School of Economics and Finance

A Bayesian Local Likelihood Method for Modelling Parameter Time Variation in DSGE Models

Ana Beatriz Galvão, Liudas Giraitis, George Kapetanios and Katerina Petrova

Working Paper No. 770 December 2015 ISSN 1473-0278

Queen Mary
University of London

A Bayesian Local Likelihood Method for Modelling Parameter Time Variation in DSGE Models[*]

Ana Beatriz Galvão[†] Liudas Giraitis[‡] George Kapetanios[‡§] Katerina Petrova[‡]

June 23, 2015

Abstract

DSGE models have recently received considerable attention in macroeconomic analysis and forecasting. They are usually estimated using Bayesian methods, which require the computation of the likelihood function under the assumption that the parameters of the model remain fixed throughout the sample. This paper presents a Local Bayesian Likelihood method suitable for estimation of DSGE models that can accommodate time variation in all parameters of the model. There are two advantages in allowing the parameters to vary over time. The first is that it enables us to assess the possibilities of regime changes, caused by shifts in the policy preferences or the volatility of shocks, as well as the possibility of misspecification in the design of DSGE models. The second advantage is that we can compute predictive densities based on the most recent parameters' values that could provide us with more accurate forecasts. The novel Bayesian Local Likelihood method applied to the Smets and Wouters (2007) model provides evidence of time variation in the policy parameters of the model as well as the volatility of the shocks. We also show that allowing for time variation improves considerably density forecasts in comparison to the fixed parameter model and we interpret this result as evidence for the presence of stochastic volatility in the structural shocks.

JEL codes: C11, C53, E27, E52

Keywords: DSGE models, local likelihood, Bayesian methods, time varying parameters

[*]Galvão, Kapetanios and Petrova acknowledge financial support from the ESRC grant No ES/K010611/1.

[†]Warwick Business School, University of Warwick

[‡]School of Economics and Finance, Queen Mary University London.

[§]Corresponding author, Tel: +442078828698, Email: G.Kapetanios@qmul.ac.uk

1 Introduction

DSGE models are extensively used in both academic work and macroeconomic policy making. Their success is a result of their capacity to combine economic microfoundations derived from optimisation decisions of agents with rational expectations and business cycle fluctuations. Traditionally, the consensus in the macroeconomic literature has been that there exists an apparent trade-off between theoretical coherence, whereby a model's outcomes can be explained by well-established theory, and empirical coherence, whereby a model can fit and explain macroeconomic data well, but its outcomes are often difficult to interpret or justify from a theoretical standpoint. Models that exhibit theoretical and empirical coherence simultaneously were deemed infeasible. DSGE models were alleged to be at the theoretical end of this trade-off curve. On the other hand, reduced-form models, such as VAR models, exploiting correlations in time series with little reliance on macroeconomic theory, were put at the empirical end. It was the work of Smets and Wouters (2003, 2005, 2007), based on earlier work of Rotemberg and Woodford (1997) and Christiano, Eichenbaum and Evans (2005), that changed this perception and demonstrated that medium-sized DSGE models can be successfully taken to the data and produce superior forecasts to standard BVAR models. Following Smets and Wouters, the literature on DSGE model estimation and forecasting has become a vibrant area of research with considerable progress in the development of the underlying economic theory and the design of numerical solution and estimation algorithms.

At the heart of DSGE models are so called deep parameters that define the preferences and technological environment of the economy. These are kept constant and are structural in the sense that they are not subject to the Lucas critique - they are invariant to both policy and structural shocks. There are two issues related to these parameters that this paper will address. First, it is important to recognise the possibility of parameter drift in order to re-evaluate the usefulness and relevance of DSGE models. If substantial evidence is found that some of these structural parameters are in fact not constant, this could be interpreted as a need to revise existing models in order to account for such variation. It is possible that slow time variation is the outcome of long term cultural or technological shifts in the economy that DSGE models are ill-equipped to model, since they focus primarily on business cycle fluctuations. Nevertheless, taking into account such slow variation is paramount for the effective use of DSGE models. Furthermore, time variation in these parameters can be a signal for misspecification in existing models and hence a guide to amend and improve them. Second, these models are widely used in forecasting both by academics and official institutions. Hence, allowing the structural parameters to change and using only their most recent values for generating predictions seems like a useful modification that would be expected to improve forecasting performance, possibly at the cost of making the separation between structural and reduced-form models less clear.

To accommodate such time variation in DSGE model parameters, this paper applies a Bayesian Local Likelihood (BLL) method, developed in a general reduced-form setting in Galvão, Giraitis, Kapetanios and Petrova (2015). BLL estimates parameters at each point in time, appropriately weighting the sum of log likelihoods of the sample, with weights generated by a kernel function. The method is general and can be applied to any DSGE model. Furthermore, for generating forecasts, it is no more computationally intensive than estimating a DSGE model with fixed parameters.

This paper contributes to a small but expanding literature on estimating DSGE models with time variation in the parameters which has two strands. Fernandez-Villaverde and Rubio-Ramirez (2008) and Justiniano and Primiceri (2008) model time variation by assuming stochastic processes for a subset of the parameters and include these to the set of state equations. Fernandez-Villaverde and Rubio-Ramirez (2008) assume that the agents, in the model, take into account current and future parameter variation, utilising the parameters' representation as stochastic processes when computing their expectations. A similar assumption is made by Schorfheide (2005), but there it is assumed that the parameters follow Markov-Switching processes. In contrast, Canova (2006), Canova and Sala (2009) and Giacomini and Rossi (2009) assess parameter time variation by estimating DSGE models over rolling samples. A similar strategy was followed by Castelnuovo (2012), Cantore, Levine and Melina (2012) and Canova and Ferroni (2012). It is useful to contrast our work with both these strands. This first strand makes parametric assumptions about the variation in the parameters. These assumptions are not microfounded but have a reduced form flavour. Instead, our method is agnostic about the source of the variation apart from assuming that it is slow, although given some time it can track more abrupt forms of change. Given the considerable likelihood that any changes are the result of long term cultural and technological shifts that no mainstream business cycle model is well equipped to explain, this agnostic approach has merit. A further issue is that computational complexity restricts the ability of allowing for time variation to only a small subset of the model parameters whereas our approach is scaleable to the full set of parameters. Concerning the second strand, our proposed method employs a nonparametric kernel-based procedure that encompasses rolling window estimation as a special case. Evidence provided in Giraitis, Kapetanios and Yates (2014), where our approach originates, suggests that other kernel functions may have more desirable propeties than the flat kernel underlying the rolling window. Our approach is related to the one in Giraitis, Kapetanios, Theodoridis and Yates (2013), but they apply the local kernel estimator developed by Giraitis et al. (2014) to the minimum distance estimator that matches DSGE and VAR impulse responses, to provide a frequentist estimation approach.

One aim of this paper is to improve the accuracy of DSGE models in forecasting. Smets and Wouters (2007) show that their medium-sized DSGE model can generate forecasts for seven US macro variables that are superior to those obtained from a BVAR model. The gains of the structural model over the reduced-form model are substantial especially at longer horizons. Additional evidence that DSGE models may deliver competitive forecasts in comparison with statistical models and survey of professional forecasters is provided by Rubaszek and Skrzypczynski (2008), Steinbach, Mathuloe and Smith (2009), Edge, Kiley and Laforte (2009), Edge and Guerkaynak (2010), Wieland and Wolters (2011) and Del Negro and Schorfheide (2013b). To the best of our knowledge, there is no documented evidence of the forecasting performance of a DSGE model with time variation in the parameters. The closest to ours is the working paper by Edge, Guerkaynak and Kisacikoglu (2013) who use rolling window scheme to assess the forecasting record of a DSGE model.

The paper is organised as follows. Section 2 introduces the Bayesian Local Likelihood approach, Section 3 presents an empirical application based on the Smets and Wouters (2007) model, Section 4 provides a forecasting comparison and Section 5 concludes.

2 The Bayesian Local Likelihood Method

In the context of DSGE model estimation, there are advantages of adopting a Bayesian approach. Bayesian methods provide a natural way of combining econometric estimation with information provided by calibration methods widely used in the previous generation of models (see Kydland and Prescott (1996)). For example, by construction, we know that the discount factor, β, that consumers use to discount expected future utility cannot take negative values, is bounded between zero and one and a typical value based on the assumption of a 4% annual discount rate is 0.99. Adding a probability mass in the form of a prior is a natural way to incorporate such additional information which is not contained in the data and serves as augmenting the likelihood with artificial observations. In addition, the likelihood of DSGE models may often be ill-identified or not globally concave. Adding a prior can resolve such issues and make the problem well-defined (see, Lindley (1971)). Finally, a Bayesian approach can deal in a natural way with model misspecification. Instead of assuming that there is a unique parameter vector that contains the "true" values of all parameters, the Bayesian approach considers the parameters as random variables and the estimation procedure as a learning process with respect to the characteristics of these random variables, after incorporating information on the available data. An and Schorfheide (2007) and Del Negro and Schorfheide (2013a) offer a detailed review of Bayesian inference in the context of DSGE models.

We begin by showing how a vector of time varying parameters can be estimated by a local likelihood estimation in Section 2.1. Then we discuss the main reasoning behind the Bayesian Local Likelihood (BLL) method, including its advantages and disadvantages in Section 2.2. Finally, Section 2.3 outlines how to apply the BLL method to DSGE estimation, including the description of the algorithms to obtain the posterior distributions and predictive densities.

2.1 The Frequentist Local Likelihood Estimator

Let y_t, $t = 1, .., T$, be an observed time series with an assumed log-likelihood, given by $l_{\theta,t,T}$, that is a function of a time varying finite dimensional vector of parameters, θ_t. θ_t can be either a deterministic function of time given by

$$\theta_t = \theta\left(\frac{t}{T}\right), \quad \theta(.) \in C^1[0,1] \tag{1}$$

where $\theta(.)$ is piecewise differentiable, or a stochastic function of time, satisfying

$$\sup_{j:|j-t|\leq h} ||\theta_t - \theta_j||^2 = O_p\left(h/t\right). \tag{2}$$

Both (1) and (2) imply that the parameters drift slowly and this is necessary for consistent estimation of θ_t. We wish to provide an extremum estimator for θ_t of the form

$$\hat{\theta}_t = \arg\min_{\theta} l_{\theta,t,T}, \qquad l_{\theta,t,T} := \sum_{j=1}^{T} w_{tj} l_j \left(y_j | y^{j-1}, \theta_t\right)$$

where $l_j\left(y_j|y^{j-1},\theta_t\right)$ is the conditional log-likelihood for observation j and parameter vector θ_t. The weights are given by $w_{tj} = \tilde{w}_{tj}/\left(\sum_{j=1}^{T}\tilde{w}_{tj}\right)$ with $\tilde{w}_{tj} = K\big((t-j)/H\big)$ where $K(x) \geq 0$, $x \in R$, K is a continuous bounded function and H is a bandwidth parameter such that $H \to \infty$, $H = o(T/\log T)$. Then, as discussed in Giraitis, Kapetanios, Wetherilt and Zikes (2015), under certain regularity conditions, $\hat{\theta}_t$ is an $H^{1/2} + (T/H)^{1/2}$- consistent estimator of θ_t for all $t = [\tau T]$, $0 < \tau < 1$. Furthermore, defining

$$\widehat{\Sigma}_t := W_{tT}\left(-\frac{\partial^2 l_{\hat{\theta}_t,t,T}}{\partial\theta\partial\theta'}\right)^{-1}, \qquad W_{tT} := \sum_{j=1}^{T} w_{tj}^2,$$

where $-(\partial^2/\partial\theta\partial\theta')l_{\hat{\theta},t,T}$ is positive definite, leads to the asymptotic normality result

$$\widehat{\Sigma}_t^{-1/2}(\hat{\theta}_t - \theta_t) \to_D \mathcal{N}(0, I),$$

when $H = o(T^{1/2})$, where $N(0, I)$ denotes the k-variate standard normal distribution.

2.2 The Bayesian Local Likelihood (BLL) Estimator

Recently, there has been increased interest in providing a Bayesian treatment to the problem of estimating time varing parameters in likelihood-based methods such as in Cogley and Sargent (2002) and Primiceri (2005). However, the prevailing solution is to provide a linear parametric model for θ_t whose law of motion is fully specified up to a unknown finite dimensional vector and recasts the model into a state space form. In fact, the usual practice is for y_t itself to be an affine, up to θ_t, function of other observed and unobserved variables. Consider a simple model:

$$\begin{aligned} y_t &= \theta_t f_t + v_{1t} \\ f_t &= a f_{t-1} + v_{2t} \\ \theta_t &= \theta_{t-1} + v_{3t} \end{aligned}$$

where θ_t and f_t are unobserved and v_{kt}, $k = 1, 2, 3$ are martingale difference sequences. Such a model becomes difficult to estimate for large dimensions of θ_t and f_t and is restrictive in a number of ways such as the choice of the law of motion for θ_t (for recent examples in empirical macroeconomics, see Cogley and Sbordone (2008), Benati and Surico (2009), Gali and Gambetti (2009), Canova and Gambetti (2009) and Mumtaz and Surico (2009)).

We suggest an alternative that is related to the frequentist work of Giraitis et al. (2015) and is both easier to estimate and shares the level of generality of treatment implied in their paper, while providing outputs that are useful from a Bayesian point of view such as posterior distribution for θ_t at each point in time t, which standard methods cannot deliver due to their treatment of the time varying parameters as latent state variables rather than a parameter with a fully specified prior and posterior distribution.

Let $p_t(\theta_t)$ denote a prior distribution for θ_t at time t. Then, the posterior $p_t(\theta_t|Y)$ is given by:

$$p_t(\theta_t|Y) = \frac{p_t(\theta_t)L_t(Y|\theta_t)}{\int_\Theta p_t(\theta_t)L_t(Y|\theta_t)d\theta} \propto p_t(\theta_t)L_t(Y|\theta_t)$$

where $L_t(Y|\theta_t) = \prod_{j=1}^{T} L(y_j|y^{j-1}, \theta_t)^{w_{tj}}$ for $t = 1, .., T$, where $Y = (y_1, .., y_T)'$, and $L(y_j|y^{j-1}, \theta_t)$ denotes the likelihood for observation j, conditional on the history y^{j-1}. This provides a generic Bayesian principle for estimating general time varying coefficient models that require little more than standard Bayesian numerical techniques applicable to fixed coefficient models. More details, standard conjugacy results, Monte Carlo evidence and illustrative empirical applications for this general method are provided in Galvão et al. (2015).

Empirical macroeconomic studies typically impose priors that conveniently deliver conditionally conjugate posterior distributions. The resulting posteriors are easy to draw from using well-known Gibbs sampling methods. In this context, our method requires running T independent Markov chains for the $k \times 1$ vector θ_t, which is still considerably easier and computationally cheaper[1] than the single chain drawing a $k \times T$-dimensional matrix of time varying parameters, e.g., as in Cogley and Sargent (2002) or Primiceri (2005)[2]. Another advantage of our method is that it is applicable without conjugacy: if the posterior did not belong to a known distributional family, other MCMC methods can be used to generate draws from that posterior. Finally, our method includes the rolling window weights employed by Canova (2006), Canova and Sala (2009), Giacomini and Rossi (2009), Castelnuovo (2012), Cantore et al. (2012), Canova and Ferroni (2012) as a special case. The flat weights w_{tj} implied by rolling windows might not be optimal since we expect the change in the DSGE

[1]Note that since the T chains are independent, computation time can be further reduced by exploiting parallel pool tools.

[2]This dimensionality issue is the reason why time varying VAR models estimated with the latter technique are only limited to at most three or four variables.

parameters to be gradual and the weights based on normal density tend to accommodate such change well. For further discussion of the advantages of exponential kernels over the flat kernel for introducing time variation, refer to Giraitis et al. (2014) and Giraitis, Kapetanios and Price (2013).

The previous discussion related chiefly to time variation in reduced form models. The literature on time varying DSGE models is less developed. The alternative approach of specifying processes for the drifting parameters in a DSGE context was applied for instance by Fernandez-Villaverde and Rubio-Ramirez (2008). The main advantage of their approach is that agents populating the model take into account the parameters' stochastic processes when forming their expectations about the future. Our econometric approach, on the other hand, does not incorporate a law of motion for the parameters when solving the agents' rational expectation problem. However, if the parameters of the model are driven by either a time varying deterministic or slowly moving stochastic process, as the popular random walk assumption in the literature, then our econometric approach does not violate the rational expectation assumption in a linearised model because future changes in the parameters are unpredictable by both agents and the econometrician. This implies that current parameter values are the best prediction for future values anyway. A disadvantage of modelling parameters' variation by explicitly specifying a stochastic process is that it is subject to the curse of dimensionality. The state vector needs to be augmented for each parameter allowed to vary and an additional shock is introduced. Because of this dimensionality problem, all parameters cannot be modelled simultaneously in this way. For instance, Fernandez-Villaverde and Rubio-Ramirez (2008) do not allow both Taylor rule and price rigidity parameters to vary simultaneously when estimating their DSGE model. Our alternative econometric approach does not suffer from such dimensionality issues.

In addition, the modelling approach of Fernandez-Villaverde and Rubio-Ramirez (2008) imposes an additional structure by relying on the assumption that the law of motion for the parameters' time variation is correctly specified. Our nonparametric approach performs well for many different parameters' laws of motion. Galvão et al. (2015) show in a Monte Carlo exercise that if the law of motion is misspecified, inconsistent estimates of the parameters' time variation are obtained if they are treated as unobserved state variables as in Fernandez-Villaverde and Rubio-Ramirez (2008). In contrast, Galvão et al. (2015) suggest that our non-parametric alternative is consistent even when the prior for the drifting parameters is poor. Schorfheide (2007) argues that by treating time varying parameters as unobserved state variables as in Fernandez-Villaverde and Rubio-Ramirez (2008), identification issues which are attenuated by the use of priors, as argued earlier in this section, may arise[3]. Our approach has the advantage of being able to incorporate prior information about the time varying parameters to solve possible identification issues.

2.3 The BLL Method for DSGE Models

In this subsection, we show how to apply the BLL approach described previously to a DSGE model with linear state-space representation. Note, however, that the BLL method could also be applied to models that have a non-linear state space representation such as in Fernandez-Villaverde and Rubio-Ramirez (2007).

The linearized rational expectation model with time varying parameters can be written in the form:

$$A(\theta_t)\mathbb{E}_t x_{t+1} = B(\theta_t)x_t + C(\theta_t)v_t, \qquad v_t \sim N(0, Q(\theta_t))$$

where x_t is a $n \times 1$ the vector of model's variables, v_t is a $k \times 1$ vector of structural shocks, θ_t is a vector of parameters, including parameters governing preferences and the shocks' stochastic processes, A, B and C are matrices, which are functions of θ_t, $Q(\theta_t)$ is a diagonal covariance matrix, and E_t is the expectation operator conditional on information available at time t. Observe that we have one such equation at each point in time $t = 1, ..., T$.

A numerical solution of the rational expectation model can be obtained by one of the available methods (for instance, Blanchard and Kahn (1980) or Sims (2002)). The resulting state equation is given by:

$$x_t = F(\theta_t)x_{t-1} + G(\theta_t)v_t \tag{3}$$

where the $n \times n$ matrix F and $n \times k$ matrix G can be computed numerically for a given parameter vector θ_t. The system is augmented with a measurement equation:

$$y_t = D(\theta_t) + Z(\theta_t)x_t \tag{4}$$

where y_t is an $m \times 1$ vector of observables, typically of a smaller dimension than x_t (i.e. $m < n$) and Z is a $m \times n$ matrix that links those observables to the latent variables in the model x_t.

[3] For instance, Fernandez-Villaverde and Rubio-Ramirez (2008) obtain values of their Taylor rule parameters for which the Taylor principle is not satisfied.

Equations 3 and 4 provide the state-space representation of the model, which is linear and Gaussian at each successive set of parameters θ_t for $t = 1, ..., T$. Therefore, the Kalman filter can be employed to recursively build the likelihood of the sample of observables $\{y_t\}_{t=1}^T$. The appropriately weighted likelihood of the sample is given by:

$$L_t(Y|\theta_t) = \prod_{j=1}^{T} L(y_j|y^{j-1}, \theta_t)^{w_{tj}} \text{ for } t = 1, .., T$$

where w_{tj} is an element of the $T \times T$ weighting matrix $W = [w_{tj}]_{t,j=1}^T$, computed using a kernel function:

$$w_{tj} = K\left(\frac{t-j}{H}\right) \quad \text{for} \quad t, j = 1, .., T \tag{5}$$

with a bandwidth H.

In the fixed parameter case, the weights on each likelihood sum up to T. In our case, each row of W is normalised to sum up to $2H + 1$, such that:

$$\sum_{j=1}^{T} w_{tj} = 2H + 1 \qquad t = 1, ..., T.$$

This normalisation is employed in order to maintain the relative weights between the likelihood and the prior.

For the application presented in this paper, the Normal kernel function is used to generate the weights:

$$w_{tj} = (1/\sqrt{2\pi}) \exp((-1/2)((t-j)/H)^2) \quad \text{for} \quad t, j = 1, .., T. \tag{6}$$

If the bandwidth H goes to infinity, the likelihood would collapse to the fixed parameter case, where each likelihood is weighted equally. If H is small, the weights are concentrated around a single observation. Our choice of bandwidth is $H = T^{0.5}$, motivated by the optimal bandwidth choice used for inference in time varying random coefficient models (see Giraitis et al. (2014)).

In the Bayesian framework, the local likelihood of the DSGE model at point t, denoted $L_t(Y|\theta_t)$, is augmented with the prior distributions for the parameters, $p_t(\theta_t)$, to get the posterior at time t, $p_t(\theta_t|Y)$:

$$p_t(\theta_t|Y) = \frac{L_t(Y|\theta_t)p_t(\theta_t)}{p(Y)} \propto \prod_{j=1}^{T} L(y_j|y_{j-1}, \theta_t)^{w_{tj}} p_t(\theta_t).$$

It should be noted that, for our DSGE application, we assume the prior $p_t(\theta_t)$ to be fixed over time, i.e., $p_t(\theta_t) = p(\theta_t)$ for all t. One could potentially allow the prior to be time varying, exploring the idea that the posterior yesterday can be used for a prior today. However, since we would like to explore only the possibility of parameter drift, we choose to be agnostic about time variation in the parameters before the estimation and keep the prior values fixed over time.

2.3.1 Characterising the Posterior Distributions

To obtain the joint posterior distribution of the parameters, we need numerical methods because the matrices F and G are non-linear functions of θ, and hence the posterior does not fall in known families of distributions with moments that could be derived analytically. The most commonly used procedure to generate draws from the posterior distribution of θ is the Metropolis-Hastings (MH) algorithm, proposed by Metropolis et al. (1953) and generalised by Hastings (1970). Although the posterior distribution could be obtained by other methods, such as the Importance Sampling (IS) algorithm, the MH algorithm delivers good convergence under fairly general regularity condition (see Geweke (1999, 2005)) and asymptotically normal posterior distribution (see Walker (1969), Crowder (1988) and Kim (1998)).

The algorithm described here is version of Schorfheide (2000)'s Random Walk Metropolis (RWM) algorithm, modified to include the kernel weighting scheme. Our aim is to obtain a sequence of posterior distributions $p_t(\theta_t|Y)$ for each point in time $t = 1, .., T$. At each t the algorithm implements the following steps.

Step 1: The posterior is log-linearised and passed to a numerical optimisation routine. Optimisation with respect to θ is performed to obtain the posterior mode:

$$\widehat{\theta}_t = \arg\min_{\theta} \left(-\sum_{j=1}^{T} w_{tj} \log L(y_j|y^{j-1}, \theta_t) - \log p(\theta_t) \right).$$

Step 2: Numerically compute $\widehat{\Sigma}_t$, the inverse of the (negative) Hessian, evaluated at the posterior mode, $\widehat{\theta}_t$.

Step 3: Draw an initial value θ_t^0 from $N(\widehat{\theta}_t, c_0^2\widehat{\Sigma}_t)$.

Step 4: For $i = 1, ..., n_{sim}$, draw ζ_t from proposal distribution $N(\theta_t^{(i-1)}, c^2\widehat{\Sigma}_t)$. Compute

$$r(\theta_t^{i-1}, \zeta_t|Y_{1:T}) = \frac{\prod_{j=1}^{T} L(y_j|y^{j-1}, \zeta_t)^{w_{tj}} p(\zeta_t)}{\prod_{j=1}^{T} L(y_j|y^{j-1}, \theta_t^{i-1})^{w_{tj}} p(\theta_t^{i-1})},$$

which is the ratio between the weighted posterior at the proposal ζ_j and θ_t^{i-1}.

The draw $\theta_t^{(i-1)}$ is accepted (setting $\theta_t^i = \zeta_t$) with probability $\tau_t^i = min\{1, r(\theta_t^{(i-1)}, \zeta_t|y_{1:T})\}$ and rejected ($\theta_t^{i-1} = \theta_t^i$) with probability $1 - \tau_t^i$. c_0^2 and c^2 are scaling parameters adjusting the step size of the MH algorithm in order to get desirable rejection rates such that we achieve convergence. The literature supports setting the scaling parameters such that acceptance rates of between 20% and 40% are achieved[4].

2.3.2 Computing Forecasts

We can compute forecasts for the observables y employing time varying the posterior distributions of the parameters. For this task, we only need the posterior distribution at the end of the sample, $p(\theta_{t=T}|Y)$, which contains the most recent values of the model's parameters and hence the most relevant information for predicting the future. Therefore, for generating DSGE-based predictions, our method is as computationally intensive as forecasting with standard fixed parameter DSGE models: it requires the computation of the posterior only once.

The predictive distribution of the sample $p(y_{T+h}|y_{1:T})$, h horizons ahead, is given by the conditional probability of the forecasts, averaged over all possible values of the parameters, the unobservables at the end of the sample x_T, and all possible future paths of the unobservables $x_{T+1:T+h}$: $p(y_{T+h}|y_{1:T}) =$

$$\int_{(x_T,\theta_T)} \left(\int_{x_{T+h}} p(y_{T+h}|x_{T+h}) p(x_{T+h}|x_T, \theta_T, y_{1:T}) dx_{T+h} \right) p(x_T|\theta_T, y_{1:T}) p(\theta_T|y_{1:T}) d(x_T, \theta_T)$$

where $p(\theta_T|y_{1:T})$ is the posterior of the parameters at the end point of the in-sample period, T. We use a slightly modified version of the algorithm for generating draws from the predictive distribution outlined in Del Negro and Schorfheide (2013b). It follows the steps:

[4]In particular, Roberts, Gelman and Gilks (1997) show that, under some conditions, the optimal asymptotic acceptance rate is 23.4%.

Step 1: Using the saved draws from the posterior at the end of the sample $p(\theta_T|y_{1:T})$, for every draw $i = 1, .., n_{sim}$ (or for every n-th draw if thinning is required), apply the Kalman filter to compute the moments of the unobserved variables at T using the density $p(x_T|\theta_T^i, y_{1:T})$.

Step 2: Draw a sequence of shocks $v_{T+1:T+h}^i$ from a $N(0, Q(\theta_T^i))$, where $Q(\theta_T^i)$ is a draw from the estimated posterior distribution of the diagonal variance-covariance matrix of the shocks at T. For each draw i from $p(\theta_T|y_{1:T})$ and from $p(x_T|\theta_T^i, y_{1:T})$, use the state equation to obtain forecasts for the unobserved variables:

$$\widehat{x}_{T+1:T+h}^i = F(\theta_T^i)x_{T:T+h-1}^i + G(\theta_T^i)v_{T+1:T+h}^i.$$

Step 3: Use the forecast simulations for the latent variables in the measurement equation:

$$\widehat{y}_{T+1:T+h}^i = D(\theta_T^i) + Z(\theta_T^i)\widehat{x}_{T+1:T+h}^i.$$

Once the simulated forecasts $\widehat{y}_{T+1:T+h}^i$ are obtained, they can be used to obtain numerical approximations of moments, quantiles and densities of the forecasts.

Point forecasts are obtained by computing the mean of the distribution of $\widehat{y}_{T+1:T+h}^i$ for each forecasting horizon.

3 A time varying DSGE Model

3.1 Model and Data

The DSGE model to which we apply our Bayesian Local Likelihood approach is the model from Smets and Wouters (2007), which is an extension of a small-scale monetary RBC model with sticky prices (such as Goodfriend and King (1997), Rotemberg and Woodford (1997), Woodford (2003), Ireland (2004) and Christiano et al. (2005)). In addition to the sticky prices, the model also contains some additional shocks and frictions, including sticky nominal price and wage settings with backward inflation indexation, investment adjustment costs, fixed costs in production, habit formation in consumption and capital utilization. It also features seven exogenous shocks that drive the stochastic dynamics of the model. The foundations of the model are derived from the intertemporal optimisation problems of different agents. In particular, there are seven types of agents in the model: consumers that supply labour, choose consumption level, hold bonds and make investment decisions;

intermediate goods producers which are in a monopolistically competitive market and cannot adjust prices at each period and final goods producers, who buy intermediate goods, package them and resell them to consumers in a perfectly competitive market. In addition, there is a labour market with a similar structure: there are labour unions with market power that buy the homogenous labour from households, differentiate it, set wages and sell it to the labour packers, who package it and resell it to intermediate goods producers in a perfectly competitive environment. Finally, there is a central bank that follows a nominal interest rate rule, adjusting the policy instrument in response to deviations of inflation or output from their target levels and a government that collects lump-sum taxes which appear in the consumer's budget constraint and whose spending is exogenously driven.

The model is log-linearised around its steady state and trended variables are detrended with a deterministic trend[5]. The model is estimated using seven macroeconomic quarterly time series for the United States for the period of 1964Q3 to 2012Q4 as observables. The variables are the ones used in Smets and Wouters (2007), namely, output, consumption, investment and wages per capita growth; inflation, hours and the interest rate (see Appendix for more details).

In this section we present results for the fixed parameter model and also for the version with time varying parameters estimated with BLL method described in the previous section. In both cases, we employ the priors from Smets and Wouters (2007), with a number of MH draws of $220,000$, from which we drop the first $20,000$. We set the scaling parameters such that acceptance rates are around 25%. We apply the BLL method using the Normal kernel function in equation (6) with a bandwidth size of $T^{0.5}$.

3.2 Results

In this section, we discuss the parameter estimates (Figures 1-3). We employ Figures 1-3 to judge informally whether a parameter's variation is substantial by checking whether the BLL estimates are outside the confidence bands of the fixed-parameter estimates. We adopt Fernandez-Villaverde and Rubio-Ramirez (2008)'s definition of 'structural' parameters: these are preference and technology parameters which are invariant to both policy and shocks. If

[5] The linearised model is presented in the Appendix. For full derivations from the non-linear first order conditons, please refer to the Technical Appendix in Smets and Wouters (2007) available at: http://www.aeaweb.org/aer/data/june07/20041254_app.pdf

a parameter is found to be within the fixed-parameter 68% bands, we conclude that it is in fact 'structural'. If a parameter varies smoothly over time, following a clear pattern, we infer that it has been a subject to structural change. On the other hand, if a parameter exhibits an erratic time variation, we would point to a possible misspecification of that parameter. The solid blue line are the time varying estimates obtained with the BLL method and 68% confidence bands are represented by the dotted blue lines. The green line represents the full sample fixed-parameter estimates, with the dotted lines around it - 68% confidence bands.

The top panel of Figure 1 assesses results for the policy preference parameters. Our results are broadly consistent with previous studies (Clarida, Gali and Gertler (2000), Cogley and Sargent (2002), Fernandez-Villaverde and Rubio-Ramirez (2008)) that found evidence of structural changes in Taylor rule parameters. The Federal Reserve has shifted its policy priority from output towards inflation since the parameter that measures the reaction to inflation increases between 1979 and 1996, while the reaction to the output gap decreases over this period. The interest rate smoothing parameter is lower in the 1980s than in later periods, while the steady state inflation rate decreases between 1985 and 1996.

The second panel of Figure 1 provides evidence of changes in the steady-state growth rate of per capita output (as well as consumption, investment and the real wage, which share the same trend). During most of the period and up to 2005, the BLL posterior mean for this parameter is around 0.4, that is, an annual growth rate of 1.6%, however, this decreases to 1% annually in period of the 2007-8 financial crisis. In contrast, fixed-parameter estimates under-estimate these values over most of the sample, but over-estimate it during the recent period. This parameter is important for generating forecasts as it appears in several of the measurement equations and Kolasa and Rubaszek (2015) bring attention to the importance of this parameter in reducing forecasting bias.

The findings on the price rigidity parameters are consistent with the evidence presented in Fernandez-Villaverde and Rubio-Ramirez (2008). In particular, we document a negative relation between the price indexation (last panel in Figure 1) and price stickiness (second panel in Figure 1) parameters after the mid-1970s; hence, periods characterised by high Calvo probability parameter are also of low indexation and vice versa. The fall in inflation indexation during the Great Moderation is consistent with findings of Gali and Gertler (1999) and could be explained by the decreased need to adjust prices frequently due to low and stable inflation leading to longer length of contracts[6] and hence higher Calvo probability parameter and lower indexation. The variation uncovered in the price rigidity parameters suggests that there is no stable predictive relation between inflation and output gap over time (i.e. the Phillips curve has become flatter in the low volatility period of the Great Moderation) which cast doubt on the ability of Calvo pricing models to adequately capture pricing behavior of firms and unions in the economy.

There are parameters that appear to move very little over the entire sample period or seem to remain within the confidence bands of the fixed parameter estimates throughout most of the sample, such as the elasticity of intertemporal substitution and the household discount factor (last panel of Figure 1), and the elasticity of labour supply or the fixed production costs (first panel of Figure 2). We draw comfort on those results, as they could be interpreted as evidence of the structural nature of these parameters.

On the other hand, the moving average (MA) coefficient, the persistence coefficient (last panel of Figure 2) and the standard deviation (last panel in Figure 3) of the wage mark-up shock process, as well as the Calvo parameter in labour markets (top panel in Figure 2) all appear very volatile and this could be evidence that they are seriously misspecified and should not be kept fixed. Interestingly enough, all four are parameters that govern labour market dynamics through the wage equation and all become very unstable during the Great Moderation period. This could be interpreted as evidence that during the Great Moderation, a Calvo model with an ARMA wage shock may not have been an adequate model to characterise the dynamics of the labour market in the US. An alternative interpretation is that there might be insufficient information in the data in order to jointly identify all four parameters during the Great Moderation.

The standard deviations of the structural shocks (panel 2 and 3 of Figure 3) also move in the expected direction, consistent with findings of low stochastic volatility during the Great Moderation (e.g. Primiceri (2005) and Sims and Zha (2006)). In particular, all shocks' volatilities fall in the late 1980s and remain low throughout the 1990s. Moreover, their posterior distributions are narrower during that period implying that there is less uncertainty about the possible values they can take. The standard deviation of the monetary policy shock, for instance, peaks in the 1980s, implying a larger role of the shock throughout that period and falls considerably after the 1990s, having lesser impact on the business cycle

[6] The average price duration is given by $\frac{1}{1-\xi_p}$, where ξ_p is the Calvo probability in the goods market.

as a consequence of the more adequate policy. For some shocks' standard deviations (e.g. TFP, investment-specific technology and price-mark up shocks) we observe an increase in the end of the sample leading to the recent financial crisis. Due to the considerable time variation we uncover in the volatilities of the shocks, using the most recent values of the estimated volatility parameters when generating forecasts is expected to improve the density of the forecasts compared to simply using the fixed parameter estimates that average these over the entire in-sample period. Finally, the autoregressive coefficients for the stochastic processes (panels 1 and 2 of Figure 3) seem to move considerably, which is unsurprising as these are designed to capture dynamics in the data. They are not truly structural, in the sense that there is no underlying macroeconomic theory that implies that they are not subjected to shocks or policy. Most of the shocks' persistence coefficients display a U-shape with low persistence towards the end of the Oil Crises and higher persistence during the Great Moderation. For instance, the TFP shock becomes very persistent during the recent crisis with AR coefficient very close to one, implying almost permanent shock to productivity.

3.3 Time Varying Impulse Response Functions

In this section, we turn to the estimated impulse response functions over time. We investigate whether there is evidence for structural change in the transmission mechanism of important variables to macroeconomic shocks, resulting from the documented time variation in the paramerers.

Figure 4 displays the impulse response functions of output, inflation and the interest rate to a monetary policy shock. Since the response is to a unit of the shock, it measures only changes in the transmission of the monetary policy shock over time without taking into account changes in the volatility of the shock, as documented in the previous subsection.

First, the response of the Fed rate to the monetary policy shock is roughly the same throughout the sample period and it is around half a percentage point. The response of output, in contrast, shows a clear trend over time, with responses increasing from around 2.5% to 4% on impact. The response of inflation, on the other hand, displays a U-shape, with inflation being quite responsive to policy in the late 1960s and early 1970s, and in the more recent period. This results are at odds with the findings in Boivin and Giannoni (2006), who find a considerable decrease in the responses of output, inflation and the interest rate to a policy shock in their post-1980 sub-sample, using minimum distance estimator between a structural VAR and a small DSGE model. They attribute this result mainly to the higher estimate of the inflation targeting parameter in their policy rule over the second period.

Figure 5 presents responses to a one standard deviation of the shock and these incorporate the decrease in the volatility of monetary policy shock over the second half of the sample. Instead of increasing responsiveness of output, we now observe somewhat constant response over time with an increase in the end of the sample due to the aggressive policy during the crisis. Interestingly, once one allows for changing size of the shock over time, inflation's response to policy is actually decreasing over time. Furthermore, one standard deviation policy shock results into considerably higher response of the interest rate during Volcker's years than in any other period.

Figures 6 and 7 display the responses to a price mark-up shock. The picture that emerges is that both output and investment become much more responsive to an inflation shock during the Great Moderation period, implying that the same shock to inflation has a relatively more harmful effect on these variables in that period than during the Oil Crises, when inflation was record high. It is also evident from the policy rate response that policy makers retaliated more to a unit of inflation shock after Paul Volcker's appointment as a Federal Reserve chairman.

Figure 8 and 9 are the IRFs of selected variables to a unit and a standard deviation of the TFP shock. The most intriguing result that emerges is that the response of all output, consumption and investment, whether one allows for the size of the shock to vary over time or not, is considerably larger in periods characterised by recessions such as the Oil Crisis in early 1970s and the recent crisis, implying an asymmetrically larger effects of TFP shocks in recessions than in booms. The decreasing responsiveness of the interest rate over time could be explained by the Federal Reserve responding less aggressively to output and more aggressively to inflation which is less affected by productivity shocks. The response of hours worked to productivity shock and the resulting implications for the relevance and relative importance of this shock for the business cycle is a much debated topic (Gali (1999)). Once we allow for time variation in the TFP responses, the response of hours remains negative in all periods for all horizons except several periods in the early 1990s when the response changes sign and becomes positive after less than 10 quarters. This could be attributed to the increased persistence of the shock that we uncover during this period. Furthermore, as argued in Smets and Wouters (2007), the habit coefficient is important for explaining the negative effect of TFP on hours and as shown in Figure 1, we obtain low habit persistence in the 1990s which contributes to the weakened duration of the negative effect.

Figures 10 and 11 display the responses to a preference shock. It appears that, after allowing for the changing size of the shock over time, output, consumption and investment are more responsive to the shock, both on impact and in duration, in periods characterised by recessions, suggesting asymmetric responses to the preference shock in the model.

Finally, Figures 13 and 14 display the responses to a unit and a standard deviation of wage mark up shock respectively. The parameters characterising the labour market during the Great Moderation period, which we discussed in the previous Section, are the reason for the misbehaved impulse response functions during the same period. It is clear that the responses are not smooth over time and the resulting response per unit of the shock of output, inflation and hours becomes essentially zero for all horizons after the beginning of the 1990s. Once we allow for the changing size of the shock, the picture becomes even more distorted, since the standard deviation of the wage mark up shock is itself one of the parameters that are misspecified during the Great Moderation period[7].

4 Forecasting with a time varying DSGE Model

As we discussed in the introduction, the literature has documented evidence of the forecasting accuracy of fixed-parameter DSGE models (Smets and Wouters (2007), Edge et al. (2009), Edge and Guerkaynak (2010), Del Negro and Schorfheide (2013b), Del Negro, Giannoni and Schorfheide (2014)). In this section we evaluate the relative forecasting performance of our time varying DSGE model. In addition to the fixed-parameter Smets and Wouters (2007) specification, we also compare the forecasting record of the time varying DSGE model against univariate models (AR(1), a Random Walk (RW) and a time varying AR(1)) and multivariate reduced-form models (a BVAR and a time varying stochastic volatility BVAR (TV-SV BVAR)).

The BVAR uses a standard Normal-inverted-Wishart conjugate prior with optimal shrinkage and optimal lag selection as in Carriero, Clark and Marcellino (2015). The TV-SV BVAR features time varing autoregressive coefficients as in Cogley and Sargent (2002) and stochas-

[7]The IRFs of selected variables to the remaining shocks (namely, Investment Technology and Government Spending Shocks) can be found in the Appendix, see Figures 12-17.

tic volatility as in Primiceri (2005)[8]. Since it is burdensome to estimate this model for more than three variables and obtain stationary draws from the posterior distribution of the autoregressive coefficients, we limit our TV-SV BVAR to only output growth, inflation and the Fed Funds rate. The TV-AR model is computed using the non-parametric kernel based method, as in Giraitis et al. (2014)[9].

We employ the algorithm outlined in Section 2.3.2 to generate density forecasts for the observables of the time varying DSGE model.

Since real-time data is limited and only available after 1991[10], we perform the out-of-sample forecasting on final revised data as we would like to be able to assess performance across different periods. Our forecast origins range from 1974Q3 up to 2010Q1 and we compute forecasts for one up to twelve quarters ahead.

We measure accuracy of point forecasts using the root mean squared forecast error (RMSFE) and forecast bias. The accuracy of density forecasts are measured by log predictive scores. We compute the logscore with the help of a nonparametric estimator to smooth the draws from the predictive density obtained for each forecast and horizon. We test whether a model is statistically more accurate than the benchmark with the Diebold and Mariano (1995) statistic computed with Newey-West estimator to obtain standard errors. We provide the results of the Diebold-Mariano test for the RMSFEs and logscores. For the bias, we simply test whether the models' bias is statistically different from zero.

4.1 Point Forecasts

Table 1 presents the absolute performance of the our TV DSGE model (in RMSFEs) and the relative performance of our approach to alternative models over different horizons (numbers smaller than one imply superior performance of the TV DSGE relative to the alternatives).

[8]The TV-SV BVAR is of lag order one and uses random walk processes for both the autoregressive coefficients and the log volatility. For more details, see for instance Cogley, Primiceri and Sargent (2010) or Benati and Mumtaz (2007).

[9]The model is estimated in each point in time t : $\widehat{\beta}_t = (X'D_tX)^{-1}X'D_tY$ where X contains the lagged dependant variable Y and D_t is a diagonal matrix with the kernel weights of the t^{th} row of the weighting matrix in equation (5) in its main diagonal. The variance of the residuals is also time varying and computed in point t as $\widehat{\sigma}_t^2 = \varepsilon' D_t \varepsilon / tr(D_t)$. Density forecasts are then generated, using wild bootstrap and the last period values $\widehat{\beta}_T$ and $\widehat{\sigma}_T^2$.

[10]Real-time data on compensation is not available prior to 1991.

One, two and three stars indicate that we reject the null of equal accuracy in favour of the better performing model at significance levels of 10%, 5% and 1% respectively.

There are some gains from using the time varying model over the standard fixed parameter DSGE for most variables but the differences are small and rarely significant. One exception is inflation: the time varying model performs significantly worse than the fixed-parameter specification. The model also outperforms the standard BVAR, which confirms previous findings (e.g. Smets and Wouters (2007), Adolfson, Andersson, Linde, Villani and Vredin (2007), Christoffel, Coenen and Warne (2010)). Moreover, we find superior performance for output growth over the TV-SV BVAR. Finally, the TV DSGE model strongly outperforms the univariate models.

In order to better understand the strengths and weaknesses of our approach, we further investigate the point forecast accuracy by splitting our sample of forecasts into subsamples corresponding roughly to three distinct periods in U.S. recent economic history: namely, the Oil Crisis or Great Inflation period (at least the end of it, ranging 1974Q2:1982Q4), the Great Moderation (1983Q1:2005Q4) and the recent financial crisis (2006Q1:2011Q3). Table 2 presents the relative RMSFE performance of our approach and Table 3 displays the forecast bias for per capita GDP growth, inflation and the interest rate during the three periods. For the Oil Crisis period, our method is superior to the two BVARs and comparative to the standard DSGE approach for output. When it comes to forecast bias in the Great Inflation Period, both DSGE models strongly and significantly underestimate inflation, but in relative terms our model does poorly, resulting in significantly worse RMSFE performance. The reason for this result is the relatively small sample size at this point and hence, little advantage in down-weighting past data. Interestingly, the two BVARs deliver unbiased inflation forecasts, but systematically underestimate output growth.

Although, as argued earlier, our method does on average worse for inflation over the full forecast sample, during the Great Moderation, we obtain better performance. The Great Moderation was characterised by low and stable inflation and very low business cycle volatility. It is clear from the forecast bias, that both BVAR models and the standard DSGE model significantly overestimate inflation during this period. This is the case, since the standard DSGE and the fixed coefficient BVAR, in order to generate projections, use samples which contain the entire Oil Crisis period, characterised by very high inflation. Our method, on the other hand, obtains better performance and remains virtually unbiased because of its way of down-weighting distant data. The TV-SV BVAR features time variation in the autoregressive coefficients of the VAR, so it is surprising that is fails to capture the structural change in inflation dynamics during the Great Moderation and also systematically overestimates inflation. This could be due to the way time variation enters the model; our approach models time variation non-parametrically, and it is more robust to misspecifications in the stochastic processes for the time varying parameters that the TV-SV BVAR utilises (for further discussion and Monte Carlo evidence, see Galvão et al. (2015)). All models deliver similar, in terms of RMSFEs, output forecasts during the Great Moderation; however, both DSGE models underestimate output.

The period of the recent financial crisis (2007-2009) has been a subject to many discussions. This crisis generated serious critiques for the forecasting literature, for instance, Wieland and Wolters (2011), Del Negro and Schorfheide (2013b) provide evidence that DSGE models not only failed to predict it, but even once the crisis had started, failed to forecast the trough and quickly returned the economy to positive growth rates. During the 2006-2011 period, which overlaps with the recent crisis period and subsequent recovery, our model outperforms the standard DSGE model and both BVARs for all horizons and even with a small sample size of 23 observations, manages to deliver some statistically significant improvements. This could also be seen from the forecast bias where all alternative models considerably overestimate output, but our approach impressively delivers unbiased output forecasts at one step ahead, compared to a bias of around 0.31% quarterly GDP growth for the fixed parameter DSGE model, 0.33% for the BVAR and 0.40% for the TV-SV BVAR. Our interpretation of this result is similar to before; both the BVAR model and the standard DSGE model use as in-sample period data from Oil Crisis and the Great Moderation in order to generate forecasts for the recent crisis, while our method also makes use of these data but discounts it progressively fast. The resulting trend coefficient, $\overline{\gamma}$, as seen in Section 4, falls considerably after 2007. This parameter is important and can have substantial effect on the model's forecasts, as it enters as an intercept in the measurement equation for output, consumption, investment and wage growth. For inflation, while RMSFE performance during the same period is relatively similar to the standard DSGE model (worse in the short run and better at longer horizons), the forecast bias is negative for the TV DSGE model while positive for the standard DSGE. Another interesting result is that, our method, while delivering similar interest rates forecasts during the previous subsamples, delivers very large and statistically significant improvements over all models during the crisis. Our TV DSGE model contains a Taylor rule with changing parameters and in particular, during the crisis period with interest rates close to the Zero Lower Bound (ZLB), our interest rate smoothing coefficient jumps to levels near 0.9. This delivers interest rate forecasts that are close to a random walk model while inflation targeting and output gap values have a reduced effect.

To summarise, in periods, in which serious structural change is present, such as the recent one, the forecast errors of the TV DSGE model are relatively smaller and the model is more robust to forecast bias resulting from the presence of the structural change in the in-sample period.

4.2 Density Forecasts

Table 4 accesses the quality of the density forecasts measured by logscores of the predictive density. The table displays absolute log predictive score for the TV DSGE model and differences in logscores over the alternative models, so numbers greater than zero imply superior performance of our approach.

A few comments are in order. First, it is clear that overall our method outperforms considerably and statistically significantly the standard DSGE for most variables and horizons. Interestingly, our simple univariate TV AR(1) also delivers density performance superior to that of the standard DSGE model. These results are important as they imply that, while allowing for parameter drift results in moderate gains for point accuracy and only for some variables and periods, it results in significantly improved density forecasts. One way to look at this is to infer about the importance of stochastic volatility. As seen in Section 3, the uncovered time variation in the standard deviations of the shocks in the structural model is substantial and previous studies have confirmed this result (Primiceri (2005), Sims and Zha (2006), Justiniano and Primiceri (2008)). Since volatility is inherently time varying and subjected to structural change, it is clear that conditioning on the most recent values of the shocks when simulating the density of forecasts, as outlined in Section 2.3.2, will deliver better forecast uncertainty. Since the TV-SV BVAR also allows for changing volatility, it is unsurprising that it delivers very similar density forecast performance.

Table 5 further investigates the relative density performance of our approach over the sub-sample periods. It is clear that our method delivers better forecast uncertainty than the standard DSGE and the standard BVAR over the Great Moderation period, which is unsurprising. The in-sample that the two fixed coefficient models contain is the high volatility period of the Oil Crises, hence, as anticipated, density forecasts during the Great Moderation generated with these in-samples are worse. On the other hand, our TV DSGE model as well as the TV-SV BVAR account for this reduced uncertainty (our approach - by kernel down-weighting of Oil Crises data and the TV-SV BVAR - by fitting random walk state equations for the volatility paths) and hence deliver similar and superior performance over the fixed parameter models.

4.3 Robustness Checks

In this section, we investigate the impact of some of our assumptions on the forecasting performance of the time varying DSGE model. We exploit the impact of different bandwidth sizes and the use of the rolling windows method. Figure 14 plots the RMSFEs[11] for the fixed-parameter DSGE model and the time varying DSGE estimated under different assumptions. We include our baseline case with the normal kernel method and bandwidth equals to $T^{0.5}$, and also the case the bandwidth of $T^{0.55}$. We also consider flat kernels which are equivalent to rolling windows of 40 and 60 observations. The results in Figure 14 support our baseline estimation method since they imply a forecasting performance that is superior to alternative specifications.

5 Conclusion

This paper develops a Bayesian local likelihood method to accommodate time variation in DSGE models' parameters, appropriately weighting the sum of log likelihoods of the sample with weights generated by a kernel function. The method can be applied to any DSGE model that can be cast in a state-space representation. The empirical application presented uncovers some time variation in the Smets and Wouters (2007) model's parameters and points to potential misspecification in the labour market of the model during to the low volatility environment of the Great Moderation period. When it comes to forecasting, our estimation procedure is no more computationally intensive than estimating a DSGE model

[11]For computational time considerations, the robustness checks have only been computed at the mode of the parameter posterior. Furthermore, the sample size of the forecasts is smaller than in the comparison in the previous Section due to use of larger in-sample periods by the wider rolling windows.

with fixed parameters and, as demonstrated in the forecasting exercise, can deliver some gains in the forecast performance both for point and density projections especially in the presence of serious structural change such as the recent financial crisis.

Figure 1: The DSGE parameters over time

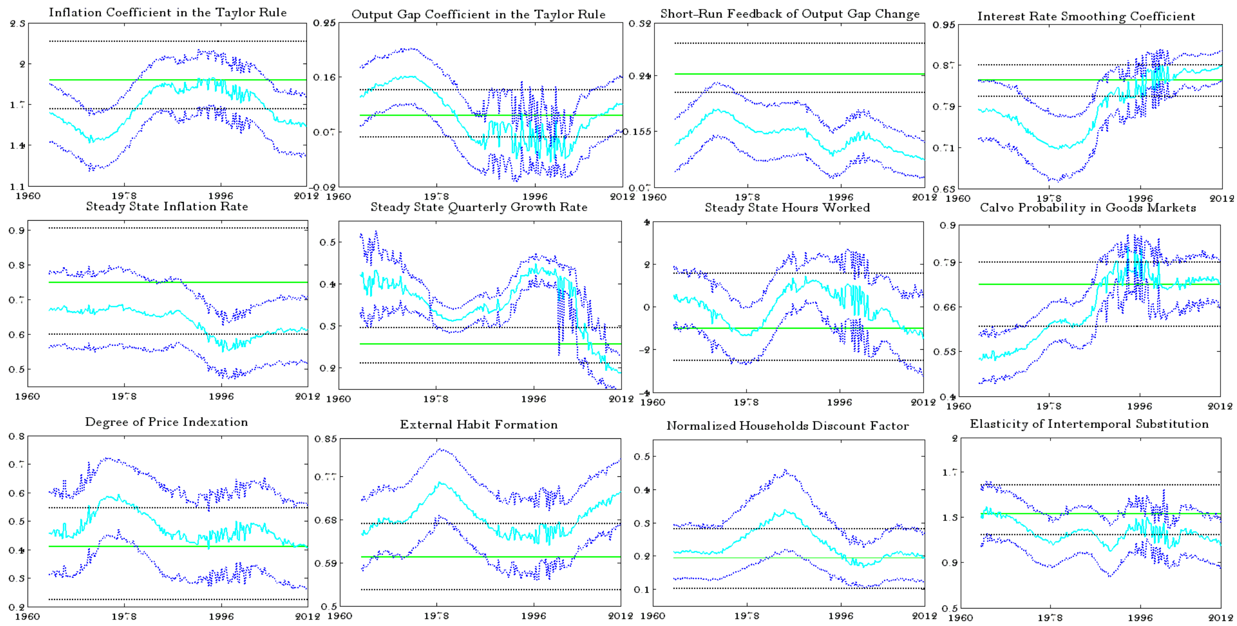


Figure 2: The DSGE parameters over time

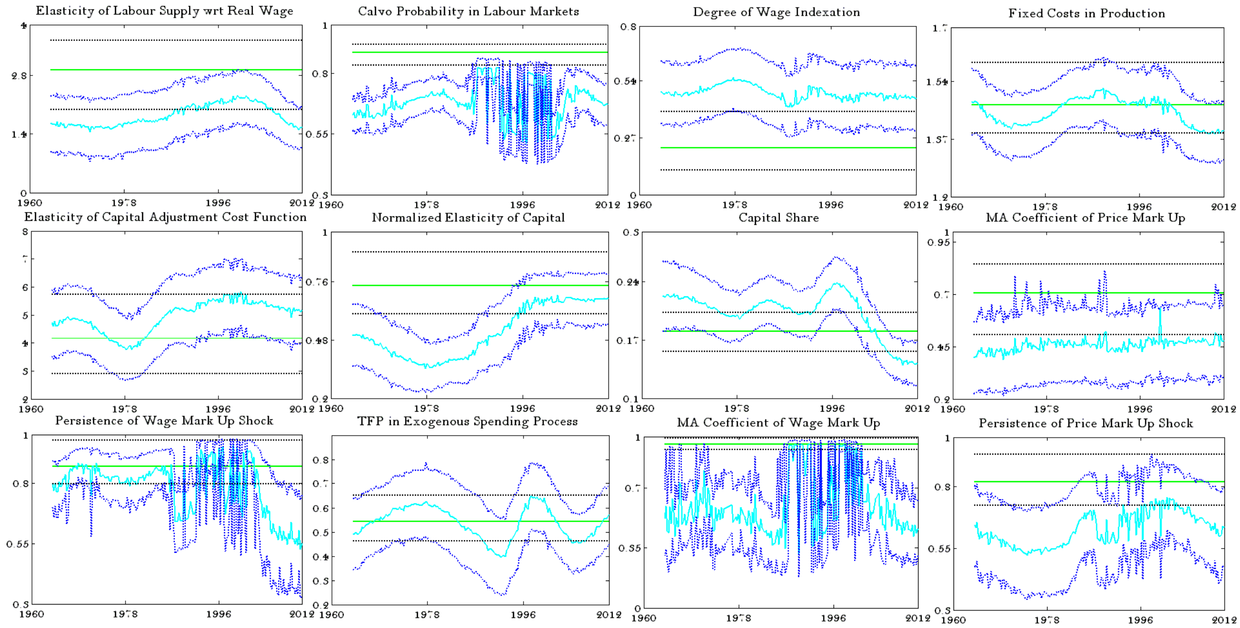


Figure 3: The DSGE parameters over time

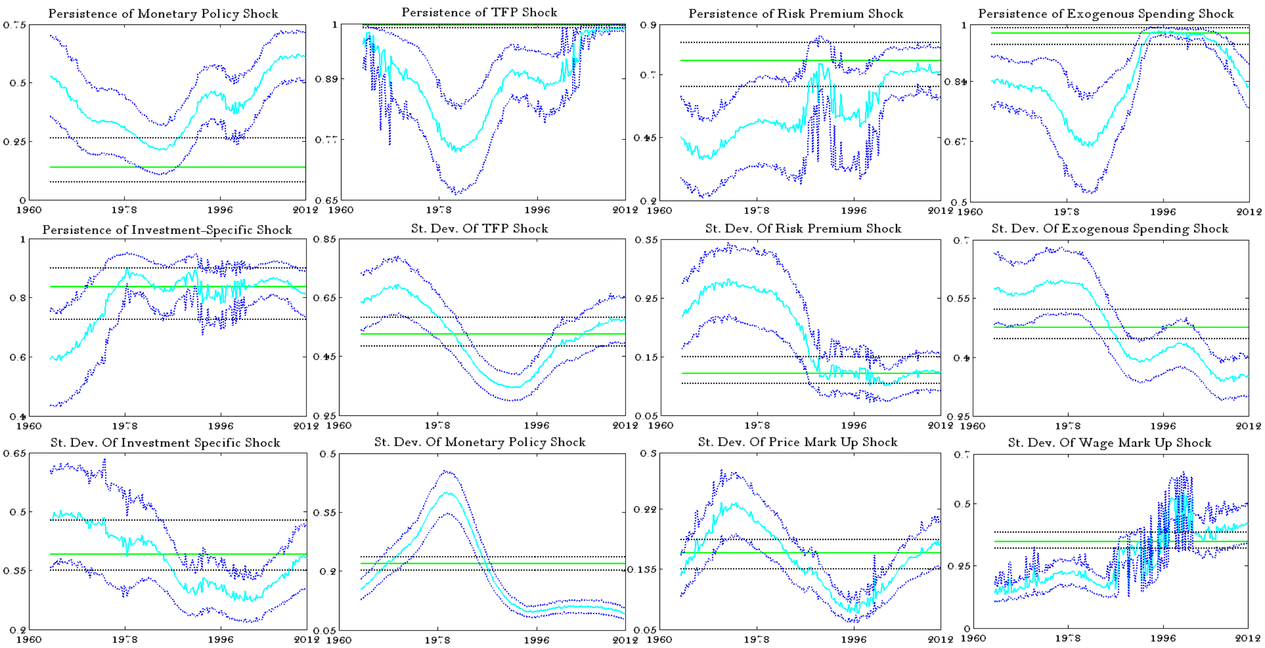


Figure 4: IRFs to 1 unit monetary policy shock

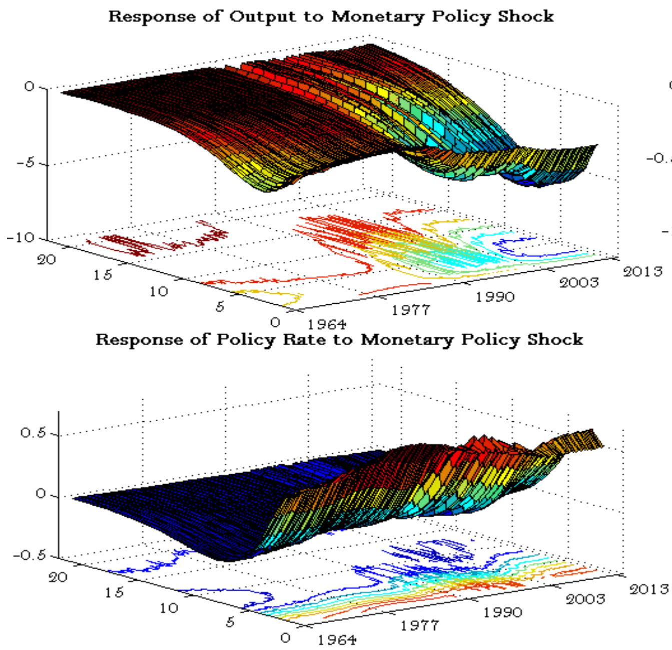


Figure 5: IRFs to 1 st. dev. monetary policy shock

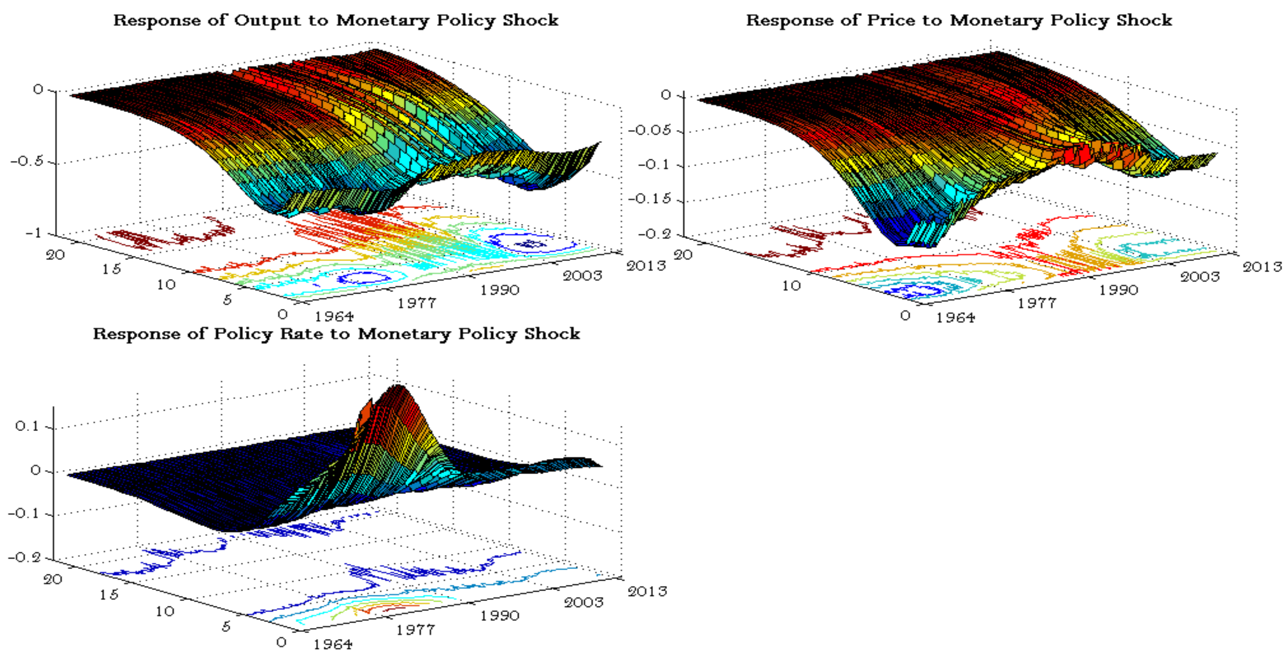


Figure 6: IRFs to 1 unit price mark up shock

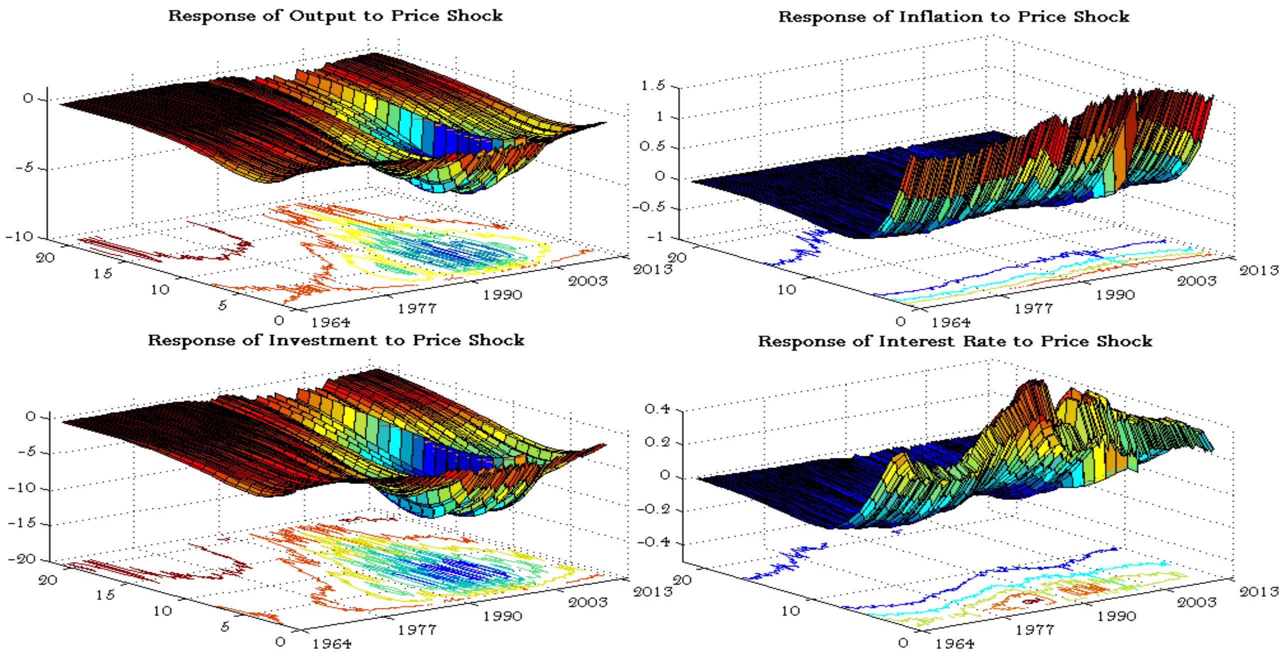


Figure 7: IRFs to 1 st. dev. price mark up shock

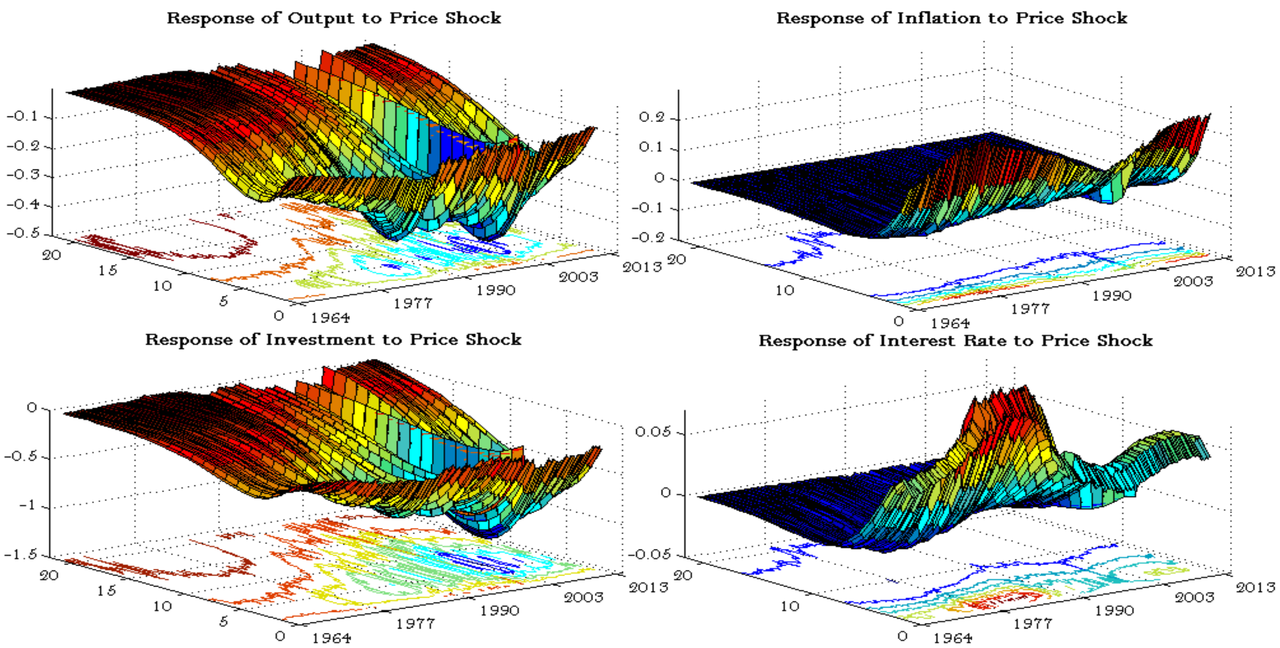


Figure 8: IRFs to 1 unit TFP shock

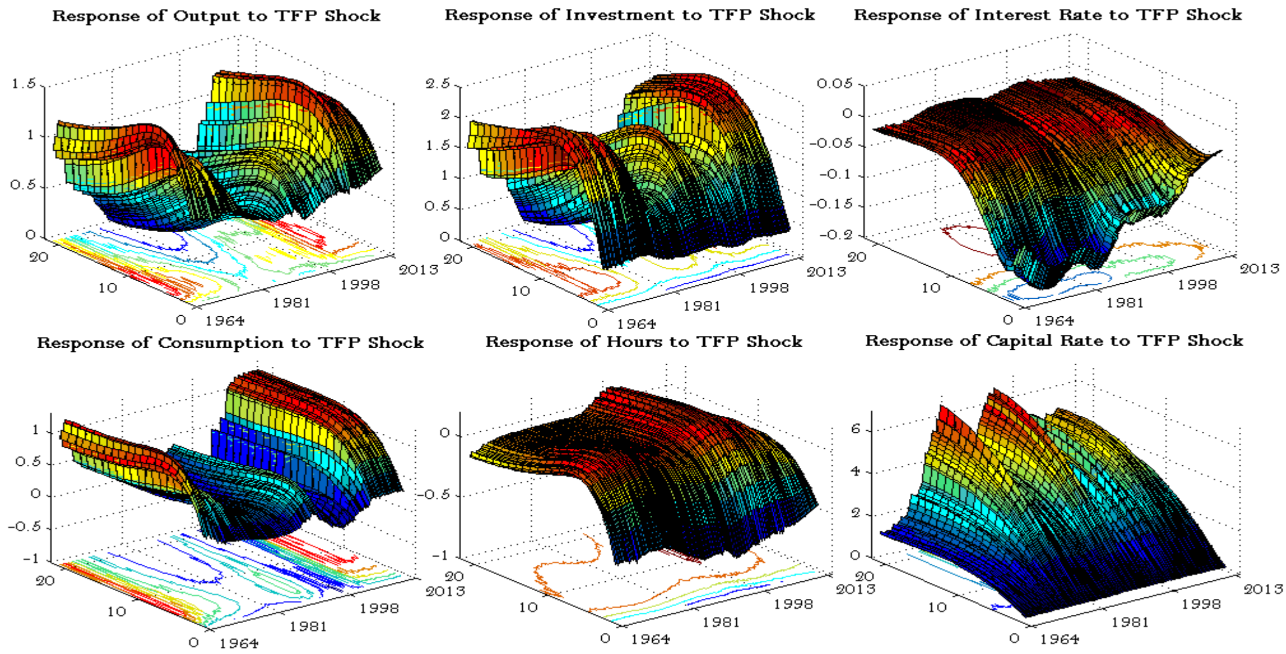


Figure 9: IRFs to 1 st. dev. TFP shock

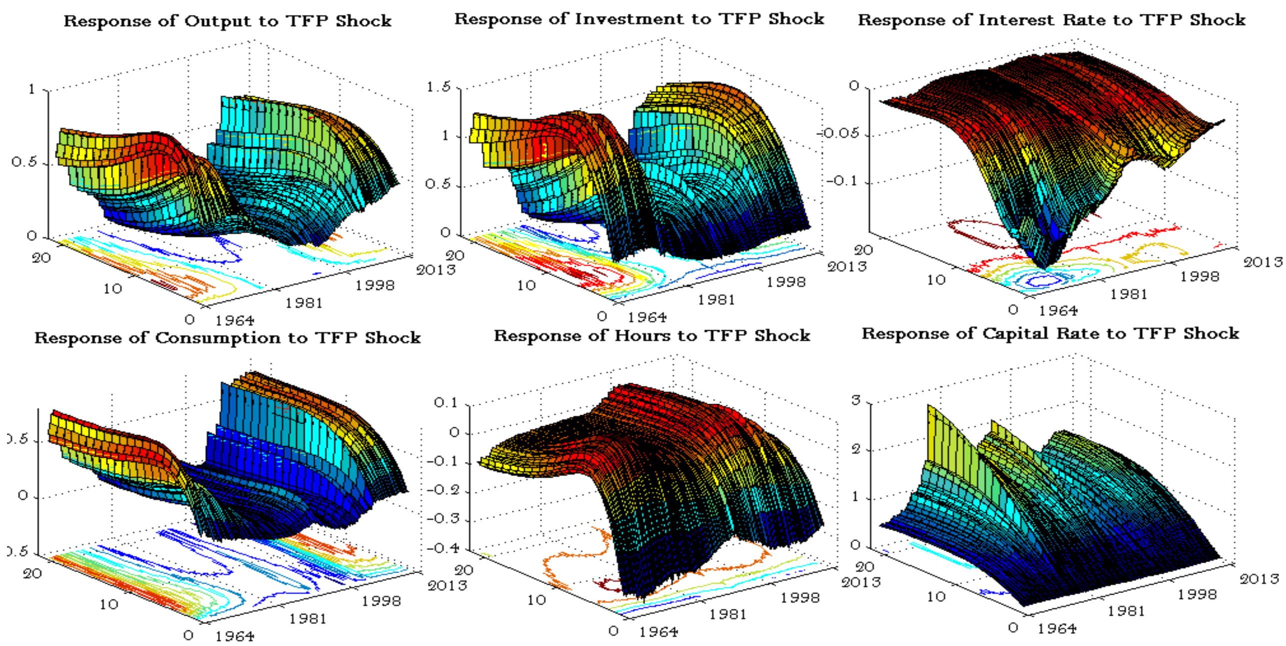


Figure 10: IRFs to 1 unit preference shock

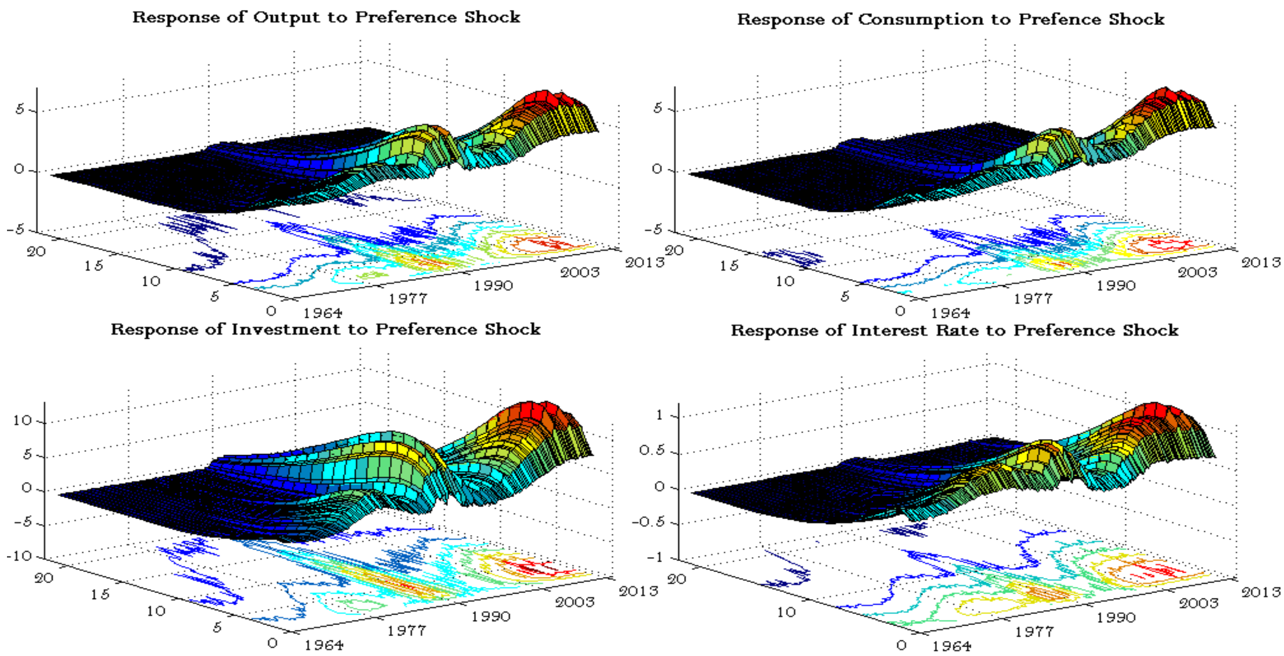
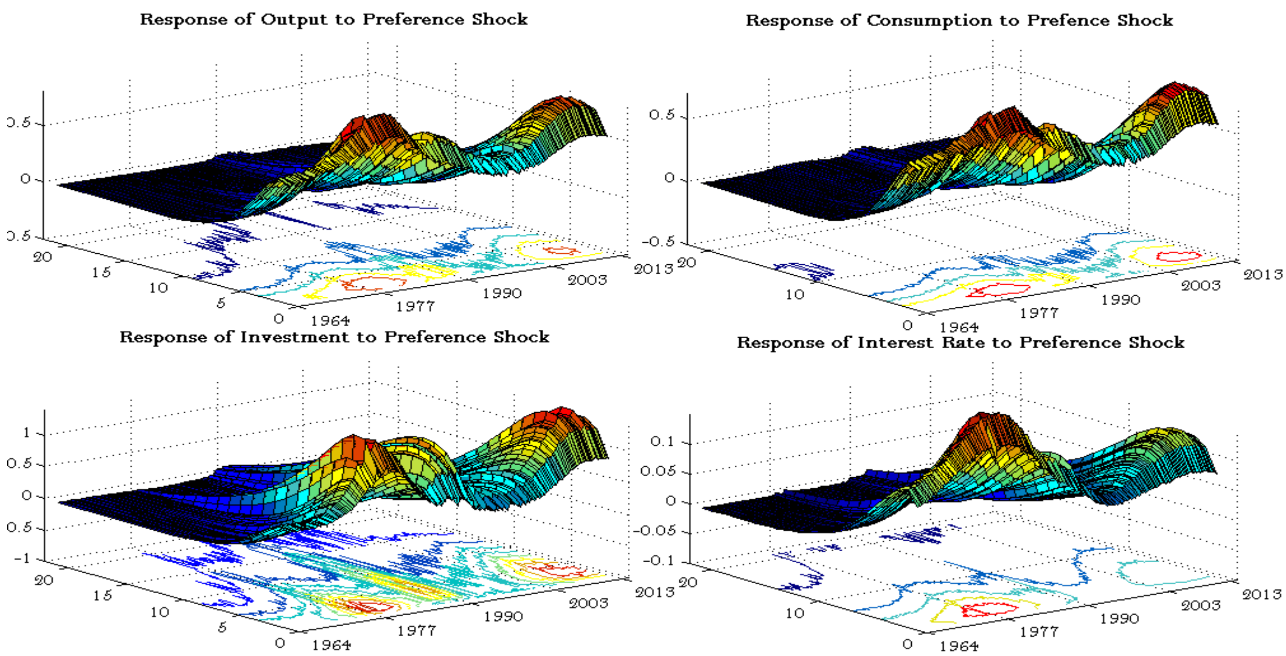


Figure 11: IRFs to 1 st. dev. preference shock



	Horizon	Output	Consumption	Investment	Wages	Hours	Inflation	Interest Rate
Forecast Performance - RMSFEs: 1974Q3-2011Q3								
TV-DSGE	1	0.72	0.66	1.88	0.74	0.63	0.28	0.26
	2	0.79	0.69	2.14	0.74	1.07	0.38	0.40
	4	0.82	0.69	2.30	0.73	1.83	0.48	0.55
	8	0.77	0.67	2.17	0.74	2.69	0.48	0.72
	12	0.77	0.68	2.15	0.78	3.15	0.49	0.81
TV-DSGE / F-DSGE	1	0.99	0.91**	1.04*	1.01	1.04	1.09***	1.01
	2	0.98	0.87**	1.07*	1.00	1.00	1.14***	1.01
	4	0.98	0.88**	1.08	1.00	0.97	1.17**	0.98
	8	0.97	0.93*	1.05	1.00	0.91	1.07	0.94
	12	0.99	0.96	1.03	0.99	0.89	1.00	0.90
TV-DSGE / BVAR	1	0.89	0.92	0.99	0.97	1.07*	0.97	0.99
	2	0.94	0.92	0.99	0.99	1.05	0.99	0.97
	4	0.90	0.90	0.98	1.00	0.98	1.02	0.99
	8	0.92	0.90	0.98	0.97	0.87	0.79	0.94
	12	0.93	0.94	0.97	0.98*	0.80	0.72	0.90
TV-DSGE/TV-SV BVAR	1	0.88	-	-	-	-	0.98	1.03
	2	0.97	-	-	-	-	1.00	1.08
	4	0.92	-	-	-	-	1.04	1.05
	8	0.83	-	-	-	-	0.82	0.94
	12	0.81	-	-	-	-	0.74	0.90
TV-DSGE / AR (1)	1	0.90*	0.94	1.02	1.00	0.88*	1.03	0.96
	2	0.94	0.94	1.01	1.02	0.86*	1.07	0.98
	4	0.97	0.96	1.00	0.99	0.86*	1.08	0.95
	8	0.97	0.97	0.97	0.98	0.82**	0.93	0.84*
	12	0.98	0.99	0.97	0.99	0.80*	0.84	0.80**
TV-DSGE / RW	1	0.70**	0.72**	0.79*	0.74**	0.52**	0.86*	0.63**
	2	0.71**	0.74**	0.79**	0.72**	0.64**	1.02	0.84
	4	0.69**	0.72**	0.76**	0.69**	0.76**	1.08	0.86
	8	0.70**	0.67**	0.67**	0.74**	0.80**	0.96	0.83*
	12	0.63**	0.66**	0.59**	0.72**	0.80**	0.90	0.80**
TV-DSGE / TV-AR	1	0.86**	0.87**	0.93	0.97	0.86**	1.01	0.94
	2	0.89*	0.84*	0.86*	0.98	0.82**	1.04	0.98
	4	0.90*	0.82	0.81*	0.96	0.80**	1.07	0.86
	8	0.88**	0.74	0.71**	0.95**	0.73**	0.94	0.73**
	12	0.87**	0.73	0.66*	0.94**	0.68**	0.89	0.64**

Table 1: RMSFEs. The figures under TV-DSGE model are absolute RMSFEs, the figures under the remaining models are ratios of RMSFEs of TV-DSGE over the alternatives. '*', '**' and '***' indicate rejection of the null of equal performance against the two-sided alternative at 10%, 5% and 1% significance level respectively, using a Diebold - Mariano test.

Forecast Performance - RMSFEs										
		Oil Crisis			Great Moderation			Recent Crisis		
		1974Q3-1982Q4			1983Q3-2005Q4			2006Q1-2011Q3		
	Horizon	Output	Inflation	Interest Rate	Output	Inflation	Interest Rate	Output	Inflation	Interest Rate
TV-DSGE	1	0.96	0.46	0.46	0.65	0.18	0.16	0.60	0.26	0.10
	2	1.16	0.67	0.71	0.62	0.20	0.27	0.72	0.29	0.18
	4	1.22	0.90	0.91	0.64	0.22	0.39	0.70	0.35	0.33
	8	1.13	0.89	1.17	0.61	0.25	0.53	0.69	0.28	0.48
	12	1.18	0.91	1.34	0.57	0.27	0.56	0.71	0.19	0.53
TV-DSGE / F-DSGE	1	0.98	1.15***	1.03	1.02	0.97	0.99	0.90	1.13*	0.76**
	2	1.08**	1.19***	1.05	0.92**	0.96	0.96	0.86	1.21	0.79*
	4	1.06*	1.19*	1.01	0.98	0.98	0.96	0.79	1.34	0.79
	8	1.01	1.11	0.99	1.00	0.98	0.99	0.82	0.91	0.63
	12	1.03	1.05	0.97	0.97	0.96	1.04	0.88	0.54	0.52
TV-DSGE / BVAR	1	0.81	1.07	1.04	1.04	0.76**	1.03	0.76	1.13	0.47*
	2	1.04	1.12	0.98	0.95	0.69**	1.13*	0.72	1.14	0.53
	4	0.97	1.26	0.99	1.01	0.57**	1.16*	0.62*	1.05	0.65
	8	0.96	1.00	0.96	1.09	0.50**	1.03	0.64	0.62	0.66
	12	0.91	0.95	1.00	1.09	0.46**	0.92	0.77	0.37	0.54
TV-DSGE/TV-SV BVAR	1	0.83	1.03	1.02	0.97	0.84*	1.06	0.80	1.09	0.86*
	2	1.12	1.10	1.10	0.89	0.73*	1.06	0.81	1.22	0.82
	4	1.02	1.21	1.09	0.88	0.61	1.02	0.76	1.32	0.83
	8	0.81	0.88	0.95	0.91	0.62	1.01	0.75	1.02	0.73
	12	0.76	0.77	0.94	0.93	0.62	0.91	0.77	0.78	0.67

Table 2: RMSFEs. The figures under TV-DSGE model are absolute RMSFEs, the figures under the remaining models are ratios of RMSFEs of TV-DSGE model over the alternatives. '*', '**' and '***' indicate rejection of the null of equal performance against the two-sided alternative at 10%, 5% and 1% significance level respectively, using a Diebold - Mariano test.

Forecast Performance - Bias										
		Oil Crisis			Great Moderation			Recent Crisis		
		1974Q3-1982Q4			1983Q3-2005Q4			2006Q1-2011Q3		
	Horizon	Output	Inflation	Interest Rate	Output	Inflation	Interest Rate	Output	Inflation	Interest Rate
TV-DSGE	1	0.11	-0.23**	-0.06	-0.10	0.00	-0.01	-0.04	-0.08	0.01
	2	0.15	-0.39**	-0.14	-0.19**	0.01	0.00	0.07	-0.13*	0.03
	4	0.18	-0.60**	-0.34	-0.25**	0.03	0.02	0.16	-0.17	0.09
	8	-0.02	-0.63	-0.58	-0.25**	0.02	0.03	0.18	-0.12	0.24
	12	-0.06	-0.66	-0.88	-0.18	0.01	0.05	0.13	-0.04	0.45
F-DSGE	1	0.06	-0.18**	-0.07	-0.18**	0.02	-0.01	0.31**	-0.01	0.04
	2	0.04	-0.32**	-0.17	-0.30**	0.04	-0.01	0.48**	0.00	0.11*
	4	0.09	-0.52**	-0.41	-0.34**	0.08**	-0.01	0.55*	0.06	0.28*
	8	-0.02	-0.58*	-0.69	-0.27**	0.11*	0.03	0.49	0.19	0.64
	12	-0.04	-0.63	-1.00	-0.20	0.13*	0.10	0.39	0.29	0.96
BVAR	1	-0.41**	0.03	-0.06	0.02	0.11***	-0.03	0.33**	0.02	-0.11**
	2	-0.47**	0.05	-0.18	0.03	0.17***	-0.01	0.61**	0.03	-0.13
	4	-0.28	0.05	-0.44*	0.01	0.27***	0.06	0.86**	0.05	-0.04
	8	-0.16	0.13	-0.62	-0.07	0.42***	0.22	0.84*	0.21	0.35
	12	-0.13	0.15	-0.78	-0.09	0.51**	0.37*	0.59	0.38	0.79
TV-SV BVAR	1	-0.25	-0.05	-0.01	-0.05	0.07***	0.02	0.40***	-0.01	0.00
	2	-0.36*	-0.04	-0.03	-0.06	0.12***	0.04	0.55**	0.00	0.03
	4	-0.42*	0.01	-0.07	-0.06	0.17**	0.09	0.61*	0.01	0.11
	8	-0.67*	0.22	-0.03	-0.02	0.19*	0.15	0.61	0.07	0.34
	12	-0.73	0.33	-0.14	0.04	0.20	0.22	0.56	0.14	0.60

Table 3: Forecast Bias. The table reports forecast bias, computed as the mean forecast error. '*', '**' and '***' indicate rejection of the null of zero bias against the two-sided alternative at 10%, 5% and 1% significance level respectively, using a Diebold - Mariano test.

Forecast Performance - Log Predictive Score: 1974Q3-2011Q3								
	Horizon	**Output**	**Consumption**	**Investment**	**Wages**	**Hours**	**Inflation**	**Interest Rate**
	1	-1.14	-1.05	-2.00	-1.28	-0.97	-0.02	0.30
	2	-1.20	-1.06	-2.17	-1.10	-1.50	-0.25	-0.30
TV-DSGE	4	-1.23	-1.10	-2.28	-1.09	-2.15	-0.45	-0.82
	8	-1.22	-1.07	-2.22	-1.11	-2.65	-0.52	-1.19
	12	-1.21	-1.05	-2.23	-1.19	-2.81	-0.53	-1.36
	1	0.05**	0.05	0.03	0.00	-0.04	0.02	0.34***
	2	0.06**	0.12***	0.02	0.12	0.03	-0.01	0.25**
TV-DSGE - F-DSGE	4	0.06**	0.11***	-0.03	0.05	0.06	0.00	0.13
	8	0.07*	0.06	-0.04	0.05	0.19	0.07	0.08
	12	0.07	0.07*	0.00	0.03	0.42	0.15**	0.18
	1	0.09	0.03	0.06	-0.02	-0.10**	0.14**	0.39***
	2	0.08	0.05	0.10	0.15	-0.01	0.16*	0.22**
TV-DSGE - BVAR	4	0.11	0.07	0.09	0.11	0.02	0.20*	0.01
	8	0.05	0.06	0.03	0.15	0.13	0.34**	0.01
	12	0.04	0.08*	0.07	0.12	0.42	0.45**	0.09
	1	0.04	-	-	-	-	-0.01	-0.05
	2	0.03	-	-	-	-	-0.06	-0.06
TV-DSGE-TV-SV BVAR	4	0.00	-	-	-	-	-0.09	-0.07
	8	0.09	-	-	-	-	0.01	0.05
	12	0.10	-	-	-	-	0.08	0.06
	1	0.03	-0.05	0.09*	0.05	0.16**	0.12*	0.28***
	2	-0.04	-0.09	-0.04	0.19***	0.25*	0.11*	0.19**
TV-DSGE - AR (1)	4	-0.05	-0.12	-0.06	0.24***	0.30	0.18**	0.10
	8	-0.09	-0.14**	-0.05	0.26***	0.43*	0.29**	0.24
	12	-0.11	-0.13*	-0.05	0.23**	0.48**	0.40**	0.37
	1	0.25***	0.23**	0.35***	0.39***	1.02***	0.34***	0.79***
	2	0.33***	0.26**	0.21	0.37***	0.69***	0.13*	0.41***
TV-DSGE - RW	4	0.51***	0.49***	0.26	0.39***	0.41**	0.17*	0.23
	8	0.84***	0.84***	0.57***	0.55***	0.44**	0.38***	0.15
	12	1.06***	1.07***	0.74***	0.66***	0.29	0.57***	0.13
	1	0.12*	0.11	0.20**	0.12	0.23***	0.13**	0.29***
	2	0.10	0.06	0.12	0.27***	0.42***	0.10	0.34**
TV-DSGE - TV-AR	4	0.09	0.06	0.20	0.25***	0.70**	0.14	0.40**
	8	0.08	0.05	0.34**	0.26***	1.42***	0.16	0.82**
	12	0.04	0.00	0.34**	0.21**	1.93***	0.16	1.12**

Table 4: Log Predictive Score. The figures under TV-DSGE model are absolute log predictive scores, the figures under the remaining models are differences of RMSFEs over the TV-DSGE model. '*', '**' and '***' indicate rejection of the null of equal performance against the two-sided alternative at 10%, 5% and 1% significance level respectively, using a Diebold - Mariano test.

Forecast Performance - Log Predictive Score										
		Oil Crisis			Great Moderation			Recent Crisis		
		1974Q3-1982Q4			1983Q3-2005Q4			2006Q1-2011Q3		
	Horizon	Output	Inflation	Interest Rate	Output	Inflation	Interest Rate	Output	Inflation	Interest Rate
TV-DSGE	1	-1.44	-0.71	-0.77	-1.04	0.27	0.55	-1.07	-0.09	0.84
	2	-1.57	-1.13	-1.31	-1.05	0.06	-0.07	-1.25	-0.20	0.22
	4	-1.65	-1.44	-1.55	-1.09	-0.10	-0.65	-1.21	-0.37	-0.46
	8	-1.58	-1.38	-1.98	-1.10	-0.26	-0.99	-1.15	-0.32	-0.86
	12	-1.60	-1.38	-2.02	-1.10	-0.29	-1.23	-1.11	-0.26	-0.94
TV-DSGE - F-DSGE	1	-0.04	-0.17**	0.21	0.09***	0.13**	0.35***	0.02	-0.12	0.48***
	2	-0.08**	-0.21*	0.42	0.13***	0.09**	0.17	0.02	-0.12	0.28*
	4	-0.09**	-0.16	0.59	0.10***	0.11**	-0.04	0.13*	-0.18	0.10
	8	-0.06*	-0.07	0.31	0.10**	0.14**	-0.12	0.13	0.02	0.45
	12	-0.05*	-0.02	0.85	0.10	0.19**	-0.32	0.11	0.21	1.02
TV-DSGE - BVAR	1	0.42	-0.01	0.34	-0.05	0.27***	0.33***	0.12	-0.10	0.66**
	2	-0.01	0.02	0.22	0.00	0.28***	0.14	0.49	-0.08	0.48
	4	0.06	-0.13	0.04	-0.02	0.39***	-0.07	0.66	-0.02	0.21
	8	0.00	-0.10	0.04	-0.03	0.50***	-0.07	0.38	0.34*	0.26
	12	0.08	-0.05	0.40	-0.03	0.60**	-0.18	0.23	0.60	0.61
TV-DSGE-TV-SV BVAR	1	0.19	-0.07	-0.24	-0.05	0.02	-0.03	0.18	-0.02	0.10
	2	-0.01	-0.20	-0.25	-0.07	0.01	-0.07	0.42	-0.11	0.23
	4	-0.05	-0.38	-0.16	-0.10	0.05	-0.16	0.38	-0.17	0.37
	8	0.10	-0.05	-0.08	-0.09	0.09	-0.08	0.70	-0.18	0.67
	12	0.32	0.17	0.06	-0.05	0.09	-0.14	0.33	-0.07	0.76

Table 5: Log Predictive Score. The figures under TV-DSGE model are absolute log predictive scores, the figures under the remaining models are differences of RMSFEs over the TV-DSGE model. '*', '**' and '***' indicate rejection of the null of equal performance against the two-sided alternative at 10%, 5% and 1% significance level respectively, using a Diebold - Mariano test.

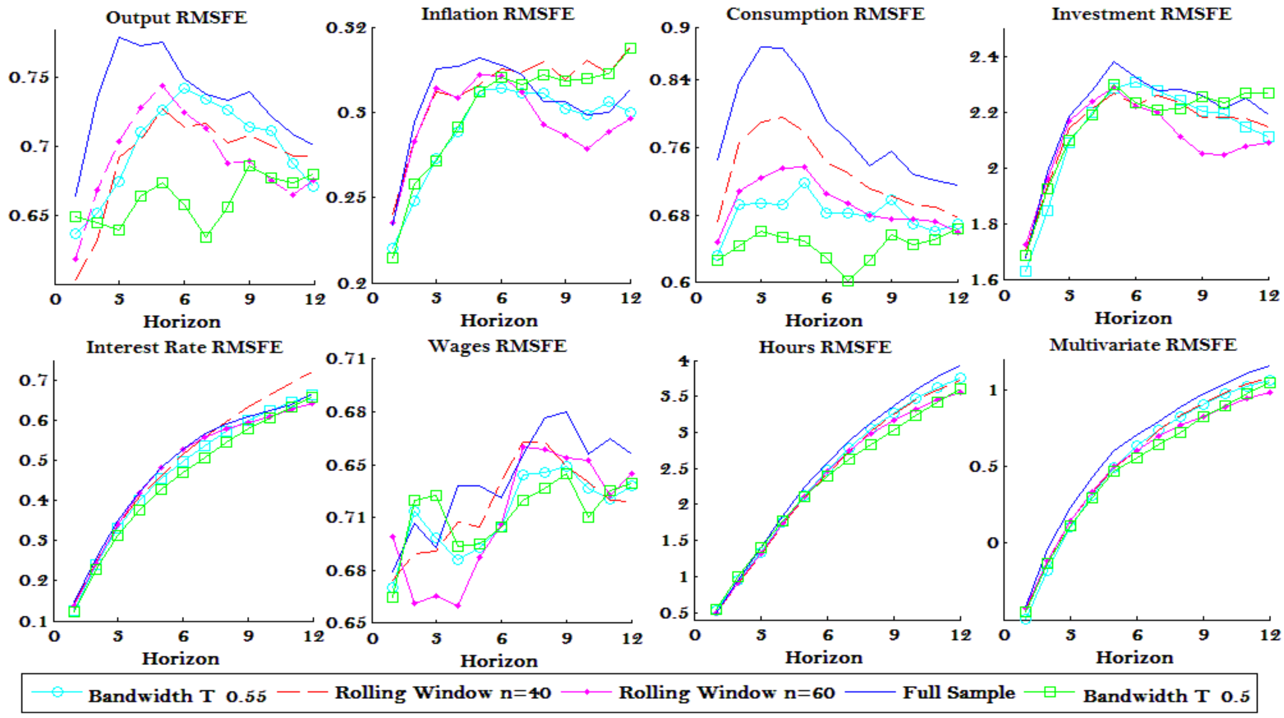


Figure 14: Robustness check: Comparison of RMSFEs obtained with bandwidths $H = T^{0.5}$, $T^{0.55}$ and rolling windows of size 40, 60.

References

Adolfson, M., Andersson, M., Linde, J., Villani, M. and Vredin, A. (2007). Modern forecasting models in action: Improving macroeconomic analyses at Central Banks, *International Journal of Central Banking* **3**(4): 111–144.

An, S. and Schorfheide, F. (2007). Bayesian analysis of DSGE models, *Econometric Reviews* **26**(2-4): 113–172.

Benati, L. and Mumtaz, H. (2007). U.S. evolving macroeconomic dynamics: a structural investigation, *ECB: Working Paper Series* (746).

Benati, L. and Surico, P. (2009). VAR analysis and the Great Moderation, *American Economic Review* **99**(4): 1636–1652.

Blanchard, O. J. and Kahn, C. M. (1980). The solution of linear difference models under rational expectations, *Econometica* **48**(5): 1305–1312.

Boivin, J. and Giannoni, M. (2006). Has monetary policy become more effective?, *Review of Economics and Statistics* **88**(3): 445–462.

Born, B. and Pfeifer, J. (2011). Policy risk and the business cycle, *Bonn Econ Discussion Papers, University of Bonn, Germany* .

Canova, F. (2006). Monetary policy and the evolution of the US economy, *CEPR Discussion Papers 5467* .

Canova, F. and Ferroni, F. (2012). The dynamics of US inflation: Can monetary policy explain the changes, *Journal of Econometrics* **167**(1): 47–60.

Canova, F. and Gambetti, L. (2009). Structural changes in the US economy: Is there a role for monetary policy, *Journal of Economic Dynamics and Control* **33**(2): 477–490.

Canova, F. and Sala, L. (2009). Back to square one: Identifcation issues in DSGE models, *Journal of Monetary Economics* **56**(4): 431–449.

Cantore, C., Levine, P. and Melina, G. (2012). Fiscal-monetary policy: A DSGE model fit for purpose, *Working Paper, University of Surrey* .

Carriero, A., Clark, T. and Marcellino, M. (2015). Bayesian VARs: specification, choices and forecast accuracy, *Journal of Applied Econometrics* **30**: 46–73.

Castelnuovo, E. (2012). Estimating the evolution of money's role in the U.S. monetary business cycle, *Journal of Money, Credit and Banking* **44**(1): 23–52.

Christiano, L., Eichenbaum, M. and Evans, C. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy, *Journal of Political Economy* **113**(1): 1–45.

Christoffel, K., Coenen, G. and Warne, A. (2010). Forecasting with DSGE models, *in* M. Clements and D. Hendry (eds), *Oxford Handbook on Economic Forecasting*, Oxford University Press: Oxford, pp. 89–128.

Clarida, R., Gali, J. and Gertler, M. (2000). Monetary policy rules and macroeconomic stability: Evidence and some theory, *Quarterly Journal of Economics* **115**(1): 147–180.

Cogley, T., Primiceri, G. E. and Sargent, T. J. (2010). Inflation-gap persistence in the US, *American Economic Journal: Macroeconomics* **2(1)**: 43–69.

Cogley, T. and Sargent, T. J. (2002). Evolving post-World War II U.S. inflation dynamics, *in* B. S. Bernanke and K. Rogoff (eds), *NBER Macroeconomics Annual*, MIT Press: Cambridge, pp. 331–88.

Cogley, T. and Sbordone, A. M. (2008). Trend inflation, indexation, and inflation persistence in the new Keynesian Phillips curve, *American Economic Review* **95**(5): 2101–2126.

Crowder, M. (1988). Asymptotic expansions of posterior expectations, distributions, and densities for stochastic processes, *Annuals of the Institute for Statistical Mathematics* **40**: 297–309.

Del Negro, M., Giannoni, M. and Schorfheide, F. (2014). Inflation in the Great Recession and new Keynesian models, *NBER Working Paper 20055* .

Del Negro, M. and Schorfheide, F. (2013a). Bayesian econometrics, *in* J. Geweke, G. Koop and H. Van Dijk (eds), *The Oxford Handbook of Bayesian Econometrics*, Vol. 2, Oxford University Press: Oxford, pp. 293–390.

Del Negro, M. and Schorfheide, F. (2013b). DSGE model-based forecasting, *in* G. Elliott and A. Timmermann (eds), *Handbook of Economic Forecasting*, Vol. 2, Elsevier, pp. 57–140.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy, *Journal of Business and Economic Statistics* **13**: 253–263. Reprinted in Mills, T. C. (ed.) (1999), *Economic Forecasting. The International Library of Critical Writings in Economics.* Cheltenham: Edward Elgar.

Edge, R. and Guerkaynak, R. (2010). How useful are estimated DSGE model forecasts for central bankers, *Brookings Papers of Economic Activity* .

Edge, R., Guerkaynak, R. and Kisacikoglu, B. (2013). Judging the DSGE model by its forecasts, *Working Paper* .

Edge, R., Kiley, M. and Laforte, J. (2009). A comparison of forecast performance between Federal Reserve staff forecasts, simple reduced-form models, and a DSGE model, *Federal Reserve Board of Governors Finance and Economics, Discussion Paper Series* .

Fernandez-Villaverde, J. and Rubio-Ramirez, J. F. (2007). Estimating macroeconomic models: a likelihood approach, *Review of Economic Studies* **74**: 1059–1087.

Fernandez-Villaverde, J. and Rubio-Ramirez, J. F. (2008). How structural are the structural parameters, *in* K. R. D. Acemoglu and M. Woodford (eds), *NBER Macroeconomics Annual 2007*, Vol. 22, Chicago: University of Chicago Press.

Gali, J. (1999). Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations?, *American Economic Review* **89**(1): 249–271.

Gali, J. and Gambetti, L. (2009). On the sources of the Great Moderation, *American Economic Journal* **1**(1): 26–57.

Gali, J. and Gertler, M. (1999). Inflation dynamics: A structural econometric analysis, *Journal of Monetary Economics* **44**(2): 195–222.

Galvão, A. B., Giraitis, L., Kapetanios, G. and Petrova, K. S. (2015). Local Bayesian likelihood estimation, *Working Paper* .

Geweke, J. (1999). Computational experiments and reality, *University of Iowa,Manuscript* .

Geweke, J. (2005). *Contemporary Bayesian Econmetrics and Statistics*, Hoboken: John Wiley and Sons.

Giacomini, R. and Rossi, B. (2009). Model comparisons in unstable environments, *Cemmap Working Paper CWP13/12* .

Giraitis, L., Kapetanios, G. and Price, S. (2013). Adaptive forecasting in the presence of recent and ongoing structural change, *Journal of Econometrics* **177**: 153–170.

Giraitis, L., Kapetanios, G., Theodoridis, K. and Yates, T. (2013). Estimating time-varying DSGE models using minimum distance methods, *Working Paper* .

Giraitis, L., Kapetanios, G., Wetherilt, A. and Zikes, F. (2015). Estimating the dynamics and persistence of financial networks, with an application to the Sterling money market, *Journal of Applied Econometrics (Forthcoming)* .

Giraitis, L., Kapetanios, G. and Yates, T. (2014). Inference on stochastic time-varying coeffcient models, *Journal of Econometrics* **179**(1): 46–65.

Goodfriend, M. and King, R. (1997). The new neoclassical synthesis and the role of monetary policy, *in* B. Bernanke and J. Rotemberg (eds), *NBER Macroeconomics Annual*, MIT Press: Cambridge, pp. 231–83.

Hastings, W. (1970). Monte Carlo sampling methods using Markov chain and their applications, *Biometrika* **57**: 97–109.

Ireland, P. (2004). Technology shocks in the new Keynesian model, *The Review of Economics and Statistics* **86**(4): 923–936.

Justiniano, A. and Primiceri, G. E. (2008). The time-varying volatility of macroeconomic fluctuations, *American Economic Review* **98**(3): 604–641.

Kim, J. Y. (1998). Large sample properties of posterior densities, Bayesian information criterion, and the likelihood principle in non-stationary time series models, *Econometrica* **62**(2): 359–380.

Kolasa, M. and Rubaszek, M. (2015). Forecasting with DSGE models with financial frictions, *International Journal of Forecasting* **31**: 1–19.

Kydland, F. and Prescott, E. (1996). The computational experiment: an econometric tool, *Journal of Economic Perspectives* **10**: 69–85.

Lindley, D. (1971). *Bayesian Statistics: A Review*, Philadelphia, SIAM.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A. and Teller, E. (1953). Equation of state calculations by fast computing machines, *Journal of Chemical Physics* **21**(6): 1087–1092.

Mumtaz, H. and Surico, P. (2009). Time-varying yield curve dynamics and monetary policy, *Journal of Applied Econometrics* **24**(6): 895–913.

Primiceri, G. (2005). Time-varying structural vector autoregressions and monetary policy, *Review of Economic Studies* **72**(3): 821–852.

Roberts, G., Gelman, A. and Gilks, W. (1997). Weak convergence and optimal scaling of random walk metropolis algorithms, *The Annals of Applied Probability* **7**(1): 110–120.

Rotemberg, J. and Woodford, M. (1997). An optimization-based econometric framework for the evaluation of monetary policy, *in* B. S. Bernanke and K. Rogoff (eds), *NBER Macroeconomics Annual*, MIT Press: Cambridge, pp. 297–346.

Rubaszek, M. and Skrzypczynski, P. (2008). On the forecasting performance of a small-scale DSGE model, *International Journal of Forecasting* **24**(3): 498–512.

Schorfheide, F. (2000). Loss function-based evaluation of DSGE model, *Journal of Applied Econometrics* **15**(6): 645–670.

Schorfheide, F. (2005). Learning and monetary policy shifts, *Review of Economic Dynamics* **8**(2): 392–419.

Schorfheide, F. (2007). Comment on 'how structural are structural parameters?', *Working Paper* .

Sims, C. (2002). Solving linear rational expectations models, *Computational Economics* **20**(1-2): 1–20.

Sims, C. A. and Zha, T. (2006). Were there regime switches in U.S. monetary policy?, *American Economic Review* **96**: 1193–1224.

Smets, F. and Wouters, R. (2003). An estimated dynamic stochastic general equilibrium model of the Euro Area, *Journal of the European Economic Association* **1**(5): 1123–1175.

Smets, F. and Wouters, R. (2005). Comparing shocks and frictions in U.S. and Euro Area business cycles: a Bayesian DSGE approach, *Journal of Applied Econometrics* **20**(2): 161–83.

Smets, F. and Wouters, R. (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach, *American Economic Review* **97**(3): 586–606.

Steinbach, R., Mathuloe, P. and Smith, B. (2009). An open economy new Keynesian DSGE model of the South African economy, *South African Journal of Economics* **77**: 207–227.

Walker, A. M. (1969). On the asymptotic behavior of posterior distributions, *Journal of Royal Statistical Society* **31**(1): 80–88.

Wieland, V. and Wolters, M. (2011). The diversity of forecasts from macroeconomic models of the US economy, *Economic Theory* **47**: 247–292.

Woodford, M. (2003). *Interest and Prices: Foundation of a Theory of Monetary Policy*, Princeton University Press: Princeton.

6 Appendix

6.1 The Smets and Wouters (2007) Model

The resource constraint is given by:

$$y_t = \underbrace{(1 - g_y - i_y)}_{\substack{\text{steady state} \\ \text{consumption-output ratio}}} c_t + \underbrace{((\gamma - 1 - \delta)k_y)}_{\substack{\text{steady state} \\ \text{investement-output ratio}}} i_t + (R_*^k k_y) z_t + \varepsilon_t^g.$$

Output, y_t, is absorbed by consumption c_t, investment i_t, capital utilization z_t and government spending ε_t^g. g_y, i_y and k_y are steady state government-output, investment-output and capital-output ratios respectively and R_*^k is the steady state rental rate of capital. γ is the steady state growth rate of output, used to detrend all non-stationary variables in the model and δ is the depreciation rate of capital. Exogenous government spending follows an AR(1) stochastic process with an autoregressive coefficient ρ_g and an iid-Normal error term η_t^g with variance σ_g^2 :

$$\varepsilon_t^g = \rho_g \varepsilon_{t-1}^g + \eta_t^g + \rho_{ga} \eta_t^a$$

where η_t^a is the iid-Normal shock to TFP and is motivated by Smets and Wouters (2007) as the model at hand is a closed economy, with ε_t^g also including data on exports/imports, which could depend on domestic productivity η_t^a.

The Euler equation for consumption is:

$$c_t = \frac{(\lambda/\gamma)}{(1+\lambda/\gamma)} c_{t-1} + \frac{1}{(1+\lambda/\gamma)} \mathbb{E}_t c_{t+1} + \frac{(\sigma_c - 1) W_*^h L_*/C_*}{\sigma_c (1+\lambda/\gamma)} \mathbb{E}_t (l_t - l_{t+1}) - \frac{(1-\lambda/\gamma)}{(1+\lambda/\gamma)\sigma_c} (r_t - \mathbb{E}_t \pi_{t+1} + \varepsilon_t^b)$$

and implies that consumption c_t is a weighted average between past consumption c_{t-1} and expected future consumption $E_t c_{t+1}$. It also depends on expected growth in the hours worked $E_t(l_t - l_{t+1})$ and ex-ante real interest rate $r_t - E_t \pi_{t+1}$ and a risk premium shock ε_t^b representing a wedge between the instrument controlled by the central bank and the rate of return on assets faced by households. It follows an AR(1) stochastic process with an autoregressive coefficient ρ_b and an iid-Normal error term η_t^b with variance σ_b^2:

$$\varepsilon_t^b = \rho_b \varepsilon_{t-1}^b + \eta_t^b.$$

In the absence of habit formation, $\lambda = 0$, the first term drops out and the Euler equation becomes entirely forward-looking. When the elasticity of intertemporal substitution, $\sigma_c = 1$, the household is facing log utility in consumption and the labour supply term drops out.

The investment Euler equation is:

$$i_t = \frac{1}{1+\beta\gamma^{1-\sigma_c}} i_{t-1} + (1 - \frac{1}{1+\beta\gamma^{1-\sigma_c}})\mathbb{E}_t i_{t+1} + \frac{1}{(1+\beta\gamma^{1-\sigma_c})\gamma^2\varphi} q_t + \varepsilon_t^i$$

implying that current investment i_t is a weighted average of past investment i_{t-1} and expected future investment $E_t i_{t+1}$. It also depends on the real value of capital q_t and an investment-specific technology disturbance term ε_t^i that captures the relative efficiency of investment expenditure and follows an AR(1) stochastic process with an autoregressive coefficient ρ_i and an iid-Normal error term η_t^i with variance σ_i^2 :

$$\varepsilon_t^i = \rho_i \varepsilon_{t-1}^i + \eta_t^i.$$

φ is the steady state elasticity of capital adjustment cost and β is the household's discount factor.

The arbitrage condition between the return to capital and the riskless rate is given by:

$$q_t = \frac{(1-\delta)}{R_*^k + (1-\delta)}\mathbb{E}_t q_{t+1} + (1 - \frac{(1-\delta)}{R_*^k + (1-\delta)})\mathbb{E}_t r_{t+1}^k - (r_t - \mathbb{E}_t \pi_{t+1} + \varepsilon_t^b)$$

where the current capital value q_t is a weighted average of expected future value $E_t q_{t+1}$ and expected real rental rate of capital $E_t r_{t+1}^k$ and depends also negatively on the ex-ante real interest rate $r_t - E_t \pi_{t+1}$ and the risk-premium disturbance ε_t^b. The aggregate production function is:

$$y_t = \phi(\alpha k_t^s + (1-\alpha) l_t + \varepsilon_t^a)$$

with output being produced with standard factors of production, capital k_t^s, labour l_t and technology ε_t^a, assumed to follow an AR(1) stochastic process with an autoregressive coefficient ρ_a and an iid-Normal error term η_t^a with variance σ_a^2:

$$\varepsilon_t^a = \rho_a \varepsilon_{t-1}^a + \eta_t^a.$$

The parameter ϕ measures one plus the fixed costs in production and α is the capital share after mark-ups and fixed costs. Since the capital installed today takes a period to become effective, current capital used in production k_t^s is a sum of capital installed the previous period k_{t-1} and the degree of capital utilization z_t:

$$k_t^s = k_{t-1} + z_t.$$

The degree of capital utilization z_t itself depends positively in the rental rate of capital r_t^k and ψ is a function of the elasticity of capital utilization adjustment cost function, normalised to take valued between zero and one:

$$z_t = \frac{1-\psi}{\psi} r_t^k.$$

The capital accumulation equation is given by:

$$k_t = \frac{1-\delta}{\gamma} k_{t-1} + (1 - \frac{1-\delta}{\gamma}) i_t + (1 - \frac{1-\delta}{\gamma})((1 + \beta\gamma^{1-\sigma_c})\gamma^2 \varphi)\varepsilon_t^i$$

where installed capital k_t is a function of previously installed capital k_{t-1}, investment flow i_t and the investment-specific disturbance ε_t^i.

The price mark-up μ_t^p in the monopolistic goods market is given by the difference between marginal product of labour mpl_t, which depends on TFP and the capital-labour ratio, and the real wage w_t:

$$\mu_t^p = \underbrace{\alpha(k_t^s - l_t) + \varepsilon_t^a}_{mpl_t} - w_t.$$

The Phillips curve is given by:

$$\pi_t = \frac{\iota_p}{1 + \beta\gamma^{(1-\sigma_c)}\iota_p} \pi_{t-1} + \frac{\beta\gamma^{(1-\sigma_c)}}{1 + \beta\gamma^{(1-\sigma_c)}\iota_p} \mathbb{E}_t \pi_{t+1} - \frac{1}{1 + \beta\gamma^{(1-\sigma_c)}\iota_p} \left\{ \frac{(1 - \beta\gamma^{(1-\sigma_c)}\xi_p)(1 - \xi_p)}{\xi_p((\phi - 1)\varepsilon_p + 1)} \right\} \mu_t^p + \varepsilon_t^p$$

and implies that current level of inflation π_t is a function of past inflation π_{t-1} and expected future inflation $E_t \pi_{t+1}$, price mark-up μ_t^p and a price mark-up shock ε_t^p. Without degree of indexation, $\iota_p = 0$, the expression reduces to purely forward-looking Phillips curve. ξ_p is the Calvo price stickiness, ε_p is the Kimball aggregator in the goods market that measures the degree of strategic interaction between price-setters and $\phi - 1$ is the steady state price mark-up, which depends on the fixed cost parameter ϕ. The price mark-up error term ε_t^p follows an ARMA(1,1) process with an autoregressive coefficient ρ_p, a moving average coefficient μ_p and an iid-Normal error term η_t^p with variance σ_p^2, motivated by the desire to capture more of the dynamics in the data on inflation fluctuations:

$$\varepsilon_t^p = \rho_p \, \varepsilon_{t-1}^p + \eta_t^p + \mu_p \eta_{t-1}^p.$$

Rental rate of capital is a function of the capital-labour ratio $(k_t - l_t)$ and the real wage w_t:

$$r_t^k = -(k_t - l_t) + w_t.$$

The labour market is characterised by similar conditions to the goods market. In particular, there is a wage mark-up equation:

$$\mu_t^w = w_t - \underbrace{(\sigma_l l_t + \frac{1}{1-\lambda/\gamma}(c_t - \lambda/\gamma c_{t-1}))}_{mrs_t}$$

where the wage mark-up μ_t^w is the difference between the real wage w_t and the marginal rate of substitution between working and consuming, mrs_t, that is the disutility of work, with σ_l capturing the elasticity of labour supply with respect to the wage. The corresponding wage equation is given by:

$$w_t = \frac{1}{1+\beta\gamma^{(1-\sigma_c)}} w_{t-1} + (1 - \frac{1}{1+\beta\gamma^{(1-\sigma_c)}})(\mathbb{E}_t w_{t+1} + \mathbb{E}_t \pi_{t+1}) - \frac{1+\beta\gamma^{(1-\sigma_c)}\iota_w}{1+\beta\gamma^{(1-\sigma_c)}} \pi_t$$
$$+ \frac{\iota_w}{1+\beta\gamma^{(1-\sigma_c)}} \pi_{t-1} - \frac{1}{1+\beta\gamma^{(1-\sigma_c)}} \left\{ \frac{(1-\beta\gamma^{(1-\sigma_c)}\xi_w)(1-\xi_w)}{\xi_w((\phi_w - 1)\varepsilon_w + 1)} \right\} \mu_t^w + \varepsilon_t^w.$$

The real wage w_t is a weighted average between past wage w_{t-1} and expected future real wage $(E_t w_{t+1} + E_t \pi_{t+1})$, depends on wage mark-up, current inflation π_t, wage mark-up shock ε_t^w and partially indexed to past inflation π_{t-1}. Similarly to the goods market, ι_w captures the degree of indexation, ξ_w is the Calvo wage stickiness, ε_w is the Kimball aggregator in the labour market and $\phi_w - 1$ is the steady state wage mark-up. Finally, the wage disturbance also follows an ARMA(1,1) process with an autoregressive coefficient ρ_w, a moving average coefficient μ_w and an iid-Normal error term η_t^w with variance σ_w^2, with the MA term added as explained by Smets and Wouters (2007) to capture more of the high frequency wage fluctuations observed in the data:

$$\varepsilon_t^w = \rho_w \ \varepsilon_{t-1}^w + \eta_t^w + \mu_w \eta_{t-1}^w.$$

The central bank in the model follows a nominal interest rate rule of the form:

$$r_t = \rho r_{t-1} + (1-\rho)\{r_\pi \pi_t + r_y (y_t - y_t^p)\} + r_{\Delta y}\left((y_t - y_t^p) - (y_{t-1} - y_{t-1}^p)\right) + \varepsilon_t^r$$

by gradually adjusting the policy rate r_t in response to fluctuations in inflation π_t, output gap $(y_t - y_t^p)$ and output gap growth $(y_t - y_t^p) - (y_{t-1} - y_{t-1}^p)$. The policy parameters ρ, r_π, r_y and $r_{\Delta y}$ capture the degree of interest rate smoothing, the level of inflation and output gap targetting and the short-run feedback from output gap change respectively. The monetary policy shock ε_t^r follow an AR(1) stochastic process with an autoregressive coefficient ρ_r and an iid-Normal error term η_t^r with variance σ_r^2:

$$\varepsilon_t^r = \rho_r \ \varepsilon_{t-1}^r + \eta_t^r.$$

The measurement equation takes the form:

$$X_t = \begin{bmatrix} 100 \times \Delta \log GDP_t \\ 100 \times \Delta \log C_t \\ 100 \times \Delta \log I_t \\ 100 \times \Delta \log W_t \\ 100 \times \log H_t \\ 100 \times \Delta \log P_t \\ FFR_t \end{bmatrix} = \begin{bmatrix} \overline{\gamma} \\ \overline{\gamma} \\ \overline{\gamma} \\ \overline{\gamma} \\ \overline{l} \\ \overline{\pi} \\ \overline{r} \end{bmatrix} + \begin{bmatrix} y_t - y_{t-1} \\ c_t - c_{t-1} \\ i_t - i_{t-1} \\ w_t - w_{t-1} \\ l_t \\ \pi_t \\ r_t \end{bmatrix}$$

6.2 Time Varying IRFs

Figure 12: IRFs to 1 unit wage mark up shock

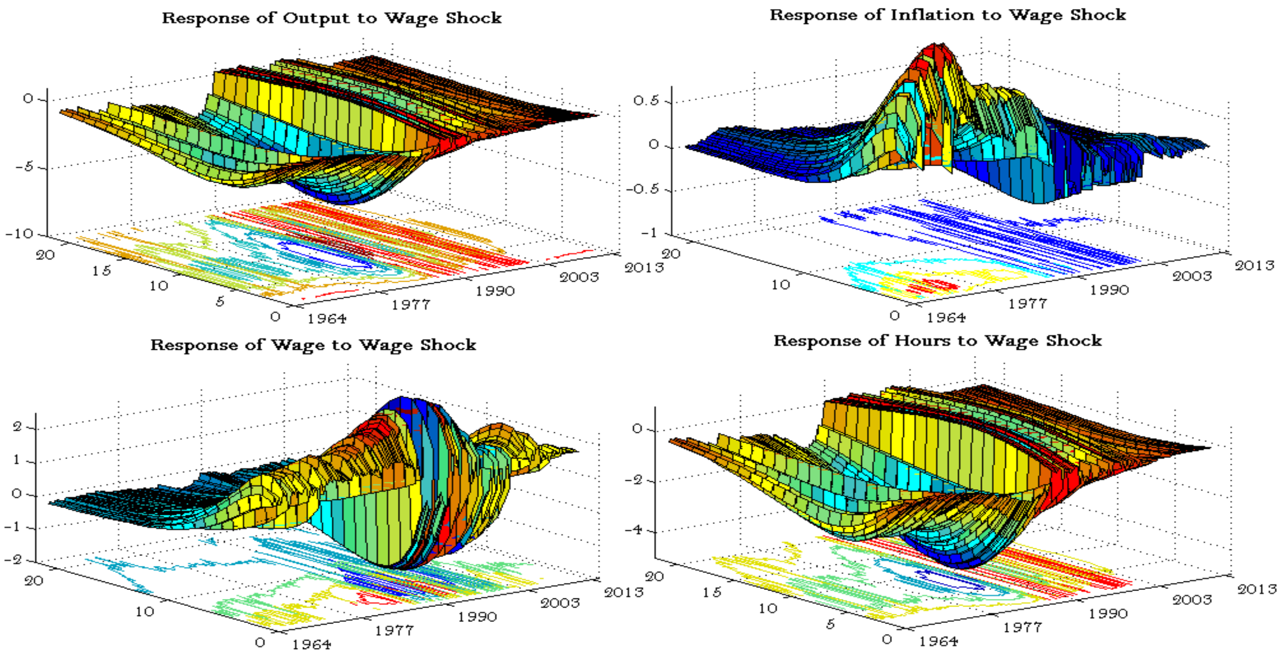


Figure 13: IRFs to 1 st. dev. wage mark up shock

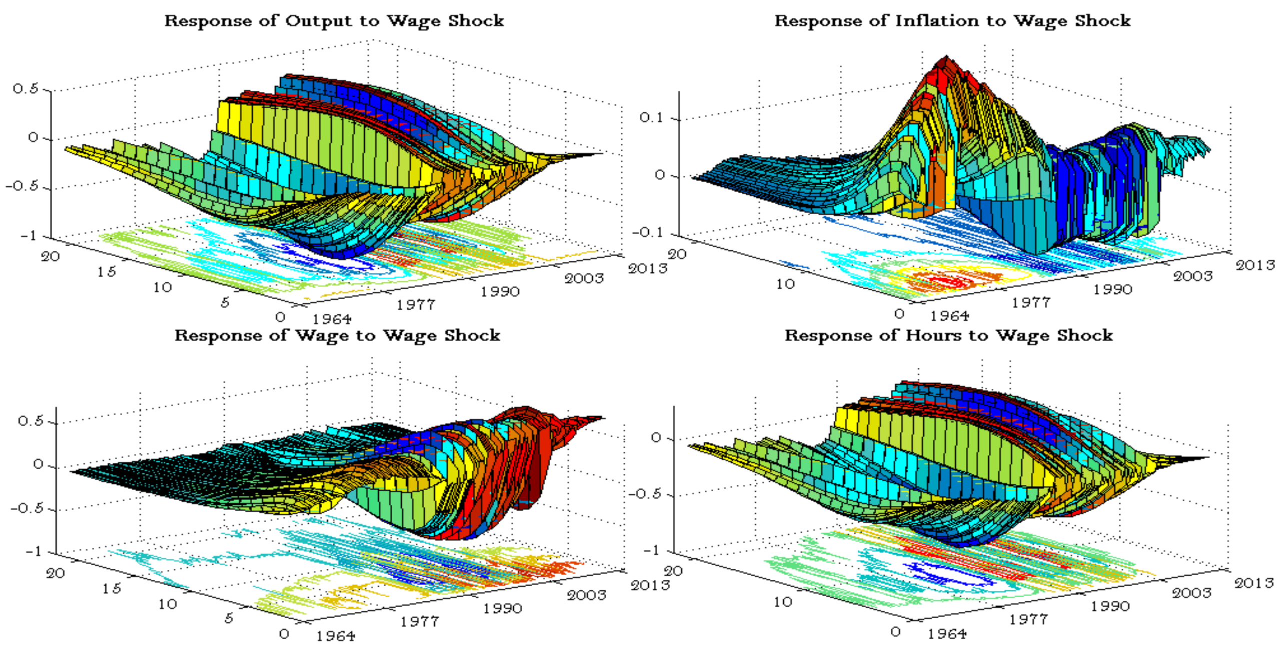


Figure 14: IRFs to 1 unit investment technology shock

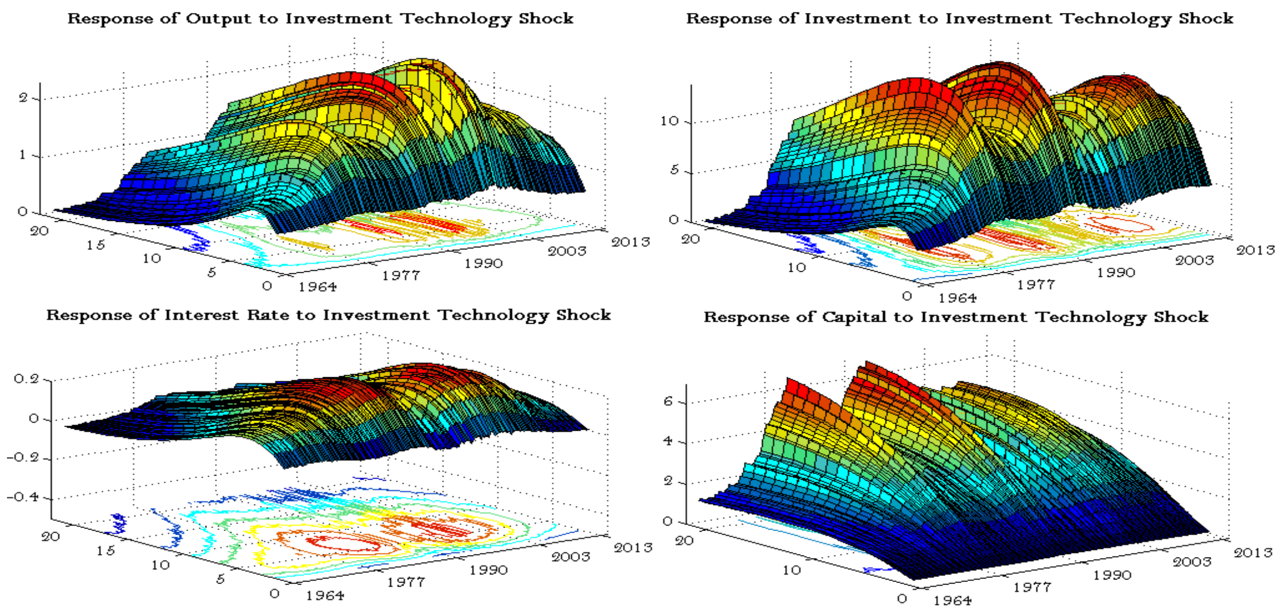


Figure 15: IRFs to 1 st. dev. investment technology shock

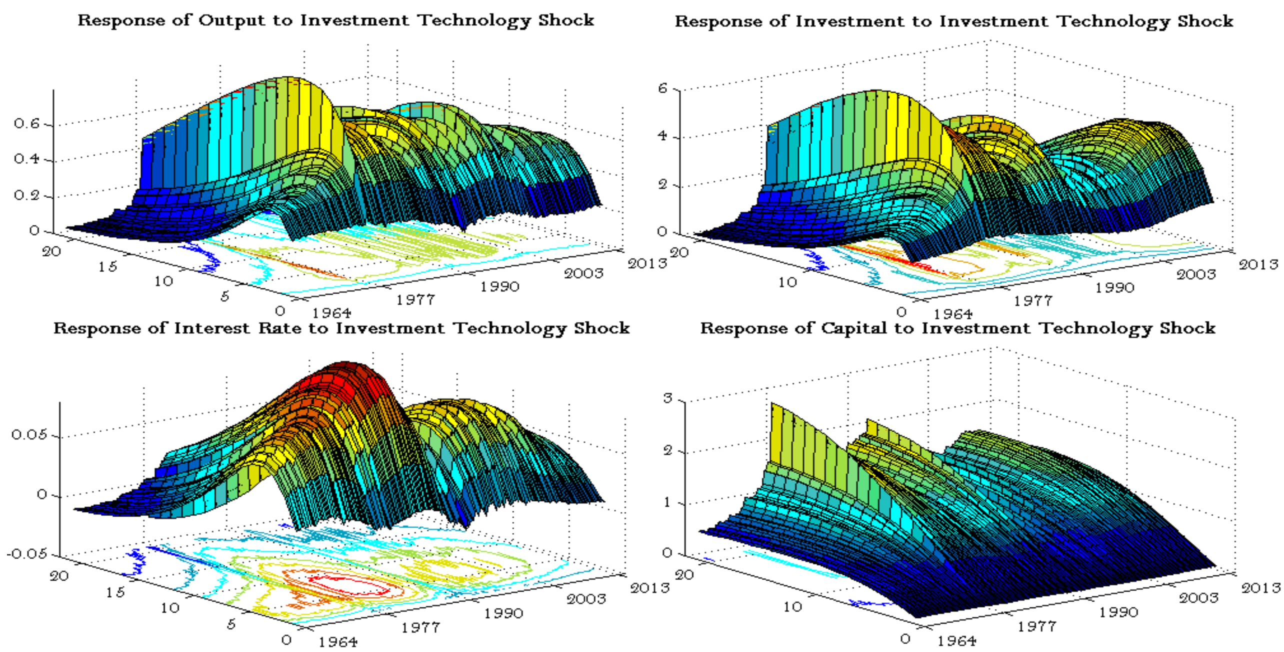


Figure 16: IRFs to 1 unit government spending shock

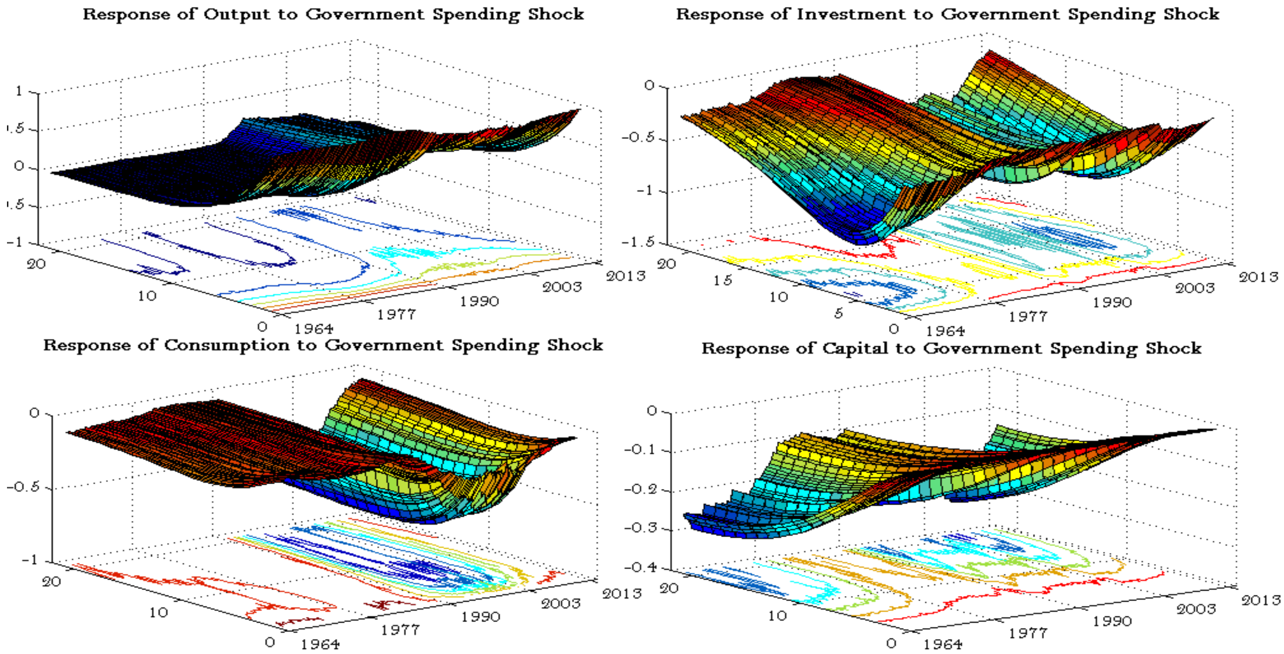
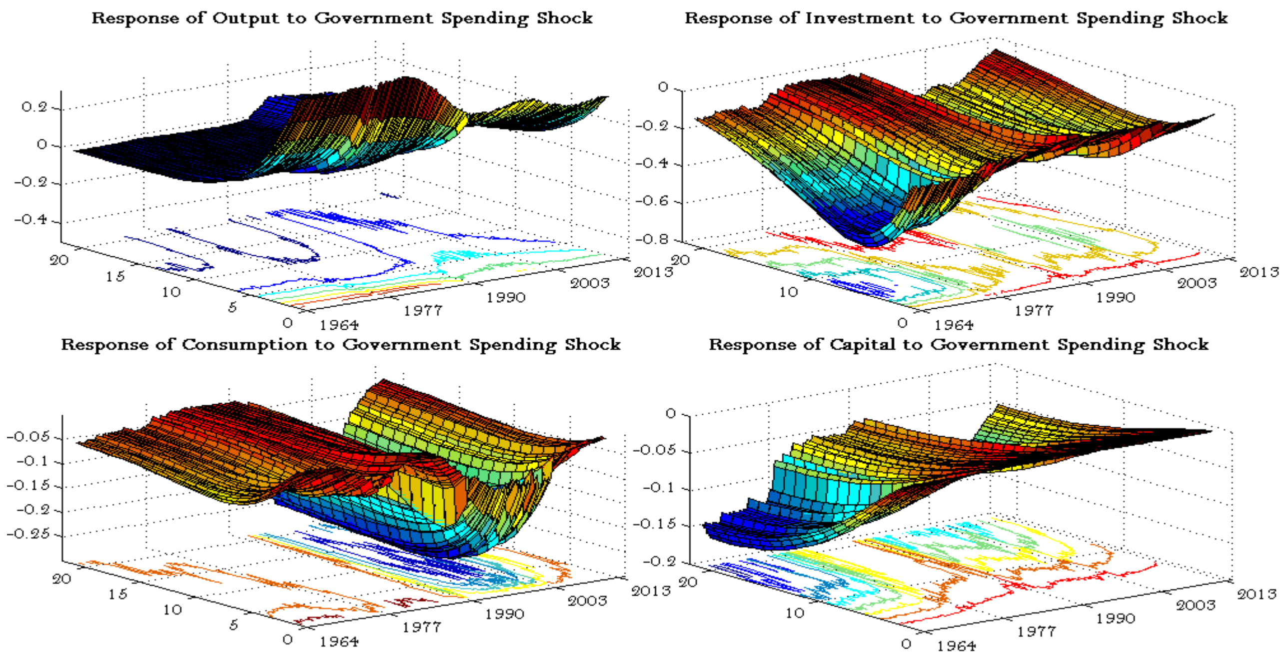


Figure 17: IRFs to 1 st. dev. government spending shock



Queen Mary
University of London

This working paper has been produced by
the School of Economics and Finance at
Queen Mary University of London

School of Economics and Finance
Queen Mary University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 7356
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/research/workingpapers/