially, since a Taylor rule regression is generally unidentified in single-equation estimation, including the estimate of it into a DSGE model and seeing it as the underlying policy may result in model misspecification that leads to quite distorted implications (again, a point discussed earlier in the last section). It follows that to use theory as such to explain the Great Moderation one has to ensure that the model he specifies is not strongly rejected by the data. This, too, requires formal testing of the whole model, including the supposed policy, for the selected data sample, just as what would be needed for identifying the underlying policy. Unfortunately, by focusing on the mimicry of counterfactual experiments this point seems to have been much overlooked by the existing studies.

1.2. 3 Uncovering the true causes of the Great Moderation: some notes on the desired methodology and the role of monetary policy

The above has suggested that the choice of decomposition methodology and one's understanding of the existing monetary policy are vital in implying the essence of the Great Moderation. The structural VAR approach builds on the facts reproduced by a time-series model and works from data to theory; it points to the environment. The New Keynesian-Taylor rule alternative exploits the mimicry of DSGE model simulations, with which it harmonizes theory with the facts; it points to the policy. This choice of analytical basis has determined that the two methods both bear clear strength and limitation: the former, while ensuring that the facts are well accounted for, has failed to conquer the pervasive problem of identification when it connects the facts to theory; the latter, where explanations are well justified by theory, is by itself incompetent in ensuring that they are too justified by the facts. Since causality can only be established by theory, which is valid only if it fits the facts, a fully satisfactory evaluation would require one to build on a DSGE model whose time-series properties are also compatible with the data so that the analysis is founded both in theory and in facts; it also suggests the two conventional methods just reviewed are effectively complements to each other for more comprehensive evaluations.

Another concealed issue pervaded the preceding is the subtle connection between one's understanding about the Fed and his understanding about the Great Moderation. Existing literature seems to have taken little heed of this point; but what is implied for the latter is generally developed upon what is presumed for the former. In particular, the structural VAR approach decomposes the causes of change in data variability using theoretical restrictions including those suggested by the supposed monetary policy; the New Keynesian-Taylor rule alternative, however, for reproducing the resemblance in counterfactual experiments, has imposed that the Fed was pursuing a Taylor rule. The fact that the presumed monetary behaviour is directive in pointing the causes of the Great Moderation has determined that ensuring the truthfulness of the former is a matter of non-trivial importance. Thus a better description of monetary policy (whose representation is not rejected by the data) must be found. This goes back to what was discussed earlier in section 1.1. Indeed, since monetary policy is part of the structure of a DSGE model, this point was also implicitly advocated in the last concern about the desired methodology.

1.3 To Understand the Fed and the Great Moderation Better—what is the way forward?

The foregone thus suggests that understanding the Fed's monetary behaviour and understanding the causes of the Great Moderation is virtually an identical problem of different scale. This is because investigation of the latter must be built upon a model description of the economy that spontaneously suggests the former; the problem in the heart is whether a model that is least rejected by the data (and hence closer to the truth) can be found. In other words it reduces to one of testing and ranking potential model candidates.

Thus, on uncovering the Fed's monetary behaviour one could include into a DSGE model with differing policies to construct competing model versions and test which of these, with its over-identifying restrictions, could best mimic the features shown by the data. For example, one could build on the New Keynesian three-equation framework and feed the same system of 'IS' and Phillips curves with differing Taylor rules or other kinds of policy to detect which is the best description of the Fed's behaviour. The advantage of fixing the demand and supply equations is that it ensures all the difference between competing models are caused singly by the assumed policies so that a better description of the Fed's behaviour could be identified simply by ranking the models' fit to data. Any regime version that one considers possible could in principle be included and tested in the same way. Of course, one could also adopt other demand/supply structure and test for the Fed's policy based on that; but he should ensure that differing policies are evaluated on the same basis to preserve comparability.

This method can also be extended easily to chasing the causes of the Great Moderation since it is essentially a way of uncovering the better model description of the economy. Yet in this context one would be required to test and rank competing model versions in both episodes to find the best representation of each. In particular, unless structural variations are allowed for, one should ensure that the baseline framework he uses is also consistent across samples such that once the least rejected models for both episodes are found he can decompose the data variability purely between policy and environment. But this restriction could be released when one considers there is a need for taking int o account the possible impact caused by 'structural breaks'.

1.4 Conclusion

In this chapter we have reviewed the recent literature in modelling the US post-war monetary policy and the causes of the Great Moderation; the focus is on the evaluation methods of these and the connection between understanding of the former and implication of the latter. As it stands the main difficulty pervading the research along these lines is the problem of identification that typically occurs in the estimation of existing monetary policy using Taylor rule regressions and in decomposing the causes of the Great Moderation basing on structural VARs. The problem also spreads to the alternative method of building the analysis on a complete DSGE model abstracted directly from the theory via the inclusion of the supposed monetary policy. However, since DSG E models are helpf ul both in over-identifying the supposed policy and in establishing causality of data variability, building the analysis on these is an option that kills two birds with one stone. But one must ensure that the model he has chosen is not rejected the data, thus also accounts for the facts. This reduces the problem to one of testing and ranking competing DSGE models with differing monetary policies. Once the best representation of the economy is found, the Fed's underlying behaviour would be readily identified as part of the least rejected model. One can also develop this further to decompose the causes of change in data variability while ensuring the analysis is founded both in theory and in facts. This exercise is performed in the chapters to follow.

Chapter 2

Taylor Rule or Optimal Timeless Policy? Reconsidering the Fed's monetary behaviour since the early 1980s

Introduction

The last chapter has suggested that to circumvent the pervasive problem of identification encountered the conventional method of modelling monetary policy one has to evaluate possible candidates in the context of a complete DSGE model and compare their performance via formal testing against the data. This chapter takes the Moderation episode as an example to demonstrate how this method is used. In particular, it reconsiders the Fed's monetary behaviour since the early 1980s by comparing several popular regime versions that earlier efforts are mostly concerned; these being the optimal timeless policy, the original Taylor rule and its 'interest-rates-smoothed' version. These candidates are fitted into a simple New Keynesian model with standard calibrations and distinguished according to their capacity in replicating the dynamics and size of the data based on the method of indirect inference. The exercise suggests the only model version that fails to be strongly rejected is the optimal timeless policy. This version is also shown to account for the widespread finding of apparent 'Taylor rules' and 'interest rates smoothing' in the data, even though neither represents the Fed's true policy. This last illustrates the identification problem of Taylor rule estimation as pointed in chapter 1.

The rest of this chapter is organised as follows: section 2.1 outlines the baseline New Keynesian model and the rules to be tested; section 2.2 explains the method of testing using indirect inference and sets out the finding that the Fed pursued an optimal timeless policy; section 2.3 discusses how this policy can be misinterpreted as Taylor rules under the perspective of single-equation estimation; section 2.4 concludes.

2.1 A Simple New Keynesian Model for Interest Rates, Output Gap and Inflation Determination

A common practice among New Keynesian authors in monetary policy analysis is to set up a full DSGE model with representative agents and reduce it to a three-equation framework that consists of an 'IS' curve, a Phillips curve and a monetary policy rule

(Clarida, Gali and Gertler (1999, 2000), Rotemberg and Woodford (1997, 1998), Walsh (2003)). This is also the approach taken here.

Under rational expectations, the 'IS' curve derived from the household's optimisation problem and the Phillips curve derived from the firm's optimal price-setting behaviour given Calvo (1983) contract can be shown as:

$$x_t = E_t x_{t+1} - (\frac{1}{\sigma})(\tilde{i}_t - E_t \pi_{t+1}) + v_t \qquad [2.1]$$

$$\pi_t = \beta E_t \pi_{t+1} + \gamma x_t + \kappa u_t^w \qquad [2.2]$$

where x_t is the output gap, $\tilde{i}_t$ is the deviation of interest rates from its steady-state value, π_t is the price inflation, and v_t and u_t^w are interpreted as 'demand disturbance' and 'supply disturbance', respectively[7].

This model can be closed by adding to it a monetary policy equation; this normally being a Taylor rule in the New Keynesian literature, but other policy alternatives is completely possible. This chapter selects the three popular regime versions widely suggested for the US economy; these are the optimal timeless policy when the Fed commits to minimize a typical social welfare loss function, the original Taylor rule [1.1], and its 'interest-rates-smoothed' version [1.2].

Many normative monetary policy studies in the literature focusing on Taylor rules have argued that policies of the sort are roughly 'optimal'. Here, the optimal timeless rule is introduced and compared to these as by pursuing the timeless optimum the Fed is assumed acting more precisely an optimizing role. This optimal timeless rule is derived by minimizing the typical quadratic social welfare loss function $SWL_t = \frac{\psi}{2}[\alpha x_t^2 + \pi_t^2]$ with respective to the constraint of the economy summarised by the Phillips curve [2.2]. Following the idea of Woodford (1999) of ignoring the initial conditions confronting the Fed at the regime's inception, it shows, if the Fed

[7] γ and κ are functions of other structural parameters and some steady-state relations—see table 2.3 for calibrations in the next section. Full derivation of the baseline model is available in Supporting Annex.

was a strict, consistent optimizer, it would have kept inflation always equal to a fixed fraction of the first difference of the output gap such that

$$\pi_t = -\frac{\alpha}{\gamma}(x_t - x_{t-1}) \qquad [2.3]$$

where α is the relative weight it put on loss from output variations against inflation variations, γ reflects the Phillips curve constraint (regarding stickiness) it faced; a full derivation of this is shown in the Supporting Annex.

Note that this optimal timeless rule, compared to Taylor rules of the standard sort, is implicit. Unlike typical Taylor rules that specify systematic policy instrument response to economic conditions, the optimal timeless rule distinguishes itself by setting an 'optimal trade-off' between the economic outcomes—here, excess inflation is 'punished' by a fall in the output growth rate. Svensson and Woodford (2004) categorized this kind of 'targeting rule' as 'high-level monetary policy'; they argued that by connecting the central bank's monetary principle to its ultimate policy objectives this rule has the advantage of being more transparent and robust, although compared to the 'low-level' instrument rules—like Taylor rules—it has more difficulty in ensuring determinacy (a diagrammatic illustration of the working of the timeless rule model is shown in the next chapter; a comparison of this with Taylor rule based on model impulse responses is revealed in chapter 4).

Thus, implementation of the optimal timeless rule in practice would require the Fed to fully understand the model (including the shocks hitting the economy) and set the policy instrument (the Fed rates e.g.) to whatever supports such an optimal trade-off. Yet in practice it would be reasonable to allow for the Fed to make a 'trembling hand' managerial error ξ_t in execution, like any policy mistake it would make in pursuit of typical Taylor rules[8].

---

[8] However, given that the connotation of the optimal timeless policy differs substantially from that of typical Taylor-type policies, interpretations of such policy error in these two different cases are slightly different.

We have said in chapter 1 (pp.7) that the trembling-hand error in a Taylor rule could reflect the Fed's occasional discretionary policy response to something unaccounted in the announced rule or its pure managerial error in execution of such policy. Under the optimal timeless rule this trembling-hand error has a broader sense. The main difference is in that in pursuit of the optimal timeless policy the Fed is assumed to first solve the model and then set the Fed rates to whatever is required for realization of the optimal trade-off between inflation and output growth shown by [2.3]. 'Policy error' in this case, Thus the three pseudo economies with differing monetary policies are readily constructed and comparable. These are summarised in table 2.1[9]:

Table 2.1: Models to be tested

Model one	(with optimal timeless policy)	
'IS' curve	$x_t = E_t x_{t+1} - (\frac{1}{\sigma})(\tilde{i}_t - E_t \pi_{t+1}) + v_t$	$v_t = \rho_v v_{t-1} + \varepsilon_t^v$
Phillips curve	$\pi_t = \beta E_t \pi_{t+1} + \gamma x_t + \kappa u_t^w$	$u_t^w = \rho_{u^w} u_{t-1}^w + \varepsilon_t^{u^w}$
Policy rule	$\pi_t = -\frac{\alpha}{\gamma}(x_t - x_{t-1}) + \xi_t$	$\xi_t = \rho_\xi \xi_{t-1} + \varepsilon_t^\xi$
Model two	(with the original Taylor rule)	
'IS' curve	$x_t = E_t x_{t+1} - (\frac{1}{\sigma})(\tilde{i}_t - E_t \pi_{t+1}) + v_t$	$v_t = \rho_v v_{t-1} + \varepsilon_t^v$
Phillips curve	$\pi_t = \beta E_t \pi_{t+1} + \gamma x_t + \kappa u_t^w$	$u_t^w = \rho_{u^w} u_{t-1}^w + \varepsilon_t^{u^w}$
Policy rule	$i_t^A = \pi_t^A + 0.5x_t + 0.5(\pi_t^A - 0.02) + 0.02 + \xi_t$	
The transformed policy rule	$\tilde{i}_t = 1.5\pi_t + 0.125x_t + \xi_t'$	$\xi_t' = \rho_\xi \xi_{t-1}' + \varepsilon_t^\xi$
Model three	(with 'interest-rates-smoothed' Taylor rule [1.2])	
'IS' curve	$x_t = E_t x_{t+1} - (\frac{1}{\sigma})(\tilde{i}_t - E_t \pi_{t+1}) + v_t$	$v_t = \rho_v v_{t-1} + \varepsilon_t^v$
Phillips curve	$\pi_t = \beta E_t \pi_{t+1} + \gamma x_t + \kappa u_t^w$	$u_t^w = \rho_{u^w} u_{t-1}^w + \varepsilon_t^{u^w}$
Policy rule	$i_t^A = (1-\rho)[\alpha + \gamma_\pi(\pi_t^A - \pi^*) + \gamma_X x_t] + \rho i_{t-1}^A + \xi_t$	
The transformed policy rule	$\tilde{i}_t = (1-\rho)[\gamma_\pi \pi_t + \gamma_x' x_t] + \rho \tilde{i}_{t-1} + \xi_t'$	$\xi_t' = \rho_\xi \xi_{t-1}' + \varepsilon_t^\xi$

besides reflecting the disturbances just mentioned, thus means more broadly to include the Fed's general failure in getting the model and hence the required Fed rates correctly solved. In practice this can mean the Fed's imperfect understanding of the model or in the case where the model is well understood its inability to correctly identify and react to the demand and supply shocks.

[9] Note it has assumed an AR(1) process for all disturbances to the structural equations to capture possible omitted variables. It also transforms the Taylor rules to quarterly versions so that the frequency of interest rates and inflation is consistent with other variables in the model—all constants are dropped as demeaned, detrended data will be used (See 'data' part in section 2.2.2 to follow).

Note that these models are different solely in the policies being followed. Hence by comparing their capacity to fit the real data, one should be able to tell which rule, when included in a simple New Keynesian framework, provides the best explanation of the facts and therefore the most appropriate description of the underlying policy. This exercise is performed in section 2.2 in what follows.

2.2 Confronting the Models with Facts

2.2. 1 The method of indirect inference

To evaluate the models' performance in fitting the real data this thesis uses the method of indirect inference proposed in Minford, Theodoridis and Meenagh (2009)[10]. The approach involves using an auxiliary model that is completely independent of the theoretical one to produce a description of the data against which the performance of the theory is evaluated indirectly. Such a description can be summarised either by the estimated parameters of the auxiliary model or by functions of these; these are called the descriptors of the data and are treated as the 'reality'; the theoretical model being evaluated is then simulated to find its implied values for these.

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)). Yet here a different use of indirect inference is made as our aim is to evaluate models that are already calibrated. The common element is the use of an auxiliary time series model. In estimation the parameters of the structural model are chosen such that when this model is simulated it generates estimates of the auxiliary model similar to those obtained from the actual data. The optimal choices of parameters for the structural model are those that minimise the distance between a given function of the two sets of estimated coefficients of the auxiliary model. Common choices of this function are the actual coefficients, the

[10] See Meenagh, Minford and Wickens (2009) and Le, et al. (2009, 2011) for more applications of this approach. Le et al. (2011) deals with a wide range of practical issues raised by this approach.

scores or the impulse response functions. In model evaluation the parameters of the structural model are taken as given. The aim is to compare the performance of the auxiliary model estimated on simulated data derived from the given estimates of a structural model—which is taken as a true model of the economy, the null hypothesis—with the performance of the auxiliary model when estimated from the actual data. If the structural model is correct then its predictions about the impulse responses, moments and time series properties of the data should statistically match those based on the actual data. The comparison is based on the distributions of the two sets of parameter estimates of the auxiliary model, or of functions of these estimates.

The testing procedure thus involves first constructing the errors implied by the previously estimated/calibrated structural model and the data. These are called the structural errors and are backed out directly from the equations and the data[11]. These errors are then bootstrapped and used to generate for each bootstrap new data based on the structural model. An auxiliary time series model is then fitted to each set of data and the sampling distribution of the coefficients of the auxiliary time series model is obtained from these estimates of the auxiliary model. A Wald statistic is then computed to determine whether functions of the parameters of the time series model estimated on the actual data lie in some confidence interval implied by this sampling distribution.

Following Minford, Theodoridis and Meenagh (2009), this thesis takes a VAR(1) for the three macro variables (interest rates, output gap and inflation) as the appropriate auxiliary model and treats as the descriptors of the data the nine VAR(1) coefficients and the three variances of the involved variables. The Wald statistic is computed from these[12]. Thus effectively it is testing whether the observed dynamics and volatility of the chosen variables are explained by the simulated joint distribution of these at a given confidence level. The Wald statistic is given by:

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

[12] Note that the VAR impulse response functions, the co-variances, as well as the auto/cross correlations of the left-hand-side variables will all be implicitly examined when the VAR coefficient matrix is considered, since the former are functions of the latter.

$$(\Phi-\overline{\Phi})'\Sigma_{(\Phi\Phi)}{}^{-1}(\Phi-\overline{\Phi}) \qquad [2.4]$$

where Φ is the vector of VAR estimates of the chosen data descriptors just mentioned, with $\overline{\Phi}$ and $\Sigma_{(\Phi\Phi)}$ representing, respectively, the mean and variance-covariance matrix of these implied by bootstrap simulations[13]. This whole test procedure can be illustrated diagrammatically in Figure 2.1 as follows:

Figure 2.1: The Principle of Testing using Indirect Inference

Panel A:

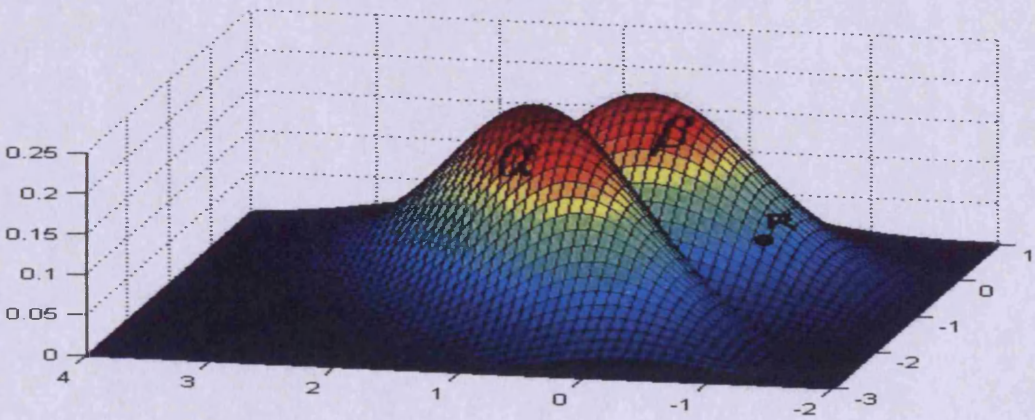
Model(s) to be tested

↓ (Bootstrap simulations)

Actual data		Simulated data
↓		↓
VAR representation		VAR representation
↓		↓
The VAR inference (the 'reality')	*vs.*	Distribution(s) of the VAR inference

Wald statistic

Panel B:



[13] Smith (1993), for his demonstration of model estimation, originally used VAR(2) as the auxiliary model. His VAR included the logged output and the logged investment and he tried to maximize the model's capacity in fitting the dynamic relation between these. To this end he included the ten VAR coefficients (including two constants) in his vector of data descriptors. Here, since a VAR(1) is chosen to provide a parsimonious description of the data and the models are tested against their capacity in fitting the data's dynamic relations *and* size, the vector of chosen data descriptors includes nine VAR(1) coefficients and the three data variances. No constant is included since demeaned de-trended data are used (See 'data' part in 2.2.2 below). Chapter 4 (4.3.1), for checking robustness, also extends the exercise to one in which differing orders of VAR are tried. It turns out that the choice of VAR order in the context here is really merely a matter of setting the test's rejection power.

While the first panel in Figure 2.1 summarises the main steps of the methodology just described, the 'mountain-shaped' diagram in panel B gives an example of how the 'reality' is compared to model predictions using the Wald test when two parameters of the auxiliary model are considered. Suppose the real-data estimates of these are given at R and there are two models to be tested, each implies a joint distribution of these parameters shown by the 'mountains' (α and β). Since R lies outside the 95% contour of α, it would reject this model at 95% confidence level; the other model that generated β would not be rejected, however, since R lies inside. In practice there are usually more than two parameters to be considered and henceforth the test is carried out by the Wald statistic of [2.4].

The joint distribution mentioned above is a bootstrap distribution simulated from bootstrapping the innovations implied by the data and the theoretical model and it is therefore an estimate of the small sample distribution[14]. Such a distribution is generally more accurate for small samples than the asymptotic distribution; it is also shown to be consistent by Le, et al. (2011) given that the Wald statistic is 'asymptotically pivotal'; they also showed it had quite good accuracy in small sample Montecarlo experiments[15].

2.2.2 Data and calibration

Data

To test the Fed's monetary policy in the Great Moderation this chapter employs the quarterly data published by the Federal Reserve Bank of St. Louis from 1982 to 2007[16]. Most discussions of the Fed's behaviour (especially those based on Taylor rules) are concerned with the periods that begin sometime in the 1980s but here 1982

[14] The bootstraps in the tests here are all drawn as time vectors so that any contemporaneous correlation between the innovations will be preserved.

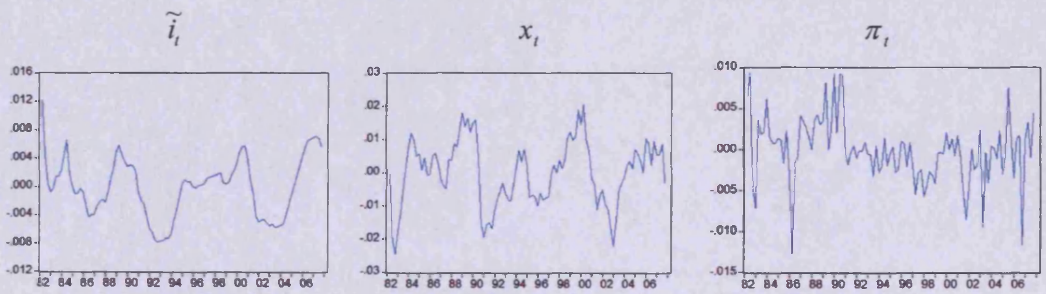
[15] Specifically, they found that the bias due to bootstrapping was just over 2% at the 95% confidence level and 0.6% at the 99% level.

[16] Data base of Federal Bank of St. Louis: http://research.stlouisfed.org/fred2/

is chosen as the starting point because many (including Bernanke and Mihov, 1998, and Clarida, Gali and Gertler, 2000) have argued that it was around then that the Fed switched from using non-borrowed reserves to setting the Funds rate as the instrument of monetary policy. Taylor (1993) originally suggested a later starting point for his specification and plainly one could choose a variety of different sample periods and test for that; a robustness check regarding this point is deferred to chapter 4.

The tests measure $\tilde{i}_t$ as the deviation of current Fed rate from the steady-state value which is interpreted here as a linear trend (at a quarterly rate for compatibility with the quarterly inflation rate); output gap x_t is approximated by the percentage deviation of real GDP from its HP-trend value[17]; quarterly inflation π_t is defined as the log difference between current CPI and the one captured in the last quarter. The data are also demeaned for simplicity. These are plotted in figure 2.2; the unit root test results are reported in table 2.2.

Figure 2.2: Demeaned-detrended Data of Interest Rates, Output Gap and Inflation



[17] Note by defining the output gap as the HP-filtered log output it has 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 the three-equation representation here, one 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 lying well beyond the scope of this thesis, though something worth pursuing in future work.

Le et al. (2011) test the Smets and Wouters (2003) US model by the same methods as used here. This has a Taylor rule that responds to flexible-price output. It is also close to the timeless optimum rule 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 flexprice 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.

Table 2.2: Unit Root Test for Stationarity

Time series	5% critical value	ADF test statistics	p-values*
$\tilde{i}_t$	-1.94	-2.8	0.0053
x_t	-1.94	-2.95	0.0035
π_t	-1.94	-3.60	0.0004

Note: '*' denotes the Mackinnon (1996) one-sided p-values; H_0: the time series has a unit root.

Since the data are mean-deviations, a VAR(1) representation of them contains no constant but only nine parameters in the autoregressive coefficient matrix. Also, the use of such data requires dropping the constants in any equation of the models as well. This explains why the two transformed Taylor rules involved in model two and three have no constant at all.

Calibration

The values of model parameters chosen are those commonly calibrated and accepted for the US economy in the literature. These are listed in table 2.3 as follows:

Table 2.3: Calibration of Model Parameters

β	time discount factor	0.99	
σ	inverse of elasticity of intertemporal consumption	2	
η	inverse of elasticity of labour	3	
ω	Calvo contract price non-adjusting probability	0.53	
G/Y	steady-state government expenditure to output ratio	0.23	
Y/C	steady-state output to consumption ratio	1/0.77	(implied)
κ	$\kappa = \frac{(1-\omega)(1-\omega\beta)}{\omega}$	0.42	(implied)
γ	$\gamma = \kappa(\eta + \sigma\frac{Y}{C})$	2.36	(implied)
θ	price elasticity of demand	6	

$\alpha/\gamma \equiv \theta^{-1}$	parameter driving the optimal timeless policy[18]	1/6	(implied)
ρ	degree of interest rates smoothness	0.76	
γ_π	interest rates response to inflation	1.44	
$\gamma_x^{'}$	interest rates response to output gap	0.14	
ρ_v	autoregressive coefficient of demand disturbance	0.91	(sample estimate)
ρ_{u^w}	autoregressive coefficient of supply disturbance	0.82	(sample estimate)
ρ_ξ	autoregressive coefficient of policy disturbance: model one	0.35	(sample estimate)
ρ_ξ	autoregressive coefficient of policy disturbance: model two	0.37	(sample estimate)
ρ_ξ	autoregressive coefficient of policy disturbance: model three	0.31	(sample estimate)

As table 2.3 shows, the quarterly time discount rate is calibrated as 0.99, implying an approximately 1% quarterly (or equivalently 4% annual) rate of interest in steady state. σ and η are set to as high as 2 and 3 respectively as in Carlstrom and Fuerst (2008), who emphasized on the values' consistency with the inelasticity of both intertemporal consumption decision and labour supply shown by the US data. The Calvo price stickiness of 0.53 and the price elasticity of demand of 6 are both taken from Kuester, Muller and Stolting (2009). These values imply a contract length of more than three quarters[19], while the constant mark-up of price to nominal marginal cost is 1.2. The implied steady-state output-consumption ratio of 1/0.77 is calculated based on the steady-state ratio of government expenditure over output of 0.23 calibrated in Foley and Taylor (2004). The Taylor rule tested in model three follows the calibration in Carlstrom and Fuerst (2008), where the interest rates' response to a unit change in inflation and output gap are 1.44 and 0.14 respectively, with the degree of 'smoothness' of 0.76. The last five lines in table 2.3 also report the autoregressive coefficients of the structural errors implied by the models, which are all sample estimates based on the real data[20]. Notice that both of the demand and supply shocks

[18] Nistico (2007) showed the relative weight α is equal to the ratio of the slope of the Phillips curve to the price elasticity of demand, namely, $\alpha = \gamma/\theta$.

[19] To be accurate, $2(1-\omega)^{-1} - 1 \approx 3.26$.

[20] These estimates are all significant at 5% significance level.

are shown to be highly persistent, in contrast to the policy shocks reflected in all the three models.

2.2. 3 Evaluating the models' performance—the test results

This section presents the test results for the three models considered; these are based on the VAR parameters and the data variances. Since there are three endogenous variables, the VAR(1) representation generates twelve components: the nine VAR coefficients and the three variances[21]. The tests calculate two kinds of Wald statistic (called 'directed' Wald statistics) according to the aspects of the data the models are asked to fit: here the dynamics and the volatility of the data. Another 'full' Wald statistic where the two properties are simultaneously considered is also calculated to measure the models' overall performance. In both cases the Wald statistic is reported as a percentile, i.e. the percentage point where the data value comes in the bootstrap distribution. The test results in detail are as follows:

Model one (with optimal timeless policy)

Table 2.4 below summarises the test results regarding the dynamic properties of model one:

Table 2.4: Individual VAR Coefficients and the Directed Wald Statistic

VAR(1) Coefficients	95% lower bound	95% upper bound	Values estimated with real data	In/Out
β_{11}	0.6454	0.9420	0.8017	In
β_{12}	- 0.0844	0.0439	0.0834	**Out**
β_{13}	- 0.1774	0.0991	0.0112	In
β_{21}	- 0.2589	0.2578	- 0.2711	**Out**
β_{22}	0.6685	0.9105	0.9009	In

[21] The VAR(1) is assumed as follows: $\begin{bmatrix} \tilde{i}_t \\ x_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{31} & \beta_{32} & \beta_{33} \end{bmatrix} \begin{bmatrix} \tilde{i}_{t-1} \\ x_{t-1} \\ \pi_{t-1} \end{bmatrix} + \Sigma_t$

β_{23}	- 0.4037	0.1871	- 0.1090	In
β_{31}	- 0.1821	0.1595	-0.0187	In
β_{32}	- 0.0434	0.1361	0.1428	**Out**
β_{33}	0.1010	0.4976	0.2552	In
Directed Wald percentile (for dynamics)			98.2	

According to table 2.4, three out of the nine real-data-based estimates of the VAR coefficients that reflect the actual dynamics are found to lie outside their corresponding 95% bounds implied by the theoretical model. Specifically, the response of interest rates to the lagged output gap and the response of output gap to the lagged interest rates, as well as the response of inflation to the lagged output gap, are all shown to be more aggressive than what the theoretical model would predict. In particular, the interest rates' response to the lagged output gap in reality is more than twice as great as what could be generated from model simulations. Overall, the directed Wald statistic is reported as 98.2; this indicates the model's success in capturing the actual dynamics at the 99% confidence level, although it clearly fails at the more conventional 95% level. Clearly, all the DSGE models here have problems fitting the data closely; yet the main purpose here is to rank these and to see if one of these stands out as relatively acceptable.

Turning to the volatility of the data, table 2.5 below shows the extent to which this is explained by the theoretical model:

Table 2.5: Volatility of the Endogenous Variables and the Directed Wald Statistic

Volatility of the endogenous variables	95% lower bound	95% upper bound	Values calculated with real data	In/Out
$\mathrm{var}(\tilde{i})$	0.0102	0.0450	0.0171	In
$\mathrm{var}(x)$	0.0411	0.1601	0.0951	In
$\mathrm{var}(\pi)$	0.0094	0.0206	0.0153	In
Directed Wald percentile (for volatility)			10.4	

Note: Values reported in table 2.5 are magnified by 1000 times as their original values.

As table 2.5 shows, not only are all three variances within their individual 95% bounds but also the directed Wald percentile is 10.4. That is, at the confidence level of 95%, the observed volatility is not only individually, but also jointly explained by the theoretical model—with such a low Wald statistic, they are very close to the joint means of the variances.

Note that by using the directed Wald the above have been examining the theoretical model's *partial* capacities in explaining the data. To evaluate the model's *overall* performance, however, the full Wald statistic needs to be calculated. This is reported in table 2.6 as 96.5; hence the null hypothesis that the theoretical model explains both the actual dynamics and volatilities is easily accepted at the 99% confidence level and only *marginally* rejected at 95%.

Table 2.6: The Full Wald Statistic

The concerned model properties	Full Wald percentile
Dynamics + Volatility	96.5

To summarise, model one does not only provide a rough explanation for the actual dynamics, but also precisely captures the volatility shown by the real data; its overall fitness in explaining the data is fairly good as DSGE models go and we may consider it as a reasonable approximation to the real-world economy.

Model two (with the original Taylor rule)

Leaving the economic environment (i.e., the 'IS' and Phillips curves) unchanged, model two replaces the optimal timeless rule assumed in model one with the original Taylor rule, widely regarded as a good description of the Fed's monetary policy from the late 1980s until at least the early 1990s. The rule's performance in mimicking the real dynamics for our sample chosen is reported as follows:

Table 2.7: Individual VAR Coefficients and the Directed Wald Statistic

VAR(1) Coefficients	95% lower bound	95% upper bound	Values estimated with real data	In/Out
β_{11}	0.6139	1.1165	0.8017	In
β_{12}	- 0.0743	0.2385	0.0834	In
β_{13}	- 0.3098	0.2977	0.0112	In
β_{21}	- 0.1571	0.3175	- 0.2711	**Out**
β_{22}	0.6112	0.8960	0.9009	**Out**
β_{23}	- 0.4316	0.1654	- 0.1090	In
β_{31}	- 0.1055	0.6202	-0.0187	In
β_{32}	- 0.1457	0.1983	0.1428	In
β_{33}	-0.0043	0.6596	0.2552	In
Directed Wald percentile (for dynamics)				100

Table 2.7 reveals that, while most of the real-data-based estimates of the VAR coefficients are individually captured by the 95% bounds implied by model simulations, the output gap's responses to the lagged interest rates and to its own lagged value are found to exceed their corresponding lower bound and upper bound, respectively. Overall, the directed Wald statistic is reported as 100, suggesting there is no hope at all for the theoretical model to generate a joint distribution of the VAR coefficients that simultaneously explains the ones observed in reality. The theoretical model thus is totally rejected by the Wald test for the dynamics.

Yet the model can still explain the data volatility reasonably well, as shown in table 2.8. It generates slight excessive interest rates and inflation variances, but ideally matches series the variance of the output gap. The directed Wald statistic for the variances is 91.5, comfortably accepted therefore at 95%.

Table 2.8: Volatility of the Endogenous Variables and the Directed Wald Statistic

Volatility of the endogenous variables	95% lower bound	95% upper bound	Values calculated with real data	In/Out
$var(\tilde{i})$	0.0604	0.2790	0.0171	**Out**
$var(x)$	0.0400	0.1527	0.0951	In
$var(\pi)$	0.0475	0.1672	0.0153	**Out**
Directed Wald percentile (for volatility)				91.5

Note: Values reported in table 2.8 are magnified by 1000 times as their original values.

Lastly, table 2.9 shows the full Wald statistic as 100. This is hardly surprising since it fails so badly to capture the dynamics of the data.

Table 2.9: The Full Wald Statistic

The concerned model properties	Full Wald percentile
Dynamics + Volatility	100

Thus the results above suggest model two, where the original Taylor rule is set as the fundamental monetary policy, has only partially captured the characteristics shown by the data; unless the discussions are focused exclusively on the 'size' of the economy's fluctuations, such a model is not to be taken as a realistic description of the prevailing economic reality.

Model three (with 'interest-rates-smoothed' Taylor rule [1.2])

In this last model a calibrated Taylor rule version whose specification reflects the feature of 'interest rates smoothing' is assumed to be the underlying policy reaction function. This rule suggests to set interest rates as a weighted average of what was set in the last period and what would be required had the original Taylor rule been followed, with the weights being the degree of 'policy inertia' and its complement, respectively. While 'interest-rates-smoothed' Taylor rules of this sort are commonly claimed to be supported by empirical evidence as representing the Fed's underlying policy (e.g., Clarida, Gali and Gertler (1999, 2000), Rotemberg and Woodford (1997, 1998)), the test results of this model version are revealed as follows:

Table 2.10: Individual VAR Coefficients and the Directed Wald Statistic

VAR(1) Coefficients	95% lower bound	95% upper bound	Values estimated with real data	In/Out
β_{11}	0.7228	0.9470	0.8017	In
β_{12}	- 0.0168	0.1287	0.0834	In
β_{13}	- 0.0029	0.1553	0.0112	In
β_{21}	- 0.1424	0.2095	- 0.2711	**Out**
β_{22}	0.6551	0.8971	0.9009	**Out**
β_{23}	- 0.2840	-0.0046	- 0.1090	In
β_{31}	- 0.1668	0.4706	-0.0187	In
β_{32}	- 0.1260	0.2655	0.1428	In
β_{33}	0.0830	0.5427	0.2552	In
Directed Wald percentile (for dynamics)				99.9

Table 2.10 summarises how the actual dynamics are explained by the theoretical model. Again, except for the output gap's responses to the lagged interest rates and to its own lagged value, all dynamic relationships shown by the real data are individually captured by the simulated 95% bounds. Yet, the directed Wald statistic reported is as high as 99.9, indicating that the theoretical model can hardly be used for explaining the observed dynamics, as the set of real-data-based estimates of the VAR coefficients is not captured by the joint distribution of these across model simulations, even at a 99% confidence level[22].

Turning to the data volatility, table 2.11 shows the model has merely correctly mimicked the variance of the output gap but has evoked too much for both the interest

[22] These are rather similar to the results for model two.

rates and inflation; the directed Wald statistic is reported as 99.4, implying the model is not a proper explanation for the observed volatility, either.

Table 2.11: Volatility of the Endogenous Variables and the Directed Wald Statistic

Volatility of the endogenous variables	95% lower bound	95% upper bound	Values calculated with real data	In/Out
$\mathrm{var}(\tilde{i})$	0.0229	0.1174	0.0171	**Out**
$\mathrm{var}(x)$	0.0380	0.1430	0.0951	In
$\mathrm{var}(\pi)$	0.0532	0.1158	0.0153	**Out**
Directed Wald percentile (for volatility)			99.4	

Note: Values reported in table 2.11 are magnified by 1000 times as their original values.

Indeed, the poor explanatory power of model three is not only detected by the directed Wald statistics but also the full Wald statistic when its overall fit to data is evaluated: note table 2.12 suggests a full Wald statistic of 99.9; this is another way of saying that it is almost impossible for the model to resemble the dynamics and 'size' of the data simultaneously. The implication is that this version with 'interest-rates-smoothed' Taylor rule is, in general, not a good proxy for the real-world economy.

Table 2.12: The Full Wald Statistic

The concerned model properties	Full Wald percentile
Dynamics + Volatility	99.9

2.3 Reconsidering the Prevailing Monetary Policy in the Light of the Test Results

2.3. 1 The best-fitting monetary policy rule in the US

What the last section has evaluated are three model versions that differ solely in monetary policy being pursued. It follows by ranking these models performance in fitting the data one will be effectively considering whether such facts are more likely to have been generated with the optimal timeless policy or the original Taylor rule, or with a Taylor-type policy where the interest rates are 'smoothed'[23]. The test results above summarised and ranked as follows:

Table 2.13: Summary of the Test Results

NK models	Directed Wald percentiles (for dynamics)	Directed Wald percentiles (for volatility)	Full Wald percentiles
Model one	98.2	10.4	96.5
Model two	100	91.5	100
Model three	99.9	99.4	99.9

Comparison by columns in table 2.13 immediately shows the first model, combined with the optimal timeless policy, is generally superior to its rivals in fitting US data as it consistently yields the lowest Wald statistics. This model is, too, the only version capable of explaining the dynamics and volatility of the data not only separately but also jointly. In the cases where Taylor rules are incorporated into exactly the same economic environments, by contrast, model two is only able to capture the scale of the economy's volatility, whereas model three is completely rejected by the data in all dimensions.

2.3. 2 Taylor rules as statistical relationships

The above suggests that the widespread success reported in single-equation Taylor rule regressions on US data could simply represent some sort of *statistical* relationships emerging from the model with the optimal timeless policy. This possibility can be examined by treating the optimal timeless rule model again as the true model, the null hypothesis and test whether the existence of empirical Taylor rules would be consistent with it.

[23] Thus the 'true' monetary policy rule is identified as a part of the 'true' model *in a relative sense*.

Suppose an arbitrarily specified Taylor-type regression is estimated to infer the potential 'Taylor rule' of the US economy. For simplicity, let the regression equation be:

$$\tilde{i}_t = \gamma_\pi \pi_t + \gamma_x x_t + \rho \tilde{i}_{t-1} + \xi_t \qquad [2.5]$$

where variables have their usual meanings. [2.5] can be estimated either using OLS if we assume the basic requirements for an OLS estimator are fulfilled, or via the IV approach to allow for possible correlations between the explanatory variables and the error term. The OLS and IV estimates based on the US data between 1982Q2 and 2007Q4 are summarised in table 2.14[24].

Table 2.14: Estimates of Taylor-type Regression [2.5]

	γ_π	γ_x	ρ	Adjusted R^2
OLS estimates	0.0453	0.0922	0.8233	0.92
IV estimates	0.0376	0.1003	0.8017	0.90

Now, use the technique of indirect inference to test if the observed 'Taylor rule' can be explained by model one based on the data simulated for the same period[25]. The test results are revealed as follows:

[24] The IV estimation here takes the lagged inflation and lagged output gap as instruments for their corresponding current values, respectively.

[25] Note: a) While one may expect the estimates of γ_π reported in table 2.14 be greater than one such that the 'Taylor principle' would be found, note that most existing literature has *treated* the interest rates series that is I(1) as a stationary series (See Carare and Tchaidze (2005), pp.17, footnote 17), whereas stationarity is obtained here by de-trending the data; Indeed, the 'Taylor principle' could be retrieved if the original I(1) interest rates series were used for estimation. b) In terms of the method of indirect inference, here it takes the Taylor-type regression [2.5] as the auxiliary model and sees the estimates of this (as reported in table 2.14) as the 'reality'.

Table 2.15: Individual Taylor Rule Coefficients and the Directed Wald Statistic

Panel A: Test for the OLS Estimates

Taylor rule coefficients	95% lower bound	95% upper bound	Values calculated with real data	In/Out
γ_π	0.0514	0.3436	0.0453	**Out**
γ_x	-0.0702	0.0650	0.0922	**Out**
ρ	0.6330	0.9198	0.8233	In
Directed Wald percentile (for Taylor rule coefficients)			97.1	

Panel B: Test for the IV Estimates

Taylor rule coefficients	95% lower bound	95% upper bound	Values calculated with real data	In/Out
γ_π	-0.8867	0.3062	0.0376	In
γ_x	-0.1072	0.0514	0.1003	**Out**
ρ	0.6454	0.9420	0.8017	In
Directed Wald percentile (for Taylor rule coefficients)			97.8	

According to table 2.15, although the real-data-based estimates of the 'Taylor rule' coefficients are not all individually captured by the model-implied 95% bounds, they are indeed explained as a set by the joint distribution of their simulation-based counterparts at the 99% confidence level, since the directed Wald statistics are reported as 97.1 and 97.8 (panel A and B), respectively, indicating that it is statistically possible for model one to imply such 'Taylor rules' observed from both OLS and IV estimations as table 2.14 shows.

This illustrates the identification problem with which this thesis began in chapter 1: a Taylor-type relation that has a good fit to the data may well be generated by a model where there is no *structural* Taylor rule at all[26]. Hence, any estimated or calibrated Taylor rule, no matter how well it might predict the actual movements of the nominal interest rates, is not by itself evidence that monetary policy follows this rule.

Table 2.16 below also summarises the Wald statistics when the optimal timeless rule model is used to explain several popular variants of the Taylor rule estimated with OLS. The reported Wald statistics suggest these are all well captured by the model. The model is thus *robust* in generating essentially the whole range of Taylor rules that have been estimated on US data.

Table 2.16: Model One in Explaining Different Taylor Rules (by OLS)

Taylor-type regressions	Adjusted R^2	Directed Wald percentiles (for Taylor rule coefficients)
$\tilde{i}_t = \gamma_\pi \pi_t + \gamma_x x_t + \xi_t$ $\xi_t = \rho_\xi \xi_{t-1} + \varepsilon_t$	0.89	92.9
$\tilde{i}_t = \gamma_\pi \pi_{t-1} + \gamma_x x_{t-1} + \xi_t$	0.40	87.0
$\tilde{i}_t = \rho \tilde{i}_{t-1} + \gamma_\pi \pi_{t-1} + \gamma_x x_{t-1} + \xi_t$	0.90	97.9

2.3. 3 The 'interest rates smoothing' illusion: an implication

Another issue on which the test results and analysis above sheds light is related to 'interest rates smoothing'. In an early paper Clarida, Gali and Gertler (1999) claimed that a 'puzzle' regarding the central banks' behaviour was yet to be solved, as the timeless rule generally derived from a standard NK model as optimal policy response to changes of macro variables would imply once-and-for-all adjustments of the nominal interest rates, whereas empirical 'evidence' from typical Taylor-type regressions estimated with the data usually displayed a high degree of 'interest rates

[26] Note that the adjusted R^2's reported in table 2.14 are as high as 0.92 for the OLS estimates and 0.90 for the IV estimates.

smoothing', in which case the sluggishness of interest rates variations could not be rationalized in terms of optimal behaviour.

While various authors explain such a discrepancy either at a theoretical level (e.g., Rotemberg and Woodford (1997, 1998), Woodford (1999, 2003a, 2003b)) or at an empirical level (e.g., Sack and Wieland (2000), Rudebusch (2002)), the tests here support the optimal timeless rule but reject the Taylor rule with 'interest rates smoothing'—implying the Fed has been responding to economic changes optimally without deliberately smoothing the interest rates. It is the persistence in the shocks themselves that induced the appearance of inertia in interest rates setting. Furthermore the above suggests one would find regressions of 'interest-smoothing Taylor rules' successfully fit the data even though this was being produced by the optimal timeless rule model. This last explains how the Fed's optimal responses could have been mis-interpreted as 'policy inertia' due to these misleading regressions.

2.4 Conclusion

This chapter has attempted to identify the principles governing the Fed's monetary behaviour since the early 1980s. It gets around the identification problem plaguing the earlier efforts to estimate these as chapter 1 reviewed by setting up three models, each with the same New Keynesian structure but differing only in the monetary policy being followed. These include an optimal timeless rule, a standard Taylor rule and another with 'interest rates smoothing'. Using statistical inference based on the method of indirect inference, the tests here show that only the optimal timeless rule can replicate both the dynamics and the volatility of the data. The tests also show that if the optimal timeless rule model was operating it would have produced data in which regressions of an interest-rate-smoothed Taylor rule would have been found. These suggest the Fed's policy in this period has been approximately optimal and the fact that its behaviour looks like a Taylor rule with interest-rates smoothing is a statistical artefact.

Chapter 3

On the Causes of the Great Moderation: is there a story of improved policy?

Introduction

Based on the method of indirect inference chapter 2 has shown that the optimal timeless policy actually outperforms Taylor rules of the standard sort in representing the Fed's monetary behaviour in the Great Moderation episode under the settings of standard New Keynesian model. This finding does not only undermine the 'good policy' explanation to the Great Moderation based on DSGE model simulation as chapter 1 reviewed as this relies on the validity of Taylor rules, it would also be challenging to the 'good shock' explanation where support is from structural VAR analysis with an identification scheme involving a Taylor-type policy. Indeed, the fact that the optimal timeless policy, when operating since 1982, looks like and may be misinterpreted as a Taylor rule as chapter 2 showed suggests we may have too misinterpreted the causes of the Great Moderation since earlier decomposition of these has failed to build on the true policy being followed.

This chapter extends the exercise carried out in chapter 2 to the US data in the 1970s. The aim of this is to uncover the Fed's monetary behaviour in the Great Acceleration episode and re-decompose the causes of the Great Moderation based on the extended results. It starts with the optimal timeless policy, the only model version that survived in the Great Moderation. It then turns to several 'weak' Taylor rule versions nearest to the testable case of the 'good policy' story and finally a non-standard Taylor rule by Ireland (2007). The tests suggest the optimal timeless policy and the (unrestricted version of) Ireland rule that effectively enforcing the former are not rejected by the data, whereas the 'weak' Taylor rules are strongly rejected, implying that the Fed's post-war policy was rather stable and roughly optimal. It further implies that the US economy was improved upon improved environment instead of improved policy. Variance decomposition suggests smaller demand disturbance accounted for the Moderation in interest rates, smaller supply disturbance for that in output and smaller policy errors for that in inflation. However, the optimal timeless policy, when implemented in the Great Acceleration, could have generated data in which Taylor rule regressions could have been found as it would in the Great Moderation. This could create an illusion that monetary policy was following such rules and that the regime switch story was plausible.

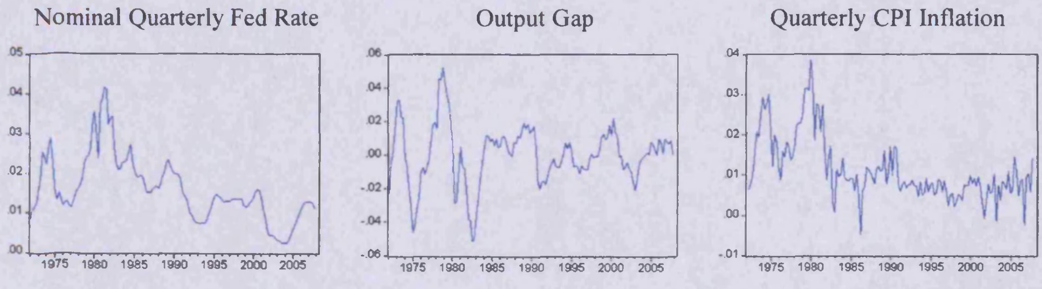
The rest of this chapter is organised as follows: section 3.1 briefly recaps the debate over the causes of the Great Moderation as chapter 1 summarised; section 3.2 builds on the effort of chapter 2, extends it to the Great Acceleration episode and sets out the argument that it was the improved environment that caused the Great Moderation; this is followed by section 3.3 that decomposes the impacts of shocks; section 3.4 returns to the story of 'improved policy' and explains how illusion of this could arise; section 3.5 concludes.

3.1 The Great Moderation in US and Its Determinants

Referring to the period during which the volatility of the main economic variables was relatively modest, the Great Moderation began in the US around the early 1980s although there is no consensus on the exact date. Figure 3.1 below shows the time paths of three main macro variables of US from 1972 to 2007; these are the nominal Fed interest rates, the output gap[27] and CPI inflation—the variables involved in the baseline model used in the last chapter. It shows that the massive fluctuations of the 1970s ceased after the early 1980s, indicating the economy's transition from the Great Acceleration to the Great Moderation.

Figure 3.1: Time Paths of Main Macro Variables of the US Economy

(Quarterly Data, 1972-2007)



Data source: the Federal Reserve Bank of St. Louis (http://research.stlouisfed.org/fred2/, accessed Nov. 2009). Fed rate and inflation unfiltered; output gap defined as log deviation of real GDP from HP trend.

[27] Recall that the output gap is defined as the percentage deviation of real GDP from its HP trend as in chapter 2. Note this series is plotted here in substitution of the actual output usually discussed in the context of Great Moderation as this is what predicted by the baseline model. Yet the actual data show these two series are highly correlated, with a correlation coefficient being as high as 0.98.

Previous efforts to explain this phenomenon have focused mainly on the monetary policy being followed and the environment (as shocks) affecting the economy; decomposition between these was based either on a time-series model building on the data or on a DSGE model building on theory. The former has tended to point to the environment, whereas the latter to the policy. However, as the review in chapter 1 suggested, the time-series method, while ensuring that the facts are well accounted for, has been intruded pervasively by the problem of identification, whereas the DSGE alternative, in which models are in general over-identified, cannot ensure that the facts are well accounted for. The review suggested the two methods in the literature were essentially complements to each other, and that a method of evaluation well-founded both in facts and in theory must be found. This pointed to DSGE models that failed to be strongly rejected by the data.

The optimal timeless rule model evaluated in the last chapter is one of these kinds. There the test used the ability of the DSGE model to replicate the description of the data provided by a time-series model—a VAR as it did. The DSGE model constructed causality, and the VAR described the facts. It showed when both theory and data were used in harness in such a way, the model stood out as the best representation of the US economy since the early 1980s; once this was replaced by other standard Taylor rule versions (on which typical 'good policy' story relied), the model was totally rejected. This points the way to a possible way forward as one could start with the timeless rule model—the 'true' model in the Great Moderation—and test for its consistency with data in the Great Acceleration; of course, other model versions are completely possible. One could then compare the two episodes using the best-fitting model of each as the test suggests and decompose the causes of change in data variability between policy and environment[28]. This approach is taken as it goes in what follows.

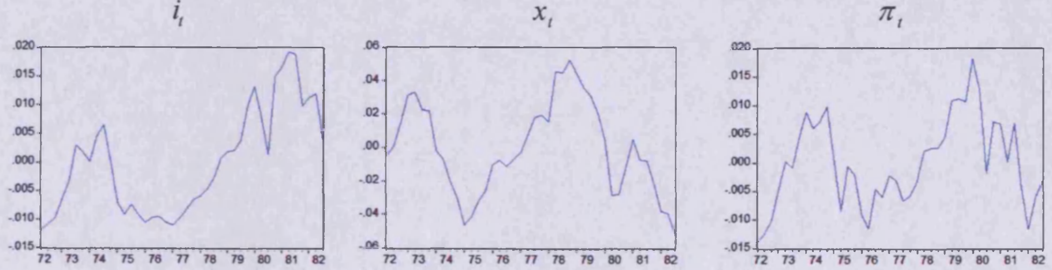
[28] Note the 'true' model also identifies the 'true' policy.

3.2 The Fed's Behaviour in the Great Acceleration: how different is it from that in the Moderation?

To identify the underlying policy of the Fed in the episode of the Great Acceleration, this section goes on using the method of indirect inference introduced in chapter 2 to test and rank competing models' performance in fitting the dynamics and volatility of the data. It traces back to the US economy since the breakdown of the Bretton Woods system in 1972 and defines the episode between this and 1982 as the Great Acceleration and the later the Great Moderation. This makes the Moderation episode consistent with the data sample used in the last chapter and is also supported by the Qu and Perron (2007) test that suggests a break in the data between 1980 and 1984 (See table 3A.1 in chapter appendix for the Qu-Perron estimate). It also follows the earlier practice of using a VAR(1) fitted with demeaned-detrended data to describe the reality and evaluate competing models against it[29]. Figure 3.2 below plots the stationarised time series involved on the latest data revised by the Fred®; the unit root test results are reported in table 3.1 to follow[30].

Figure 3.2: Stationarised Time Series on the Latest Data

Panel A: The Acceleration Subsample (1972Q2-1982Q3)

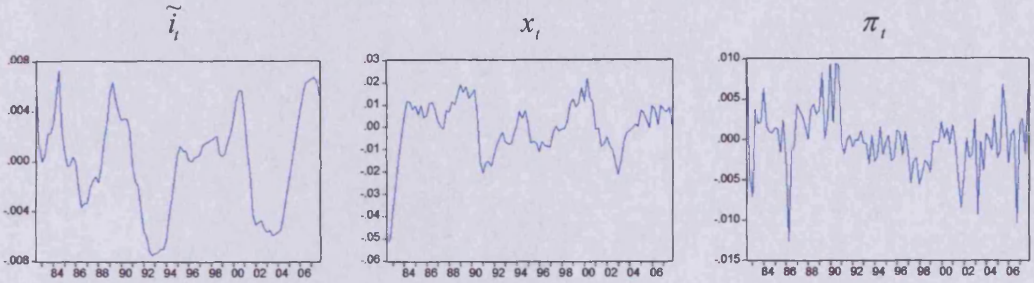


[29] Recall the VAR(1) representation takes the form:

$$\begin{bmatrix} \tilde{i}_t \\ x_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{31} & \beta_{32} & \beta_{33} \end{bmatrix} \begin{bmatrix} \tilde{i}_{t-1} \\ x_{t-1} \\ \pi_{t-1} \end{bmatrix} + \Sigma_t \cdot$$

[30] This data revision mainly involves the time series of real GDP and therefore the output gap in the 1980s. I show in what follows that the argument in chapter 2 is not qualitatively affected by this change. The rest of this thesis will all be built on these data.

Panel B: The Moderation Subsample (1982Q3-2007Q4)



Note: $\tilde{i}_t$ ≡ deviation of quarterly Fed rate from steady-state value; x_t ≡ log difference of quarterly real GDP from HP trend; π_t ≡ quarterly CPI inflation.

Table 3.1: Unit Root Test for Stationarity

Panel A: The Acceleration Subsample

Time series	5% critical value	10% critical value	ADF test statistics	p-values*
$\tilde{i}_t$	-1.95	-1.61	-1.71	0.0818
x_t	-1.95	-1.61	-1.67	0.0901
π_t	-1.95	-1.61	-2.86	0.0053

Panel B: The Moderation Subsample

Time series	5% critical value	10% critical value	ADF test statistics	p-values*
$\tilde{i}_t$	-1.94	-1.61	-2.91	0.0040
x_t	-1.94	-1.61	-4.42	0.0000
π_t	-1.94	-1.61	-3.34	0.0010

Note: 1. '*' denotes the Mackinnon (1996) one-sided p-values.
2. H_0: the time series has a unit root.
3. Adjusted observation sample for Great Acceleration: 1972Q3—1982Q3.
Adjusted observation sample for Great Moderation: 1982Q4—2007Q4.

Since the purpose of this extended exercise is to compare the policy and environment in the Great Acceleration to those in the Great Moderation, it assumes that only changes in policy and in the shocks' dynamics are possible. It therefore adopts the same baseline IS-Phillips Curve framework [2.1 and 2.2] as in chapter 2 and assumes the *structural* parameters calibrated for the post-break sample (table 2.3) are also applicable to the pre-break sample. Yet the policy response parameters and the shocks' persistency (ρ 's) are remained unrestricted. The latter, in particular, are left freely determined by the model and sample data under evaluation as the method of indirect inference requires.

The optimal timeless rule model in the Great Acceleration:

The extension here starts with the optimal timeless rule model [2.1] to [2.3] which is found to have outperformed the other Taylor rule versions widely accepted in representing the Fed's behaviour in the Great Moderation. Table 3.2 replicates this earlier finding using the updated data just described. It shows the timeless rule model is the only model version failed to be rejected by the data at normal confidence levels[31]. This model explains the chosen features, the dynamics and volatility, of the data not only separately but also jointly at 95% (or indeed, even at 90%). When the optimal timeless policy is replaced with the original Taylor rule or its interest-rate-smoothed version with commonly accepted calibrations, the model is strongly rejected at 99%.

Table 3.2: Review of the Best-fitting Model in the Great Moderation

Chosen features	Optimal timeless rule model	Taylor rule models	
		original	interest-rate-smoothed
Directed Wald (dynamics only)	86.4	100	99.8
Directed Wald (volatilities only)	89.6	99.2	99
Full Wald (dynamics + volatilities)	77.1	100	99.7

Table 3.3 in what follows reveals how the same model would behave in the episode of the Great Acceleration.

[31] Note with the updated data the earlier implication of table 2.13 (chapter 2) is substantially strengthened with clearer acceptance of the optimal timeless rule model.

Table 3.3: The Optimal Timeless Rule Model in the Great Acceleration

Panel A: Individual VAR Coefficients—Directed Wald Statistic

VAR(1) Coefficients	95% lower bound	95% upper bound	Values estimated with real data	In/Out
β_{11}	0.4146	1.0629	0.9519	In
β_{12}	-0.2505	0.1274	0.0592	In
β_{13}	-0.8794	0.5251	-0.1089	In
β_{21}	-0.3401	0.3581	-0.5006	**Out**
β_{22}	0.6090	0.9994	0.9474	In
β_{23}	-0.8439	0.7108	-0.4702	In
β_{31}	-0.1360	0.1962	0.1398	In
β_{32}	-0.0551	0.1566	0.0865	In
β_{33}	-0.0147	0.7576	0.5490	In
Directed Wald percentile (for dynamics)				98.2

Panel B: Volatilities of the Endogenous Variables—Directed Wald Statistic

Volatilities of the endogenous variables	95% lower bound	95% upper bound	Values calculated with real data	In/Out
var($\tilde{i}$)	0.0905	0.6543	0.0841	**Out**
var(x)	0.1559	1.4	0.7420	In
var(π)	0.0262	0.0722	0.0586	In
Directed Wald percentile (for volatilities)				89.6

Note: Estimates reported in panel B are magnified by 1000 times as their original values.

Panel C: Full Wald Statistic

The concerned model properties	Full Wald percentile
Dynamics + Volatilities	97.3

As the first panel in table 3.3 shows, the VAR(1) coefficients estimated with the actual data are all well captured by their 95% bounds implied by the model, apart from β_{21} which lies below its lower limit—thus at 95% confidence level the model overpredicts this partial response of the output gap to the lagged interest rates. The directed Wald percentile at 98 suggests that these estimates, though individually almost all within their 95% bounds, are jointly rejected at 95% but not at 99%.

Turning to the model's performance in fitting the data volatility, panel B suggests that except for the variance of the interest rates, which is slightly overpredicted by the model, the variances lie well within the model-implied 95% bounds. The directed Wald percentile at 89.6 indicates the model cannot be rejected even at 90% when it is used to explain the actual volatility.

Overall, when all features of the data are combined, the full Wald statistic in panel C is 97.3 and so fails to be rejected at confidence levels between 95 and 99%. So while the model fits the facts less well than in the case of the Moderation subsample, it still fits those of the turbulent Acceleration episode reasonably well.

Taylor rule models in the Great Acceleration:

The review in chapter 1 suggested most authors using theory to explain the Great Moderation had counted on DSGE models with a Taylor rule that was 'passive' before the break and was 'active' thereafter. T hey argued the improvement was caused by improved policy, a regime shift shown by the change in Taylor rule parameters. Indeed, although such an assertion has been partly rejected by our earlier examination that suggested active Taylor rule models of the standard sort were incompatible with the data in the Moderation, it would still be interesting to know how passive v ersions of these would perform in the Acceleration. Unfortunately DSGE models with the generally proposed pre-1982 Taylor rules are technically untestable because the solution is indeterminate, the models 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. But 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, one 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 one would have to impose an infinite variance; by implication all possible outcomes would be embraced by the simulations of the model and hence the model cannot be falsified. Thus the pre-1982 Taylor rule DSGE models are in general not a testable theory of this period.

However, testing the model with a pre-1982 Taylor rule that gives a determinate solution is completely possible. To make it analogous to the untestable case just described, one could set the Taylor rule as unresponsive to inflation as is consistent with determinacy, implying a long-run inflation response of just above unity (so the Taylor Principle is *just* satisfied). Such a rule would show considerably more monetary 'weakness' than the rule typically used for the post-1982 period, when the long-run response of interest rates to inflation was 1.5 in the original rule [1.1] without smoothing and as high as 6 in [1.2] with smoothing which was the one that fitted the data least badly.

The following implements this weak Taylor rule across a spectrum of combinations of smoothing parameter and short-run response to inflation, with in all cases the long-run coefficient equalling 1.001. The suggested Wald statistics are revealed in table 3.4 to follow.

What we see here is that with a low smoothing parameter 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 parameter, the model is rejected less strongly by the data dynamics and also overall but it is then increasingly at odds with the data variances. In all cases the model is rejected strongly overall by the data, though least badly with the highest smoothing parameter. Thus the testable model that gets nearest to the position that the shift in data variability was due to the shift in Taylor rule parameters is rejected most conclusively.

Table 3.4: Wald Statistics for 'Weak' Taylor Rule Models in the Great Acceleration (with 'weak' rule defined as having a long-run interest-rates response to inflation equalling 1.001)

Taylor rule: $\tilde{i}_t = \rho\tilde{i}_{t-1} + \gamma_\pi \pi_t + \xi_t$		Wald percentiles for chosen features *(Normalized t-values in parenthesis[32])*		
Rule parameters	Dynamics of error process estimated from data	Directed Wald for dynamics	Directed Wald for volatilities	Full Wald for dyn. & vol.
$\rho = 0, \gamma_\pi = 1.001$	$\xi_t \sim AR(1)$	100 *(39.81)*	78.9 *(0.22)*	100 *(40.24)*
$\rho = 0.3, \gamma_\pi = 0.7007$	$\xi_t \sim AR(1)$	100 *(30.26)*	92 *(1.08)*	100 *(28.01)*
$\rho = 0.5, \gamma_\pi = 0.5005$	$\xi_t \sim AR(1)$	100 *(22.69)*	95.9 *(1.77)*	100 *(21.98)*
$\rho = 0.7, \gamma_\pi = 0.3003$	$\xi_t \sim iid$	100 *(19.26)*	98.2 *(2.73)*	100 *(18.24)*
$\rho = 0.9, \gamma_\pi = 0.1001$	$\xi_t \sim iid$	100 *(9.09)*	99 *(3.56)*	100 *(9.03)*

Ireland's alternative Taylor rule representation of Fed policy [1.5]:

A recent paper by Ireland (2007), unlike the conventional New Keynesian approach of assuming differing Taylor rules respectively in the Great Acceleration and Great moderation, estimates a model in which there is a non-standard Taylor rule that is held constant across both post-war episodes. His rule always satisfies the Taylor Principle because unusually it is the change in interest rates 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 changes in the rule's parameters but by a time-varying inflation target which he treats under

[32] T-value normalization of the Wald percentiles is calculated based on Wilson and Hilferty (1931)'s method of transforming a chi-squared distribution into a standard normal distribution. The formula used here takes the form: $Z = \{[(2M^{squ})^{1/2} - (2n)^{1/2}]/[(2M^{squ^{95th}})^{1/2} - (2n)^{1/2}]\} \times 1.645$, where M^{squ} is the square of the Mahalanobis distance calculated from equation [2.4] (chapter 2) with actual 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.

the assumptions of 'opportunism' largely as a function of the shocks hitting the economy. Ireland showed that his model, when estimated using data from 1959 to 2004, implied that the Fed had an implicit inflation target trending upwards before 1980s which then reversed afterwards. Based on the similarity of this to the actual path of inflation, he argued that the Great Acceleration was caused by the Fed's decision to translate short-run price/inflation pressures into persistent movements in inflation; these pressures then ceased from the early 1980s and the Fed exploited this by setting the inflation target low, ultimately bringing actual inflation down.

Ireland's model, like the timeless rule model, essentially implies that the cause of the Great Moderation is the fall in the shocks' variance. However the difference is that it attributes the policy variance change partly to the change in the variance of the inflation target, whereas the timeless optimum attributes it entirely to the change in the variance of the policy ('trembling hand') error.

A full test of Ireland's model by indirect inference cannot be carried out here because his model restricts the target-related part of the error in his Taylor rule to be a function of the other errors in his model according to his opportunistic theory of policy target choice; as the baseline model here is different from his in a variety of ways, these restrictions are unable to be tested[33]. However, one can test his model in unrestricted form where the error in his particular Taylor rule is let freely determined by the data. Table 3.5 shows the results of this exercise.

---

[33] In particular, Ireland set $\widetilde{i}_t = \widetilde{i}_{t-1} + \gamma_\pi \pi_t + \gamma_g (g_t - g) - \gamma_\pi \pi_t^* - \Delta\pi_t^* + \Omega_t$, where (while variables have their usual meanings) π_t^* , the (unobservable, implicit) inflation target, was assumed to follow $\pi_t^* = \pi_{t-1}^* - \delta_\theta \varepsilon_{\theta t} - \delta_Z \varepsilon_{Zt} + \sigma_\pi \varepsilon_{\pi t}$, where $\varepsilon_{\theta t}$ and ε_{Zt} are the cost-push shock and technology shock, respectively, and $\delta_\theta > 0$ and $\delta_Z > 0$ are the central bank's opportunistic response of policy target choice. Ireland claimed that by reacting to 'supply shocks' of this sort the Fed translated short-run price/inflation pressures into persistent movements in inflation; he also assumed the inflation target could vary exogenously due to innovation denoted by $\varepsilon_{\pi t}$.

Ireland's original specification of Taylor rule cannot be meaningfully tested here for two reasons. First, he used non-filtered and so in general non-stationary data, whereas here HP-filtered and thus stationary data are used. Second, the inflation target in his rule is set to be determined by two structural supply shocks (i.e., $\varepsilon_{\theta t}$ and ε_{Zt}) that cannot be identified in our 3-equation reduction of the DSGE model where the supply shock is an aggregate shock to the Phillips curve. Ireland in his paper estimated this unobservable target by using the Kalman filter based on Maximum Likelihood. While we could use indirect inference to evaluate and estimate Ireland's original specification, we cannot do so in the set-up we have here.

Table 3.5: Wald Statistics for the Baseline Model with Ireland-type Policy

Wald statistics for chosen features	Ireland's rule in unrestricted form: $\tilde{i}_t = \tilde{i}_{t-1} + \gamma_\pi \pi_t + \gamma_g (g_t - g) + \xi_t$ & equivalent transformation[34]: $\tilde{i}_t = \tilde{i}_{t-1} + \gamma_\pi \pi_t + \gamma_g (x_t - x_{t-1}) + \xi_t$	
	pre-1982 sample	post-1982 sample
Directed Wald for dynamics	98.9	79
Directed Wald for volatilities	78.8	89.4
Full Wald for dynamics & volatilities	98.1	71.1

Note: 1. Ireland (2007)'s ML estimates suggest $\gamma_\pi = 0.91, \gamma_g = 0.23$.

2. All equation errors follow an AR(1) process according to the data and model.

It turns out that the unrestricted Ireland model is hardly distinguishable from the optimal timeless rule model. The unrestricted Ireland rule changes interest rates until the optimal timeless policy is satisfied, in effect forcing it on the economy. Alternatively one can write the Ireland rule as a rule for inflation determination (i.e. with inflation on the left hand side), which reveals that it is identical to the timeless rule's setting of inflation apart from the term in the change in interest rates and some slight difference in the coefficient on output gap change[35]. Since the Ireland rule is so similar to the timeless optimum, it is not surprising that the Wald percentiles for it are hardly different: 71.1 in the Great Moderation (against 77.1 for the optimal timeless rule model) and 98.1 in the Great Acceleration (against 97.3).

[34] While Ireland originally specified $\tilde{i}_t = \tilde{i}_{t-1} + \gamma_\pi \pi_t + \gamma_g (g_t - g) - \gamma_\pi \pi_t^* - \Delta\pi_t^* + \Omega_t$, the exercise here tests its unrestricted form, where $\tilde{i}_t = \tilde{i}_{t-1} + \gamma_\pi \pi_t + \gamma_g (g_t - g) + \xi_t$ and $\xi_t = -\gamma_\pi \pi_t^* - \Delta\pi_t^* + \Omega_t$. In particular, this unrestricted Ireland rule is rewritten as $\tilde{i}_t = \tilde{i}_{t-1} + \gamma_\pi \pi_t + \gamma_g (x_t - x_{t-1}) + \xi_t$ so that it can be evaluated within the baseline framework; such an equivalent transformation is derived by writing: $g_t - g = \ln y_t - \ln y_{t-1} - (\ln y_t^{hptr} - \ln y_{t-1}^{hptr}) = \ln y_t - \ln y_t^{hptr} - (\ln y_{t-1} - \ln y_{t-1}^{hptr}) = x_t - x_{t-1}$.

[35] Note the Ireland rule $\tilde{i}_t = \tilde{i}_{t-1} + \gamma_\pi \pi_t + \gamma_g (x_t - x_{t-1}) + \xi_t$ can be rewritten as $\pi_t = \frac{1}{\gamma_\pi}(\tilde{i}_t - \tilde{i}_{t-1}) - \frac{\gamma_g}{\gamma_\pi}(x_t - x_{t-1}) - \frac{1}{\gamma_\pi}\xi_t$ that mimics the optimal timeless policy [2.3]; its coefficient on output gap change, according to Ireland's estimation, is 0.25, close to that of 0.17 in the latter.

Ireland's Taylor rule can in principle be distinguished from the optimal timeless policy via his restriction on the rule's error. As noted earlier, however, this restriction is unavailable within the baseline model being employed so that the Ireland rule in its unrestricted form here only differs materially from the optimal timeless policy in the interpretation of the error. But 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.

It might be argued that the success of Ireland's rule reveals that a type of Taylor rule does after all explain the data. This would be true. But in the context of the debate over the cause of the Great Moderation it is to be firmly distinguished from what we call the 'standard Taylor rule' under which policy shifts in the rule are regarded as the cause. In Ireland's rule there are no such shifts and as we have seen the behaviour under it is essentially identical to that from the optimal timeless policy. This finding and its corollaries are the key contributions of this chapter, however one chooses to describe the rule.

Concluding remarks on the comparison of the optimal timeless rule model and Taylor rule models:

While by contrast the baseline DSGE model with the optimal timeless policy has more trouble explaining the pre-1982 period than the post-1982, it is therefore not rejected at reasonable levels of confidence. This has important causal implication about the shift in the behaviour of the post-war US economy. In particular if this model is the true data-generating mechanism of the US history since the early 1970s, it does of course imply that there was no structural shift in the parameters—especially in those of monetary policy—between the two periods since it is the same model that has been used to fit both periods. Accordingly it also implies that the changes were due to the errors. The next section goes on to investigate in more detail how the errors changed according to the timeless rule model.

3.3 How Has Environment Affected the US Economy since 1970s? A Variance Decomposition Analysis

What the above has shown is that the change in post-war US economy was more likely to be an issue of change in economic environment instead of change in economic policy. Table 3.6 below summarises the size of structural errors in both the Great Acceleration and Great Moderation according to the timeless rule model and the actual data. It shows the shocks have all fallen sharply after the break in 1982, with the demand and policy shocks each by 60%, while the supply shock a massive 80% (and of this 80% just under half was due to the fall in the shock's persistency).

Table 3.7 decomposes the impact of these shocks on the data's volatility. It can be seen that in both episodes the US economy was operating in a recursive manner: the output gap was dominated by the Phillips curve ('supply') shocks, while inflation was dominated by monetary policy shocks; with output gap and inflation set entirely independently of demand shocks, interest rates moved to offset these as well as reacting to output and inflation.

Table 3.6: The Shrinking Size of Shocks

Standard deviation of	Pre-1982	Post-1982
Demand shock	0.0358	0.0143
	(0.0043)	*(0.0010)*
Supply shock	0.7867	0.1595
	(0.0708)	*(0.0319)*
Policy shock	0.0132	0.0053
	(0.0054)	*(0.0033)*

Note: 1. Values in parenthesis are sample estimates of standard deviation of innovation.
2. Standard deviation of shocks is calculated using formula sd(err.)=sd(innov.)/(1-rho); rho is the sample estimate of AR(1) coefficient of the errors.

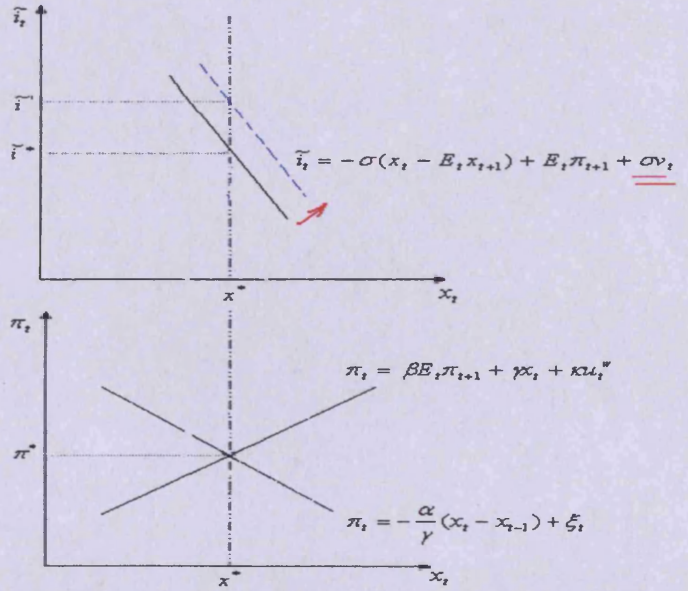
Table 3.7: Variance Decomposition of the Optimal Timeless Rule Model

Variables / Shocks	$\tilde{i}_t$	x_t (pre-break)	π_t	$\tilde{i}_t$	x_t (post-break)	π_t
Demand shock	91.3%	0%	0%	75.4%	0%	0%
Supply shock	5.9%	99.9%	8.4%	24%	99.1%	6.6%
Policy shock	2.8%	0.1%	91.6%	0.5%	0.9%	93.4%
Total contribution	100%	100%	100%	100%	100%	100%

To understand this recursiveness, recall that pursuing the optimal timeless policy [2.3] requires keeping inflation equal to a fixed fraction of the first difference of the output gap. In effect, such a policy constitutes a simultaneous pair with the Phillips curve in the model that pins down the equilibrium output gap and inflation; in the optimal timeless rule inflation responds to the first difference in the output gap, while in the Phillips curve something close to the first difference of future inflation is negatively related to the level of the output gap. Given that both inflation and the output gap are highly autoregressive both because of the errors and the model dynamics, these first differences will be rather small; hence in the Phillips curve the level of the output gap will largely be set by the equation (supply) error, while in the inflation rule the level of inflation will largely be set by that equation's (trembling hand) policy error. If we now turn to the IS curve, with inflation and the output gap already set, the equilibrium interest rates are then recursively set in its turn by the IS curve alone. In other words, under the optimal policy any innovation to the demand side will be fully neutralized by the adjustment of the real interest rates, leaving the output gap and inflation unaffected. The real interest rates also respond to the expected change in output gap but this is small because of output gap autocorrelation. The nominal interest rates also respond to expected future inflation; but this is dominated by the policy error which dies away quickly and so moves little also. Hence the dominance of the demand shock on nominal interest rates; the supply shock intrudes more on interest rates in the Great Moderation period because it is less persistent and so the expected change in the output gap is larger, affecting the real interest rates more. This structure is illustrated in figure 3.3; derivations are shown in appendix.

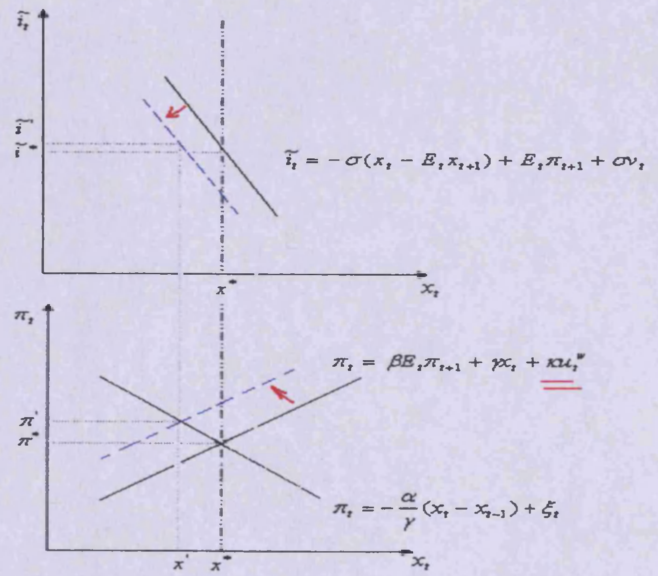
Figure 3.3: Working of the Optimal Timeless Rule Model in Face of Shocks

Panel A: When a (positive) Demand Shock Occurs



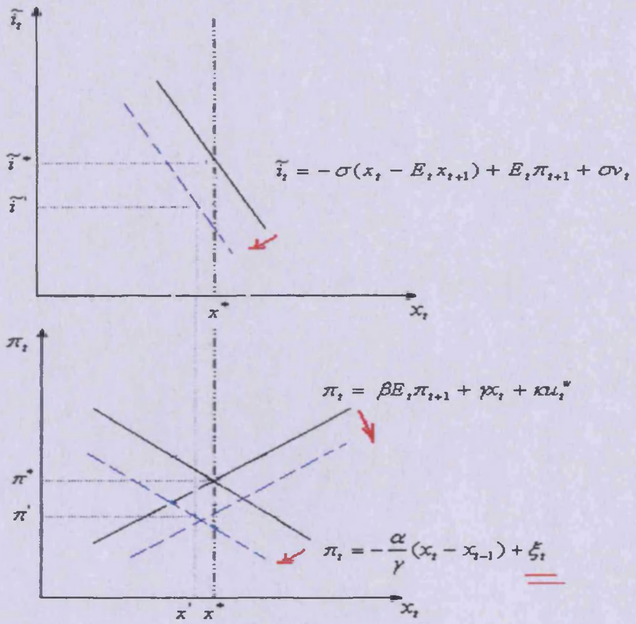
Note: a positive demand shock is totally offset by the movement of nominal rates of interest, leaving the output gap and inflation intact.

Panel B: When a (positive) Supply Shock Occurs



Note: a positive supply shock shifts the Phillips curve up along the policy curve; this is a joint effect of the dominating upward movement caused by the shock and the downward movement caused by reduced expected *future* inflation. The latter also works with lower expected future output gap to shift down the 'IS' curve; it could cause the interest rates to be higher or lower.

Panel C: When a (tightening) Policy Shock Occurs



Note: a tightening policy shock tightens monetary conditions, causing a fall in both expected future output gap and expected future inflation so all the curves shift downwards. The extent to which the 'IS' curve shifts determines whether nominal interest rates will rise or fall.

To summarise, while a decline in the variances of all shocks brought about the switch from the Great Acceleration to the Great Moderation, the variance decomposition here shows that the relative impact of these shocks on the economy has been fairly similar over time, apart from an increase in the role of the supply shock in interest rates. Smaller demand disturbance has stabilized the interest-rates fluctuation. Smaller supply disturbance has stabilized output, and smaller policy errors have stabilized inflation.

3.4 The 'Good Policy' Explanation Revisited: a victim of Taylor rule illusion?

Now it has been clear that the optimal timeless rule model can explain the Great Moderation, not by any change in the policy regime but rather purely through the changing variances of the shocks. How, then, can it be that economists have observed different Taylor rules across the two episodes and concluded from these that policy regime changes were at work? The answer suggested here is that such apparent rules were statistically observable because produced by the behaviour of the economy in conjunction with the (same) timeless optimum.

The typical 'good policy' explanation to the Moderation relies on evidence from an estimated Taylor rule that is presumed to describe the true behaviour of the Fed, and interprets the corresponding change in rule parameters estimated with different subsamples as shifts in monetary policy. However, the fact that a Taylor-type relation between the data may well be representing something else of a DSGE model instead of the structural policy being followed—the identification problem reviewed in chapter 1—has determined that regression evidence of this kind is fundamentally untenable. Indeed, such changing Taylor rule estimates could simply be an illusion arising from alterations in statistical relationships within the data that the true, unchanged policy has incurred.

This was partly demonstrated earlier in chapter 2 using the timeless rule model with data in the post-1982 sample. This exercise can be extended too to the pre-1982 sample[36]. Table 3.8 below shows several variants of Taylor rule the updated data may display before and after the break based on OLS and the extent to which these can be explained by the timeless rule model.

To compare the regression results it finds here with those commonly found in the US Taylor rule literature it should be emphasized that for the post-82 subsample a linear trend is taken out of the interest rates series to ensure stationarity. When estimated on the stationary data the exercise has used here, the Taylor rules obtained generally fail to satisfy the Taylor Principle, in much the same way as those pre-1982. Thus econometrically the standard estimates of the long-run Taylor rule response to inflation post-1982 are biased upwards. There is little statistical difference in the data of the two periods for estimated long-run Taylor rule responses to inflation. The reported Wald percentiles show that this is exactly what the timeless rule model of Fed behaviour implies: in both panel A and panel B the four hypothetical Taylor rules estimated with the data are all explained by the timeless rule model, at varying

[36] Recall that this involves using the method of indirect inference and treating the 'Taylor rule' specified as the auxiliary model and its parameters estimated with real data the 'reality'.

confidence levels. This indicates in both episodes if timeless rule model is the true data-generating mechanism we would find such relationships in the data exactly as the data says we do.

Table 3.8: 'Taylor Rules' Shown by Real Data (with OLS) and Explanatory Power of the Timeless Rule Model

Panel A: 'Taylor rules' in the Great Acceleration

Taylor rule estimated	γ_π	γ_x	ρ	Adjusted R^2	Wald percentile
$\tilde{i}_t = \gamma_\pi \pi_t + \gamma_x x_t + \rho \tilde{i}_{t-1} + \xi_t$	0.09	0.06	0.90	0.84	97.2
$\tilde{i}_t = \gamma_\pi \pi_t + \gamma_x x_t + \xi_t$ $\xi_t = \rho_\xi \xi_{t-1} + \varepsilon_t$	0.30	0.07	0.92	0.85	96.7
$\tilde{i}_t = \gamma_\pi \pi_{t-1} + \gamma_x x_{t-1} + \xi_t$	0.60	-0.01	N/A	0.24	36.1
$\tilde{i}_t = \rho \tilde{i}_{t-1} + \gamma_\pi \pi_{t-1} + \gamma_x x_{t-1} + \xi_t$	-0.11	0.06	0.82	0.83	65.6

Panel B: 'Taylor rules' in the Great Moderation

Taylor rule estimated	γ_π	γ_x	ρ	Adjusted R^2	Wald percentile
$\tilde{i}_t = \gamma_\pi \pi_t + \gamma_x x_t + \rho \tilde{i}_{t-1} + \xi_t$	0.08	0.05	0.89	0.92	21.7
$\tilde{i}_t = \gamma_\pi \pi_t + \gamma_x x_t + \xi_t$ $\xi_t = \rho_\xi \xi_{t-1} + \varepsilon_t$	0.07	0.06	0.93	0.90	47.4
$\tilde{i}_t = \gamma_\pi \pi_{t-1} + \gamma_x x_{t-1} + \xi_t$	0.26	0.13	N/A	0.24	10.9
$\tilde{i}_t = \rho \tilde{i}_{t-1} + \gamma_\pi \pi_{t-1} + \gamma_x x_{t-1} + \xi_t$	0.03	0.04	0.89	0.91	85

The implication from this last exercise is first that econometrically Taylor rules changed little between the two episodes once non-stationarity of the data is allowed for; second, that the Taylor rules found in the data could have been generated by the completely different monetary policy, the optimal timeless rule, that we found fits the data in general. This last was illustrated in chapter 2 for demonstration of the non-identifiability of Taylor rules; here it shows such Taylor rule illusion could actually arise in both episodes, suggesting the 'good policy' assertion might be a victim from this.

3.5 Conclusion

This chapter has attempted a fresh investigation of the reason for the shift to the US Great Moderation from its predecessor period, the Great Acceleration. The conventional DSGE approach to these episodes mostly starts with a New Keynesian model including a standard Taylor rule where the level of interest rates responds to inflation; the output gap may also enter, and so may the lagged interest rates as a smoothing mechanism. It goes on to claim that the shift was the result of improved policy in the form of higher Taylor rule responses to inflation. This chapter challenges this view. It builds on the earlier finding that the Fed's behaviour was better understood as the timeless optimum in the Great Moderation and extends it to the Great Acceleration. It shows that the same monetary principle also accounts for the data in this turbulent episode. From this it suggests that the Great Moderation was due to much reduced volatility of shocks.

The findings here, like the earlier, are based on the method of indirect inference in which the simulated behaviour from the DSGE model is compared with a VAR estimated on the actual data. The standard New Keynesian model with this optimal timeless policy instead of the Taylor rule explains the dynamics and volatility of US economy both before and after 1982. It also explains the existence of Taylor rule regressions found in the data and, thus, how the illusion of a regime switch could statistically arise.

In short, the extended exercise in this chapter implies that in that it followed the optimal timeless policy the Fed did a good job in both the Great Acceleration and the Great Moderation by ensuring that the economy was at its best possible state subject to the occurrence of shocks—the optimal trade-off between inflation and output growth required for social welfare loss minimization under timeless perspective was roughly satisfied. Given that the Fed's monetary behaviour was unchanged, it can be concluded that the Great Moderation was, in the final analysis, caused by the reduction of in the size of shocks after 1982: smaller demand shocks caused the moderation in interest rates, smaller supply shocks caused the moderation in output and smaller monetary policy shocks the moderation in inflation. While the reduced size of demand and supply shocks suggest the Fed was lucky that the economic environment improved, the reduced policy shock shows that the Fed itself also contributed to the greater economic stability by improving its own performance in monetary management.

Appendix to Chapter 3

A. Cut-off between the Great Acceleration and Great Moderation Suggested by the Data

Table 3A.1: Qu-Perron Test for Structural Break

Estimated break date	95% confidence interval lower	upper	supLR test statistic for a fixed number of breaks	5% critical value
1984Q3	1980Q1	1984Q4	164.84	31.85

Note: 1. Time series model: VAR(1) consisting of nominal Fed rates, output gap and inflation with no constant.
2. H_0: there is no structural break; H_1: there is one structural break.
3. Adjusted observation sample: 1972Q2—2007Q4.

B. Derivation of Impulse Responses of the Timeless Rule Model to shocks

a. Impulse response of inflation to shocks:

Given rational expectations and equations:

$$\pi_t = \beta E_t \pi_{t+1} + \gamma x_t + \kappa u_t^w \qquad [2.2]$$

$$\pi_t = -\frac{\alpha}{\gamma}(x_t - x_{t-1}) + \xi_t \qquad [2.3]$$

Rewrite [3.2] as:

$$x_t = \frac{(1-\beta B^{-1})\pi_t - \kappa u_t^w}{\gamma} \qquad ①$$

Also, write [2.3] as:

$$x_t = \frac{(\pi_t - \xi_t)\gamma}{(L-1)\alpha} \qquad ②$$

Equate ① to ② such that:

$$\frac{(1-\beta B^{-1})\pi_t - \kappa u_t^w}{\gamma} = \frac{(\pi_t - \xi_t)\gamma}{(L-1)\alpha} \qquad ③$$

Solve for π_t from ③ to obtain:

$$\pi_t = \frac{\gamma^2\alpha^{-1}\xi_t + \kappa u_t^w - \kappa u_{t-1}^w}{(1-L)[1-\beta B^{-1} + \gamma^2\alpha^{-1}(1-L)^{-1}]} \qquad ④$$

Now note that the supply error has a high autocorrelation so that the terms in it nearly cancel, while also the coefficient on it (kappa) is small, leaving the policy error as the dominant factor in inflation.

b. Impulse response of output gap to shocks:

Given rational expectations and equations:

$$\pi_t = \beta E_t\pi_{t+1} + \gamma x_t + \kappa u_t^w \qquad [2.2]$$

$$\pi_t = -\frac{\alpha}{\gamma}(x_t - x_{t-1}) + \xi_t \qquad [2.3]$$

Rewrite [2.2] as:

$$\pi_t = \frac{\gamma x_t + \kappa u_t^w}{1-\beta B^{-1}} \qquad ①'$$

Put ①' into [2.3] to obtain:

$$x_t = \frac{(1-\beta B^{-1})\xi_t - \kappa u_t^w}{\gamma + (1-\beta B^{-1})\frac{\alpha}{\gamma}(1-L)} \qquad ②'$$

Since $\xi_t = \rho_\xi \xi_{t-1} + \varepsilon_t^\xi$, the term in the policy error is small and as the standard deviation of the supply error is also massively larger than that of the policy error, this supply error then dominates the output gap.

c. Impulse response of interest rates to shocks:

Given 'IS' curve:

$$x_t = E_t x_{t+1} - (\frac{1}{\sigma})(\tilde{i}_t - E_t\pi_{t+1}) + v_t \qquad [2.1]$$

Rewrite [2.1] as:

$$\tilde{i}_t = \sigma(E_t x_{t+1} - x_t) + E_t\pi_{t+1} + \sigma v_t \qquad ①''$$

Now lead the targeting rule [2.3] for one period and take expectation at t to get:

$$E_t\pi_{t+1} = -\frac{\alpha}{\gamma}(E_t x_{t+1} - x_t) + E_t\xi_{t+1} \qquad ②''$$

Substitute ②'' into ①'' to obtain:

$$\tilde{i}_t = (\sigma - \frac{\alpha}{\gamma})(E_t x_{t+1} - x_t) + E_t \xi_{t+1} + \sigma v_t \qquad ③''$$

Since $\xi_t = \rho_\xi \xi_{t-1} + \varepsilon_t^\xi$ and therefore $E_t \xi_{t+1} = \rho_\xi \xi_t$, the above equals:

$$\tilde{i}_t = (\sigma - \frac{\alpha}{\gamma})(E_t x_{t+1} - x_t) + \rho_\xi \xi_t + \sigma v_t \qquad ④''$$

Note the expected change in output gap dominated by the supply error is small due to high autocorrelation, the standard deviation of demand error is some three times that of the policy error and σ is large, this demand error then dominates the interest rates.

Chapter 4

Further Comparison of Models basing on Estimation: can Taylor rule models get the upper hand?

Introduction

Having compared their capacity in mimicking the dynamics and volatility of the actual data, the last two chapters have shown that the optimal timeless rule model under the standard New Keynesian settings is superior to typical Taylor rule alternatives in representing the US economy since the early 1970s. This comparison, by using consensus calibrations in the literature, has effectively pinned down the structure, i.e., the 'IS' and Phillips curves, of the economy and fitted to it a series of competing monetary policies so that if the parameters assumed are true it would identify the marginal contribution and therefore the most precise specification of the operating policy; the decomposition exercise regarding the causes of the Great Moderation that followed was conducted on the basis of this.

However, fixing model parameters in such a way is a fairly 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 was tested in the previous chapters, this could mean that the Taylor rule models were rejected not because the policy specified was incorrect but that the calibrated 'IS' and Phillips curves had failed to reflect the true structure of the economy. Equally of course it could mean that the optimal timeless policy was untrue but this, when incorporated in the complete model, was saved by a set of 'good' structural parameters. 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 he 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. In other words the models are to be fully estimated before they are tested and evaluated against each other.

This chapter uses the method of indirect inference to estimate the timeless rule model and Taylor rule models discussed earlier in both the Great Acceleration and Great Moderation. The aim is to find the best possible versions of these in each episode as the data suggest and test the validity for that. It shows for both episodes there are versions with the optimal timeless policy and 'interest-rates smoothed' Taylor rule that the data fail to strongly reject. But in either case the former remains significantly less rejected compared to the latter, and the ranking of this is robust to a range of evaluation basis, including the choice of auxiliary model, the cut-off between the Acceleration and the Moderation and the method of data stationarization.

This chapter is organized as follows: section 4.1 explains the method of model estimation basing on indirect inference; section 4.2 sets out the benchmark estimates of the timeless rule model and Taylor rule model and re-compares the explanatory power of these; section 4.3 looks into several issues regarding robustness of the models' ranking, and section 4.4 compares the dynamic properties of these; section 4.5 concludes.

4.1 Indirect Inference as a Method of Estimation

Section 2.2.1 in chapter 2 has suggested that the method of indirect inference was originally designed for structural model estimation before it was recently developed for model evaluation. As the foregone has mentioned, this method distinguishes itself from other methodological alternatives by using an auxiliary model that is completely independent of the theory to generate descriptors of the data against which the theoretical model is estimated or evaluated indirectly. We have seen that in model evaluation as in the previous chapters where parameters were taken as given, the method calculated the Wald statistic [2.4] to see if the real-data-based estimates of the auxiliary model were captured by the joint distribution of these suggested by the theoretical model; the purpose was to see if the theory was 'close enough' to the data such that from the statistical viewpoint it could be taken as the true data-generating mechanism. In model estimation as it goes in this chapter, however, indirect inference is used in a different way, as the aim is no longer to measure the 'distance' between the theory and the data but to find a set of parameters that minimizes such distance when the theoretical model is taken as true. The common element is to calculate the Wald statistic based on the estimates of the auxiliary model. But the underlying question is with what structural parameters the real-data-based estimates are closest to the (joint) mean of these as model simulations would predict. Yet just as in the case of testing these estimates can be the estimated parameters of the auxiliary model or functions of these.

The estimation procedure can be summarised with the five steps as follows:

Step 1: *Select an auxiliary model and estimate it on the real data to produce benchmark descriptors of the reality.*

Step 2: *Assign initial values of structural parameters to be estimated and use these to generate a number of pseudo samples of simulated data with the theoretical model.*

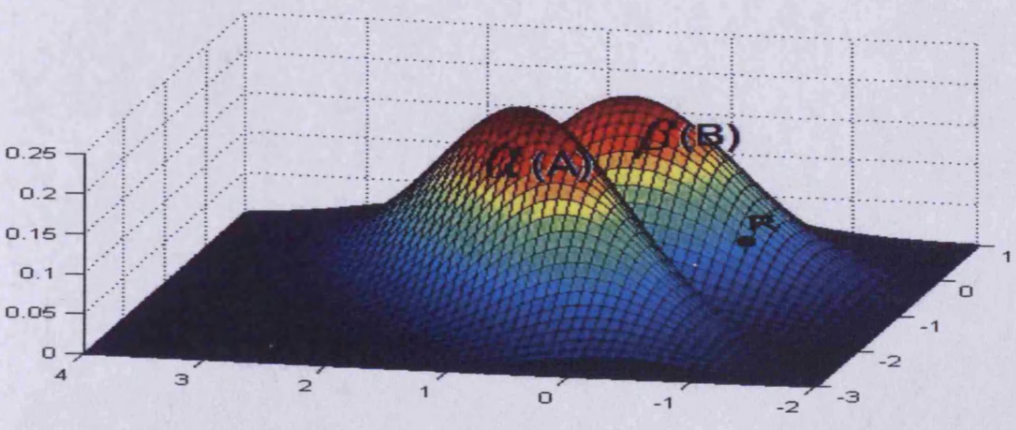
Step 3: *Estimate the selected auxiliary model on the simulated data obtained in step 2 to produce the (joint) distribution of the chosen descriptors and the mean of this.*

Step 4: *Calculate the Wald statistic formula [2.4], the square of 'Mahalanobis distance' to formally measure the distance between the data descriptors obtained in step 1 and the mean of these implied in step 3.*

Step 5: *Repeat steps 2, 3 and 4 until the calculated Wald statistic is minimized.*

The last step of the above can be illustrated using the second panel of figure 2.1 (chapter 2) replicated as follows:

Figure 4.1: The Principle of Estimation using Indirect Inference



Suppose, as in the case of testing, that we have chosen two parameters of the auxiliary model to describe the reality and the real-data-based estimates of these are given at R. Suppose for now the structural model under estimation has two potential sets of parameter values (vectors A and B), each accordingly implies a joint distribution of the descriptive parameters of the auxiliary model shown by the 'mountains' (α and β). Since the contours of these distributions show the mean of β is closer to R compared to that of α, B is the more preferred parameter set to A for the structural model. Of course in practice there are numerous potential sets of structural parameters and normally one would consider more than two parameters of the auxiliary model. Thus the full estimation process would typically involve large number iterations based on the calculation of the Wald statistic.

To preserve comparability of the implications of the estimated models to those of the calibrated models revealed in the previous exercise, this chapter goes on using a VAR(1) as the auxiliary model and chooses as descriptors the coefficient matrix of this and the variances of the data. Of course other auxiliary models, such as VAR(2) or VAR(3), or others such as ARIMAs, are possible, both for testing and for estimation; but to produce a parsimonious description of the data a VAR(1) is generally acceptable. The pseudo data used for implying the joint distribution of the VAR(1) estimates are simulated by bootstrapping the 'structural errors' as before. When such distribution is found and compared to the real-data-based estimates the Wald statistic [2.4] is calculated and the method of Simulated Annealing is used to find the structural parameter values that deliver the (global) minimum of this. Note minimization of the non-linear formula [2.4] requires initial input values of structural parameters for numerical iterations. The exercise here uses the calibrated values (as in table 2.3, chapter 2) for these. It also restricts the parameters under estimation to be around $\pm$ 50% of their calibrated values[37]. These estimates are reported in section 4.2 in what follows.

[37] This excludes the discount factor β, the steady-state debt-to-GDP ratio and therefore the consumption-output ratio, and other parameters specifically restricted by the theory. See estimation results in the section to follow for more details.

4.2 The Performance of the Optimal Timeless Rule Model and Taylor Rule Model with Estimated Parameters

By restricting the structural parameters to be the commonly accepted calibrated values in the literature, the earlier chapters (2 and 3) have shown that the optimal timeless rule model is significantly superior to Taylor rule alternatives of the standard sort in representing the US economy, both in the Great Acceleration and in the Great Moderation. This section releases such restriction by using the method of indirect inference to estimate the most fitting values of these. The aim is to re-evaluate the competing models in the two post-war episodes on their best possible versions according to the data. Effectively it is also a check for robustness of the earlier findings under calibration.

The optimal timeless rule model under estimation:

Recall that given the baseline 'IS'-Phillips Curve setup [2.1] and [2.2] the optimal timeless rule model is closed by the optimality condition [2.3] implicitly derived under the principle of social welfare loss minimization. This model was shown to fit the data well in both post-war episodes when the parameters of it were calibrated. Yet it would be right to expect the estimated version of it building on indirect inference and Simulated Annealing as just described would perform no worse than this. This is because when calibrated values are set as initial guess for the structural parameters, the SA method will start searching from these values and replace them with more appropriate ones whenever a smaller Wald statistic can be found. The process will terminate only when the Wald statistic cannot be anymore smaller, and the best estimates of structural parameters are found from this. Hence, if with calibrated parameters the model is not rejected, of course it would not be rejected either with the improved, estimated values. Indeed, since the SA method here stands effectively as a way of 'fine-tuning' the calibrations, it is expected that the estimated model, being more precise from the data's point of view, would be less rejected. Table 4.1 below reports the estimates of the timeless rule model in both the pre-break and post-break episodes. Evaluation of these is shown in table 4.2.

Table 4.1: SA Estimates of the Optimal Timeless Rule Model

Parameters	Definitions	Calibrations	SA Estimates Pre-1982	SA Estimates Post-1982
β	time discount factor	0.99	fixed	fixed
σ	inverse of elasticity of intertemporal consumption	2	1.01	1.46
η	inverse of elasticity of labour	3	2.04	3.23
ω	Calvo contract price non-adjusting probability	0.53	0.79	0.54
G/Y	steady-state government expenditure to output ratio	0.23	fixed	fixed
Y/C	steady-state output to consumption ratio	1/0.77	fixed	fixed
κ	$\kappa = \frac{(1-\omega)(1-\omega\beta)}{\omega}$	0.42	0.06	0.40
γ	$\gamma = \kappa(\eta + \sigma\frac{Y}{C})$	2.36	0.19	2.06
α	relative weight of loss assigned to output variations (against inflation)	0.39	0.20	0.58
$\alpha/\gamma \equiv \theta^{-1}$	parameter driving the optimal timeless policy	1/6	1/0.95	1/3.6
θ	price elasticity of demand	6	0.95	3.6
ρ_v	autoregressive coefficient of demand disturbance (pre-)	0.88	0.92	----
ρ_{u^w}	autoregressive coefficient of supply disturbance (pre-)	0.91	0.86	----
ρ_ξ	autoregressive coefficient of policy disturbance (pre-)	0.59	0.14	----
ρ_v	autoregressive coefficient of demand disturbance (post-)	0.93	----	0.94
ρ_{u^w}	autoregressive coefficient of supply disturbance (post-)	0.80	-----	0.79
ρ_ξ	autoregressive coefficient of policy disturbance (post-)	0.38	-----	0.42

Table 4.2: Performance of the Timeless Rule Model under Calibration and Estimation

	Pre-1982 under calibration	Pre-1982 under estimation	Post-1982 under calibration	Post-1982 under 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

Overall, table 4.1 shows the estimated timeless rule model is reasonably close to its calibrated version, especially in the episode of the Great Moderation. Yet compared to the earlier episode, the estimation suggests the elasticity of intertemporal consumption (the inverse of σ) and that of labour (the inverse of η) were both slightly lowered after the break, and the same was true for the Calvo contract non-adjusting parameter (ω). Yet via the cumulating effect the latter implies a much steeper Phillips curve in contrast to the first episode as γ had deepened. The estimates also suggest the relative weight the Fed put on output variations (α) was doubled in the Great Moderation. However given that nominal rigidity has significantly reduced the Fed was forced to trade off even greater output growth when excess inflation was shown (since α/γ falls); this could also imply a surge in the price elasticity of demand (θ)[38]. The full-model based estimates of shock persistency generally resemble those calculated basing on calibrated parameters, although on this occasion the data suggest an even quicker die-out of policy shocks pre-break.

Table 4.2 confirms that this estimated model can fit the data even better as expected, as the reported Wald percentiles of chosen features (except that accounts solely for data volatility in the Moderation episode that appears equally well as under calibration) have all decreased significantly both before and after the break. Overall the Full Wald statistic shows the model fails to be strongly rejected at 95% and would well explain the US data post Bretton Woods up to around 70% (Full Wald of 71.7 in the Great Acceleration and 68.6 when the Great Moderation prevailed).

The above thus implies if the Fed was constantly pursuing the optimal timeless policy its behaviour would look 'tougher' against inflation relative to output variation in the second episode as a result of adaptation to structural economic change. Clearly this is somewhat different from what was assumed in the earlier evaluation exercise (chapter 3) where under calibration no parametric movement was allowed for. However, given that both structural parameters and policy and the size of shocks could change across episodes as the data suggest they did counterfactual experiment would still show the fall in data variability was dominated by the fall in shocks. This result is reported in

[38] Footnote 19, chapter 2.

table 4.3 as follows; it suggests the reduction of shocks dominated the pro-cyclical movement of interest rates induced by the change in the structure and policy and caused the Great Moderation in this; this f all also accounted for 76% of output moderation and 89% of inflation moderation, whereas the tougher policy only contributed to 11% of this last. Hence the best-fitting timeless rule model where parametric variation is allowed for would too attribute the reduced data variability mostly to reduced shocks, much in the same vein as it did with fixed calibrated parameters where such reduction accounted for all.

Table 4.3: Accountability of Factor Variation for Reduced Data Volatility
(Timeless rule model under estimation)

Reduced data volatility caused by	Interest rates	Output gap	Inflation
reduced shocks	100%	76.4%	89.2%
chg in policy paras	pro-cyclical	pro-cyclical	10.8%
chg in structural paras	pro-cyclical	23.6%	pro-cyclical

Taylor rule model under estimation:

The above can be compared to the more widespread model version where a Taylor-type policy is substituted for. To this end this subsection replaces the optimal timeless policy with equation [4.1] as follows and estimates this under the identical framework of 'IS' and Phillips Curve equations and evaluates for that.

$$\tilde{i}_t = (1-\rho)[\gamma_\pi \pi_t + \gamma_x x_t] + \rho \tilde{i}_{t-1} + \xi_t \qquad [4.1]$$

Note [4.1] is essentially the specification of the 'interest-rate-smoothed' Taylor rule [1.2] the earlier chapters have considered. Yet as estimation goes it also covers the other standard Taylor-type policies the preceding have reviewed, as when ρ is zero it reduces to [1.1] as Taylor (1993) originally suggested while when γ_π is just above unity it becomes a 'weak' Taylor rule. The foregone have shown that under calibration the US data strongly reject the 'weak' Taylor rule (as specified in table 3.4, chapter 3) in the Great Acceleration and the original Taylor rule and its 'interest-rate-smoothed' version in the Great Moderation. Table 4.4 in what follows uses the SA method to search for each episode the best model version of this type. In particular it allows ρ to take any value between 0 and 1 such that both the original and the 'interest-rate-smoothed' Taylor rules are possible. It also sets the lower bound of γ_π to 1.001 to account for the possibility of 'weak' Taylor rule. The estimated model is evaluated in table 4.5.

Table 4.4: SA Estimates of the Taylor Rule Model

Parameters	Definitions	Calibrations	SA Estimates Pre-1982	Post-1982
β	time discount factor	0.99	fixed	fixed
σ	inverse of elasticity of intertemporal consumption	2	1.15	1.16
η	inverse of elasticity of labour	3	2.66	3.85
ω	Calvo contract price non-adjusting probability	0.53	0.79	0.61
G/Y	steady-state government expenditure to output ratio	0.23	fixed	fixed
Y/C	steady-state output to consumption ratio	1/0.77	fixed	fixed
κ	$\kappa = \frac{(1-\omega)(1-\omega\beta)}{\omega}$	0.42	0.06	0.25
γ	$\gamma = \kappa(\eta + \sigma\frac{Y}{C})$	2.36	0.23	1.33
γ_π	interest rates response to inflation	1.44	2.03	2.06
γ_x	interest rates response to output gap	0.14	0.001	0.06
ρ	interest-rate-smoothing parameter	0.76	0.42	0.89
ρ_v	autoregressive coefficient of demand disturbance (pre-)	n/a	0.91	-----
ρ_{u^w}	autoregressive coefficient of supply disturbance (pre-)	n/a	0.87	-----
ρ_ξ	autoregressive coefficient of policy disturbance (pre-)	n/a	0.58	-----
ρ_v	autoregressive coefficient of demand disturbance (post-)	0.93	-----	0.95
ρ_{u^w}	autoregressive coefficient of supply disturbance (post-)	0.80	-----	0.77
ρ_ξ	autoregressive coefficient of policy disturbance (post-)	0.39	-----	0.40

Table 4.5: Performance of the Taylor Rule Model under Calibration and Estimation

	Pre-1982 under		Post-1982 under	
	calibration	estimation	calibration	estimation
Directed Wald (for dynamics)	n/a	98	99.8	89.6
Directed Wald (for volatilities)	n/a	40.6	99	94.9
Full Wald (for dynamics & volatilities)	n/a	96.1	99.7	87.6

Table 4.4 shows in general the estimated Taylor rule model is not very different from its calibrated version, either. The elasticity of intertemporal consumption (the inverse of σ) and that of labour (the inverse of η) are found similar to what just estimated with the optimal timeless policy, although in this case the data suggest no significant movement of the former. The Calvo contract non-adjusting probability (ω) is still high in the first episode compared to the second. Again, this through the cumulating effect implies a much steeper Phillips curve when the Great Moderation prevailed (γ). The estimate of ρ shows if a Taylor rule was operating it must be one where interest rates were 'smoothed'; but compared to the earlier episode the degree of 'policy inertia' in the later was doubled. Yet the Fed's response to inflation (γ_π) was essentially unaltered and remained strong. This precludes indeterminacy in either episode for that the policy was far from being 'weak' or 'passive'[39]. The estimation also suggests quite persistent demand and supply shocks in contrast to policy errors both before and after the break. This last is consistent with what the benchmark calibrations would predict.

Table 4.5 shows under estimation the Taylor rule model can, too, explain the US data post-1972. This can be seen from the reported Wald percentiles that given the SA estimates the model is saved from being strongly rejected in the Great Moderation

[39] Recall in estimation γ_π is allowed to take as low as 1.001.

(with Full Wald of 87.6) while in the Acceleration episode just marginally rejected at 95% (with Full Wald of 96.1). Hence loosely one might still argue that there exists a Taylor rule model whose dynamic behaviour could mimic the truth.

This is not supporting the conventional 'good policy' explanation to the Great Moderation, however, even it does show the Fed's post-war policy could well be one of Taylor-type. This is because the strong Taylor rule the data suggest in the Acceleration episode has completely violated the underlying assumption of the good-policy story in that indeterminacy being caused by passive policy would not ever arise. Indeed, given that the Fed's policy was so 'active' both before and after the break and that the change in economic structure had failed to cause fundamental contraction of data variability, this estimated model—while implying that a Taylor rule was possible—would too suggest the Great Moderation was mainly caused by reduced shocks as table 4.6 shows it was[40].

Table 4.6: Accountability of Factor Variation for Reduced Data Volatility
(Taylor rule model under estimation)

Reduced data volatility caused by	Interest rates	Output gap	Inflation
reduced shocks	87.4%	67.6%	19.3%
chg in policy paras			
γ_π	0.1%	pro-cyclical	5.1%
γ_x	pro-cyclical	0.4%	pro-cyclical
ρ	12.5%	2.9%	75.7%
chg in structural paras	pro-cyclical	29.1%	pro-cyclical

[40] While the last column of table 4.6 shows over 75% of inflation moderation was caused by the Fed's deepened degree of policy inertia, it shall be emphasized that in the conventional good policy/good shocks debate over the causes of the Great Moderation the former is referred particularly to the Fed's *feedback* responses such as γ_π and γ_x. Clearly these compared to the reduced shocks could hardly be decisive in lowering the data's variability according to the table, although in the last column it shows the change of γ_π would play a non-negligible role.

Concluding remarks on the comparison of the estimated timeless rule model and Taylor rule model:

What the above have shown is that even when estimated parameters are allowed for the optimal timeless rule model is still overwhelmingly superior to the Taylor rule alternative in explaining the US post-war data, although in this case the latter is not shown to be strongly rejected. This extended exercise thus supports the earlier findings basing on standard calibrations that the Fed's post-war monetary policy was better understood as the timeless optimum instead of a Taylor rule and that it was the improved environment, the reduced shocks that fundamentally caused the Great Moderation. The estimation also shows if a Taylor rule were operating, for both episodes it must be one in which active interest-rate response to inflation could have been found. Thus from the data's viewpoint evidence in favour of the good-policy story that requires a weak/passive policy in the first episode does not exist.

4.3 Other Issues on the Point of Robustness

4.3. 1 The choice of auxiliary model

Clearly, the above comparison between the optimal timeless rule model and Taylor rule model, like any other empirical evaluation based on data, would involve inevitably various kinds of issues regarding robustness. One of these in the context here where the method of indirect inference is used is the choice of auxiliary time-series model that provides benchmark description of the data against which theoretical models are estimated and evaluated indirectly. This has been a VAR(1) in the foregone exercise where the descriptors chosen are the estimate of its coefficient matrix and the data's volatility. Yet as stated a VAR of higher order or time-series models of other types are completely possible, depending on what and the extent to which one requires the model to fit.

This subsection proceeds by checking how robust the above ranking of models, i.e., the superiority of the timeless rule model over the Taylor rule alternative, is to the choice of differing orders of a VAR. This assessment can in principle be done in regards both of model estimation and of model evaluation, but here it restricts itself to the latter of these for simplicity basing on the estimates just obtained. Clearly, using a VAR of higher orders as auxiliary model will make the test more demanding as effectively more detailed features of the data are asked to fit. Thus practically this is also a way of further discriminating between competing theories whose performances are hardly distinguishable under parsimonious auxiliaries. This is not the purpose on this occasion, though, since in the above it showed the timeless rule model was significantly better than the Taylor rule alternative when a VAR(1) was used. Yet as robustness is concerned the focus here is whether such ranking will be overturned when a VAR of higher orders is substituted for.

This is shown not to be the case, however, according to table 4.7 when the auxiliary model chosen is VAR(2) or VAR(3). Indeed, while the reported Wald percentiles suggest increasing the order of VAR would render strong rejection in most cases both for the timeless rule model and for the Taylor rule alternative due to surged burden laid on the models[41], it is shown by the normalized t statistics (in parentheses) that in all cases the former is always less rejected, thus more preferred compared to the latter regardless of the order of the VAR. In other words the ranking of these models is robust to such choice.

Table 4.7: Performance of Models under Differing Auxiliaries

(Panel A: pre-1982, the Great Acceleration)

	VAR(2)		VAR(3)	
	Timeless optimum	Taylor rule	Timeless optimum	Taylor rule
Directed Wald for dynamics	99.7	100	100	100
(Normalized t-stat)	*(3.92)*	*(14.0)*	*(4.86)*	*(15.8)*
Directed Wald for volatilities	66.5	84.8	86.2	81.2
(Normalized t-stat)	*(-0.12)*	*(0.44)*	*(0.48)*	*(0.25)*
Full Wald for both	99.9	100	100	100
(Normalized t-stat)	*(4.19)*	*(13.0)*	*(4.78)*	*(14.6)*

41 Most likely the 'extra loan' related to the dynamic features, caused by extra lags of the VAR.

(Panel B: post-1982, the Great Moderation)

	VAR(2)		VAR(3)	
	Timeless optimum	Taylor rule	Timeless optimum	Taylor rule
Directed Wald for dynamics	99.9	100	100	100
(Normalized t-stat)	*(4.33)*	*(9.38)*	*(10.1)*	*(13.7)*
Directed Wald for volatilities	93.7	99.9	90.7	100
(Normalized t-stat)	*(1.41)*	*(6.59)*	*(1.06)*	*(6.37)*
Full Wald for both	100	100	100	100
(Normalized t-stat)	*(4.87)*	*(12.1)*	*(9.80)*	*(15.0)*

4.3.2 The choice of cut-off between the Great Acceleration and the Great Moderation

Besides the choice of auxiliary model, another factor by which the above result of episode comparison might have been affected is the choice of breakpoint on the data, i.e., the cut-off between the Great Acceleration and the Great Moderation. The forgone has followed the common practice of setting this to 1982—the year around which many (including Bernanke and Mihov, 1998, and Clarida, Gali and Gertler, 2000) believe the Fed switched from using non-borrowed reserves to setting the Funds rate as the instrument of monetary policy; yet a choice also supported by the Qu-Perron test that found the 95% interval between 1980 and 1984 (Table 3A.1 in the last chapter). Indeed, given that historically the US economy improved in a continuous manner during the start of the 1980s, any point within this range should in principle be considered appropriate from the data's point of view. Yet few would see 1980 as the real start of the new era as it was around when Paul Volcker ploughed into the business of combating the stagflation crisis of the 1970s. 1984 is a widely accepted alternative of this, as by then inflation was successfully controlled at around 3.5% (annual rate). This is also consistent with the Qu-Perron estimate of the best breakpoint at the third quarter of the year.

The purpose of this subsection is to examine how robust the earlier result of episode/model comparison is to the choice of this break. Table 4.8 reveals the SA estimates of both the timeless rule model and the Taylor rule model (panel A for the first and panel B for the second) when the cut-off is made at 1984Q3 and compares these to the benchmark estimates (the parenthesized) when 1982 was used. It shows in neither case the change of cut-off would cause substantial difference in terms of model estimation, although as evaluation is concerned the Taylor rule model would outperform the timeless rule version in fitting the data's volatility in the first episode and their dynamic features in the second (table 4.9). However according to the Full Wald statistics the timeless rule model is still more preferred overall (92.6 vs 93.9 in the Acceleration and 92.3 vs 96.9 in the Moderation), and its superiority over the Taylor rule alternative would be made more apparent once VAR(2) or VAR(3) is chosen as the auxiliary for more precise comparison/evaluation—this last is verified by the normalized t statistics reported in table 4.10 that show for most criteria much less rejection is found under the timeless rule. Hence a lthough compared to the benchmark result splitting the data in 1984 is less in favour of the timeless rule model and more of the Taylor rule alternative, the extended exercise here shows that strictly the previous ranking of models is robust, too, to this choice.

Table 4.8: SA Estimates of Models when Cut-off at 1984

(Panel A: model with optimal timeless policy)

Parameters	Definitions	SA Estimates Pre-1984 *(-82)*		SA Estimates Post-1984 *(-82)*	
β	time discount factor	---------------fixed at 0.99---------------			
σ	inverse of elasticity of intertemporal consumption	1.01	*(1.01)*	2.67	*(1.46)*
η	inverse of elasticity of labour	1.54	*(2.04)*	2.53	*(3.23)*
ω	Calvo contract price non-adjusting probability	0.79	*(0.79)*	0.48	*(0.54)*
G/Y	steady-state government expenditure to output ratio	---------------fixed at 0.23---------------			
Y/C	steady-state output to consumption ratio	-------------fixed at 1/0.77-------------			
κ	$\kappa = \frac{(1-\omega)(1-\omega\beta)}{\omega}$	0.06	*(0.06)*	0.57	*(0.40)*
γ	$\gamma = \kappa(\eta + \sigma\frac{Y}{C})$	0.17	*(0.19)*	3.41	*(2.06)*
α	relative weight of loss assigned to output variations (against inflation)	0.20	*(0.20)*	0.58	*(0.58)*
$\alpha/\gamma \equiv \theta^{-1}$	parameter driving the optimal timeless policy	1/0.85	*(1/0.95)*	1/5.9	*(1/3.6)*
θ	price elasticity of demand	0.85	*(0.95)*	5.9	*(3.6)*

ρ_v	autoregressive coefficient of demand disturbance	0.89 *(0.92)*	0.94 *(0.94)*
ρ_{u^w}	autoregressive coefficient of supply disturbance	0.87 *(0.86)*	0.84 *(0.79)*
ρ_ξ	autoregressive coefficient of policy disturbance	0.18 *(0.14)*	0.36 *(0.42)*

(Panel B: model with Taylor rule)

Parameters	Definitions	SA Estimates Pre-1984 *(-82)*	Post-1984 *(-82)*
β	time discount factor	--------------fixed at 0.99--------------	
σ	inverse of elasticity of intertemporal consumption	1.00 *(1.15)*	2.83 *(1.16)*
η	inverse of elasticity of labour	2.42 *(2.66)*	3.40 *(3.85)*
ω	Calvo contract price non-adjusting probability	0.79 *(0.79)*	0.64 *(0.61)*
G/Y	steady-state government expenditure to output ratio	--------------fixed at 0.23--------------	
Y/C	steady-state output to consumption ratio	-------------fixed at 1/0.77------------	
κ	$\kappa = \frac{(1-\omega)(1-\omega\beta)}{\omega}$	0.06 *(0.06)*	0.21 *(0.25)*
γ	$\gamma = \kappa(\eta + \sigma\frac{Y}{C})$	0.22 *(0.23)*	1.46 *(1.33)*
γ_π	interest rates response to inflation	2.10 *(2.03)*	1.34 *(2.06)*
γ_x	interest rates response to output gap	0.006 *(0.001)*	0.09 *(0.06)*
ρ	interest-rate-smoothing parameter	0.63 *(0.42)*	0.83 *(0.89)*
ρ_v	autoregressive coefficient of demand disturbance	0.89 *(0.91)*	0.94 *(0.95)*
ρ_{u^w}	autoregressive coefficient of supply disturbance	0.88 *(0.87)*	0.80 *(0.77)*
ρ_ξ	autoregressive coefficient of policy disturbance	0.60 *(0.58)*	0.51 *(0.40)*

Table 4.9: Performance of Models when Cut-off at 1984

	Pre-1984 with		Post-1984 with	
	Timeless optimum	Taylor rule	Timeless optimum	Taylor rule
Directed Wald (for dynamics)	92.4	95.8	96.4	88.2
Directed Wald (for volatilities)	82.8	27.1	2.3	99.3
Full Wald (for dynamics & volatilities)	92.6	93.9	92.3	96.9

Table 4.10: Performance of Models under Differing Auxiliaries (II)

(Panel A: pre-1984, the Great Acceleration)

	VAR(2)		VAR(3)	
	Timeless optimum	Taylor rule	Timeless optimum	Taylor rule
Directed Wald for dynamics	99.9	100	100	100
(Normalized t-stat)	*(6.03)*	*(14.3)*	*(5.56)*	*(16.4)*
Directed Wald for volatilities	73.6	88.8	90.8	89.5
(Normalized t-stat)	*(0.05)*	*(0.86)*	*(1.02)*	*(0.90)*
Full Wald for both	99.9	100	100	100
(Normalized t-stat)	*(6.01)*	*(15.0)*	*(5.61)*	*(15.1)*

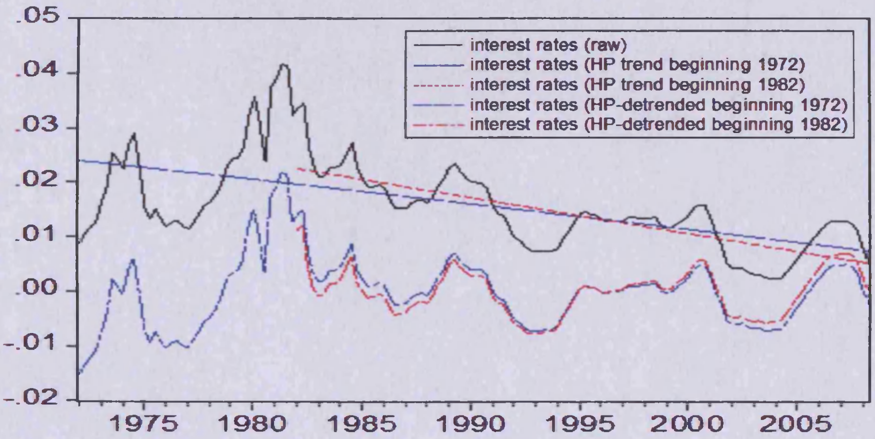
(Panel B: post-1984, the Great Moderation)

	VAR(2)		VAR(3)	
	Timeless optimum	Taylor rule	Timeless optimum	Taylor rule
Directed Wald for dynamics	100	100	100	100
(Normalized t-stat)	*(5.12)*	*(14.8)*	*(11.2)*	*(33.3)*
Directed Wald for volatilities	42	99.9	58.9	99.7
(Normalized t-stat)	*(-0.56)*	*(5.04)*	*(-0.16)*	*(4.16)*
Full Wald for both	100	100	100	100
(Normalized t-stat)	*(5.08)*	*(14.7)*	*(11.1)*	*(31.0)*

4.3. 3 The choice of method of data stationarization

So far the preceding have shown that the earlier finding of superiority of the timeless rule model over the Taylor rule alternative is robust both to the choice of auxiliary model and to differing cut-offs of the data. These exercises were conducted using stationarised time series where a linear trend, if any, was filtered away. Known as 'linear detrending', this method has the advantage of preserving data information that might otherwise be excluded as trend component when other alternatives (such as HP-filtering or first-order differencing) are used. The drawback of this, however, is that the trend and therefore the stationarised time series for an interval given can look quite different as the observation window varies, creating inconsistency in the data practically used in empirical work—the robustness check above where differing breakpoints were considered was one victim of this. Figure 4.2 uses the interest rates series starting respectively in 1972 and in 1982 as an example to illustrate the problem. It shows the two observation windows have suggested quite different linear trends of the interest rates; the result is that for comparable episode the detrended series are in general inconsistent.

Figure 4.2: Inconsistency of Linear-detrended Time-series (Interest Rates)



One possible solution to the above problem (and indeed a popular alternative to the method of linear-detrending) is to use the Hodrick-Prescott filter where when the original data are dealt with a smoothed, time-varying trend (instead of a deterministic one) is filtered away. In general the smoother the time-varying trend is assumed the less the detrended data would be intruded by the observation window; the cost, however, is that the less information they would contain. So there is a trade-off[42]. Figure 4.3 follows the common practice of setting (for quarterly data) the 'smoothing

[42] Note at one extreme when the degree of smoothness is infinitely small the trend it yields converges to a deterministic, thus a linear one.

parameter' λ to 1600 and replicates the detrending exercise just done; it is clear that the difference between the two filtered series caused by the observation windows chosen is negligible.

Figure 4.3: Inconsistency of HP-detrended Time-series (Interest Rates)

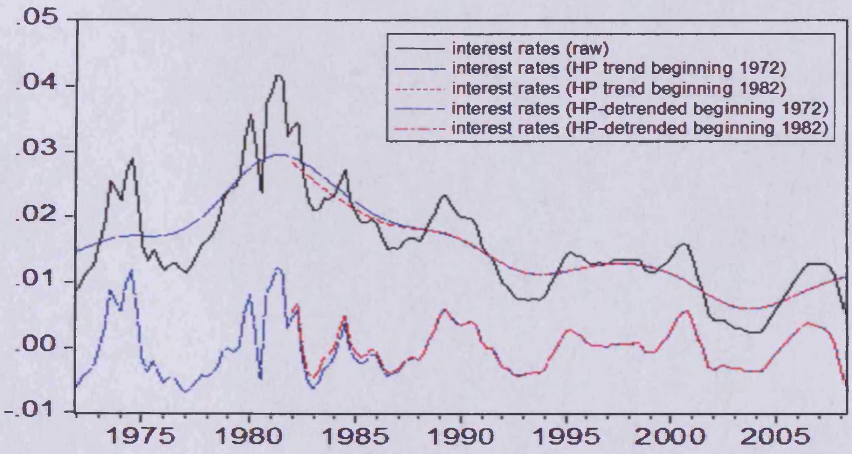


Table 4.11 in what follows re-compares the models' performance for the cut-offs considered (1982 in panel A and 1984 in panel B) with the original data stationarised using the HP trend[43]. It shows once this method is chosen both the timeless rule model and the Taylor rule alternative would have more difficulty in fitting the data, mostly because their dynamic performance exacerbates. Yet with either cut-off this fails to change the earlier ranking of these in the Moderation episode as the Wald statistics (or more clearly, the normalized t statistics) show, although given the standard auxiliary model of VAR(1) here their difference in the first episode is hardly distinguishable. However one can further discriminate between these using higher orders of VAR, and once a VAR(2) or VAR(3) is substituted for it shows the timeless rule model in both episodes significantly outperforms (See extended results of these in tables 4A.2 and 4A.3). This shows the foregone model ranking basing on linear-detrended data is also robust to data stationarized using the HP filter.

[43] SA estimates of models reported in table 4A.1 in chapter appendix.

Table 4.11: Performance of Models when Data Are HP-filtered

(Panel A: with breakpoint chosen at 1982)

	Pre-1982 with		Post-1982 with	
	Timeless optimum	Taylor rule	Timeless optimum	Taylor rule
Directed Wald for dynamics	97.3	97.8	99.8	99.9
(Normalized t-stat)	*(2.14)*	*(2.27)*	*(3.77)*	*(4.19)*
Directed Wald for volatilities	34.8	29.1	59.5	99.5
(Normalized t-stat)	*(-0.65)*	*(-0.81)*	*(-0.13)*	*(4.52)*
Full Wald for both	96.5	96.6	99.7	100
(Normalized t-stat)	*(1.94)*	*(1.90)*	*(3.34)*	*(5.76)*

(Panel B: with breakpoint chosen at 1984)

	Pre-1984 with		Post-1984 with	
	Timeless optimum	Taylor rule	Timeless optimum	Taylor rule
Directed Wald for dynamics	99.7	99.7	88.5	96.0
(Normalized t-stat)	*(3.51)*	*(3.24)*	*(1.08)*	*(1.81)*
Directed Wald for volatilities	43.8	97.1	3.7	90.1
(Normalized t-stat)	*(-0.59)*	*(2.14)*	*(-1.61)*	*(0.94)*
Full Wald for both	99.2	99.4	79.7	95.8
(Normalized t-stat)	*(3.23)*	*(3.11)*	*(0.64)*	*(1.76)*

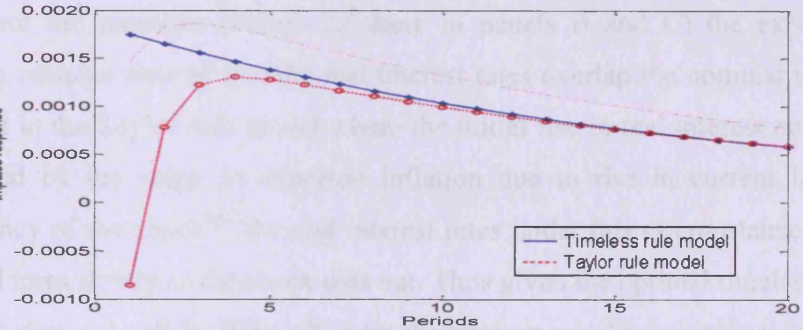
4.4 Implementation of the Optimal Timeless Policy: how different is this from a Taylor rule?

All the above thus have re-endorsed from the data's viewpoint that the Fed's post-war monetary policy was better understood as the timeless optimum rather than a Taylor rule. We have seen in chapter 3 (figure 3.3) that implementing the optimal timeless policy would result in recursiveness in interest rates determination as to maintain the optimality condition [2.3] the Fed is forced to set real interest rates to an adaptive level that supports the equilibrium trade-off between inflation and the output gap (or growth). This generates quite different economic dynamics that would otherwise be caused if interest rates were set actively using a Taylor rule to determine equilibrium.

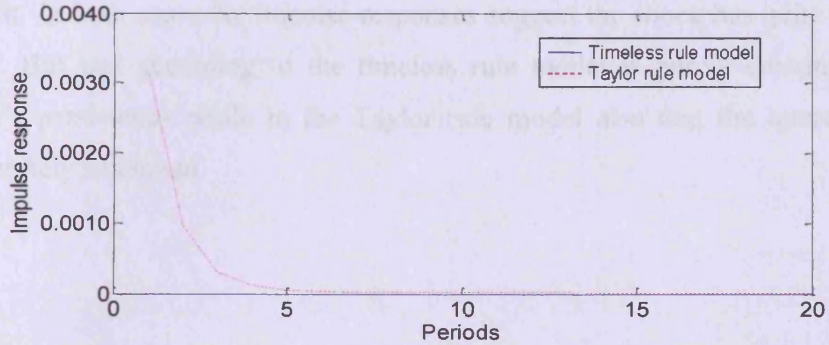
Figure 4.4 to 4.6 in what follows use the model versions estimated for the post-1982 episode (table 4.1 with optimal timeless policy and table 4.2 with Taylor rule) as an example to illustrate how differently the economy would response to a (one-standard-error) unit shock caused by the demand side, the supply side and the monetary authority. The responses of the timeless rule model are shown with solid lines; the dashed lines indicate those of the Taylor rule model.

Figure 4.4: Impulse Responses to a Unit Shock to Demand

Panel A: Interest rates (nominal, unmarked; real, marked)



Panel B: Output gap



Panel C: Inflation

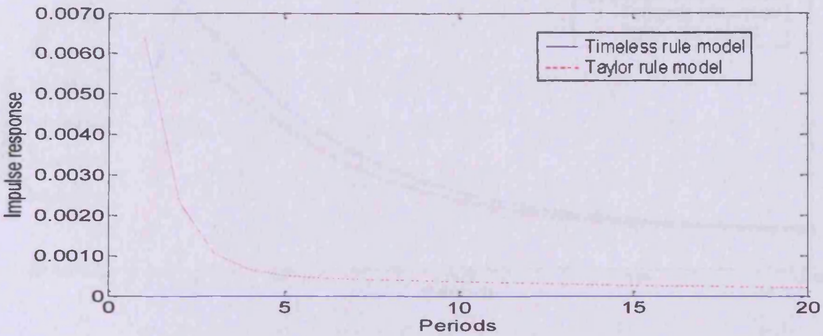
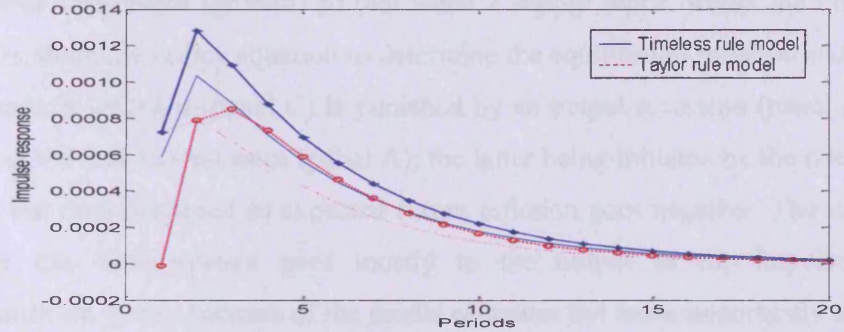


Figure 4.4 shows an increase in aggregate demand raises nominal interest rates both in the timeless rule model and in the Taylor rule model to similar levels. Under the optimal timeless policy where output and inflation are determined solely by the Phillips curve and the policy (the recursiveness feature just mentioned, and yet can be seen from the unresponsiveness of these in panels B and C) the expected future inflation remains zero so that the real interest rates overlap the nominal rates. This is different in the Taylor rule model where the initial rise in real interest rates is largely weakened by the surge in expected inflation due to rise in current inflation and persistency of the shock[44]; the real interest rates under this circumstance pick up the nominal rates slowly as the shock dies out. Thus given the optimal timeless policy any shock to demand will be fully offset by the adjustment of nominal/real interest rates, leaving the rest of the system intact, whereas when a Taylor-type policy is substituted for the shock spreads out as a result of inadequate movements of the real rates of interest. In both cases the impulse responses suggest the shock has quite long-lasting effect. But this according to the timeless rule model is purely determined by the shock's persistency while in the Taylor rule model also that the interest rates are deliberately smoothed.

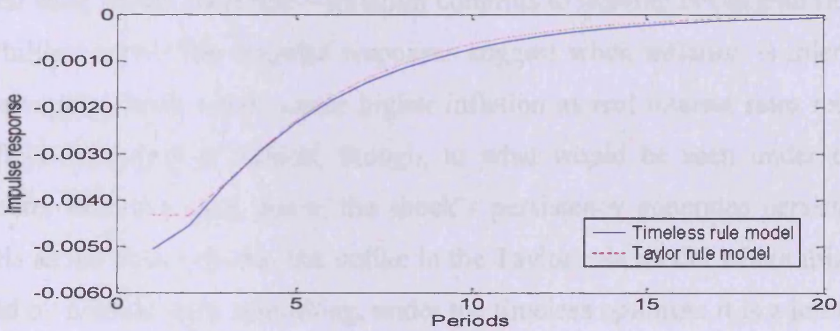
[44] Note output also rises according to the Phillips curve, the dashed line in panel B.

Figure 4.5: Impulse Responses to a Unit Shock to Supply

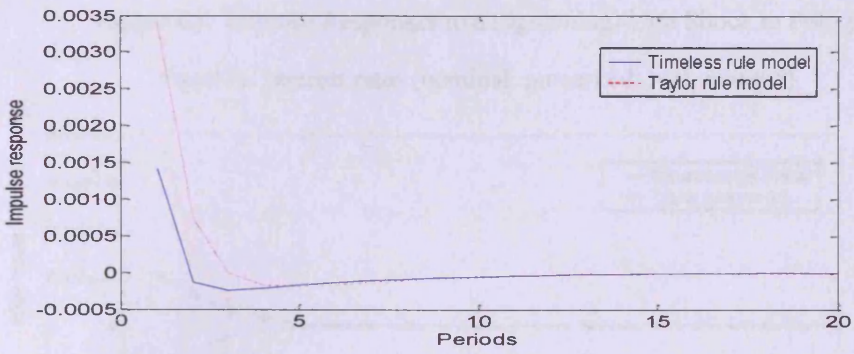
Panel A: Interest rates (nominal, unmarked; real, marked)



Panel B: Output gap



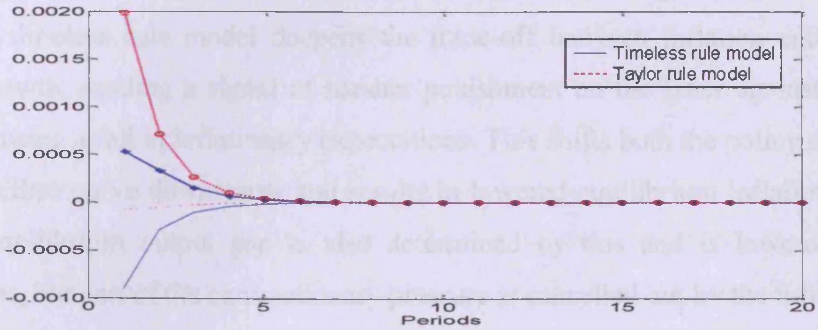
Panel C: Inflation



A shock to aggregate supply shifts the Phillips curve upwards, worsening the trade-off between inflation and the output gap. In either model this raises inflation and interest rates (both nominal and real) and causes an output recession as figure 4.5 illustrates. Yet with differing policies the shock exhibits clear distributional difference according to the magnitude of the impulse responses: under the optimal timeless policy commitment of [2.3] requires keeping inflation a fixed fraction of (the first difference of) the output gap. This in effect constitutes another 'optimal trade-off' between inflation and output (growth) so that when a supply shock occurs the Phillips curve moves along the policy equation to determine the equilibrium inflation and output; the increase in inflation (panel C) is punished by an output recession (panel B) made by raising the real interest rates (panel A); the latter being initiated by the rise in nominal rates but then deepened as expected future inflation goes negative. The supply shock under this circumstance goes mostly to the output as the impulse responses demonstrate, partly because of the model estimates but more importantly that inflation is bound by the optimal plan. When this is replaced by the Taylor rule, however, the shock spreads out more evenly as—except being suppressed primly by the real interest rates set by such rule—inflation commits to nothing but determined solely by the Phillips curve. The impulse responses suggest when inflation is tolerated in this way a supply shock would cause higher inflation as real interest rates response less; its effect on output is similar, though, to what would be seen under the timeless optimum. In either case, again, the shock's persistency generates persistency of the models as the figure shows, but unlike in the Taylor rule model where this is partially caused by interest rates smoothing, under the timeless optimum it is a joint result with the optimal trade-off[45].

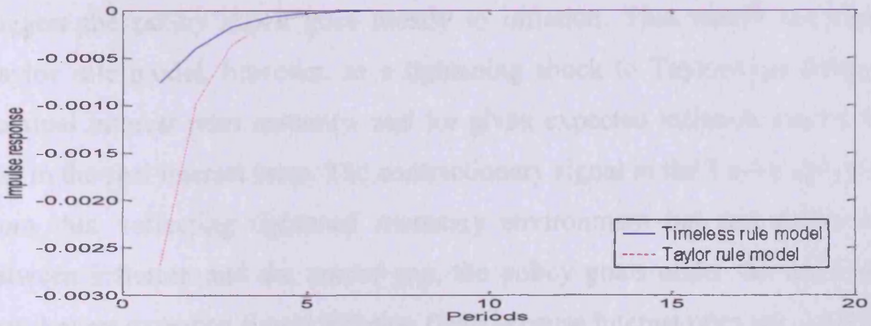
Figure 4.6: Impulse Responses to a (tightening) Unit Shock to Policy

Panel A: Interest rates (nominal, unmarked; real, marked)



[45] This last does not hold if the optimal trade-off were between inflation and the *level* of output rather than its growth, i.e., the optimality condition under discretion where the lag of output is not involved (Walsh (2003) provides a neat discussion on this).

Panel B: Output gap



Panel C: Inflation

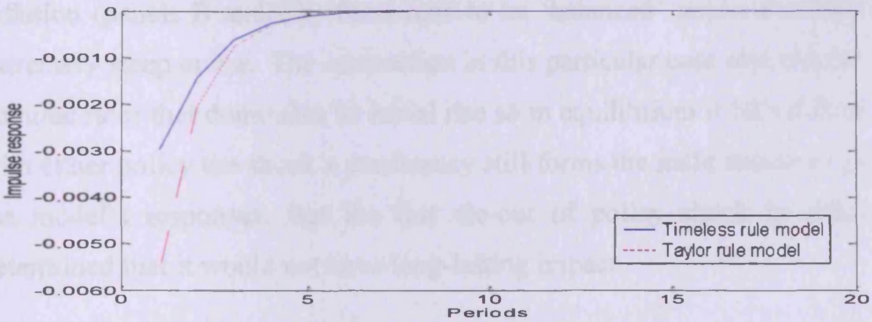


Figure 4.6 shows finally the models' impulse responses to a tightening shock to the monetary policy. In the foregone context this has been interpreted as a 'trembling hand' error made by the policy maker in execution of the preset monetary rule. But plainly due to differing natures of the optimal timeless policy and the Taylor-type alternative its connotation in the two models is different. A tightening monetary shock to the timeless rule model deepens the trade-off between inflation and the output gap/growth, sending a signal of harsher punishment on the latter against the former and causing a fall in inflationary expectations. This shifts both the policy equation and the Phillips curve downwards and results in lowered equilibrium inflation (panel C). The equilibrium output gap is also determined by this and is lowered as policy tightens, but part of the contractionary pressure is cancelled out by the fall of expected future inflation that encourages current production so the actual fall of it is small (panel B). Panel A shows to support this equilibrium the real interest rates must rise. But according to the fall of inflationary expectations this is made not by raising but

slightly lowering the nominal interest rates[46]. The impulse responses in this case thus suggest the policy shock goes mostly to inflation. This would not happen to the Taylor rule model, however, as a tightening shock to Taylor-type policy raises the nominal interest rates instantly, and for given expected inflation causes a temporary rise in the real interest rates. The contractionary signal in the Taylor rule model is sent from this, reflecting tightened monetary environment but not deepened trade-off between inflation and the output gap, the policy goals under the optimal rule. This then lowers expected future inflation (here because interest rates are smoothed and the shock is persistent) and further raises the real interest rates (panel A), causing a strong reduction in equilibrium output and correspondingly a strong reduction in equilibrium inflation (panels B and C)—these tend to be 'balanced' unless the Phillips curve is extremely steep or flat. The contraction in this particular case also causes a fall in the nominal rates that dominates its initial rise so in equilibrium it falls a little[47]. However, with either policy the shock's persistency still forms the main source of persistency in the model's responses. But the fast die-out of policy shock in either sense has determined that it would not have long-lasting impact.

To sum up, implementation of the optimal timeless policy has helped directing differing shocks into different sectors of the economy, facilitating the Fed in stabilization in that the causes of instability are easier to be identified and eliminated. Compared to a Taylor rule that specifies systematic interest-rates responses, the timeless optimum advocates active adjustment of these to ensure the policy outcome is always at the least cost. Effectively this trades the volatility of policy instrument, the nominal interest rates, with those of the policy objectives, i.e., output gap and inflation, that would otherwise be less stabilized as the impulse responses illustrate. Our earlier empirical assessments have suggested for both post-war episodes the Fed's behaviour came closest to the optimal timeless rule. Thus from the point of view of the history, there is reason to believe by committing to the timeless optimum the Fed had successfully circumvented some costs of monetary management and that

---

[46] Thus accompanied by an extensive downward movement of the 'IS' curve caused by the fall of expected inflation and output gap.

[47] This is largely determined by the extent to which inflationary expectations fall in response to the shock.

the US economy would have suffered greater volatility loss if a Taylor rule were substituted for.

4.5 Conclusion

This chapter extends the partial comparison of the optimal timeless rule model and Taylor rule model in chapters 2 and 3 basing on calibration to a full comparison of these bas ing on estimation. Using the method of Simulated Annealing based on indirect inference, it finds if the models' full capacity is allowed for, for each version there is a set of 'best-fitting' parameters that significantly improves its performance although compared to the calibrated values these are not of significant difference. However, this fails to change the fact that in both the episodes of the Great Acceleration and the Great Moderation the optimal timeless rule model remains overwhelmingly superior to the Taylor rule alternative, so that even under estimation it still suggests the Fed's post-war monetary policy was better understood as the optimal timeless rule. Of course, as the estimates of model vary across subsamples it would no longer imply the improvement was caused totally by improved environment. But given that the latter was still dominating its occurrence the underlying argument is yet intact. The estimation also suggests if a Taylor rule was operating, for both episodes it must be one in which the interest-rates response to inflation was so strong that the 'good policy' story of Great Moderation would not be true; it would also have generated redundant economic instability that has been successfully avoided under the optimal timeless policy. In short, the Fed's post-Bretton-Woods monetary behaviour was roughly optimal, and the US economy improved in the early 1980s because the Fed had managed to commit with less error to the optimal plan and that the economic environment had improved; this, clearly, had nothing to do with a Taylor rule.

Appendix to Chapter 4

A. Some Supplementary Exercises of Comparison between the Timeless Rule Model and Taylor Rule Model

Table 4A.1: SA Estimates of Models under HP-filtered Data

(Panel A: model with optimal timeless policy)

Parameters	Definitions	SA Estimates Pre-1982 *(-84)*		SA Estimates Post-1982 *(-84)*	
β	time discount factor	---fixed at 0.99---			
σ	inverse of elasticity of intertemporal consumption	1.01	*(1.00)*	1.10	*(1.59)*
η	inverse of elasticity of labour	1.66	*(4.41)*	2.15	*(3.79)*
ω	Calvo contract price non-adjusting probability	0.75	*(0.78)*	0.51	*(0.41)*
G/Y	steady-state government expenditure to output ratio	---fixed at 0.23---			
Y/C	steady-state output to consumption ratio	---fixed at 1/0.77---			
κ	$\kappa = \frac{(1-\omega)(1-\omega\beta)}{\omega}$	0.09	*(0.06)*	0.48	*(0.85)*
γ	$\gamma = \kappa(\eta + \sigma\frac{Y}{C})$	0.27	*(0.34)*	1.72	*(4.98)*
α	relative weight of loss assigned to output variations (against inflation)	0.24	*(0.31)*	0.53	*(0.49)*
$\alpha/\gamma \equiv \theta^{-1}$	parameter driving the optimal timeless policy	0.89	*(0.91)*	0.31	*(0.10)*
θ	price elasticity of demand	1.12	*(1.10)*	3.23	*(10)*
ρ_v	autoregressive coefficient of demand disturbance	0.79	*(0.78)*	0.91	*(0.89)*
ρ_{u^w}	autoregressive coefficient of supply disturbance	0.80	*(0.82)*	0.71	*(0.76)*
ρ_ξ	autoregressive coefficient of policy disturbance	0.07	*(0.12)*	0.22	*(0.13)*

(Panel B: model with Taylor rule)

Parameters	Definitions	SA Estimates Pre-1982 *(-84)*		Post-1982 *(-84)*	
β	time discount factor	--------------fixed at 0.99--------------			
σ	inverse of elasticity of intertemporal consumption	1.00	*(1.02)*	2.57	*(1.87)*
η	inverse of elasticity of labour	3.48	*(3.09)*	2.86	*(2.88)*
ω	Calvo contract price non-adjusting probability	0.79	*(0.55)*	0.47	*(0.49)*
G/Y	steady-state government expenditure to output ratio	--------------fixed at 0.23--------------			
Y/C	steady-state output to consumption ratio	-------------fixed at 1/0.77------------			
κ	$\kappa = \frac{(1-\omega)(1-\omega\beta)}{\omega}$	0.06	*(0.37)*	0.60	*(0.54)*
γ	$\gamma = \kappa(\eta + \sigma\frac{Y}{C})$	0.29	*(1.63)*	3.72	*(2.87)*
γ_π	interest rates response to inflation	2.13	*(2.12)*	1.80	*(1.66)*
γ_x	interest rates response to output gap	0.001	*(0.04)*	0.05	*(0.005)*
ρ	interest-rate-smoothing parameter	0.72	*(0.85)*	0.84	*(0.82)*
ρ_v	autoregressive coefficient of demand disturbance	0.79	*(0.78)*	0.90	*(0.89)*
ρ_{u^*}	autoregressive coefficient of supply disturbance	0.81	*(0.87)*	0.75	*(0.74)*
ρ_ξ	autoregressive coefficient of policy disturbance	0.22	*(0.07)*	0.14	*(0.18)*

Table 4A.2: Performance of Models under Differing Auxiliaries (HP-filtered data)

(Panel A: pre-1982, the Great Acceleration)

	VAR(2)		VAR(3)	
	Timeless optimum	Taylor rule	Timeless optimum	Taylor rule
Directed Wald for dynamics	100	100	100	100
(Normalized t-stat)	*(5.15)*	*(9.90)*	*(6.03)*	*(7.76)*
Directed Wald for volatilities	79.8	97.7	93.4	98.9
(Normalized t-stat)	*(0.35)*	*(2.34)*	*(1.33)*	*(3.24)*
Full Wald for both	100	100	100	100
(Normalized t-stat)	*(5.45)*	*(9.77)*	*(6.12)*	*(7.72)*

(Panel B: post-1982, the Great Moderation)

	VAR(2)		VAR(3)	
	Timeless optimum	Taylor rule	Timeless optimum	Taylor rule
Directed Wald for dynamics	100	100	100	100
(Normalized t-stat)	*(4.56)*	*(11.7)*	*(6.31)*	*(19.4)*
Directed Wald for volatilities	84.8	99.9	81.2	99.9
(Normalized t-stat)	*(0.78)*	*(6.29)*	*(0.49)*	*(5.10)*
Full Wald for both	100	100	100	100
(Normalized t-stat)	*(4.81)*	*(13.4)*	*(6.93)*	*(18.7)*

Table 4A.3: Performance of Models under Differing Auxiliaries (HP-filtered data) (II)

(Panel A: pre-1984, the Great Acceleration)

	VAR(2)		VAR(3)	
	Timeless optimum	Taylor rule	Timeless optimum	Taylor rule
Directed Wald for dynamics	100	100	100	100
(Normalized t-stat)	*(6.56)*	*(12.7)*	*(6.94)*	*(9.00)*
Directed Wald for volatilities	62.9	99.3	84.5	98.9
(Normalized t-stat)	*(-0.16)*	*(3.29)*	*(0.68)*	*(3.35)*
Full Wald for both	100	100	100	100
(Normalized t-stat)	*(6.39)*	*(12.0)*	*(6.88)*	*(8.50)*

(Panel B: post-1984, the Great Moderation)

	VAR(2)		VAR(3)	
	Timeless optimum	Taylor rule	Timeless optimum	Taylor rule
Directed Wald for dynamics	99.3	100	100	100
(Normalized t-stat)	*(3.41)*	*(7.68)*	*(8.20)*	*(10.9)*
Directed Wald for volatilities	33.2	99.8	59.1	99.9
(Normalized t-stat)	*(-0.70)*	*(4.50)*	*(-0.16)*	*(5.38)*
Full Wald for both	99.6	100	100	100
(Normalized t-stat)	*(3.32)*	*(8.82)*	*(8.19)*	*(11.3)*

References

Benati, Luca and Surico, Paolo (2009), 'VAR Analysis and the Great Moderation', in *American Economic Review*, American Economic Association, vol. 99(4), pages 1636-52, September

Bernanke, Ben, S. and Mihov, Ilian (1998), 'Measuring Monetary Policy', in *Quarterly Journal of Economics*, MIT Press, vol. 113(3), pp. 869-902, August

Boivin, Jean and Giannoni, Marc P. (2006), 'Has Monetary Policy Become More Effective?', in *The Review of Economics and Statistics*, MIT Press, vol. 88(3), pages 445-462, October

Brayton, F., Levin, A., Tryon, R. and Williams, J. C. (1997), 'The Evolution of Macro Models at the Federal Reserve Board', in *Carnegie-Rochester Conference Series on Public Policy*, Elsevier, vol. 47(1), pp. 43-81, December

Calvo, Guillermo, A. (1983), 'Staggered Prices in a Utility Maximising Framework', in *Journal of Monetary Economics*, vol. 12, pp. 383-398

Campbell, John Y. (1995), 'Some Lessons from the Yield Curve', in *Journal of Economic Perspectives*, American Economic Association, vol. 9(3), pp. 129-52, Summer

Canova, Fabio (2005), *Methods for Applied Macroeconomic Research*, Princeton University Press, Princeton

Carare, A. and Tchaidze, R. (2005), 'The Use and Abuse of Taylor Rules: How Precisely Can We Estimate Them?', IMF Working Paper, No. 05/148, July

Carlstrom, C. T. and Fuerst, T. S. (2008), 'Inertial Taylor Rules: the Benefit of Signalling Future Policy', in *Federal Reserve Bank of St. Louis Review*, May/June 2008, 90(3, part 2), pp. 193-203

Castelnuovo, Efrem (2003), 'Taylor Rules, Omitted Variables, and Interest Rate Smoothing in the US', in *Economics Letters*, 81 (2003), pp.55-59

Clarida, Richard, Galí, Jordi and Gertler, Mark L. (1999), 'The Science of Monetary Policy: A New Keynesian Perspective', in *Journal of Economic Literature*, 37(4), pp.1661-1707

Clarida, Richard, Galí, Jordi and Gertler, Mark L. (2000), 'Monetary Policy Rules And Macroeconomic Stability: Evidence And Some Theory', in *The Quarterly Journal of Economics*, MIT Press, vol. 115(1), pages 147-180, February

Cochrane, John, H. (2007), 'Identification with Taylor Rules: A Critical Review', NBER Working Papers 13410, National Bureau of Economic Research, Inc.

Currie, David and Levin, Paul (1985), 'Simple Macropolicy Rules for the Open Economy', in *Economic Journal*, Vol. 95, Supplement, pp. 60-70, Blackwell Publishing

English, W. B., Nelson, W. R. and Sack, B. P. (2002), 'Interpreting the Significance of the Lagged Interest Rate in Estimated Monetary Policy Rules', Federal Reserve Bank of Washington, Finance and Economics Discussion Series Working Paper No. 2002-24

Fair, R. C. (2000), 'Estimated, Calibrated, and Optimal Interest Rate Rules', Cowles Foundation Discussion Papers, No. 1258, Cowles Foundation for Research in Economics, Yale University

Foley, K. Duncan and Taylor, Lance (2004), 'A Heterodox Growth and Distribution Model', in *Economic Growth and Distribution*, Neri Salvadori, ed., Growth and Distribution Conference, University of Pisa, June 16-19, 2004

Friedman, Milton (1960), *A Program For Monetary Stability*, Fordham University Press, New York

Gambetti, Luca, Pappa, Evi and Canova, Fabio (2008), 'The Structural Dynamics of U.S. Output and Inflation: What Explains the Changes?', in *Journal of Money, Credit and Banking*, Blackwell Publishing, vol. 40(2-3), pages 369-388, 03

Giannoni, M. P. and Woodford, M. (2005), 'Optimal Interest Rate Targeting Rules', in *Inflation Targeting*, Benjamin S. Bernanke and Michael Woodford, eds., University of Chicago Press, Chicago, US

Gillman, M., Le, V. P. M. and Minford, A. P. L. (2007), 'An Endogenous Taylor Condition in an Endogenous Growth Monetary Policy Model', Cardiff University Working Paper Series, E2007/29, November 2007

Goodfriend, Marvin (1991), 'Interest Rates and the Conduct of Monetary Policy', in *Carnegie-Rochester Conference Series on Public Policy*, Elsevier, vol. 34(1), pp. 7-30, January

Gourieroux, Christian and Monfort, Alain (1996), *Simulation Based Econometric Methods*, CORE Lectures Series, Oxford University Press, Oxford

Gourieroux, Christian, Monfort, Alain and Renault, Eric (1993), 'Indirect inference', in *Journal of Applied Econometrics*, John Wiley & Sons, Ltd., vol. 8(S), pages S85-118, Suppl. De

Gregory, Allan W. and Smith, Gregor W. (1991), 'Calibration as Testing: Inference in Simulated Macroeconomic Models', in *Journal of Business and Economic Statistics*, American Statistical Association, vol. 9(3), pages 297-303, July

Gregory, Allan W. and Smith, Gregor W. (1993), 'Calibration in macroeconomics', in *Handbook of Statistics*, G. S. Maddla, ed., 11: pp.703-719, St. Louis, MO.: Elsevier

Henderson, Dale W. and McKibbin, W. J. (1993a), 'An Assessment of Some Basic Monetary Policy Regime Pairs: Analytical and Simulation Results from Simple Multi-Region Macroeconomic Models', in *Evaluating Monetary Policy Regimes: New Research in Empirical Macroeconomics,* Ralph Bryant, Peter Hooper and Catherine Mann eds., Chapter 2, pp. 45-218, Brookings Institution Press, Washington, DC

Henderson, Dale W. and McKibbin, W. J. (1993b), 'A Comparison of Some Basic Monetary Policy Regimes for Open Economies: Implications of Different Degrees of Instrument Adjustment and Wage Persistence', in *Carnegie-Rochester Conference Series on Public Policy,* December, *39,* pp. 221-317

Ireland, Peter N. (2007), 'Changes in the Federal Reserve's Inflation Target: Causes and Consequences', in *Journal of Money, Credit and Banking,* Blackwell Publishing, vol. 39(8), pages 1851-1882, December

Judd, J. P. and Rudebusch, G. D. (1998), 'Taylor's Rule and the Fed, 1970-1997', in *Economic Review,* Federal Reserve Bank of San Francisco, No. 3, pp. 3-16

Kuester, K., Muller, G. J. and Stolting, S. (2009), 'Is the New Keynesian Phillips Curve Flat?', in *Economics Letters,* Elsevier, vol. 103(1), pages 39-41, April

Le, V. P. M., Meenagh, D., Minford, A. P. L. and Wickens, M. R. (2009), 'Two Orthogonal Continents: Testing a Two-country DSGE Model of the US and the EU Using Indirect Inference', Cardiff University Working Paper Series, E2009/3, March 2009

Le, V. P. M., Meenagh, D., Minford, A. P. L. and Wickens, M. R. (2011), 'How Much Nominal Rigidity Is There in the US Economy? Testing a New Keynesian DSGE Model Using Indirect Inference', Cardiff University Working Paper Series, E2008/32, December 2008, updated March 2011

Lubik, Thomas A. and Schorfheide, Frank (2004), 'Testing for Indeterminacy: An Application to U.S. Monetary Policy', in *American Economic Review,* American Economic Association, vol. 94(1), pages 190-217, March

Lucas, R. E. (1976), 'Econometric policy evaluation: A critique', in *Carnegie-Rochester Conference Series on Public Policy,* Elsevier, vol. 1(1), pp. 19-46, January

McCallum, B. T. (1988), 'Robustness properties of a rule for monetary policy', in *Carnegie-Rochester Conference Series on Public Policy,* Elsevier, vol. 29(1), pp. 173-203, January

McCallum, B. T. (1976), 'Rational Expectations and the Natural Rate Hypothesis: some consistent estimates', in *Econometrica,* Econometric Society, vol. 44(1), pp. 43-52, January

McCallum, B. T. and Nelson, E. (2004), 'Timeless Perspective versus Discretionary Monetary Policy in Forward-looking Models', in *Review,* Federal Reserve Bank of St. Louis, issue Mar, pp.43-56

Meenagh, D., Minford, A. P. L., Nowell, E., Sofat, P. and Srinivasan, N. (2009), 'Can the Facts of UK Inflation Persistence be Explained by Nominal Rigidity?', in *Economic Modelling*, Elsevier, vol. 26(5), pp. 978-992, September

Meenagh, D., Minford, A. P. L. and Wickens, M. R. (2009), 'Testing a DSGE Model of the EU Using Indirect Inference', in *Open Economies Review*, Springer, vol. 20(4), pp. 435-471, September

Minford, A. P. L. (2008), 'Commentary on 'Economic projections and rules of thumb for monetary policy'', in *Review*, Federal Reserve Bank of St. Louis, issue Jul, pp.331-338

Minford, A. P. L., Perugini, F. and Srinivasan, N. (2001), 'The Observational Equivalence of Taylor Rule and Taylor-type Rules', CEPR discussion papers 2959, September

Minford, A. P. L., Perugini, F. and Srinivasan, N. (2002), 'Are Interest Rate Regressions Evidence for a Taylor Rule?', in *Economic Letters*, Vol. 76, Issue 1, June, pp.145-150

Minford, A. P. L., Theodoridis, K. and Meenagh, D. (2009), 'Testing a Model of the UK by the Method of Indirect Inference', in *Open Economies Review*, Springer, vol. 20(2), pages 265-291, April

Nistico, Salvatore (2007), 'The Welfare Loss from Unstable Inflation', in *Economics Letters*, Elsevier, vol. 96(1), pages 51-57, July

Onatski, Alexei and Williams, Noah (2004), 'Empirical and Policy Performance of a Forward-looking Monetary Model', in *Manuscript*, Columbia University and Princeton University

Orphanides, Athanasios (2001), 'Monetary Policy Rules Based on Real-Time Data', in *American Economic Review*, American Economic Association, vol. 91(4), pp. 964-985, September

Orphanides, Athanasios (2002), 'Activist Stabilization Policy and Inflation: The Taylor Rule in the 1970s', CFS working paper, No. 2002/15

Primiceri, Giorgio E. (2005), 'Time Varying Structural Vector Autoregressions and Monetary Policy', in *Review of Economic Studies*, 72, pages 821-825, The Review of Economic Studies Limited

Qu, Zhongjun and Perron, Pierre (2007), 'Estimating and Testing Structural Changes in Multivariate Regressions', in *Econometrica*, 75(2), 459-502, 03

Rotemberg, Julio J. (1982), 'Sticky Prices in the United States', in *Journal of Political Economy*, vol. 90(6), pp.1187-1211, University of Chicago Press

Rotemberg, Julio J. and Woodford, M. (1997), 'An Optimization-based Econometric Model for the Evaluation of Monetary Policy', in *NEBR Macroeconomics Annual*, 12, pp.297-346

Rotemberg, Julio J. and Woodford, M. (1998), 'An Optimization-based Econometric Model for the Evaluation of Monetary Policy', in *NEBR Technical Working Paper*, No.233

Rudebusch, Glenn D. (2002), 'Term Structure Evidence on Interest Rate Smoothing and Monetary Policy Inertia', in *Journal of Monetary Economics*, Elsevier, vol. 49(6), pages 1161-1187, September

Sack, B. and Wieland, V. (2000), 'Interest-rate Smoothing and Optimal Monetary Policy: A Review of Recent Empirical Evidence', in *Journal of Economics and Business*, 52, pp.205-228

Sims, Christopher A. and Zha, Tao (2006), 'Were There Regime Switches in U.S. Monetary Policy?', in *American Economic Review*, American Economic Association, vol. 96(1), pages 54-81, March

Smets, Frank and Wouters, Raf (2003), 'An Estimated Dynamic Stochastic General Equilibrium Model for the Euro Area', in *Journal of European Economic Association 1(5)*: pp. 1123-1175

Smith, Anthony (1993), 'Estimating nonlinear time-series models using simulated vector autoregressions', in *Journal of Applied Econometrics*, 8: S63—S84

Stock, James H. and Watson, Mark W. (2002), 'Has the Business Cycle Changed and Why?', NBER Working Papers 9127, National Bureau of Economic Research, Inc

Stuart, Alison (1996), 'Simple Monetary Policy Rules', in *Bank of England Quarterly Bulletin*, August 1996, [WWW], <http://www.bankofengland.co.uk/publications/quarterlybulletin/qb960301.pdf>, (accessed January 2011)

Svensson, Lars E.O. and Woodford, Michael (2004), 'Implementing Optimal Policy through Inflation-Forecast Targeting', in *Inflation Targeting*, Ben S. Bernanke and Michael Woodford, eds. University of Chicago Press, Chicago

Taylor, J. B. (1993), 'Discretion versus Policy Rules in Practice', in Carnegie-Rochester Conference Series on Public Policy 39, pp.195-214

Walsh, Carl E. (2003), *Monetary Theory and Policy*, 2nd ed., the MIT Press, London, England

Wickens, Michael, R. (1982), 'The Efficient Estimation of Econometric Models with Rational Expectations', in *Review of Economic Studies*, Blackwell Publishing, vol. 49(1), pages 55-67, January

Wilson, Edwin B. and Hilferty, Margaret M. (1931), 'The Distribution of Chi-Square', in *Proceedings of the National Academy of Sciences*, 17(12), 684–688, December

Woodford, Michael (1999), 'Optimal Monetary Policy Inertia', in *Manchester School*, University of Manchester, vol. 67(0), pages 1-35, Supplement

Woodford, Michael (2003a), 'Optimal Interest Rate Smoothing', in *Review of Economic Studies*, vol. 70, issue 4, pp. 861-886

Woodford, Michael (2003b), *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press, Princeton

Woodford, Michael (2010), 'Robustly Optimal Monetary Policy with Near-Rational Expectations', in *American Economic Review*, American Economic Association, vol. 100(1), pp. 274-303, March

Supporting Annex

A. Micro-foundations and Derivations of the Baseline Model and the Optimal Timeless Rule

The main context has followed the common practice among New Keynesian authors (including Clarida, Gali and Gertler (1999, 2000), Rotemberg and Woodford (1997, 1998), Walsh (2003) and Holmberg (2006)) of reducing a full DSGE model to a three-equation framework consisting of an 'IS' curve, a Phillips curve and a monetary policy rule that summarises the Fed's behaviour. The following outlines the micro-foundations and the derivation process of these.

The Households

Representative households are assumed to consume a composite of differentiated goods produced by monopolistically competitive firms that make up of a continuum of measure 1. The composite consumption that enters the utility function in each period is:

$$C_t = [\int_0^1 c_{jt}^{\frac{\theta-1}{\theta}} dj]^{\frac{\theta-1}{\theta}} \qquad [A.1]$$

where $\theta(>1)$ is the price elasticity of demand for good j. The cost minimization process of representative households implies the demand for good j is:

$$c_{jt} = (\frac{p_{jt}}{P_t})^{-\theta} C_t \qquad [A.2]$$

where p_{jt} is the price of j and P_t is the general price level in that period[48].

Assume for simplicity that the representative agents care only about leisure and the level of composite consumption such that the life-time utility function takes the form:

$$U_t = E_t \sum_{i=0}^{\infty} \beta^i [\frac{C_{t+i}^{1-\sigma}}{1-\sigma} - \chi \frac{N_{t+i}^{1+\eta}}{1+\eta}] \qquad [A.3]$$

where σ is the inverse of intertemporal elasticity of substitution of consumption, whereas η is the inverse of elasticity of labour[49].

---

[48] Details of this could be found in Walsh (2003, pp.232).

[49] The utility function here is deliberately assumed to take the same form as in Woodford and Rotemberg (1998) and Nistico (2007) such that the utility-based micro-founded quadratic social

Suppose further that the agents also own the firms and work at the same time as employees. They then face a real budget constraint as [A.4]:

$$C_t + \frac{M_{t+1}}{P_t} + \frac{B_{t+1}}{P_t} = \frac{W_t}{P_t} N_t + \frac{M_t}{P_t} + (1+i_t)\frac{B_t}{P_t} + \Pi_t \qquad [A.4]$$

where M_t and B_t are respectively the initial stocks of money and nominal bond in each period, W_t is the nominal wage income, Π_t denotes the profit from running the firm and i_t is the nominal interest rates. The bond market is introduced here to give interest rates a role; labour is made the only factor that goes into the production process so the wage income constitutes 100% of the households' disposal income.

By assuming the Cash-in-advance constraint, the utility maximization problem can be described using the Lagrangean function as follows:

$$\begin{aligned}
\underset{C_t, N_t, M_{t+1}, B_{t+1}}{Max} L_0 = E_0 \sum_{t=0}^{\infty} \beta^t \{ & [\frac{C_t^{1-\sigma}}{1-\sigma} - \chi \frac{N_t^{1+\eta}}{1+\eta}] \\
& - \lambda_t [C_t + \frac{M_{t+1}}{P_t} + \frac{B_{t+1}}{P_t} - \frac{W_t}{P_t} N_t - \frac{M_t}{P_t} - (1+i_t)\frac{B_t}{P_t} - \Pi_t] \\
& - \mu_t [C_t - \frac{M_t}{P_t}] \} \qquad [A.5]
\end{aligned}$$

The first order conditions suggest:

$$\begin{aligned}
C_t: &\quad C_t^{-\sigma} = (\lambda_t + \mu_t) \\
N_t: &\quad \chi N_t^{\eta} = \lambda_t \frac{W_t}{P_t} \\
M_{t+1}: &\quad \lambda_t (1+\pi_{t+1}) = \beta(\lambda_{t+1} + \mu_{t+1}) \\
B_{t+1}: &\quad \lambda_t (1+\pi_{t+1}) = \lambda_{t+1} \beta (1+i_{t+1})
\end{aligned}$$

These further imply:

$$C_t^{-\sigma} = \beta(1+i_t) E_t \frac{P_t}{P_{t+1}} C_{t+1}^{-\sigma} \qquad [A.6]$$

$$\frac{\chi N_t^{\eta}}{C_t^{-\sigma}}(1+i_t) = \frac{W_t}{P_t} \qquad [A.7]$$

i.e., the 'Euler's equation' [A.6] and the optimal intratemporal substitution between labour and consumption [A.7].

---

welfare loss function they have suggested are also applicable here. In contrast to Walsh (2003) where Money-in-utility is assumed, this model retains the role of money through the CIA constraint.

Log-linearization of [A.6] around zero-inflation steady state suggests:

$$\tilde{c}_t = E_t\tilde{c}_{t+1} - (\frac{1}{\sigma})(\tilde{i}_t - E_t\pi_{t+1}) \qquad \text{[A.6]'}$$

where '~' denotes 'percentage deviation from steady state'[50].

Since no physical capital (and therefore no investment) is assumed, log-linearising the market clearing condition $Y_t=C_t+G_t$ returns:

$$\tilde{c}_t = \tilde{y}_t + \ln(1-\frac{G_t}{Y_t}) - \ln\frac{C}{Y} \qquad \text{[A.8]}$$

Combining [A.6]' and [A.8] then gives the commonly found 'IS' curve:

$$x_t = E_t x_{t+1} - (\frac{1}{\sigma})(\tilde{i}_t - E_t\pi_{t+1}) + v_t \qquad \text{[A.9]}$$

where $x_t \equiv \tilde{y}_t - \tilde{y}_t^f$, $v_t = (E_t\tilde{y}_{t+1}^f - \tilde{y}_t^f) + (E_t\hat{g}_{t+1} - \hat{g}_t)$ with $\hat{g}_t \equiv \ln(1-\frac{G_t}{Y_t})$[51].

Note the 'output gap' x_t here is, by theory, defined as the (log) difference between actual output and the output that would prevail under perfect flexibility. This has been approximated using the log difference of actual output from its HP trend, however, as the main context has explained. Note also in this case the 'demand shock' v_t is a combination of shocks to both technology and government expenditure.

The Firms

As explained in this model the representative agents also own the firms. Under a monopolistically competitive environment each firm has production function:

$$y_{jt} = A_t N_{jt} \qquad \text{[A.10]}$$

where 'j' denotes the jth firm; A_t is technology with $\log A_t = \xi \log A_{t-1} + z_t$, where z_t indicates the i.i.d. productivity shock.

Under the Calvo (1983) contract, for any given period t, only a fraction, $1-\omega$, of these are able to reset their prices to the optimal level, whereas the rest, ω, have to keep these unchanged due to 'menu cost'[52].

---

[50] In particular, $\tilde{i}_t \equiv i_t - i$.

[51] Walsh (2003) assumes Y_t=Ct so that $v_t = E_t\tilde{y}_{t+1}^f - \tilde{y}_t^f$.

[52] For simplicity, nominal wages in the labour market are presumed to be fully flexible.

Equation [A.2] implies the demand curve faced by each firm is:

$$y_{jt} = (\frac{p_{jt}}{P_t})^{-\theta} Y_t \qquad [A.11]$$

So in each period firms producing differentiated goods but processing identical pricing strategy would set individual prices p_{jt} , subject to the production constraint [A.10], the Calvo contract resetting probability $1-\omega$ and the demand curve [A.11], to maximize (the discounted) real profits.

Let φ denotes the real marginal cost to each firms' production. The cost minimization process suggests

$$\varphi_t = \frac{W_t / P_t}{A_t} \qquad [A.12]$$

so that the profit maximization problem for each firm becomes:

$$\underset{p_{jt}}{Max} E_t \sum_{i=0}^{\infty} \omega^i \beta^i V_{i,t+i} [(\frac{p_{jt}}{P_{t+i}}) y_{j,t+i} - \varphi_{t+i} y_{j,t+i}] \qquad [A.13]$$

where $V_{i,t+i}$ is a discount factor, indicating the ratio of marginal utilities of consumption between periods.

Using the demand curve [A.11] to substitute away $y_{j,t+i}$, equation [A.13] can be rewritten as:

$$\underset{p_{jt}}{Max} E_t \sum_{i=0}^{\infty} \omega^i \beta^i V_{i,t+i} [(\frac{p_{jt}}{P_{t+i}})^{1-\theta} - \varphi_{t+i} (\frac{p_{jt}}{P_{t+i}})^{-\theta}] Y_{t+i} \qquad [A.13]'$$

The first order condition of [A.13]' with respect to individual price p_{jt} implies:

$$E_t \sum_{i=0}^{\infty} \omega^i \beta^i V_{i,t+i} Y_{t+i} [(1-\theta)(\frac{p_{jt}}{P_{t+i}}) + \theta \varphi_{t+i}] \frac{1}{p_{jt}} (\frac{p_{jt}}{P_{t+i}})^{-\theta} = 0 \qquad [A.14]$$

Log-linearization of [A.14] around zero inflation steady state yields the optimal reset price for each firm as follows:

$$\widetilde{p}_{jt}^* = (1-\omega\beta) \sum_{i=0}^{\infty} \omega^i \beta^i (E_t \widetilde{\varphi}_{t+i} + E_t \widetilde{P}_{t+i}) \qquad [A.15]$$

The general price level in each period given the Calvo contract can be written as the weighted average of this up-to-date reset prices and the unchanged, with the weights being the reset probability, $1-\omega$, and its opposite, ω, respectively[53]. That is:

$$P_t = (1-\omega) p_{jt}^* + \omega P_{t-1} \qquad [A.16]$$

---

[53] Note that individual firms have exactly the same pricing strategy p_{jt}^* (or equivalently, $\widetilde{p}_{jt}^*$).

Log-linearization of [A.16] implies:

$$\pi_t = (1-\omega)\widetilde{p}^*_{jt} + (\omega-1)\widetilde{P}_{t-1} \qquad [A.17]$$

Combining [A.15] and [A.17] gives:

$$\pi_t = \beta E_t \pi_{t+1} + \frac{(1-\omega)(1-\omega\beta)}{\omega}\widetilde{\varphi}_t \qquad [A.18]^{54}$$

or more conveniently, the standard forward-looking *New Keynesian Phillips curve*:

$$\pi_t = \beta E_t \pi_{t+1} + \kappa\widetilde{\varphi}_t \qquad [A.18]'$$

where $\kappa = \frac{(1-\omega)(1-\omega\beta)}{\omega}$.

This can be further transformed to the version in which inflation is related to output gap by log-linearising the real marginal cost equation [A.12] and the labour supply equation [A.7] (which was implied by the household's problem) and combining the results. After some tedious algebra it can be shown that:

$$\widetilde{\varphi}_t = (\eta + \sigma\frac{Y}{C})(\widetilde{y}_t - \widetilde{y}_t^f) = (\eta + \sigma\frac{Y}{C})x_t \qquad [A.19]^{55}$$

In the spirit of Clarida, Gali and Gertler (2002), suppose further the labour market is *not* perfectly competitive such that the wage mark-up over intratemporal substitution between consumption and labour is subject to stochastic errors and so [A.19] becomes:

$$\widetilde{\varphi}_t = (\eta + \sigma\frac{Y}{C})x_t + u_t^w \qquad [A.19]'$$

where u_t^w is interpreted as the disturbance causing bias to the wage mark-ups. The Phillips curve can then be rewritten as:

$$\pi_t = \beta E_t \pi_{t+1} + \gamma x_t + u_t \qquad [A.20]$$

where $\gamma = \kappa(\eta + \sigma\frac{Y}{C})$, $u_t = k u_t^w$, and $\kappa = \frac{(1-\omega)(1-\omega\beta)}{\omega}$.

Monetary Policy

To close the model most New Keynesian authors have employed an exogenously-specified Taylor rule similar to [1.2] such that a model for inflation, output gap and interest rates determinations is complete. The baseline model in this thesis, however,

---

[54] Note [A.15] can be conveniently written as $\widetilde{p}^*_{jt} = \frac{(1-\omega\beta)}{(1-\omega\beta B^{-1})}(\widetilde{\varphi}_t + \widetilde{P}_t)$ under rational expectations.

[55] This result is obtained when the market-clearing condition is $Y_t=C_t+G_t$. Had it been defined as $Y_t=C_t$ as in Walsh (2003), it would imply $\widetilde{\varphi}_t = (\eta+\sigma)x_t$.

has taken the alternative of assuming the optimal timeless policy implicitly determined by the economy.

Following Rotemberg and Woodford (1998) and Nistico (2007), it defines the 'social welfare loss' as 'the loss in units of consumption as a percentage of steady-state output' so that:

$$SWL_t \equiv \frac{U - U_t}{MU_c \cdot Y}$$

Under the Calvo (1983) pricing mechanism and given the utility function [A.3], Rotemberg and Woodford (1998) showed the social welfare loss function can approximately be expressed in terms of the variance of inflation and output through second-order Taylor expansion. The transformed social welfare loss takes the form:

$$SWL_t = \frac{\psi}{2}[\alpha x_t^2 + \pi_t{}^2] \qquad \text{[A.21]}^{56}$$

where ψ is some measure of stickiness, α indicates the relative weight that central banks put on loss from output variation against inflation variation[57].

The social planner's problem then involves minimizing [A.21] in each period subject to the Phillips curve. One can write this in terms of a Lagrangean equation as in McCallum and Nelson (2004) as:

$$\underset{\pi_{t+i}, x_{t+i}}{Min}\ L_t = E_t \sum_{i=0}^{\infty} \beta^i \{ \frac{\psi}{2}(\pi_{t+i}^2 + \alpha x_{t+i}^2) + \lambda_{t+i}(\gamma x_{t+i} + \beta \pi_{t+i+1} + u_{t+i} - \pi_{t+i}) \} \qquad \text{[A.22]}$$

Suppose the problem starts from period '1', the first order conditions with respect to π_t's and x_t's are:

$$\pi_1: \quad \psi\pi_1 - \lambda_1 = 0 \qquad \text{(the initial period)} \qquad \text{[A.23]}$$

$$\pi_t: \quad E_1(\psi\pi_t + \lambda_{t-1} - \lambda_t) = 0 \qquad t=2,3,\ldots\ldots \qquad \text{[A.24]}$$

$$x_t: \quad E_1(\psi\alpha x_t + \gamma\lambda_t) = 0 \qquad t=1,2,3,\ldots\ldots. \qquad \text{[A.25]}$$

Under the 'timeless perspective' that 'ignores the conditions that prevail at the regime's inception' (McCallum and Nelson, 2004, pp.44), the optimal response is derived by combining [A.24] and [A.25] while [A.23] is dropped. Hence:

[56] Note it has implicitly assumed that the steady state inflation is zero. The social welfare loss function in this case is *not* ad hoc. The same expression is also derived by Nistico (2007) who assumed the Rotemberg (1982) pricing mechanism. In particular, Nistico showed the relative weight α is equal to the ratio of the slope of the Phillips curve to the price elasticity of demand so α=γ/θ.

[57] Note ψ does not affect the implied optimal response as it will be cancelled out in combining the F.O.Cs.

$$\pi_t = -\frac{\alpha}{\gamma}(x_t - x_{t-1}) \qquad \text{[A.26]}$$

If a ‘trembling hand’ is assumed in implementation of the policy, the above becomes:

$$\pi_t = -\frac{\alpha}{\gamma}(x_t - x_{t-1}) + \xi_t \qquad \text{[A.26]'}$$

where ξ_t denotes the ‘policy shock’ that causes bias to [A.26].