

On the analytical complexity of the
likelihood for a simple DSGE model.

Tore Bersvendsen

Master thesis in business administration

Agder University College
Faculty of Economics and Social Sciences

2008

Preface

This master thesis is a part of the master program in business and administration at University of Agder in Kristiansand. The master thesis is mandatory in the last spring semester of the two year long master program. The total master program is given 120 ECTS, and this thesis is 30 of them.

I choose to write about macro economy because I have great interest in this subject, and I know the importance of understanding how world economy is connected, and especial how the Norwegian economy are connected. This thesis also touches subjects like micro economy, econometrics and statistics. I choice to write about DSGE models because it is the “new” thing among macroeconomist. Many central banks are using this technology, and the Norwegian central bank is one of them. My supervisor motivated me for this subject, and one of the reasons is that it could be wise to know about these models, because they play an important role for central banks. I had never heard about these models, so it was much work needed just to understand on how these models works and the structures behind them. It was also much work needed to learn some new mathematical techniques.

I wish to express my gratitude to my supervisor Jochen Jungeilges, for needed, essential and helpful guidance during this whole process. I will also thank him for good ideas and important feedback on my thesis. I will also use the opportunity to thank my family and my friends for the support and for keeping me motivated during the whole studying period, and especially during this thesis.

Kristiansand, June 4, 2008

Tore Bersvendsen

Contents

| | |
|---|----|
| 1. Introduction | 1 |
| 2. The Background for DSGE models | 2 |
| 2.1 The Beginning | 2 |
| 2.2 Evolution | 4 |
| 3. The model | |
| 3.1 Households | 7 |
| 3.2 Firms | 8 |
| 3.3 Linearization of the model | 9 |
| 4. Estimation Method | 15 |
| 4.1 Maximum likelihood: General Theory | 15 |
| 4.1.1 Maximum likelihood: For our model | 17 |
| 4.2 Kalman filter | 21 |
| 5. Conclusion | 22 |
| Appendix A | 23 |
| Appendix B | 27 |
| Appendix C | 30 |
| Appendix D | 32 |
| Appendix E | 33 |
| References | 36 |

1. Introduction

In this paper I will present a one-sector real business cycle (RBC) model, and see how the complexity and the structure behind a dynamic stochastic general equilibrium (DSGE) model. The motivation of this paper is to learn how these models work. One reason for this motivation is that these models are beginning to be central in different decision making processes in economics. Especially central banks are using DSGE models in their decisions making process. Since the central banks are using these types of models, this thesis will see in short terms on how these models have developed. The RBC model that will be presented in this paper comes from Ruge-Murcia (2007). Since central part of the model is analyzed numerically much of the underlying structure is hidden. Is my goal to ravel some of the underlying structure that is important for the understanding on the complexity of the likelihood function.

Maximum likelihood estimation (MLE) is often used when DSGE models are estimated. There is also three other methods: Generalized method of moments, simulated method of moments and the indirect procedure proposed by Smith (1993). Ruge-Murcia (2007) uses all these methods and compares them. I will present the all the preparation behind the likelihood function so we can be able to use MLE. To use MLE in the Ruge-Murcia (2007) model we have to estimate values for an unknown parameter. This is done with a technique called Kalman filter. In this paper I will outline an introduction to this technique, and also connect it with our specific case. When presenting this model analytically we will easily see how complex these models are. The main result for this paper is that it reveals the DSGE model interdependency of the structural variables in a very non-linear way. The maximum likelihood function presented in Ruge-Murcia (2007) seems really straight forward due to the notation he uses, but the likelihood depends on the structural parameters in a very complex way. I will present five additional results to the Ruge-Murcia (2007) article. The first four results are needed for computing the likelihood function in just structural parameters. For the sake of readability the proof for the additional results, have been organized in an appendix. The text will then not be so interrupted with mathematical expressions. Major bulk of work went into the analytical proofs. Still they are organized in appendix for the readability.

The paper will be organized as follow: Section 2 will outline the historic background for DSGE models and also some facts about these models. Section 3 will present the model of

Ruge-Murcia with the additional proofs and results. Section 4 describes the MLE in general and for this model. In this section I will also provide a short sketch of the Kalman filter. Section 5 contains the conclusion part.

2. The Background for DSGE models

DSGE stands for “dynamic stochastic general equilibrium” model and it is a modern macroeconomic model. This is a long-term general equilibrium model, and is derived from microeconomic principals. These models play an important role for economist and actors that are interesting in analyzing monetary policies at a macroeconomic perspective. Many central banks including Norway is already using these types of models in their decision making process¹.

One aspect of these models is that it includes one or more random shocks. This is where the stochastic element is included in the model. But because of the fluctuation over time and randomness these models are more difficult to use for people with little mathematic background. One typical feature of DSGE models is that the behavior of the different economic agents is modeled explicitly and founded on choice-theoretic assumptions. What’s making DSGE models singular are because they use a small number of structural shocks to generate predictions about a large number of observable variables.

The DSGE model that will be presented later in this paper is a typical earlier DSGE model. Here it will be more focus on the supply side, because it will contain a technology shock that affects the firms. These models have feature from the new classical macro view and are often called “real business cycle” models, because it’s assumed that the economy have perfect competition and fully flexible prices.

2.1 The Beginning

In 1936 J.M. Keynes² published a monograph that revolutionized the thinking on economy, and especially macro economy. He provided a model that could be used to give a better understanding of macro economy, and how different policies would affect the economy. Even

¹ See for example Bank of England (2004) and Fenton and Murchison (2006)

² Keynes, J.M. (1936). *The General Theory of Emploment, Interest, and Money*. Macmillan. London

in the late 1970s and 1980s central banks were relying on the Keynesian paradigm. But in the late 60s a famous economist called Milton Friedman argued against Keynes, and tried to start a new revolution in macroeconomic thinking. This was the start of the monetarist. The biggest discussions between Keynes and the monetarists were on how they interpret interest sensitivity and monetary policy. According to Keynes monetary policy will not help, because the additional money will simply be absorbed by investors with no noticeable effect on interest rate. Fiscal policy, on the other hand, will work really well³. The monetarist thinks that monetary policy has an effect, but the policy maker is however not very good at timing monetary policy. Under monetarist assumptions fiscal policy is unable to influence employment and output⁴. This is the main cause monetarists are against the Keynesians thinking.

In 1961 John Muth published an article that argued that the modeling of expectations wasn't good enough. Muth proposed that: expectations, since they are informed predictions of future events, are essentially the same as the predictions of the relevant economic theory⁵. Muth formulated hypothesis that claimed that the economy generally does not waste information, and the formulation of the expectations depends specifically on the structure of the entire system describing the economy. The two main conclusions from his study of expectations data can be summarized: The average expectations in an industry is quite accurate, and reported expectations seems to underestimate the actual change that takes place.

Then Robert E. Lucas in 1976 presented a radical new way of thinking in econometric models. He used Muth's (1961) thinking about rational expectations in his discussion, and argued that the popular macroeconomic models that economists were using at that time was totally useless. In evaluation the effect of different type of economic policy they could not be used. He meant that older standard models do not match several important characteristics of econometric practice. By adding a new general structure that includes stochastic parameters, the models will be much closer to these characteristics. The work of Lucas convinced many in the game that using rational expectations would require large adjustment in the models they already were using, and that it will deliver different theoretical outcomes. This also led to more micro based macroeconomic models. He argued that we have to look at the structure of the model, not just the parameter value. The method that often was used is to look at historical

³ Heijdra, B.J. and F. van der Ploeg (2002): The Foundations of Modern Macroeconomics, page 20

⁴ Heijdra, B.J. and F. van der Ploeg (2002): The Foundations of Modern Macroeconomics, page 23

⁵ J. Muth: Rational Expectations and the Theory of Price Movements, page 316

data, look after correlations and then try to formulate a function that fits this data. The new way of thinking was to try and formulate functions for a representative agent that can include expectations. Both on the supply- and demand side of the economy. We can summarize the Lucas critique like this: If the structure of a model is affected by political methods of operations, then it will be useless to analyze changes in the same methods of operations⁶.

2.2 Evolution

After the Lucas critique there has been a large development in these types of models. New ways of solving and estimate complicated mathematical functions gives us more opportunities. But the idéé behind DESG models we already can find back in the 1970s. According to the Norwegian DSGE model NEMO (Brubakk, Husebø, Maih, Olsen and Østnor 2006) the new DEGE models also contains some New Keynesian aspects in the short-run. This because they have two additional features: Nominal rigidities and Monopolistic competition. New Keynesian economists argued that economy contained imperfect competition, and because of this there are also sticky prices. There were also many new Keynesian economists argued and used very rigid wages. Because of these two additional features the model acts in a different way. Even if there are nominal rigidities in the short run, the prices and wages are assumed to be fully adjusted in the long run. Monetary policy can affect the economy in the short run, but in the long run it can only affect nominal variables.

Rational expectation policy models was introduced in the 1980s, this gave the macroeconomic models ha deeper insight in important economical issues, like exchange rate overshooting. Macroeconomic models that where developed in the late 1980s and 1990s were focusing on rational expectations, but also trying to use more micro-based relationships in modeling the different agents. These models were often called real-business-cycle models where prices were fully flexible (Kydland and Prescott 1982). These models developed, and an important improvement was that they started to include some form of normal inertia. To find an exact date when DSGE models came is quite difficult. But the last ten years the DSGE models development is quite enormous. In 1992 Bernanke and Blinder presented a model that tried to see how monetary policy affects the real economy. This had already been done in an IS-LM view, not in a more structural micro based setting. Here they try to measure the effects of a

⁶ Mork, K.A. (2008): Modeller og pengepolitikk. *Samfunnsøkonomene*, nr. 1, page 12

policy change by modeling a policy shock that could be measured. One catch is that they isolate the direct measure of Federal Reserve policy, so it does not become completely stochastic. Leeper and Sims (1994) presented a dynamic general equilibrium model. They thought their model was a potential competitor to the standard IS-LM based models. This model has many shocks and stochastic elements. It already has many elements of a modern DSGE model, for example: different agents are explicitly modeled. There are people who identify Rotemberg and Woodford (1997) as the start of a New Keynesian DSGE model.

They try to derive a complete structural model to answer the Lucas critique. They use an optimization-based approach, and have a variable for monetary policy shock. They also use monopolistic competition. If you go through this model we can see elements that are used in later articles. An example of this is the use of a monetary policy shock. Other articles that modify this are Rotemberg and Woodford (1999) and Gali (2000). An important article for the development of today's DSGE models is Christiano, Eichenbaum and Evans (2001). They use both Leeper and Sims (1994) and Rotemberg and Woodford (1997) as a starting point. Their model tries to prevent a large rise in marginal cost after an expansionary shock to monetary policy. They seek to understand the observed inertial behavior in inflation and persistence in aggregate quantities. The model GEM (Bayoumi with assistance 2003) uses the CEE (2001) model as a starting point. But making it more international, not just focusing so much on one US. Another model that uses CEE (2001) very much is Smets and Wouters (2003). They focus on the euro area. They add more shocks to their model, and also they have even more specified functions for the different agents in the economy. This model is often used when different central banks formulate their DSGE models. In 2005 Christiano, Eichenbaum and Evans (CEE) took their old model and added some more empirical work and tests. They try to answer the question: Can models with moderate degree of nominal rigidities generate inertial inflation and persistence output movements in response to a monetary policy shock? Their answer is yes. Belaygorod and Dueker (2005) use CEE (2005) article, but focus more on the central banks with using inflation targeting. They mean that the promise of estimated DSGE models is that one can take the parameter estimates, plug them into the underlying optimizing model, and perform welfare calculations. This is why policymakers should know the benefits of interest rate smoothing. They try to get a sharper specification of interest rate smoothing into a DSGE model. The NEMO model used in the central bank of Norway is estimated using techniques proposed in Bayoumi (2003) and Smets and Wouters (2003).

The use of models in monetary policy decisions is important. A model cannot give us the correct answer. But they are very helpful for the central-banks. Norway uses these models as a guideline for where the economy is heading in the future. It is important to take into consideration that different models have different properties. Many central-banks use many different models. In Holmsen, Nicolaisen and Røisland (2007) paper they recommend not using just one model for the monetary policy analysis. It is of course important that a decision is not just based on the output from the model(s). General economic theory and experience from other countries is also important input in a decision making process. Models are a good tool for helping us to understand different aspects of the economy. But they always have to continue their development, so that they could be more accurate. This will make decisions a lot easier.

According to Mork (2008) models that the central bank of Norway uses have to give us the most correct picture on how important economic variables like inflation will be affected by different political decisions⁷. He means that the only way to achieve this is by using totally structured models. These models have to explicitly use the representative agent's behavior, expectations and decisions as a starting point. An important part is to get information and do the calculations over time. DSGE models have all these properties. DSGE models need to continue their development even further because they are not perfected. A model will never be perfect, and obvious reason is that it is impossible to predict the future. But models like DSGE is today's future models. The application and interpretation of different models requires a sound understanding of the structural differences between models. In the rest of this thesis I will present, a prototype DSGE model and I will study and reveal its structural complexity.

3. The Model

The model I am going to use was first presented by G. D Hansen in 1985. This is a one-sector RBC model with indivisible labor. The more specific version comes from Ruge-Murcia (2007).

⁷ Mork, K.A. (2008): Modeller og pengepolitikk. *Samfunnsøkonomene*, nr. 1, page 12

3.1 Households

The representative agent will maximize the expected lifetime utility function.

$$E_t \sum_{i=t}^{\infty} \beta^{i-t} (\ln(c_i) + \psi(1 - n_i))$$

Here $\beta \in (0, 1)$ is the discount factor, c_t is the consumption, n_t is hours worked and ψ is a utility weight. ψ can be interpreted as a parameter that regulates the agent valuing of spare time ($1 - n_t$). If this parameter is high then spare time is highly valued for the agent, and the opposite for a low value. In this economy the population growth is zero, and the size and time endowment are normalized to one. The budget constraint consist the agent's income on the right-hand side and the expenditure on the left-hand side.

$$c_t + x_t = w_t n_t + r_t k_t$$

Here x_t is investment, w_t it the real wage, r_t is the real rental rate of capital, and k_t is the capital stock. The right-hand side includes wages ($w_t n_t$) and rents received from selling labor and renting capital to firms ($r_t k_t$). This is allocated into the two left-hand side variables consumption and investment. The amount of investment at time t increases the capital stock at time $t+1$.

$$k_{t+1} = (1 - \delta)k_t + x_t$$

The parameter $\delta \in (0, 1)$ is the depreciation rate. In addition to the transversality condition it is necessary to have some first-order condition associated with the optimal choice of c_t , n_t and k_t . The transversality condition fixes the behavior of some variables in the far future. We work backwards by fixing the future values and then compute backwards to our state of time. The first-order conditions have to be fulfilled if we are going to maximize the utility function. In Ruge-Murcia (2007) he presents these two:

$$1/c_t = \beta E_t((1/c_{t+1})(1 + r_{t+1} - \delta)) \quad (1)$$

$$\psi c_t = w_t. \quad (2)$$

To find these equations Hansen (1985) use a function called Bellman's equation. The Bellman equation is a technique that can be used to find necessary conditions for an optimization problem. The man behind this technique is Richard Bellman. The Bellman equation is also called dynamic programming equation. This is a very useful technique when we are dealing with a dynamic optimization problem. This is so in our case, because we have movement in every time t . It is a recursive technique working backward. The backward recursion technique always solves the original problem if a solution exists⁸. For a good introduction to the Bellman equation I recommend reading Sargent (1987 chapter 1). Equation (2) is not that complicated to find. Since it is not lagged, you could just form a one-period Lagrange function and solve for c_t and n_t . This is because the condition has to hold in every period.

$$L(c_t, n_t, \lambda) = \beta E_t (\ln(c_t) + \psi(1 - n_t)) - \lambda(w_t n_t + r_t k_t - c_t - x_t)$$

Then find the partial derivatives for c_t and n_t .

$$\frac{\partial L}{\partial c_t} = \beta E_t \left(\frac{1}{c_t} \right) - \lambda = 0$$

$$\frac{\partial L}{\partial n_t} = -\beta E_t \psi + \lambda w_t = 0$$

$$\frac{\partial L}{\partial \lambda} = w_t n_t + r_t k_t - c_t - x_t$$

Here we can see that there are some expectation operators left. But E_t applied to a constant like ψ is just the constant himself. Since c_t is known in time t , $1/c_t$ is predetermined. The expectation in time t of a predetermined variable in time t , will turn out to be a constant. So there are no difficulties with expectations sign. After the Lagrange multiplier (λ) is eliminated we will get the last first-order condition (2).

3.2 Firms

In this economy there is only one good, and this good is produced by perfectly competitive firms. The representative firm rents labor and capital from the agent and combines them using

⁸ Sargent, T.J. (1987): *Dynamic Macroeconomic Theory*. *Harvard University Press*, page 18

the constant returns to scale technology. This is a Cobb-Douglas function that includes a technology shock, where labor and capital is perfect substitutes.

$$y_t = z_t k_t^\alpha n_t^{1-\alpha}$$

Here $\alpha \in (0, 1)$, y_t is the output and z_t is a technology shock. The technology shock follows the exogenous stochastic process

$$\ln z_{t+1} = \rho \ln z_t + \varepsilon_{t+1}$$

where $\rho \in (-1, 1)$ and ε_t is the random element. This random element can be interpreted as innovation in technology next period. It is assumed that this element is independently, identically and normally distributed (i.i.d.N) with zero mean and variance σ^2 . The level of input is chosen such that the firms maximize their profit and equates the marginal product of labor (capital) to the real wage (rental rate). This procedure happens in every period. The equilibrium for this economy is the sequence of prices $\{w_t, r_t\}_{t=0}^\infty$ and allocations $\{c_t, n_t, x_t, k_{t+1}, y_t\}_{t=0}^\infty$ such that firms maximize profits, agents maximize utility, and all markets clear⁹.

3.3 Linearization of the Model

It can be showed that this economy converges to a steady state. A common strategy to solve DSGE models is to determinate a steady state and move the whole system to this deterministic steady state, in a next step we linearize the first-order condition and constraints by means of a first-order Taylor series expansion around the deterministic steady state. Often this deterministic steady state is origin. This process can contain complicated math, Ruge-Murcia (2007) gives this linearization in percentage deviation from its steady state¹⁰. Since these are given in percentage deviation it is not that difficult to check whether the central equations are right or not. Because we could use this combination,

$$\ln(c_t) \approx \frac{c_t - c}{c} = \hat{c}_t$$

⁹ Ruge-Murcia F.J (2007): Methods to estimate dynamic stochastic general equilibrium models. page 2602

¹⁰ Ruge-Murcia F.J (2007): Methods to estimate dynamic stochastic general equilibrium models. page 2634

where the parameter without subscript are the steady state values and the variable with the hat denotes the percentage deviation from steady state. For the two first-order conditions it is important to remember that w_t is equal to the marginal product of labor, and r_t is equal the marginal product of capital. First we start with the first-order condition. Remember that:

$$r_{t+1} = \alpha z_{t+1} (k_{t+1})^{\alpha-1} (n_{t+1})^{1-\alpha}.$$

To rewrite equation (1) is not that straight forward as it seems. The rule I am going to use can be found in Mood, Graybill and Boes (1974)¹¹. This rule is stated in the following way:

$$E \left(\frac{X}{Y} \right) \approx \frac{\mu_x}{\mu_y} - \frac{1}{\mu_y^2} Cov(X, Y) + \frac{\mu_x}{\mu_y^3} Var(Y).$$

Here μ_x and μ_y describes the expectations for X and Y. Our covariance is equal to zero to that term can be zeroed out. The last term is a little more difficult to interpret. But because μ_y raised to the third will be a very larger number, we will get a small number divided by a large number. This will go towards zero. If our data sample is large then $Var(Y)$ also will go towards zero. So we can rewrite (1) into this:

$$\frac{1}{c_t} = \frac{\beta(1 - \delta) + \beta E_t r_{t+1}}{E_t c_{t+1}}.$$

After some manipulations we end up with an equation that is easy to work with if we are going to find the linearized function of this first-order condition.

$$E_t c_{t+1} = (\beta(1 - \delta) + \beta E_t r_{t+1}) c_t$$

Then we take the log to this equation and insert the log of r_{t+1} .

$$E_t \ln c_{t+1} = \ln \beta - \ln \delta + \beta \ln \alpha + \ln c_t + \beta(\alpha - 1) E_t \ln k_{t+1} + \beta(1 - \alpha) E_t \ln n_{t+1} + \beta E_t \ln z_{t+1}$$

This equation has the same structure as presented in Ruge-Murcia (2007). The part which has been typed in bold face will disappear when we determine the steady state equilibrium to

¹¹ Mood, A.M., F.A. Graybill and D.C. Boes (1974): Introduction to the Theory of Statistics, 3th , page 181

origin. I will discuss how this is done later one in this paper. Then the second one starts with equation (2), and substitute w_t with the marginal product of labor.

$$\psi c_t = (1 - \alpha)z_t(k_t)^\alpha(n_t)^{-\alpha}$$

Here the right-hand side has been substituted with the marginal product of labor. Then take the log to this function and solve for $\ln n_t$.

$$\begin{aligned} \ln c_t &= \ln \{(1/\psi)(1 - \alpha)z_t(k_t)^\alpha(n_t)^{-\alpha}\} \\ \ln n_t &= (\mathbf{1}/\alpha)(\ln(\mathbf{1}/\psi) + \ln(\mathbf{1} - \alpha)) + (1/\alpha) \ln z_t + \ln k_t - (1/\alpha) \ln c_t \end{aligned}$$

We can see that this has the same structure as given in the Ruge-Murcia (2007) article. The part that is written in fat will disappear when we move the steady state to origin. This is because we do not have a fixed point in origin. Comparing for the product function.

$$\begin{aligned} \ln y_t &= \ln(z_t(k_t)^\alpha(n_t)^{1-\alpha}) \\ \ln y_t &= \ln z_t + \alpha \ln k_t + (1 - \alpha) \ln n_t \end{aligned}$$

The meaning of these equations is that they form a dynamic system that determines the path for output, hours worked, technology shock, capital, consumption and investment. If we not linearize the model it will be very difficult, if not impossible to estimate the model. After some manipulations Ruge-Murcia (2007)¹² presents these linearized equations in equation (3).

$$\begin{bmatrix} \hat{k}_{t+1} \\ E_t \hat{c}_{t+1} \end{bmatrix} = A \begin{bmatrix} \hat{k}_t \\ \hat{c}_t \end{bmatrix} + B \hat{z}_t \quad (3)$$

$$\begin{aligned} A &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 + \delta\gamma/(1 - \gamma) & -\delta(1 + \alpha\gamma - \alpha)/(\alpha - \alpha\gamma) \\ 0 & \alpha/(\zeta + \alpha - \alpha\zeta) \end{bmatrix} \\ B &= \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \delta/(\alpha - \alpha\gamma) \\ \zeta\rho/(\zeta + \alpha - \alpha\zeta) \end{bmatrix} \end{aligned}$$

According to Ruge-Murcia (2007) the $\zeta = \alpha\beta(k/n)^{\alpha-1}$, $k/n = ((1/\beta + \delta - 1)/\alpha)^{1/(\alpha-1)}$ is the steady-state capital-labor ratio, $\gamma = 1 - \delta(k/n)^{1-\alpha}$ is the steady-state consumption-output ratio. All the variables that do not have a time subscript are the steady state values. If we will use this

¹² Ruge-Murcia F.J (2007): Methods to estimate dynamic stochastic general equilibrium models. Page 2602

model for calculations, the A matrix has to be converted into a matrix that only includes the structural parameters. This is done by inserting the equations for ζ , γ and k/n , after some manipulations the result of this is quite easy.

Result 1:

If $\zeta = \alpha\beta(k/n)^{\alpha-1}$, $k/n = ((1/\beta + \delta - 1)/\alpha)^{1/(\alpha-1)}$, $\gamma = 1 - \delta(k/n)^{1-\alpha}$ and matrix A and B holds, then matrix A and B can be written in this way:

$$A = \begin{bmatrix} 1 + \frac{\frac{1}{\beta} + \delta - 1}{\alpha} + \delta & -\frac{1}{\beta} - 1 \\ 0 & \frac{1}{\frac{1 + \beta(\delta - 1)}{\alpha} + \beta(1 - \delta)} \end{bmatrix} \quad B = \begin{bmatrix} \frac{1}{\beta} + \delta - 1 \\ \frac{\rho(\frac{1}{\delta - 1} + \beta)}{\frac{1}{\delta - 1} + \beta(1 - \alpha)} \end{bmatrix}.$$

Proof for result 1: See appendix A

These types of model are moving around the steady state in all kind of different directions and are difficult to read. What Ruge-Murcia (2007) has done, is that he has made equations that tells us on how fare from the steady state equilibrium these movements are. This has been done in percentage deviations. To understand equation (3) a bit more we could look at a more general example. This example can be found in Shone (1997 page 180), but I going to modify it so it will be closer to the case in the Ruge-Murcia article. Suppose we have a nonlinear system of equations,

$$\begin{aligned} x_{t+1} &= f(x_t, y_t) \\ y_{t+1} &= g(x_t, y_t) \end{aligned}$$

where we have just one period time lag. In order to be able to investigate the stability properties of this nonlinear system in the neighborhood of the steady state, the steady state have to exist for this system. Another important condition is that f and g have to be continuous and differentiable. The steady state (x, y) exist if it satisfies these conditions.

$$\begin{aligned} x &= f(x, y) \\ y &= g(x, y) \end{aligned}$$

If these conditions are fulfilled then we can use a Taylor expansion in the steady sate (x, y) .

$$\frac{x_{t+1} - x}{x} = \frac{\partial f(x, y)}{\partial x_t} \frac{x_t - x}{x} + \frac{\partial f(x, y)}{\partial y_t} \frac{y_t - y}{y}$$

$$\frac{y_{t+1} - y}{y} = \frac{\partial f(x, y)}{\partial x} \frac{x_t - x}{x} + \frac{\partial f(x, y)}{\partial y_t} \frac{y_t - y}{y}$$

And let:

$$a_{11} = \frac{\partial f(x, y)}{\partial x_t}, a_{12} = \frac{\partial f(x, y)}{\partial y_t}$$

$$a_{21} = \frac{\partial g(x, y)}{\partial x_t}, a_{22} = \frac{\partial f g(x, y)}{\partial y_t}.$$

If we let $\hat{y}_t = (y_t - y)/y$ which is the percentage deviation from the steady state, then the system can be written like this:

$$\begin{bmatrix} \hat{x}_{t+1} \\ \hat{y}_{t+1} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \hat{x}_t \\ \hat{y}_t \end{bmatrix}.$$

This gives us the same structure as equation (3). The structure is just a first-order linear system. If we replace x_t with k_t and y_t with c_t you can see that this is almost the same as equation (3). We have now transformed the nonlinear system into a tractable linear system.

We are transforming the system such that the linear system has the same qualitative properties in a neighborhood around the steady-state as the nonlinear system.

To find a solution for the system (3) is not straight forward because it includes expectations. There are many methods for solving linear differences models with rational expectation. The method I am going to use was first presented by Blanchard and Kahn (1980). This is also the method that Ruge-Murcia (2007) uses. This technique requires that we first compute the eigenvalues and eigenvectors for matrix A. But since we are in a situation where we only have one predetermined and one non-predetermined variable we do not need the eigenvectors to compute the solution¹³. The eigenvalues and eigenvector are presented in *result 2*.

Result 2:

If the form of matrix A that Ruge-Murcia (2007) presents holds, then we get two eigenvalues noted as λ_1 and λ_2 :

¹³ Blanchard, O.J. and C.M. Kahn (1980): The solution of linear difference models under rational expectations. Page 1309

$$\lambda_1 = a_{22} = \alpha / (\zeta + \alpha - \alpha\zeta)$$

$$\lambda_2 = a_{11} = 1 + \delta\gamma / (1 - \gamma).$$

We can observe that $|\lambda_1| < 1$ and $|\lambda_2| > 1$. The eigenvector V can be noted like this:

$$V = \begin{bmatrix} -\frac{a_{12}}{a_{11}-a_{22}} & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -\frac{\frac{\delta(1+\alpha\gamma-\alpha)}{\alpha-\alpha\gamma}}{1+\frac{\delta\gamma}{1-\gamma}\left(\frac{\alpha}{\zeta+\alpha-\alpha\zeta}\right)} & 1 \\ 1 & 0 \end{bmatrix}.$$

Proof for result 2: See appendix B. B.1 describes the eigenvalue and B.2 the eigenvector.

The solutions we get from the Blanchard and Kahn approach is given in equation (4) and (5).

$$\hat{k}_{t+1} = a_{11}\hat{k}_t + a_{12}\hat{c}_t + b_1\hat{z}_t \quad (4)$$

$$\hat{c}_t = \phi_{ck}\hat{k}_t + \phi_{cz}\hat{z}_t \quad (5)$$

Here ϕ_{ck} and ϕ_{cz} are combinations of eigenvalues and eigenvectors of the matrix A. These parameters depend on the structural parameters in a nonlinearly way. These equations have been printed in the Ruge-Murcia (2007), but the author does not reveal the nature of their functional dependency. I provide a result (*Result 3*) which exhibits the full complexity of the dependency of ϕ_{ck} and ϕ_{cz} on the structural parameters δ , α and β .

Result 3:

If the matrix A that Ruge-Murcia gives us holds, and that the Blanchard and Kahn (1980) method hold. Then ϕ_{ck} and ϕ_{cz} are equal to these equations.

$$\phi_{ck} = \frac{a_{22} - a_{11}}{a_{12}}$$

$$\phi_{cz} = \frac{a_{22}(a_{22} - a_{11})}{a_{11}a_{12}} - 1$$

Proof for result 3: See appendix C.

The state variables of this system are the capital stock k_t and technology shock z_t . The observable variables (y_t, n_t, c_t) are the variables that will be used to estimate the model. We can then form a model using the linearized equations.

$$s_t = \begin{bmatrix} \hat{y}_t \\ \hat{n}_t \\ \hat{c}_t \end{bmatrix} = \Phi \xi_t = \begin{bmatrix} \phi_{yk} & \phi_{yz} \\ \phi_{nk} & \phi_{nz} \\ \phi_{ck} & \phi_{cz} \end{bmatrix} \begin{bmatrix} \hat{k}_t \\ \hat{z}_t \end{bmatrix} \quad (6)$$

Here the state- and observable variables are written in percentage deviation from the steady state. The vector Φ is nonlinear functions of the structural parameters, and ξ_t is a 2×1 vector that contains our state variables. This equation (6) is the state-space representation of the model. This model uses the predetermined level of capital and one exogenous shock as the input. If this input is multiplied by Φ it provides us with predictions about our endogenous observables: output, consumption and hours worked. The matrix Φ can be presented in the way *Result 4* shows us.

Result 4:

If the condition \hat{y}_t, \hat{n}_t and equation (5) that Ruge-Murcia (2007)¹⁴ presents holds then we have this result.

$$\Phi = \begin{bmatrix} \phi_{yk} & \phi_{yz} \\ \phi_{nk} & \phi_{nz} \\ \phi_{ck} & \phi_{cz} \end{bmatrix} = \begin{bmatrix} 1 + \left(1 - \frac{1}{\alpha}\right) \phi_{ck} & \frac{1}{\alpha} + \left(1 - \frac{1}{\alpha}\right) \phi_{cz} \\ 1 - \frac{1}{\alpha} \phi_{ck} & \frac{1}{\alpha} - \frac{1}{\alpha} \phi_{cz} \\ \phi_{ck} & \phi_{cz} \end{bmatrix}$$

Proof for result 4: See appendix D.

4. Estimation method

4.1 Maximum likelihood: General theory

The estimation method that is often used to solve DSGE models is called Maximum Likelihood (ML). The principle of maximum likelihood provides a means of choosing an asymptotically efficient estimator for a parameter or a set of parameters¹⁵. The general result

¹⁴ Ruge-Murcia F.J (2007): Methods to estimate dynamic stochastic general equilibrium models. Page 2634

¹⁵ Green, W.H (2003): *Econometric Analysis. Pearson Education International*. 5th edition page 470

is that the ML estimator is a stationary point of the likelihood function. In our model there are many parameters that need to be estimated. These parameters are collected in the $q \times 1$ vector θ . The observables (y, c, n) are then used to estimate θ . Then we can write the probability density function as $f(x / \theta)$ where x is the data that is observed. After observing n random draws from the density $f(x / \theta)$ that are assumed to be independent and identically distributed (iid), we can determine the likelihood of the occurrence of the sample as the product of individual densities:

$$f(x_T, x_{T-1}, \dots, x_1 | \theta) = \prod_{t=1}^T f(x_t | \theta) = L(\theta | x).$$

The probability of the occurrence of a sample is given by this likelihood function. The maximum likelihood estimate of θ is the value for which this sample is most likely to have been observed; that it is the value of θ that maximizes $L(\theta | x) = f_{x_T, x_{T-1}, \dots, x_1}(x_T, x_{T-1}, \dots, x_1 | \theta)$ ¹⁶. Finding maximum likelihood estimates can be split into two parts. One is to find the likelihood function, second is to find the values of θ that maximizes the likelihood function. It is often easier to work with the log of the likelihood function, because it avoids technical problems when finding the first order conditions. The maximum of a function $f(x)$ is the same as the maximum of $\ln f(x)$. This gives us a sum of T expressions involving the log of the densities over time.

$$\begin{aligned} \ln L(\theta | x) &= \ln \prod_{t=1}^T f(x_t | \theta) \\ \ln L(\theta | x) &= \sum_{t=1}^T \ln f(x_t | \theta) \end{aligned}$$

An important aspect is that maximum likelihood estimators are consistent, asymptotically normally distributed, and therefore efficient among estimators that have these properties. This is the case if MLE satisfies certain standard regularity conditions (see Green 2003 5th edition page 473 for these conditions). A drawback to this technique is that the density of the observed variable has to be known. That is one has to be able to give a detailed specification

¹⁶ Hamilton, J.D (1994): Time Series Analysis. Princeton University Press, page 117

of the data generating process.

4.1.1 Maximum likelihood: For our model

The representation of the system (6) has to have a state and an observation equation. To find the state equation we use (4) and (5). If we substitute (5) into (4) we will find \hat{k}_{t+1} described by \hat{k}_t and \hat{z}_t . The other state variable move from time t to $t+1$ are already described by the linearization. Then we have:

$$\begin{bmatrix} \hat{k}_{t+1} \\ \hat{z}_{t+1} \end{bmatrix} = \begin{bmatrix} a_{11} + a_{12}\phi_{ck} & a_{12}\phi_{cz} + b_1 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} \hat{k}_t \\ \hat{z}_t \end{bmatrix} + \begin{bmatrix} 0 \\ \varepsilon_{t+1} \end{bmatrix}.$$

Which can be formulated in another way, where F is a 2×2 matrix and v_t is a 2×1 vector. This will be the state equation.

$$\xi_{t+1} = F\xi_t + v_{t+1} \quad (7)$$

To be able to us ML estimation we have to rewrite equation (6) by adding an extra variable.

$$x_t = hs_t = h\Phi\xi_t = H\xi_t \quad (8)$$

This is our observation equation. Here we have just multiplied equation (6) with vector h . This vector can be interpreted as a selection vector. We do this because when we estimate DSGE models with MLE there cannot be more observables then structural shocks. In our model we just have one structural shock, so we can only estimate with one observable at each time. The selection vector will be an 1×3 vector, so if this vector is like this (0,1,0) we are estimating using hours worked n_t alone.

Let all the parameters of the model be denoted in the $q \times 1$ vector θ . The past observations of x_t can be collected in \mathfrak{N}_{t-1} . At time $t-1$ we use \mathfrak{N}_{t-1} as a basis for the forecast of ξ_t noted as $\tilde{\xi}_{t|t-1}$, and the mean square error for this forecast is $P_{t|t-1}$. If we take the assumption that technology innovation is normal distributed we get this:

$$f(x_t | \mathfrak{R}_{t-1}; \theta) = N(H\tilde{\xi}_{t|t-1}, HP_{t|t-1}H').$$

Here we have used a result that is often used for any constant. If we let L be our constant in this example:

$$X \sim N(\mu, P_{t|t-1}) \rightarrow LX \sim N(L\mu, LP_{t|t-1}L').$$

Here L' denotes the transpose of L . Then we can start the ML estimation.

$$\tilde{\theta}_{ml} = \underset{\{\theta\}}{\operatorname{argmax}} L(\theta).$$

Have to find the log likelihood function. Since we assume normal distribution we can use the general form of the normal distribution as a starting point.

$$f(x|\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2[(x-\mu)^2/\sigma^2]}$$

Here the mean is μ and σ is the standard deviation. A large sample date is preferred because we are assuming normal distribution. So when we take the log to the standard normal distribution function and take into consideration that our sample is large we get this function.

$$\ln L(\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln \sigma^2 - \frac{1}{2}\sum_{t=1}^n \left[\frac{(x_i - \mu)^2}{\sigma^2} \right]$$

Since the mean and variance is given earlier we just have to insert it in. We change n to T which is equal to the sample size and insert our mean and variance. We will then get the log likelihood function. Here $L(\theta)$ denotes the log likelihood function.

$$L(\theta) = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln|HP_{t|t-1}H'| - \frac{1}{2}\sum_{t=1}^T (x_t - H\tilde{\xi}_{t|t-1})'(HP_{t|t-1}H')^{-1}(x_t - H\tilde{\xi}_{t|t-1})$$

Since we do not have a value for ζ we need an estimated a value for ζ . The technique that can be used is called Kalman filtering. Since we have only one shock in this model we cannot use

all the observables when we are estimating. We can use just one. To solve this problem we have three options. The first is of course to estimate one observable at each time. The second alternative will be to add an error term to the observation equation of the state-space representation of the model, and the last option is to add more structural shocks. As mentioned earlier the ML estimator under standard regularity conditions is consistent and asymptotically normal. In our case we have

$$\sqrt{T}(\tilde{\theta}_{ml} - \theta) \rightarrow N(0, (\mathcal{J}/T)^{-1}).$$

Green (2003 5th edition page 478) has sketched a proof, a good proof for this outline. Here we have an unbiased estimator, our $\mu = 0$. If we look at the variance term we can observe that a larger T that means a larger data sample, we get a smaller variance. In mathematical terms: When $T \rightarrow \infty$ then $(\mathcal{J}/T)^{-1} \rightarrow 0$. We can interpret θ as the true value of the parameter vector, and $\mathcal{J} = -E(\partial^2 L(\theta) / \partial \theta \partial \theta')$ as the information matrix. When computing this we can use the Hessian, this is just a matrix of second derivatives.

The likelihood function that is mentioned earlier in the text looks quite easy. But it is in fact a very complex function of the structural parameters. In *Result 5* I present this complex likelihood function. Since we have three observables there will be three likelihood functions. They will approximately give us the same estimation results, but there will be some differences. This has been demonstrated in the case of Ruge-Murcia (2007) in the numerical results. In *result 5* I just present one of these three equations. That is the one when we set the selection vector h to be equal to $(1,0,0)$. Then we are estimating using \hat{y}_t . To find the other two equations we have to set the selection vector h to $(0,1,0)$ and $(0,0,1)$. According to Ruge-Murcia (2007) the α and δ can be fixed then θ will be equal to (β, ρ, σ) . In *result 5* we can see p_{11} , p_{12} , p_{21} and p_{22} , they are only the variance that comes from the Kalman filter. The estimated values from the Kalman filter are $\tilde{k}_{(t|t-1)}$ and $\tilde{z}_{(t|t-1)}$. If we take a look at the function *result 5* gives us it can be difficult to see connection to the parameters ρ and σ . We can find σ in the p_{xx} and ρ are hidden in the estimated value for \hat{z}_t . We can also find ρ in the F matrix.

Result 5:

If the result 2, 3 and the likelihood function holds, we can write the likelihood functions dependency to the structural parameters in the following way.

$$\begin{aligned}
 L(\theta) = & -\frac{T}{2} \ln(2\pi) - \frac{T}{2} \ln \left(1 + \frac{p_{11}}{\alpha} \left(\frac{(\alpha-1) \left(\frac{1}{1+\beta(\delta-1)+\beta(1-\delta)} - \frac{1+\frac{1}{\beta}+\delta-1}{\alpha} \right)}{-\alpha \left(\frac{1}{\beta} + 1 \right)} \right) \right)^2 + (p_{12} + p_{21}) \left(\left(1 + \left(\frac{(\alpha-1) \left(\frac{1}{1+\beta(\delta-1)+\beta(1-\delta)} - \frac{1+\frac{1}{\beta}+\delta-1}{\alpha} \right)}{-\alpha \left(\frac{1}{\beta} + 1 \right)} \right) \right) \left(\frac{2}{\alpha} + \left(1 - \frac{1}{\alpha} \right) \left(\frac{\frac{1}{1+\beta(\delta-1)+\beta(1-\delta)} - \frac{1+\frac{1}{\beta}+\delta-1}{\alpha}}{\left(1 + \frac{1}{\beta} + \delta \right) \left(-\frac{1}{\beta} - 1 \right)} \right) - 1 \right) \right) + \\
 & p_{22} \left(\frac{2}{\alpha} + \left(1 - \frac{1}{\alpha} \right) \left(\frac{\frac{1}{1+\beta(\delta-1)+\beta(1-\delta)} - \frac{1+\frac{1}{\beta}+\delta-1}{\alpha}}{\left(1 + \frac{1}{\beta} + \delta \right) \left(-\frac{1}{\beta} - 1 \right)} \right) \right)^2 - \\
 & \left(\left(\left(\left(\frac{1}{1+\beta(\delta-1)+\beta(1-\delta)} - \frac{1+\frac{1}{\beta}+\delta-1}{\alpha} \right) \right) - \frac{1+\frac{1}{\beta}+\delta-1}{\alpha} \right) \right)^2 - 1 \right) - \\
 & \left(\left(\left(\left(\frac{(\alpha-1) \left(\frac{1}{1+\beta(\delta-1)+\beta(1-\delta)} - \frac{1+\frac{1}{\beta}+\delta-1}{\alpha} \right)}{-\alpha \left(\frac{1}{\beta} + 1 \right)} \right) \right) \tilde{K}_{(t|t-1)} \left(\frac{1+\frac{1}{\beta}+\delta-1}{\alpha} \right) \right) \right)^2 - 1 \tilde{Z}_{(t|t-1)} \right) \right)^2 \\
 & \frac{1}{2} \sum_{t=1}^{\infty} \frac{\left(\left(\frac{(\alpha-1) \left(\frac{1}{1+\beta(\delta-1)+\beta(1-\delta)} - \frac{1+\frac{1}{\beta}+\delta-1}{\alpha} \right)}{-\alpha \left(\frac{1}{\beta} + 1 \right)} \right) \right)^2 + (p_{12} + p_{21}) \left(\left(\left(\frac{(\alpha-1) \left(\frac{1}{1+\beta(\delta-1)+\beta(1-\delta)} - \frac{1+\frac{1}{\beta}+\delta-1}{\alpha} \right)}{-\alpha \left(\frac{1}{\beta} + 1 \right)} \right) \right) \left(\frac{2}{\alpha} + \left(1 - \frac{1}{\alpha} \right) \left(\frac{\frac{1}{1+\beta(\delta-1)+\beta(1-\delta)} - \frac{1+\frac{1}{\beta}+\delta-1}{\alpha}}{\left(1 + \frac{1}{\beta} + \delta \right) \left(-\frac{1}{\beta} - 1 \right)} \right) - 1 \right) \right) \right)^2 + p_{22} \left(\frac{2}{\alpha} + \left(1 - \frac{1}{\alpha} \right) \left(\frac{\frac{1}{1+\beta(\delta-1)+\beta(1-\delta)} - \frac{1+\frac{1}{\beta}+\delta-1}{\alpha}}{\left(1 + \frac{1}{\beta} + \delta \right) \left(-\frac{1}{\beta} - 1 \right)} \right) \right)^2 - 1 \right)^2}{2}
 \end{aligned}$$

Proof: See appendix E.

4.2 Kalman filter

Kalman filtering technique was developed by R.E. Kalman (Kalman R.E. 1960, Kalman R.E. and R.S. Bucy 1961). This technique was in the beginning just used by control engineers and other physical scientists, and one of the main reasons for this is that the Kalman article was first printed in an engineering journal. Kalman filter technique is a mathematical estimation technique that can be used to tracking, prediction or forecasting. Of course with these applications statisticians and economist could use this technique. Especially statisticians who are interested in linear regression models and time series analysis. The Kalman filter is an algorithm for sequentially updating a linear projection for the system¹⁷. It is a recursive technique using information from the past. One of the main advantages with this technique is that it is quite easy to use.

To get a better understanding of the Kalman filter I would describe the concept in a little more technical aspect. If we observe $y_{t-1}, y_{t-2}, \dots, y_{t-T}$ and let this be our data. This data can be in vector or scalar. We assume the y_t depends on the variable τ_t . This is an unobservable variable, and is the state vector/variable. It is important to notice that if y_t and τ_t are vectors, the dimension of τ_t is independent form y_t . The relationship between these two is linear:

$$y_t = F_t \tau_t + v_t.$$

This is the observation equation. The connection between τ_t and τ_{t-1} can be written like this:

$$\tau_t = G_t \tau_{t-1} + w_t.$$

This is the state equation also known as the system equation. It's called system equation because it's not fixed, but moves over time like a dynamic system. F_t and G_t in our case is known, but there are methods to estimate this through data, see Hamilton (1994, Chapter 13.4). The observation error v_t and the system equation error w_t are both assumed to be normal distributed with mean zero and a known variance. They are also assumed to be independent of each other. Kalman filter is a recursive technique, and the mean squared error (MSE) for each of these forecast in matrix notation (this because it is closer to our case) can be written like this:

¹⁷ Hamilton, J.D (1994): Time Series Analysis. *Princeton University Press*, page 372

$$P_{t|t-1} \equiv E \left[(\tau_t - \tilde{\tau}_{t|t-1})(\tau_t - \tilde{\tau}_{t|t-1})' \right].$$

Here the tilde denotes the expected value. The recursive procedure starts at $\tilde{\tau}_{1|0}$. We focus on time $t-1$, where $t = 1, 2, \dots, T$ and we use ore observed data up to time $t-1$. First we have to set a starting value for $\tilde{\tau}$ and for P . With them given $\tilde{\tau}_{1|0}$ and $P_{1|0}$, the next step is then to calculate same type of values to the next period. This process is continued on. In general terms we can write it like this: Given $\tilde{\tau}_{t|t-1}$ and $P_{t|t-1}$ the goal is to calculate $\tilde{\tau}_{t+1|t}$ and $P_{t+1|t}$. When we have these results they can easily be inserted in our starting equations.

This explanation of the Kalman filter is obtained by reading Meinhold and Singpurwalla (1983) and Hamilton (1994 chapter 13). For future understanding of the Kalman filter I recommend you to read both. Hamilton (1994) is very close related to the Ruge-Murcia (2007) article. In this paper equation (7) is the state equation and (8) is the observation equation. In our model the Kalman filter recursion will be set to $\tilde{\xi}_{1|0} = E(\xi_t) = (0,0)'$ and $P_{1|0} = E(\xi_t \xi_t')$.

5. Conclusion

In this paper I have presented the Ruge-Murcia (2007) model in more analytical way. The five results that have been presented in this paper are all important to see the underlying structure of this model. It is also easier to understand how the model works, and how the different agents are connected. If we use the first four results we can find the likelihood function, and we get a very complex equation. To estimate the DSGE model, we need to maximize the likelihood. This requires partial differentiation of the function. This will be a very difficult to do analytical. DSGE models are already affecting us, because central banks are using them in monetary policy decisions, but the models that the central banks are even more complicated. They are more precise in their formulation of the different agents in the economy. The Ruge-Murcia (2007) model together with the additional results, are a good model when it comes to understanding how DSGE models work. In the Ruge-Murcia (2007) model there are different important agents that are not modeled, example is the government. Since these models are so complex it will maybe difficult to do the calculation with many agents. So these DSGE models are best when we focus on a small part of the economy. DSGE models and the model are complex and very detailed models. Since these models will be used in the future, it is

important to understand their underlying structure. It will then be easier to interpret estimation results from different DSGE models. Practical economical work involving the DSGE modeling approach is barely feasible without the intensive use of numerical techniques.

Appendix A: Proof for Result 1

A.1: For a_{11} we have this combination:

$$a_{11} = 1 + \frac{\delta\gamma}{1-\gamma}.$$

Then we insert $\gamma = 1 - \delta(k/n)^{1-\alpha}$.

$$\begin{aligned} a_{11} &= 1 + \frac{\delta(1 - \delta(k/n)^{1-\alpha})}{1 - (1 - \delta(k/n)^{1-\alpha})} \\ &= 1 + \frac{\delta}{\delta(k/n)^{1-\alpha}} - \frac{\delta^2(k/n)^{1-\alpha}}{\delta(k/n)^{1-\alpha}} \\ &= 1 + \frac{1}{(k/n)^{1-\alpha}} - \delta \end{aligned}$$

Then we insert $k/n = ((1/\beta + \delta - 1)/\alpha)^{1/(\alpha-1)}$.

$$\begin{aligned} a_{11} &= 1 + \frac{1}{(((1/\beta + \delta - 1)/\alpha)^{1/(\alpha-1)})^{1-\alpha}} - \delta \\ &= 1 + \frac{1}{((1/\beta + \delta - 1)/\alpha)^{-1}} - \delta \\ a_{11} &= 1 + \frac{\frac{1}{\beta} + \delta - 1}{\alpha} - \delta \end{aligned}$$

A.2: For a_{12} we have:

$$a_{12} = \frac{-\delta(1 + \alpha\gamma - \alpha)}{\alpha - \alpha\gamma}.$$

Insert $\gamma = 1 - \delta(k/n)^{1-\alpha}$.

$$\begin{aligned}
 a_{12} &= \frac{-\delta(1 + \alpha(1 - \delta(k/n)^{1-\alpha}) - \alpha)}{\alpha - \alpha(1 - \delta(k/n)^{1-\alpha})} \\
 &= \frac{-\delta - \alpha\delta + \alpha\delta^2(k/n)^{1-\alpha} + \alpha\delta}{\alpha\delta(k/n)^{1-\alpha}} \\
 &= -\frac{\delta}{\alpha\delta(k/n)^{1-\alpha}} + \frac{\alpha\delta^2(k/n)^{1-\alpha}}{\alpha\delta(k/n)^{1-\alpha}} \\
 &= -\frac{1}{\alpha(k/n)^{1-\alpha}} + \delta
 \end{aligned}$$

Insert $k/n = ((1/\beta + \delta - 1)/\alpha)^{1/(\alpha-1)}$.

$$\begin{aligned}
 a_{12} &= -\frac{1}{(1/\beta + \delta - 1)^{-1}} + \delta \\
 &= -\left(\frac{1}{\beta} + \delta - 1\right) + \delta = -\frac{1}{\beta} - 1
 \end{aligned}$$

A.3: The equation for a_{22} is given in the following way:

$$a_{22} = \frac{\alpha}{\zeta + \alpha - \alpha\zeta}.$$

Insert $\zeta = \alpha\beta(k/n)^{\alpha-1}$.

$$a_{22} = \frac{\alpha}{\alpha\beta(k/n)^{\alpha-1} + \alpha - \alpha^2\beta(k/n)^{\alpha-1}}$$

Here in this case it is easier to focus on the denominator first. We can express the denominator in the following way.

$$\beta(k/n)^{\alpha-1} + 1 - \alpha\beta(k/n)^{\alpha-1}$$

Insert $k/n = ((1/\beta + \delta - 1)/\alpha)^{1/(\alpha-1)}$.

$$\begin{aligned}
 &\beta\left(\frac{1/\beta + \delta - 1}{\alpha}\right) + 1 - \beta(1/\beta + \delta - 1) \\
 &\quad \frac{1 + \beta(\delta - 1)}{\alpha} + \beta(1 - \delta)
 \end{aligned}$$

Then we have this result

$$a_{22} = \frac{1}{\frac{1 + \beta(\delta - 1)}{\alpha} + \beta(1 - \delta)}$$

A.4: For b_1 we are given this combination:

$$b_1 = \frac{\delta}{\alpha - \alpha\gamma}$$

Insert $\gamma = 1 - \delta(k/n)^{1-\alpha}$.

$$\begin{aligned} b_1 &= \frac{\delta}{\alpha - \alpha(1 - \delta(k/n)^{1-\alpha})} \\ &= \frac{1}{\alpha(k/n)^{1-\alpha}} \end{aligned}$$

Then insert $k/n = ((1/\beta + \delta - 1)/\alpha)^{1/(\alpha-1)}$.

$$\begin{aligned} b_1 &= \frac{1}{((1/\beta + \delta - 1)/\alpha)^{-1}} \\ &= \frac{1}{\beta} + \delta - 1 \end{aligned}$$

A.5: The last one, b_2 is given:

$$b_2 = \frac{\zeta\rho}{\zeta + \alpha - \alpha\zeta}$$

Insert $\zeta = \alpha\beta(k/n)^{\alpha-1}$.

$$b_2 = \frac{\rho\alpha\beta(k/n)^{\alpha-1}}{\alpha\beta(k/n)^{\alpha-1} + \alpha - \alpha^2\beta(k/n)^{\alpha-1}}$$

Eliminate α and insert $k/n = ((1/\beta + \delta - 1)/\alpha)^{1/(\alpha-1)}$.

$$\begin{aligned}
 b_2 &= \frac{\rho\beta\left(\frac{1/\beta + \delta - 1}{\alpha}\right)}{\beta\left(\frac{1/\beta + \delta - 1}{\alpha}\right) + 1 - \alpha\beta\left(\frac{1/\beta + \delta - 1}{\alpha}\right)} \\
 &= \frac{\frac{\rho + \rho\beta(\delta - 1)}{\alpha}}{\frac{1 + \beta(\delta - 1)}{\alpha} - \beta(\delta - 1)} \\
 &= \frac{\rho + \rho\beta(\delta - 1)}{1 + \beta(\delta - 1) - \alpha\beta(\delta - 1)} \\
 &= \frac{\frac{\rho}{\delta - 1} + \rho\beta}{\frac{1}{\delta - 1} + \beta - \alpha\beta} \\
 &= \frac{\rho\left(\frac{1}{\delta - 1} + \beta\right)}{\frac{1}{\delta - 1} + \beta(1 - \alpha)}
 \end{aligned}$$

Appendix B: Proof for Result 2

B.1: For the eigenvalues

The eigenvalues of the matrix \mathbf{A} are obtained by solving $|\mathbf{A} - \lambda\mathbf{I}| = 0$ ¹⁸. Then we can write our situation like this:

$$\left| \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| = 0.$$

Then we have to find the determinant and set it equal to zero:

$$\begin{aligned} (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} &= 0 \\ \lambda^2 - \lambda(a_{11} + a_{22}) + a_{11}a_{22} - a_{12}a_{21} &= 0 \\ \lambda^2 - \lambda p + q &= 0 \end{aligned}$$

As we can see this is a polynomial equation, and we could use a standard polynomial formal to solve our problem. I us this one:

$$\lambda_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

For our problem we get:

$$\begin{aligned} \lambda_{1,2} &= \frac{1}{2} \left(p \pm \sqrt{p^2 - 4q} \right) \\ \lambda_{1,2} &= \frac{1}{2} \left((a_{11} + a_{22}) \pm \sqrt{(a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})} \right). \end{aligned}$$

Here there is important to remember that $a_{21} = 0$. We get this from the matrix \mathbf{A} . Then the result can be written in the following way:

$$\begin{aligned} \lambda_{1,2} &= \frac{1}{2} \left((a_{11} + a_{22}) \pm \sqrt{(a_{11} - a_{22})^2} \right) \\ \lambda_1 &= \frac{1}{2} \left((a_{11} + a_{22}) - (a_{11} - a_{22}) \right) = a_{22} \end{aligned}$$

¹⁸ Shone R. (1997): *Economic Dynamics*. Cambridge University Press, page 159

$$\lambda_2 = \frac{1}{2}((a_{11} + a_{22}) + (a_{11} - a_{22})) = a_{11}$$

Is important to remember that we assume $p^2 - 4q \geq 0$.

B.2: For the eigenvector

If we denote the eigenvector as $V = [v^{\lambda_1}, v^{\lambda_2}]$ we have these two combinations:

$$(A - \lambda_1 I)v^{\lambda_1} = 0$$

$$(A - \lambda_2 I)v^{\lambda_2} = 0.$$

Then we just solve these equations.

$$\begin{bmatrix} (a_{11} & a_{12}) \\ (a_{21} & a_{22}) \end{bmatrix} - \begin{bmatrix} a_{22} & 0 \\ 0 & a_{22} \end{bmatrix} \begin{bmatrix} v_1^{\lambda_1} \\ v_2^{\lambda_1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} (a_{11} - a_{22})v_1^{\lambda_1} + a_{12}v_2^{\lambda_1} \\ a_{21}v_1^{\lambda_1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

If we fix $v_2^{\lambda_1}$ equal to 1 and solve $v_1^{\lambda_1}$ we get the last necessary result we need to form the eigenvector. We know that a_{21} is equal to zero. So the last element of this vector will always be zero.

$$v^{\lambda_1} = \begin{bmatrix} -\frac{a_{12}}{a_{11} - a_{22}} \\ 1 \end{bmatrix}$$

The second eigenvector have the same procedure.

$$\begin{bmatrix} (a_{11} & a_{12}) \\ (a_{21} & a_{22}) \end{bmatrix} - \begin{bmatrix} a_{11} & 0 \\ 0 & a_{11} \end{bmatrix} \begin{bmatrix} v_1^{\lambda_2} \\ v_2^{\lambda_2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} a_{12}v_2^{\lambda_2} \\ a_{21}v_1^{\lambda_2} + (a_{22} - a_{11})v_2^{\lambda_2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Since we know that a_{21} is equal to zero, we can fix $v_1^{\lambda_2}$ to 1 and still it will fulfill the condition. There is no other solution to $v_2^{\lambda_2}$ than setting it equal to 0. Then we get this result.

$$v^{\lambda_2} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

This give us the eigenvector that has been presented in *Result x*. If we would like to check if this vector is correct we could solve the combination $D=V^{-1}AV$, where matrix D is the eigenvalue matrix.

$$V^{-1}AV = \begin{bmatrix} 0 & 1 \\ 1 & \frac{a_{12}}{a_{11}-a_{22}} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} -\frac{a_{12}}{a_{11}-a_{22}} & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} a_{22} & 0 \\ 0 & a_{11} \end{bmatrix}$$

It is important to remember that a_{21} is equal to zero.

Appendix C: Outline for equation (4) and (5). Proof for result 3

The solution presented by Blanchard and Kahn (1980) page 1309 is the starting point. We have that matrix A is 2×2 , and $|\lambda_1| < 1$ and $|\lambda_2| > 1$. The solution is given like this:

$$\begin{aligned}\hat{k}_t &= \hat{k}_0 \text{ for } t = 0 \\ \hat{k}_{t+1} &= \lambda_1 \hat{k}_t + b_1 \hat{z}_t + \mu \sum_{i=0}^{\infty} \lambda_2^{-i-1} E(\hat{z}_{t+i} | \Omega_t) \text{ for } t > 0, \\ \hat{c}_t &= a_{12}^{-1} [(\lambda_1 - a_{11}) \hat{k}_t + \mu \sum_{i=0}^{\infty} \lambda_2^{-i-1} E(\hat{z}_{t+i} | \Omega_t)] \text{ for } t \geq 0.\end{aligned}$$

Here $\mu = (\lambda_1 - a_{11})\lambda_1 - a_{12}\lambda_2$. Then we insert the eigenvalues and do the computations.

$$\begin{aligned}\hat{c}_t &= a_{12}^{-1} [(a_{22} - a_{11}) \hat{k}_t + \mu \sum_{i=0}^{\infty} a_{11}^{-i-1} E(\hat{z}_{t+i} | \Omega_t)] \\ \hat{c}_t &= \frac{a_{22} + a_{11}}{a_{12}} \hat{k}_t + a_{12}^{-1} \mu [a_{11}^{-1} E(\hat{z}_t | \Omega_t) + a_{11}^{-1-1} E(\hat{z}_{t+1} | \Omega_t) + a_{11}^{-2-1} E(\hat{z}_{t+2} | \Omega_t) + \dots]\end{aligned}$$

Here Ω_t is all the information we have available at time t. One of our main assumptions is that $E(\hat{z}_{t+1} | \Omega_t)$ is equal to zero. This means that at time t we don't expect any technology shock to accrue in the future. Another important thing to remember it that the information we have available at time t, includes \hat{z}_t . If we take these assumptions into consideration we get a much easier equation to work with.

$$\begin{aligned}\hat{c}_t &= \frac{a_{22} + a_{11}}{a_{12}} \hat{k}_t + a_{12}^{-1} \mu [a_{11}^{-1} \hat{z}_t + a_{11}^{-1-1} * 0 + a_{11}^{-2-1} * 0 + \dots] \\ &= \frac{a_{22} + a_{11}}{a_{12}} \hat{k}_t + \frac{1}{a_{12}} \left(\frac{\mu}{a_{11}} \hat{z}_t \right) \\ &= \frac{a_{22} + a_{11}}{a_{12}} \hat{k}_t + \frac{(a_{22} - a_{11})a_{22} - a_{12}a_{11}}{a_{12}a_{11}} \hat{z}_t \\ &= \frac{a_{22} + a_{11}}{a_{12}} \hat{k}_t + \left(\frac{(a_{22} - a_{11})a_{22}}{a_{12}a_{11}} - 1 \right) \hat{z}_t \\ \hat{c}_t &= \phi_{ck} \hat{k}_t + \phi_{cz} \hat{z}_t\end{aligned} \tag{9}$$

$$\begin{aligned}\hat{k}_{t+1} &= a_{22}\hat{k}_t + b_1\hat{z}_t + \mu \sum_{i=0}^{\infty} a_{11}^{-i-1} E(\hat{z}_{t+i}|\Omega_t) \\ &= a_{22}\hat{k}_t + b_1\hat{z}_t + \mu[a_{11}^{-1}E(\hat{z}_t|\Omega_t) + a_{11}^{-1-1}E(\hat{z}_{t+1}|\Omega_t) + a_{11}^{-2-1}E(\hat{z}_{t+2}|\Omega_t) + \dots]\end{aligned}$$

We have the same reasoning as above in this case.

$$\begin{aligned}\hat{k}_{t+1} &= a_{22}\hat{k}_t + b_1\hat{z}_t + \mu[a_{11}^{-1}\hat{z}_t + a_{11}^{-1-1} * 0 + a_{11}^{-2-1} * 0 + \dots] \\ &= a_{22}\hat{k}_t + b_1\hat{z}_t + \frac{\mu}{a_{11}}\hat{z}_t\end{aligned}$$

We can rewrite (9) and get an equation that could be replace $\frac{\mu}{a_{11}}\hat{z}_t$.

$$\frac{\mu}{a_{11}}\hat{z}_t = a_{12}\hat{c}_t - (a_{22} - a_{11})\hat{k}_t$$

This gives us the final result.

$$\begin{aligned}\hat{k}_{t+1} &= a_{22}\hat{k}_t + b_1\hat{z}_t + a_{12}\hat{c}_t - (a_{22} - a_{11})\hat{k}_t \\ &= a_{11}\hat{k}_t + a_{12}\hat{c}_t + b_1\hat{z}_t\end{aligned}$$

Appendix D: Proof for Result 4

Given equation (5) and these two linearized equations from Ruge-Murcia (2007)¹⁹:

$$\begin{aligned}\hat{y}_t &= \alpha \hat{k}_t + (1 - \alpha) \hat{n}_t + \hat{z}_t \\ \hat{n}_t &= -\frac{1}{\alpha} \hat{c}_t + \hat{k}_t + \frac{1}{\alpha} \hat{z}_t\end{aligned}$$

We insert equation (5) in \hat{n}_t

$$\begin{aligned}\hat{n}_t &= -\frac{1}{\alpha} (\phi_{ck} \hat{k}_t + \phi_{cz} \hat{z}_t) + \hat{k}_t + \frac{1}{\alpha} \hat{z}_t \\ &= -\frac{1}{\alpha} \phi_{ck} \hat{k}_t - \frac{1}{\alpha} \phi_{cz} \hat{z}_t + \hat{k}_t + \frac{1}{\alpha} \hat{z}_t \\ &= \left(1 - \frac{1}{\alpha} \phi_{ck}\right) \hat{k}_t + \left(\frac{1}{\alpha} - \frac{1}{\alpha} \phi_{cz}\right) \hat{z}_t.\end{aligned}$$

Then we can insert this into \hat{y}_t

$$\begin{aligned}\hat{y}_t &= \alpha \hat{k}_t + (1 - \alpha) \left[\left(1 - \frac{1}{\alpha} \phi_{ck}\right) \hat{k}_t + \left(\frac{1}{\alpha} - \frac{1}{\alpha} \phi_{cz}\right) \hat{z}_t \right] + \hat{z}_t \\ &= \alpha \hat{k}_t + \left(1 - \frac{1}{\alpha} \phi_{ck} - \alpha + \phi_{ck}\right) \hat{k}_t + \left(\frac{1}{\alpha} - \frac{1}{\alpha} \phi_{cz} - 1 + \phi_{cz}\right) \hat{z}_t + \hat{z}_t \\ &= \hat{k}_t - \frac{1}{\alpha} \phi_{ck} \hat{k}_t + \phi_{ck} \hat{k}_t + \frac{1}{\alpha} \hat{z}_t - \frac{1}{\alpha} \phi_{cz} \hat{z}_t + \phi_{cz} \hat{z}_t \\ &= \left(1 + \left(1 - \frac{1}{\alpha}\right) \phi_{ck}\right) \hat{k}_t + \left(\frac{1}{\alpha} + \left(1 - \frac{1}{\alpha}\right) \phi_{cz}\right) \hat{z}_t.\end{aligned}$$

¹⁹ Ruge-Murcia F.J (2007): Methods to estimate dynamic stochastic general equilibrium models. Page 2634

Appendix E: Proof for result 5

We are going to estimate the likelihood function with using \hat{y}_t , the h vector is (1,0,0). To this gives us the following x_t :

$$x_t = [1 \quad 0 \quad 0] \begin{bmatrix} \hat{y}_t \\ \hat{n}_t \\ \hat{c}_t \end{bmatrix} = \hat{y}_t.$$

The H matrix will be looking like this.

$$H = [1 \quad 0 \quad 0] \begin{bmatrix} \phi_{yk} & \phi_{yz} \\ \phi_{nk} & \phi_{nz} \\ \phi_{ck} & \phi_{cz} \end{bmatrix} = [\phi_{yk} \quad \phi_{yz}].$$

Then we can compute the variance $HP_{t|t-1}H'$. In the outline of this variance $P_{t|t-1}$ is just a 2×2 matrix. The values for this matrix come from the kalman filter. So we just use standard notation for a 2×2 matrix.

$$\begin{aligned} HP_{t|t-1}H' &= [\phi_{yk} \quad \phi_{yz}] \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} \phi_{yk} \\ \phi_{yz} \end{bmatrix} = [\phi_{yk} \quad \phi_{yz}] \begin{bmatrix} p_{11}\phi_{yk} + p_{12}\phi_{yz} \\ p_{21}\phi_{yk} + p_{22}\phi_{yz} \end{bmatrix} \\ &= \phi_{yk}(p_{11}\phi_{yk} + p_{12}\phi_{yz}) + \phi_{yz}(p_{21}\phi_{yk} + p_{22}\phi_{yz}) \\ &= p_{11}\phi_{yk}^2 + (p_{12} + p_{21})\phi_{yk}\phi_{yz} + p_{22}\phi_{yz}^2 \end{aligned}$$

The likelihood function that are given:

$$L(\theta) = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln|HP_{t|t-1}H'| - \frac{1}{2}\sum_{t=1}^T (x_t - H\xi_{t|t-1})'(HP_{t|t-1}H')^{-1}(x_t - H\xi_{t|t-1}).$$

If we then insert the variance and H matrix we get this result.

$$L(\theta) =$$

$$-\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln\left(p_{11}\phi_{yk}^2 + (p_{12} + p_{21})\phi_{yk}\phi_{yz} + p_{22}\phi_{yz}^2\right) - \frac{1}{2}\sum_{t=1}^T \frac{\left(\hat{y}_t - (\phi_{yk}\tilde{k}_{(t|t-1)} + \phi_{yz}\tilde{z}_{(t|t-1)})\right)^2}{p_{11}\phi_{yk}^2 + (p_{12} + p_{21})\phi_{yk}\phi_{yz} + p_{22}\phi_{yz}^2}$$

Form result 4 we have a combination for ϕ_{yk} and ϕ_{yz} . If we insert them we get the following result.

$$\begin{aligned}
 L(\theta) = & -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln \left(p_{11} \left(1 + \left(1 - \frac{1}{\alpha} \right) \phi_{ck} \right)^2 + (p_{12} + p_{21}) \left(\left(1 + \left(1 - \frac{1}{\alpha} \right) \phi_{ck} \right) \left(\frac{1}{\alpha} + \right. \right. \right. \\
 & \left. \left. \left. \left(1 - \frac{1}{\alpha} \right) \phi_{cz} \right) \right) + p_{22} \left(\frac{1}{\alpha} + \left(1 - \frac{1}{\alpha} \right) \phi_{cz} \right)^2 \right) - \\
 & \frac{1}{2} \sum_{t=1}^{\infty} \frac{\left(\hat{y}_t - \left(\left(1 + \left(1 - \frac{1}{\alpha} \right) \phi_{ck} \right) \bar{k}_{(t|t-1)} + \left(\frac{1}{\alpha} + \left(1 - \frac{1}{\alpha} \right) \phi_{cz} \right) \bar{z}_{(t|t-1)} \right) \right)^2}{p_{11} \left(1 + \left(1 - \frac{1}{\alpha} \right) \phi_{ck} \right)^2 + (p_{12} + p_{21}) \left(\left(1 + \left(1 - \frac{1}{\alpha} \right) \phi_{ck} \right) \left(\frac{1}{\alpha} + \left(1 - \frac{1}{\alpha} \right) \phi_{cz} \right) \right) + p_{22} \left(\frac{1}{\alpha} + \left(1 - \frac{1}{\alpha} \right) \phi_{cz} \right)^2}
 \end{aligned}$$

If we then use *result 2* that explains the A matrix in just structural parameters and insert them into *result 3* which gives us the combinations behind ϕ_{ck} and ϕ_{cz} , we have the following combinations.

$$1 + \left(1 - \frac{1}{\alpha} \right) \phi_{ck} = 1 + \left(1 - \frac{1}{\alpha} \right) \left(\frac{\left(\frac{1}{\frac{1+\beta(\delta-1)}{\alpha} + \beta(1-\delta)} - \left(1 + \frac{\frac{1}{\beta} + \delta - 1}{\alpha} + \delta \right) \right)}{\frac{-\frac{1}{\beta} - 1}{\alpha}} \right) = 1 + \left(\frac{(\alpha-1) \left(\frac{1}{\frac{1+\beta(\delta-1)}{\alpha} + \beta(1-\delta)} - \left(1 + \frac{\frac{1}{\beta} + \delta - 1}{\alpha} + \delta \right) \right)}{-\alpha \left(\frac{1}{\beta} + 1 \right)} \right)$$

$$\frac{1}{\alpha} + \left(1 - \frac{1}{\alpha} \right) \phi_{cz} = \frac{1}{\alpha} + \left(1 - \frac{1}{\alpha} \right) \left(\frac{\left(\frac{1}{\frac{1+\beta(\delta-1)}{\alpha} + \beta(1-\delta)} \left(\left(\frac{1}{\frac{1+\beta(\delta-1)}{\alpha} + \beta(1-\delta)} - \left(1 + \frac{\frac{1}{\beta} + \delta - 1}{\alpha} + \delta \right) \right) \right)}{\left(1 + \frac{\frac{1}{\beta} + \delta - 1}{\alpha} + \delta \right) \left(\frac{-\frac{1}{\beta} - 1}{\alpha} \right)} - 1 \right) = \frac{2}{\alpha} +$$

$$\left(1 - \frac{1}{\alpha} \right) \left(\frac{\left(\frac{1}{\frac{1+\beta(\delta-1)}{\alpha} + \beta(1-\delta)} \left(\left(\frac{1}{\frac{1+\beta(\delta-1)}{\alpha} + \beta(1-\delta)} - \left(1 + \frac{\frac{1}{\beta} + \delta - 1}{\alpha} + \delta \right) \right) \right)}{\left(1 + \frac{\frac{1}{\beta} + \delta - 1}{\alpha} + \delta \right) \left(\frac{-\frac{1}{\beta} - 1}{\alpha} \right)} \right) - 1$$

If we insert these combinations into the likelihood function we get our final result.

References:

- Bank of England (2004): The new Bank of England Quarterly Model. *Quarterly Bulletin* 2, Bank of England
- Bayoumi, T. (2004): GEM: A new international macroeconomic model. *Occasional Paper* 239, International Monetary Fund
- Bernanke, B.S. and A.S. Blinder (1992): The Federal Funds Rate and the Channels of Monetary Transmission. *The American Economic Review*, vol. 82, no. 4, 901-921
- Bjørnstad, R. and R. Nymoen: Den norske modellen for lønnsdannelse: Enda viktigere etter ny politikk. *Økonomisk forum*, nr. 7, 4-13
- Blanchard, O.J. and C.M. Kahn (1980): The solution of linear difference models under rational expectations. *Econometrica*, vol. 48, 1305-1311
- Brubakk, L, T.A. Husebø, J. Maih, K.Olsen and M. Østnor (2006): Finding NEMO: Documentation of the Norwegian economy model. *Norges Bank Staff Memo Monetary Policy No. 2006/6*
- Christiano, L.J., M. Eichenbaum and C. Evans (2001): Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Federal Reserve Bank of Cleveland Working Paper*
- Christiano, L.J., M. Eichenbaum and C. Evans (2005): Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy*, vol. 113, no.1
- Fenton, P. and S. Murchison (2006): ToTEM: The Bank of Canada's New Projection and Policy-Analysis model. *Bank of Canada Review*, vol. 3, 2006
- Gali, J. (2000): *New Perspectives on Monetary Policy, Inflation, and the Business Cycle*. Barcelona: Universitat Pompeu Fabra
- Green W.H. (2003): *Econometric Analysis*, 5th Edition. *Pearson Education, Prentice Hall*

Hamilton, J.D (1994): Time Series Analysis. *Princeton University Press*

Hansen, G.D. (1985): Indivisible labor and business cycle. *Journal of Monetary Economics* 16, 309-327

Heijdra, B.J. and F. van der Ploeg (2002): The Foundations of Modern Macroeconomics. *Oxford University Press*

Holmsen, A., J. Nicolaisen and Ø. Røisland (2007): Pengepolitikens rolle og økonomiske modeller. *Økonimisk forum*, nr. 9, 4-8

Kalman, R.E. (1960): A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82, 34-45

Kalman, R.E. and R.S. Bucy (1961): New Results in Linear Filtering and Prediction Theory. *Journal of Basic Engineering*, 83, 95-108

Keynes, J.M. (1936): The General Theory of Employment, Interest, and Money. Macmillan, London

Kydland, F.E. and E.C. Prescott (1982): Time to Build and Aggregate Fluctuations. *Econometrica*, vol. 50, 1345-1370

Leeper, E.M. and C.A. Sims (1994): Toward a Modern Macroeconomic Model Usable for Policy Analysis. *NBER Macroeconomics Annual*, vol. 9, 81-118

Lucas, R.E. (1976): Econometric policy analysis: A critique. *Carnegie-Rochester Conference on Public Policy*, 1 19-46

Mork, K.A. (2008): Modeller og pengepolitikk. *Samfunnsøkonomene*, nr. 1, s. 12-15

Meinhold, R.J. and N.D. Singpurwalla (1983): Understanding the Kalman Filter. *The American Statistician*, vol. 37, no.2

Muth, J.F. (1961): Rational Expectations and the Theory of Price Movements. *Econometrica*, vol. 29, no. 3, 315-335

Nymoene, R. and E. Tveter (2007): Er Norges Banks pengepolitiske modell en god nok modell for norsk økonomi. *Økonomisk Forum*, nr. 5, 45-54

Rotemberg, J.J. and M. Woodford (1997): An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy. *NBER Macroeconomics Annual*, vol. 12, 297-346

Rotemberg, J.J. and M. Woodford (1999): Interest Rate Rule in an Estimated Sticky Price Model. In John B. Taylor, ed., *Monetary Policy Rules*. Chicago: University of Chicago press

Ruge-Murcia, F.J. (2007): Methods to estimate dynamic stochastic general equilibrium models. *Journal of Economic Dynamics & Control* 31 (2007), 2599-2636

Sargent, T.J. (1987): *Dynamic Macroeconomic Theory*. Harvard University Press, Cambridge, Massachusetts and London, England

Sargent, T.J. (1996): Expectations and the nonneutrality of Lucas. *Journal of Monetary Economics* 37, 535-548

Smith, A.A (1993): Estimating nonlinear time-series models using simulated vector autoregressions. *Journal of Applied Econometrics* 8, 63-84

Shone R. (1997): *Economic Dynamics*. Cambridge University Press

Snowdon, B. and H.R. Vane (1996): The Development of Modern Macroeconomics: Reflections in the Light of Johnson's Analysis after Twenty-Five Years. *Journal of Macroeconomics*, vol 18, No. 3, 381-401