

Memoire

Lukas Gahl

1 Introduction

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 macroeconomic 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 RBC model

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 ressources

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 succesfully 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 emperical evidence (Gali, 2008, p 15)

Assumptions - perfect compitions - flexible prices - technology shocks - infinitely lived households and firms - identical technology across firms

Agents - representative households, where the sum of all households is normed to one - Firms, the representative firm normed to unity

Allocation decisions HH - intertemporal consumption and leisure choice - intertemporal consumption savings

Firm - static optimisation of profit

the model - Non-linearity arises from multiplicative Cobb-Douglas production and additive law of motion of e.g. capital (Campbell) - This raises the need for linear approximations, overview of possible methods: - Models become more complex as researchers are trying to use more realistic functions of utility and incorporate heterogeneity (Taylor Uhlig, 1990) - 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)

Methods of linearisation - This work builds on models that can be solved analytically - All models are solved analytically and then log-linearised with 1st order Taylor approximation following (Campbell) - Based on the log-linearisation the predictions are performed - Log-linear quadratic approximation Rebello (1987) - Under a deterministic singular solution it solves correctly (Campbell) This work follows Campbell and uses the capital stock as the state variable As there is no multiple steady states, this work solves the models analytically

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" [CE92]

3 NK

Additional NK assumptions (Gali) - monopolistic competition, inputs are set by private agents according to their own optimisation problem - nominal rigidity, price setting is limited in frequency - this results in short-run neutrality of monetary policy where changes in interest are not directly matched by changes in expected inflation - this is the source of short-run fluctuations - However, in the long-run prices adjust - This causes the response of the economy to be inefficient, unlike RBC where cycles are result of efficient adjustments

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

Small scale new keynesian model solved with first order log-approximation [HS14]

Define stability condition of the linear state-space, Blanchard-Kahn

4 Petrol

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 the inclusion of petrol as an exogenous shock series [KL92]

Part I

Bringing the model to data

5 The Data

Data sources References for data sources

6 Preprocessing

per capita log HP filter - what is natural GDP - what is cyclicity - why does it yield log-deviations

7 Descriptives

Descriptives Plot data

8 Analysis of models

Compare model covariance matrices, to actual covariance in data

Show theoretical impulse response functions, differences between NK and RBC

9 Estimation methods

9.1 Frequentist estimation

Traditional calibration of DSGE models relied on guessing or taking averages.

Early attempts of linking DSGE models to data have mainly relied on frequentists statistics in using maximum-likelihood estimators (MLE) or general method of moments (GMM). The purpose of these methods lied in identifying the parameters of structural economic models (ref). These parameters are expected to be stable over time, wherefore they can be estimated off large time series.

One such attempt has been pioneered by [CE92] who employ GMM to discriminate between two model specification. Comparing the estimated parameters of the model with real data values then serves as inspection of model fit to data.

Short-coming of GMM is that with the growing number of parameters the moment conditions are no longer sufficient to estimate model fit [GQN13]. Likewise, the ML estimator is limited in its ability to estimate the entirety of parameters, requiring manual calibration or assumptions [GQN13].

While frequentists estimation knew to overcome the identification problem using theoretical movements as suggested by Smith (1993) the frequentist approach was soon deemed inappropriate to DSGE estimation. - Frequentists assume the model to be the 'true' model, if not the estimation process is biased and needs adaption - Many adaptations have been suggested by e.g. Smith (1993) or Dridi, Guay, and Renault (2007) who select the parameters relevant to overall model moments for the econometric estimation. The non-relevant parameters are then kept constant throughout the estimation, yet it is argued that this approach fails to overcome the conditional bias introduced by non-relevant parameters [GQN13]. - It is thus a question of modelling philosophy that lead to frequentists methods being surpassed by Bayesian estimation [GQN13].

9.2 Bayesian estimation

For the majority of DSGE models the prior distribution cannot be directly analysed [HS14]

Bayesian estimation as an alternative [GQN13]. - Poirier (1998) also suggests a separation of parameters into two groups - Bayesian statistics circumvents any assumption about a singular true model in referring to conditional likelihood of a model specification

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

9.2.1 Sampler

There are local and global identification problems [HS14]

Mapping a DSGE model to its posterior is non-linear in Θ , wherefore the posterior distribution cannot be evaluated analytically [HS14]. It thus requires a numerical solution [GQN13].

Metropolis Hastings belongs to the Markov Chain Monte Carlo (MCMC) group of algorithms. It generates a stable Markov chain, which is equivalent to the posterior distribution [HS14]

Metropolis Hastings sampling algorithm - a multivariate random walk to update the posterior distribution - the posterior then is updated based on improvements in the overall model likelihood - 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 [CL19]

MH-MCMC sampler - predominant estimation method in linear Bayesian estimation literature [GQN13]

- the MH sampler is applicable to small scale New Keynesian models and usually delivers good results on this scale. For larger models a different kind of sampling is required [HS14]. As this thesis is concerned with such small scale model it will follow the Metropolis Hastings sampling procedure.

The algorithm proceeds as follows [HS14]

1. A new draw ζ is suggested based on a random draw from the prior distribution of parameters Θ . In order to determine whether this draw is admissible to the posterior $\Gamma_{1:i}$ the following steps are followed.
2. The new draw ζ is evaluated in likelihood using the DSGE model's likelihood function given the data Y , yielding $p(Y|\zeta)$.
3. In a two stage procedure the algorithm then either accepts or rejects the new draw ζ as part of the Markov chain of posterior distributions.
4. First the posterior likelihood of ζ is calculated based on Bayes theorem. It is then compared to the the posterior likelihood of the last accepted draw, that is now part of the posterior Γ_{i-1} .

$$\omega_i = \min\left\{\frac{L(Y|\zeta, \Gamma) * P(\zeta)}{L(Y|\Gamma_{i-1}, \Gamma)P(\Gamma_{i-1})}, 1\right\}$$

5. The second step of the decision algorithm draws from a uniform random variable $\psi \sim U(0, 1)$. The new draw ζ is accepted if its likelihood ration ω_i is accepted if $\omega_i \geq \psi$. This step is equivalent to moving to a higher point on the likelihood surface [HS14]
6. If accepted ζ will become Γ_i
7. The algorithm is then repeated

How many runs

As illustrated in the above MH-MCMC is a recursive method and as such requires an initial value for the posterior Γ_0 . 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

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 [HS14]

9.2.2 The log-likelihood

The Kalman filter serves as means of evaluating the log-likelihood of a linear state space system. The filter is a good remedy, as for a given system the Kalman Filter is the optimal forecast (ref). This of course is conditional on the belief that the underlying system is normally distributed. - Linear Gaussian state space system for the Kalman filter, this requires shocks to be normally distributed [HS14] - For the non-normal case a Kalman-Particle filter is used, which however is not discussed here

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

9.3 Priors

The assumption for prior selection is that priors are obtained from some past knowledge. If this knowledge is independent of the data used to evaluate the likelihood function then the prior holds. This is fulfilled if the data for prior building at least predates the data for likelihood evaluation [HS14].

Priors are statistical distributions which either reflect beliefs about the structural mechanism or are derived from data, external to the estimation process [DNS08] - Priors thus reflect microeconomic evidence, for example regarding labour supply and elasticity - However, they are a simplification of the real joint likelihood function of the structural model. Such function is difficult and sometimes impossible to evaluate

Priors [DNS08] - Priors are usually assumed to be independent for simplicities sake - Priors are often calibrated for the model to match the covariance of real data

[DNS08] suggest two kinds of data sources. Pre-sample data, which is not included in the model evaluation process or in-sample data, that is explicitly excluded from the model evaluating likelihood function. As the model parameters are thought of as structural, hence relatively time-invariant, more dated data can be used to estimate them.

Priors are divided into three groups, following [DNS08]

9.3.1 Steady state priors

The 'great ratios' or long-run relationships as [KP82] referred to them pin down the steady state

$$\theta_{ss} = \{\alpha, \beta, \delta, \pi^*, \eta_p\}$$

When determining the steady-state priors θ_{ss} and assuming independence it is possible that the resulting joint prior distributions assigns high probability to unrealistic steady state values [DNS08]. To circumvent this issue other method

This priors are usually estimated from pre-sample data [HS14] this work follows this approach and thus

9.3.2 Exogenous priors

The exogenous propagation parameters these are the exogenous shock parameters, their autoregressive coefficient and their standard deviation

$$\theta_{exog} = \{\rho_a, \rho_s, \sigma_a, \sigma_s\}$$

Exogenous priors can be evaluated from the model specification in keeping the influence of other priors minimal. Del Negro suggests, to set all other parameters to zero where possible. The system of equations then reduces to variables that are depend on the exogenous shocks. This reduced system can then be used to evaluate the likelihood of prior values, using out-of sample data of the later calibration. An evaluation based on the likelihood of the reduced system is not necessary if the variation in observable variables can be directly retraced to individual shocks. In a simple model where shocks do not intersect in their influence on observable variables, e.g. in the case of a singular technology shock the variance of the shock can be directly inferred from the observable variable. In case of intersecting multiple shocks the likelihood of the reduced system needs to be consulted to choose appropriate priors, as their influence on observables cannot be directly inferred.

In case of exogenous observable shocks, such as energy prices, the above procedure is not required. Other than the level of technology energy price shocks are measurable, wherefore one can rely on the observable first and second order moments (ref).

9.3.3 Endogenous priors

The endogenous propagation parameters, including the shock autoregressive coefficients are parameters involved with the propagation of shocks through

Table 1: Prior distributions

model param	mean			std		
	NK	NK Energy	RBC	NK	NK Energy	RBC
alpha	0.285714		0.285714	0.02551		0.02551
alpha_m		0.285714			0.02551	
alpha_n		0.285714			0.02551	
epsilon		0.5			0.073529	
epsilon_A	0.099099		0.099099	0.007378		0.007378
epsilon_R	0.099099			0.007378		
epsilon_T	0.099099			0.007378		
epsilon_Y	0.099099			0.007378		
epsilon_pi	0.099099			0.007378		
epsilon_s		0.099099			0.007378	
eta_p	0.746269			0.013149		
gamma_R	2.0			1.0		
gamma_Y	2.0			1.0		
gamma_pi	2.0	2.0		1.0	1.0	
mu		9.440416			56.352936	
sigma_C	2.0	2.0	2.0	4.0	4.0	4.0
sigma_L	2.0	2.0	2.0	4.0	4.0	4.0

the system. These parameters heavily rely on micro-evidence from external data sets [DNS08] All remaining parameters are stacked into the the following

$$\theta_{endog} = \{\sigma_C, \sigma_L, \}$$

} Priors of the remaining variables are chosen in order to reflect believes or micro evidence. This work draws from the choice of [DNS08] in their endogenous priors, given the similarity in models

Priors are chosen in order to match first and second order moments individually as well as in a joint distribution across all parameters [DNS08]

9.3.4 Reference Model BVar

[Sch00]

BVar with normally independently distributed lag parameters [CL19]

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 [CL19]

9.4 Forecasts

Relevant to policy recommendations How well can models capture the dynamics of the economy, e.g. as a spectrum of possible outcomes

Including priors into bayesian VAR makes the model better at forecasting [CL19]

To compare models this work follows a bayesian averaging approach as in [CL19] - (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

BMA - the weights of each model are calculated from its posterior model probability [CL19] - This is traditionally done for structurally similar models in order to discriminate between different priors - the posterior probability is thereby obtained from [CL19]

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

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

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" [CL19]

Comparing forecasting performance Diebold-Mariano test for point forecast comparison following [CL19] - MSE based point forecast comparison Amisano-Giacomini test for density forecast performance following [CL19]

References

- [CE92] Lawrence J. Christiano and Martin Eichenbaum. Current real-business-cycle theories and aggregate labor-market fluctuations. *The American Economic Review*, pages 430–450, 1992. Publisher: JSTOR.
- [CL19] Kuo-Hsuan Chin and Xue Li. Bayesian forecast combination in VAR-DSGE models. *Journal of Macroeconomics*, 59:278–298, March 2019.
- [DNS08] Marco Del Negro and Frank Schorfheide. Forming priors for DSGE models (and how it affects the assessment of nominal rigidities). *Journal of Monetary Economics*, 55(7):1191–1208, October 2008.
- [GQN13] Pablo A. Guerrón-Quintana and James M. Nason. Bayesian estimation of DSGE models. In *Handbook of research methods and applications in empirical macroeconomics*, pages 486–512. Edward Elgar Publishing, 2013.
- [HS14] Edward Herbst and Frank Schorfheide. Bayesian inference for DSGE models. *Manuscript. Board of Governors of the Federal Reserve System, Washington, DC*, 2014.
- [KL92] In-Moo Kim and Prakash Loungani. The role of energy in real business cycle models. *Journal of Monetary Economics*, 29(2):173–189, April 1992.
- [KP82] Finn E. Kydland and Edward C. Prescott. Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, pages 1345–1370, 1982. Publisher: JSTOR.
- [Sch00] Frank Schorfheide. Loss function-based evaluation of DSGE models. *Journal of Applied Econometrics*, 15(6):645–670, 2000. Publisher: Wiley Online Library.

Part II

Appendix

10 Appendix A

10.1 RBC

