

Memoire

Lukas Gahl

May 23, 2023

Abstract

Contents

1	Introduction	5
2	Literature Review	6
3	DSGE Models	7
3.1	Real Business Cycle model	8
3.1.1	The Household	8
3.1.2	The firm	9
3.1.3	Equilibrium	10
3.2	NK	11
3.2.1	Consumption bundle	12
3.2.2	Household	13
3.2.3	Price setting	14
3.2.4	The firm	15
3.2.5	Equilibrium conditions	16
3.2.6	Log linearisation	17
3.2.7	New Keynesian Phillips and IS curve	18
3.2.8	Monetary policy	19
3.2.9	State Space	20
3.3	Petrol	20
3.3.1	Bundler	20
3.3.2	Household	21
3.3.3	The Firm	22
3.3.4	Equilibrium	22
3.3.5	Price setting	23
3.3.6	Log linearisation	24
4	Data	25
4.1	Preprocessing	25
4.2	Analysis of models	26
5	Bayesian Estimation	30
5.1	The log-likelihood	32

5.2	Sampler	35
5.3	Priors	37
5.3.1	Steady state priors	38
5.3.2	Exogenous prior	39
5.3.3	Endogenous priors	40
6	Forecast evaluation	42
6.1	DSGE forecasts	42
6.1.1	RBC	42
6.1.2	NK	43
6.1.3	NKE	44
6.1.4	BVAR	44
6.1.5	Forecast comparison	45
7	Conclusion	47
7.0.1	Conclusion on MH MCMC	47
	References	49
7.1	Appendix E	61

1 Introduction

High correlation of hours worked and real wage in the RBC model, an fact that is not matched by reality. The aim of RBC research thus must be finding new kinds of shocks, that allow for a more realistic representation of labour supply (Christiano Eichenbaum 1991).

One such 'new' shock is the inclusion of petrol as an exogenous shock series [Kim and Loungani, 1992]

- The method of solving and approximating the model has significant impact on simulated data, hence is relevant when relating models to real data (Taylor Uhlig, 1990)

This work is joining 3 strains of literature - Literature on the RBC, the origin of all new keynesian models - Literature on bayesian estimation of structural macroeconomics models - Literature on energy price shocks and their role in the - This refers to the nature of the shock - Shocks as theoretical origin on business cycles are at the origine of this research, identifying suitable shocks that explain more thus adds to the literature on business cycles

The Lucas 1976 critique:

Intertamperol evolution of macreoconomic variables

Kydland and Prescott (1982) introduced the RBC

- became main model of macroeconomy - rational expectation model

The New Keynesian model extends the RBC by inflation -

Rational expectation models build on two corner stones Chair et al - Structural parameters which are unaffected by policy changes - shocks that have a relevant economic interpretation

Kocherlakota (2007) adjustment principle - the better a model fits to data the better it is for policy advise - in assessing this fit one needs to be wary of over fitting specific data

2 Literature Review

RBC Criticism the correlation of real wage and hours worked does not correspond to reality, this is the lacmus test of RBC models (Christiano Eichenbaum 1992) "Robert E. Lucas (1981 p. 226) says that "observed real wages are not constant over the cycle, but neither do they exhibit consistent pro- or countercyclical tendencies. This suggests that any attempt to assign systematic real wage movements a central role in an explanation of business cycles is doomed to failure" [Christiano and Eichenbaum, 1992]

NK models are suitable to comparing alternative policy regimes without being subject to the Lucas (1976) critique (Gali)

3 DSGE Models

Before diving into the derivation of each individual model this section provides some general remarks applicable to all models across this work. All models discussed in the below have analytical solutions to their steady state, in which the economy is at its natural optimum. However, the system describing the economy is non-linear due to the multiplicative production relationship arising from the Cobb-Douglas production function [Campbell, 1994]. This gives rise to the need of some form of linear approximation around the steady state in order to be able to analyse the system's behaviour. The most common approach adopted across literature is a log-linear Taylor approximation, this work will follow suit [Campbell, 1994]. Such linearisation is applicable under the existence of a deterministic singular solution, which is the case for the small scale DSGE models discussed in the below [Herbst and Schorfheide, 2016].

The theoretical benchmark model a classic RBC The aim of analysing business cycles

The micro foundation, instead of clear behavioural rules agents optimise their behaviour according to utility optimisation

Conceptual ideas: The efficiency of business cycles (Gali, 2008) in a world of perfect competition and lack of nominal rigidities business cycles might be efficient they are the response to changes in productivity, and are actually the result of a "correcting" force towards an efficient equilibrium this raises questions on the desirability of policy interventions This goes against (Keynes, 1936) who regarded recession as inefficient, due to under-utilisation of resources

Technology shocks Technological shocks for a correctly calibrated model allowed to simulate cycles similar to actual business cycles in first and second order moment (ref). This shed new light on the assumption that technology was solely a driver of long-term growth with negligible impact on business cycles (Gali, 2008, p.3).

RBCs successfully abstracted from monetary rigidities in their explanation of business cycles

The assumptions of the RBC and their consequences, namely the non-existence of money and monetary rigidities in the economy are greatly contrasted by empirical

evidence (Gali, 2008, p 15)

Assumptions - perfect compitions - flexible prices - technology shocks - identical technology across firms

3.1 Real Business Cycle model

The Real Business Cycle model (RBC) was first introduced by [Prescott, 1986] in the aim of providing an explanation to business cycles. Departing from Schumpeterian growth driven by advances in technology the RBC model is at its heart a de-trended growth model. It emphasis the role of technology growth and resulting changes in total factor productivity in influencing business cycles. In doing so it makes the following assumptions:

- perfect competition
- flexible prices
- technology shocks
- identical technology across firms

This leads to the interpretation of business cycles as efficient adjustments of economy towards a new equilibrium. While this idea is antithetical to the New-Keynesian literature considering business cycles as inefficient both models share a the mathematical foundation. The following will therefore derive the RBC model

3.1.1 The Household

The households in the RBC model are infinitely lived agents with homogeneous preferences. They simultaneously solve the same threefold optimisation problem. Faced with an income consisting of labour wage, returns on last periods capital and dividends from firms. The income is divided between consumption, capital investment and leisure. The leisure expenditure is implicitly made by adjusting the hours of work, factoring negatively into household utility. Dividends are drawing from the fact that in the RBC firms are owned by the consumer. However, the assumption of full price flexibility and perfect competition implies zero firm profits at equilibrium, dividend incomes are thus null in the steady state. The following will assume a separable utility

function of the form:

$$U(C_t, L_t) = \frac{C_t^{1-\sigma}}{1-\sigma} - \frac{L_t^{1+\phi}}{1+\phi} \quad (1)$$

In this specification σ is a coefficient of risk aversion and ϕ is the marginal disutility of labour. While other utility specification have been employed across the RBC literature this work will remain with a separable utility function in order to be consistent with the New-Keynesian literature. The function $U(C_t, L_t)$ exhibits the usual characteristics of well behaved utility function $\frac{\partial U}{\partial C_t} > 0$, $\frac{\partial U}{\partial L_t} \leq 0$ and $\frac{\partial^2 U}{\partial C_t^2} \leq 0$, $\frac{\partial^2 U}{\partial L_t^2} \leq 0$. The household maximisation problem then reads as follows

$$\begin{aligned} & \max_{B_t, C_t, L_t} \sum_{t=0}^{\infty} \beta^t U(C_t, L_t) \\ \text{s.t. } & C_t + I_t = R_t K_{t-1} + W_t N_t + \Pi_t \\ & K_t = (1 - \delta) K_{t-1} + I_t \\ & \text{FOC:} \\ & -\frac{U_{n,t}}{U_{c,t}} = W_t \end{aligned} \quad (2)$$

$$\mathbb{E}_t \left[\frac{U_{c,t}}{U_{c,t+1}} \right]^\sigma = \beta [1 - \delta + \mathbb{E}_t[R_{t+1}]]$$

The problem yields two main equations. The Euler equation of intertemporal consumption trade-off and the labour supply equation.

$$\begin{aligned} & C_t^\sigma L_t^\phi = W_t \\ & \mathbb{E}_t \left[\frac{C_{t+1}}{C_t} \right]^\sigma = \beta [1 - \delta + \mathbb{E}_t[R_{t+1}]] \end{aligned} \quad (3)$$

3.1.2 The firm

Firms in the RBC are also assumed to be homogenous, producing with the same production technology. As such they solve the same static optimisation problem across its production inputs capital and labour. Due to the assumption of perfect competition and flexible prices firms pay the marginal product of labour and capital to its respective inputs. This implies a zero-profit condition, as mentioned in the above. The two first

order conditions thus correspond to capital and labour demand.

$$\begin{aligned}
& \max_{N_t} A_t K_t^\alpha N_t^{1-\alpha} \\
& \text{s.t. } Y_t = W_t N_t + K_t R_t \\
& \text{FOC:} \\
& W_t = (1 - \alpha) A_t K_t^\alpha N_t^{-\alpha} \\
& K_t = \alpha A_t K_t^{\alpha-1} N_t^{1-\alpha}
\end{aligned} \tag{4}$$

3.1.3 Equilibrium

Having derived the above optimality condition an analysis of the equilibrium requires two further equations. These are the equilibrium resource constraint arising from the output accounting identity and the exogenous law of motion of technology.

$$\begin{aligned}
Y_t &= C_t + I_t \\
\log(A_t) &= (1 - \rho_A) \log(A^*) + \rho_A \log(A_{t-1}) + \epsilon_t^A
\end{aligned} \tag{5}$$

As the case with all DSGE models the system in its current state is non-linear. This is due to the from multiplicative Cobb-Douglas production and additive law of motion without full depreciation $\delta \leq 1$ [Campbell, 1994]. This raises the need for log-linear approximations which yields the following system of equations.

Equation name	Log-linear expression
Euler equation	$\hat{c}_t = \hat{c}_{t+1} - \hat{r}_{t+1}$
Capital demand	$\hat{y}_t - \hat{k}_t = \hat{r}_t$
Capital supply	$\hat{i}_t = \delta \hat{k}_{t+1} - \frac{1-\delta}{\delta} \hat{k}_t$
Labour supply	$\hat{c}_t = \hat{w}_t - \frac{\bar{l}}{1-\bar{l}} \hat{l}_t$
Labour demand	$\hat{y}_t - \hat{l}_t = \hat{w}_t$
Production function	$\hat{y}_t = \hat{a}_t + \alpha \hat{k}_t + (1 - \alpha) \hat{l}_t$
Equilibrium condition	$\hat{y}_t = \frac{\bar{y}}{\bar{c}} \hat{c}_t + \frac{\bar{l}}{\bar{y}} \hat{l}_t$
Technology law of motion	$\hat{a}_t = \rho_a \hat{a}_{t-1} + \epsilon_{a,t}$

The above linearisation still contains values of the steady state indicated by the \bar{x} notation. In order to express the model as a linear system these have to be replaced with their respective deterministic values. However, the RBC model under separable

consumption labour utility, does not possess a deterministic steady state for labour (ref). This issues has often been circumvented by relying on a logarithmic specification of household utility [Campbell, 1994]¹. However, the purpose of this work is a comparison between an RBC and New-Keynesian models. The latter traditionally rely on separable utility and so will this work will in order to allow for a fair comparison of results. The usual solution path for an RBC with separable utility is the computation of the steady state as ratios of labour. These 'great ratios' are deterministic and allow for some analysis [Prescott, 1986].

Variable	Deterministic steady state
Rate of return	$\bar{R} = \frac{1}{\beta} + \delta - 1$
Capital labour ratio	$\frac{\bar{K}}{\bar{L}} = \left[\frac{\alpha \bar{A}}{\frac{1}{\beta} + \delta - 1} \right]^{\frac{1}{1-\alpha}}$
Wage	$\bar{W} = (1 - \alpha) \bar{A} \left(\left[\frac{\alpha \bar{A}}{\frac{1}{\beta} + \delta - 1} \right]^{\frac{1}{1-\alpha}} \right)^{\alpha}$
Investment labour ratio	$\frac{\bar{I}}{\bar{L}} = \delta \left[\frac{\alpha \bar{A}}{\frac{1}{\beta} + \delta - 1} \right]^{\frac{1}{1-\alpha}}$
Output labour ratio	$\frac{\bar{Y}}{\bar{L}} = \bar{A} \left(\left[\frac{\alpha \bar{A}}{\frac{1}{\beta} + \delta - 1} \right]^{\frac{1}{1-\alpha}} \right)^{\alpha}$
Consumption labour ratio	$\frac{\bar{C}}{\bar{L}} = \frac{\bar{Y}}{\bar{L}} - \frac{\bar{I}}{\bar{L}}$

Obtaining a linear state-space representation of the RBC model requires the numerical approximation of the steady state value of labour \bar{l} . This work relies on the software package gEconpy (ref) for this numerical approximation. The specification of the input file, can be found in 7.1. The resulting steady-state values and the state-space transition matrix can be found in the 7.0.1.

3.2 NK

The New-Keynesian model developed in this section is based on the work of [Galí, 2008]. It diverges from the RBC model in three key aspects. Firstly, the NK model in its simplest form analyses an economy without capital, wherefore its intertemporal instrument is the output gap instead of the capital stock (ref). Secondly, the NK model introduces nominal rigidities through monopolistic competition. Since firms are endowed with

¹In the work of Campbell the analytical derivation relies on a log-utility function

some market power, they can reset prices above marginal cost, introducing a gap between real and nominal marginal cost. This dynamic results in inflation. This extension has several consequences. For one monetary policy is neutral in the short-run, thus changes in interest are not directly matched by changes in expected inflation (ref). This gap in expectations becomes the driver of short-run fluctuation, or business cycles in the economy. The main driver of business cycles is thus no-longer technological progress through a change in total factor productivity. Instead an adjustment or shock to the interest rates drives business cycles.

This of course has important economic implications. Instead of efficient direct adjustments of the economy to new production technology business cycles no longer exhibit efficiency in the NK model. Instead, they are an inefficient divergence caused by central bank interventions. This is also the reason why the model is called New-Keynesian. The model returns Keynes's perspective, regarding business cycles as inefficient. The main divergence from RBC literature consists in this change of perspective.

3.2.1 Consumption bundle

The New-Keynesian framework introduces monopolistic competition in order to equip firms with price setting power (ref). The market power arises from the fact that goods are not perfect substitutes. In this setting, the simplified assumption of a unique good of the RBC holds no longer and the household's allocation problem becomes a two stage process. First, the household needs to decide on the basket of products that optimises its utility. Only then the household is able to solve for the optimal allocation of income in the consumption-leisure trade-off. This requires the consumption basket composition problem to be solved prior to the consumption-leisure trade-off. Across the DSGE literature the aggregation of consumption is also sometimes referred to as the output of a good bundling firm. This intermediary party determines the optimal basket of goods according to the household preferences. In a perfectly competitive environment the household buys the optimal bundle at no mark up beyond the producer's mark-up. Determining the optimal basket requires some function of aggregation across the continuum of goods. The most commonly employed aggregator is the Dixit-Stiglitz function. This function has the appealing property of a constant elasticity of substitu-

tion (CES). This is important later in deriving marginal cost and its equilibrium value. The optimisation problem reads as follows where each good C_{it} is matched by its price P_{it} purchased with budget Z_t .

$$\begin{aligned} \max_{C_{it}} \quad & \left(\int_0^1 C_{it}^{\frac{\epsilon-1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}} \\ \text{s.t.} \quad & \int_0^1 P_{it} C_{it} di \leq Z_t \end{aligned} \quad (6)$$

Solving the above optimisation problem results in three important expression. First, the aggregated price level P_t of the optimal consumption bundle is derived. This corresponds to the consumer price index (CPI).

$$P_t = \left(\int_0^1 P_{it}^{1-\epsilon} di \right)^{\frac{1}{1-\epsilon}} \quad (7)$$

Further manipulation leads to the share of product C_{it} in the consumption basket C_t . This share is defined by the goods price relative to the price level P_t and the CES between goods ϵ .

$$C_{it} = \left(\frac{P_{it}}{P_t} \right)^{-\epsilon} C_t \quad (8)$$

Lastly, using the two above it can be demonstrated that under a binding budget constraint individual goods prices correspond to the CPI times the consumption bundle. This is the case as goods are weighted according to their relative price. This property is important as it allows to solve the household and firm optimisation problem as a representative agent problem.

$$\int_0^1 P_{it} C_{it} di = Z_t = P_t C_t \quad (9)$$

3.2.2 Household

Having found C_t and its price P_t the household's consumption-leisure allocation can be formulated. It follows straightforwardly from the RBC model with one extension, bond holdings. While zero in equilibrium bond holdings introduces the interest rate to the households problem. This is an important extension in order to allow for a monetary authority. The interest rate is contained in the discount factor $Q_t = \frac{1}{1+i_t}$. Similar to the RBC households are infinitely lived solving this problem for all future

periods.

$$\begin{aligned}
& \max_{B_t, C_t, L_t} \sum_{t=0}^{\infty} \beta^t U(C_t, L_t) \\
& P_t C_t + Q_t B_t = B_{t-1} + W_t N_t \\
& \text{FOC:} \\
& -\frac{U_{n,t}}{U_{c,t}} = \frac{W_t}{P_t} \\
& Q_t = \beta \mathbb{E}_t \left[\frac{U_{c,t+1}}{U_{c,t}} \frac{P_t}{P_{t+1}} \right]
\end{aligned} \tag{10}$$

Solving the optimisation problems yields the labour supply, this time explicitly depending on the real wage and the Euler equation. Assuming the function form of $U(C_t, L_t) = \frac{C_t^{1-\sigma}}{1-\sigma} - \frac{L_t^{1+\phi}}{1+\phi}$ and plugging its derivatives into the first order condition yields.

$$\begin{aligned}
& C_t^\sigma L_t^\phi = \frac{W_t}{P_t} \\
& Q_{t,t+k} = \mathbb{E}_t \left[\frac{C_{t+k}}{C_t} \right]^\sigma \frac{P_t}{P_{t+1}}
\end{aligned} \tag{11}$$

3.2.3 Price setting

The main innovation of the NK model lies in modelling inflation. As mentioned in the above inflation arises from the imperfect substitutability consumption goods. This results in monopolistic competition endowing firm's price setting power. Modelling the price setting process is thus at the heart of of the NK model. In doing so the standard model relies on the Calvo-pricing mechanism. Firms are able to reset prices in any period with the probability θ . This is equal to saying that every period a share θ of firms can reset prices to their optimal price \hat{P}_t while $1 - \theta$ firms have to remain with last periods price P_{t-1} . This allows to describe the evolution of the price level as an in the following intertemporal equation, which if divided by P_{t-1} yields an expression of inflation.

$$P_t = \left[\theta P_{t-1}^{1-\epsilon} + (1 - \theta) \hat{P}_t^{1-\epsilon} \right]^{\frac{1}{1-\epsilon}} \tag{12}$$

$$\Pi_t^{1-\epsilon} = \theta + (1 - \theta) \left(\frac{\hat{P}_t}{P_{t-1}} \right)^{1-\epsilon} \tag{13}$$

The key aspect to price evolution thus lies in the optimal price \hat{P}_t . The further it is away from last periods price, the higher is inflation.

3.2.4 The firm

In the RBC the firm solved a static optimisation problem providing labour and capital demand. In the NK model they are instead faced with setting the optimal price \hat{P}_t . For this the NK model assumes that a continuum of firms each producing one good provide factoring into the continuum of goods as outlined in (section). Setting the optimal price becomes an intertemporal problem as firms trade-off between two opposing forces. As illustrated by (ref) the share of any good in the overall basket depends on its price relative to the CPI. This implies that setting a price far above P_t leads to a lower demand for a firm's single good $C_{i,t}$. On the other hand, as firms are limited in their price frequency they need to set a price which they would be comfortable with in the future. Under the Calvo-scheme a price once set will be unchanged with probability θ^k for k periods. Consequently, if the price once set is too low it will at some point allow for no further profits potentially even incur losses. Firms thus try to set a price that allows them an acceptable share in the basket of goods while also guaranteeing profitability across future periods. In this decision the size of $\theta \in [0, 1]$ combined with the time discount factor $Q_t\beta$ weights the second force against the first. The problem can be formalised as a future discounted sum of revenue minus total cost. The notation of $Y_{t+k|t}$ hereby refers to the income in period $t+k$ at price of period t . The notation of $Q_{t,t+k}$ is the stochastic discount factor of today's discount carried forward into period $t+k$. The budget constraint makes use of the fact that under equilibrium assumption all output is consumed as part of the consumption basket $Y_{it} = C_{it}$. (section) will deliver a more in-depth perspective on this.

$$\begin{aligned}
& \max_{\hat{P}_t} \sum_{k=0}^{\infty} \theta^k \mathbb{E}_t \left[Q_{t,t+k} \left(\hat{P}_t Y_{t+k|t} - TC_{t+k|t}^n(Y_{t+k|t}) \right) \right] \\
& \text{s.t.} \quad Y_{it+k|t} = \left(\frac{\hat{P}_t}{P_{t+k}} \right)^{-\epsilon} C_{t+k} \\
& \text{FOC:} \quad \sum_{k=0}^{\infty} \theta^k \mathbb{E}_t \left[Q_{t,t+k} Y_{t+k|t} \left(\hat{P}_t - \frac{\epsilon}{1-\epsilon} MC_{t+k|t}^n \right) \right] = 0
\end{aligned} \tag{14}$$

Inserting the budget constraint and the discount factor as defined by the Euler condi-

tion 11 results in:

$$\frac{\hat{P}_t}{P_t} = \frac{\epsilon}{\epsilon - 1} \frac{\mathbb{E}_t \sum_{k=0}^{\infty} \theta^k \beta^k C_{t+k}^{1-\sigma} \left(\frac{P_{t+k}}{P_t}\right)^\epsilon MC_{t+k}^r}{\mathbb{E}_t \sum_{k=0}^{\infty} \theta^k \beta^k C_{t+k}^{1-\sigma} \left(\frac{P_{t+k}}{P_t}\right)^{\epsilon-1}} \quad (15)$$

The above describes the mark-up of the optimal price \hat{P}_t beyond the price level CPI. The difference between the two directly drives inflation as it implies a higher price level as can be seen in ???. In the steady state inflation is zero by definition, implying that no firm resets prices. This is equivalent to saying that $\theta = 0$. Multiplying the above by P_t and setting $\theta = 0$ results in $\hat{P}_t = \frac{\epsilon}{\epsilon-1} MC_t^n$. This allows to derive an important aspect to the NK model, the price-mark up over nominal marginal cost in the steady state. Firms are endowed with market power, allowing them to set a mark up above marginal cost. The real marginal defined as $MC_t^r \equiv \frac{MC_t^n}{P_t}$ cost can thus be derived. It is important to note that MC^r are the real marginal cost in the steady state, wherefore they are time independent. They are complemented by MC_t^r , the time depended marginal cost, which can deviate from their steady state value leading to the gap of $\hat{MC}_t^r \equiv MC_t^r - MC^r$.

$$MC^r = \frac{\epsilon}{\epsilon - 1} \quad (16)$$

3.2.5 Equilibrium conditions

To this point a number of optimal choices have been derived. The households decides on the consumption basket and consumption-leisure trade-off. The firm sets the optimal price and consequently mark-up over nominal marginal costs, yielding an expression for inflation. Having derived the optimality conditions for all agents the model now requires market clearing condition to describe their interaction. This is based on accounting identity on individual good and production level $Y_{it} = C_{it}$. In order to close the model one needs to demonstrate, that this property on individual goods level translates in the global property $Y_t = C_t$.

Output Y_t is defined as an aggregator function. The accounting identity combined with 8 allows to link Y_t to the the price level, allowing to demonstrate that $Y_t = C_t$ holds.

$$Y_t = \left(\int_0^1 Y_{it}^{\frac{\epsilon-1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}} = \left(\int_0^1 \left[\left(\frac{P_{it}}{P_t} \right)^\epsilon C_t \right]^{\frac{\epsilon-1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}} = P_t^\epsilon C_t P_t^{-\epsilon} = C_t \quad (17)$$

A similar reasoning is required in order to derive an economy wide production function an expression from the firm's individual production functions $Y_{it} = A_t N_{it}$. This yields the below expression, which will be kept unchanged for now.

$$N_t = \int_0^1 \left(\frac{Y_t}{N_t} \right)^{\frac{1}{1-\alpha}} \int_0^1 \left(\frac{P_{it}}{P_t} \right)^{-\frac{\epsilon}{1-\alpha}} di \quad (18)$$

3.2.6 Log linearisation

This completes the number of equations required for the derivation of the New Keynesian model. As the above equations are a non-linear system without a closed form solution some method of linearisation is required. The below will follow the procedure of [Galí, 2008] and apply log-linearisation first order Taylor approximations. Throughout the log-linearisation several manipulations are required. The Euler equation relies on reasoning about the non-inflationary steady state so does the expression for N_t . Explanation and proofs can be found in 7.0.1. Lower-case expressions are the log-equivalent to previous upper case variables.

Euler equation

$$c_t = \mathbb{E}_t[c_{t+1}] - \frac{1}{\sigma}(i_t - \rho - \mathbb{E}_t[\pi_{t+1}]) \quad (19)$$

Labour supply

$$w_t - p_t = \sigma c_t + \phi n_t \quad (20)$$

Inflation

$$\pi_t = (1 - \theta)(\hat{p}_t - p_{t-1}) \quad (21)$$

Optimal price

$$\hat{p}_t = (1 - \theta\beta)\mathbb{E}_t \left[\sum_{k=0}^{\infty} \theta^k \beta^k (mc_{t+k|t}^r - mc^r + p_{t+k}) \right] \quad (22)$$

Production function

$$y_t = a_t + (1 - \alpha)n_t \quad (23)$$

Equilibrium condition

$$y_t = c_t \quad (24)$$

3.2.7 New Keynesian Phillips and IS curve

The next step is the derivation of an expression for real marginal cost mr_t^r contained in the optimal price setting. In the absence of other production factors the marginal product of labour $MPN_t = \frac{\partial Y}{\partial N} = A_t(1 - \alpha)N_t^{-\alpha}$ corresponds to nominal marginal cost mc_t^n . Real marginal cost can thus be obtained from $MC_t^r = \frac{W_t}{P_t MCN_t}$. This allows to formulate the below property in logs by including the production function 23.

$$\begin{aligned} mc_t^r &= w_t - p_t - a_t - mpn_t \\ mc_t^r &= w_t - p_t - \frac{a_t - \alpha y_t}{1 - \alpha} - \ln(1 - \alpha) \end{aligned} \quad (25)$$

This expression is at the provides the basis for two essential conclusions. Firstly, it allows to derive a relation between mc_t^r and output y_t . Using labour supply 20 and the production function 23 one can rephrase the above as.

$$\begin{aligned} mc_t^r &= \sigma c_t + \phi n_t - [a_t + \alpha n_t + \ln(1 - \alpha)] \\ mc_t^r &= \frac{\sigma(1 - \alpha) + \phi + \alpha}{(1 - \alpha)} y_t - \frac{(1 + \phi)}{(1 - \alpha)} a_t + \ln(1 - \alpha) \end{aligned} \quad (26)$$

In the steady state marginal cost should be equal to its natural level and $mc_t^r = mc^r$. In this case the above describes natural output y_t^n . This fact can be used to derive an expression for the gap between marginal cost and its steady state value in logs $\hat{mc}_t^r \equiv mc_t^r - mc^r$. In doing so one Reformulating the above by replacing mc_t^r by mc^r yields an expression of the output gap in terms of the marginal cost gap.

$$\hat{mc}_t^r = \frac{\sigma(1 - \alpha) + \phi + \alpha}{1 - \alpha} (y_t - y_t^n) \quad (27)$$

The second insight arising from the real marginal cost is its translation in period $t + k$. Combining this with the share of consumption good i in the overall basket as well as the market clearing condition 24 logs gives

$$y_{t+k|t} = -\epsilon(p_{t+k|t} - p_{t+k}) + y_{t+k} \quad (28)$$

Combining the two yields an expression of marginal cost and the price level.

$$mc_{i,t+k|t}^r = mc_{t+k}^r - \frac{\epsilon\alpha}{1 - \alpha} (\hat{p}_t - p_{t+k}) \quad (29)$$

The above combined with the expression for the optimal price level.

$$\begin{aligned}
\hat{p}_t - p_{t-1} &= (1 - \theta\beta)\mathbb{E}_t \left[\sum_{k=0}^{\infty} \theta^k \beta^k (mc_{t+k|t}^r - mc^r + p_{t+k} - p_{t-1}) \right] \\
&= (1 - \theta\beta)\Theta \mathbb{E}_t \sum_{k=0}^{\infty} \theta^k \beta^k (mc_{t+k}^r - mc^r) + \sum_{k=0}^{\infty} \theta^k \pi_{t+k} \\
&= \theta\beta\mathbb{E}_t(\hat{p}_{t+1} - p_t) + (1 - \theta\beta)\Theta(mc_{t+k}^r - mc^r) + (1 - \theta)(\hat{p}_t - p_{t-1})
\end{aligned} \tag{30}$$

This can be rewritten in more compact form as ^{2 3}

$$\pi_t = \beta E_t[\pi_{t+1}] + \lambda \hat{m}c_t \tag{31}$$

The above combined with 27 allows to construct the New-Keynesian Phillips curve, on of the two core equations of the NK model.⁴

$$\pi_t = \beta E_t[\pi_{t+1}] + \kappa \tilde{y}_t \tag{32}$$

The second core equation of the New-Keynesian model is derived from the Euler equation 19, the equilibrium condition 24 and the log linear Fisher identity $r_t \equiv i_t - \mathbb{E}_t[\pi_{t+1}]$. Inserting y_t for c_t and introducing the difference between output and its natural level as $\tilde{y}_t \equiv y_t - y_t^n$ allows to write

$$\tilde{y}_t = \left[\mathbb{E}_t[y_{t+1}] - \frac{1}{\sigma}(i_t - \rho - \mathbb{E}_t[\pi_{t+1}]) \right] - \left[\mathbb{E}_t[y_{t+1}^n] - \frac{1}{\sigma}(r_t^n - \rho) \right] \tag{33}$$

This provides the New-Keynesian IS curve.

$$\tilde{y}_t = \mathbb{E}_t[\tilde{y}_{t+1}] - \frac{1}{\sigma}(i_t - r_t^n - \mathbb{E}_t[\pi_{t+1}]) \tag{34}$$

3.2.8 Monetary policy

Closing the model requires an interest rate rule, describing the behaviour of the central bank. This work will rely on a traditional Taylor rule with an exogenous interest rate shock following [Galí, 2008].

$$i_t = \rho + \phi_\pi \pi_t + \phi_y \tilde{y}_t + v_t \tag{35}$$

$$v_t = \rho_v v_{t-1} + \epsilon_t^v$$

²where $\Theta \equiv \frac{1-\alpha}{1-\alpha+\alpha\epsilon} \leq 1$

³where $\lambda \equiv \frac{(1-\theta)(1-\beta\theta)}{\theta} \Theta$

⁴where $\kappa = \lambda \frac{\sigma(1-\alpha)+\phi+\alpha}{1-\alpha}$

3.2.9 State Space

Having derived the New-Keynesian Phillips curve, New-Keynesian IS equation as well as a monetary policy rule allows to formulate the model as a linear state space. For this purpose

$$\begin{bmatrix} \tilde{y}_t \\ \pi_t \end{bmatrix} = \Omega \begin{bmatrix} \sigma & 1 - \beta\phi_\pi \\ \sigma\kappa & \kappa + \beta(\sigma + \phi_y) \end{bmatrix} * \begin{bmatrix} \tilde{y}_{t-1} \\ \pi_{t-1} \end{bmatrix} + \Omega \begin{bmatrix} 1 \\ \kappa \end{bmatrix} (\hat{r}_t^n - v_t) \quad (36)$$

where $\Omega \equiv \frac{1}{\sigma + \phi_y + \kappa\phi_\pi}$ and $\hat{r}_t^n \equiv r_t^n - \rho$.

3.3 Petrol

The below is an extension of the standard NK model by petrol as a consumption good as well as factor of production. There is no domestic production of petrol, instead it is imported and its price is assumed to be exogenous. Beyond the import of petrol there is no foreign trade, all closed economy assumptions thus hold. Given the exogenous nature of petrol the model is quite similar to the standard NK model. The below will therefore depart from (section) and only outline the changes to the model from petrol production.

3.3.1 Bundler

Petrol is considered a consumption good and as such enters into the good basket C_t . This inclusion is modelled across the bundling process, deriving the optimal basket across two classes of goods, domestic and foreign. The bundling of domestic goods follows from the same procedure as outlined in (section). Variables $C_{q,t}$ and $P_{q,t}$ are aggregated consumption and price respectively. The optimisation across the continuum of goods results in the same properties as above. These are share of good $C_{i,q,t}$ in the consumption basket $C_{q,t}$ and the aggregated price index $P_{q,t}$ as well as an expression of aggregate consumption.

$$P_{q,t} = \left(\int_0^1 P_{i,q,t}^{1-\epsilon} di \right)^{\frac{1}{1-\epsilon}} \quad (37)$$

$$C_{i,q,t} = \left(\frac{P_{i,q,t}}{P_{q,t}} \right)^{-\epsilon} C_{q,t} \quad (38)$$

$$\int_0^1 P_{it} C_{it} di = Z_t = P_t C_t \quad (39)$$

Foreign goods consumption $C_{m,t}$ exclusively consists of petrol and its share in overall consumption C_t is assumed to be constant. Parameter χ represents this share. This assumption, though a multiplication, is viable as consumption demand for petrol is relatively inelastic in the short-run. Empirical evidence suggests that due to the inability to replace durable goods (e.g. cars) that usually are at heart of consumption demand this demand is inelastic (ref). Overall consumption this is defined as ⁵

$$C_t \equiv \Theta_\chi C_{m,t}^\chi C_{q,t}^{1-\chi} \quad (40)$$

From this the overall price level of the consumption basket $P_{c,t}$ is derived, introducing $S_t \equiv \frac{P_{m,t}}{P_{q,t}}$ the real price of oil.

$$P_{c,t} \equiv P_{m,t}^\chi P_{q,t}^{1-\chi} = P_{q,t} \frac{P_{m,t}^\chi}{P_{q,t}^{1-\chi}} = P_{q,t} S_t^\chi \quad (41)$$

Following this reasoning and using 39 overall consumption can be written as

$$C_t P_{c,t} = C_{m,t} P_{m,t} + C_{q,t} P_{q,t} \quad (42)$$

3.3.2 Household

Having derived expressions for the consumption basket and its price the household's optimisation problem can be solved. As in the standard NK model this reads as follows.

$$\begin{aligned} & \max_{B_t, C_t, L_t} \sum_{t=0}^{\infty} \beta^t U(C_t, L_t) \\ \text{s.t. } & P_{c,t} C_t + Q_t B_t = B_{t-1} + W_t N_t \\ & \text{FOC:} \end{aligned} \quad (43)$$

$$\begin{aligned} & -\frac{U_{n,t}}{U_{c,t}} = \frac{W_t}{P_{c,t}} \\ & Q_t = \beta \mathbb{E}_t \left[\left(\frac{U_{c,t+1}}{U_{c,t}} \right)^\sigma \frac{P_{c,t}}{P_{c,t+1}} \right] \end{aligned}$$

The authors use a log-utility function $U(C_t, L_t) \equiv \log(C_t) - \frac{N_t^{1+\phi}}{1+\phi}$ which results in the following first order conditions.

$$\begin{aligned} & Q_t = \beta \mathbb{E}_t \left[\frac{C_t}{C_{t+1}} \frac{P_{c,t}}{P_{c,t+1}} \right] \\ & \frac{W_t}{P_{c,t}} = C_t N_t^\phi \end{aligned} \quad (44)$$

⁵where $\Theta = \chi^{-\chi}(1-\chi)^{(\chi-1)}$

3.3.3 The Firm

Petrol serving as a second factor of production the firm's optimisation problem is extended from the standard NK model. As in the RBC the firm obtains its production factor demand by equalising their marginal cost.

$$\begin{aligned} \max_{M_t} \quad & A_t M_{it}^{\alpha_m} N_{it}^{\alpha_n} \\ \text{s.t.} \quad & Y_{i,q,t} = W_t N_{i,t} + M_{i,t} P_{m,t} \\ \text{FOC:} \end{aligned} \tag{45}$$

$$\begin{aligned} W_{i,t} &= \alpha_n A_t M_{i,t}^{\alpha_m} N_{i,t}^{\alpha_n-1} \\ K_{i,t} &= \alpha_m A_t M_{i,t}^{\alpha_m-1} N_{i,t}^{\alpha_n} \end{aligned}$$

The real marginal cost can thus be expressed at the equality of marginal input cost of labour and petrol. This constitutes a deviation from the standard NK model, The the firms nominal marginal cost MC_t^n is therefore.

$$MC_t^n = \frac{W_{i,t}}{\alpha_n Y_{i,q,t} N_{i,t}^{-1}} = \frac{P_{m,t}}{\alpha_m Y_{i,q,t} M_{i,t}^{-1}} \tag{46}$$

Having obtained nominal marginal cost an expression for the firm's mark-up can be defined as $\xi_{i,t} \equiv \frac{P_{q,t}}{MC_{i,t}^n}$. With some further manipulation [Blanchard and Gali, 2007] uses this to obtain the demand for petrol.

$$\begin{aligned} \xi_{i,t} \frac{P_{m,t}}{P_{q,t}} M_{i,t} &= \alpha_m Q_{i,t} \frac{P_{i,q,t}}{P_{q,t}} \\ M_{i,t} &= \alpha_m \frac{Q_{i,t}}{S_t \xi_{i,t}} \frac{P_{i,q,t}}{P_{q,t}} \end{aligned} \tag{47}$$

3.3.4 Equilibrium

The equilibrium conditions follow the standard NK model in assuming the ressource constraint $Y_{i,q,t} = C_{i,q,t}$ on individual firm level. As showed in (section) of the standard NK model overall domestic this property also holds on the overall output level $Y_{q,t} = \left(\int_0^1 Y_{it}^{1-\frac{1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}} = C_{q,t}$.

Using this fact the

$$M_t = \frac{\alpha_m Y_{q,t}}{\xi_t^n S_t} \tag{48}$$

where ξ_t^n is the individual firm mark-up weighted by its share in output $Y_{q,t}$ which is given by the ratio of $\frac{P_{i,q,t}}{P_{q,t}}$.

As bond holdings are zero in equilibrium all allows to derive an expression for the overall price level, factoring out petrol.

$$P_{c,t}C_t = P_{q,t}C_{q,t} - P_{m,t}M_t = \left(1 - \frac{\alpha_m}{\xi_t^n}\right) P_{q,t}Y_{q,t} \quad (49)$$

[Blanchard and Gali, 2007] then go ahead and define the GDP deflator $P_{y,t}$ as

$$P_{q,t} \equiv P_{y,t}^{1-\alpha_m} P_{m,t}^{\alpha_m} \quad (50)$$

Moreover, they define the gross domestic product as

$$P_{y,t}Y_t \equiv P_{q,t}Y_{q,t} - P_{m,t}M_t = \left(1 - \frac{\alpha_m}{\xi_t^n}\right) P_{q,t}Y_{q,t} \quad (51)$$

This yields: and expression for petrol inputs M_t and expression for time dependent real mark-up $\log(\xi_t) \equiv \mu_t \equiv mc_t^r$

3.3.5 Price setting

The price setting problem of the firm is equivalent to the price setting in the standard NK model. It diverges in only one notion, the real marginal cost MC_t^r .

The expression for inflation is obtained as in the standard model

$$P_{q,t} = \left[\theta P_{q,t-1}^{1-\epsilon} + (1-\theta) \hat{P}_t^{1-\epsilon} \right]^{\frac{1}{1-\epsilon}} \quad (52)$$

The firm solves the same optimisation problem as in the above.

$$\begin{aligned} \max_{\hat{P}_t} \sum_{k=0}^{\infty} \theta^k \mathbb{E}_t \left[Q_{t,t+k} \left(\hat{P}_t Y_{t+k|t} - TC_{t+k|t}^n(Y_{t+k|t}) \right) \right] \\ \text{s.t.} \end{aligned} \quad (53)$$

$$Y_{it+k|t} = \left(\frac{\hat{P}_t}{P_{t+k}} \right)^{-\epsilon} C_{t+k}$$

with the first order condition

$$\sum_{k=0}^{\infty} \theta^k \mathbb{E}_t \left[Q_{t,t+k} Y_{q,t+k|t} \left(\hat{P}_{q,t} - \frac{\epsilon}{1-\epsilon} MC_{t+k|t}^n \right) \right] = 0 \quad (54)$$

$$\pi_{q,t} = \beta E_t[\pi_{q,t+1}] + \lambda \hat{m}c_t \quad (55)$$

The difference in the New-Keynesian Phillips curve lies in $\hat{m}c_t \equiv mc_t^p - mc^p =$. Instead of the output gap the mark-up gap becomes the instrument at the heart of the model.

3.3.6 Log linearisation

The above is then log-linearised to the following form

Equation name	Equation	Log-linear Equation
Euler equation	$Q_t = \beta \mathbb{E}_t \left[\frac{C_t}{C_{t+1}} \frac{P_t}{P_{t+1}} \right]$	$c_t = \mathbb{E}_t[c_{t+1}] - (i_t - \mathbb{E}_t[\pi_{t+1}] - \rho)$ ⁶
Labour supply	$\frac{W_t}{P_{c,t}} = C_t N_t^\phi$	$w_t - p_t = c_t + n_t \phi$
Oil factor demand	$M_t = \frac{\alpha_m Y_{q,t}}{\xi_t^n S_t}$	$m_t = -\mu_t^n - s_t + y_{q,t}$ ⁷
Production function	$Y_{q,t} = A_t N_t^{\alpha_n} \left(\frac{\alpha_m Y_{q,t}}{\xi_t^n S_t} \right)^{\alpha_m}$	$y_{q,t} = \frac{1}{1-\alpha_m} (a_t + \alpha_n n_t - \alpha_m s_t - \alpha_m \mu_t^n)$
Gross output	$P_{c,t} C_t = (1 - \frac{\alpha_m}{\xi_t^n}) P_{q,t} Y_{q,t}$	$c_t = y_{q,t} - \chi s_t + \eta \mu_t^n$ ⁸
GDP deflator	$P_{q,t} \equiv P_{y,t}^{1-\alpha_m} P_{m,t}^{\alpha_m}$	$p_{y,t} = p_{q,t} - \frac{\alpha_m}{1-\alpha_m} s_t$
GDP	$P_{y,t} Y_t = \left(1 - \frac{\alpha_m}{\xi_t^n} \right) P_{q,t} Y_{q,t}$	$y_t = y_{q,t} + \frac{\alpha_m}{1-\alpha_m} s_t + \eta \mu_t^n$

The expressions for gross output and output can be combined into

$$y_t = \frac{1}{1-\alpha_m} (a_t + \alpha_n n_t) \quad (56)$$

which combined with the expression for consumption

$$c_t = y_t - \left(\frac{\alpha_m}{1-\alpha_m} + \chi \right) s_t \quad (57)$$

⁶where $\rho \equiv -\ln(\beta)$ and $i_t = \ln(Q_t)$

⁷where $\mu_t^n \equiv \ln(\xi_t^n)$

⁸where $\eta \equiv \frac{\alpha_m}{\xi_t^n - \alpha_m}$

4 Data

All data is gathered from St. Louis Federal Reserve data base of seasonally adjusted deflated macroeconomic time series (ref). The specific data codes can be found in the (appendix). Most data used across the following is standard to the DSGE literature. Some adjustments haven made to allow for the inclusion of a petrol sector. The below will therefore briefly outline the use of standard data, while providing more detail on the data specific to this work. The common variables for DSGE model evaluation are nominal interest rates, consumption, wage and labour hours. The nominal interest rate is proxied by the federal funds effective rate (ref). Consumption is measured by household expenditure data. Labour hours and wage refer to non-farm business total hours worked and average hourly wage. The exclusion of the agricultural sector is generally recommended for the US economy (ref). Moreover, this work employs the WTI Crude Oil Index as proxy for the price of petrol. The WTI dates back longer than the more recent Brent Crude Oil Index, wherefore it is generally used across works in energy economics (ref). With regards to inflation π this work relies on the median consumer price index for the RBC and the NK model. The NK energy model's aim is to analyse inflation under the inclusion of petrol in the economy. Relying on the CPI, which includes petrol prices, would thus be counterproductive as the model should generate inflation resulting from petrol endogenously. In order to accurately reflect the models ability energy price need to be removed from inflation data. In the NK energy model therefore relies on the "Sticky Price Consumer Price Index less Food and Energy". Similar reasoning applies to consumption expenditure which for the NK energy model specifically excludes expenditure for energy and utilities.

4.1 Preprocessing

Economic data is usually considered in per capita terms in order to account for different population dynamics (ref). Population dynamics are not continuously collected and usually a mix of actual measurement and imputation. As a result population data exhibits significant irregularities [Pfeifer, 2021]. Using such noisy records would distort the data of more carefully collected economic variables. Population data thus requires some preprocessing on its own. Smoothing the data with the Hodrick-Prescott (HP)

filter's at a smoothing factor of 10,000 is recommended in order to de-noise the data [Edge et al., 2013]. The smoothed time series captures the real population dynamics and can then be used to obtain per capita variables. At this point it is important to mention that the St. Louis FED data is in chained dollars, thus not concerned by inflation of monetary measurements (ref).

Having removed population dynamic and inflation the next step is the separation in two key components, trend and cyclical. DSGE literature generally studies variables in their deviation from their steady state. This requires an identification of the steady state and its evolution over time. Separating the long-term trend of GDP from its cyclical deviations commonly relies on the HP filter. The filter separates trend and deviation. In the case of quarterly variables a smoothing parameter of 1600 is recommended [Ravn and Uhlig, 2002]. As DSGE models are log-linearised they describe deviation in terms of logarithm, which for small values approximates percentage deviations. In order to transform data into log-deviations from their underlying trend the HP filter is fed with log-transformed time series. This procedure by construction yields stationary time series of cyclical. However, to comply with common standards stationarity is formally confirmed by the Augmented-Dickey Fuller test (ref). Having transformed all trending variables have into stationary log-deviation from their natural level only the percentage variables remain, namely inflation and interest rate. They require no such treatment, however need translation from annual to quarterly rates. This has been done according to the following formula $(1 + \frac{\mathbf{X}}{100})^{1/4} - 1$.

An overview over the transformed variables as well as their descriptive statistics can be found in the (appendix).

4.2 Analysis of models

Compare model covariance matrices, to actual covariance in data Show theoretical impulse response functions, differences between NK and RBC

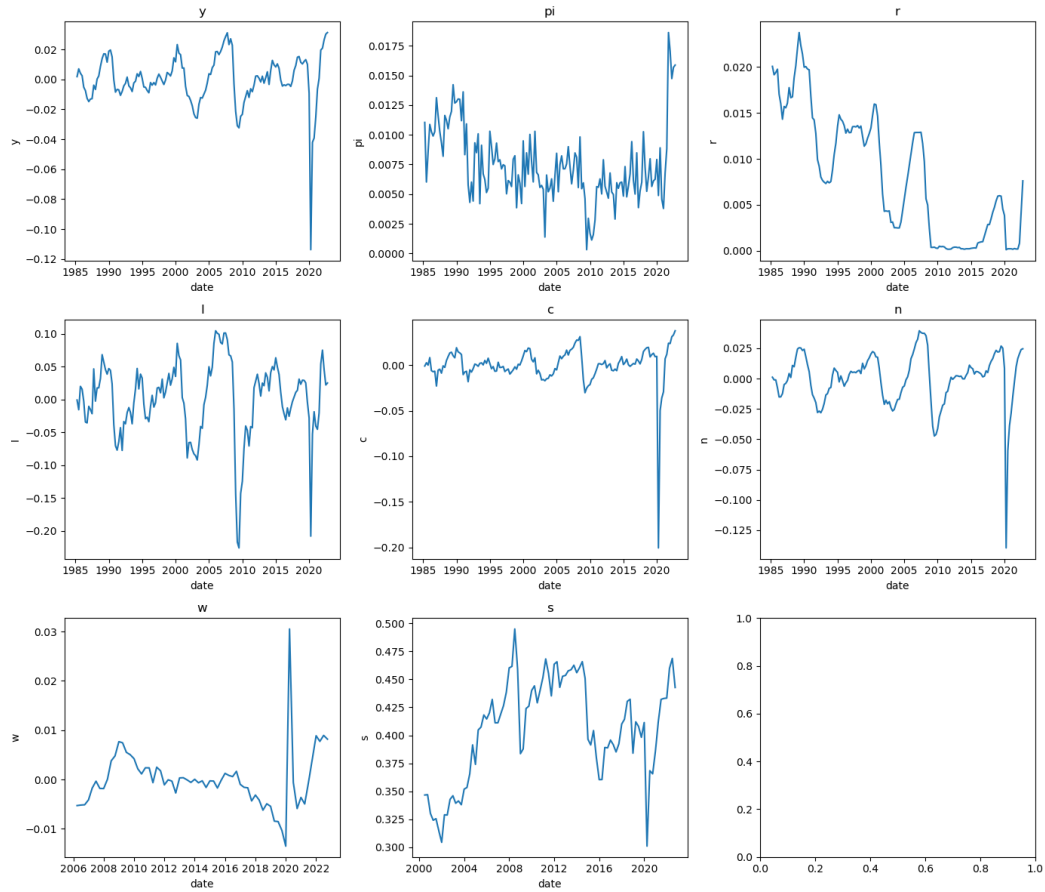


Figure 1: Bayesian VAR 2 lags

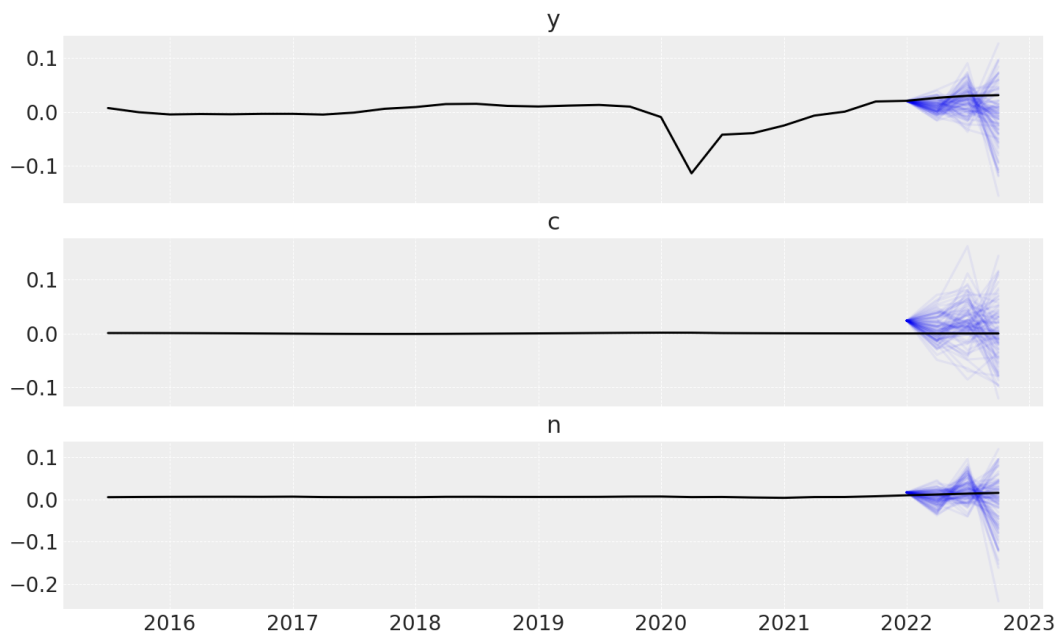


Figure 2: Bayesian VAR 2 lags

Table 1: Descriptive statistics

	count	mean	std	min	25%	50%	75%	max
y	151.0	-0.000	0.017	-0.114	-0.007	-0.000	0.010	0.031
y_p	151.0	-0.000	0.002	-0.006	-0.001	0.000	0.001	0.007
pi_s	151.0	0.007	0.003	0.002	0.006	0.007	0.009	0.016
pi	151.0	0.007	0.003	0.000	0.006	0.007	0.009	0.019
r	151.0	0.008	0.007	0.000	0.001	0.007	0.014	0.024
I	151.0	0.000	0.058	-0.226	-0.030	0.010	0.038	0.105
c_s	151.0	-0.000	0.004	-0.014	-0.002	0.000	0.002	0.021
c	151.0	-0.000	0.021	-0.201	-0.006	0.001	0.009	0.038
n	151.0	0.000	0.022	-0.140	-0.012	0.002	0.012	0.039
w	67.0	-0.000	0.006	-0.014	-0.003	-0.000	0.002	0.031
s	90.0	0.405	0.046	0.301	0.376	0.411	0.442	0.495

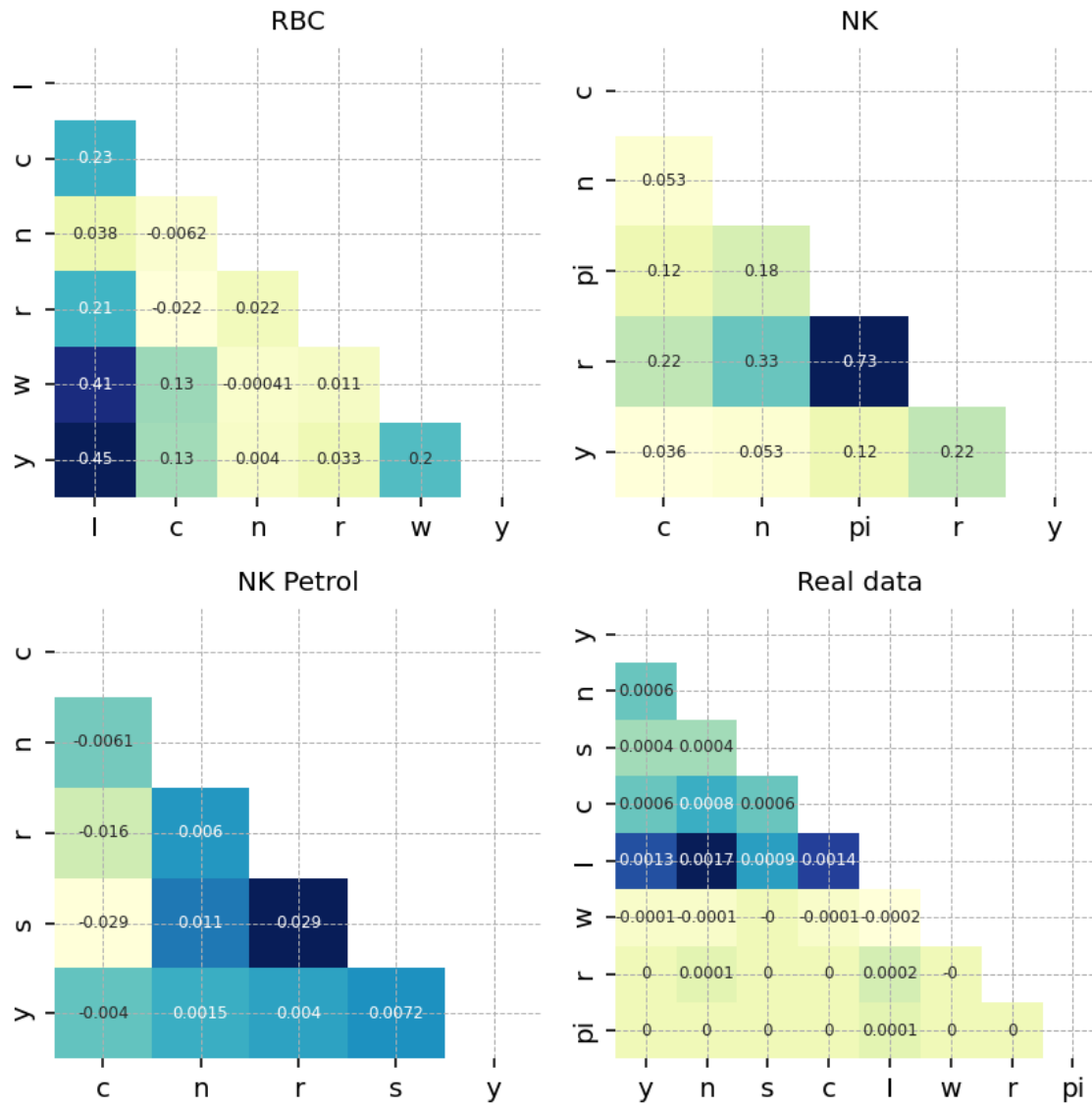


Figure 3: Covariance matrices

5 Bayesian Estimation

As explained in the above the RBC and later NK models are based on structural parameters of the economy. As such they are not subject to the Lucas 1976 critique as their parameters do not directly depend on the choices of agents (ref). Instead, they represent the deep-rooted dynamics of the dynamic system describing the economy, as such they should exhibit a certain stability over time (ref).

This perspective makes an implicit but important assumption: DSGE model corresponds to the real dynamic system describing the economy. A far reaching assumption which has been considered at least partially validated based on the ability of early DSGE models to generate data matching 1st and 2nd order moments of real variables such as output, consumption and investment (ref). In an attempt to tweak this capability early attempts to parameter to identifying the true parameters of the economy were made. In doing so academics first relied on what [Prescott, 1986] referred to as great ratios of the economy. Calculating for example the real interest rate allowed to derive the discount factor β . Other parameters had fewer empirical counterfactuals and were thus not estimated by manually calibrated to sensible values. These inaccuracy gave rise to approaches more deeply rooted in econometric analysis. They were undertaken in the aim of reliably estimating the economy's true structural parameters.

Such early attempts of econometric identification relied on standard frequentists statistics such as the maximum-likelihood estimators (MLE) (ref). Another approach pioneered by Blanchard & Kahn (ref) referred to as simulation method of moments, later progressed into the general method of moments (GMM) attempted to minimize the distance between simulated and real data moments. The GMM allowed to discriminate between several competing specification of the same structural model [Christiano and Eichenbaum, 1992]. Comparing the estimated parameters of the model with real data values then served as inspection of model fit to data. However, as models grew in size a major short-coming of GMM became apparent. The method of moment conditions were no longer sufficient to estimate model fit across the variety of parameters [Guerrón-Quintana and Nason, 2013]. Likewise, the ML estimator turned out to be limited in its ability to estimated the entirety of parameters, requiring manual calibration of some parameters [Guerrón-Quintana and Nason, 2013].

While frequentists statistics knew to overcome the identification problem in introducing theoretical moments Smith (1993) it was soon challenged by a different perspective. At the heart of all frequentist statistics lies the assumption that a given model corresponds to the true structural process, it is the 'true' model (ref). If this assumption cannot be met all estimation will be biased and thus need to be corrected for. Again remedies were found (ref) Smith (1993) or Dridi, Guay, and Renault (2007) in varying parameters most relevant to overall model moments and keeping less relevant ones constant throughout econometric estimation. Yet it is argued that this approach at least partly failed in overcoming the conditional bias introduced by non-relevant parameters [Guerrón-Quintana and Nason, 2013].

Consequently, a new philosophy of linking models to data was identified. Bayesian statistics replaces the assumption of a single true model by the convenient formulation of conditional likelihood making it more appealing to the purposes of DSGE modelling [Guerrón-Quintana and Nason, 2013]. In doing so the Bayesian statistics relies on Bayes law linking the conditional likelihood of an event to the prior distribution.

$$P(\Theta|Y_T) = L(Y_T|\Theta)P(\Theta)$$

A given model specification can be linked to a likelihood evaluated against the backdrop of data the system is meant to generate (ref). This allows to assess a model's suitability. The frequentist's idea of a single true parameter is replaced by distributions assigning different probabilities to parameter values. These prior distribution reflect previous knowledge allowing to narrow down the possible space of parameters. Instead of identifying a true parameter Bayesian statistics, aims at evaluating these beliefs and at decreasing uncertainty around the parameter space i.e. the variance of the distribution. This process translates the prior distribution into a posterior distribution, not necessarily belonging to the same class. This process of evaluation is referred to as sampling and builds on numerical and stochastic calculus to approximate the true posterior distribution (ref). Furthermore, Bayesian statistics naturally provides great forecasting properties. Once uncertainty around the space of the real system has been reduced it allows for make informed forecasts, naturally providing confidence bands.

The following section will therefore be concerned with these four concepts, conditional likelihood, prior and posterior distribution, sampling and forecasting. DSGE literature has developed a rich body of literature covering the case of the canonical DSGE model rather well (ref). The state of research is currently concerned with the estimation of larger models and non-linear and not normally distributed posterior distributions (ref). However, given the simplicity of the above discussed models current techniques are more than sufficient for this thesis purposes.

5.1 The log-likelihood

Bayesian analysis of DSGE models is concerned with the transition from a prior distribution reflecting previous beliefs to the true posterior distribution (ref). This process relies on Bayes law as illustrated above, which requires the prior distribution and the marginal likelihood of data given the representation of the system $L(Y|\Theta)$ (ref). This section is concerned with obtaining the latter.

The marginal likelihood for a linear state system requires a technique allowing to assess the fit of data to the system. In doing so one relies on the fact that DSGE models when expressed as linear dynamic systems are data generating processes. This allows to formulate the question of how likely a given set of data has been generated by the system. A method allowing to do so is the Kalman filter (ref). This filter originates from engineering and has originally been developed to track the position of an air plane off signals from different sensor. As sensor signals are inherently noisy the filter is designed to exploit one's knowledge about the measurement noise in order to reduce the variance on its believe about the actual position or state of the plane. In doing so the filter imposes the important assumption of Gaussian distributed noise, exploiting the fact that any multiple of two Gaussian distributions will again yield a Gaussian distribution. The resulting Gaussian then has a lower variance than its parents, which is equivalent to saying that the resulting distribution has narrowed down the believe on the current state of the system. This assumption is especially important for the purposes of DSGE modelling, as the likelihood function of the Kalman filter relies on the assumption of Gaussian noise. Applied to DSGE models this implies that the joint distribution of shocks must be Gaussian [Herbst and Schorfheide, 2016]. This is usually satisfied for small scale New Keynesian models if the individual shocks are

Gaussian (ref). For larger DSGE models this assumption does not necessarily hold and shocks are highly correlated. As mentioned above the is state of research and demands other methods of evaluation such as the Kalman-Particle filter (ref). As this work is concerned with the canonical model the Kalman filter will be sufficient.

Given such linear Gaussian system the Kalman filter becomes especially desirable as it provides optimal forecast (ref), making any other method of evaluation obsolete. The below will now briefly explain how the Kalman filter is constructed off the linear system and end with deriving its likelihood function.

Given the generic linear system below wit \mathbf{F} as the transition and \mathbf{R} as the shock noise matrix the Kalman filter proceeds in two steps, a predict and an update step:

$$\begin{aligned}\mathbf{X}_t &= \mathbf{F}\mathbf{X}_{t-1} + \mathbf{R}\epsilon_t \\ \epsilon_t &\sim N(0, \sigma_\epsilon) \\ \begin{bmatrix} y_t \\ z_t \end{bmatrix} &= \begin{bmatrix} \rho_y & \rho_{yz} \\ \rho_{zy} & \rho_z \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{yt} \\ \epsilon_{zt} \end{bmatrix} \\ \begin{bmatrix} \epsilon_{y,t} \\ \epsilon_{z,t} \end{bmatrix} &\sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_y^2 & 0 \\ 0 & \sigma_z^2 \end{bmatrix} \right)\end{aligned}$$

Predict step

In the predict step the Kalman filter propagates the system of state variables \mathbf{X}_t according to the transition matrix \mathbf{F} . This matrix defines the system and as such is time independent. The propagation then yields an estimate for the X_{t+1} system state. This could be supplemented by directions or external influences on the system that one is aware of. However, this is not the case for this work.

$$\hat{\mathbf{X}}_{k|k-1} = \mathbf{F}\mathbf{X}_{k-1|k-1}$$

The second key aspect to the Kalman Filter is the system state variable's covariance matrix \mathbf{P}_t , which is also time dependent. As pointed out in the above, this matrix is assumed to be Guassian, thus semi-definit positive (ref). The covariance matrix and the state \mathbf{X} essentially build a Guassian distribution summarising the uncertainty around the actual state.

$$\mathbf{P}_{k|k-1} = \mathbf{F}\mathbf{P}_{k-1|k-1}\mathbf{F}^T + \mathbf{Q}$$

Filter step

Once the predicted state \mathbf{X} and covariance matrix \mathbf{P} have been determined the Kalman filter proceeds to identifying the most likely system state. For this it takes into account the state variables measurement referred to as \mathbf{z}_t , which due to measurement noise is not considered to always reflect the actual state of the system. The actual state is instead thought of to lie in-between the predicted state $\mathbf{X}_{k|k-1}$ and the measurement \mathbf{z}_k . The objective of the filtering step is therefore to identify this true state in-between the two. For this purpose the Kalman filter derives the Kalman gain $\mathbf{K}_{k|k-1}$, which can be thought of as a weighting function between $\mathbf{X}_{k|k-1}$ and \mathbf{z}_k . In building this average the filter takes into account the process noise and the covariance matrix as well as the accuracy of past iterations.

\mathbf{y}_k is the difference between measurement and prediction. The matrix \mathbf{H} is the measurement matrix, ensuring that predicted state $\mathbf{X}_{k|k-1}$ and \mathbf{z}_k are in the same units.

$$\mathbf{y}_k = \mathbf{z}_k - \mathbf{H}\mathbf{X}_{k|k-1}$$

The \mathbf{S}_k captures the filter's uncertainty by combining the covariance around the predicted state $\mathbf{X}_{k|k-1}$ with the process noise \mathbf{R} . In DSGE terminology the process noise is derived from the shock covariance matrix (ref).

$$\mathbf{S}_k = \mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^T + \mathbf{R}$$

The uncertainty term \mathbf{S}_k is then used to calculate the weighting function between measurement and predicted state the Kalman gain.

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}^T\mathbf{S}_k^{-1}$$

The filtered state which will be next iterations start then is calculated off the Kalman gain and \mathbf{y}_k . The covariance matrix is updated in a similar fashion. This yields a new narrower distribution summarising the uncertainty around the true system's state.

$$\mathbf{X}_{k|k} = \mathbf{X}_{k|k-1} + \mathbf{K}_k\mathbf{y}_k$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H})\mathbf{P}_{k|k-1}$$

Likelihood function

Having introduced the Kalman filter algorithm the setting of the starting values remain. For the Kalman filter these are less decisive, has the filter is very capable in converging to the true state. However, as the models at hand are describing deviations from the steady state this work has set $\mathbf{X}_{0|0}$ to zero, assuming that the system is in its steady state to begin with, an approach generally followed by the literature. Secondly, this section has initially been dedicated to the Kalman filter's likelihood function. Having explored the propagation one can now derive the likelihood function. This function is heavily reliant on the assumption of Gaussian system noise, as it draws from the normal distribution in evaluating the likelihood for each iterations. In the below when referring to likelihood this thesis will however be referring not to the individual likelihood of each iteration but to the cumulated likelihood across all iterations.

The probabilitiy desinty function for any normal distribution is given by:

$$\mathcal{L} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2} \frac{x - \mu}{\sigma}\right]^2$$

The Kalman filter relies on the multivariate where the standard deviation is given by system uncertainty \mathbf{S}_k and the difference between mean and measurement are expressed through \mathbf{y}_k .

$$\mathcal{L}_k = \frac{1}{\mathbf{S}_k\sqrt{2\pi}} \exp\left[-\frac{1}{2} \mathbf{y}_k^T \mathbf{S}_k^{-1} \mathbf{y}_k\right]$$

This work has drawn from filterpy in order to perform the above calculations (ref).

5.2 Sampler

There are local and global identification problems

The key aspect of Bayesian estimation across DSGE models is the identification of the true parameter distribution, the posterior. However, mapping a DSGE model's parameters to its posterior is non-linear in the parameter vector Θ . Therefore the posterior distribution cannot be evaluated analytically [Herbst and Schorfheide, 2016]. Instead a numerical approximation mechanism is require, this is referred to as the sampler [Guerrón-Quintana and Nason, 2013]. The Metropolis-Hastings Monte Carlo Markov Chain (MH-MCMC) sampler is the predominant sampling method in linear Bayesian estimation literature [Guerrón-Quintana and Nason, 2013]. The MH-MCMC sampler is applicable to small scale New Keynesian models, for which it usually delivers good results as the joint parameter distribution is well-behaved and elliptical

[Herbst and Schorfheide, 2016]. This reduces the posterior identification problem to a global problem. Larger NK models exhibit no-elliptic parameter distributions and thus require samplers which are able to distinguish local and global likelihood maxima in their identification of the posterior [Herbst and Schorfheide, 2016]. As this thesis is concerned with the small scale canonical model the Metropolis-Hastings sampler is sufficiently accurate and will be used.

The MH-MCMC sampler suggests new posterior candidates according to a multivariate random walk. The posterior is extended by a suggestion if it provides an improvement in the likelihood as described in (section). The algorithm generates a stable Markov chain, which is meant to exhibit low autocorrelation and low variance of the MH estimator. If this is fulfilled the resulting chain is equivalent to the posterior distribution [Herbst and Schorfheide, 2016].

Metropolis Hastings sampling algorithm - it is important to note that a draw from the shocks prior, factors into the drawing process - procedure also followed by, with 300,000 draws of which 20,000 are considered burn in [Chin and Li, 2019]

The below will be explaining this thesis implementation of the MH sampler and methods of evaluating its convergence to the true posterior distribution. The algorithm proceeds in three main steps. First a candidate for the posterior $\hat{\Theta}_{k|k-1}$ is suggested based on the random-walk law of motion. The random-walk departs from the last accepted posterior candidate and suggests a new candidate by departing from the $\Theta_{k-1|k-1}$ according to the zero-mean Gaussian distribution ϵ . The distribution's variance Σ can be set to either the identity matrix \mathbf{I} or to contain the parameter prior variances on its diagonal. The latter is recommended if well-defined priors are at play [Herbst and Schorfheide, 2016]. To complete the random-walk the MH sampler also employs a gain parameter η that essentially scales the distance of the new posterior candidate away from the last accepted posterior $\Theta_{k-1|k-1}$. This value has been calibrated by several studies, which recommend an $\eta = 0.234$ for the special case of a multivariate normal distribution (ref). Generally the value of η should correspond to an acceptance rate between 20% and 40% of suggested draws [Herbst and Schorfheide, 2016].

$$\hat{\Theta}_{k|k-1} = \Theta_{k-1|k-1} + \eta\epsilon$$

$$\epsilon = N(0, \Sigma)$$

The posterior candidate $\hat{\Theta}_{k|k-1}$ is then evaluated based on the Kalman Filter likelihood function introduced in (section). This yields the likelihood of posterior candidate given the data \mathbf{Y} .

$$\mathcal{L}(Y|\hat{\Theta}_{k|k-1})$$

Once the likelihood is obtained the algorithm proceeds into the two step acceptance procedure. First the likelihood of the candidate is compared to last accepted posterior in order to determine whether this draw constitutes an improvement. This is done according to Bayes' law.

$$\omega_k = \min\left\{\frac{\mathcal{L}(Y|\hat{\Theta}_{k|k-1}) * P(\hat{\Theta}_{k|k-1})}{\mathcal{L}(Y|\Theta_{k-1|k-1})P(\Theta_{k-1|k-1})}, 1\right\}$$

The likelihood ratio is then fed into the second step the random acceptance. The above obtained ω_k is compared the uniform variable $\phi \sim U(0, 1)$. If $\omega_k \leq \phi_k$ the draw is accepted and the candidate is assumed into the posterior. This process is repeated N times and thus generates the posterior distribution. As illustrated in the above MH-MCMC is a recursive method and as such requires an initial value for the posterior Θ . This value is usually obtained throughout a so-called burn in period corresponding to 10 % of total iterations (ref). The total number of iterations depends on the speed of convergence of the chain but usually ranges between 20,000 and 50,000 iterations (ref).

The following will now describe the evaluation steps to insure convergence of the chain. Moreover, once the algorithm has run one needs to evaluate this performance. A method beyond visual inspection is suggested by (ref).

The MH random walk is highly susceptible to mis specification of the posterior of the prior distribution is far away from the true posterior [Herbst and Schorfheide, 2016]

5.3 Priors

In response to the Lucas critique (ref) DSGE models were built around structural time invariant parameters. They summarise agent preferences impacting behaviour and do not describe time dependent dynamics in the short-run. Parameters thus are a statistical distributions of preferences, forming the joint distribution of preferences. This joint distribution is usually impossible to evaluate. Instead parameters

are individually analysed and distributions reflecting independent and prior knowledge are assigned [Del Negro and Schorfheide, 2008]. The notion of independence is important in two ways. Firstly, only if the data used to construct the priors is independent of the likelihood mechanism the resulting posterior is unbiased. This condition is fulfilled if the data for prior formation at least pre-dates the of the likelihood evaluation [Herbst and Schorfheide, 2016]. Secondly, assuming independence between parameters is reasonable within limits. Especially, shocks do not necessarily occur independently [Herbst and Schorfheide, 2016]. However, this is usually only a problem for larger models with several shocks. As this work analyses the canonical model shock independence is of no concern. The later will therefore focus on identifying priors based on independent information. The below will follow the approach of [Del Negro and Schorfheide, 2008], dividing priors into three groups in order to obtain eligible distributions. An overview on all prior distributions can be found in der appendix 7.0.1.

5.3.1 Steady state priors

The steady state priors can be easily identified from the steady state or long-run relationship of model variables. They are also referred to as great ratios' [Kydlan and Prescott, 1982].

$$\begin{aligned}\Theta_{ss}^{RBC} &= \{\alpha, \beta, \delta\} \\ \Theta_{ss}^{NK} &= \{\alpha, \beta, \theta\} \\ \Theta_{ss}^{NKE} &= \{\alpha_m, \alpha_n, \beta, \theta, \chi\}\end{aligned}\tag{58}$$

The time discount factor β is usually not calibrated, as it results directly from the inverse of the interest rate. The pre-1985 data used to estimate prior distributions puts β to 0.995.

The capital share can be derived from the long-run share in industrialised nations. For smaller models using a uniform distribution covering the possible parameter interval $\alpha \in [0, 1[$ is suggested [Del Negro and Schorfheide, 2008]. In the RBC α 's posterior converges from a uniform to a centred distribution 4. While this is still rather broad and far off the empirical value of 33% it is indication of convergence towards the true posterior 4. For both New-Keynesian models a more targeted distribution has been used to ensure steady state existence [Guerrón-Quintana and Nason, 2013]. Consequently the values for α are more left skewed 5. Interestingly enough the cal-

ibration of the NKE greatly overestimates the share of oil in production which by [Blanchard and Gali, 2007] is reckoned at below 5% 6. The labour share in production is more reasonable but also off the empirical estimate of 0.7 [Blanchard and Gali, 2007].

The depreciation rate δ while rather high in the posterior has yet assumed reasonable values in limiting its posterior to below 0.8 4. This indicates some convergence to the real prior, while still being far from empirically estimated values of 2.5% per quarter [Campbell, 1994].

The price setting probability $\theta \in [0, 1]$ has reasonable estimates in the NK model 5. Its mean is around 0.8 which corresponds to an average price duration of roughly four month, close to the empirical estimate [Blanchard and Gali, 2007]. This is not the case for the NKE model where θ exceeds one, which as a probability is impossible. The NKE estimates are much more accurate for the share of petrol in consumption, whose posterior remains close to the empirical mean of 1.7% [Blanchard and Gali, 2007].

5.3.2 Exogenous prior

Exogenous priors contain the exogenous law of motion and standard deviation of shocks. In a simple model with singular shocks the shock first and second order moments can be directly from the economic variables that are affected by it [Del Negro and Schorfheide, 2007]. This work thus calibrated the shock standard deviation based on the variance of output. Moreover, in setting the autoregressive parameter this work draws from the distributions suggested by [Vásconez,]. Looking across the posterior all estimates for ρ are bounded at one. While this might appear to be a successful identification this is merely due to the stability requirement of the dynamic system. If $\rho \in]-1, 1[$ is violated the dynamic system does not exhibit stability and the Blanchard-Kahn conditions of stability are violated. The high variance of the posterior thus indicates problems in the identification of the posterior. However, its skewness towards higher values seems to confirm the prior assumption of relatively time persistent shocks.

$$\begin{aligned}\Theta_{exog}^{RBC} &= \{\rho_A\} \\ \Theta_{exog}^{NK} &= \{\rho_v\} \\ \Theta_{exog}^{NKE} &= \{\rho_s\}\end{aligned}\tag{59}$$

5.3.3 Endogenous priors

All remaining parameters make up the group of endogenous priors. These parameters cannot be measured directly through economic variables' long-term relationships. Instead they rely on micro-evidence from external data sets [Del Negro and Schorfheide, 2008].

$$\begin{aligned}\Theta_{endog}^{RBC} &= \{\sigma_C, \sigma_L, \} \\ \Theta_{endog}^{NK} &= \{\sigma_C, \sigma_L, \epsilon, \phi_y, \phi_\pi\} \\ \Theta_{endog}^{NKE} &= \{\epsilon, \phi_y\}\end{aligned}\tag{60}$$

The elasticity of substitution σ_C and σ_L belong in this category. In setting their prior distribution this work relied on the recommendation from [Del Negro and Schorfheide, 2008].

For the RBC the posterior values demonstrate large spreads indicating that the MH sampler did not identify the posterior correctly ⁴. Even larger dispersion can be observed for the NK model ⁵ ⁹. This occurs frequently as the consumption and labour elasticity are parameters with low influence on the transition dynamics [Guerrón-Quintana and Nason, 2013]. Consequently they are dominated by stronger parameters in the likelihood function and the posterior identification fails for these variables.

The elasticity of substitution ϵ originates from the Dixit-Stiglitz aggregator and factors into the steady-state mark-up of firms. As such it is one of the driving parameters behind inflation and strong in its influence on the transition dynamics. In setting the prior distribution one can rely on information on the steady-state mark-up $mc^r = \frac{\epsilon}{1-\epsilon}$. Following this a value of ϵ slightly above 0.5 appears realistic. Consulting the posterior both put the mean around 0.5 ⁵ ⁶. While the standard deviation is still large the posterior specification does appear to be within reasonable bounds.

The central bank's preference parameters ϕ_y and ϕ_π also appear within reasonable bounds. As suggested by [Galí, 2008] the central bank has a stronger response to inflation than to the output gap. This has been incorporated into priors and the resulting posterior seems to confirm this assumption in the NK model ⁵. The NKE model relies on output and not the output gap as intertemporal instrument, wherefore ϕ_π appears. However, the posterior estimate appear to confirm the same reasoning ⁶.

⁹In the NKE model the labour and consumption elasticities do not enter into the log-linear model specification due to the log-utility function

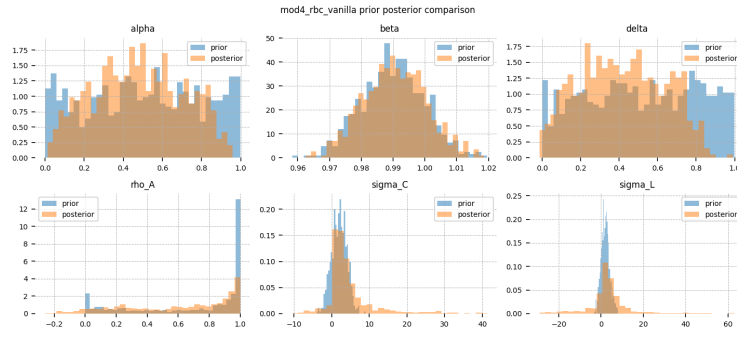


Figure 4: RBC model prior posterior comparison.

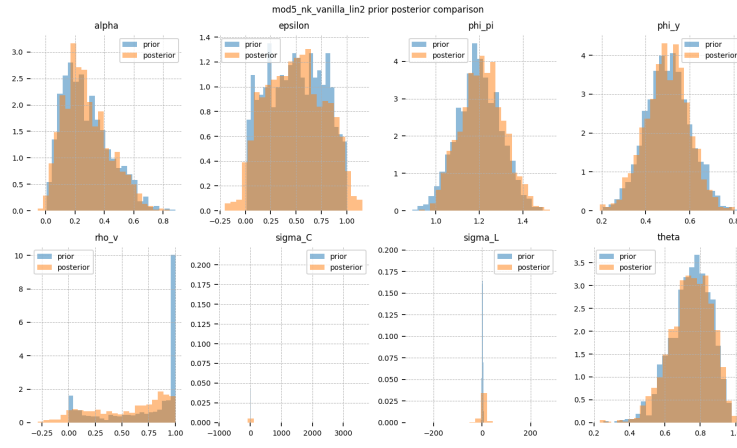


Figure 5: New Keynesian model prior posterior comparison.

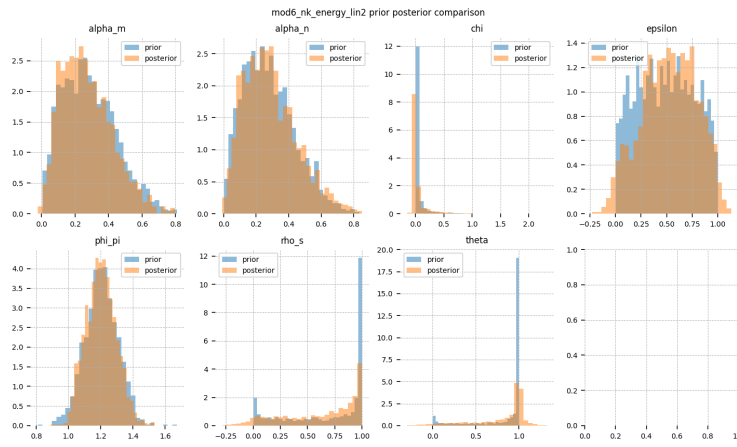


Figure 6: New Keynesian Energy model prior posterior comparison.

6 Forecast evaluation

Having identified the parameter posterior distribution this section will now again rely on the Kalman filter to make a three quarter ahead forecast for each model. In doing so the mean posterior will be used across the model specification as this is the most likely model given the underlying data. A first section will assess the individual forecast of each model. It will primarily consider output as the central variable of the economy, while briefly commenting other variables as well. A full overview on the Kalman filter forecasts can be found in 7.0.1. A latter section will then introduce a Bayesian Vector Autoregression (BVAR) model and compare its performance to the DSGE model forecasts.

6.1 DSGE forecasts

While the Kalman filter for the posterior evaluation has been run on the entire data set starting in 1985 the following has been run on post 2010 data. This allows to assess the data's fit to recent events.

6.1.1 RBC

The Kalman filter fed data on output and investment provided the best fit to data 7. Looking at the filter output output and investment are closely traced and the real data lies within the 95% confidence interval after 2010. This is also true for the recession in early 2020. This indicates a good fit of the model to data.

Consumption, while not contained as an observable in the filter, also is relatively closely traced with similar accuracy of the confidence intervals. This points to even better model accuracy, as the dynamic system is able to replicate a non-included variable to decent extend.

The interest rate is less closely traced and predicted to be much more volatile by the Kalman filter than the case in reality. While this indicates that problems in the model specification it is important to note that the FED has pursued a notorious low interest rate policy for the sample duration. As such interest rate was not set by the dynamics contained in the model, and a good fit is unlikely.

Labour is the only variable with bad fit. The 90% confidence interval far-off reality and the filter even suggests a rise labour hours during the 2020 crisis. While in the case of Covid-19 one could argue that particular dynamics where at play ?? illustrates that even in earlier period recessionary unemployment is not well captured. The historic overview sheds a better light on the overall model fit. While post 2010 dynamics are well captured by the model, fit prior to this date is less accurate. Output and consumption data escapes the filter's 95% confidence interval before and after the 2008 financial crisis.

Forecasts

The RBC forecasts seem reasonable in direction and as well as confidence. The 95% confidence intervals naturally increase with distance to the point of prediction. However, especially at the one period ahead forecast they exhibit reasonable bounds. Moreover, the point forecast with exception of labour is closely tracing the real variable evolution with no systematic bias. Labour as explained in the above appears to be incorrectly specified in the model's functional form, yet not impeding the model in functioning.

6.1.2 NK

At first glance it is evident that the New-Keynesian filter traces data less accurately. As with the RBC output is closely traced after 2010 while confidence intervals exclude actual data around the 2008 crisis. However, other than with the RBC filter the NK filter does not manage to track the 2020 crisis as closely, even deems it outside its 95% confidence interval. A similar dynamic can be observed with consumption. Again filter fit is much better after 2010 and does not capture the 2020 crisis. Consumption essentially remodels output data, due to the absence of investment the NK model assumes consumption and output to be equal in equilibrium. Thus the consumption time series is less of a performance indicator as in the RBC model. Labour on the other hand is more closely traced and differently from the RBC no longer exhibits counter intuitive movements during crisis, modelling recessionary unemployment correctly.

The main contribution of the New-Keynesian literature is the inclusion of inflation. Looking at the actual inflationary dynamics and the filtered inflation series the large

divergence is obvious. The model puts inflation at much higher rates than actually the case in the economy. The same goes for the real interest rate, which is not surprising due to the close tie between the two $r_t = i_t + \mathbb{E}[\pi_{t+1}]$. Moreover, the confidence intervals are rather broad on both inflation and interest rate. This while no sign of good fit at leasts shows that the model is uncertain about the actual behaviour of inflation and does not deem it erroneously close to its filter result.

While the filtered data provided some good results the forecasts of the NK model are far from informative. Output and consumption are not contained in the forecasts' 95% interval. Moreover, the broad confidence intervals on inflation and interest rate conceal the fact that the inflation forecasts actually moves opposite to real data. Only labour is proxied traced to reasonable extend by the confidence interval, while the point forecast is systematically off.

6.1.3 NKE

The filtered data for the New-Keynesian Energy model are the furthest off real data. The dynamics of output and labour are poorly tracked while spanned by erroneously narrow confidence intervals. Including inflation in the filter does not improve this inability to track output or consumption. Neither is the real interest rate, even-though closely tied to inflation, traced by the confidence interval or point forecast. The reason for including petrol in the model specification has been to supplement the analysis of inflation by an important factor. Including the petrol price in the filter does not provide the theorised improvement. Instead confidence intervals narrow across all variables with no improvement in point forecasts.

The poor filter fit is obviously reflected across the forecasts as well. While confidence bands are narrow the forecast often points into opposite direction, excluding the real data for consumption, inflation and the real interest rate. Output and labour are in no way traced reasonably.

6.1.4 BVAR

Using a Vector Autoregression as reference model for comparing DSGE model forecasts has been commonly done across literature [Schorfheide, 2000, Chin and Li, 2019].

Naturally the BVAR does not provide confidence intervals like frequentist methods of forecasting. Instead model parameters are defined through a sampler, which in this case ran in four iterations. Sampling from these iterations that represent the posterior coefficient distribution allows to derive information about the range of possible outcomes. This range can then be interpreted as a confidence interval (ref). A sample from the posterior draws is illustrated in blue 14. The range of outcomes for output suggest bias in the forecast. They are not evenly distributed around the actual data series. Instead the range of outcomes underestimates output and also introduces additional volatility not matched by the actual observations. This is also the case for all other variables, additional volatility or classicality is introduced systematically across all forecasts. The forecast for the real interest rate after over correcting even moves in the opposite direction. The BVAR thus represents on first account not better forecasts than for example the RBC model.

This work uses a frequentist VAR to determine the optimal number of lags, which are found to be 4. This corresponds to the choice of four lags in [Chin and Li, 2019] Including priors into bayesian VAR makes the model better at forecasting [Chin and Li, 2019]

6.1.5 Forecast comparison

Comparing several models across their forecasting performance can be done in two ways, a frequentist and bayesian approach. The bayesian averaging approach compares models in their log-likelihood based on an information criterion. For the purposes of this work the information criterion will be the widely applicable information criterion (WAIC) as suggested in [Chin and Li, 2019]. In Bayesian model averaging the weights of each model is calculated from its posterior model probability [Chin and Li, 2019]. This is traditionally done for structurally similar models in order to discriminate between different priors. However this procedure has also been employed in comparing different model classes, then usually relying on the WAIC [Chin and Li, 2019].

- (Stock Watson 2003) argue that a simple average of models provides the best forecast
- (.)Jacobson and Karlsson (2004) and Wright (2009)) argue in favour of BMA

$$p(M_k|y) = \frac{p(y|M_k)p(M_k)}{\sum_{k=1}^K p(y|M_k)p(M_k)}$$

$$p(y|M_k) = \int p(y|\theta_k, M_k)p(\theta_k|M_k)d\theta_k$$

where $p(M_k)$ is the prior density, and $p(y|M_k)$ is the model's marginal likelihood

Anderson (2008) suggest to use the predictive likelihood instead of the marginal model likelihood

"Simply speaking, our Bayesian combination forecast (point forecast) is a weighted average of the K individual point forecasts, wherein the weights are determined by their predictive likelihoods" [Chin and Li, 2019]

Comparing forecasting performance Diebold-Mariano test for point forecast comparison following [Chin and Li, 2019] - MSE based point forecast comparison Amisano–Giacomini test for density forecast performance following [Chin and Li, 2019]

7 Conclusion

7.0.1 Conclusion on MH MCMC

The MH MCMC sampler used across this work is one of the simpler likelihood sampling algorithms. In more comprehensive DSGE models it fails to converge, but for the small scale models analysed in this work it is still deemed appropriate (ref). However, as the above illustrates some identification problems especially regarding 'weaker' parameters have become apparent. Despite providing reasonable estimates for most parameters these should also be handled with care.

Possible explanations for the deterioration of model fit across the NKE could lie in its specification. The model was originally build around domestic output which combined with petrol consumption generates overall consumption. Overall output was then derived off nominal output by including a GDP deflator. This extension was not necessary for closing the model but a way of including the correct output specification to make the model comparable to data. The NKE at its core does thus not describe output as it is specified in the data, wherefore it is no surprise that it does badly in tracing the real data. In that the model does deliver further insight in possible links between petrol and inflation. However, it does so in a way that cannot be verified with observable data.

References

- [Blanchard and Gali, 2007] Blanchard, O. and Gali, J. (2007). The Macroeconomic Effects of Oil Shocks: Why are the 2000s So Different from the 1970s? Technical Report w13368, National Bureau of Economic Research, Cambridge, MA.
- [Campbell, 1994] Campbell, J. Y. (1994). Inspecting the mechanism: An analytical approach to the stochastic growth model. *Journal of Monetary Economics*, 33(3):463–506.
- [Chin and Li, 2019] Chin, K.-H. and Li, X. (2019). Bayesian forecast combination in VAR-DSGE models. *Journal of Macroeconomics*, 59:278–298.
- [Christiano and Eichenbaum, 1992] Christiano, L. J. and Eichenbaum, M. (1992). Current real-business-cycle theories and aggregate labor-market fluctuations. *The American Economic Review*, pages 430–450. Publisher: JSTOR.
- [Del Negro and Schorfheide, 2008] Del Negro, M. and Schorfheide, F. (2008). Forming priors for DSGE models (and how it affects the assessment of nominal rigidities). *Journal of Monetary Economics*, 55(7):1191–1208.
- [Edge et al., 2013] Edge, R. M., Gürkaynak, R. S., and Kisacikoglu, B. (2013). Judging the DSGE model by its forecast. Technical report, mimeo.
- [Galí, 2008] Galí, J. (2008). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework*. Princeton University Press, Princeton, N.J. OCLC: ocn177826088.
- [Guerrón-Quintana and Nason, 2013] Guerrón-Quintana, P. A. and Nason, J. M. (2013). Bayesian estimation of DSGE models. In *Handbook of research methods and applications in empirical macroeconomics*, pages 486–512. Edward Elgar Publishing.
- [Herbst and Schorfheide, 2016] Herbst, E. P. and Schorfheide, F. (2016). *Bayesian estimation of DSGE models*. Princeton University Press.
- [Kim and Loungani, 1992] Kim, I.-M. and Loungani, P. (1992). The role of energy in real business cycle models. *Journal of Monetary Economics*, 29(2):173–189.

- [Kydland and Prescott, 1982] Kydland, F. E. and Prescott, E. C. (1982). Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, pages 1345–1370. Publisher: JSTOR.
- [Pfeifer, 2021] Pfeifer, J. (2021). A Guide to Specifying Observation Equations for the Estimation of DSGE Models.
- [Prescott, 1986] Prescott, E. C. (1986). Theory ahead of business-cycle measurement. In *Carnegie-Rochester conference series on public policy*, volume 25, pages 11–44. Elsevier.
- [Ravn and Uhlig, 2002] Ravn, M. O. and Uhlig, H. (2002). On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations. *Review of Economics and Statistics*, 84(2):371–376.
- [Schorfheide, 2000] Schorfheide, F. (2000). Loss function-based evaluation of DSGE models. *Journal of Applied Econometrics*, 15(6):645–670. Publisher: Wiley Online Library.
- [Vásconez,] Vásconez, V. A. What if oil is less substitutable? A New-Keynesian Model with Oil, Price and Wage Stickiness including Capital Accumulation.

Appendix

Appendix A

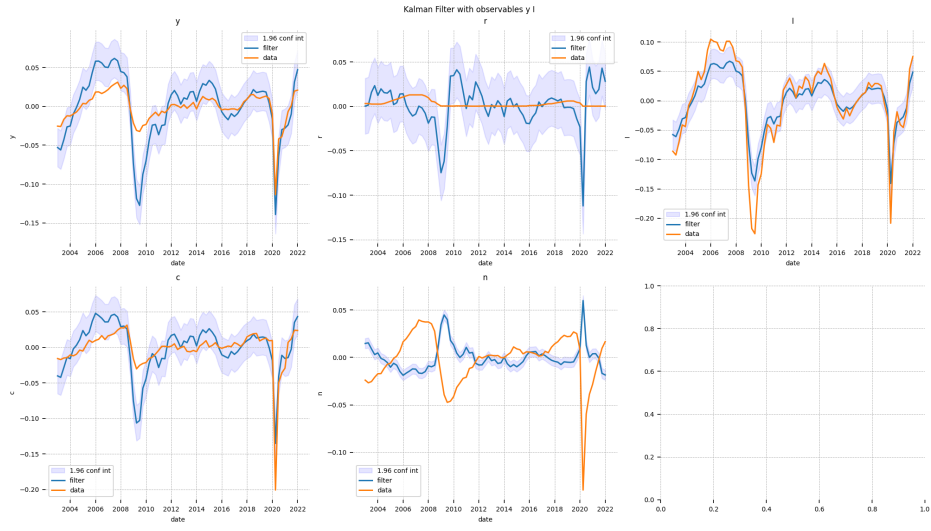


Figure 7: NK model Kalman Filter with variables c, n

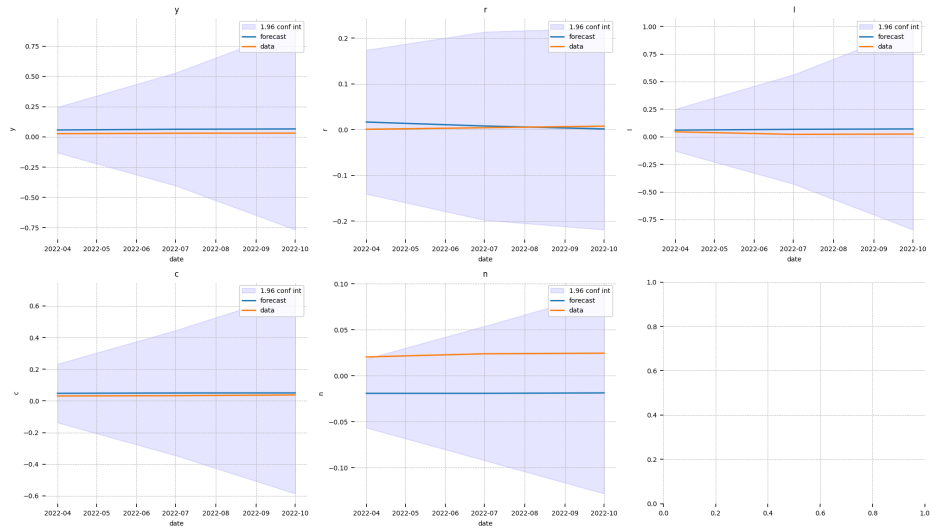


Figure 8: NK model Kalman Filter with variables y

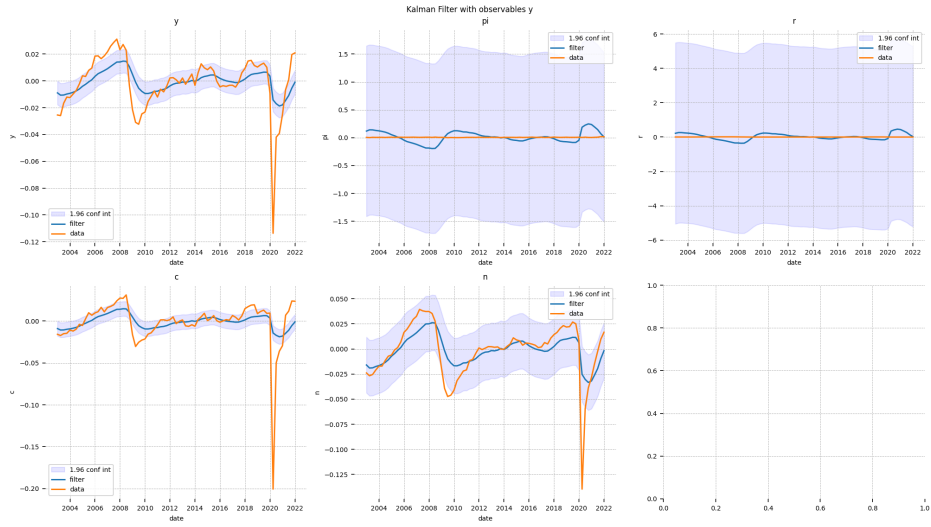


Figure 9: NK model Kalman Filter with variables c, n

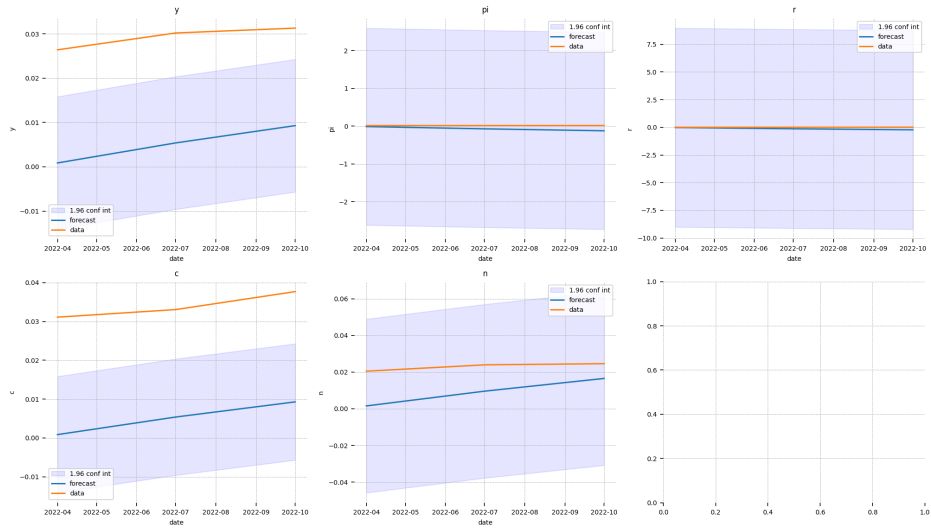


Figure 10: NK model Kalman Filter with variables y

Appendix B



Figure 11: NK model Kalman Filter with variables c , n

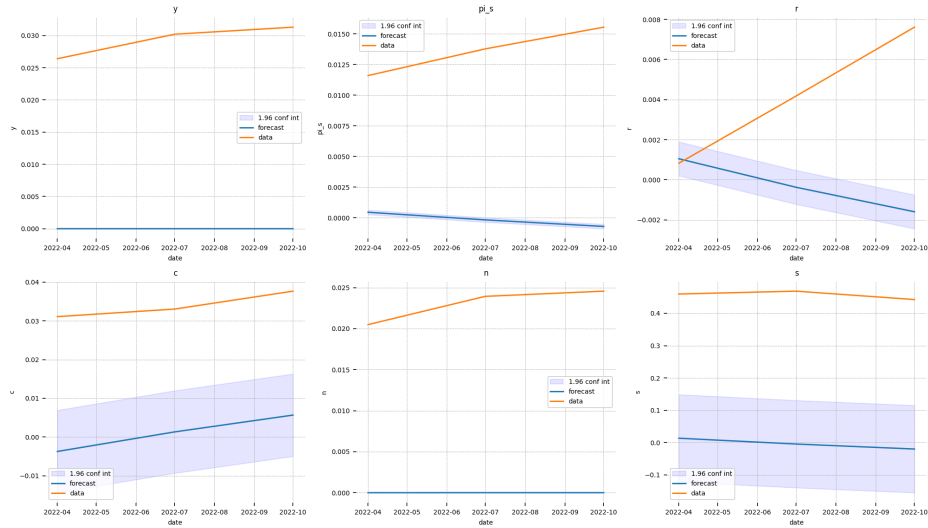


Figure 12: NK model Kalman Filter with variables y

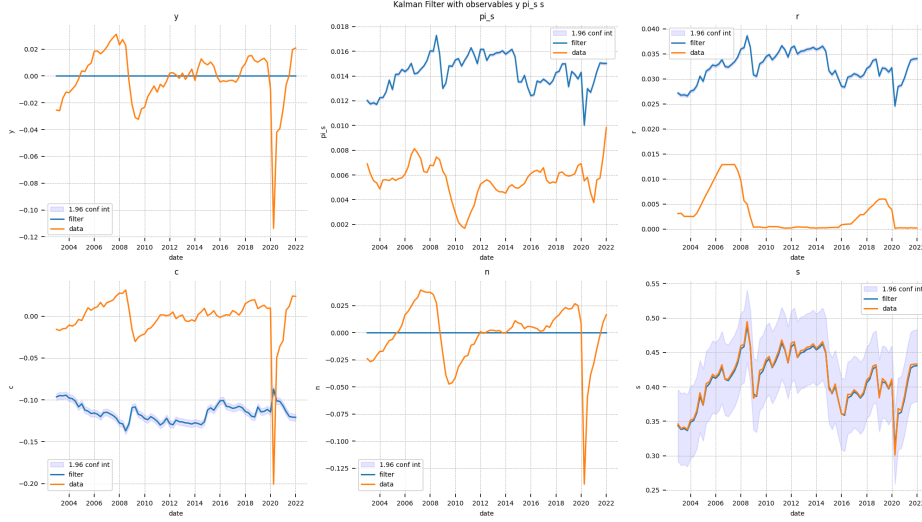


Figure 13: NK model Kalman Filter with variables y

Appendix C

Log-linearisation

Euler equations

Expression for N_t from N_{it} It can be shown that $\log \left(\int_0^1 \left(\frac{P_{it}}{P_t} \right)^{-\frac{\epsilon}{1-\alpha}} di \right) \approx 0$ which then allows to log linearise the production function to $y_t = a_t + (1 - \alpha)n_t$.

RBC all equations

Equation name	Log-linear expression
Euler equation	$\mathbb{E}_t \left[\frac{C_{t+1}}{C_t} \right]^\sigma = \beta [(1 - \delta) + \mathbb{E}_t[R_{t+1}]]$
Capital supply	$K_t = (1 - \text{delta})K_{t-1} + I_t$
Capital demand	$K_t = \alpha A_t K_t^{\alpha-1} N_t^{1-\alpha}$
Labour supply	$-\frac{U_{n,t}}{U_{c,t}} = W_t$
Labour demand	$W_t = (1 - \alpha)A_t K_t^\alpha N_t^{-\alpha}$
Production function	$Y_t = A_t K_t^\alpha N_t^{1-\alpha}$
Equilibrium condition	$Y_t = C_t + I_t$
Technology law of motion	$\log(A_t) = (1 - \rho_A) \log(A^*) + \rho_A \log(A_{t-1}) + \epsilon_t^A$

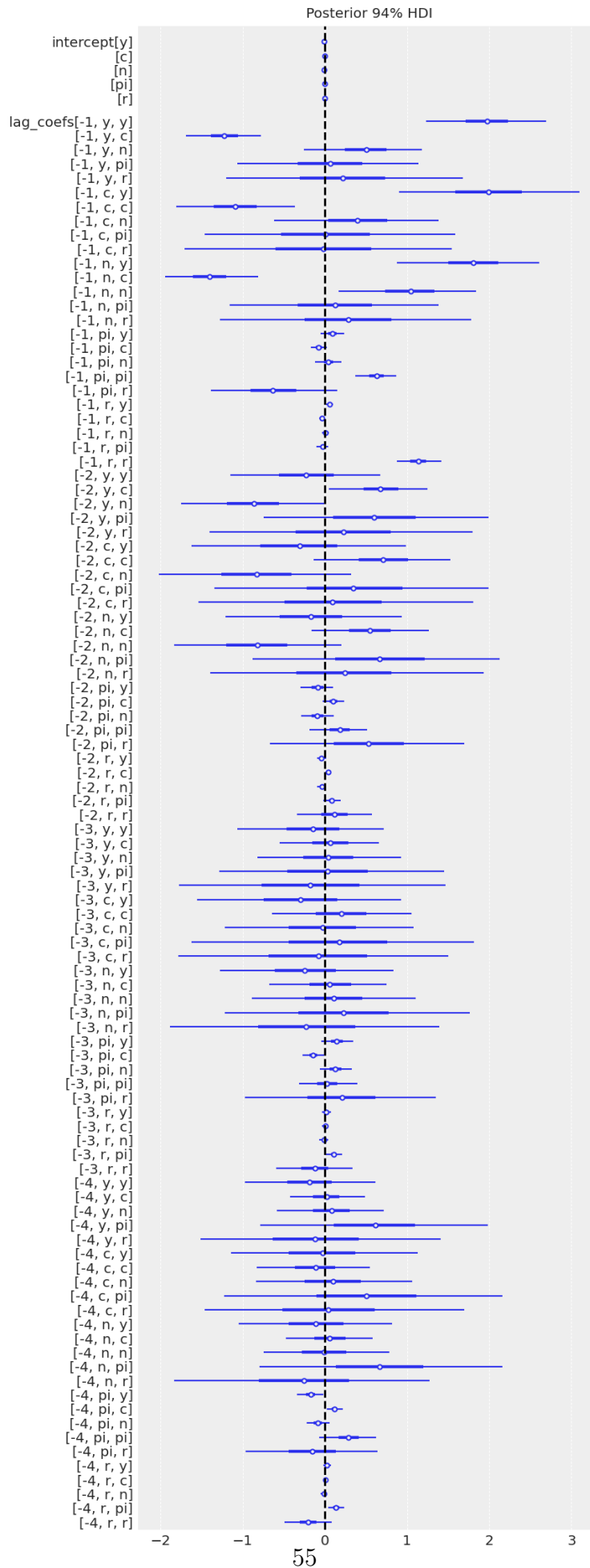


Figure 14: NK model Kalman Filter with variables y

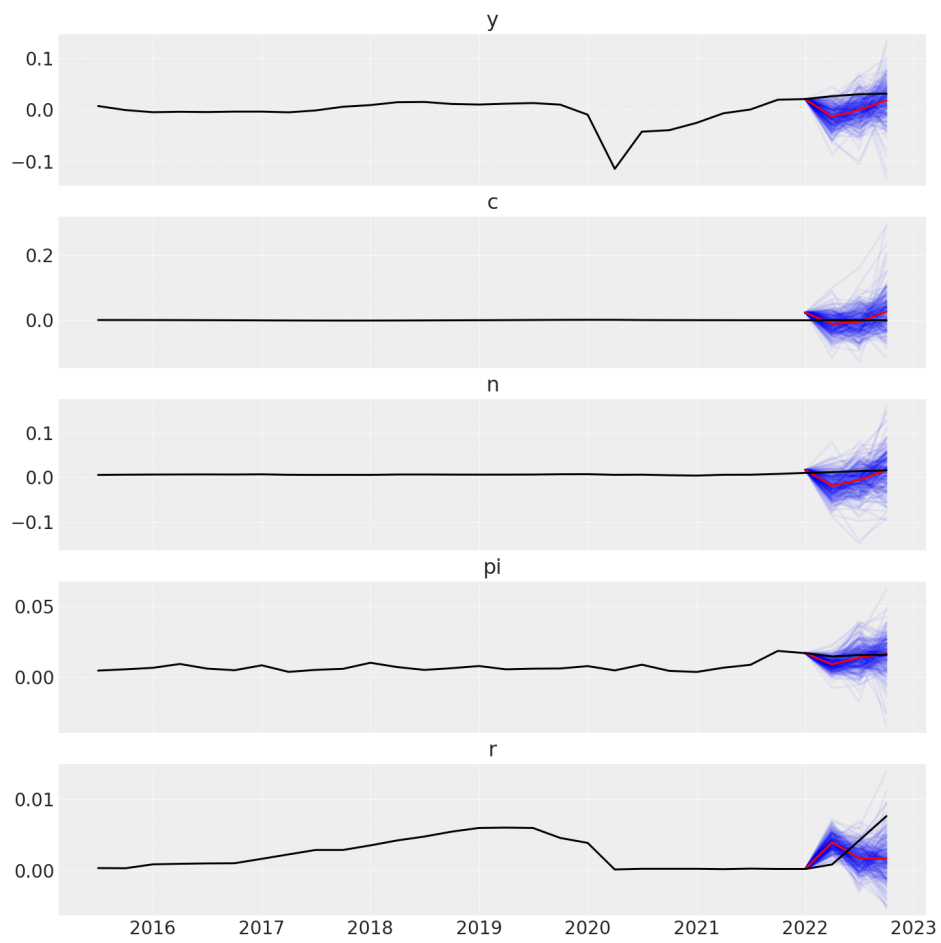


Figure 15: NK model Kalman Filter with variables y

Table 2: RBC steady state values

	0	1
0	A _{ss}	1.000
1	C _{ss}	2.358
2	L _{ss}	0.715
3	K _{ss}	35.732
4	L _{ss}	0.820
5	Y _{ss}	3.073
7	r _{ss}	0.030
8	w _{ss}	2.436

Table 3: RBC transition matrix

	A	C	I	K	L	Y	r	w
A	0.95000	-0.0	-0.0	-0.00000	-0.0	-0.0	-0.0	-0.0
C	0.30966	0.0	0.0	0.47875	0.0	0.0	0.0	0.0
I	3.64070	-0.0	-0.0	-0.51273	-0.0	-0.0	-0.0	-0.0
K	0.07281	-0.0	-0.0	0.96975	-0.0	-0.0	-0.0	-0.0
L	0.20660	0.0	0.0	-0.15665	0.0	0.0	0.0	0.0
Y	1.08429	0.0	0.0	0.24818	0.0	0.0	0.0	0.0
r	1.08429	0.0	0.0	-0.75182	0.0	0.0	0.0	0.0
w	0.87769	0.0	0.0	0.40483	0.0	0.0	0.0	0.0

Appendix D

Metropolis Hastings MCMC sampler code

Deriving the multivariate covariance matrix G from prior the prior distributions standard deviation

```
1 # get sigma and scaling factor for the random walk law of motion
2 p, s = get_Sigma(prior_dist, mod_params, shock_names)
3 G_params = np.array(list(p.values())) #
4 G_shocks = np.array(list(s.values()))
5
6 G_params = np.zeros(len(p)) + 1
7 G_shocks = np.zeros(len(s)) + 1
8
9
10 scaling_factor = .4
11
12 # get final output
```

Suggesting new priors based on the random walk law of motion

```
1 # sample from priors according to random walk law of motion
2 # a multivariate normal distribution wit G as standard deviation
3
4 # prior for parameters
5 prior = np.array(list(old_posterior_p.values())) +
multivariate_normal(
6     list([0] * len(old_posterior_p)), scaling_factor * G_params).
    rvs())
7
8 # prior for shocks
9 shocks = np.array(list(old_posterior_s.values())) +
multivariate_normal(
10     list([0] * len(old_posterior_s)), scaling_factor * G_shocks).
    rvs())
11
12 # put priors into dictionary for further process
13 prior, shocks = dict(zip(old_posterior_p.keys(), prior)),\
14     dict(zip(old_posterior_s.keys(), shocks))
```

MH MCMC acceptance procedure

```

1      # compare current prior likelihood with the likelihood of the most
      recent draw accepted into the posterior
2      ratio = (new_loglike * prior_pdf / old_loglike * mask_old_post).
mean()
3      omega = min([ratio, 1])
4      # save ratios
5      draw_dict['ratio'] = ratio
6      draw_dict['omega'] = omega
7
8      # MH sampler random acceptance, based on uniform distribution U
(0,1)
9      random = np.random.uniform(0, 1)
10     if random <= omega:
11         counter_accp += 1.
12         draw_dict['is_accepted'] = True
13
14     # update param posterior
15     for key in prior.keys():
16         param_posterior_list[key].append(prior[key])
17
18     # update shock posterior
19     for key in shock_posterior_list:
20         shock_posterior_list[key].append(shocks[key])

```

7.1 Appendix E

```
#####
# Real Business Cycle model
#####

tryreduce
{
    U[], Tc[];
};

block HOUSEHOLD
{
    definitions
    {
        u[] = (c[] ^ (1 - sigma_C) / (1 - sigma_C) - n[] ^ (1 + sigma_L) / (1 + sigma_L));
    };

    controls
    {
        c[], n[], I[], k[];
    };

    objective
    {
        U[] = u[] + beta * E[] [U[1]];
    };

    constraints
    {
        c[] + I[] = r[] * k[-1] + w[] * n[] : lambda[];
        k[] = (1 - delta) * k[-1] + I[];
    };

    calibration
    {
        beta = 0.99;
        delta = 0.02;

        sigma_C = 1.5;
        sigma_L = 2.0;
    };
};

block FIRM
{
    controls
    {
        k[-1], n[];
    };

    objective
    {
        Tc[] = -(r[] * k[-1] + w[] * n[]);
    };

    constraints
    {
        y[] = a[] * k[-1] ^ alpha * n[] ^ (1 - alpha) : mc[];
    };

    identities
    {
        # Perfect competition
        mc[] = 1;
    };
};
```

```

    };

    calibration
    {
        alpha = 0.35;
    };
};

block TECHNOLOGY_SHOCKS
{
    identities
    {
        log(a[]) = rho_A * log(a[-1]) + epsilon_a[];
    };

    shocks
    {
        epsilon_a[] ;
    };

    calibration
    {
        rho_A = 0.95;
        sigma_epsilon_A = 0.05;
    };
};

#####
# New Keynesian model (Gali, 2008)
#####

block LINEAR_MODEL
{
    identities
    {
        # NKP
        pi[] = beta * E[] [ pi[1] ] + kappa * y[] ;

        # NKIS gap and real
        y[] = E[] [ y[1] ] - 1 / sigma_C * (i[] - rn - E[] [ pi[1] ] ) ;

        # labour
        y[] = a + (1 - alpha) * n[] ;

        # wage
        # w[] - p[] = sigma_C * y[] + sigma_L * n[] ;

        # CB
        i[] = rho + phi_pi * pi[] + phi_y * y[] + v[] ;

        # real interest rate
        r[] = i[] + pi[] ;

        # equilibrium
        c[] = y[] ;
    };

    calibration
    {
        # natural rate of interest could be exogenous law of motion
        rn = .01 ;

        # technology
        a = 1;

        # discount factor

```

```

beta = .99 ;
rho = -log(beta) ;

# capital share - N_t^(1-alpha) even in absence of K_t
alpha = .33 ;

# CB preference
phi_pi = 1.1 ;
phi_y = .4 ;

# Dixit-Stiglitz aggregator CES
epsilon = .8 ;

# Calvo pricing
theta = .75 ;

# labour mrs
sigma_L = 1.5 ;

# consumption mrs
sigma_C = 1.5 ;

# aggregated parameters

Theta = (1 - alpha) / (1 - alpha + alpha * epsilon) ;
param1 = (1 - theta) * (1 - beta * theta) / theta ;

kappa = param1 * (sigma_C + (sigma_L + alpha) / (1 - alpha)) ;

};
};

block SHOCKS
{
    identities
    {
        # monetary policy shock
        v[] = rho_v * v[-1] + epsilon_v[] ;
    };
    shocks
    {
        epsilon_v[];
    };
    calibration
    {
        rho_v = .9 ; # AR(1) monetary policy
    };
};

#####
# New Keynesian model with petrol sector (Blanchard & Gali, 2007)
#####

block LINEAR_MODEL
{
    identities
    {

# NKP
pi_q[] = beta * E[] [ pi_q[1] ] + kappa * y[] + lambda_p * Gamma_s * s[] ;

# NKIS
y[] = E[] [ y[1] ] - (i[] - E[] [ pi_q[1] ]) + alpha_m * (1 - rho_s) / (1 - alpha_m) * s[] ;

```

```

# monetary policy
i[] = phi_pi * pi_q[] ;

# real interest rate
r[] = i[] + pi_q[] ;

# consumer price index
pi_s[] = pi_q[] + chi * (s[] - s[-1]) ;

# labour
n[] = (y[] * (1-alpha_m) - 1) / alpha_n ;

# consumption
c[] = y[] - s[] * (alpha_m / (1 - alpha_m) + chi) ;

};

calibration
{

alpha_m = .4 ;
alpha_n = .4 ;

beta = .995 ; # time discount factor

gamma = 0 ; # real wage rigidity, at gamma = 0 -> w[] = mrs

phi_pi = 1.1 ; # CB monetary policy

theta = .75 ; # Calvo pricing

chi = 0.03 ; # natural share of oil in consumption

epsilon = 1.5 ; # elasticity of substitution between domestic goods and consumption
mu = epsilon / (1 - epsilon) ; # average mark-up

# Aggregated parameters

eta = alpha_m / (mu - alpha_m) ;

Gamma_n = ((1 - alpha_n - alpha_m) * gamma + (1 - alpha_m) * (1 - gamma) * (1 + phi_pi)) / (1 - (1 - gamma) * (alpha_m - (1 - alpha_m) * eta)) ;

Gamma_a = gamma / (1 - (1 - gamma) * (alpha_m - (1 - alpha_m)) * eta) ;

Gamma_s = (gamma * (alpha_m + (1 - alpha_m) * chi)) / (1 - (1 - gamma) * (alpha_m - (1 - alpha_m)) * eta) ;

lambda_p = ((1 - theta) * (1 - beta * theta) / theta) / (1 + (1 - alpha_m + alpha_n) * (epsilon - 1)) ;

kappa = lambda_p * Gamma_n * (1 - alpha_m) / alpha_m ;

};

};

block OIL_PRICE_SHOCK
{
    identities
    {
        # real price of oil law of motion
        s[] = rho_s * s[-1] + epsilon_s[] ;
    };
    shocks

```



```
{
  epsilon_s[];
};
calibration
{
  rho_s = .9 ; # AR(1) parameter real oil price
};
};
```

Test