University of Economics in Prague

Faculty of Finance and Accounting

Department of Monetary Theory and Policy

An Analysis of Forecasting Abilities of DSGE and Threshold VAR Models

Master Thesis

January 11, 2018

Author:

Bc. Michal Kuchta

Supervisor:

Ing. et Ing. Quang van Tran

Ph.D. et Ph.D.

I, Bc. Michal Kuchta, hereby declare that this thesis titled, "An Analysis of Forecasting Abilities of DSGE and Threshold VAR Models" and the work presented in it are my own and that thesis was written with utilization of listed literature.

Signed:

Date:

Acknowledgements

At this place, I would like to thank my supervisor for this thesis, Ing. et Ing. Quang van Tran Ph.D. et Ph.D., for his guidance, dedicated time and useful comments that improved its content.

Abstract

In this Thesis, small to medium scale DSGE model for Czech republic incorporating set of frictions and rigidities is derived, and estimated by Bayesian techniques. Subsequent Impulse response and Shock decomposition analysis evaluate properties of this model. Model is then matched to empirical data and estimated on Czech major economic time series. Forecasting performance of this models is opposed by Bayesian Threshold VAR and plain Bayesian VAR Models. Forecasting exercise considers a variety of settings and its evaluation is judged upon RMSE providing simple ranking of inquired models.

Key Words: DSGE, Bayesian estimation, forecasting, Bayesian VAR

JEL classification: E37, C11, C51, C53

Abstrakt

V tejto diplomovej práci je odvodený a Bazesovskými metódami odvadnutý malý, až stredne veľký DSGE model pre Českú republiku zahrňujúci tržné frikcie a rigidity. Nasledujúca analýza odozvy a dokompozícia šokov popisuju a hodnotí vlastnosti tohto modelu. Ten je potom upravený tak, aby zodpovedal empiricky pozorvatelným datám a následne odhadnutý na hlavných časových radoch popisujúcich Českú ekonomiku. Predikčné schopnosti modelu sú porovnané s Bayesovským režim-prepinajúcim VAR a obyčajným Bayesovským VAR modelom. Výpočet predpovedí pozostáva z celého setu scenárov a výsledky sú ohodnetené na základe RMSE, ktorá poskytuje jednoduché zoradenie modelov.

Key Words: DSGE, Bayesovské odhady, predpovede, Bayesovký VAR

JEL klasifikácia: E37, C11, C51, C53

Contents

Declaration of Authorship **iii**

1 Introduction **1**

2 Brief History of Macroeconomic Modeling **3**

3 DSGE Model **7**

3.1 Households 7

3.2 Labour Market 11

3.3 Firms 14

3.4 Government 21

3.5 Aggregation 21

3.6 Equilibrium 24

3.7 Model Solution 25

3.8 Estimation 30

3.8.1 Data and Observation Equations 32

3.8.2 Priors 40

3.8.3 Estimation results and Convergence 41

3.9 Model Properties 45

3.10 Forecasting 51

3.11 Model Implementation 53

4 Bayesian Threshold VAR **55**

4.1 Bayesian approach to linear models 55

4.2 Bayesian VAR 56

4.3 Threshold Model 59

4.4 Minnesota Prior 61

4.5 Forecasting 64

4.6 Implementation 67

5 Results Summary and Comparison **69**

6 Conclusion **83**

A DSGE Implementation **91**

A.1 *Dynare* Implementation 91

A.2 Steady state file 95

B Appendix **99**

C BVAR Implementations in Matlab **105**

C.1 Setting of Minnesota Prior 105

C.2 Sampling of Threshold 106

C.3 Forecasting 108

List of Figures

3.1 Final data for the estimation 33

3.2 Multivariate Convergence Diagnostics 42

3.3 Prior and Posterior Distributions (1) 45

3.4 Prior and Posterior Distributions (2) 46

3.5 Smoothed past realization of shocks 46

3.6 Response to Monetary shock 47

3.7 Response to TFP shock 48

3.8 Response to investment shock 48

3.9 Response to consumption preference shock 49

3.10 Response to labour supply shock 50

3.11 Shock decomposition - Output 51

5.1 Unfiltered and Filtered forecasts of output by SETAR(2,2,2) 74

5.2 1-step out-of-sample forecast of Ouutput by SETAR(2,2,2) 75

5.3 12-step out-of-sample forecasts of output by SETAR(2,2,2) 76

5.4 12-step out-of-sample forecasts of output by DSGE 77

5.5 1-step out-of-sample forecasts of output by DSGE 78

5.6 1-step out-of-sample forecasts of inflation by BVAR(1) 79

5.7 1-step out-of-sample forecasts of inflation by DSGE 79

5.8 12-step out-of-sample forecasts of inflation by DSGE 80

5.9 12-step out-of-sample forecasts of inflation by SETAR(2,2,2) 81

B.1 Shock decomposition - Consumption 99

B.2 Shock decomposition - Investments 100

B.3 Shock decomposition - Hours worked 100

B.4 Shock decomposition - Inflation 101

B.5 Shock decomposition - Gross nominal IR 101

B.6 In-sample ft of output by BVAR(2) 102

B.7 Mode Check Plots (1) 102

B.8 Mode Check Plots (2) 103

List of Tables

3.1 Descriptive Statistics 34

3.2 Estimated Parameters 43

3.3 Estimated standard deviation of Structural Shocks 44

3.4 Estimated correlation of Structural Shocks 44

5.1 Marginal likelihood based model comparison 72

5.2 Results Comparison (RMSE) 72

1 Introduction

This thesis aims to develop and practically implement progressive and powerful forecasting models able to from precise prognosis about future realization of major macroeconomic time series, which are usual interests of financial institutions, central banks or government. Since economic development is a dynamic process subject to temporary fluctuations or permanent structural breaks, there is considerable uncertainty related to almost every aspects of economic activity. High quality forecasts are able to reduce this uncertainty and turn it into quantifiable risk.

In relation to that, they are necessary condition for successful implementation of various risk models requiring short and long term forecasts as an basic input and for simulations of artificial scenarios and evaluation of their impact. Widely used risk management models such as Value at Risk, portfolio (or even individual) credit risk modeling, provisioning under IFRS9 or BASEL III, all contain macroeconomic part to produce either thought-the-cycle or point-in-time estimates. In particular, strong forecasting models are appreciated in stress testing, when assessing overall health of financial institutions.

For this purpose, two main types of models are constructed: a small to medium scale DSGE model describing Czech economy, which performance is compared to Threshold VAR model, both estimated by Bayesian techniques. For the former, intention is to develop structural macroeconomic model suitable not only for forecasting under different scenarios, but also for conducting inference in policy evaluation and counter-factual analysis. For the latter, goal is to implement modern Bayesian techniques to improve performance and build stable roots for inference of already powerful VAR models.

Derivation of the DSGE model follows mainly influential work of Smets and Wouters (2002), Fernandez-Villaverde and Rubio-Ramirez (2005) and Adolfson et al. (2005). Partially following New Keynesian tradition, estimated DSGE model incorporates short run rigidities of prices and wages, but still perceives perfectly mobile and cost-free adjustments of capital stock typical for RBC models as in King and Rebelo (1999) or in basic NKE models. In addition, it is a money-in-the-utility model with habit persistence formation, indexation to past inflation and two sectors of firms. Role of government in model is reduced to setting of gross nominal interest rate through Taylor rule. As such, model does not provide only unconditional forecast, but also identifies drivers of structural changes and responses to past supply, demand and monetary shocks hitting the economy and considered in the model. Moreover, with little adjustments, it is well-suitable to produce conditional forecasts for structural analysis, although this matter is not subject of this thesis. Solved model is estimated by Metropolis-Hastings algorithm and implemented in *Dynare*.

Next, constructed VAR models and estimated on the same data challenge performance of this DSGE. VAR as non-structural models are capable of producing relatively precise unconditional forecasts. In the past, they were widely used and into certain degree still are, for economic or monetary policy assessment (Walsch, 2010). For estimation of VAR parameters, Gibbs sampling algorithm is employed extended by Metropolis-Hastings algorithm to sample threshold as in Blake and Mumtaz(2012). Utilization and implementation of Bayesian techniques follows most recent trend in econometrics and mainly macroeconomic modeling. Advantages of these methods are discussed right in the subsequent period.

Forecasting exercise consist of out-of-sample 1 and 12-step ahead forecasts with expanding window and two different in-sample periods. Forecasting performance of models is judged upon and RMSE statistics computed from out-of-sample forecasts and in sample-fit assessment is done upon marginal likelihood. Thesis also includes illustrations of all practical implementations in Matlab and Dynare.

2 Brief History of Macroeconomic Modeling

Tradition of macroeconomic modeling is characterized by two approaches: structural and non-structural. In the former, econometric techniques are tighten to theory and as such depends on currently dominating economic thinking. The former abstract themselves from macroeconomic theory and postulate unrestricted reduced-form models intended to abstract correlations in observed time series. (Diebold, 2007)

The rise of structural modeling is closely related to Keynes' (1936) General Theory and Hicks' (1937) IS-LM model. Subsequent period is characterized by development of models based on system of decision rules used for conditional forecasting and for evaluation of economic policy. Klein's (1946) "revolutionary" approach introduced simultaneous equation modeling. These determine equilibrium of the system as a whole, but often are poorly identified due to lack of instruments in macroeconomic time series and criticized for arbitrary selection of exogenous variables and ad-hoc formulation of decision rules, such as consumption of investment functions. From today's perspective, they are even subject to spurious regression.

In the 70's, a period characterized by empirical failure of such famous Phillips curve, Lucas (1976) formally showed that conditional forecasts produced by systems of simultaneous equations models are false as the change of policy changes also parameters of decision rules employed in the model. This gave a rise to models based on rational expectations and optimizing agents of Kydlan and Prescot (1982), able to produce artificial time series relatively close to observed ones. This models are built on micro foundations incorporating several features of neoclassical economics.

RBC research program is built on intertemporal substitution of labour and states that fluctuations are caused by the fluctuations of the potential output itself due to technological shocks. Further incorporation of nominal rigidities and imperfect adjustments created New Keynesian economics and todays mainstream. In a recent discussion, DSGE models are doubted to meet requirements for structural forecasting imposed by Lucas critique, as larger-scale models are able to produce the same outcomes with completely different sets of parameters.

On the other hand, non-structural models are intended to produce unconditional forecast as a likely future path and therefore and not subject to Lucas critique (Diebold, 2007). Let begin with Box and Jenkins (1970) methodology for univariate time series analysis, i.e. ARMA models. They abandoned to that time generally accepted deterministic trend modeling, and introduced stochastic trends generated by cumulation of past shocks.

Breakthrough in non-structural modeling was Sims (i) formulation of famous forecasting equation $y_{t+1} = \beta_0 + \beta_1 x_t + \varepsilon_t$ and (ii) relaxation of arbitrary selection of exogenous variables in simultaneous equation modeling, resulting in formulation of Vector Autoregressive models. Advantage over ARMA models is in description of multivariate relationships. VAR models can be easily manipulated to produce even conditional forecast (Lutkepohl, 2005) and still are often used for policy analysis (Walsch, 2010).

Cointegration analysis of Engle and Granger (1987) as a response to spurious regression allowed for modeling of long term relationships without differencing time series and thus loosing information.

Threshold models, firstly introduced by Tong (1980) view economic fluctuations as different regimes determined by certain threshold value of exogenous or lagged endogenous variable. Markov-switching models are basically generalization of threshold models with threshold being determined by unobserved indicator.

Current rise of computational power enhanced utilization of Bayesian methods in estimation of both, structural and non-structural models. Especially for the large scale models it is easier to integrate likelihood function then to maximize it (Fernández-Villaverde; 2009). Moreover, modern algorithms such as Gibbs sampling or Metropolis-Hastings algorithm allow for approximation of posterior distributions of parameters of interest combining prior distributions with likelihood function devoid of tedious integration of multidimensional functions.

Key difference between classical and Bayesian approach is in the interpretation of the relationship between parameters and data: while classical approach (MLE) takes data as random realization of likelihood function and parameters intended to maximize it as fixed, yet exhibiting uncertainty connected to random realization of data, Bayesians consider data as given and parameters are treated as being random, with objective to provide conditional probabilistic statements (DeJong and Dave, 2007). It is the probabilistic treatment that enables to incorporate prior beliefs about parameters to be estimated in form of prior distribution.

In non-structural modeling, main advantage of Bayesian econometrics in estimation of linear models over classical approach is that no limiting assumption have to be employed. Standard VAR model requires stationarity of all variables (or stability of a model) and testing for co-integration relationships, while only limitation of BVAR is convergence. Second, standard methods utilizing OLS or maximum likelihood estimation provides point estimates of $\hat{\beta}$ parameters and (probably biased) estimate of error variance $\hat{\sigma}^2$ relying only on the information contained in the data (Blake and Mumtaz, 2012). Bayesian approach produce not only mode, but entire posterior distribution and again allows for incorporation of prior beliefs.

Availability of entire posterior distribution for parameters and endogenous variables of interest, or possible impulse response functions, represents better cornerstone for inference about significance, efficiency, unbiasedness and confidence bands (percentiles of posterior distribution) of estimated parameters themselves or predicted (forecasted) values in both, structural and non-structural modeling.

3 DSGE Model

In this section, DSGE model For Czech economy is derived and solved. Later on, linkage of model to empirical data is demonstrated and subsequent estimation methodology, setting of priors and results are presented. In this thesis, Czech republic is modeled as a closed economy with corresponding corrections to the data and without government sector of major influence. Derived DSGE is a mixture of NKE with labour market a price rigidities, and RBC with perfect capital mobility. Price rigidities are incorporated into model though intermediate goods producers and their monopolistic power in price setting.

3.1 Households

Domestic economy is populated by the continuum of identical forever living households indexed over interval $(i) \in [0,1]$. The unitary preferences are represented by instantaneous utility function separable in consumption, labour supply and real money balances:

$$U\left(C_t(i), N_t(i), \frac{M_t(i)}{P_t}\right) = \left(\frac{\mathrm{e}^{\epsilon_t^c}\left(C_t(i) - H_t\right)^{1-\sigma}}{1-\sigma} - \psi\frac{\mathrm{e}^{\epsilon_t^l} N_t(i)^{1+\eta}}{1+\eta} + \frac{\left(\frac{M_t(i)}{P_t}\right)^{1-\mu}}{1-\mu}\right) \tag{3.1}$$

where $C_t(i)$ is the per capita composite consumption index, $N_t(i)$ differentiated per capita labour supply in terms of hours worked and $\left(\frac{M_t(i)}{P_t}\right)$ are per capita real money balances, i.e. real money demand. Next, σ is the inverse of intertemporal elasticity of substitution in consumption parameter, i.e. a consumption smoothing or relative risk aversion parameter, η is the inverse of Frisch labour supply elasticity or work effort with respect to real wage (Smets and Wouters, 2007), ψ is labour disutility parameters and μ is money demand smoothing parameter.

Consumption pattern is subject to the habit persistence $H_t = hC_{t-1}(i)$ with h representing habit formation parameter and $h \in [0, 1]$. It is a function of lagged consumption and thus is not affected by current household decision Its purpose is to incorporate desired persistence of consumption seen in empirical data and "keeping up wit Joneses" effect (Abel, 1990). Motivation for money holding is of transactional character and provided utility is rather of available free liquidity to finance consumption. ϵ_t^c and ϵ_t^l are, respectively, consumption preference and labour shock processes defined as:

$$\epsilon_t^c = \rho_c \epsilon_{t-1}^c + \varepsilon_t^c \tag{3.2}$$

$$\epsilon_t^l = \rho_l \epsilon_{t-1}^l + \varepsilon_t^l \tag{3.3}$$

Households are capital owners and thus provide capital services to production sector for rental rate r_t and decide about its level in subsequent period through current period investments. They also form savings by holding bond and receiving interest payments corresponding to nominal interest rate i_t. Law of motion for capital is defined as:

$$K_{t+1}(i) = \mathrm{e}^{b_t} I_t(i) + (1-\delta)K_{t-1}(i) \tag{3.4}$$

with

$$b_t = \rho_b b_{t-1} + \varepsilon_t^i \ \text{, where } \varepsilon_t^i \sim N(0, \sigma_i^2) \tag{3.5}$$

Discounted lifetime utility is maximized subject to a series of intertemporal budget constraints defined in real and per capita (household) terms and after substitution for investments defined as:

$$\begin{aligned} C_t(i) + \frac{B_t(i)}{P_t} + \frac{M_t(i)}{P_t} + \mathrm{e}^{-b_t} K_{t+1}(i) \leq ... \\ W_t(i)N_t(i) + r_t K_t(i) + (1-\delta)\mathrm{e}^{-b_t} K_t(i) + \frac{(1+i_{t-1})B_{t-1}(i)}{P_t} + \frac{M_{t-1}(i)}{P_t} \end{aligned} \tag{3.6}$$

Stochastic utility maximization problem is taking form of:

$$\underset{C_t(i),N_t(i),M_t(i),B_t(i,)K_{t+1}(i)}{max} : U\left(C_t(i), N_t(i), \frac{M_t(i)}{P_t}\right) \quad \text{s. t.} \quad \text{BC}$$

This problem can be solved by forming and solving a Lagrangian or Bellman equation, providing the same first order conditions. Whereas solving Lagrangian is simpler, Bellman equation allows for direct cardinal evaluation of preferences. Nonetheless, for this problem is Lagrangian more then sufficient and is utilized in all optimization problems.

$$\mathcal{L}_h = \mathbb{E}_t\beta^t\Bigg[\sum_{t=0}^{\infty} U\left(C_t(i), N_t(i), \frac{M_t(i)}{P_t}\right) + \lambda_t^h\Big(W_t(i)N_t(i) + r_tK_t(i) + ...$$
$$+(1-\delta)K_t(i)\mathrm{e}^{-b_t} + \frac{(1+i_{t-1})B_{t-1}(i)}{P_t} + \frac{M_{t-1}(i)}{P_t} - C_t(i) - \frac{B_t(i)}{P_t} - \frac{M_t(i)}{P_t} - K_{t+1}(i)\mathrm{e}^{-b_t}\Big)\Bigg]$$

Lagrangian is differentiated by choice variables and first order conditions associated with this problem are:

$$\frac{\partial\mathcal{L}_h}{\partial C_t(i)} : \ \mathrm{e}^{\epsilon_t^c}\beta^t\Big(C_t(i) - hC_{t-1}(i)\Big)^{-\sigma} + \mathrm{e}^{\epsilon_{t+1}^c}\beta^{t+1}\Big(C_{t+1}(i) - hC_t(i)\Big)^{-\sigma}(-h) = \beta^t\lambda_t^h$$

$$\frac{\partial\mathcal{L}_h}{\partial N_t(i)} : \ \psi\mathrm{e}^{\epsilon_t^l}N_t(i)^{\eta} = \lambda_t^hW_t(i)$$

$$\frac{\partial\mathcal{L}_h}{\partial M_t(i)} : \ \beta^t\left[\left(\frac{M_t(i)}{P_t}\right)^{-\mu}\frac{1}{P_t}\right] = \beta^t\lambda_t^h\frac{1}{P_t} - \beta^{t+1}\lambda_{t+1}^h\frac{1}{P_{t+1}}$$

$$\frac{\partial\mathcal{L}_h}{\partial B_t(i)} : \ \beta^t\lambda_t^h\frac{1}{P_t} = \beta^{t+1}\lambda_{t+1}^h\frac{(1+i_t)}{P_{t+1}}$$

$$\frac{\partial\mathcal{L}_h}{\partial K_{t+1}(i)} : \ \beta^t\lambda_t^h\mathrm{e}^{-b_t} = \beta^{t+1}\lambda_{t+1}^h((1-\delta)\mathrm{e}^{-b_{t+1}} + r_{t+1})$$

with K_{t+1} as a predetermined variable. Defining marginal utility of consumption at time t as $u_c^t = \lambda_t^h$ and in time $t+1$ as $u_c^{t+1} = \lambda_{t+1}^h$ as and in similar fashion u_n^t as marginal disutility from work at time t and u_m^t as marginal utility from real money holdings at time t, first order conditions can be rewritten and manipulated as follows:

$$C_t(i) : u_c^t = \lambda_t^h = \mathrm{e}^{\epsilon_t^c}\Big(C_t(i) - hC_{t-1}(i)\Big)^{-\sigma} - \mathrm{e}^{\epsilon_{t+1}^c}h\beta\Big(C_{t+1}(i) - hC_t(i)\Big)^{-\sigma} \tag{3.7}$$

$$N_t(i) : u_l^t = u_c^t W_t(i) = u_c^t MPL_t \tag{3.8}$$

$$M_t(i) : u_m^t = u_c^t - \beta u_c^{t+1} \frac{1}{1+\pi_{t+1}} \tag{3.9}$$

$$B_t(i) : u_c^t = \beta u_c^{t+1} \frac{1+i_t}{1+\pi_{t+1}} \tag{3.10}$$

$$K_{t+1}(i) : u_c^t = \beta u_c^{t+1} \frac{\mathrm{e}^{b_t}}{\mathrm{e}^{b_{t+1}}} \Big(1 - \delta + \mathrm{e}^{b_{t+1}} r_{t+1}\Big) \tag{3.11}$$

Where $(1+\pi_{t+1}) = \dfrac{P_{t+1}}{P_t}$ is inflation index. Note the equivalence of equation (3.10) and (3.11) through the Fisher relation[1]:

$$1 + r_{t+1}^{real} = \frac{1+i_t}{1+\pi_{t+1}} \tag{3.12}$$

Utilizing the relationship from intermediate firm's optimization where it will be derived that marginal return from capital is equal to $R_{t+1} = MPK_{t+1}$, real rate of return from capital r_t is given by:

$$r_{t+1}^{real} = r_{t+1} - \delta \tag{3.13}$$

which follows to:

$$1 + r_{t+1}^{real} = 1 + r_{t+1} - \delta = \frac{1+i_t}{1+\pi_{t+1}} \tag{3.14}$$

This result states that real rate of return from holding bonds and capital ownership has to be equal across the entire economy. Substituting (3.11) into (3.9) and (3.14) into (3.11), first order conditions for $C_t(i)$ and $M_t(i)$ are taking form of:

$$u_c^t = \beta u_c^{t+1} \Big(\frac{1+i_t}{1+\pi_{t+1}}\Big) \tag{3.15}$$

$$u_m^t = \beta u_c^{t+1} \Big(\frac{1+i_t}{1+\pi_{t+1}} - \frac{1}{1+\pi_{t+1}}\Big) \tag{3.16}$$

and (3.17) is then the marginal rate of substitution between real money holdings and consumption. It also states that money demand is a function of nominal interest rate, i.e. the opportunity costs from holding money, as i_t can be interpreted as difference in

[1] With $\mathbb{E}b_t = \mathbb{E}c_t = 0$

real return from capital and money holdings[2]. Friedman rule of zero nominal interest rate, implying stable deflationary environment where $r_t = -\pi_{t+1}$ and resulting in non-substitutability between money and consumption, is the most effective outcome in this model as well.

$$\frac{u^t_m}{u^t_c} = \frac{i_t}{1+i_t} \tag{3.17}$$

3.2 Labour Market

There exists a labour packer firm, or a labour union, that packs and aggregate differentiated labour services supplied by households into homogeneous labour service later provided to intermediate producers via a Dixit-Stiglitz type of aggregator:

$$N^d_t \equiv \left[\int_0^1 N_t(i)^{\frac{\epsilon_l - 1}{\epsilon_l}} di \right]^{\frac{\epsilon_l}{\epsilon_l - 1}} \tag{3.18}$$

where N^d_t denotes aggregate labour demand and $N_t(i)$ differentiated labour supply. Labour packer takes all individual differentiated wages $W_t(i)$ and aggregate (average) wage W_t as given and maximizes its profit subject to (3.18):

$$\underset{N_t(i)}{max}\ W_t N_t - \int_0^1 W_t(i) N_t(i) di$$

$$\mathcal{L}_w = W_t N_t - \lambda^w_t \left(\int_0^1 W_t(i) N_t(i) di \right)$$

First order condition for this problem is:

$$\frac{\partial \mathcal{L}_w}{\partial N_t(i)} :\ W_t \frac{\epsilon_l}{\epsilon_l - 1} \left(\int_0^1 N_t(i)^{\frac{\epsilon_l - 1}{\epsilon_l}} di \right)^{\frac{\epsilon_l}{\epsilon_l - 1} - 1} \frac{\epsilon_l - 1}{\epsilon_l} N_t(i)^{\frac{\epsilon_l}{\epsilon_l - 1} - 1} - \lambda^w_t W_t(i) = 0$$

Considering equivalent FOC for labour i and j and by perfect competition imposed zero-profit condition $W_t N^d_t = \int_0^1 W_t(i) N_t(i) di$, demand function for per capita labour

[2]With approximation of Fisher relation as $i_t \approx r^{real}_t + \pi_{t+1}$; where r_t is the real return to capital and $-\pi_{t+1}$ is real return of money

type i is obtained:

$$N_t(i) = \left(\frac{W_t(i)}{W_t}\right)^{-\epsilon_l} N_t^d \tag{3.19}$$

Aggregate wage is also found by utilization of zero-profit condition such that:

$$W_t N_t^d = \int_0^1 W_t(i) \left(\frac{W_t(i)}{W_t}\right)^{-\epsilon_l} N_t^d \, di$$

$$W_t^{1-\epsilon_l} = \int_0^1 W_t(i)(W_t(i))^{-\epsilon_l} \, di$$

to again obtain Dixit-Stiglitz type of aggregator function describing relationship for aggregate (average) wage in this economy:

$$W_t \equiv \left[\int_0^1 W_t(i)^{1-\epsilon_l} di\right]^{\frac{1}{1-\epsilon_l}} \tag{3.20}$$

Wage Setting

Each household have a certain monopolistic power when setting their wage, which is subject to a Calvo (1983) pricing. In each period, each household faces a probability of $(1-\theta_w)$ that it will be allowed to optimally adjust its wage, and at the same time probability θ_w that it will remain stacked with old wage from previous period. However, households not allowed to optimize can partially index their wage to past inflation as in Smets and Wouters (2003) or Adolfson et al. (2005). Rigorously:

$$W_t(i) = \begin{cases} W_t^*(i), & \text{with probability } (1-\theta_w) \\ \Pi_{t-1}^{\kappa_w} W_{t-1}(i), & \text{with probability } \theta_w \end{cases}$$

where W_t^* is the optimal rest wage for period t, Π_{t-1} is gross inflation index from previous period and $\kappa_w \in [0,1]$ is indexation controlling parameter, with 0 meaning no indexation and 1 full indexation to past inflation. Solution of wage setting problem follows (Fernandez-Villaverde, Rubio-Ramirez; 2006). With above imposed conditions, relevant part of Lagrangian for household optimization becomes:

$$\underset{W_t(i)}{max}\ \mathbb{E}_t \sum_{k=0}^{\infty} \Big(\theta_w \beta\Big)^k \Big(-\psi e^{\epsilon^l_{t+k}} \frac{N_{t+k}(i)^{1+\eta}}{1+\eta} + \lambda^h_{t+k} \prod_{s=1}^{k} \frac{\Pi^{\kappa w}_{t+s-1}}{\Pi_{t+s}} W_t(i) N_{t+k}(i) \Big)$$

s.t.

$$N_{t+k}(i) = \left(\prod_{s=1}^{k} \frac{\Pi^{\kappa w}_{t+s-1}}{\Pi_{t+s}} \frac{W_t(i)}{W_{t+k}} \right)^{-\epsilon_l} N_{t+k}$$

Pre-multiplying objective function by $1 = \dfrac{W_{t+k}}{W_{t+k}}$ and substituting constraint to objective function gives:

$$\underset{W_t(i)}{max}\ \mathbb{E}_t \sum_{k=0}^{\infty} \Big(\theta_w \beta\Big)^k \Bigg\{ \frac{-\psi e^{\epsilon^l_{t+k}}}{1+\eta} \left(\prod_{s=1}^{k} \frac{\Pi^{\kappa w}_{t+s-1}}{\Pi_{t+s}} \frac{W_t(i)}{W_{t+k}} \right)^{-\epsilon_l(1+\eta)} N^{1+\eta}_{t+k} + \ldots$$
$$\lambda^h_{t+k} \left(\prod_{s=1}^{k} \frac{\Pi^{\kappa w}_{t+s-1}}{\Pi_{t+s}} \frac{W_t(i)}{W_{t+k}} \right)^{1-\epsilon_l} N_{t+k} W_{t+k} \Bigg\} \quad (3.21)$$

First order condition is:

$$\mathbb{E}_t \sum_{k=0}^{\infty} \Big(\theta_w \beta\Big)^k \Bigg\{ \frac{\epsilon_l \psi e^{\epsilon^l_{t+k}}}{W^*_t} \left(\prod_{s=1}^{k} \frac{\Pi^{\kappa w}_{t+s-1}}{\Pi_{t+s}} \frac{W^*_t}{W_{t+k}} \right)^{-\epsilon_l(1+\eta)} N^{1+\eta}_{t+k} + \ldots$$
$$\lambda^h_{t+k}(1-\epsilon_l) \left(\prod_{s=1}^{k} \frac{\Pi^{\kappa w}_{t+s-1}}{\Pi_{t+s}} \right)^{1-\epsilon_l} \left(\frac{W^*_t}{W_{t+k}} \right)^{-\epsilon_l} N_{t+k} \Bigg\}$$

Due to complete markets property allowing for risk sharing in timing of wage setting, all household set the same wage W^*_t and index (i) is then dropped (Fernandez-Villaverde, Rubio-Ramirez; 2006). Defining equality $w^1_t = w^2_t$ for first order condition, it can be rewritten as:

$$w^1_t = \mathbb{E}_t \sum_{k=0}^{\infty} \Big(\theta_w \beta\Big)^k \psi e^{\epsilon^l_{t+k}} \left(\prod_{s=1}^{k} \frac{\Pi^{\kappa w}_{t+s-1}}{\Pi_{t+s}} \right)^{-\epsilon_l(1+\eta)} \left(\frac{W_{t+k}}{W^*_t} \right)^{\epsilon_l(1+\eta)} N^{1+\eta}_{t+k} \quad (3.22)$$

and

$$w_t^2 = \frac{\epsilon_l - 1}{\epsilon_l} W_t^* \mathbb{E}_t \sum_{k=0}^{\infty} \left(\theta_w \beta\right)^k \lambda_{t+k}^h \left(\prod_{s=1}^{k} \frac{\Pi_{t+s-1}^{\kappa w}}{\Pi_{t+s}} \right)^{1-\epsilon_l} \left(\frac{W_{t+k}}{W_t^*} \right)^{\epsilon_l} N_{t+k} \tag{3.23}$$

which recursive solution eliminating sums and products operators leads to:

$$w_t^1 = \psi e^{\epsilon_t^l} \left(\frac{W_t}{W_t^*} \right)^{\epsilon_l (1+\eta)} N_t^{1+\eta} + \beta \theta_w \mathbb{E}_t \left(\frac{\Pi_t^{\kappa w}}{\Pi_{t+1}} \right)^{-\epsilon_l (1+\eta)} \left(\frac{W_{t+1}^*}{W_t^*} \right)^{\epsilon_l (1+\eta)} w_{t+1}^1 \tag{3.24}$$

and

$$w_t^2 = \frac{\epsilon_l - 1}{\epsilon_l} \left(W_t^* \right)^{1-\epsilon_l} \lambda_t^h W_t^{\epsilon_l} N_t + \beta \theta_w \mathbb{E}_t \left(\frac{\Pi_t^{\kappa w}}{\Pi_{t+1}} \right)^{1-\epsilon_l} \left(\frac{W_{t+1}^*}{W_t^*} \right)^{\epsilon_l - 1} w_{t+1}^2 \tag{3.25}$$

but since $w_t = w_t^1 = w_t^2$ we can write that:

$$w_t = \psi e^{\epsilon_t^l} \left(\frac{W_t}{W_t^*} \right)^{\epsilon_l (1+\eta)} N_t^{1+\eta} + \beta \theta_w \mathbb{E}_t \left(\frac{\Pi_t^{\kappa w}}{\Pi_{t+1}} \right)^{-\epsilon_l (1+\eta)} \left(\frac{W_{t+1}^*}{W_t^*} \right)^{\epsilon_l (1+\eta)} w_{t+1} \tag{3.26}$$

and

$$w_t = \frac{\epsilon_l - 1}{\epsilon_l} \left(W_t^* \right)^{1-\epsilon_l} \lambda_t^h W_t^{\epsilon_l} N_t + \beta \theta_w \mathbb{E}_t \left(\frac{\Pi_t^{\kappa w}}{\Pi_{t+1}} \right)^{1-\epsilon_l} \left(\frac{W_{t+1}^*}{W_t^*} \right)^{\epsilon_l - 1} w_{t+1} \tag{3.27}$$

3.3 Firms

There are two types of firms: intermediate and final producers. Final goods producers operate on perfectly competitive market and produce single homogeneous good, while intermediate markets exhibits monopolistic competition property implied by heterogeneous nature of their production.

Final Goods Producers

Producers of final goods buy heterogeneous intermediate producers output $Y_t(i)$ and use it as an input to produce single homogeneous product Y_t designed for final consumption of households. Production function is basically a CES aggregate over a continuum of intermediate goods (i) distributed on the unit interval:

$$Y_t \equiv \left(\int_0^1 Y_t(i)^{\frac{\epsilon-1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}} \tag{3.28}$$

where ϵ denotes elasticity of substitution between consumption goods. Final production sector is perfectly competitive and cost minimization problem subject to a downward sloping demand curve takes form of:

$$\underset{Y_t(i)}{min} : P_t \left(\int_0^1 Y_t(i)^{\frac{\epsilon-1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}} \quad \text{s. t} \quad \left(P_t Y_t = \int_0^1 Y_t(i) P_t(i) \right)$$

Optimization is performed by solving the Lagrangian:

$$\mathcal{L}_f = P_t \left(\int_0^1 Y_t(i)^{\frac{\epsilon-1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}} - \lambda_t^f \left(\int_0^1 Y_t(i) P_t(i) di \right)$$

First order condition for this problem is:

$$\frac{\partial \mathcal{L}_f}{\partial Y_t(i)} : \; P_t \frac{\epsilon}{\epsilon - 1} \left(\int_0^1 Y_t(i)^{\frac{\epsilon-1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}-1} \frac{\epsilon - 1}{\epsilon} Y_t(i)^{\frac{\epsilon}{\epsilon-1}-1} - \lambda_t^f P_t(i) = 0$$

Downward sloping demand curve for each intermediate good is then:

$$Y_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\epsilon} Y_t \tag{3.29}$$

Using the aggregate sum of total expenditures, overall composite price index is determined as:

$$P_t \equiv \left[\int_0^1 P_t(i)^{1-\epsilon} di \right]^{\frac{1}{1-\epsilon}} \tag{3.30}$$

Intermediate Goods Producers

In the intermediate producer‘s sector, there is a continuum of these firms defined over unit interval $(i) \in [0,1]$ producing corresponding number of heterogeneous products. Intermediate sector operates in monopolistic competition market structure providing monopolistic power over their products for each intermediate producer. Production technology is of Cobb-Douglas type:

$$Y_t(i) = \mathrm{e}^{a_t} K_t^{\alpha}(i)(N_t^d(i))^{1-\alpha} \tag{3.31}$$

where:

$$a_t = \rho_a a_{t-1} + \varepsilon_t^a \quad \text{, where } \varepsilon_t^a \sim N(0, \sigma_a^2) \tag{3.32}$$

is the TFP shock process defined as AR(1) and written in terms of level, i.e. aggregate level of technology, $K_t(i)$ is the amount of rented capital from households and $N_t^d(i)$ is the amount of labour services rented from labour packer by firm i. Cost minimization problem is subject to production technology and downward sloping demand curve of final goods producers for intermediate output:

$$\underset{K_t(i),N_t(i)}{min} : W_t N_t(i) + r_t K_t(i) \quad \text{s. t.} \quad \mathrm{e}^{a_t} K_t(i)^{\alpha} N_t(i)^{1-\alpha} \geq \left(\frac{P_t(i)}{P_t}\right)^{-\epsilon_j} Y_t$$

in the form of Lagrangian as:

$$\mathcal{L}_i = W_t N_t(i) + r_t K_t(i) - \lambda_t^i \Big(\mathrm{e}^{a_t} K_t(i)^{\alpha} N_t(i)^{1-\alpha} - \left(\frac{P_t(i)}{P_t}\right)^{-\epsilon_j} Y_t\Big)$$

Optimization of intermediate producers consists of two stages. At, first optimal production inputs, capital $K_t(i)$ and labour demand $N^d(i)$ are chosen. First order conditions for this stage are:

$$\frac{\partial \mathcal{L}_i}{\partial N_t} : \; W_t = (1-\alpha)\lambda_t^i \mathrm{e}^{a_t} \left(\frac{K_t(i)}{N_t(i)}\right)^{\alpha}$$

$$\frac{\partial \mathcal{L}_i}{\partial K_t} : \; r_t = \alpha \lambda_t^i \mathrm{e}^{a_t} \left(\frac{N_t(i)}{K_t(i)}\right)^{1-\alpha}$$

By dividing first order conditions by each other, capital-labour ratio (3.33) representing optimal inputs equation is obtained. Due to symmetric equilibrium nature of this model, subscript (i) can be dropped, meaning that capital-labour ratio is identical across entire sector of intermediate firms.

$$\frac{W_t}{r_t} = \frac{1-\alpha}{\alpha}\left(\frac{K_t}{N_t}\right) \tag{3.33}$$

To derive marginal cost, definition of unconstrained problem by solving the constrain for $K_t = \left(\frac{Y_t}{e^{a_t} N_t^{1-\alpha}}\right)^{\frac{1}{\alpha}}$, optimal utilization of labour and capital is so that satisfy:

$$N_t^* = \left(\frac{1-\alpha}{\alpha}\frac{r_t}{W_t}\right)^{\alpha}\frac{Y_t}{e^{a_t}}$$

$$K_t^* = \left(\frac{\alpha}{1-\alpha}\frac{W_t}{r_t}\right)^{1-\alpha}\frac{Y_t}{e^{a_t}}$$

Plugging into original minimization problem:

$$\left[\left(\frac{1-\alpha}{\alpha}\frac{r_t}{W_t}\right)^{\alpha} W_t + \left(\frac{\alpha}{1-\alpha}\frac{W_t}{r_t}\right)^{1-\alpha} r_t\right]\frac{Y_t}{e^{a_t}}$$

and taking derivative w.r.t. Y_t, marginal cost are obtained and defined as:

$$MC_t = \left(\frac{r_t}{\alpha}\right)^{\alpha}\left(\frac{W_t}{1-\alpha}\right)^{1-\alpha}\frac{1}{e^{a_t}} \tag{3.34}$$

Price Setting

In the second stage, given the optimal inputs ration, firms have to choose their optimal price. Pricing mechanism of intermediate producers follows Calvo (1983) and solution procedure Fernandez-Villaverde and Rubio-Ramirez (2006). In each period there is a probability θ that firm will not be able to adjust its price and remain stacked with old price from previous period. On the other hand, there is probability $(1-\theta)$ that firm will have an opportunity to reset price for its product and set it optimally. However, non-optimizing firms can partially index their price to previous period inflation Π_{t-1}. Recalling composite price index (3.30), the price evolution equation is derived as:

$$P_t = \left(\int_0^{1-\theta} P_t^{*1-\epsilon}(i)di + \int_{1-\theta}^{1} (\Pi_{t-1}^{\kappa})^{1-\epsilon} P_{t-1}^{1-\epsilon}(i)di \right)^{\frac{1}{1-\epsilon}} \tag{3.35}$$

where $\Pi_{t-1} = \dfrac{P_{t-1}}{P_{t-2}}$ is an inflation index, $\kappa \in [0,1]$ is an controlling inflation-indexation parameter, with $\kappa = 0$ meaning no indexation and $\kappa = 1$ resulting in perfect indexation to past inflation. Gross inflation index exhibits can be factorized to $\Pi_t = \dfrac{P_t}{P_{t-1}}$ and $\Pi_t = 1 + \pi_t = 1 + \dfrac{P_t}{P_{t-1}} - 1$.

In continuous setting, i.e. with continuum of firms, or in discrete setting with limit case of infinite number of firms, evaluation of these integrals by properties law of large numbers and random selection or firms being able to reset their price, results in exact portion of $(1-\theta)$ setting optimal price and θ firms stacked with old prices.

$$P_t = \left(\theta \left(\Pi_{t-1}^{\kappa}\right)^{1-\epsilon} P_{t-1}^{1-\epsilon} + (1-\theta) P_t^{*1-\epsilon} \right)^{\frac{1}{1-\epsilon}} \tag{3.36}$$

or after dividing by $P_t^{1-\epsilon}$ equivalently expressed as the law of motion of prices as:

$$1 = \theta \left(\frac{\Pi_{t-1}^{\kappa}}{\Pi_t} \right)^{1-\epsilon} + (1-\theta) \left(\frac{P_t^*}{P_t} \right)^{1-\epsilon} \tag{3.37}$$

Firms optimization problem is to choose price maximizing the sum of its currently expected discounted flow of future profits. These are given by the difference in set price per (i) unit of production in given time period and its marginal cost, i.e. cost for one additional (i) unit of production and scaled by overall quantity of product (i) in that period:

$$\underset{P_t(i)}{max}\ \mathbb{E}_t \sum_{k=0}^{\infty} (\theta\,\beta)^k \lambda_{t+k}^h \left(\prod_{s=1}^{k} \Pi_{t+s-1}^{\kappa} \frac{P_t(i)}{P_{t+k}} - MC_{t+k}(i) \right) Y_{t+k}(i)$$

s.t.

$$Y_{t+k}(i) = \left(\prod_{s=1}^{k} \Pi_{t+s-1}^{\kappa} \frac{P_t(i)}{P_{t+k}} \right)^{-\epsilon} Y_{t+k}$$

Unconstrained problem is obtained by inserting demand function into objective function.

$$\underset{P_t(i)}{max}\, \mathbb{E}_t \sum_{k=0}^{\infty} \Big(\theta\beta\Big)^k \lambda_{t+k}^h \Bigg\{ \left(\prod_{s=1}^{k} \Pi_{t+s-1}^{\kappa} \frac{P_t(i)}{P_{t+k}} \right)^{1-\epsilon} - \left(\prod_{s=1}^{k} \Pi_{t+s-1}^{\kappa} \frac{P_t(i)}{P_{t+k}} \right)^{-\epsilon_j} MC_{t+k}(i) \Bigg\} Y_{t+k}$$

In the next step, price evolution relationship given by equation (3.38) is utilized to eliminate P_{t+k} from (3.3) and so the optimization depends only on currently expected evolution of future inflation. It follows that price level P_{t+k} is given by current price level P_t and basis inflation index $\frac{P_{t+k}}{P_t}$.

$$P_{t+k} = P_t \prod_{s=1}^{k} \Pi_{t+s} = P_t \frac{P_{t+1}}{P_t} \frac{P_{t+2}}{P_{t+1}} \cdots \frac{P_{t+k}}{P_{t+k-1}} = P_t \frac{P_{t+k}}{P_t} \tag{3.38}$$

$$\mathbb{E}_t \sum_{k=0}^{\infty} \Big(\theta\,\beta\Big)^k \lambda_{t+k}^h \Bigg\{ \left(\prod_{s=1}^{k} \frac{\Pi_{t+s-1}^{\kappa}}{\Pi_{t+s}} \frac{P_t(i)}{P_t} \right)^{1-\epsilon} - \left(\prod_{s=1}^{k} \frac{\Pi_{t+s-1}^{\kappa}}{\Pi_{t+s}} \frac{P_t(i)}{P_t} \right)^{-\epsilon} MC_{t+k}(i) \Bigg\} Y_{t+k} \tag{3.39}$$

Also note, that with perfect price indexation to past inflation, i.e. when $\kappa_j = 1$, product of inflation indexes collapses $\dfrac{\Pi_t}{\Pi_{t+k}} = \dfrac{P_t P_{t+k-1}}{P_{t-1} P_{t+k}}$. Optimization with respect to $P_{J,t}(i)$, i.e. $\dfrac{\partial(\cdot)}{\partial P_{J,t}(i)}$, first order conditions after multiplication by $1 = \dfrac{P_t^*(i)}{P_t^*(i)}$ can be manipulated to:

$$\mathbb{E}_t \sum_{k=0}^{\infty} \Big(\theta\beta\Big)^k \lambda_{t+k}^h \Bigg\{ \left((1-\epsilon) \prod_{s=1}^{k} \frac{\Pi_{t+s-1}^{\kappa}}{\Pi_{t+s}} \frac{P_t^*(i)}{P_t} \right)^{1-\epsilon} \frac{1}{P_t^*(i)} + \ldots$$
$$\epsilon \left(\prod_{s=1}^{k} \frac{\Pi_{t+s-1}^{\kappa}}{\Pi_{t+s}} \frac{P_t^*(i)}{P_t} \right)^{-\epsilon} \frac{MC_{t+k}(i)}{P_t^*(i)} \Bigg\} Y_{t+k}$$

Since only symmetric equilibrium is considered, $P_t^*(i) = P_t^*$ and as in (3.34) $MC_t(i) = MC_t$, i.e. it is assumed that in equilibrium all firms are identical and therefore all face the same optimization problem and prices. After rearranging above stated first order conditions, dropping irrelevant constant and substituting for $P^*_{J,t}(i)$ we get:

$$\mathbb{E}_t \sum_{k=0}^{\infty} \Big(\theta\beta\Big)^k \lambda^h_{t+k} \Bigg\{ \Bigg((1-\epsilon) \prod_{s=1}^{k} \frac{\Pi^{\kappa}_{t+s-1}}{\Pi_{t+s}} \Bigg)^{1-\epsilon} \frac{P^*_t}{P_t} + \epsilon \Bigg(\prod_{s=1}^{k} \frac{\Pi^{\kappa}_{t+s-1}}{\Pi_{t+s}} \Bigg)^{-\epsilon} MC_{t+k} \Bigg\} Y_{J,t+k} \tag{3.40}$$

Devoid of price rigidity in the price setting mechanism, implying that $\theta_{_J} = 0$, (3.40) can be evaluated only in current period $k = 0$ (otherwise its value is zero) and simplifies to usual expression for mark-up representing monopolistic power of a firm.

$$P^*_t = \frac{\epsilon}{\epsilon - 1} P_t MC_t$$

Now, let define equality $\epsilon f^1_t = (\epsilon - 1) f^2_t$ to equalize previously defined first order condition. The purpose is to express FOC recursively and by rewriting it in the form of difference equation to eliminate sums and products operators.

$$f^1_t = \mathbb{E}_t \sum_{k=0}^{\infty} \Big(\theta\beta\Big)^k \lambda^h_{t+k} \Bigg(\prod_{s=1}^{k} \frac{\Pi^{\kappa}_{t+s-1}}{\Pi_{t+s}} \Bigg)^{1-\epsilon} \frac{P^*_t}{P_t} Y_{t+k} \tag{3.41}$$

and

$$f^2_t = \mathbb{E}_t \sum_{k=0}^{\infty} \Big(\theta\beta\Big)^k \lambda^h_{t+k} \Bigg(\prod_{s=1}^{k} \frac{\Pi^{\kappa}_{t+s-1}}{\Pi_{t+s}} \Bigg)^{-\epsilon} MC_{t+k}\, Y_{t+k} \tag{3.42}$$

To ensure well definiteness and stationarity of these sums, and thus the solution of this maximization problem, $\Big(\theta\beta\Big)^k$ has to converge to zero in expectations faster than $\prod_{s=1}^{k} \frac{\Pi^{\kappa_j}_{t+s-1}}{\Pi_{t+s}}$ goes to infinity (Fernandez-Villaverde, Rubio-Ramirez; 2006). Recursive expression is taking form of:

$$f^1_t = \lambda^h_t \Pi^*_t Y_t + \theta\beta\mathbb{E} \Bigg(\frac{\Pi^{\kappa}_t}{\Pi_{t+1}} \Bigg)^{1-\epsilon} \Bigg(\frac{\Pi^*_t}{\Pi^*_{t+1}} \Bigg) f^1_{t+1} \tag{3.43}$$

$$f^2_t = \lambda^h_t MC_t Y_t + \theta_{_J}\beta\mathbb{E} \Bigg(\frac{\Pi^{\kappa}_t}{\Pi_{t+1}} \Bigg)^{-\epsilon} f^2_{t+1} \tag{3.44}$$

where previous FOC is:

$$\epsilon f^1_t = (\epsilon - 1) f^2_t$$

and optimal reset price in terms of gross inflation index is:

$$\Pi_t^* = \frac{P_t^*}{P_t}$$

3.4 Government

Role of the government int his model is marginal and consists only of setting of nominal interest rate according to Taylor rule:

$$\frac{R_t}{\bar{R}} = \left(\frac{R_{t-1}}{\bar{R}}\right)^{\gamma_R} \left(\left(\frac{\Pi_t}{\bar{\Pi}}\right)^{\gamma_\Pi} \left(\frac{Y_t}{Y_{t-1}}\right)^{\gamma_Y}\right)^{1-\gamma_R} \mathrm{e}^{\epsilon_t^m} \tag{3.45}$$

where R_t is the gross nominal interest rate $R_t = 1 + i_t$ and $\bar{R}$ is the corresponding equilibrium rate. Next, ϵ_t^m represents monetary shock defined as AR(1) stationary process:

$$\epsilon_t^m = \rho_m \epsilon_{t-1}^m + \varepsilon_t^m \tag{3.46}$$

There is no money growth equation, i.e. real demand for money and amount of money in this economy is identical. This imply strictly endogenous property of money supply. Government does not impose any taxes (for individuals, consumption, firms or profits) and have no income from seigniorage or inflation tax. Thus, it cannot provide transfers and cannot have its own expenditures.

3.5 Aggregation

Symmetric equilibrium properties of this model define $C_t(i) = C_t$, $I_t(i) = I_t$, $K_t(i) = K_t$ and $W_t^*(i) = W_t^*$. At first, expression for aggregate demand is derived:

$$Y_t = C_t + I_t \tag{3.47}$$

Individual demand for products of each intermediate producer is:

$$Y_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\epsilon}(C_t + I_t) \tag{3.48}$$

and substituting production function leads to:

$$\mathrm{e}^{a_t} K_t^{\alpha}(i) N_t^{1-\alpha}(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\epsilon}(C_t + I_t) \tag{3.49}$$

Now, let define $\upsilon_t^p = \int_0^1 \left(\frac{P_t(i)}{P_t}\right)^{-\epsilon} di$ to get:

$$C_t + I_t = \frac{\mathrm{e}^{a_t} K_t^{\alpha}(i) N_t^{1-\alpha}(i)}{\upsilon_t^p} \tag{3.50}$$

Deflator index υ_t^p is the measure of dispersion of intermediate producer‘s relative prices. Next to the distortion associated with monopolistic competition and corresponding monopolistic power of intermediate producers, this is another distortion associated with fluctuations of relative prices due to price stickiness (Sims, 2016). Essentially, in flexible price models, this term is always equal to one, i.e. $\upsilon_t^p = 1$, as all firms choose the same price. Degree of price dispersion depends on stickiness parameter θ, degree of indexation to past inflation κ and inflation targeted by central bank $\bar{\Pi}$. This equation can be break down by Calvo pricing properties to:

$$\upsilon_t^p = \int_0^{1-\theta} \left(\frac{P_t^*}{P_t}\right)^{-\epsilon} di + \int_{1-\theta}^{1} \left(\frac{\Pi_{t-1}^{\kappa} P_{t-1}(i)}{P_t}\right)^{-\epsilon} di \tag{3.51}$$

Pre-multiplying second term by $\frac{P_{t-1}}{P_{t-1}} = 1$ and substituting the definition of optimal reset price $\Pi_t^* = \dfrac{P_t^*}{P_t}$ leads to:

$$\upsilon_t^p = (1-\theta)(\Pi_t^*)^{-\epsilon} + \int_{1-\theta}^{1} \Pi_{t-1}^{-\kappa\epsilon} \left(\frac{P_{t-1}(i)}{P_{t-1}} \frac{P_{t-1}}{P_t}\right)^{-\epsilon} di \tag{3.52}$$

where the evaluation of the first integral gives by the law of large numbers exactly

$(1-\theta)$ portion of firms allowed to optimally rest their price. Evaluation of second integral leaves θ portion of firms allowed only to partially index their prices to past inflation:

$$\upsilon_t^p = (1-\theta)(\Pi_t^*)^{-\epsilon} + \theta\left(\frac{\Pi_{t-1}^{\kappa_j}}{\Pi_t}\right)^{-\epsilon_j} \int_0^1 \left(\frac{P_{t-1}}{P_t}\right)^{-\epsilon} di \tag{3.53}$$

Change of lower bound of remaining integral form $(1-\theta)$ to 0 follows from the random selection of these firms. Further evaluation fo this integral result in obtaining υ_{t-1}^p. Thus, getting all together, the law of motion of price dispersion is:

$$\upsilon_t^p = (1-\theta)(\Pi_t^*)^{-\epsilon} + \theta\left(\frac{\Pi_{t-1}^{\kappa_j}}{\Pi_t}\right)^{-\epsilon_j} \upsilon_{t-1}^p \tag{3.54}$$

Defining aggregate labour supply as an integral over all households i, $N_t = \int_0^1 N_t(i)di$, aggregate labour demand N_t^d is derived from (3.19) as:

$$N_t = \int_0^1 N_t(i)di = \int_0^1 \left(\frac{W_t(i)}{W_t}\right)^{-\epsilon_l} di\ N_t^d \tag{3.55}$$

Following exactly the same steps as in case of aggregate output definition:

$$\upsilon_t^w = \int_0^1 \left(\frac{W_t(i)}{W_t}\right)^{-\epsilon_l} di = \theta_w \left(\frac{W_{t-1}}{W_t}\frac{\Pi_{t-1}^{\kappa_w}}{\Pi_t}\right)^{-\epsilon_l} \upsilon_{t-1}^w + (1-\theta_W)(\Pi_t^{*w})^{-\epsilon_l} \tag{3.56}$$

is the measure of wage dispersion cased by imperfect wage setting mechanism. Relationship between aggregate labour demand is supply is then:

$$N_t^d = \frac{N_t}{\upsilon_t^w} \tag{3.57}$$

closing last part of this model.

3.6 Equilibrium

Model is given by the following set of equations. Each of them is implemented in *Dynare*

$$\lambda_t^h = \Big(C_t(i) - hC_{t-1}(i)\Big)^{-\sigma} - h\beta\Big(C_{t+1}(i) - hC_t(i)\Big)^{-\sigma} \tag{3.58}$$

$$m^{-\mu} = \lambda_t^h - \beta\lambda_{t+1}^h \frac{1}{\Pi_{t+1}} \tag{3.59}$$

$$\lambda_t^h = \beta\lambda_{t+1}^h \frac{1+i_t}{\Pi_{t+1}} \tag{3.60}$$

$$\lambda_t^h = \beta\lambda_{t+1}^h \frac{\mathrm{e}^{b_t}}{\mathrm{e}^{b_{t+1}}}\Big(1-\delta+\mathrm{e}^{b_{t+1}} r_{t+1}\Big) \tag{3.61}$$

$$w_t = \psi \mathrm{e}^{\epsilon_t^l l}\left(\frac{W_t}{W_t^*}\right)^{\epsilon_l(1+\eta)} N_t^{1+\eta} + \beta\theta_w \mathbb{E}_t\left(\frac{\Pi_t^{\kappa w}}{\Pi_{t+1}}\right)^{-\epsilon_l(1+\eta)}\left(\frac{W_{t+1}^*}{W_t^*}\right)^{\epsilon_l(1+\eta)} w_{t+1} \tag{3.62}$$

$$w_t = \frac{\epsilon_l - 1}{\epsilon_l}\Big(W_t^*\Big)^{1-\epsilon_l}\lambda_t^h W_t^{\epsilon_l} N_t + \beta\theta_w \mathbb{E}_t\left(\frac{\Pi_t^{\kappa w}}{\Pi_{t+1}}\right)^{1-\epsilon_l}\left(\frac{W_{t+1}^*}{W_t^*}\right)^{\epsilon_l - 1} w_{t+1} \tag{3.63}$$

$$f_t^1 = \lambda_t^h \Pi_t^* Y_t + \theta\beta\mathbb{E}\left(\frac{\Pi_t^\kappa}{\Pi_{t+1}}\right)^{1-\epsilon}\left(\frac{\Pi_t^*}{\Pi_{t+1}^*}\right) f_{t+1}^1 \tag{3.64}$$

$$f_t^2 = \lambda_t^h MC_t Y_t + \theta\beta\mathbb{E}\left(\frac{\Pi_t^\kappa}{\Pi_{t+1}}\right)^{-\epsilon} f_{t+1}^2 \tag{3.65}$$

$$\epsilon_j f_t^1 = (\epsilon_j - 1) f_t^2 \tag{3.66}$$

$$Y_t = C_t + I_t \tag{3.67}$$

$$Y_t = \frac{\mathrm{e}^{a_t} K_t^\alpha N_t^{1-\alpha}}{\upsilon_t^p} \tag{3.68}$$

$$\upsilon_t^p = \theta\left(\frac{\Pi_{t-1}^\kappa}{\Pi_t}\right)^{-\epsilon}\upsilon_{t-1}^p + (1-\theta)\Pi_t^{*-\epsilon} \tag{3.69}$$

$$\upsilon_t^w = \theta_w\left(\frac{W_{t-1}}{W_t}\frac{\Pi_{t-1}^{\kappa_w}}{\Pi_t}\right)^{-\epsilon_l}\upsilon_{t-1}^w + (1-\theta_W)(\Pi_t^{*w})^{-\epsilon_l} \tag{3.70}$$

$$1 = \theta\left(\frac{\Pi_{t-1}^\kappa}{\Pi_t}\right)^{1-\epsilon} + (1-\theta)\Pi_t^{*1-\epsilon} \tag{3.71}$$

$$1 = \theta_w \left(\frac{W_{t-1}}{W_t} \frac{\Pi_{t-1}^{\kappa_w}}{\Pi_t} \right)^{1-\epsilon_l} + (1-\theta_w)(\Pi_t^{*w})^{1-\epsilon_l} \tag{3.72}$$

$$\Pi^*_{W,t} = \frac{W^*_t}{W_t} \tag{3.73}$$

$$\frac{R_t}{\bar{R}} = \left(\frac{R_{t-1}}{\bar{R}} \right)^{\gamma_R} \left(\left(\frac{\Pi_t}{\bar{\Pi}} \right)^{\gamma_\Pi} \left(\frac{Y_t}{Y_{t-1}} \right)^{\gamma_Y} \right)^{1-\gamma_R} \mathrm{e}^{\varepsilon_t^m} \tag{3.74}$$

$$N_t = \upsilon_t^w N_t^D \tag{3.75}$$

$$\frac{W_t}{r_t} = \frac{1-\alpha}{\alpha} \left(\frac{K_t}{N_t} \right) \tag{3.76}$$

$$MC_t = \left(\frac{r_t}{\alpha} \right)^{\alpha} \left(\frac{W_t}{1-\alpha} \right)^{1-\alpha} \frac{1}{\mathrm{e}^{a_t}} \tag{3.77}$$

$$K_{t+1}(i) = \mathrm{e}^{b_t} I_t(i) + (1-\delta) K_{t-1}(i) \tag{3.78}$$

$$a_t = \rho_1 a_{t-1} + \varepsilon_t^a \tag{3.79}$$

$$b_t = \rho_1 b_{t-1} + \varepsilon_t^i \tag{3.80}$$

$$\epsilon_t^c = \rho_c \epsilon_{t-1}^c + \varepsilon_t^c \tag{3.81}$$

$$\epsilon_t^l = \rho_l \epsilon_{t-1}^l + \varepsilon_t^l \tag{3.82}$$

$$\epsilon_t^m = \rho_m \epsilon_{t-1}^m + \varepsilon_t^m \tag{3.83}$$

3.7 Model Solution

Steady State

Here, analytical steady state relationships are derived. Household optimality conditions are:

$$(1-h)^{-\sigma}(1-h\beta)C^{-\sigma} = \lambda_h^{ss} \tag{3.84}$$

$$\beta^{-1} = \frac{1+i_t}{1+\pi_t} \tag{3.85}$$

Recalling (3.17), equilibrium condition for $m = \frac{M}{P}$ is:

$$m = \lambda_h^{ss}\Big(\frac{i}{1+i}\Big)^{-\frac{1}{\mu}} \tag{3.86}$$

$$\beta^{-1} = \delta + r \tag{3.87}$$

$$w = (\Pi^*_{W,t})^{-\epsilon_l(1+\eta)}(N^d)^{1+\eta} + \beta\theta_w\left(\frac{\Pi^{\kappa_w}}{\Pi}\right)^{-\epsilon_l(1+\eta)}\left(\frac{W^*}{W^*}\right)^{\epsilon_l(1+\eta)} w \tag{3.88}$$

$$w = \frac{\epsilon_l - 1}{\epsilon_l}\Big(W^*\Big)^{1-\epsilon_l}\lambda W^{\epsilon_l}N^d + \beta\theta_w\left(\frac{\Pi^{\kappa_w}}{\Pi}\right)^{1-\epsilon_l}\left(\frac{W^*}{W^*}\right)^{\epsilon_l - 1} w \tag{3.89}$$

First order condition for firms are:

$$f^1 = \lambda_t^h\Pi^*Y + \theta\beta\left(\frac{\Pi^\kappa}{\Pi}\right)^{1-\epsilon}\left(\frac{\Pi^*}{\Pi^*}\right)f^1 \tag{3.90}$$

$$f^2 = \lambda_t^h MC\,Y + \theta\beta\left(\frac{\Pi^\kappa}{\Pi}\right)^{-\epsilon} f^2 \tag{3.91}$$

$$\epsilon_t^1 = (\epsilon - 1)f^2 \tag{3.92}$$

$$\frac{W}{r} = \frac{1-\alpha}{\alpha}\left(\frac{K}{N}\right) \tag{3.93}$$

$$MC = \left(\frac{r}{\alpha}\right)^{\alpha}\left(\frac{W}{1-\alpha}\right)^{1-\alpha} \tag{3.94}$$

Equilibrium law of motion of prices and wages:

$$1 = \theta\left(\frac{\Pi^\kappa}{\Pi}\right)^{1-\epsilon} + (1-\theta)\Pi^{*1-\epsilon} \tag{3.95}$$

$$1 = \theta_W\left(\frac{\Pi^{\kappa_w}}{\Pi}\right)^{1-\epsilon_l} + (1-\theta_W)(\Pi^{*w})^{1-\epsilon_l} \tag{3.96}$$

Price and wage dispersion:

$$\upsilon^p = \theta\left(\frac{\Pi^\kappa}{\Pi}\right)^{-\epsilon}\upsilon^p + (1-\theta)\Pi^{*-\epsilon} \tag{3.97}$$

$$\upsilon^w = \theta_w \left(\frac{\Pi^{\kappa_w}}{\Pi}\right)^{-\epsilon_l} \upsilon^w + (1-\theta_w)(\Pi^{*w})^{-\epsilon_l} \tag{3.98}$$

Law of motion fo capital (investments):

$$I = \delta K \tag{3.99}$$

Market clearing conditions:

$$Y\upsilon^p = K^\alpha N^{1-\alpha} \tag{3.100}$$

$$N = \upsilon^w N^d \tag{3.101}$$

$$Y = C + I \tag{3.102}$$

And for shock processes

$$\mathrm{e}^{a_t} = \mathrm{e}^{b_t} = \mathrm{e}^{\epsilon_t^c} = \mathrm{e}^{\epsilon_t^l} = \mathrm{e}^{\epsilon_t^m} = 1 \tag{3.103}$$

Analytical Solution

To solve the model and find deterministic steady state, it is necessary to re-arrange previously stated equilibrium conditions such that (i) variables uniquely determined by parameters are computed (ii) labour demand is calculated from market clearing condition and (iii) remaining steady state values depended on labour demand are determined. Throughout the modeling gross interest and inflation rates are utilized instead of their net counterparts, particularly $R = 1 + i$ and $\Pi = 1 + \pi$, respectively. Model solution is then given by following set of equations. Gross interest rate is given by the ratio of gross inflation rate Π and discount factor β:

$$R = \frac{\Pi}{\beta}$$

Note that this relationship also defines optimal gross interest rate $\bar{R}$ targeted by central bank through Taylor rule (3.74). This is crucial to account for in later stages when model is programmed in *Dynare*. Then, real rental rate of capital is given by:

$$r = \beta - 1 + \delta$$

Next, relationships between inflation and optimal relative price and wages are:

$$\Pi^* = \Big(\frac{1 - \theta\Pi^{-(1-\epsilon)(1-\kappa)}}{1-\theta}\Big)^{\frac{1}{1-\epsilon}}$$

$$\Pi^{*w} = \Big(\frac{1 - \theta_w\Pi^{-(1-\epsilon_l)(1-\kappa_w)}}{1-\theta_w}\Big)^{\frac{1}{1-\epsilon_l}}$$

defining equilibrium price and wages dispersion

$$\upsilon^p = \frac{1-\theta}{1-\theta\Pi^{(1-\kappa)\epsilon}}\Pi^{*-\epsilon}$$

$$\upsilon^w = \frac{1-\theta_w}{1-\theta_w\Pi^{(1-\kappa_w)\epsilon_l}}\Pi^{*-\epsilon_l}$$

Marginal costs are obtained from optimal price setting equations:

$$MC = \frac{\epsilon - 1}{\epsilon}\frac{1-\theta\Pi^{(1-\kappa)\epsilon}}{1-\theta\Pi^{-(1-\kappa)(1-\epsilon)}}\Pi^*$$

Availability af marginal costs MC allows for computation of wages W from (3.94):

$$W = (1-\alpha)\Big(\frac{r}{\alpha}\Big)^{\frac{1-\alpha}{\alpha}}MC^{\frac{1}{1-\alpha}}$$

and optimal wage is:

$$W^* = W\Pi^{*w}$$

Equilibrium for wage setting equations of household wage decision are obtained by solving (3.88) and (3.89) for w and utilizing previous relationship for optimal wage W^*:

$$w = \frac{\psi(\Pi^*_w)^{-\epsilon_l(1+\eta)}(N^d)^{1+\eta}}{1-\beta\theta_w\Pi^{\epsilon_l(1+\eta)(1-\kappa_w)}}$$

and

$$w = \frac{\frac{\epsilon_l - 1}{\epsilon_l}W^*(\Pi^*_{w,t})^{-\epsilon_l}\lambda N^d}{1-\beta\theta_w\Pi^{-(1-\epsilon_l)(1-\kappa_w)}}$$

Equalizing both equations and manipulating terms, labour demand N^d is defined as a function of λ:

$$\frac{1-\beta\theta_w\Pi^{\epsilon_l(1+\eta)(1-\kappa_w)}}{1-\beta\theta_w\Pi^{-(1-\epsilon_l)(1-\kappa_w)}}=\frac{\frac{\epsilon_l-1}{\epsilon_l}W^*\lambda}{\psi(\Pi_w^*)^{-\epsilon_l\eta}(N^d)^\eta}$$

To solve for labour demand recall market clearing condition:

$$Y=C+I$$

but since $Y=\dfrac{K^\alpha N^{1-\alpha}}{\upsilon^p}$ and $I=\delta K$ it is the case that:

$$\frac{K^\alpha N^{1-\alpha}}{\upsilon^p}=C+\delta K$$

From relationship of optimal inputs allocation, it is possible to find K as a function of N^d, namely:

$$K=\frac{\alpha}{1-\alpha}\frac{W}{r}N^d=\Omega N^d$$

with defining auxiliary variable $\Omega=\dfrac{\alpha}{1-\alpha}\dfrac{W}{r}$ and substituting for each K:

$$\frac{\Omega^\alpha N^d}{\upsilon^p}=C+\delta\Omega N^d$$

or equivalently

$$C=\frac{\Omega^\alpha N^d}{\upsilon^p}-\delta\Omega N^d$$

Note that definition of λ_h^{ss} given by (3.84) is a function of known parameters and C:

$$\lambda_h^{ss}=(1-h)^{-\sigma}(1-h\beta)C^{-\sigma}$$

and thus λ_h^{ss} can by expressed as a function of parameters, already in steady state evaluated variables and labour demand N^d:

$$\lambda_h^{ss}=(1-h)^{-\sigma}(1-h\beta)\left(\frac{\Omega^\alpha N^d}{\upsilon^p}-\delta\Omega N^d\right)^{-\sigma}$$

Last step is to substitute this λ into labour demand N^d relationship defined by labour market equilibrium as a function of λ:

$$\frac{1-\beta\theta_w\Pi^{\epsilon_l(1+\eta)(1-\kappa_w)}}{1-\beta\theta_w\Pi^{-(1-\epsilon_l)(1-\kappa_w)}} = \frac{\frac{\epsilon_l-1}{\epsilon_l}W^*}{\psi(\Pi_w^*)^{-\epsilon_l\eta}(N^d)^\eta}(1-h)^{-\sigma}(1-h\beta)\left(\frac{\Omega^\alpha N^d}{\upsilon^p}-\delta\Omega N^d\right)^{-\sigma} \tag{3.104}$$

This is a non-linear equation uniquely defining N^d and is solved by Matlab root finder *fsolve*. All other equilibrium conditions are recursive to these:

$$K = \Omega N^d$$

$$I = \delta K$$

$$Y = \frac{\Omega^\alpha N^d}{\upsilon^p}$$

$$C = \frac{\Omega^\alpha N^d}{\upsilon^p} - \delta\Omega N^d = Y - I$$

$$\lambda_h^{ss} = (1-h)^{-\sigma}(1-h\beta)C^{-\sigma}$$

and from (3.17) and definition of $u_c^t = \lambda_t^h$ real money demand is obtained:

$$m = \lambda_h^{ss}\left(\frac{i}{1+i}\right)^{-\frac{1}{\mu}}$$

3.8 Estimation

Original approach for bringing model into data since the rise of RBC model firstly introduced by Kydland and Prescott (1982) was calibration. It represents weak econometric approach, which goal is to simulate such artificial time series that best match the moments of empirical data. Advantage of calibration is that it incorporates micro evidence and set the properties of the model to be the most informative in the relevant area of interest (Smets and Wouters; 2007). On the other hand, one should be careful when building macroeconomic model on microeconomic evidence, since macro model is build in certain context not necessarily the same as the micro study was conducted (Fernandez-Villaverde, 2009).

As opposed to calibration, estimation representing strong econometric approach provides full characterization of empirically observed time series. DSGE models can provide good fit on data when enough shocks are specified (Smets and Wouters, 2002). Estimation can be performed utilizing Maximum likelihood or Bayesian inference, which differences were discussed in section 2. Main downturn of MLE is that likelihood function of a DSGE model is in general highly dimensional object with numerous local minima and maxima and flat, or nearly flat surfaces (Fernandez-Villaverde, 2009). On the other hand, Bayesian econometrics became popular as it is sufficient to apply Bayes rule (theorem) on the data. In fact, Bayesian estimation incorporates both: calibration for setting of prior and MLE for estimation.

Bayesian estimation of DSGE models consists of these steps (Griffoli, 2013; Herbst and Schorfheide, 2015):

1. Select vector of parameters $\boldsymbol{\vartheta}$ of model M to be estimated and set their priors $p(\boldsymbol{\vartheta}, M)$
2. Write down the likelihood function $L(\boldsymbol{\vartheta}|Y_T, M) = p(Y_T|\boldsymbol{\vartheta}, M)$ describing the density of the data approximated by Kalman filter and derive log-likelihood. Kalman filter is used because likelihood function is non-linear in deep parameters, but linear in variables.
3. Combine prior with likelihood function to get posterior density:
$$p(\boldsymbol{\vartheta}|T_T, M) = \frac{p(Y_T|\boldsymbol{\vartheta}, M)p(\boldsymbol{\vartheta}, M)}{p(Y_t|M)}$$
4. Find mode by maximization of the log-posterior kernel (numerator of posterior density) consisting from log-likelihood and priors
5. Approximate posterior distribution employing Random Walk Metropolis-Hastings algorithm

Step 4 is the most tricky and often a reason of failure of the estimation. If mode of the posterior distribution is not correctly identified, it results in not positive definite, and thus invertible, Hessian a MH algorithm cannot be initialized. Such obstacle usually appears when model parameters are (i) poorly identified due to exclusion of time series and inappropriate definition of observation equations or by natural model structure. Or (ii) when mean of prior distribution is set too far from its true value. Likelihood function is Gaussian only with respect to functions of parameters and thus is difficult to maximize. With poor initial specification, optimization algorithm is then unable to find mode and optimization fails. *Dynare* provides mode check plots as an output of every estimation where this issue can be easily identified, but solution always results in try-and-fail method. This issue is closely related to above mentioned flat likelihood of highly dimensional DSGE models. For the mode check plots see Appendix B.

3.8.1 Data and Observation Equations

For the estimation of the model, quarterly data for Czech republic ranging from 1995:Q1 to 2017:Q3 on GDP, consumption, investments, hours worked, interest rate 3M PRIBOR, inflation and GDP deflator are used. As the data sources, ARAD system for time series of Czech National Bank and OECD statistics (for comparative price level) are used. Necessary condition for estimation of any DSGE model are seasonally adjusted time series. Presence of seasonality in the time series would create fluctuations perceived by the model as a spurious economic cycle. To correct for seasonality, Census X-13 filter is employed. Data transformation and corresponding specification of measurement and observational equations is discussed for two cases: (i) non-linear model and (ii) non-linear model for log-linearization. These two differs not only in *Dynare* implementation, but also in specification of observation equations and required data transformation. Despite complete *.mod* files being presented in Appendix A, *Dynare* model block related to observation equations is demonstrated here as well.

In general, raw data for specific variables consists of:

1. **GDP, Consumption, Investment** and **Hours Worked** (Group 1): aggregate level in nominal values (current prices)

2. **Inflation** and **GDP deflator** (Group 2): evolution of comparative price level (GDP deflator) in form of basis index

3. **Interest rate**: level of 3M PRIBOR in percentage points

Descriptive statics are in Table 3.1 and visualized is Figure 3.1.

Figure 3.1: Final data for the estimation

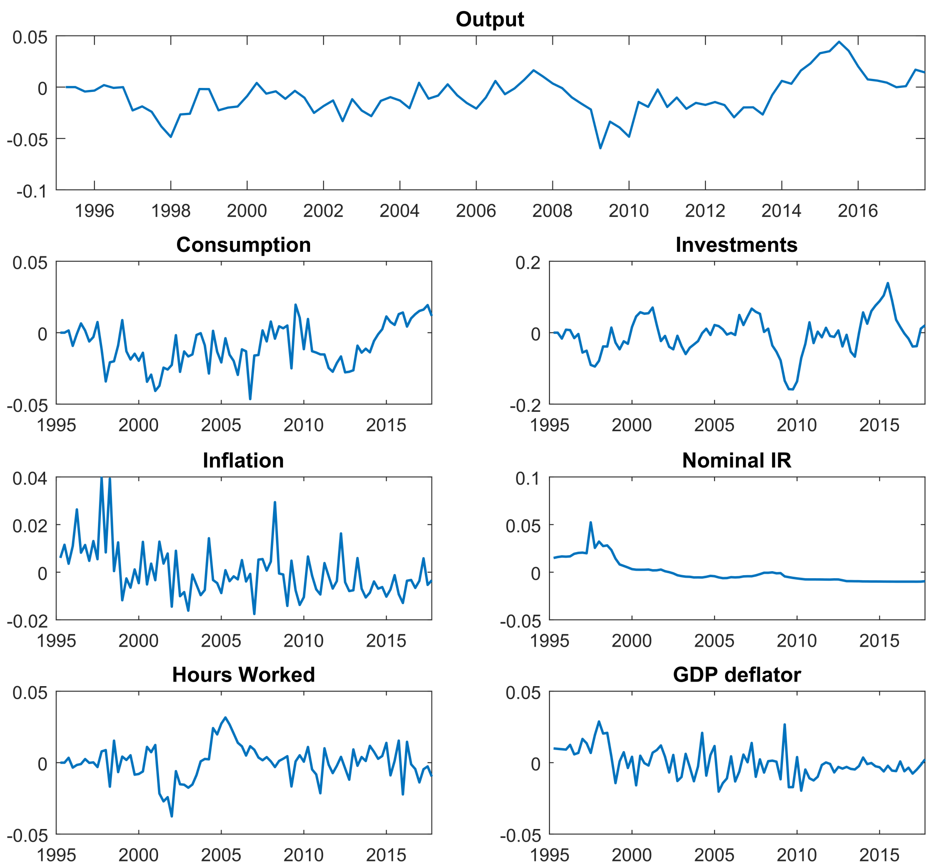


Table 3.1: Descriptive Statistics

	Mean	Std. dev.	5^{th} perc.	95^{th} perc.	min	max
Output	-0.85	1.81	-3.8	2.29	-5.95	4.42
Consumption	-0.91	1.48	-3.39	1.41	-4.65	1.96
Investments	-0.71	5.36	-9.45	7.57	-15.88	13.88
Inflation	0	1.06	-1.37	1.62	-1.75	3.97
Nominal IR	0	1.2	-0.99	2.54	-0.99	5.23
Hours Worked	0.03	1.22	-2.16	2.04	-3.76	3.16
GDP deflator	0	1.01	-1.58	2.03	-2.03	2.88

To match model properties to data, two assumptions have to be employed:

(i) **Exclusion of government from the model**: Consumption data (and therefore model consumption variable $C_t(i)$) incorporates final government consumption as well.

(ii) **Closed economy property**: Net exports are subtracted from value the of GDP and thus are completely excluded from the estimation.

To describe stationary properties of this model, let define general shock process s_t representing all model shock processes such that:

$$s_t = \rho_s s_{t-1} + \varepsilon_t^s \text{ , where } \varepsilon_t^s \sim N(0, \sigma_z^2)$$

Then, model variable $S_t = \mathrm{e}^{s_t}$ is fluctuation around its long-term, but unspecified trend. As S_t is stationary, mean-reverting and log-normal AR(1) process resulting in both specified models being stationary. Thus, model does not exhibit trend in terms of growth along balanced growth path, but describes behaviour of the economy along this path. Data transformation and specification of observation equations follows mainly Pfeifer (2013), Adjemian et al. (2011) and Griffoli (2008).

As a result of this specification, all model variables correspond to stationary, per capita values. General model variable $Z_t(i)$ therefore is the intensive form of Z_t such that $Z_t(i) = \frac{Z_t}{L_t}$, where L_t denotes overall labour force. At first, *Group 1* variables are expressed in per capita terms, when as denominator is in line with $Z_t(i)$ used number of employed, i.e. overall labour force L_t^{data}. This is at first seasonally adjusted and then cyclical component is filtered by two-sided HP filter, to prevent distractions of transformed variables due to statistical revisions regarding population statistics or measurement errors. All transformations performed on each variable in *Group 1* are demonstrated on GDP. Denote Y_t^{data} empirical GDP variable, then per capita value $Y_t^{pc,data}$ is:

$$Y_t^{pc,data} = \frac{Y_t^{data}}{L_t^{data}}$$

Subsequently, empirical gross GDP deflator $\upsilon_t^{p,data}$ with model counterpart of υ_t^p measuring price dispersion and resulting distortion is computed:

$$\upsilon_t^{p,data} = \frac{\Upsilon_t^{p,data}}{\Upsilon_{t-1}^{p,data}}$$

where $\Upsilon_t^{p,data}$ in basis deflator index. This is then used to obtain:

$$y_t^{data} = \frac{Y_t^{pc,data}}{\upsilon_t^{p,data}}$$

with y_t^{data} denoting empirically observed real per capita output. Up to this point, data transformations are the same for both *Dynare* implementation of this model.

Non-linear Model

In strict non-linear specification, model variable $Z_t(i)$ is the level of per capita value in real measures. To link data to the model 3 options are available:

1. use first difference filter and estimate level model variables: implemented
2. define observation equations as deviations from steady state and keep model

variables on level: technically possible, but non-consistent implementation. Interpretation of several model features is different.

3. use *loglinear* option is *estimation* command of the *.mod* file to let *Dynare* log-linearize model using Jacobian transformation: this requires strictly non-negative steady state to perform Jacobian transformation. In terms of estimation, this is the same as finally implemented model in Non-linear form for Log-linearization, but differs in interpretation as Impulse-response functions (IRFs) can no longer be interpreted as elasticities.

When first difference filter is applied, model relevant variable y_t^{obs} given in terms of demeaned growth rates is:

$$y_t^{obs} = log\left(\frac{Y_t^{data}}{Y_{t-1}^{data}}\right) - log\left(\overline{\frac{Y_t^{data}}{Y_{t-1}^{data}}}\right)$$

Specification for other variables is very simple:

$$\Pi_t^{obs} = \Pi_t^{data}$$

$$\upsilon_t^{p,obs} = \upsilon_t^{p,data}$$

$$R_t^{obs} = 1 + \frac{R_t^{data}}{4 \times 100}$$

Idea behind denominator for R_t^{obs} is to convert annual interest rate to quarterly ones. Before writing down observation equations (see Listing 1), note that model variables are on level and real per capita values. Demeaned growth rates for trending variables with added measurement error are:

$$y_t^{obs} = log(Y_t(i)) - log(Y_{t-1}(i)) + \varepsilon_t^y \text{ , where } \varepsilon_t^y \sim N(0, \sigma_y^2)$$

For non-trending variables it is more straightforward:

$$\Pi_t = \Pi_t^{obs} + \varepsilon_t^{\Pi}$$

$$\upsilon_t^p = \upsilon_t^{p,obs} + \varepsilon_t^{\upsilon^p}$$

$$R_t = R_t^{obs} + \varepsilon_t^R$$

Incorporation of measurement error is intended solve lack of shocks included in the model. Problem is that non-adequately tightness of model variables to observable data results in poor identification of parameters as any model is by nature misspecified. Added measurement errors allows for model variables to decouple from their empirical counterparts devoid of ruining the model. Shock decomposition analysis proves that measurement errors are especially useful when major shock not considered in the model hits the economy, such as monetary crisis of 90's and dramatic increase of nominal IR to prevent speculations.

Listing 3.1: observation equation for non-linear model

```
% Observation equations
y_obs=log(y)-log(y(-1))+e_y;
c_obs=log(c)-log(c(-1))+e_C;
l_obs=log(l)-log(l(-1))+e_l;
x_obs=log(x)-log(x(-1))+e_x;
PI_obs=PI+e_PI;
vp_obs=vp+e_vp;
R_obs=R+e_R;
```

Despite initial intention to implement this type of model for estimation, its properties are implausible:

1. After application of first difference filter, empirical data did not exhibit desired properties. Data contained very little autocorrelation and too extensive fluctuations, although typical for this filter. Usual data definition for empirical work when applying first differences is to compute quarter-on-quarter (or generally period-on-period) in opposed to period-on-period differences corresponding to model structure of DSGE.

2. simulated artificial time series did not nearly match moments of empirical data

3. filtered time series are too noisy and non-informative, translating to unprecedentedly high measurement errors incorporated in observation equations. Consequently, it is difficult to find mode of posterior density due to often flat surface of likelihood function and identification of parameters is rather unsuccessful despite inclusion of desired data

As a result, alternative solution is implemented.

Non-linear Model for Log-linearization

For this model type, data transformation and writing down the observation equations is more structured process. To obtain model relevant variable y_t^{obs}, one-sided HP filter as in Stock and Watson (1999) is applied to $log(y_t^{data})$. Logarithmic transformation persists scales of percentage changes for different level of data. Advantage of causal one-sided HP filter over its commonly used non-causal two-sided alternative comes from the state-space setup of the model Pfeifer(2013). Solution of the model is backward-looking system depending on current and past states. Filtration in two-sided HP filter is based not only on past and current, but also on future values, contradicting set up of the model. Moreover, non-causal filters are inappropriate in any forecasting analysis, as only past values of filtered variables should enter the estimation.

Empirical gross inflation Π_t^{data} is computed as a ratio of price level in two subsequent period:

$$\Pi_t^{data} = \frac{P_t^{data}}{P_{t-1^{data}}}$$

observational equivalent is then:

$$\Pi_t^{obs} = log\Big(\upsilon_t^{p,data}\Big) - log\Big(\bar{\upsilon}_t^{data}\Big)$$

and for GDP deflator:

$$\upsilon_t^{p,obs} = log\Big(\Pi_t^{data}\Big) - log\Big(\bar{\Pi}_t^{data}\Big)$$

This specification assumes that historical (long-run) well represents equilibrium (steady state) inflation rate and deflator. For gross nominal interest rate there are two alternatives. First is the same as for inflation and deflator and that is to consider historical mean to beaing equilibrium rate. Then:

$$R_t^{obs} = log\left(1 + \frac{R_t^{data}}{4 \times 100}\right) - log\overline{\left(1 + \frac{R_t^{data}}{4 \times 100}\right)}$$

All observed variables derived above are measure is deviation from their steady state and have (asymptotically) zero-mean. This setting also makes prediction and forecasting independent of empirical mean, model equilibrium target gross inflation $\bar{\Pi}$ and optimal gross nominal interest rate $\bar{R}$ determined by model parameters. Next, specification of observation equations to match transformed observable variables $y_t^{obs}, c_t^{obs}, x_t^{obs}, n_t^{obs}, \Pi_t^{obs}, R_t^{obs}, \upsilon_t^{obs}$ is required. Recall general model variable $Z_t(i)$, which is in this model specification transformed to model variable $z = log(Z_t(i))$. Observation equation has to be specified in such way that $\hat{z}_t = z_t^{obs}$ corresponds to deviations of $Z_t(i)$ from its steady state $Z_t^{ss}(i)$. This is acquired by specifying:

$$\hat{z}_t = z_t - z_t^{ss}$$

which is in this model implementation observational equation of each above stated variable. Final measurement equation is then obtained by equalizing $\hat{z}_t = z_t^{obs}$ and adding measurement error $\varepsilon_t^z \sim N(0, \sigma_z^2)$:

$$z_t^{obs} = z_t - z_t^{ss} + \varepsilon_t^z \tag{3.105}$$

Exact specification is in Listing 1.

Listing 3.2: Observation equation for non-linear model for log-linearization

```
% Observation equations
y_obs=yd$ $steady_state(yd)+e_y;
c_obs=c$ $steady_state(c)+e_C;
l_obs=l$ $steady_state(l)+e_l;
x_obs=x$ $steady_state(x)+e_x;
PI_obs=PI$ $steady_state(PI)+e_PI;
vp_obs=vp$ $steady_state(vp)+e_vp;
R_obs=R$ $steady_state(R)+e_R;
```

3.8.2 Priors

Total number of parameters included in the model counts to 35. Out of these 34+2 are included in the estimation, with +2 estimatescorrelations between shocks. It is common to exclude parameters from the estimation if they are subject to e.g. government decision such as tax rates or Taylor rule weights or when they are sufficiently identified by other micro or macro study. As this model does not really have desired properties (discussed below), it was decided to estimate model as the whole and let it find its parameters to describe empirical data and potentially maximize forecasting performance. Consequently, μ is the only parameter excluded from the estimation. It is (i) due to extremely poor identification, as model lacks financial frictions features, money growth equation or any link between inflation and money growth (or money demand), (i) unavailability of empirical data on money demand and (iii) unimportance of μ on results.

Setting of prior follows semi-conservative approach, when mean of the priors is subject to calibration, moments of empirical data, or based on macro and microeconomic theory, but, overall, set relatively loose with high standard deviations. This allows them to vary across subspace of potential values to extract maximum information contained in the data. Priors selection follows standards of DSGE literature with inverse Gamma distribution with infinite standard deviation for shocks and Beta distribution of parameters restricted on interval $[0,1]$. Likewise, Gamma distribution is used to tighten posterior mean closer to prior mean and Normal distribution when no prior information about parameters are known[3].

3.8.3 Estimation results and Convergence

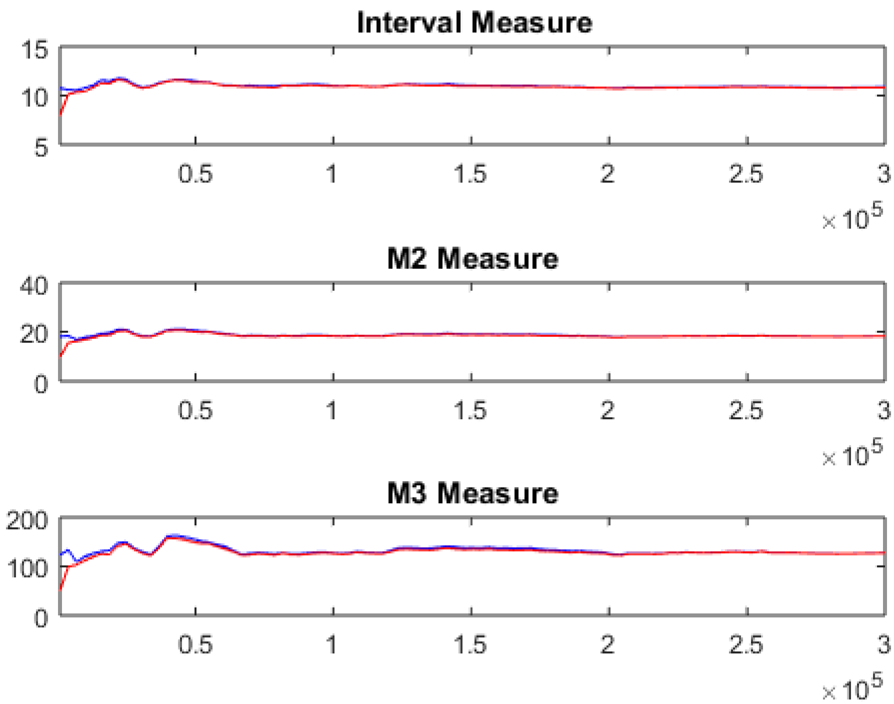
Estimation was conducted using full sample of data from 1995:Q1 to 2017:Q3 providing 91 observations in total. Model was estimated in *Dynare* by Metropolis-Hastings algorithm employing 6 MH blocks with 300 000 draws in each block and burn ratio of 0.6. Convergence of MH sampler is judged upon statistics and plots directly produced by *Dynare* on univariate and multivariate basis. Univariate describes convergence of each individual parameter being estimated, while multivariate is based on the range of posterior likelihood. Making things short, MH chains have converged if both lines are stabilized horizontally and close to each other[4]. Plots for univariate convergence diagnostics are not reported due to extensiveness, but all estimated parameters exhibit properties of successful convergence.

Estimation results reported in Tables 3.2 - 3.4 contain information on mean and standard deviations for prior and extra 5th and 95th percentiles for the posterior distributions. Results on parameters listed in Table 3.2 are reasonable. Only concern might be stated to capital share of output parameter α, which value of nearly 0.7 might appear unprecedentedly high when compared to other estimation. On the other hand, it is perfectly in line with endogenous growth theories. Depreciation parameter δ imply annual depreciation rate around 0.12, which is also consistent with theory and other estimations. Discount factor $\beta \approx 0.97$ corresponds to equilibrium net nominal interest rate of approx. 5% with inflation target being 2%. Labour dis-utility parameter, elasticities of substitution between labour and good varieties or shock persistence are similar to models estimated for European economies, Smets

[3]Typical incorporation of uniform distribution failed to find posterior mode

[4]For more details, see e.g. Pfeifer (2014) or Brooks and Gelman (1998)

Figure 3.2: Multivariate Convergence Diagnostics



and Wouters (2007) for entire Euro Area, Burriel, Fernández-Villaverde and Rubio-Ramírez (2010) for Spain or Hristov (2016) for Germany and G3 model of Czech National Bank. Problematic part are rigidities (Calvo) parameters θ, θ_w and partially indexation ones κ, κ_w, which are estimated on considerably lower level then in other studies. Possible explanation is in model structure, when capital is perfectly adjustable and is no subject to neither installation costs nor capacity utilization. Investments and capital stock therefore serve as easy channels for accommodation on shocks devoid of putting much pressure on wages and prices. Second, it might be relatively high price stability in Czech republic accompanied with traditionally strong dislike of inflation represented also by high value of γ_Π parameter of Taylor rule. Possible cause of extremely low weight on interest rate R_t smoothing parameter γ_R can be past setting of interest rate not following domestic economic fundamentals, but rather currency stability during monetary crises in 90's.

Table 3.2: Estimated Parameters

	Prior			Posterior			
	Dist.	Mean	Std. dev.	Mean	Std. dev.	HPD[5] 95%	HPD 5%
α	Beta	0.600	0.1000	0.695	0.0806	0.5692	0.8293
β	Beta	0.950	0.0200	0.968	0.0108	0.9509	0.9855
h	Beta	0.700	0.1500	0.869	0.0600	0.7892	0.9538
σ	Gamma	1.000	0.2000	1.121	0.2098	0.7772	1.4619
η	Gamma	3.500	1.5000	4.606	1.3724	2.3832	6.7323
ψ	Normal	3.000	1.0000	3.004	1.0033	1.3274	4.6408
ϵ_l	Gamma	8.000	2.0000	6.229	1.6478	3.4843	8.8111
ϵ	Gamma	5.000	2.0000	7.234	2.7321	2.9555	11.3459
δ	Beta	0.050	0.0400	0.039	0.0143	0.0161	0.0613
θ	Beta	0.500	0.1000	0.374	0.0655	0.2649	0.4788
θ_w	Beta	0.500	0.1000	0.259	0.0713	0.1413	0.3735
κ	Beta	0.400	0.1500	0.183	0.1135	0.0245	0.3448
κ_w	Beta	0.500	0.1500	0.379	0.1360	0.1588	0.6006
γ_y	Beta	0.300	0.1500	0.368	0.1257	0.1608	0.5746
γ_{Π}	Gamma	1.500	0.2000	2.025	0.1803	1.7223	2.3101
γ_R	Beta	0.500	0.1500	0.144	0.0575	0.0510	0.2335
ρ_a	Beta	0.800	0.0500	0.893	0.0290	0.8468	0.9403
ρ_b	Beta	0.700	0.1000	0.883	0.0279	0.8378	0.9283
ρ_c	Beta	0.700	0.1000	0.703	0.0895	0.5610	0.8500
ρ_l	Beta	0.700	0.1000	0.755	0.0813	0.6263	0.8851
ρ_e	Beta	0.400	0.0500	0.385	0.0499	0.3033	0.4677

Table 3.3: Estimated standard deviation of Structural Shocks

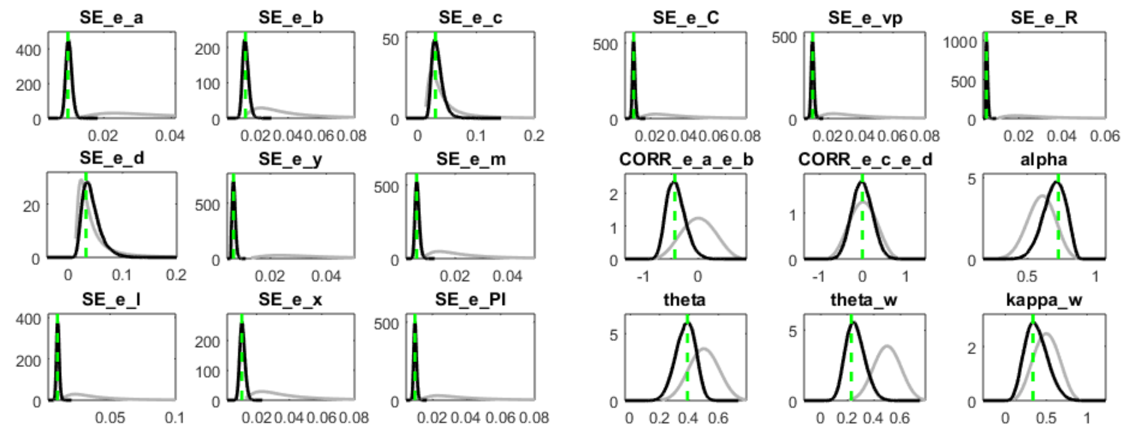
	Prior			Posterior			
	Dist.	Mean	Std. dev.	Mean	Std. dev.	HPD 95%	HPD 5%
ε^a	Inverse Gamma	0.050	Inf	0.010	0.0009	0.0081	0.0110
ε^b	Inverse Gamma	0.050	Inf	0.014	0.0019	0.0107	0.0166
ε^c	Inverse Gamma	0.050	Inf	0.033	0.0097	0.0188	0.0473
ε^l	Inverse Gamma	0.050	Inf	0.042	0.0168	0.0170	0.0665
ε^r	Inverse Gamma	0.030	Inf	0.006	0.0008	0.0043	0.0068
ε^y	Inverse Gamma	0.050	Inf	0.007	0.0006	0.0060	0.0077
ε^C	Inverse Gamma	0.050	Inf	0.009	0.0008	0.0081	0.0106
ε^x	Inverse Gamma	0.050	Inf	0.011	0.0016	0.0087	0.0138
ε^Π	Inverse Gamma	0.050	Inf	0.010	0.0008	0.0084	0.0110
ε^n	Inverse Gamma	0.050	Inf	0.010	0.0010	0.0082	0.0116
ε^R	Inverse Gamma	0.040	Inf	0.005	0.0004	0.0047	0.0059
ε^{υ^p}	Inverse Gamma	0.050	Inf	0.010	0.0008	0.0090	0.0118

Table 3.4: Estimated correlation of Structural Shocks

	Prior			Posterior			
	Dist.	Mean	Std. dev.	Mean	Std. dev.	HPD 95%	HPD 5%
$(\varepsilon^a, \varepsilon^b)$	Beta	0.000	0.3000	-0.413	0.1691	-0.6889	-0.1450
$(\varepsilon^c, \varepsilon^l)$	Beta	0.000	0.3000	-0.022	0.2303	-0.3984	0.3578

Correlation between structural shock was estimated in order to get insight about fundamentals affecting the economy. Pairs are selected according to the sector they come from. First estimated correlation between TFP and investment shock is surprising, as RBC theory suggests that positive TFP shock due to intertemporal substitution effect rise both hours worked and investments as both factors of production become more productive, thus positive correlation. Negative one appear more like an "income" effect, when higher overall productivity requires less capital stock, and thus investments, to persist current level of output. Transformation process conducted mainly in 90s-00s but still not finished for Czech economy provides explanation: it is a negative TFP shock hitting economy for long period of time, but at the same time, this period is characterized by high level of private investments required to re-build economy structure and start new private businesses. This interpretation is supported also by the Shock decompositions analysis, when supply shocks remain negative for long period of time[6]. For the second group, estimated results are considered as insignificant. Plots 3.3 - 3.4 illustrate Tables 3.2 - 3.4, with grey being prior, black posterior and with dashed green line denoting mode of the posterior.

Figure 3.3: Prior and Posterior Distributions (1)

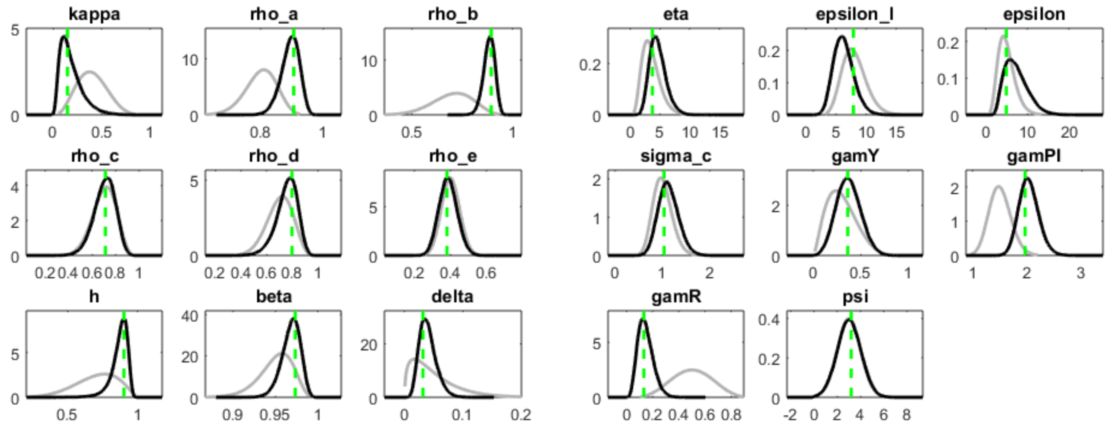


3.9 Model Properties

Model properties are evaluated by (i) IRFs calculated during MH sampling utilizing Cholesky decomposition and representing mean response to shocks (not response at the mean) and (ii) historical shock decompositions. Former shows impact of each shock on observables with 5 and 95% confidence bands, i.e. 90% Highest probability

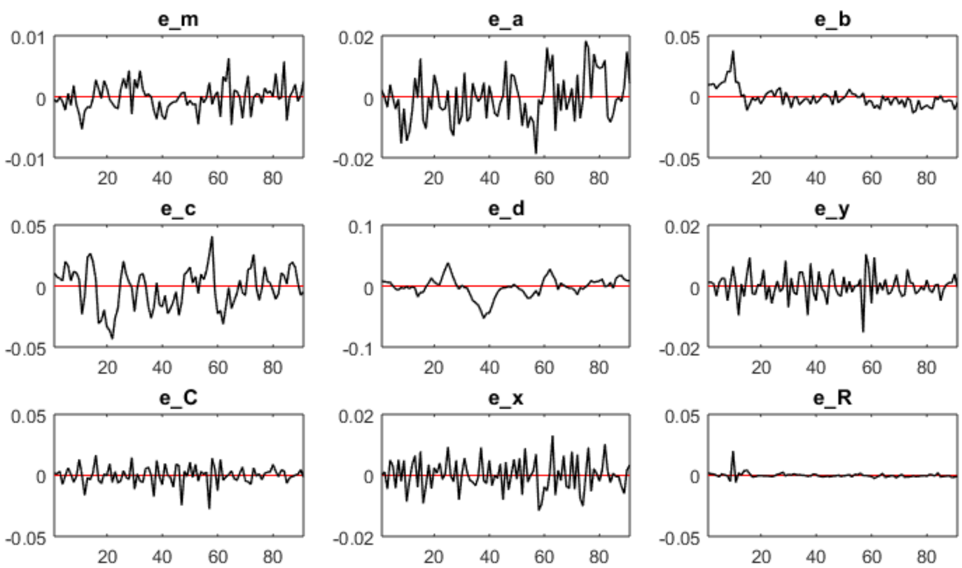
[6]On separately conducted analysis with individual shcks investment-specific shock accounted for majority of growth in that period

Figure 3.4: Prior and Posterior Distributions (2)



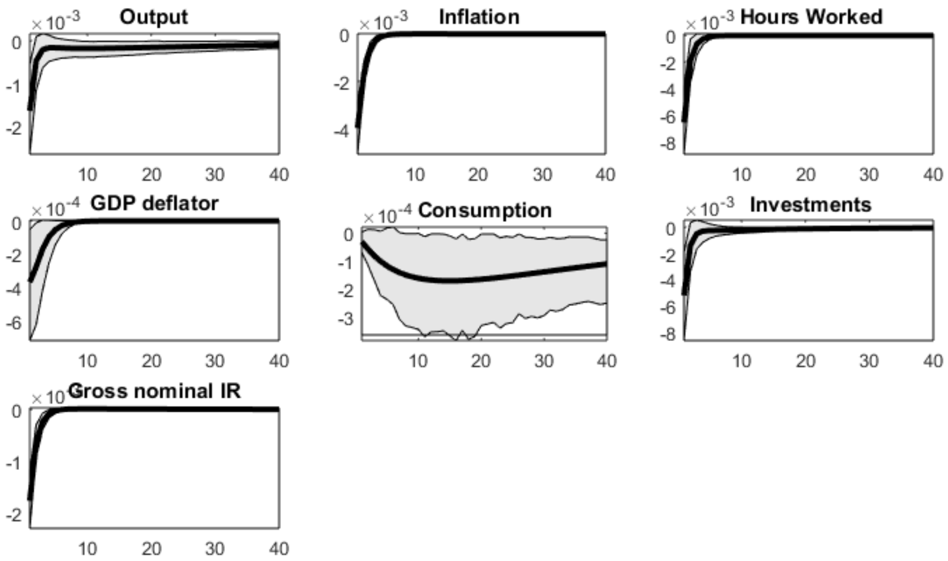
density interval of their posterior distribution, while latter decompose historical fluctuations around their steady state into individual shocks. Entire analysis is performed on empirically observable time series expressed in real per capita values linked to the model through observation equations. Model is then perfectly capable of estimating the impact of structural shocks upon these time series of main interest. Figure 3.5 shows evolution of smoothed structural shock as estimated by the model.

Figure 3.5: Smoothed past realization of shocks



IRFs are depended not only relationships between variables, but also on estimated

Figure 3.6: Response to Monetary shock



magnitude of each shock. Therefore note that responses to monetary shock ε_m (Figure 3.6) are of such negligible magnitude due to low estimated standard deviation for this shock. This is perfectly illustrated by shock decomposition for Nominal interest rate, as its movement is not primarily influenced by monetary, but supply shocks. Analysis shows that response to monetary shock is rather negative, as nominal interest rate R_t is set above its equilibrium value $\bar{R}_t$ implied by the model, and relatively quickly fading.

Response to TFP shock (Figure 3.7) is standard for RBC model: increasing output, investment, labour supply and steadily consumption while decreasing inflation (and GDP deflator). Lower nominal IR is in line with RBC theory, but its impact on IR is uncertain, mainly due to negative correlation of TFP with investment shocks. Usual response of working hours (or employment) is NKE models with rigid labour market, is the decrease of supplied labour services, as with better technology less work is required to produce the same output. However, as mentioned in Estimation section, this model contains rigid labour market, but also perfectly flexible capital formation that accommodates TFP shock very quickly, which in turn shifts labour demand up as in RBC models.

Figure 3.7: Response to TFP shock

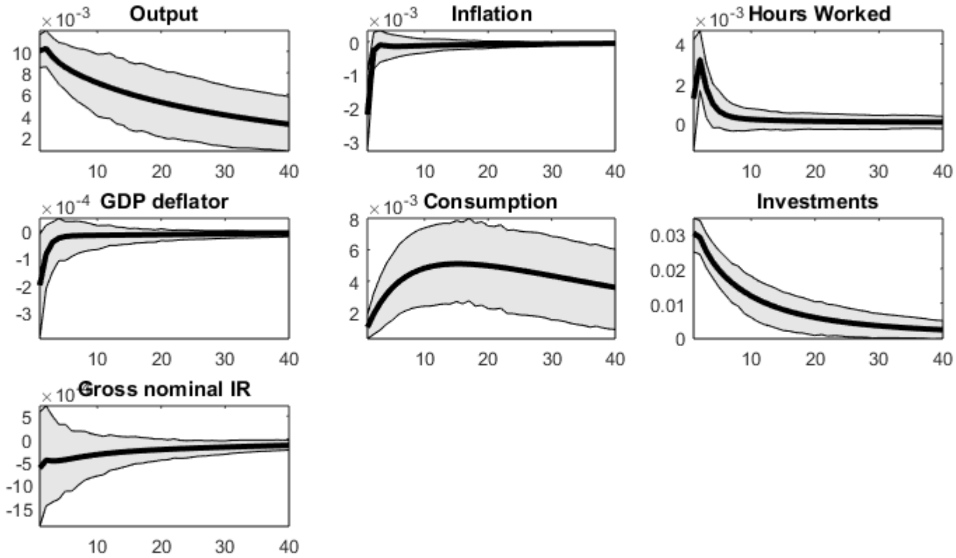


Figure 3.8: Response to investment shock

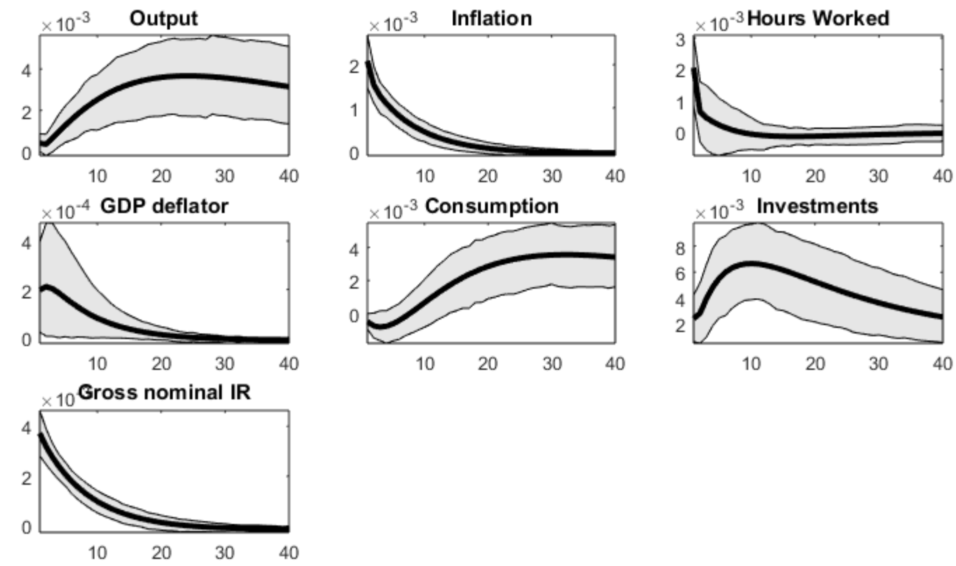


Figure 3.8 reports responses to investment specific shock. Heavy rise at impact for investments is rather expected and rising impact on consumption is also property of RBC model. Note that for later periods, slowly diminishing dynamics on investments

are substituted by stronger growth in consumption keeping growth of output relatively constant. Steep response of nominal interest rate is natural answer to increased demand for loanable funds.

Figure 3.9: Response to consumption preference shock

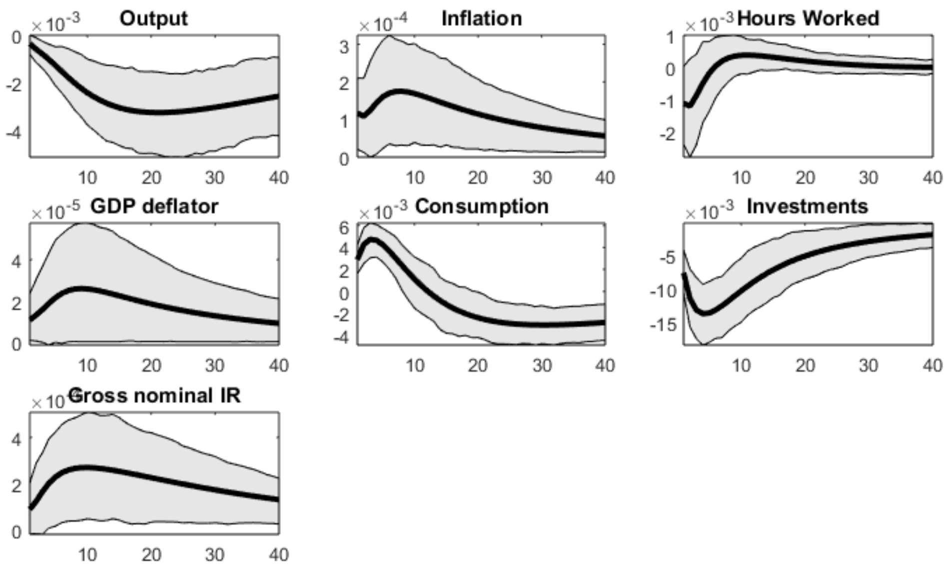
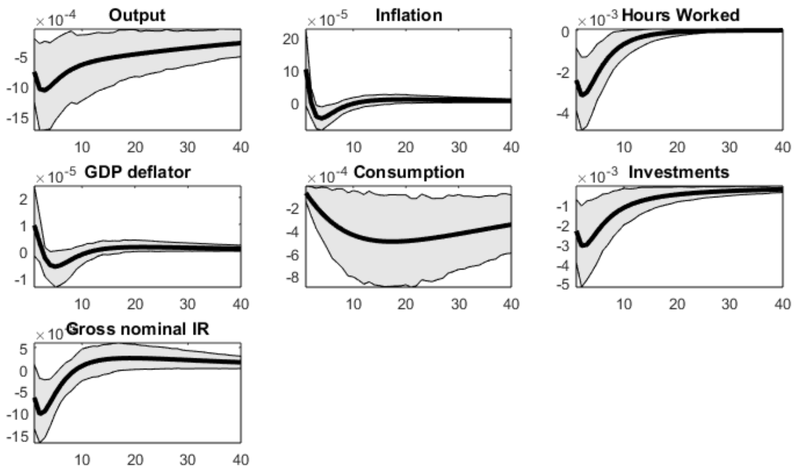


Figure 3.9 describes responses on consumption preference shock. As expected, Consumption shock is accompanied by steep growth of current consumption at the expense of future one, as investments fall down even more rapidly. Moreover, this is accompanied by decrease of employment (as lower capital stock requires less labour input in production) and thus decreasing output is just a logical outcome. Note, that in this model real consumption per capita also includes government consumption. From this point of view, higher government expenditures and support of private consumption is rather contra productive for stabilization of output.

At last, Figure 3.10 shows responses to shock into labour supply. It is characterized by substantial decrease on impact in provided labour services and this shock translates to entire economy, when worsening every indicator of economic prosperity.

For the purposes of the variance decomposition analysis, shocks are grouped into categories: (i) Supply shock consists of TFP, investment and labour supply, (ii)

Figure 3.10: Response to labour supply shock

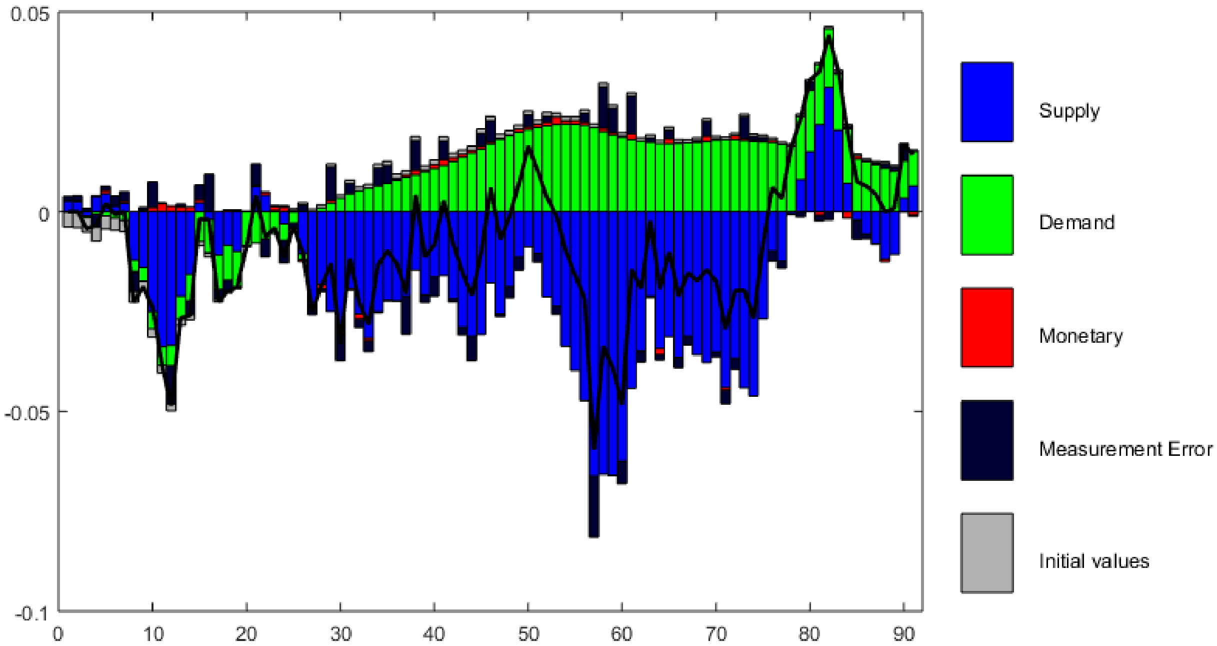


Demand being consumption preference shock, (iii) Monetary (Taylor rule shock) and (iv) Measurement errors as specified in section 3.8.1. At this place is illustrated only decomposition of output and rest is to be found in Appendix B.

Figure 3.11 shows decomposed fluctuations of output around its equilibrium value as being hit by specified shocks. Entire shock decomposition analysis for all observable time series is characterized by supply and demand sides effects opposing each other. As model is by its natural structure not capable of explaining monetary part of the economy, dynamics of nominal interest rate is identified as being subject to supply, particularly investment shocks. Initial strong wave of negative supply shocks is considered being caused by transformation and privatization process in 90's, that translates into 00s' as well. Thereafter, Czech economy was hit by economic downturn in early 2000's and then by global financial crisis. Subsequent recovery was thanks to finally positive TFP shocks. This decomposition shows that Czech economic growth was spurred by demand side, i.e. private and government consumption. Assigning extensive amount of negative impact on TFP shocks might be also caused by filtration technique: one-sided HP filter that demean time series only asymptotically (Stock and Watson, 1990). When applied to Czech data, considerable negative mean was still present after filtration (see Table 3.1).

As for inflation, model is unable to well explain steep sudden changes opposing each other from period to period. Vast majority os this movements is assigned to

Figure 3.11: Shock decomposition - Output



supply shock, while only a fraction is identified as a monetary or other shock. On top of this, degree of measurement errors, representing unexplained part of the model, well define model inability to explain inflation. As is the matter of DSGE, inflation is closely tighten to its respective equilibrium value, in the model specified as 2% as is the mean inflation target of Czech National bank.

3.10 Forecasting

Typical approach to determine forecasting ability of any model is to conduct out-of-sample forecast k-steps ahead employing expanding window. In this procedure, data are split into two subsamples (e.g. 40:60), when initial model is fit only on first period, corresponding forecast is conducted k-steps ahead and one observation is repeatedly added into in-sample period, model is re-estimated and new forecast is computed. In case of DSGE is such procedure cumberstone:

1. Estimation of model of such complexity and scale by Bayesian methods is extremely demanding on computational power. To obtain desired properties of estimated model, i.e. to find true distributions of estimated parameters and reach convergence, requires extensive amount of MH draws and blocks[7].

2. Each iteration of expanding window triggers entire estimation of DSGE from mode finding, MH sampling to posterior draws. As mentioned before, finding mode of the posterior is often non-trivial task for solution algorithms due to flat surfaces of likelihood function. As data changes over time and so does the priors, repeating estimation would requires either try-and-fail setting of new priors in each period or employment of more robust, but inefficient optimization methods such as MCMC algorithm.

To sum up, full out-of-sample forecast would consume full author's computational power for more than a month. Instead, forecasting performance in case of DSGE model is evaluated from k-step ahead filtered variables. This is a combination of in-sample and out-of-sample forecast such that model is initially estimated on entire data set, and filtered variables are obtained taking into account only data known at time T. Formally, general model filtered variable z k-step ahead is given by its expectation with respect to time T:

$$\hat{z}_{T+k} = \mathbb{E}_T z_{T+k}$$

where T denotes initial period of filtration/pseudo-forecast. This procedure is sometimes referred to as an in-sample back-testing to test for model stability over time. Forecasting performances of all models developed in this thesis are reported and summed in section 5.

[7] Only one estimation for this model took approx. 20 hours

3.11 Model Implementation

Above derived model is solved and estimated using *Dynare*. It is a software platform utilizing Matlab in-build functions to solve rational expectations DSGE models, i.e. agent's expectations about future are model consistent. *Dynare* works as a translator between Matlab *.m* functions and *.mat* files and *Dynare's .mod* files, which the model is written in. Although presented DSGE model was not explicitly log-linearized, but was entered into *Dynare* in non-linear form on levels, it. is still possible to write the model in such a way that *Dynare* log-linearize it itself. To fasten the estimation, it is recommended to write external *_steady_state.m* file that analytically computes steady state, as $60 - 70\%$ of estimation time is spent on finding steady state by simulation algorithms. From complete *.mod* files with corresponding *_steady_state.m* files for both versions of implementation of the model solved and estimated in *Matlab* see the Appendix A.

4 Bayesian Threshold VAR

4.1 Bayesian approach to linear models

In this section, Bayesian VAR and Threshold VAR models challenging performance of previously derived DSGE model are developed. Description of estimation of Bayesian linear models can be found in any textbook on Bayesian econometrics, e.g. in Koop (2003), Geweke (2005), Kilian and Lutkepohl (2016) or in comprehensive guide of Blake and Mumtaz (2012). Bayesian estimation of linear models, as usual, consists of (i) formation of prior beliefs about parameters to be estimated in the form of prior distribution, (ii) writing the likelihood function of the model representing information in the data and (iii) continual updates of prior beliefs based on the information contained in the data, i.e. combining prior distribution with likelihood function. Purpose of this procedure is the same as in DSGE estimation: to form posterior distribution $H(B, \sigma^2 \mid Y_t)$, which is the simple product of the prior $P(B, \sigma^2)$ and likelihood $F(Y_t \mid B, \sigma^2)$ scaled by $F(Y)$ representing density of the data or marginal likelihood, and is defined by the Bayes Law as:

$$H(B, \sigma^2 \mid Y_t) = \frac{F(Y_t \mid B, \sigma^2) \times P(B, \sigma^2)}{F(Y)} \tag{4.1}$$

Before turning to multivariate and threshold models, suppose empirically relevant univariate case when vector of coefficients B and variance σ are unknown. Conjugate prior for vector of parameters β is normally distributed prior $P(B) \sim N(B_0, \Sigma_0)$ and for σ^2 it is inverse Gamma distribution, or alternatively, Gamma distribution for $\frac{1}{\sigma^2}$. Definition of conjugate prior is that when combined with the likelihood function (normal PDF in case of classical linear models), resulting posterior distribution is the same as prior and thus is known. Formally, joint prior density of B and σ is:

$$p\left(B, \frac{1}{\sigma^2}\right) = P\left(\frac{1}{\sigma^2}\right) \times P\left(B \mid \frac{1}{\sigma^2}\right) \tag{4.2}$$

Resulting conditional posterior distribution $H(B, \frac{1}{\sigma^2} \mid Y_t)$ is function of both B and $\frac{1}{\sigma^2}$, but to conduct inference about individual parameters B and σ^2, marginal posterior distribution for each parameter is necessary. Joint and marginal distributions can be approximated by Gibbs sampling algorithm that utilizes draws from conditional posterior distribution available for each set of parameters. It si a Markov Chain Monte Carlo algorithm, implying that random draws are independent of past ones and it can be seen as an special case of Metropolis-Hastings algorithm (Geweke; 2005).

4.2 Bayesian VAR

Suppose a standard VAR(p) model of the following form:

$$Y_t = c + B_1 Y_{t-1} + ... + B_p Y_{t-p} + u_t \tag{4.3}$$

which can be in companion form written as

$$Y_t = X_t B + u_t \tag{4.4}$$

where $X = \{c_i, Y_{it-1} ..., Y_{it-p}\}$. This can be further rewritten in vectorized form utilizing Kronecker's product:

$$y = (I_N \otimes X)b + U \tag{4.5}$$

with $y = vec(Y_t)$, $b = vec(B)$, $U = vec(u)$.

Prior distribution of VAR coefficient is assumed to be normal, with $\tilde{b}_0$ denoting prior mean and where H is matrix with diagonal elements specifying variance of the prior distribution:

$$p(b) \sim N(\tilde{b}_0, H) \tag{4.6}$$

In turn, due to property of conjugate prior, posterior distribution of these coefficients is also normal and conditional on Σ

$$H(b \,|\, \Sigma, Y_t) \sim (M^*, V^*) \tag{4.7}$$

where

$$M^* = (H^{-1} + \Sigma^{-1} \otimes X_t' X_t)^{-1}(H^{-1}\tilde{b_0} + \Sigma^{-1} \otimes X_t' X_t \hat{b_0}) \tag{4.8}$$

$$V^* = (H^{-1} + \Sigma^{-1} \otimes X_t' X_t)^{-1} \tag{4.9}$$

where $\hat{b}$ denotes vectorized format of OLS estimates of VAR coefficients $\hat{b} = vec((X_t' X_t)^{-1} X_t' Y_t))$. Note that the mean of this posterior distribution is simply a weighted average of prior mean $\tilde{b}$ and OLS estimate $\hat{b}$. Weights are given by the inverse of corresponding variance, i.e. by H^{-1} in case of prior $\tilde{b}$ and by $\Sigma^{-1} \otimes X_t' X_t$ for OLS estimate $\hat{b}$.

For the VAR covariance matrix, conjugate prior is given by the inverse Wishart distribution, multivariate version of inverse Gamma, with scale matrix $\bar{S}$ and degrees of freedom $\alpha = N + 1$. Posterior for Σ is thus again inverse Wishart

$$H(\Sigma | b, Y_t) \sim IW(\bar{\Sigma}, T + \alpha) \tag{4.10}$$

where T is the length of sample and $\bar{\Sigma}$ is updated covariance matrix $\bar{\Sigma} = \bar{S} + (Y_t - X_t B)'(Y_t - X_t B)$.

To find posterior distribution, Gibbs Sampling algorithm is employed. For estimation of VAR models Gibbs Sampling consists of:

1. Priors for VAR coefficient and covariance matrix are set.

2. Sampling of VAR coefficients is performed from conditional posterior

distribution $H(b\,|\,\Sigma, Y_t) \sim N(M^*, V^*)$ in two steps: First M^* and V^* are calculated

$$M^* = (H^{-1} + \Sigma^{-1} \otimes X_t' X_t)^{-1}(H^{-1}\tilde{b}_0 + \Sigma^{-1} \otimes X_t' X_t \hat{b}_0) \tag{4.11}$$

$$V^* = (H^{-1} + \Sigma^{-1} \otimes X_t' X_t)^{-1} \tag{4.12}$$

Second, VAR coefficients $\bar{b}$ are drawn from the standard normal distribution $\bar{b} \sim N(0,1)$ and by adding mean M^* and multiplication by squared root of variance matrix[1] $(V^*)^{\frac{1}{2}}$ as in (4.13), draw $\bar{b}$ is transformed to $b^1 \sim N(M^*, V^*)$ and first set of estimated parameters is obtained.

$$b^1 = M^* + \left(\bar{b} \times (V^*)^{\frac{1}{2}}\right) \tag{4.13}$$

3. Σ is drawn from its conditional distribution $H(\Sigma|b, Y_t) \sim IW(\bar{\Sigma}, T+\alpha)$. Again, Sigma $\hat{\Sigma}$ is updated covariance matrix as in (4.10).

4. With increasing number of iterations, "conditional distributions converge to joint and marginal distribution at an exponential rate" (Blake, Mumtaz; 2012) allowing for approximation of marginal distributions by their empirical counterparts. By repetition of Steps 2 and 3 M-times, estimates of B^1 to B^M are obtained. In sake of robustness, estimates provided by first N_b iterations are burned and remaining $(N - N_b)$ estimates are used to form the empirical distribution of VAR coefficients. Point estimate of the mean subsequently used in forecasting or impulse response analysis is the simple average of $(N - N_b)$ retained draws:

$$\hat{\beta}_i = \frac{1}{N - N_b} \sum_{b=1}^{N-N_b} \beta_i^b$$

Convergence

Convergence check of Gibbs sampler can be partitioned into visual tests and formal test statistics. In the former sequence of retained draws in examined in three ways:

[1] multivariate equivalent to standard deviation

1. retained draws exhibit stationary, mean-reverting property, i.e. fluctuate randomly around a stable mean

2. computed recursive mean shows little fluctuations

3. sample of retianed draws is not serially correlated, i.e. there is no autocorrelation.

For the latter, Geweke (1991) propose formal tests to infer about convergence based on stability of the mean and efficiency of provided estimates. Idea behind stable mean test is similar to that in recursive mean plot. A sample of retained draws of any model parameter θ_i is divided into two subsamples N_1 and N_2 and corresponding simple averages M_1, M_2 and asymptotic variance $\frac{S_1(0)}{N_1}$, $\frac{S_2(0)}{N_2}$ for each subsample are computed. The asymptotically normally distributed test statistics is then:

$$Z = \frac{M_1 - M_2}{\sqrt{\frac{S_1(0)}{N_1} + \frac{S_2(0)}{N_2}}} \tag{4.14}$$

Convergence is judged upon values of Z, when low values indicate that there is no significant difference in means and thus Gibbs sampler has converged. To infer about efficiency, Geweke (1991) also introduced measure of relative numerical efficiency (RNE) that explicitly accounts for autocorrelation of retained draws from approximated posterior distribution. Test statistic is:

$$RNE = \frac{VA\hat{R}(\theta_1)}{S(0)} \tag{4.15}$$

when values close to unity mean convergence.

4.3 Threshold Model

Model developed in this thesis follows usual implementation of threshold models known as SETAR, i.e. Self-Exciting Threshold Autoregressive model. This is the special cases of TAR model and is typical by two properties: (i) switching between regimes is based on past realization of single time series $Y_{j,t-d}$ being also endogenous variable (self-exciting) and (ii) with exogenously set delay parameter d (Koop and Potter; 2000). Simple SETAR(2,p_1,p_2) model contains two regimes with possibly different degree of autoregressive properties p_1,p_2, but here always implemented as $p_1 = p_2 = p$:

$$Y_t = c_1 + B_{11}Y_{t-1} + ... + B_{1p}Y_{t-p} + u_{1t} \qquad \text{for } Y_{j,t-d} \leq Y^* \tag{4.16}$$

$$Y_t = c_2 + B_{21}Y_{t-1} + ... + B_{2p}Y_{t-p} + u_{2t} \qquad \text{for } Y_{j,t-d} > Y^*$$

where Y^* is the level of threshold. Fundamentally, with known threshold variable and delay parameters, these are two ordinary VAR models conditional on currently valid value of threshold (Tong, 1990). Threshold variable determines tandem allocation of all other variables. Exogenous setting of delay parameter reduces problem to determination of threshold valid for current data and estimation of parameters for both regimes. Employing Bayesian methodology, model is estimated utilizing Gibbs sampling with incorporated MH step to estimate threshold level Y^* as in Blake and Mumtaz (2012):

1. Set priors.
2. Split data into regimes according to threshold variable $Y_{j,t-d}$ and currently valid threshold level Y^* such that $Y_t^{(1)} = Y_t \mid Y_{j,t-d} \leq Y^*$ and $Y_t^{(2)} = Y_t \mid Y_{j,t-d} > Y^*$. In the first iteration, threshold level Y^* is set to the values of historical mean of threshold variable $Y^* = Y_{j,t-d}$.
3. Sample coefficients for both regimes following (4.11) - (4.13).
4. Incorporate extra RWMH step to sample Y^*. New value is drawn from the random walk:
$$Y^*_{new} = Y^* + \epsilon \quad \text{,where} \quad \epsilon \sim N(0, \Sigma)$$
Acceptance probability α given as ratio of likelihood functions for new and old threshold is then computed and compared to draw from uniform distribution

$u \sim U(0,1)$ such that:

$$u < \alpha \longrightarrow Y^{*}_{new}$$

$$u \geq \alpha \longrightarrow Y^{*}_{old}$$

Similarly to estimation of DSGE, scale Σ is tuned throughout the application in first 5000 iterations of the Gibbs sampler to keep acceptance rate $\alpha \in [0.2, 0.3]$ considered as being optimal. In SETAR models, threshold variable and lag parameter d are assumed to be known or chosen arbitrarily. Criteria for selection of threshold variable are: (i) sufficient volatility to ensure switching between regimes and (i) well representativeness of variable of main interest (Kwon, 2003). Here, as a threshold variable Y^{*} is selected GDP deviations with lag degree of 2: $Y^{*} = Y_{1,t-2}$. Performance of TVAR with delay parameter $d = 2$ is superior to that with $d = 1$ and ensures roughly equal data partition to both regimes. This is particularly important in out-of-sample forecasting (see Section 4.5) when initial in-sample period contains only 50 observations.

4.4 Minnesota Prior

Minnesota prior is the special case of Gaussian conjugate prior when it is assumed that data generating process for each endogenous variable in $vec(Y_t)$ is an AR(1) process or a random walk. Distinction between these two is in stationary properties of examined time series. Fundamentally, Minnesota prior is the most conservative setting of the prior distribution and states that nothing is known (Koop, 2003). It reduces the problem of prior setting into selection of hyperparameters specifying prior variance matrix (Kilian and Lutkepohl, 2016).

Recalling companion form presented in equation (4.4), elements of prior mean $\tilde{b} = vec(\tilde{B})$ equal one only for b_{ij} elements where $i = j$, and zero otherwise. Dimensions of $\tilde{b}$ are $[N \times (N \times L + 1)] \times 1$ with N denoting number of endogenous variables, L is the lag degree, and $+1$ stands for constant term. Scale matrix $\bar{S}$ for Minnesota prior is an identity matrix $\bar{S} = I_N$.

Prior variance is given by a square matrix with dimensions of $[N \times (N \times L + 1)] \times [N \times (N \times L + 1)]$ and its setting requires more structured approach. For VAR coefficients, it is controlled by hyperparameters $\lambda_1, ...\lambda_4$ and is defined by following set of relations:

$$\left(\frac{\lambda_1}{l^{\lambda_3}}\right) \text{ for } i = j \tag{4.17}$$

$$\left(\frac{\sigma_i \lambda_1 \lambda_2}{\sigma_j l^{\lambda_3}}\right) \text{ for } i \neq j$$

$$(\sigma_1 \lambda_4) \text{ for constant}$$

Parameters $\lambda_1, ...\lambda_4$ control for relative tightness of the prior in the estimation. They are also called shrinkage parameters, as they shrink variance of the prior. Namely, λ_1 and λ_2 control for the standard deviation of the prior on lags of particular endogenous variable b_{ij}, such that λ_1 influences std. deviation on own lags, i.e. if $i = j$, and λ_2 controls for std. deviation on other than own lags, i.e. if $i \neq j$. Values of λ_1, λ_2 approaching 0 "shrinks" final estimates of b_{ij} towards the mean of the prior distribution (Kilian and Lutkepohl, 2016). Note that if $\lambda_2 = 1$ there is no distinction among own and other lags. Increasing value of parameter λ_3 increases probability the coefficients on higher lags are equal to zero. Finally, λ_4 controls constant terms in such a way that if λ_4 approaches zero, constant terms are shrunk to zero as well. Purpose of variance ratio $\frac{\sigma_i}{\sigma_j}$ is to normalize possibly different scales of included endogenous variables and l stands for lag degree.

Although developed BVAR models contain 7 endogenous variables and are characterized by lag degree of $L = 1, 2, 3$, for illustration of Minnesota prior setting, let consider a VAR(2) with only four endogenous variables and constant terms, as dimension of these matrices grow exponentially[2]:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \\ y_{4t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \\ y_{4t-1} \end{pmatrix} + \begin{pmatrix} d_{11} & d_{12} & d_{13} & d_{14} \\ d_{21} & d_{22} & d_{23} & d_{24} \\ d_{31} & d_{32} & d_{33} & d_{34} \\ d_{41} & d_{42} & d_{43} & d_{44} \end{pmatrix} \begin{pmatrix} y_{1t-2} \\ y_{2t-2} \\ y_{3t-2} \\ y_{4t-2} \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}$$

As discussed above, under Minnesota prior are endogenous variables assumed to follow random walk, which implies:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \\ y_{4t} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} b^0_{11} & 0 & 0 & 0 \\ 0 & b^0_{22} & 0 & 0 \\ 0 & 0 & b^0_{33} & 0 \\ 0 & 0 & 0 & b^0_{44} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \\ y_{4t-1} \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_{1t-2} \\ y_{2t-2} \\ y_{3t-2} \\ y_{4t-2} \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}$$

Vectorized prior mean $\tilde{b}_0 = vec(B)$ is then simply:

$$\tilde{b}_0 = \begin{pmatrix} b^0_1 \\ b^0_2 \\ b^0_3 \\ b^0_4 \end{pmatrix} \tag{4.18}$$

where

$$b^0_1 = \begin{pmatrix} 0 \\ b^0_{11} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} ; b^0_2 = \begin{pmatrix} 0 \\ 0 \\ b^0_{22} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} ; b^0_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ b^0_{33} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} ; b^0_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ b^0_{44} \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

[2]Prior covariance matrix $H = [105 \times 105]$ for $L = 2$ and $H = [156 \times 156]$ for $L = 3$

Setting of variance matrix of prior distribution is demonstrated only on one variable y_{1t} as other are analogous to this and in case of 4 endogenous variables ($N = 4$), lag degree of 2 ($P = 2$), and with constant terms, it is a square matrix with dimensions of $[36 \times 36]$. With strict implementation of Minnesota prior it is a diagonal matrix with fixed elements, imposing independence among parameters (Kilian and Lutkepohl, 2016). Error variances of σ_i are obtain by OLS estimates of corresponding AR(1) processes.

$$H = \begin{pmatrix} (\sigma_1\lambda_4)^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & (\lambda_1)^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & (\frac{\sigma_1\lambda_1\lambda_2}{\sigma_2})^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (\frac{\sigma_1\lambda_1\lambda_2}{\sigma_3})^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & (\frac{\sigma_1\lambda_1\lambda_2}{\sigma_4})^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & (\frac{\lambda_1}{2^{\lambda_3}})^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & (\frac{\sigma_1\lambda_1\lambda_2}{\sigma_2 2^{\lambda_3}})^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & (\frac{\sigma_1\lambda_1\lambda_2}{\sigma_3 2^{\lambda_3}})^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & (\frac{\sigma_1\lambda_1\lambda_2}{\sigma_4 2^{\lambda_3}})^2 \end{pmatrix}$$

4.5 Forecasting

General in-sample (full-sample) forecast equation in matrix form is defined as follows:

$$y_{t+1} = c + X_t(L)B + u_{t+1} \tag{4.19}$$

where $t = 1, \ldots, T$, endogenous variable y_{t+1} is a matrix of one-sided HP-filtered or log-demeaned data as described in Section 3.8.1, $X_t(L) = \{y_t, y_{t-1}, ...y_{t-L}\}$ is matrix of the same, but up to degree of L lagged variables, B is corresponding matrix of estimated parameters and u_{t+1} is the forecast error term. This equation is estimated using Gibbs sampling algorithm and Minnesota prior.

Out-of-sample forecasts $\hat{y}_{t+1}$, are generated using expanding window method providing 1-step-ahead predictions. First protection against overfitting is estimation of model parameters only on in-sample data (Ashey, Granger and Schmalensee; 1980). Thus, full sample of T observations is split into in-sample and out-of-sample portions. In-sample portion consists of R observations, leaving $P = T - R$ observations for out-of-sample portion. To obtain first value of out-of-sample forecast, equation (4.19) is fitted on first R observations and matrix of lagged endogenous variables $X_t(L)$ is multiplied by estimated vector of parameters $\hat{B}$. Corresponding forecast error is acquired as the difference between forecasted and actual value. For $k = K = 1$, i.e. one step ahead forecasting, the next period $(t + 1)$, $R + 1$ observations are used as the in-sample portion of data to create $t = R + 2$ forecasts.

Multiple, k-step ahead forecast with $k = \{1, ...K\}$, is obtained by repeating estimation of $\hat{y}_{t+k}$ devoid of expanding in-sample portion of data. Instead, matrix $X_{t+k}(L)$ includes previous forecasts such that for $k > L$ only already forecasted values are used to form subsequent forecasts. Consider relevant case when $L = 2$, $K = 12$ and fixed in-sample portion $t = R$:

$$y_{t+1} = c + X_t(L)B + u_{t+1} \quad \text{, where} \quad X_t(2) = \{y_t, y_{t-1}\}$$

$$y_{t+2} = c + X_{t+1}(L)B + u_{t+2} \quad \text{, where} \quad X_{t+1}(2) = \{\hat{y}_{t+1}, y_t\}$$

$$\vdots$$

$$y_{t+K} = c + X_{t+K}(L)B + u_{t+K} \quad \text{, where} \quad X_{t+k}(2) = \{\hat{y}_{K-1}\hat{y}_{K-2}\}$$

with B being constant for one iteration of Gibbs sampler. Only after calculation of all k forecasts, in-sample is expanded by adding $R+1^{st}$ observation. General forecast K-steps ahead with R in-sample observations is calculated as :

$$\hat{y}_{R+k} = \mathbb{E}y_{R+k} = \hat{c} + X_{R+k}(L)\hat{B} \quad , \quad k = \{1, ...K\}$$

with corresponding forecast error:

$$\hat{u}_{R+k} = y_{R+k} - \hat{y}_{R+k} \quad , \quad k = \{1, ...K\}$$

Model forecasting performance is then assessed by RMSE:

$$RMSE = \sqrt{\frac{\left(\sum_{i=t+1}^{T} \hat{u}_{t+k}\right)^2}{R}} \quad , \quad k = \{1, ...K\} \tag{4.20}$$

Each iteration of Gibbs samples triggers stability check, when B is updated until stable properties are guaranteed. When estimating thresholds, scales of Σ in MH algorithm is tuned on first 5000 observations (after initialization on first 100 obs.) to keep acceptance rate in desired range $\alpha \in [0.2 - 0.25]$. Threshold estimation includes 20 000 iterations of Gibbs sampler and MH algorithm with 16 000 burns. Estimation of parameters B is done upon 50 000 draws and 40 000 burns. With last burn-in draw, convergence statistics are computed and when desired, additional 5000 draws are trigged, with maximum of 2-triggers per estimation[3]. Despite this number of iterations is not considered sufficient, limitations in computational power do not allow for more.

For threshold models, forecasting algorithm is:

1. Separate data into in-sample and out-of sample
2. Estimate threshold value $Y^{*} = Y_{1,R-2}$
3. Partition data into $Y^{(1)}$ and $Y^{(2)}$ for corresponding regimes
4. Perform K-step ahead forecast with $K = 1$ or $K = 12$.
5. Repeat Steps $1 - 4$ (P-R)-times

That is to say, value of threshold is estimated only on in-sample data and remains fixed for the entire forecasting period $K = \{1, 12\}$. 1-step ahead forecast is computed only for regime the system is currently in. For 12-step ahead forecast, forecasts are calculated for both regimes and all K periods. Subsequently, threshold variable Y^{*} is recursively induced and final forecast is retrieved from mean estimates of both regimes conditionally on fitted values of threshold (see Figure 5.1).

[3]Implying maximum possible amount of 60k draws and 50k burns

For estimation and subsequent forecasting, the same data as for estimation of DSGE are utilized, i.e. one-sided HP filtered or log-demeaned deviations from the steady state of respective variable. For reference of transformations see Section 3.8.1. Regarding out-of-sample forecast and dataset splitting, Clark and McCracken (2001) recommend using P/R ratio $\hat{\pi} = 0.4$, corresponding to $R = 65$ for in-sample portion and $P = T - R = 26$ fro out-of-sample. However, Bayesian methods and efficiency of Gibbs sampler are challenged by setting $R = 50$ and thus $\hat{\pi} = 0.8$. This forecasting exercise produces 41 out-of-sample forecast and includes period of major economic downturn of 2009. Therefore it is well-suitable for inferences about forecasting abilities and performance.

4.6 Implementation

As is the case for DSGE, all BVAR and BTVAR models are implemented in *Matlab*. Core of the codes is taken from Blake and Mumtaz (2012) and their Guide for Applied Bayesian econometrics. However, obtaining concrete solution required extensive corrections, manipulations and combinations of various parts and so the original code served rather as a guide and illustration of practical estimation than a 'running-version' of the code.

5 Results Summary and Comparison

To summarize, DSGE model is estimated on full sample of data and forecasting capabilities are assessed utilizing filtered variables such that $\hat{y}_{t+k} = \mathbb{E}_t y_{t+k}$. It is then confronted with SETAR(2,2,2) estimated by Bayesian methods with threshold, regime switching variable, $Y^* = Y_{1,t-2}$ being 2-periods lagged output. As a benchmark are selected plain Bayesian VAR models (BVAR(L)) with lag degree of $L = \{1, 2, 3\}$. Assessment of forecasting power is done upon RMSE statistics. Analyzed time series are those of usual interest in macroeconomic forecasting: output and inflation. For the comparative analysis, these exercises were performed:

1. DSGE
2. 1-step ahead out-of-sample forecast of output and inflation with plain BVAR(1) and BVAR(3)
3. 1-step and 12-steps ahead out-of-sample forecast of output and inflation with SETAR(2,2,2)
4. 1-step and 12-steps ahead out-of-sample forecast of output and 1-step ahead forecast of inflation with BVAR(2)

To evaluate the appropriateness of approximation of "forecasting" abilities of DSGE model by filtered variables approach, experiment consisting of short series of forecast is performed. Again, due to limitations in computational power, experiment is conducted only on five[1] subsequent period preceding the end of data sample. In this exercise, five one-step ahead forecasts from subsequent periods are compared

[1]incorporation of more periods was unsuccessful due to inability to find posterior mode. Moreover, even inclusion of 12 period would still provide too little information for inference about approximation of 12-step ahead forecast by filtered variables.

to filtered variables for the same period acquired from full sample of data. Differences in both approaches are within a fraction of 1 standard deviation and thus filtered variables are fair approximation of actual 1-step ahead forecasts. However, one should still consider uncertainty related to this approximation.

Model Comparison

As mentioned before, estimation by Bayesian methods provides full posterior distribution of estimated parameters. Mean of these distributions serves as a point estimate for parameters and their standard errors. Percentiles of these distributions produce posterior density intervals, and when combined, represents highest posterior density intervals. These can be used for hypothesis testing and inference.

Formal model comparison of two (or more) models, is done by comparison of their posterior probabilities, known as posterior odds ratio. With equal prior weights of all models, this reduces to Bayes factor, given by the ratio of marginal likelihoods, which depends only upon the prior and the likelihood. Bayes factor balances quality of the fit and extra model complexity, while accounting for misspecification of the model. Forecasting power of the model is judged upon its predictive density, which accounts for parameter uncertainty as well. This is generated forecasts are calculated.

Sufficient condition for comparison of models using marginal likelihood is the utilization of the same dataset (Koop, 2003). This imply:

1. Models do not need to be nested as there is natural correction for degrees of freedom
2. Prior distributions over models are not required to be the same.
3. It is possible to compare performance of models with different parameters
4. Given the validity of point (3), shocks (or error terms) do not have to be necessarily equivalent, as extra shocks included in the DSGE compared to BVARs are only subjects of different parametrization.

Marginal Likelihood is computed by Chib's (1995) method. He decomposes Bayes rule (4.1) through logarithmic transformation to:

$$log\ F(Y) = log\ F(Y \mid B, \Sigma) + log\, P(B, \Sigma) - log\ H(B, \Sigma \mid Y)$$

Thereafter, posterior density is factorized such that:

$$log\ F(Y) = log\ F(Y \mid B, \Sigma) + log\, P(B, \Sigma) - log\bigg(\ H(B^* \mid \Sigma^*, Y) \times H(\Sigma^* \mid Y)\bigg)$$

where * values denote posterior means. Note that $H(B^* \mid \Sigma^*, Y) \sim N(M^*, V^*)$ is as in (4.7) multivariate normal posterior distribution for B^* with M^* and V^* given by (4.8) - (4.9) and thus known. Last term si:

$$H(\Sigma^* \mid Y) \approx \frac{1}{N - N_b} \sum_{i=1}^{N-N_b} H(\Sigma^* \mid B_i, Y)$$

$H(\Sigma^* \mid B_i, Y)$ is drawn from inverse Wishart with degrees of freedom $T + \alpha$ and scale matrix $\bar{\Sigma} = \bar{S} + SSR_i$, where SSR_i is sum of squared residuals for VAR model for current draw of B_i. In case of DSGE, marginal likelihood is computed via Laplace approximation and is part of *Dynare* output.

As further noted by Koop (2003), one should be careful when comparing models on the basis of marginal likelihood when using non-informative priors, as Minnesota certainly is, especially when different scales of data are included in the models about to be in compared. Non-informative priors are not capable of correction for this different measurement scales and results are inherently biased. However, as all included models are estimated on exactly the same data and of the same measure, i.e. percentage deviations from their respective steady state, comparison is still conducted upon marginal likelihood, despite non-informative Minnesota prior.

Results

At first, individual marginal likelihoods assessing in-sample fit of models are presented in Table 5.1. Significant difference between values for DSGE and VARs is (i) due to different computation method and (ii) weak in-sample fit of DSGE model, visible also on shock decomposition charts[2]. On the other hand, VAR models are able to fit modeled time series almost perfectly (see Figure B.6 in Appendix B). Despite in-sample fit is one indicator for of power of the model, it does not have any further implications for forecasting performance.

Table 5.1: Marginal likelihood based model comparison

	BVAR(1)	BVAR(2)	BVAR(3)	SETAR(2,2,2)	DSGE
Marginal Likelihood	-1748	-1726	-1705	-1759	-1891

Table 5.2: Results Comparison (RMSE)

	In-sample period=50				In-sample period=60			
	Output		Inflation		Output		Inflation	
	1-step	12-step	1-step	12-step	1-step	12-step	1-step	12-step
BVAR(1)	1.22		0.8321		1.18		0.667	
BVAR(2)	1.21	1.423	0.849	0.34	1.16	1.33	0.692	0.346
BVAR(3)	1.215		0.858		1.166		0.71	
SETAR(2,2,2)	1.185	1.515	0.946		1.09	1.56	0.77	
DSGE	1.245	1.5042	0.794	0.704	1.224	1.61	0.612	0.66

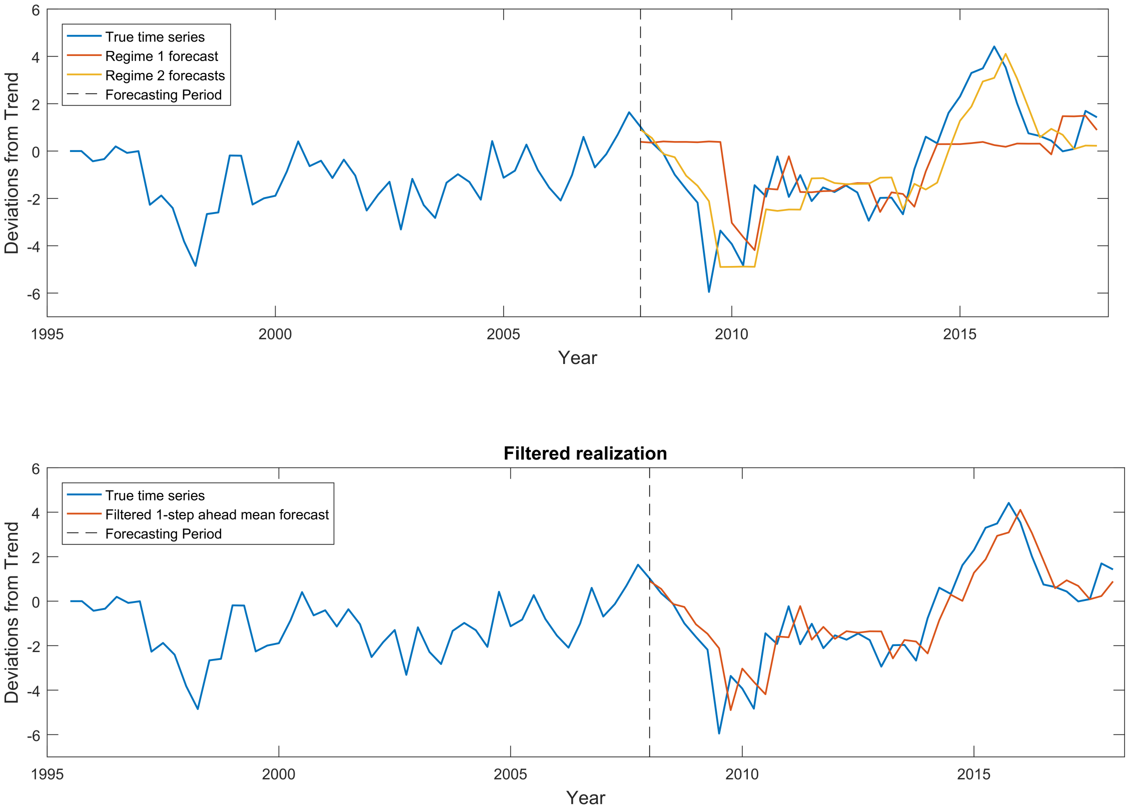
Table 5.2 summarizes out-of-sample **unconditional** forecasting performance of prescribed models on different forecasting horizon K and length of in-sample period

[2]Identifiable as high measurement errors

R. Conditional forecasting is not performed as (i) VAR models do not meet requirements of Lucas critique, and (ii) derived DSGE model does not include any economic policy for forecast to be conditional on, as monetary shocks plays negligible role in the model. In turn, this table well-represents forecasting ability of included models for 1-step ahead forecasts disregarding their properties and set up. 12-step ahead forecast results for DSGE are non-representative, but still valid for other models.

Forecasting performance regarding 1-step ahead forecasts of output with in-sample period $R = 50$ of all included models are almost identical, varying in range only of 0.06. Minimum RMSE for Threshold model perfectly advocates its inclusion in analysis and status of main benchmark model. To understand the importance of threshold on estimation and subsequent forecasting, Figure 5.1 shows 1-step ahead forecasts for both regimes of SETAR(2,2,2) before and after recursive computation of threshold and filtration of both regimes to obtain single realization.

Figure 5.1: Unfiltered and Filtered forecasts of output by SETAR(2,2,2)



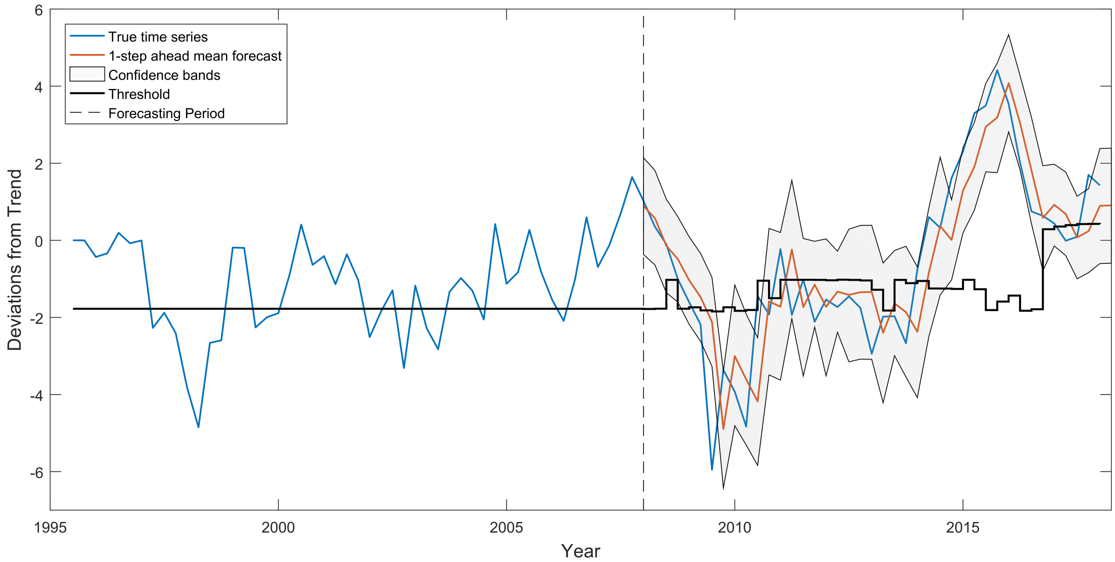
DSGE being the worst, but still relatively close to benchmarks, in not that surprising as structural models are intended to produce rather long term predictions. On longer horizon (12-step ahead forecast), SETAR(2,2,2) is outperformed by BVAR(2) as well as DSGE. Explanation of failure is the fixed level of threshold for the entire 12-quarters long forecasting period and thus its inherent non-representativeness. With increased in-sample period to $R = 60$, order of models remains the same, supporting robustness of stated results.

Following set of Figures visualizes most important results. Note that forecasting performance af all included VAR models are not close only in terms of RMSE, but also in terms of graphic representation Therefore, if not mentioned specifically, stated claims apply to all VAR models.

Output

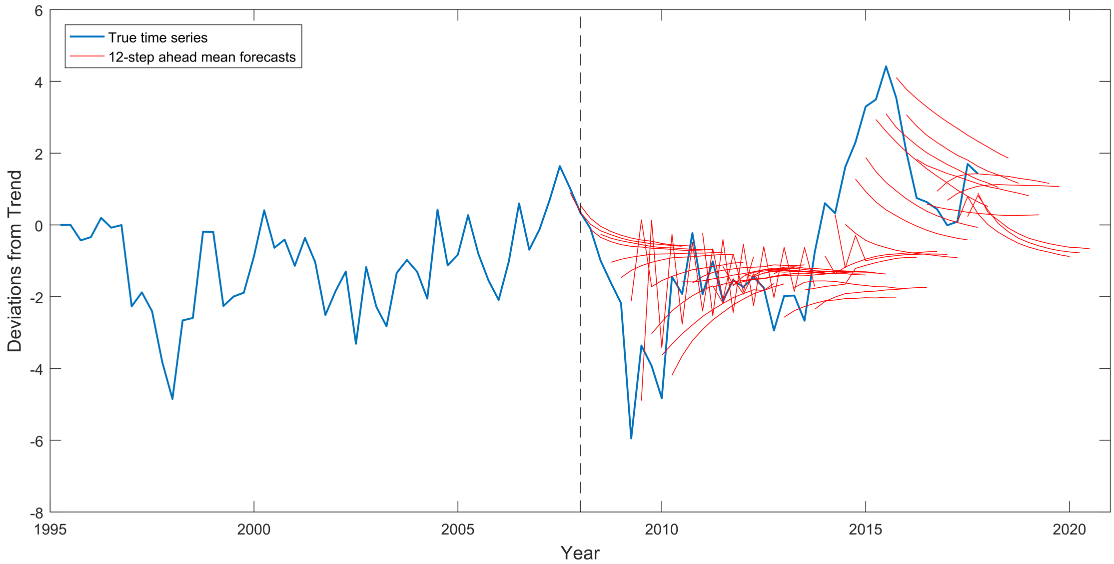
One step ahead forecast with in-sample period of $R = 50$ for SETAR(2,2,2) in Figure 5.2 is accompanied by corresponding level of threshold estimated at that period (but valid with delay of two quarters). Confidence bands are given by 80% of Highest probability density interval (HPDI), with lower bound corresponding to 10^{th} and upper to 90^{th} percentile.

Figure 5.2: 1-step out-of-sample forecast of Ouutput by SETAR(2,2,2)



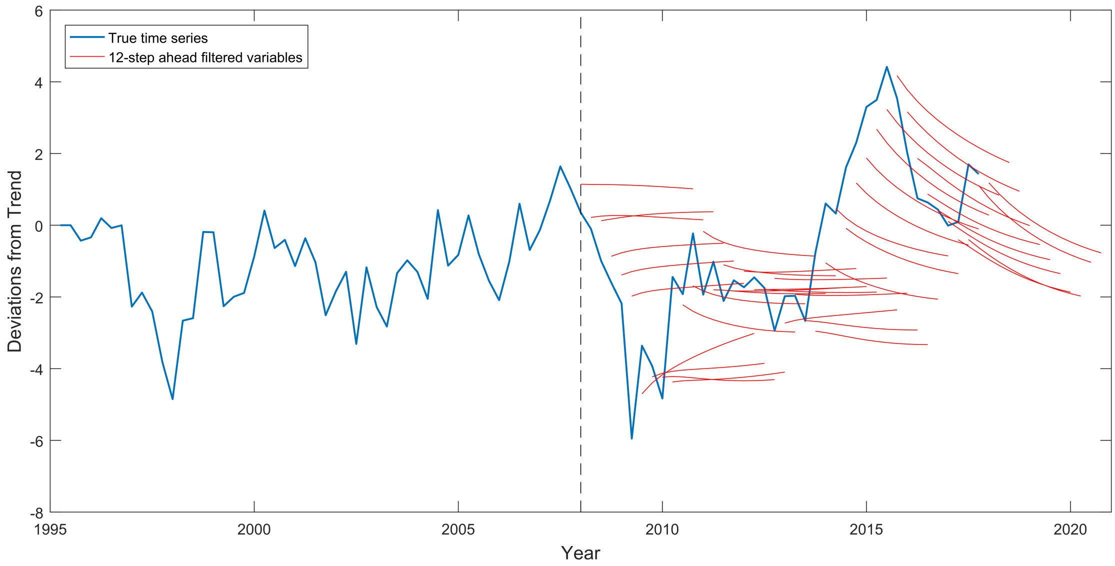
At the first view, VAR models are very adaptive and quickly responsive to past development of tracked time series. They are able to well-define dynamics of the true time series if the spread is not extensive. In vast majority of point forecast, true values lies within the confidence bands.

Figure 5.3: 12-step out-of-sample forecasts of output by SETAR(2,2,2)



In some instances, it appears that forecasted values are just shifted past realizations or shifted in-sample fitted values. By construction, it is natural that 1-step ahead out-of-sample forecasts copy the most recent past development. This phenomenon it better illustrating by plotting 12-step ahead forecasts as in Figure 5.3. This catches dynamics of forecast driven by mean-reversion property significantly better. The further is deviation of analyzed time series from its respective equilibrium rate, the stronger mean reversion property of the forecasts is. For certain periods of time, subsequent forecasts are nothing else then by a constant horizontal shifts of previous ones, creating "shifted" this property of 1-step ahead plots.

Figure 5.4: 12-step out-of-sample forecasts of output by DSGE

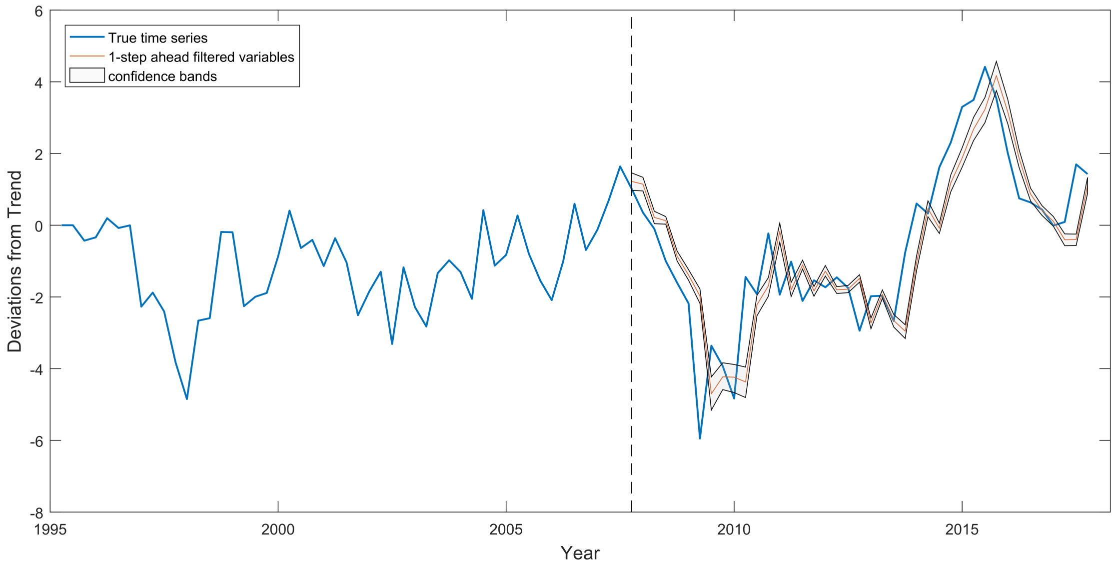


Inability of DSGE to outperform BVAR (2) in long term forecasting is surprising. As a structural model, it should be able to provide high quality forecasts for longer periods ahead. Failure of DSGE is determined in-between 2010-2015 (Figure 5.4), when it is unable to copy short term, kinky, dynamics as apposed to SETAR(2,2,2). This exercise shows that VAR models are much more reactive to current development and are able to incorporate new information faster. On the other hand, one should be really skeptical about DSGE results especially for 12-step ahead forecasts, considering the filtered variable methodology and thus non well-representativeness.

For 1-step ahead forecasting, result and visual realization of DSGE is similar to VAR models. Extremely tight confidence bands are the inherent result of filtered variables approach as there is no uncertainty related to estimated parameters, but only to future realization of shocks. Also DSGE models is estimated by significantly larger number of iterations, in total counting to 1.8 million compared to 50 000 of VAR models for each in-sample period[3], thus reduction in parameter uncertainty even for in-sample period is incorporable across models.

[3]Again, due to limitations in computational power, number of iterations for VAR was reduced in sake of producing full set of out-of-sample forecasts

Figure 5.5: 1-step out-of-sample forecasts of output by DSGE



Inflation

Inflation forecasts in Figure 5.6 corresponds to best-performing model, i.e. BVAR(1), but those are again almost identical for all VARs. It shows that if analyzed time series is not over volatile, VAR is able to form relatively precise point forecasts of its future realizations, with almost all taking place within confidence bands. Consequently, it does not just passively follow past states, but provides relatively smooth trajectory of forecasts.

Figure 5.6: 1-step out-of-sample forecasts of inflation by BVAR(1)

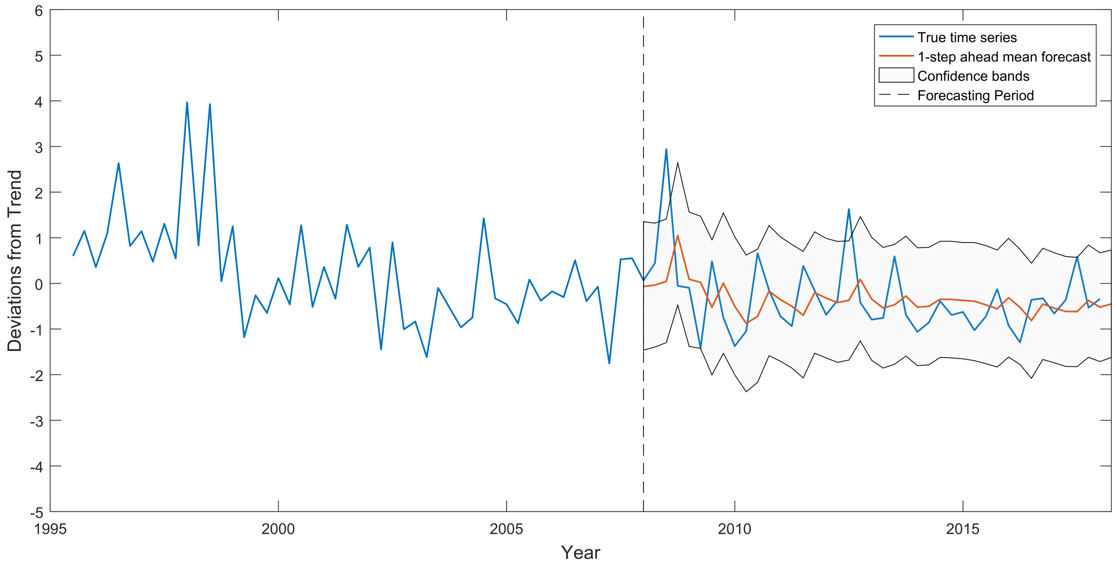
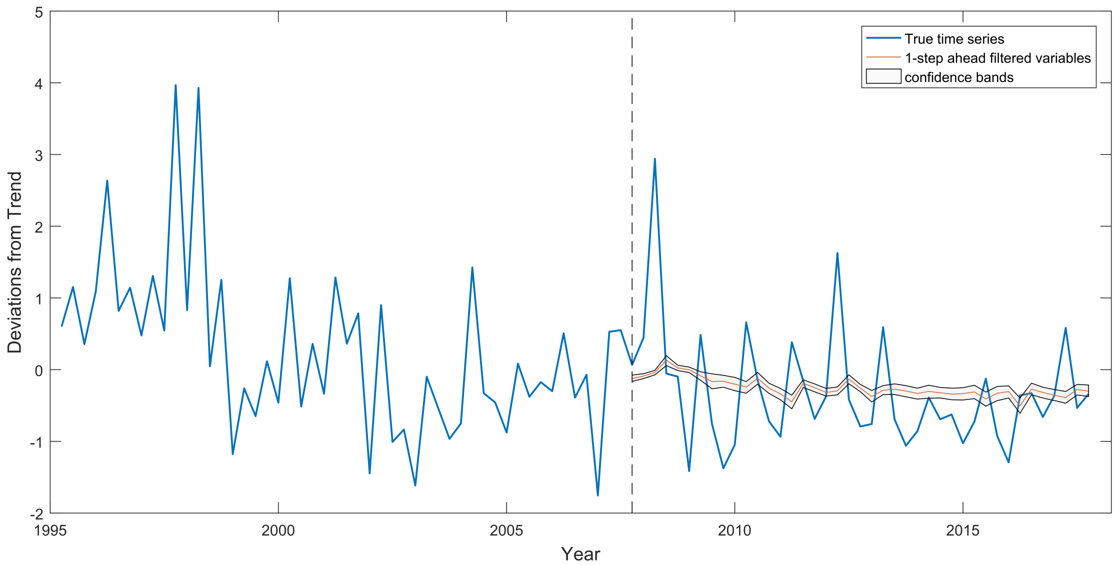
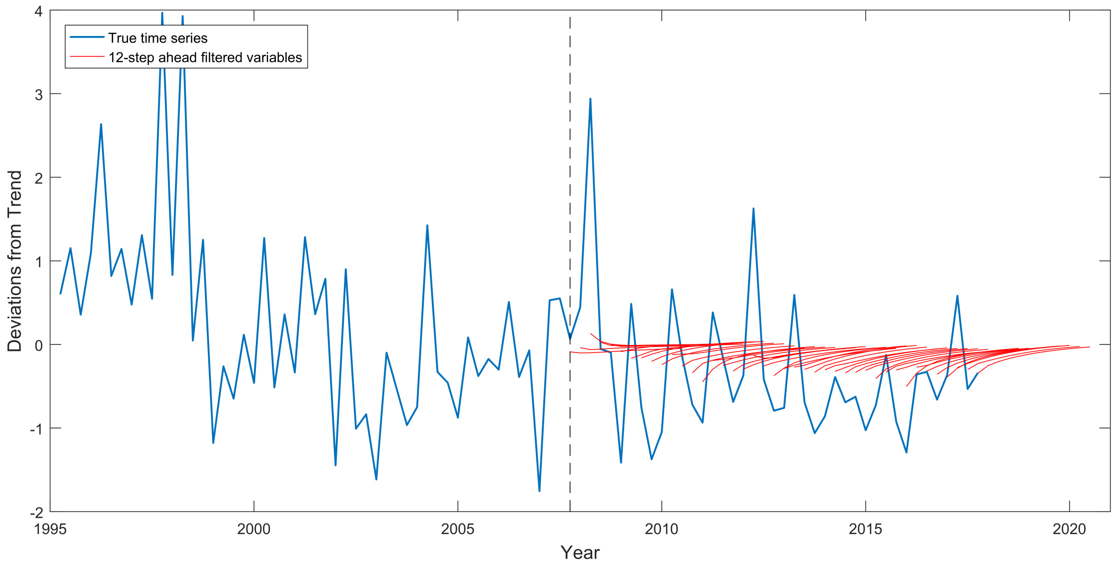


Figure 5.7: 1-step out-of-sample forecasts of inflation by DSGE



DSGE inability to explain historical changes of inflation described in Shock decomposition analysis is transformed also to forecasting capabilities. Although it is the best performing model in forecasting inflation 1-step ahead, Figure 5.7 illustrates that unconditional 12-step ahead forecast in each period practically degrades into reversion to zero mean[4], i.e. target equilibrium value. Paradoxically enough, this inability benefits DSGE in 1-period ahead forecasting, when it outperforms all other models by significant margin, simply by ignoring short run deviations and focusing on its respective equilibrium value at zero mean. Dib et al.(2008) call this sampling variability vs. inconsistencies paradox, when (probably) misspecified decisions rules of DSGE impose, in comparison to VAR, relative tightness on behaviour of the model and not allow to fully follow data. Extremely narrow confidence bands are again a cause of filtered variables approach and superior estimation procedure compared to BVARs.

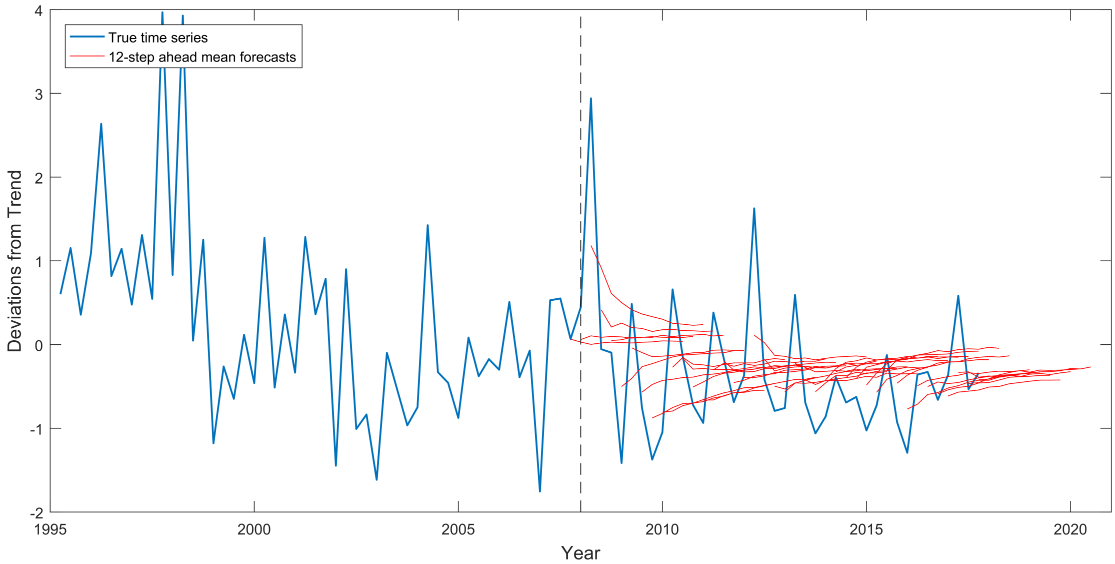
Figure 5.8: 12-step out-of-sample forecasts of inflation by DSGE



Long term inflation forecasts were performed only for BVAR(2). SETAR(2,2,2). It turns out that producing longer period forecasts by SETAR(2,2,2) is extremely computationally demanding and thus forecasts for other variables besides output were discarded. Forecasting with SETAR (2,2,2) in each in-sample period consists of sampling of threshold, sampling of parameters in both regimes, calculation of K-step ahead forecast, and repeating extension of in-sample period and consumes triple amount of time than regular BVAR.

[4]Note the similarity with Inflation forecasts of Czech National bank in last two years.

Figure 5.9: 12-step out-of-sample forecasts of inflation by SETAR(2,2,2)



Simple ranking of the models according to Table 5.2 is:

1. SETAR(2,2,2) for short term output forecasting
2. BVAR(2) for long term output forecasting
3. DSGE for forecasting inflation

Interestingly enough, to obtain presented results required approximately 5 days of intensive computations on 2 separate computers. From this perspective, conduction of such forecasting exercise is not a subject of everyday activity. On the other hand, producing single forecast for 12 subsequent period with sufficient amount of iterations, is a matter of several minutes, depending on computational power. This can be used as essential inputs to other models such as Value at Risk or stress testing. Especially for artificial scenarios analysis, availability of entire posterior distributions of forecasted variables of interest allowing for complex inference and assessment of actual probability of extreme events is considerable advantage. High quality forecasts in such applications are capable of increasing profitability of financial institutions when turning uncertainty into quantifiable risk and thus allowing fo creation of corresponding financial buffer. Given the advantages, relative preciseness and desirable properties of these models, argument for their utilization is at least considerable.

6 Conclusion

In this thesis, small to medium scale closed economy DSGE model for Czech republic was introduced, solved, implemented and with Bayesian techniques estimated on empirically observable time series. DSGE model is a mixture of NKE and RBC modeling traditions, when incorporating frictions in prices and wages, but allowing for perfect mobility of capital. Model properties were judged upon subsequent Impulse Response and Shock decompositions analysis, when identifying supply, demand, momentary shocks and measurement errors representing unexplained part of a model.

Due to modeling Czech republic as a closed economy, exclusion of government and imperfect filtration of observable time series, model is unable to fully explain temporary shocks hitting the economy and past development of variables of interest. Major downturns are (i) the estimate of significantly negative correlation between investment-specific and TFP shock, both being referred to as a supply shocks, and (ii) extremely small standard deviation of monetary shock in Taylor rule. The former results in failure to explain fluctuations in output, while the latter is responsible for unexplained part of inflation and nominal interest rate decomposed into individual shocks. Minor ones are relatively low values of Calvo and inflation indexation parameters compared to similar studies, caused by perfect capital mobility allowing for quick accommodation of shocks.

Model is estimated by Bayesian techniques on Czech major macro time series: output, consumption, investments, inflation, nominal interest rate and GDP deflator. Due to limitations in computational power, forecasted values of DSGE are approximated by filter variables, when parameters are estimated on full sample but only data corresponding to in-sample period are used. **Unconditional** forecasting performance of DSGE model is confronted by Threshold Bayesian VAR (SETAR(2,2,2)) and plain Bayesian VARs with Minnesota prior.

Forecasting exercise is a set of 1-step and 12-step ahead out-of-sample forecasts with 2 different in-sample periods. Although all models perform similarly, Evaluation by RMSE suggests that forecasts of output produced by BVARs are superior to those of DSGE, with SETAR(2,2,2) for short term and BVAR(2) for long term forecasting. On the other hand, DSGE is superior when forecasting inflation in all considered scenarios, paralytically thanks to its misspecification.

Despite some undesirable properties, this DSGE model can still be solid point of departure for extensions. First natural path is to extend model for open economy features. Next is to include actual government expenditures and taxes and rigidities to capital stock formation such as installation costs or capacity utilizations. Such model would then become perfectly suitable for conditional forecasting and economic policy evaluation in further academic studies.

On top of that, this Thesis demonstrates that despite additional demands on development, computational power and complexity, this methods are well-suitable for implementation in real application to provide essential inputs for different scenarios modeling.

Bibliography

Abel A.D. (1990), “Asset Prices under Habit Formation and Catching Up with the Joneses”, American Economic Review, 80:2, p. 38-42.

Adolfson, Malin, Stefan Laséen, Jesper Lindé, and Mattias Villani, Bayesian Estimation of an Open Economy DSGE Model with Incomplete Pass-Through, Journal of International Economics, 72/2 (2007), 481–511

An, Sungbae, and Frank Schorfheide, Bayesian Analysis of DSGE Models (2006)

Ashley, R., C.W.J. Granger, and R. Schmalensee, 1980; Advertising and Aggregate Consumption: An Analysis of Causality, Econometrica 48 (July); pp. 1149-67.

Barbara Rudolf and Mathias Zurlinden, A compact open economy DSGE model for Switzerland, SNB Economic Studies, 8 / 2014

Blake, Andrew, Mumtaz, Haroon, Technical Handbook – No. 4 Applied Bayesian econometrics for central bankers; Centre for Central Banking Studies, Bank of England, 2012

Box, G.E.P., Jenkins, G.M., Reinsel, G.C. (1994), Time Series Analysis, Forecasting and

Burriel, Pablo, Jesús Fernández-Villaverde, and Juan F. Rubio-Ramírez, MEDEA: A DSGE Model for the Spanish Economy, SERIEs, 1/1–2 (2010), 175–243

Calderón Villanueva, Sergio Alejandro, Bayesian Analysis of Multivariate Threshold Autoregressive Models with Missing Data (unpublished PhD Thesis, Universidad Nacional de Colombia, 2014)

Calvo, Guillermo A., 1983. Staggered prices in a utility-maximizing framework, Journal of Monetary Economics, Elsevier, vol. 12(3), pages 383-398, September.

Canova, Fabio, 2007, Methods for Applied Macroeconomic Research, Princeton University Press, Princeton.

Chib, Siddhartha, 1995, Marginal Likelihood from the Gibbs Output, Journal of the American Statistical Association 90(432), 1313—1321. Control, Third Edition. Englewood Cliffs, N.J.: Prentice Hall.

CLARK, Todd E., McCRACKEN, Michael; 2001; Tests of Equal Forecast Accuracy and Encompassing for Nested Models; Journal of Econometrics, 105, 85:110

Dib Ali, Gammoudi Mohamed, Moran Kevin, Forecasting Canadian time Series with New Keynesian Model, The Canadian Journal of Economics Vol. 41, No. 1, Feb.2008, pp 138-165

DeJong, David N.; Cheaten Dave; Structural Macroeconometrics; Princeton University Press; 2007; ISBN-13: 978-0-691-12648-7

Diebold, Francis X., The Past, Present, and Future of Macroeconomic Forecasting, Journal of Economic Perspectives, 12/2 (1998), 175–192

Edge, Rochelle M., Michael T. Kiley, and Jean-Philippe Laforte, A Comparison of Forecast Performance between Federal Reserve Staff Forecasts, Simple Reduced-Form Models, and a DSGE Model, Journal of Applied Econometrics, 25/4 (2010), 720–754

Engle, R.F. and Granger, C.W.J. (1987), "Co-Integration and Error Correction: Erling, M. Kravik, and S. Paulsen Kenneth, A complete documentation of Norges Bankís policy model NEMO, 2007

Fernández-Villaverde, Jesús, and Juan F. Rubio-Ramírez, A Baseline DSGE Model, Duke University and Federal Reserve Bank of Atlanta (2006)

Fernández-Villaverde, Jesús, The Econometrics of DSGE Models, SERIEs, 1/1

(2010), 3–49 Fernandez-Villaverde, Jesús; RUBIO-RAMIREZ, Juan; SCHORFHEIDE, Frank; Solution and Estimation Methods for DSGE Models; April 2007; Handbook of Macroeconomics, Volume 2

Galí, Jordi, Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and its Applications, Princeton, NJ, Princeton University Press, 2007.

Geweke, John, 1991, Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments, Technical report.

Geweke, John, Contemporary Bayesian Econometrics and Statistics, Hoboken, NJ, John Wiley and Sons, 2005

Herbst, Edward P., and Frank Schorfheide, Bayesian Estimation of DSGE Models, Princeton University Press, 2015.

Hristov, Nikolay; The Ifo DSGE Model for the German Economy; 2016; Ifo Working Paper No. 210

INOUE, A., KILIAN L.; 2004; In-sample or out-of-sample tests of predictability: Which one should we use? Econometric Reviews 23, 371-402.

Karlsson, Sune, Forecasting with Bayesian Vector Autoregressions, Handbook of Economic Forecasting, 2/Part B (2013), 791–897

King, Robert G., Rebelo, Sergio T., 1999. `Resuscitating real business cycles,` Handbook of Macroeconomics, in: J. B. Taylor & M. Woodford (ed.), Handbook of Macroeconomics, edition 1, volume 1, chapter 14, pages 927-1007 Elsevier.

Klein, L.R. (1946), The Keynesian Revolution. New York: MacMillan.

Koop, Gary, and Simon M. Potter, Bayesian Analysis of Endogenous Delay Threshold Models, Journal of Business and Economic Statistics, (2000)

Koop, Gary, Bayesian Econometrics, 1th edition., Hoboken, NJ, Wiley-Interscience, 2003. ISBN 978-0-470-84567-7

Kwon, Yongjae, Bayesian Analysis of Threshold Autoregressive Models, Doctoral Dissertations, University of Tennessee – Knoxville, 2003

Kydland, F.E. and Prescott, E.C. (1982), Time to Build and Aggregate Fluctuations, Econometrica, 50, 1345-1371.

Lubok, Thomas; Schorfheide Frank; A Bayesian Look at New Open Economy Macroeconomics, The University of Chicago Press, NBER Macroeconomics Annual, Vol. 20 (2005), pp. 313-366

Lucas, Robert E., Econometric Policy Evaluation: A Critique, in Carnegie-Rochester Conference Series on Public Policy (1976), i, 19–46

Lutkepohl, Helmut, New Introduction to Multiple Time Series Analysis, Springer, 2005, ISBN 3-540-40175-5

PESARAN, Hashem M., SCHUERMANN, Til, SMITH, Vanessa L.; 2008; Forecasting Economic and Financial Variables with Global VARs; Federal Reserve Bank of New York; Staff Reports

Pfeifer, Johannes, A Guide to Specifying Observation Equations for the Estimation of DSGE Models, Research Series, 2014, 1–150

Pfeifer, Johannes, An Introduction to Graphs in Dynare, University of Mannheim, 2014 Representation, Estimation and Testing," Econometrica, 55, 251-276.

Schmitt-Grohé, Stephanie, and Martın Uribe, Closing Small Open Economy Models, Journal of International Economics, 61/1 (2003), 163–185

Sims, C.A. (1980), "Macroeconomics and Reality," Econometrica, 48, 1-48.

Sims, Christopher A and Tao Zha, 1998, Bayesian Methods for Dynamic Multivariate Models, International Economic Review 39(4), 949—68.

Sims, Eric. Graduate Macro Theory II: A Medium-Scale New Keynesian DSGE Model, University of Notre Dame (2016)

Sims, Eric. Graduate Macro Theory II: Notes on Medium Scale DSGE Models, University of Notre Dame (2011)

Smets, Frank; Wouters, Raf; An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area; August 2002; European Central Bank Working Paper Series; Working Paper no. 171

Stock, James H., and Mark W. Watson, Forecasting inflation, Journal of Monetary Economics 44.2 (1999), 293-335

Taylor, John B., and Harald Uhlig, Handbook of Macroeconomics, Volume 2A-2B SET - 1st Edition

Tong, H; Lim K. S.; Threshold Autoregression, Limit Cycles and Cyclical Data; Journal of the Royal Statistical Society. Series B (Methodological), Vol. 42, No. 3. (1980), pp. 245-292.

Tong, Howell, Threshold Models in Time Series Analysis — 30 Years On, Statistics and Its Interface, 4/2 (2011), 107–18

Vukotiśc, Marija, Exchange Rate Dynamics in an Estimated Small Open Economy DSGE Model, Duke University, 2007

Walsch, Carl E., Monetary Theory and Policy, The MIT Press, Cambridge, Massachusetts, third edition, 2010, ISBN 978-0-262-01377-2

A DSGE Implementation

A.1 *Dynare* Implementation

Listing A.1: Non-linear model for log-linearization

```
%Specify endogenous variables prepared for Tex output

var
c        $c$ (long_name='Consumption')
lam      $\lambda$ (long_name='Shadow Price')
R        $R$ (long_name='Gross nominal IR')
PI       $\Pi$ (long_name='Gross inflation rate')
r        $r$ (long_name='Real IR')
x        $x$ (long_name='Investments')
f        $f$ (long_name='Wage setting equation')
ld       $N^d$ (long_name='Labour Demand')
w        $W$ (long_name='Wage')
wstar    $W^*$ (long_name='Optimal reset wage')
PIstar   $\Pi^*$ (long_name='Optimal reset price')
PIstarw  $\Pi_w^*$ (long_name='Optimal wage inflation')
g1       $f1$ (long_name='Price setting equation')
g2       $f2$ (long_name='Price setting equation')
yd       $y$ (long_name='Output')
mc       $mc$ (long_name='Marginal costs')
k        $k$ (long_name='Capital stock')
vp       $v^p$ (long_name='Price dispersion')
vw       $v^v$ (long_name='Wage dispersion')
l        $n$ (long_name='Labour supply')
a        $a$ (long_name='TFP shock')
b        $b$ (long_name='Investment shock')
m        $m$ (long_name='Money demand')
eps_c    $\epsilon^c$ (long_name='Consumption shock')
eps_m    $\epsilon^c$ (long_name='Consumption shock')
d        $\epsilon^l$ (long_name='Labour supply shock')
e        $\varepsilon^r$ (long_name='Monetary policy shock')
y_obs    $y^{obs}$ (long_name='Observed outuput')
PI_obs   $\Pi^{obs}$ (long_name='Observed gross inflation')
l_obs    $n^{obs}$ (long_name='Observed labour supply')
c_obs    $c^{obs}$ (long_name='Observed consumption')
vp_obs   $v^{obs}$ (long_name='Observed GDP deflator')
R_obs    $R^{obs}$ (long_name='Observed gross nominal IR')
x_obs    $x^{obs}$ (long_name='Observed investments')
;

varexo
epsmE    $\varepsilon^r$ (long_name='Monetary policy shock')
e_a      $\varepsilon^a$ (long_name='TFP shock')
e_b      $\varepsilon^b$ (long_name='Investment shock')
```

```
e_c      $\varepsilon^c$ (long_name='Consumption shock')
e_d      $\varepsilon^l$ (long_name='Labour supply shock')
e_y      $\varepsilon^y$ (long_name='Output Measurement Error')
e_C      $\varepsilon^C$ (long_name='Consumption Measurement Error')
e_x      $\varepsilon^x$ (long_name='Investment Measurement Error')
e_R      $\varepsilon^R$ (long_name='Gross nominal IR Measurement Error')
e_PI     $\varepsilon^\Pi$ (long_name='Gross inlfation Measurement Error')
e_l      $\varepsilon^n$ (long_name='Labour supply Measurement Error')
e_vp     $\varepsilon^{vp}$ (long_name='GDP deflator Measurement Error')
;

% Specify predetrmineed variables and keep their repsective timings as in derivation
predetermined_variables k;

%
parameters
bet       $\beta$ (long_name='Discoutn factor')
del       $\delta$ (long_name='Depresiation rate')
eps_l     $\epsilon_l$ (long_name='Elasticity of substitution between labours')
eps_h     $\epsilon$ (long_name='Elasticity of substitution between goods')
psy       $\psi$ (long_name='Labor disutility ')
eta       $\eta$ (long_name='Inverse Frisch elasticity')
kappa     $\kappa$ (long_name='Price indexation to inflation')
kappaw    $\kappa_w$ (long_name='Wage indexation to inflation')
theta     $\theta$ (long_name='Calvo parameter prices')
theta_w   $\theta_w$ (long_name='Calvo parameter wages')
alf       $\alpha$ (long_name='Capital share of output')
%Rbar     $\bar{R}$ (long_name='Steady state IR') %Floating parameter
PIbar     $\bar{\Pi}$ (long_name='Inflation target')
gamR      $\gamma_R$ (long_name='TR: IR smoothing')
gam_PI    $\gamma_\Pi$ (long_name='TR: inflation smoothing')
gam_y     $\gamma_y$ (long_name='TR: output gap smoothing')
sig_m     $\sigma_m$ (long_name='Consumption')
rho_e     $\rho_e$ (long_name='AR parameter MP shock')
rho_a     $\rho_a$ (long_name='AR parameter MP shock')
sig_a     $\sigma_a$ (long_name='stddev TFP')
sig_m     $\sigma_m$ (long_name='stddev monetary')
h         $h$ (long_name='Habit persistence')
sig_b     $\sigma_b$ (long_name='stddev investmetns')
sig_c     $\sigma_c$ (long_name='stddev consumption')
sig_d     $\sigma_l$ (long_name='stddev labour supply')
rho_b     $\rho_b$ (long_name='AR parameter investment shock')
rho_c     $\rho_c$ (long_name='AR parameter consumption shock')
rho_d     $\rho_l$ (long_name='AR parameter labour supply shock')
rho_m     $\rho_m$ (long_name='AR parameter monetary shock')
mu_m      $\mu$ (long_name='Consumption')
sigma_c   $\sigma$ (long_name='Inverse of intertemporal substitution in consumption')
;

del=0.05;
eps_h=7;
eps_l=7;
bet=0.96;
psy=6;
eta = 2;
alf =0.66;
theta =0.2;
theta_w=0.2;
kappa = 0.6;
kappaw = 0.6;
gamR =0.7;
gam_y =0.6;
gam_PI =1.2;
PIbar = 1.02;
sig_m =0.004;
```

```
rho_e=0.1;
rho_a=0.8;
rho_m=0.8;
sig_a=0.009;
sig_b=0.009;
sig_c=0.009;
sig_d=0.009;
h=0.6;
rho_b=0.7;
rho_c=0.2;
rho_d=0.2;
mu_m=2;
sigma_c=3;



model;
#Rbar = PIbar/bet;

%1. FOC consumption
exp((eps_c))*(exp(c)-h*exp(c(-1)))^(-sigma_c)-h*(exp(eps_c(+1)))*bet*(exp(c(+1))-h*exp(c))^(-sigma_c)=exp(lam);

%2. FOC bonds
exp(lam)=bet*exp(lam(+1))/exp(PI(+1))*exp(R);

%3. FOC capital
exp(lam)=bet*exp(lam(+1))*exp((b))/exp((b(+1)))*(1-del+(exp(b(+1)))*exp(r(+1)));

%4. FOC money
exp((m))^(-mu_m)=exp(lam)-bet*exp(lam(+1))/exp(PI(+1));

% Labour market
%5. - 6.
exp(f)=(eps_l-1)/eps_l*exp(wstar)^(1-eps_l)*exp(lam)*exp(w)^eps_l*exp(ld)+bet*theta_w*(exp(PI)^kappaw/exp(PI(+1)))^(1-eps_l)*(exp(wstar(+1))/exp(wstar))^(eps_l-1)*exp(f(+1));
exp(f)=psy*exp(d)*exp(eps_c)*exp(PIstarw)^(-eps_l*(1+eta))*exp(ld)^(1+eta)+bet*theta_w*(exp(PI)^kappaw/exp(PI(+1)))^(-eps_l*(1+eta))*(exp(wstar(+1))/exp(wstar))^(eps_l*(1+eta))*exp(f(+1));

% 7. Law of Motion of Wages
1=theta_w*(exp(PI(-1))^kappaw/exp(PI))^(1-eps_l)*(exp(w(-1))/exp(w))^(1-eps_l)+(1-theta_w)*exp(PIstarw)^(1-eps_l);

%8-10. firm's price setting
exp(g1)=exp(lam)*exp(mc)*exp(yd)+bet*theta*(exp(PI)^kappa/exp(PI(+1)))^(-eps_h)*exp(g1(+1));
exp(g2)=exp(lam)*exp(PIstar)*exp(yd)+bet*theta*(exp(PI)^kappa/exp(PI(+1)))^(1-eps_h)*exp(PIstar)/exp(PIstar(+1))*exp(g2(+1));
eps_h*exp(g1)=(eps_h-1)*exp(g2);

%11-12. optimal inputs (11-12)
exp(k)/exp(ld)=alf/(1-alf)*exp(w)/exp(r);
exp(mc)=(1/(1-alf))^(1-alf)*(1/alf)^alf*exp(w)^(1-alf)*exp(r)^alf/exp(a);

%13. law of motion prices
1=theta*(exp(PI(-1))^kappa/exp(PI))^(1-eps_h)+(1-theta)*exp(PIstar)^(1-eps_h);

%14. Law of motion for capital
exp(x)*exp(b) = exp(k(+1)) - (1-del)*exp(k);

%15. Taylor Rule
exp(R)/Rbar=(exp(R(-1))/Rbar)^gamR*((exp(PI)/PIbar)^gam_PI*(exp(yd)/exp(yd(-1)))^gam_y)^(1-gamR)*exp(e);


%16-18. Market clearing
```

```
exp(yd)=exp(c)+exp(x);
exp(yd)=(exp(a)*exp(k)^alf*exp(ld)^(1-alf))/exp(vp);
exp(l)=exp(vw)*exp(ld);

%19-20. Price and wage dispersion terms
exp(vp)=theta*(exp(PI(-1))^kappa/exp(PI))^(-eps_h)*exp(vp(-1))+(1-theta)*exp(PIstar)^(-eps_h);
exp(vw)=theta_w*(exp(w(-1))/exp(w)*exp(PI(-1))^kappaw/exp(PI))^(-eps_l)*exp(vw(-1))+(1-theta_w)*
    exp((PIstarw))^(-eps_l);

%21. Equilibrium wage
exp(PIstarw)=exp(wstar)/exp(w);

%21-26. Shock processes
a=rho_a*a(-1)+e_a;
b=rho_b*b(-1)+e_b;
eps_c=rho_c*eps_c(-1)+e_c;
eps_m=rho_m*eps_m(-1)+epsmE;
d=rho_d*d(-1)+e_d;
e=rho_e*e(-1)+epsmE;

%Observation equations
y_obs=yd-steady_state(yd)+e_y;
c_obs=c-steady_state(c)+e_C;
l_obs=l-steady_state(l)+e_l;
x_obs=x-steady_state(x)+e_x;
PI_obs=PI-steady_state(PI)+e_PI;
vp_obs=vp-steady_state(vp)+e_vp;
R_obs=R-steady_state(R)+e_R;
end;


% Shock can be ucommented for stoch_simul command
%shocks;
%var epsmE;    stderr (sig_m);
%var e_a;    stderr sig_a;
%var e_b;    stderr sig_b;
%var e_c;    stderr sig_c;
%var e_d;    stderr sig_d;
%var e_a, e_b =0.1*0.009*0.009;
%end;

%stoch_simul(order=1,periods=200,irf=0)yd y_obs R_obs c l l_obs PI PI_obs;

write_latex_original_model;
write_latex_static_model;
write_latex_dynamic_model(write_equation_tags);
write_latex_parameter_table;
write_latex_definitions;

varobs y_obs PI_obs l_obs vp_obs c_obs x_obs R_obs;

estimated_params;
alf, beta_pdf, 0.6, 0.1;
theta, beta_pdf,0.5, 0.1;
theta_w, beta_pdf,0.5, 0.1;
kappaw, beta_pdf,0.5, 0.15;
kappa, beta_pdf,0.4, 0.15;
stderr e_a, inv_gamma_pdf, 0.05, inf;
stderr e_b, inv_gamma_pdf, 0.05, inf;
stderr e_c, inv_gamma_pdf, 0.05, inf;
stderr e_d, inv_gamma_pdf, 0.05, inf;
stderr e_y, inv_gamma_pdf, 0.05, inf;
corr e_a, e_b, 0.5, , , beta_pdf, 0, 0.3, -1, 1;
corr e_c, e_d, 0.5, , , beta_pdf, 0, 0.3, -1, 1;
stderr epsmE, inv_gamma_pdf, 0.01, inf;
```

```
stderr e_l, inv_gamma_pdf, 0.05, inf;
stderr e_x, inv_gamma_pdf, 0.05, inf;
stderr e_PI, inv_gamma_pdf, 0.05, inf;
stderr e_C, inv_gamma_pdf, 0.05, inf;
stderr e_vp, inv_gamma_pdf, 0.05, inf;
stderr e_R, inv_gamma_pdf, 0.05, inf;
rho_a, beta_pdf,0.7, 0.1;
rho_b, beta_pdf,0.7, 0.1;
rho_c, beta_pdf,0.7, 0.1;
rho_d, beta_pdf,0.7, 0.1;
rho_m, beta_pdf,0.8, 0.1;
rho_e, beta_pdf,0.4, 0.05;
h, beta_pdf, 0.7, 0.15;
bet, beta_pdf, 0.95, 0.02;
del, beta_pdf, 0.05, 0.04;
eta, gamma_pdf, 3.5, 1.5;
eps_l, gamma_pdf,8,2;
eps_h, gamma_pdf,5,2;
sigma_c, gamma_pdf, 1, 0.2;
gam_y, beta_pdf,0.3, 0.15;
gam_PI, gamma_pdf,1.5, 0.2;
gamR, beta_pdf,0.5, 0.15;
psy, normal_pdf,5,1;
end;

write_latex_prior_table;

shock_groups(name=group1);
Supply = e_a, e_b, e_d;
Demand = e_c;
Monetary=epsmE;
'Measrement Error'= e_y, e_l, e_x, e_PI, e_C, e_R, e_vp;
end;

estimation(datafile=data_try2, first_obs=1, mh_replic=300000, mh_nblocks=6, mh_drop=0.6, mh_jscale=0.35, filtered_vars, logdata,
mode_compute=4, filter_step_ahead=[1:12], mode_check, bayesian_irf, tex, forecast=12) y_obs PI_obs l_obs vp_obs c_obs x_obs R_obs;

%[logdata]: prevents Dynare from logging data
%[mh_replic]: iterations per blck of estimation
%[mh_blocks]: number of blocks
%[bayesian_irf]: produces IRF
%[filter_step_ahead]: filtered varibales for perioed 1-12
%[tex]: created Tex output



%stoch_simul(order=1,periods=200)y y_obs R_obs c c_obs l l_obs ;
shock_decomposition(parameter_set=posterior_mean,use_shock_groups=group1) y_obs PI_obs l_obs vp_obs c_obs R_obs

collect_latex_files;
```

A.2 Steady state file

Listing A.2: Non−linear model for log−linearization *(_steady_state.m)*

```
function [ys,check] = DP_exp_steadystate(ys,exo)
% function [ys,check] = NK_baseline_steadystate(ys,exo)
% computes the steady state for the NK_baseline.mod and uses a numerical
% solver to do so
% Inputs:
% - ys [vector] vector of initial values for the steady state of
% the endogenous variables
% - exo [vector] vector of values for the exogenous variables
%
% Output:
% - ys [vector] vector of steady state values fpr the the endogenous variables
% - check [scalar] set to 0 if steady state computation worked and to
% 1 of not (allows to impos restriction on parameters)


global M_
(*@    @*)
% read out parameters to access them with their name
NumberOfParameters = M_.param_nbr;
for ii = 1:NumberOfParameters
paramname = deblank(M_.param_names(ii,:));
eval([ paramname ' = M_.params(' int2str(ii) ');']);
end
% initialize indicator
check = 0;


options=optimset(); % set options for numerical solver

%set equilibriu values of shocks
PI=log(PIbar);
d=1;
a=0;
b=0;
d=0;
e=0;
eps_c=0;
eps_m=0;


%Analysital computation of steady state as in Section 3.7
Rbar=PIbar/bet;
R=log(Rbar);
r=log(bet^(-1)-1+del);



PIstar=log(((1-theta*exp(PI)^( (1-eps_p)*(1-chi)))/(1-theta))^(1/(1-eps_p)));
PIstarw=log(((1-theta_w*exp(PI)^(-(1-kappaw)*(1-eps_l)))/(1-theta_w))^(1/(1-eps_l)));

mc=log((eps_p-1)/eps_p*(1-bet*theta*exp(PI)^((1-kappa)*eps_p))/(1-bet*theta*exp(PI)^(-(1-eps_p)
    *(1-kappa)))*exp(PIstar));
w=log((1-alf)*(exp(mc)*(alf/exp(r))^alf)^(1/(1-alf)));
%w=(1-alf)*(exp(mc)*(alf/r)^alf)^(1/(1-alf));

wstar=log(exp(w)*exp(PIstarw));
vp=log((1-theta)/(1-theta*exp(PI)^((1-kappa)*eps_p))*exp(PIstar)^(-eps_p));
vw=log((1-theta_w)/(1-theta_w*exp(PI)^((1-kappaw)*eps_l))*exp(PIstarw)^(-eps_l));
tempvaromega=alf/(1-alf)*exp(w)/exp(r);



%Fsolve function to solve equation 3.104
```

```
[ld,fval,exitflag]=fsolve(@(ld)(1-bet*theta_w*exp(PI)^(-(1-kappaw)*(1-eps_l)))/(1-bet*theta_w*exp(PI)^(eps_l*(1-kappaw)*(1+gam)))...
-(eps_l-1)/eps_l*exp(wstar)/(psy*exp(PIstarw)^(-eps_l*gam)*exp(ld)^gam)*((1-h*bet)*(1-h)^(-sigma_c))*...
((exp(vp)^(-1)*tempvaromega^alf-tempvaromega*del)*exp(ld))^(-sigma_c),0.25,options);
disp(ld)

l=log(exp(vw)*exp(ld));
k=log(tempvaromega*exp(ld));
x=log(del*exp(k));
yd=log((exp(k)^alf*exp(ld)^(1-alf))/exp(vp));
c=log((exp(vp)^(-1)*exp(tempvaromega)^alf-exp(tempvaromega)*del)*exp(ld));
c=log(exp(yd)-exp(x));
lam=log((1-h*bet)*(1-h)^(-sigma_c)*exp(c)^(-sigma_c));
f=log((eps_l-1)/eps_l*exp(wstar)*exp(PIstarw)^(-eps_l)*exp(lam)*exp(ld)/(1-bet*theta_w*exp(PI)^(-(1-kappaw)*(1-eps_l))));
f2=log(psy*exp(PIstarw)^(-eps_l*(1+gam))*exp(ld)^(1+gam)/(1-bet*theta_w*(exp(PI)^kappaw/exp(PI))^(-eps_l*(1+gam))*(exp(wstar)/exp(wstar))^(eps_l*(1+gam))));

g1=log(exp(lam)*exp(mc)*exp(yd)/(1-bet*theta*exp(PI)^((1-kappa)*eps_p)));
g2=log(eps_p/(eps_p-1)*exp(g1));
m=log((exp(lam)*((exp(R)-1)/exp(R)))^(-1/mu_m));
y_obs=0;
PI_obs=0;
l_obs=0;
c_obs=0;
vp_obs=0;
R_obs=0;
x_obs=0;
%% end own model equations
for iter = 1:length(M_.params) %update parameters set in the file
eval([ 'M_.params(' num2str(iter) ') = ' M_.param_names(iter,:) ';' ])
end

NumberOfEndogenousVariables = M_.orig_endo_nbr; %auxiliary variables are set automatically
for ii = 1:NumberOfEndogenousVariables
varname = deblank(M_.endo_names(ii,:));
eval(['ys(' int2str(ii) ') = ' varname ';']);
end
```

B Appendix

Figure B.1: Shock decomposition - Consumption

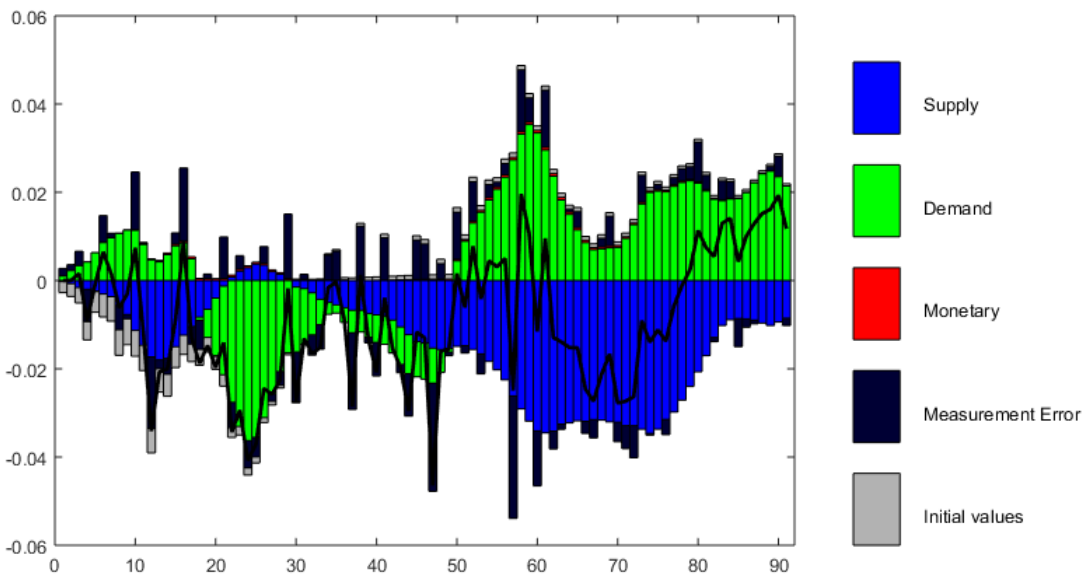


Figure B.2: Shock decomposition - Investments

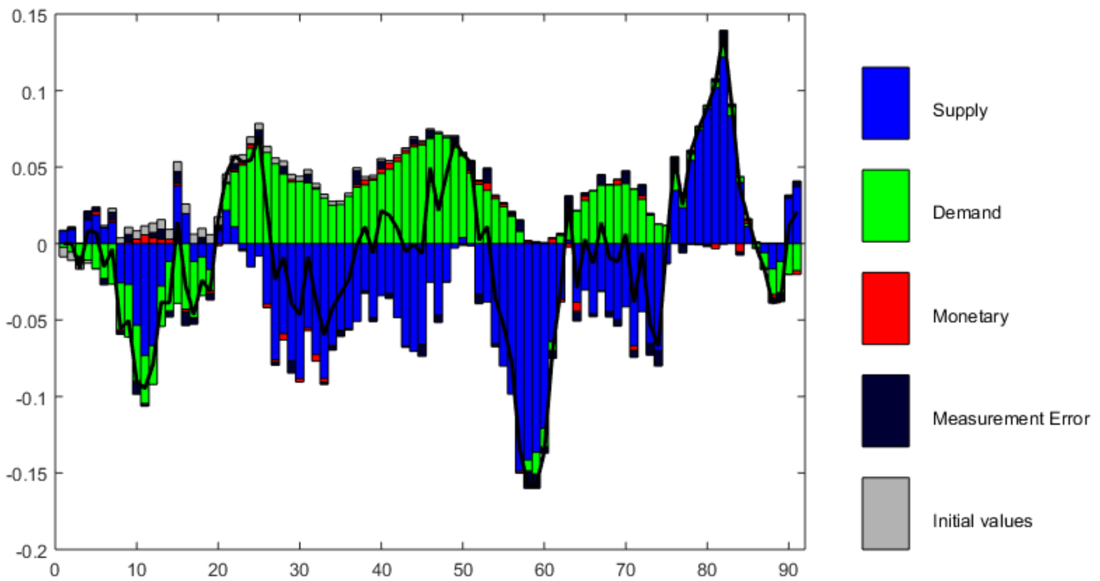


Figure B.3: Shock decomposition - Hours worked

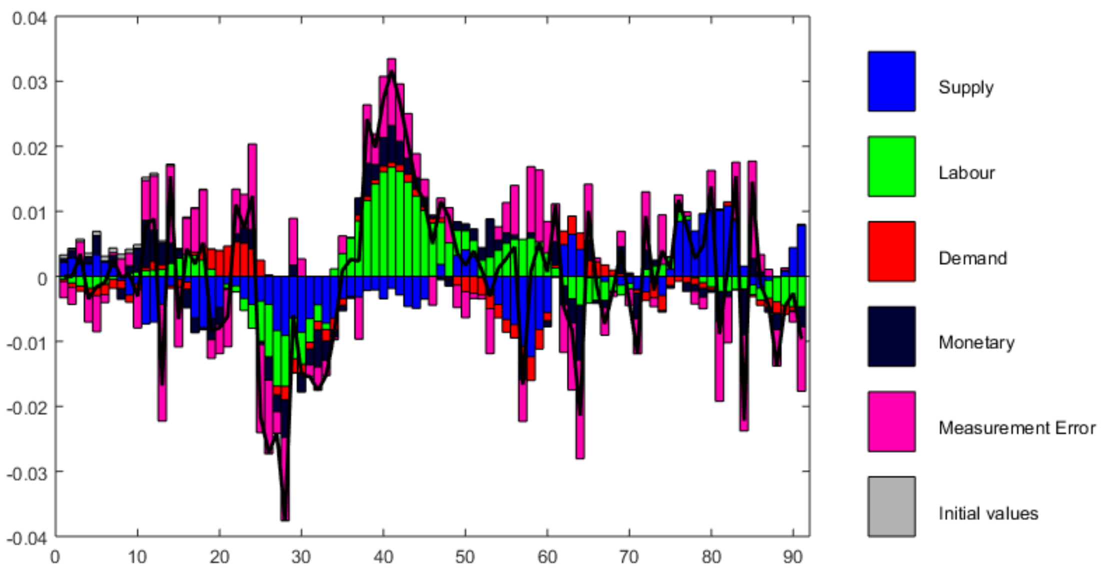


Figure B.4: Shock decomposition - Inflation

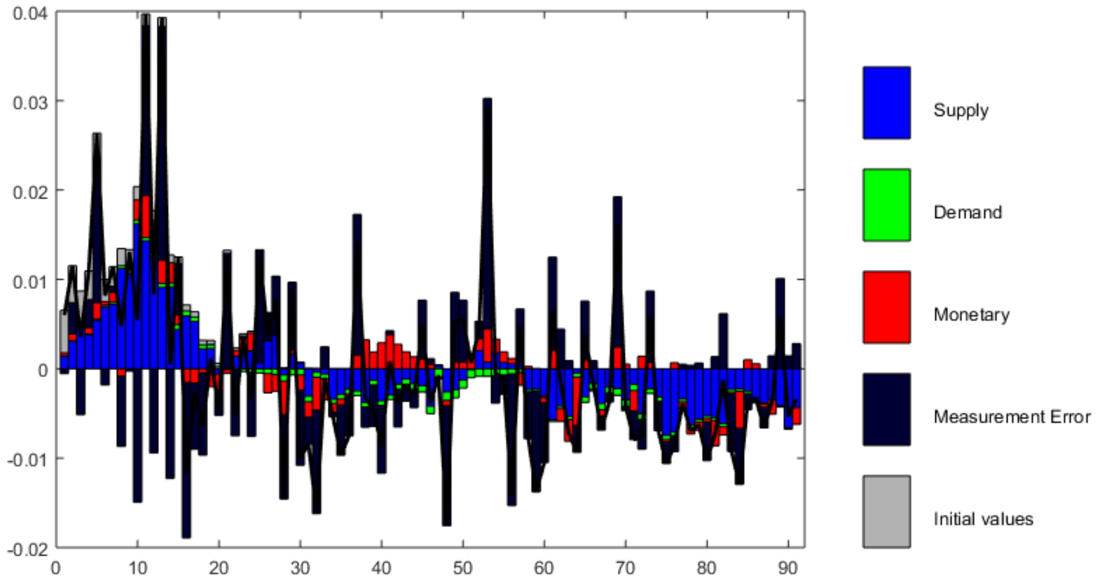


Figure B.5: Shock decomposition - Gross nominal IR

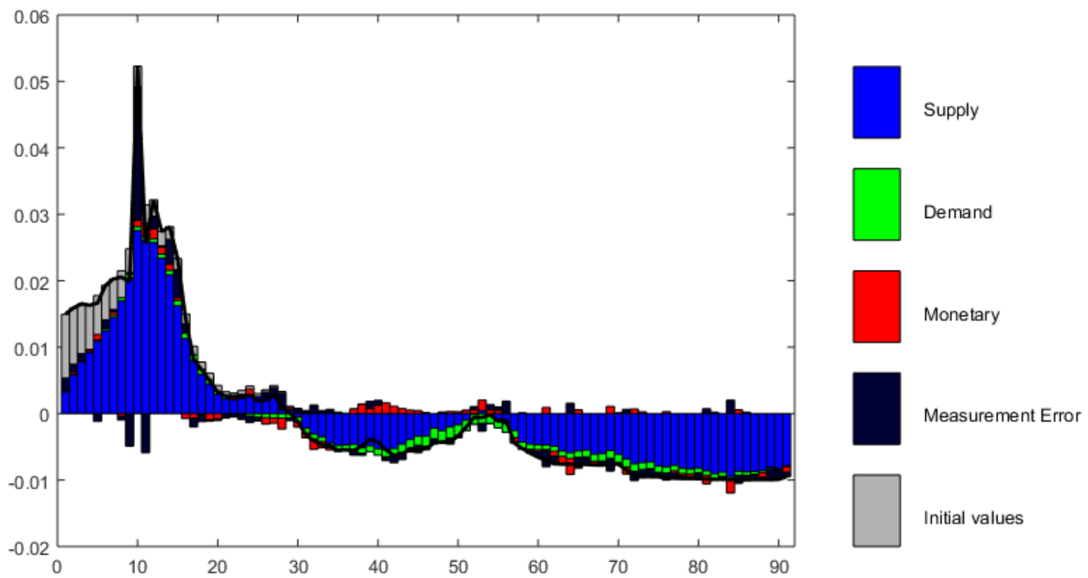
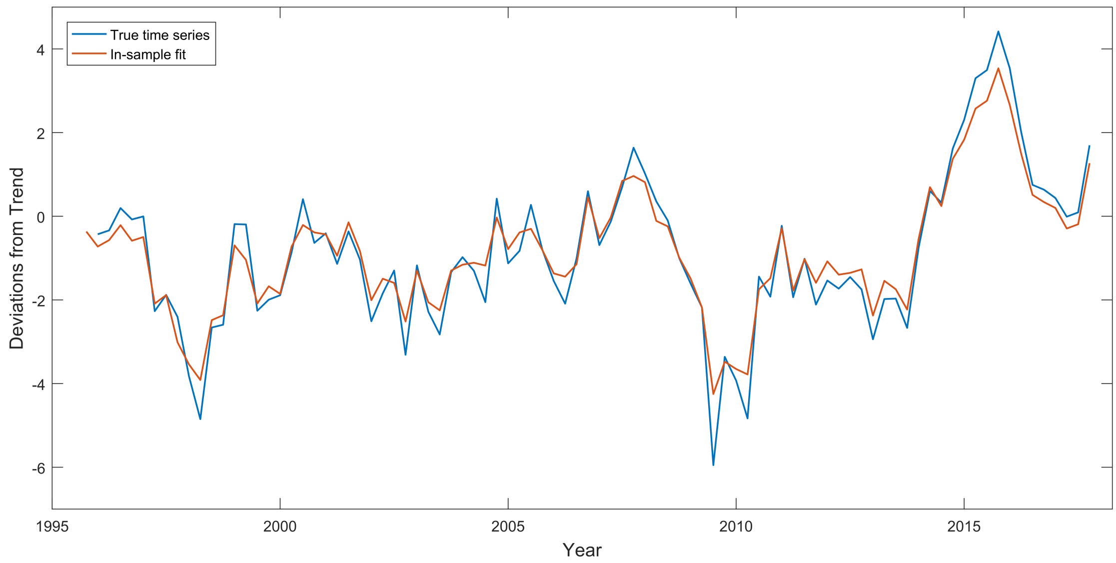


Figure B.6: In-sample ft of output by BVAR(2)



For successful initiation of MH algorithm, finding mode of posterior distribution is required. In the Figures, it is represented by highest density of blue line. Red line represents deviation of currently approximated posterior distribution from the prior. The further are these from each other, the bigger difference.

Figure B.7: Mode Check Plots (1)

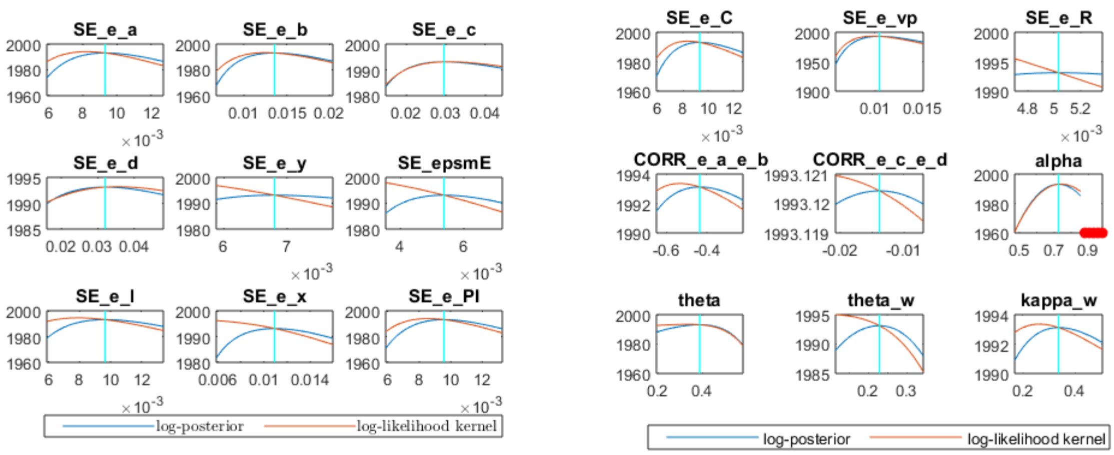
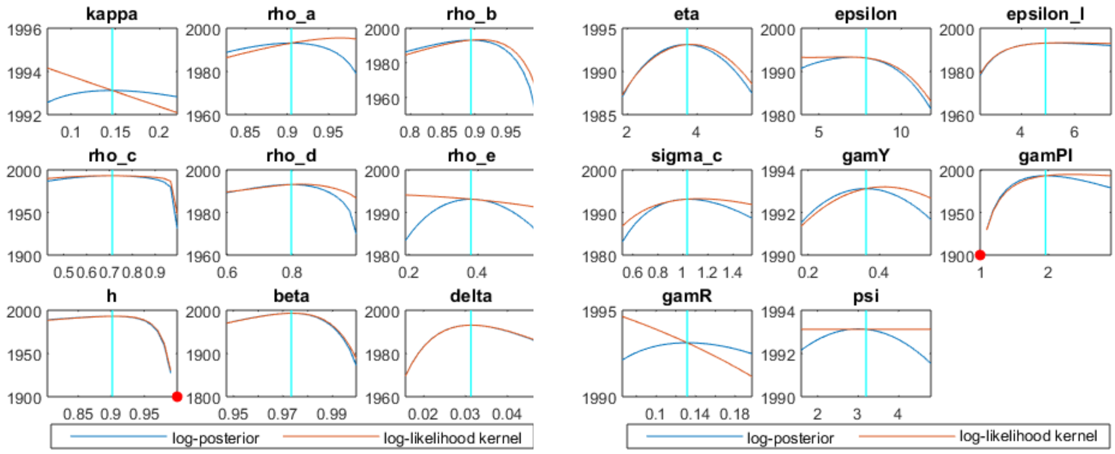


Figure B.8: Mode Check Plots (2)



C BVAR Implementations in Matlab

C.1 Setting of Minnesota Prior

Setting of prior is due to extensiveness demonstrated only on least setting, i.e. BVAR(1)

Listing C.1: Function Get prior

```
function [H, betaols]=get_prior(Y,X)

%returns prior variance B~N(b,H)

% set b
B0=zeros((N*L+1),N);
for i=1:N
B0(i+1,i)=1;
end
B0=vec(B0);

%compute standard deviation of each series residual via an ols regression
%to be used in setting the prior
betaols=vec(inv(X'*X)*(X'*Y));

%set hyperparameters
lamda1=0.2; %controls the prior on own lags
lamda2=0.5; %on other lags
lamda3=1;   %size of lags of higher degree
lamda4=10000; %for constatn term

%compute stardard error sigma(i) to be used in H
y=Y(:,1);                %select only first variable
x=X(:,1:2);      %select vector of ones for intersepts and lag of the same variabel as Y
b0=inv(x'*x)*(x'*y);              %perform OLS
s1=sqrt(((y x*b0)'*(y x*b0))/(rows(y) 2));  %std of residual standard error
%second variable
y=Y(:,2);
x=X(:,[1 3]);
b0=inv(x'*x)*(x'*y);
s2=sqrt(((y x*b0)'*(y x*b0))/(rows(y) 2));
% third variable
y=Y(:,3);
x=X(:,[1 4]);
b0=inv(x'*x)*(x'*y);
%... and analogously for remaining variables 4 7.

%Specify the prior variance of vec(B)
H=zeros(56,56);
```

```
%for equation 1  of the VAR
H(1,1)=(s1*lamda4)^2;
H(2,2)=(lamda1)^2;
H(3,3)=((s1*lamda1*lamda2)/s2)^2;
H(4,4)=((s1*lamda1*lamda2)/s3)^2;
H(5,5)=((s1*lamda1*lamda2)/s4)^2;
H(6,6)=((s1*lamda1*lamda2)/s5)^2;
H(7,7)=((s1*lamda1*lamda2)/s6)^2;
H(8,8)=((s1*lamda1*lamda2)/s7)^2;

%second equation
H(9,9)=(s2*lamda4)^2;
H(10,10)=((s2*lamda1*lamda2)/s1)^2;
H(11,11)=(lamda1)^2;
H(12,12)=((s2*lamda1*lamda2)/s3)^2;
H(13,13)=((s2*lamda1*lamda2)/s4)^2;
H(14,14)=((s2*lamda1*lamda2)/s5)^2;
H(15,15)=((s2*lamda1*lamda2)/s6)^2;
H(16,16)=((s2*lamda1*lamda2)/s7)^2;

%third equation
H(17,17)=(s3*lamda4)^2;
H(18,18)=((s3*lamda1*lamda2)/s1)^2;
H(19,19)=((s3*lamda1*lamda2)/s2)^2;
H(20,20)=(lamda1)^2;
H(21,21)=((s3*lamda1*lamda2)/s4)^2;
H(22,22)=((s3*lamda1*lamda2)/s5)^2;
H(23,23)=((s3*lamda1*lamda2)/s6)^2;
H(24,24)=((s3*lamda1*lamda2)/s7)^2;
%....and analogously for remaining

end
```

C.2 Sampling of Threshold

Listing C.2: Function Get threshold

```
function[t]=get_threshold(Y,X,Ystar)
%returns t=mean(threshold)
k=1;
Reps=20000;
burn=16000;
L=2;
N=cols(Y);
ncrit=(N*L+1);
tarscale=0.1;
tarmean=mean(Ystar);  %mean of the prior on the threshold is the mean value of the threshold
    variable
tarvariance=10;           %prior variance for threshold
%specify the prior mean of the coefficients

%prior scale matrix for sigma of the VAR covariance
S=eye(N);
%prior degrees of freedom
alpha=N+1;
tar=tarmean; %initial value of the threshold
tarold=tar;
naccept=0;
```

```
tmat=zeros(1,Reps burn);

%starting values for the Gibbs sampling algorithm (Minnesota prior)
Sigma1=eye(N);
Sigma2=eye(N);
B0=zeros((N*L+1),N);
for i=1:N
B0(i+1,i)=1;
end
B0=vec(B0);

for j=1:Reps

        %separate sample
        e1=Ystar<=tar;
        e2=Ystar>tar;
        Y1=Y(e1,:);
        X1=X(e1,:);
        Y2=Y(e2,:);
        X2=X(e2,:);

        %Regime 1
        [H,betaols]=get_prior(Y1,X1,2);
        T=rows(X1);
        M=inv(inv(H)+kron(inv(Sigma1),X1'*X1))*(inv(H)*B0+kron(inv(Sigma1),X1'*X1)*betaols);
        V=inv(inv(H)+kron(inv(Sigma1),X1'*X1));

        %check for stability of the VAR
        check= 1;
        while check<0
                %update of prior (equation 10)
                beta1=M+(randn(1,N*(N*L+1))*chol(V))';
                CH=stability(beta1,N,L);
                if CH==0
                        check=10;
                end
        end

        e=Y1 X1*reshape(beta1,N*L+1,N);
        %scale matrix
        scale=e'*e+S;
        Sigma1=IWPQ(T+alpha,inv(scale));

        %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        %and the same for regime 2 (not here)
        %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

        %continue with sampling of threshold
        tarnew=tarold+randn(1,1)*sqrt(tarscale);
        postnew=getvarpost(Y,X,beta1,beta2,Sigma1,Sigma2,L,tarnew,tarmean,tarvariance,Ystar,ncrit);
        postold=getvarpost(Y,X,beta1,beta2,Sigma1,Sigma2,L,tarold,tarmean,tarvariance,Ystar,ncrit);

        accept=exp(postnew postold);
        u=rand(1,1);
        if u<accept
                tarold=tarnew;
                naccept=naccept+1;
        end
        tar=tarold;
        arate=naccept/j;
        if j>100 && j<5100              %tuning of scale for threshold sampling
                if arate<0.2
                        tarscale=tarscale*0.99;
```

```
                elseif arate>0.4
                    tarscale=tarscale*1.01;
                end
        end
        % partinion of data according to threshold for new iteration
        if j>burn
                %smat1(:,k)=e1;
                %smat2(:,k)=e2;
                tmat(k)=tar;
                k=k+1;
        end
end
t=mean(tmat);
end
```

C.3 Forecasting

Equivalent implementation is for plain BVAR(L) models when threshold is discarded

Listing C.3: Forecast with SETAR(2,2,2)

```
%... initialize prior and data as before
L=2;  %lag length of the model
tard=2;  %delay of threshold value
tarvar=1;  % threshold variable is the column number tarvar in data
tarscale=0.1;  %scaling parameter for RW Metropolis algorithm
Reps=50000;
burn=40000;
N=size(data,2);
forecast_period=12;              %12 step ahead forecast
n=rows(data);
S=eye(N);          %prior scale matrix
alpha=N+1;         %prior degrees of freedom
%setting of prior
B0=zeros((N*L+1),N);
for i=1:N
B0(i+1,i)=1;
end
B0=vec(B0);

%for storage of output
threshold=zeros(forecast_period,n);
out1=[];
out2=[];
out11=[];
out12=[];
R1=zeros(n+forecast_period,forecast_period);
R2=zeros(n+forecast_period,forecast_period);

for period=50:n          %out of sample forecast for period 50 90
        %prepare data
        Y=data(1:period,:);
        X=[ones(size(Y,1),1) lag0(Y,1) lag0(Y,2) ];
        Ystar=lag0(Y(:,tarvar),tard);
        Y=Y(max([L,tard(1)])+1:end,:);
        X=X(max([L,tard(1)])+1:end,:);
        Ystar=Ystar(max([L,tard(1)])+1:end,:);
        [t]=get_threshold(Y,X,Ystar);              %sample threshold
```

```
        threshold(1,period)=t;
        %partition data
        e1=Ystar<=t;
        e2=Ystar>t;
        Y1=Y(e1,:);
        X1=X(e1,:);
        Y2=Y(e2,:);
        X2=X(e2,:);
        T1=rows(X1);
        T2=rows(X2);
        %get priors
        [H1,betaols1]=get_prior(Y1,X1,L);
        [H2,betaols2]=get_prior(Y2,X2,L);
        Sigma1=eye(N);
        Sigma2=eye(N);
        for j=1:Reps                    %start sampling
            M=inv(inv(H1)+kron(inv(Sigma1),X1'*X1))*(inv(H1)*B0+kron(inv(Sigma1),X1'*X1)*betaols1);
            V=inv(inv(H1)+kron(inv(Sigma1),X1'*X1));

            %check for stability of the VAR
            check= 1;
            while check<0
                %update of prior (equation 10)
                beta1=M+(randn(1,N*(N*L+1))*chol(V))';
                CH=stability(beta1,N,L);
                if CH==0
                    check=10;
                end
            end
            e=Y1 X1*reshape(beta1,N*L+1,N);
            scale=e'*e+S;
            Sigma1=IWPQ(T1+alpha,inv(scale));
            %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
            %...same for regime  2
            %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

            if j>burn       %store outup after burn in period
                %forecast GDP growth and inflation for 3 years
                A01=chol(Sigma1);
                A02=chol(Sigma2);
                yhat1=zeros(2+forecast_period,N);
                yhat1(1:2,:)=Y1(end 1:end,:);
                yhat2=zeros(2+forecast_period,N);
                yhat2(1:2,:)=Y2(end 1:end,:);
                for i=3:(2+forecast_period)             %actual computation of forecasts
                    yhat1(i,:)=[1 yhat1(i 1,:) yhat1(i 2,:)]*reshape(beta1,N*L+1,N)+randn(1,N)*A01;
                    yhat2(i,:)=[1 yhat2(i 1,:) yhat2(i 2,:)]*reshape(beta2,N*L+1,N)+randn(1,N)*A02;
                end

                out1=[out1 yhat1(3:end,1)];
                out2=[out2 yhat2(3:end,1)];
            end
        end
        for k=1:forecast_period         %computations mean forecasts throughout the estimation to save RAM
            R1(period,k)=mean(out1(k,:),2);
            R2(period,k)=mean(out2(k,:),2);
        end
        out11{period,1}=out1;
        out12{period,1}=out2;
end
```

Listing C.4: Marginal likelihood computation for Minnesota prior

```
%get mean value of parameters
betam=squeeze(mean(outbeta,1));
sigmam=squeeze(mean(outsigma,1));
%evaluate priors
b0=B0;
b01=reshape(b0,N*L+1,N);
e0=Y X*b01;
S=eye(N);
%evaluate log prior distribution for VAR coefficients
bp=multivariatenormal(betam',B0,H);
%evaluate log prior for VAR covariance
sp= invwishpdf(sigmam,S,T+alpha);
%evaluate log likelihood
lik=loglik(reshape(betam,N*L+1,N),sigmam,Y,X);
%evaluate H(Bstar\sigmastar);
vstar1=inv(inv(H)+kron(inv(sigmam),X'*X));
Mstar=inv(inv(H)+kron(inv(sigmam),X'*X))*(inv(H)*B0+kron(inv(sigmam),X'*X)*betaols);

H1=multivariatenormal(betam',Mstar,vstar1);
%evaluate H(sigmastar\beta[j])
H2i=[];
for j=1:size(outbeta,1)
        betaj=outbeta(j,:);
        e=Y X*reshape(betaj,N*L+1,N);
        scale=e'*e;
        H2i= [H2i;invwishpdf(sigmam,S+scale,T+alpha)];
end
%take mean taking care of possible underflow/overflow with exp
factor=max(H2i);
H2=exp(H2i factor);
H2m=mean(H2);
H2m=log(H2m)+factor;

mlik=lik+bp+sp H1 H2m; %marginal likelihood
```