STUDY CENTER GERZENSEE

Monetary Policy in a Small Open Economy Model: A DSGE-VAR Approach for Switzerland

Gregor Bäurle Tobias Menz

Working Paper 08.03

This discussion paper series represents research work-in-progress and is distributed with the intention to foster discussion. The views herein solely represent those of the authors. No research paper in this series implies agreement by the Study Center Gerzensee and the Swiss National Bank, nor does it imply the policy views, nor potential policy of those institutions.

Monetary Policy in a Small Open Economy Model: A DSGE-VAR Approach for Switzerland*

Gregor Bäurle† Tobias Menz‡

November 21, 2008

Abstract

We study the transmission of monetary shocks and monetary policy with a behavioral model, corrected for potential misspecification using the DSGE-VAR framework elaborated by DelNegro and Schorfheide (2004). In particular, we investigate if the central bank should react to movements in the nominal exchange rate. We contribute to the empirical literature as we use Swiss data, which is very rarely used in that context.

JEL-Classification: C11, C32, E58

Keywords: Bayesian Analysis, DSGE-VAR, Small Open Economy, Optimal Policy

---

*We thank Fabio Canova, Harris Dellas, Klaus Neusser, Frank Schorfheide and the seminar participants at the University of Bern for valuable discussions and comments. The usual disclaimer applies.

†University of Bern, gregor.baeurle@vwi.unibe.ch

‡University of Bern and Study Center Gerzensee, tobias.menz@vwi.unibe.ch

Contents

1 Introduction

During the past two decades, many countries have moved their monetary policy regime to inflation targeting.[1] As described by Svensson (2000) inflation targeting is characterized by (1) an explicit quantitative inflation target, either an interval or point target, (2) an operating procedure that uses a conditional inflation forecast as an intermediate target variable, and (3) a high degree of transparency and accountability. Furthermore, as argued in Svensson (1999), inflation targeting can be modeled as minimization of a specific loss function of the central bank. The operating procedure ensures that the first-order conditions of this minimization problem are approximately fulfilled. The role of transparency in this case is to what extent outside observers can verify if those conditions are fulfilled. A high degree of transparency therefore increases the incentives for the central bank to minimize the assigned loss function. This procedure results in an endogenous reaction function, that is the monetary policy instrument is expressed as function of all relevant information.

It has become a standard in empirical macroeconomics to use the short-term interest rate as the policy instrument and model the reaction function in the spirit of Taylor (1993). In the original version of the so called Taylor-rule, the short-term interest rate depends solely on current inflation and the output gap. In general, however, the reaction function will depend on much more information, that is in principle, it will depend on anything that has informational content for the central bank's conditional inflation forecast. In the case of open economies, it is therefore natural to ask whether variables such as the exchange rate or foreign interest rates should also be taken into account when formulating monetary policy.

In recent years, there has been extensive research on this question, namely how monetary policy should be best conducted in open economies that face movements in the nominal and real exchange rates. This issue has become particularly interesting in the aftermath of the breakdown of many fixed exchange rate regimes in the 1990's.[2] As Taylor (2001, p.267) puts it: "An important and still unsettled issue for monetary policy in open economies is how much of an interest-rate reaction there should be to the exchange rate in a monetary regime of a flexible exchange rate, an inflation

[1]Among them are Australia, Canada, Finland, New Zealand, Sweden and the UK. Switzerland constitutes a special case in that the Swiss National Bank does not consider itself an inflation targeter. Their main argument is that they do not commit to reach an inflation target within a specific horizon. For the purpose of this paper however, the differences between Swiss monetary policy implementation and inflation targeting are of definitional nature.

[2]For a comprehensive review of the literature on fixed exchange rate regimes see Garber and Svensson (1995).

target, and a monetary-policy rule." An early contribution on this debate was made by Obstfeld and Rogoff (1995). They propose a "rule of thumb" calling for a relaxing of monetary policy following a substantial appreciation of the real exchange rate. Since an appreciation in real terms makes foreign goods relatively cheaper compared to domestic goods, domestic aggregate demand is contracted through a reduction in net exports. As a consequence, interest rates should be lowered in order to mitigate the contractionary effects on domestic aggregate demand. Although intuitive, this rule of thumb remains speculative as it relies to some extent on partial equilibrium reasoning.

To gain meaningful insights on this issue, it is of interest to answer the question within the framework of a general equilibrium model. Several theoretical studies have done so. The approach termed as "new normative macroeconomic research" by Taylor (2001) has been used overwhelmingly to assess this issue. Roughly speaking, this means that the researcher builds a macroeconomic model including a monetary policy rule. The model is then solved using one of the numerous numerical algorithms[3] and the properties of the variables examined. Based on a loss-function, statements about the optimal policy rule can be made. The general finding is that including the exchange rate to the policy rule does not significantly improve, and sometimes even worsens, macroeconomic performance. Ball (1999) studies a simple small open economy model with sticky prices. He finds that the optimal policy parameters for the exchange rate are non-zero and quite large in size. However, macroeconomic performance measured by the volatility of inflation is only improved very modestly compared to a policy rule excluding the exchange rate. Using a different model with forward looking agents and more explicit microfoundations, Svensson (2000) performs a similar exercise. He also finds quite sizable optimal parameter values for the exchange rate in the policy rule. His simulations show that the central bank can indeed lower the volatility of inflation when following this rule. However, this comes at the expense of output variability because the variance of output increases. Another study of this class of policy rules is carried out by Taylor (1999). Using a seven-country model with France, Germany, and Italy joined into a single currency union representing the European Monetary Union and with Canada, Japan, the UK, and the US conducting their own monetary policy, he simulates a policy rule that includes a reaction to the exchange rate for the European Central Bank. Compared to a simple rule excluding exchange rate, he finds that the exchange rate reaction leads to a better performance for some countries in Europe, but to a poorer perfor-

[3]Among them are e.g. King and Watson (1998), Uhlig (1998), Klein (2000), or Sims (2002), all building on the principle idea of Blanchard and Kahn (1980).

mance for others. In summary, the above mentioned papers claim that the addition of the exchange rate to a standard Taylor-rule may worsen, or at best, improves only insignificantly the macroeconomic performance. Collard and Dellas (2002) even find evidence that including the exchange rate can have substantial negative effects on macroeconomic performance and welfare.

Most of this literature is based on calibrated dynamic stochastic general equilibrium (DSGE) models. That is, formal econometric methods are not invoked. Taking a first step towards filling this gap, Lubik and Schorfheide (2007) estimate a small scale DSGE model of a small open economy using a Bayesian approach. They contrast this full information system equation method to a single equation instrumental variables estimation, emphasizing the advantages of the former. The main goal is to investigate whether central banks in the UK, Canada, New Zealand and Australia actually target exchange rates. They find that the Bank of Canada and the Bank of England do, whereas there is no evidence for a reaction in New Zealand and Australia. The question of optimal policy is not adressed in the paper. Justiniano and Preston (2008) estimate a more elaborated small open economy DSGE model using data for Canada, Australia, and New Zealand. They evaluate optimal policy by minimizing a weighted objective of output and inflation variability over a set of generalized Taylor rules. They find that the optimal coefficient on the exchange rate is zero.

The aim of our paper is twofold. On the one hand, we follow Lubik and Schorfheide (2007) by empirically assessing whether the Swiss National Bank targets the nominal exchange rate in its policy rule. On the other hand, we address the question whether it is optimal for the central bank to react to movements in the nominal exchange rate by performing some policy experiments similar to those of Justiniano and Preston (2008). Methodically, we extend the analysis by employing a method that is capable of taking possible misspecification of the underlying theoretical model into account.[4] The method is also based on full information, Bayesian techniques. In a first step we therefore solve a medium scale small open economy DSGE model. Then, when taking the model to the data we use the approach proposed by DelNegro and Schorfheide (2004), which can roughly be described as confronting the DSGE model with a more general statistical model, a vector autoregressive model (VAR). The DSGE model is interpreted as prior information about parameter combinations in the unrestricted VAR. This prior information is augmented with information contained in the data,

[4]Lees, Matheson, and Smith (2007) also use this method to re-estimate the model of Lubik and Schorfheide (2007). They do investigate the question of optimal policy. However, they do not test whether the coefficients on exchange rates are positive. Furthermore, their technical implementation seems to be based on the model without taking misspecification into account.

that is, with the likelihood of the VAR. The method allows to express confidence in the exact restrictions of the DSGE model by increasing the tightness of the prior. The resulting estimated model, the DSGE-VAR, is a precise description of the data. Most importantly, depending on the tightness of the prior, it partly has a structural interpretation. Intuitively, when doing a policy experiment, the structural DSGE model predictions are 'corrected' by a non-structural component, which is policy invariant by assumption. The less confidence the researcher has in his DSGE model, the more weight is given to this statistical, only data-based correction.

We find that the Swiss National Bank did target the exchange rate. The posterior odds ratio overwhelmingly favors a model with the exchange rate included in the Taylor rule. Moreover, our optimal policy exercises shows that a non-zero reaction of the policy instrument to the nominal exchange rate lowers volatility of output and inflation. This is at odds with the findings of Justiniano and Preston (2008).

The remainder of the paper is organized as follows. In section 2 we present the theoretical model. Section 3 describes the empirical method and section 4 the implementation. Results are presented in section 5. Finally, section 6 concludes.

2 The DSGE model

The theoretical model is based on Monacelli (2005) and follows closely the specifications introduced by Justiniano and Preston (2008). The model consists of two countries, one being the small open economy and the other the rest of the world. The small open economy is populated by an infinitely-lived, representative household that consumes, supplies labor, and invests in either domestic or foreign one-period bonds. The interest rate on foreign bonds is subject to a risk-premium. On the production side of the economy, there is a sector of a continuum of monopolistically competitive firms producing a variety of domestic goods and selling them both domestically and abroad. There is also a monopolistically competitive retail sector in which a continuum of retailers import differentiated products from the rest of the world and sell them on the domestic market. Both sectors are assumed to be subject to a staggered price setting problem à la Calvo. As a result of this setup, imports are subject to local currency pricing, i.e. the law-of-one price is violated in the short-run. There is in principle another production sector in the economy: the final goods sector. However, it will not be modeled explicitly because it can be thought of as a perfectly competitive firm making zero profit that buys domestic and foreign varieties and turns them into a final consumption good that is sold to the household. The monetary policy instrument is the short-term interest rate, so we assume a generalized Taylor-type policy rule. The rest of the world is large compared to the small open economy. Therefore, although there is trade between the two countries, the imported and exported quantities are negligible relative to total foreign output. As a consequence, all foreign variables are taken as exogenous by the domestic economy.[5] Moreover we assume that foreign households are restricted from holding domestic bonds. In what follows, we briefly characterize the decision problems of each sector for the small open economy. Then we describe how the rest of the world is modeled. Finally, we summarize the linearized model economies. For the details of the derivations please consult Appendix B.

2.1 Domestic Household

The household problem is standard. The representative agent maximizes lifetime utility, subject to a budget constraint. She consumes, invests, and supplies labor to

[5]There is a subtlety to this point: technically, the model is a semi-small open economy model as the domestic producers have some market power. We circumvent this issue by assuming that the law of one price holds for those products.

the domestic firms. The utility function is specified as follows

$$E_0 \sum_{t=0}^{\infty} \beta^t \, \tilde{\varepsilon}_{g,t} \left[\frac{(C_t - H_t)^{1-\sigma}}{1-\sigma} - \frac{N_t^{1+\varphi}}{1+\varphi} \right]$$

where $\tilde{\varepsilon}_{g,t}$ is a preference shock and $H_t = hC_{t-1}$ is a habit formation term that is taken as exogenous by the household. Aggregate consumption is given by

$$C_t = \left[(1-\alpha)^{\frac{1}{\eta}} C_{H,t}^{\frac{\eta-1}{\eta}} + \alpha^{\frac{1}{\eta}} C_{F,t}^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}} \tag{1}$$

where $C_{H,t}$ and $C_{F,t}$ are the bundles of domestically produced and imported goods, respectively.[6] They are themselves aggregates of the different varieties. The elasticity of substitution between domestic and foreign goods is given by η and α is the weight of foreign goods relative to total consumption. In other words, α is a measure for trade openness. Let D_t denote bonds denominated in domestic currency and B_t bonds denominated in foreign currency. Then, the period budget constraint is given by

$$\begin{aligned} P_t C_t + D_t + \tilde{e}_t B_t = {} & D_{t-1}(1+i_{t-1}) + \tilde{e}_t B_{t-1}(1+i^*_{t-1})\phi_t \\ & + W_t N_t + \Pi_{H,t} + \Pi_{F,t} + T_t \end{aligned} \tag{2}$$

where $\tilde{e}_t$ denotes the nominal exchange rate[7], T_t is a lump-sum transfer from the government, and $\Pi_{H,t}$ and $\Pi_{F,t}$ are profits from the domestic firms and retailers, respectively. The domestic household must pay a risk-premium[8] in order to obtain funds from abroad. We follow Benigno (2001) and Schmitt-Grohé and Uribe (2003) by assuming that the gross premium ϕ_t is a function of aggregate net foreign debt B_{t-1} and a random shock $\varepsilon_{s,t}$. It proves convenient to express the risk premium in terms of the real quantity of net foreign debt denominated in domestic currency units as a fraction of steady state output. Formally, we let $A_t \equiv \tilde{e}_t B_t / (\bar{Y} P_t)$ and $\phi_{t+1} = f(A_t, \varepsilon_{s,t})$ with $f_1(\cdot) > 0$, where $\varepsilon_{s,t}$ is a risk-premium shock.

The household's optimization problem requires that the expenditures for domestic

[6] The aggregator in (1) can be interpreted as an Armington-production function of the representative final goods firm. As described above, we do not explicitly model this sector because it can be easily incorporated into the household problem.

[7] In our notation, $\tilde{e}_t$ is the domestic price of foreign currency.

[8] The risk-premium is introduced mainly for technical reasons. One the one hand, this constitutes a convenient way to avoid the unit-root problem for consumption characterizing many small open economy models (see Schmitt-Grohé and Uribe, 2003, for more Details). On the other hand, it allows us to introduce an economically interpretable shock that we need for the estimation of the model.

and foreign goods are cost-minimizing for any level of aggregate consumption. This implies the following demand functions for domestic and foreign goods

$$C_{H,t} = (1-\alpha)\left(\frac{P_{H,t}}{P_t}\right)^{-\eta} C_t, \qquad C_{F,t} = \alpha\left(\frac{P_{F,t}}{P_t}\right)^{-\eta} C_t \tag{3}$$

together with the theoretically correct consumer-price index

$$P_t = [(1-\alpha)P_{H,t}^{1-\eta} + \alpha P_{F,t}^{1-\eta}]^{\frac{1}{1-\eta}} \tag{4}$$

where $P_{H,t}$ and $P_{F,t}$ are the prices indices for the domestic and foreign consumption bundles. Given the demand functions (3), the household maximizes its lifetime utility subject to the budget constraint (2) by choosing optimally how much to consume, work, and invest. This yields the following set of optimality conditions:

$$(C_t - hC_{t-1})^\sigma N_t^\varphi = \frac{W_t}{P_t} \tag{5}$$

$$\lambda_t = \beta E_t\left[\lambda_{t+1}(1+i_t)\frac{P_t}{P_{t+1}}\right] \tag{6}$$

$$0 = E_t\left[\frac{\lambda_{t+1}}{P_{t+1}}\left((1+i_t) - (1+i_t^*)\frac{\tilde{e}_{t+1}}{\tilde{e}_t}\phi_{t+1}\right)\right] \tag{7}$$

with $\lambda_t = \tilde{\varepsilon}_{g,t}(C_t - hC_{t-1})^{-\sigma}$ being the marginal utility of consumption. Equation (5) describes the optimal labor supply schedule, equation (6) is the standard Euler equation, and equation (7) is an arbitrage condition restricting the relative movements of domestic and foreign interest rates and changes in the nominal exchange rate.

2.2 Domestic Firms

This section describes the main equations of the profit maximization problem for the domestic firms. There is a continuum of monopolistically competitive domestic firms of mass 1. Each firm i produces a differentiated good using labor as single input. The individual and aggregate production functions are given by

$$y_{H,t}(i) = \tilde{\varepsilon}_{a,t}\, N_t(i) \quad \text{and} \quad Y_{H,t} = \left(\int_0^1 y_{H,t}(i)^{\frac{\varepsilon-1}{\varepsilon}} di\right)^{\frac{\varepsilon}{\varepsilon-1}} \tag{8}$$

where ε is the elasticity of substitution between the different varieties and $\tilde{\varepsilon}_{a,t}$ represents technological innovation that is common to all firms. Because the goods are imperfect substitutes, each firm has some degree of monopolistic power when setting prices. In doing so, firms take into account that they face downward sloping demand curves. The domestic demand for variety i that emerges from the household problem is given by

$$C_{H,t}(i) = \left(\frac{P_{H,t}(i)}{P_{H,t}}\right)^{-\varepsilon} C_{H,t} \tag{9}$$

In words, this means that the quantity demanded of good i varies inversely with its price $P_{H,t}(i)$. We assume staggered price setting à la Calvo where θ_H denotes the probability that the firm cannot reset its price. The implied price duration is then $1/(1-\theta_H)$. Firms that cannot re-optimize follow an indexation rule of the form

$$P_{H,t}(i) = P_{H,t-1}(i)\left(\frac{P_{H,t-1}}{P_{H,t-2}}\right)^{\delta_H} \tag{10}$$

That means, these firms automatically adjust prices taking past inflation of domestic goods prices into account. The parameter δ_H indicates to what degree they react to past inflation. Although firms are heterogeneous ex ante, we will only consider the symmetric equilibrium in which all firms behave identically and can consequently omit the index i in what follows. Firms that can reset their prices in period t therefore all set the same one which we denote by $P'_{H,t}$. It can be shown that the price index for domestic goods evolves according to

$$P_{H,t} = \left[(1-\theta_H)\, {P'_{H,t}}^{1-\varepsilon} + \theta_H \left(P_{H,t-1}\left(\frac{P_{H,t-1}}{P_{H,t-2}}\right)^{\delta_H}\right)^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}} \tag{11}$$

A firm choosing the optimal price in period t maximizes the present discounted value of profits, taking into account the probability of not being able to re-set prices in the future. Firms sell their goods both domestically and abroad. When assuming that foreign demand is of the same functional form as domestic demand (9), the demand curve faced in period $t+\tau$ for a firm that last re-set prices optimally in period t and henceforth just adjusted prices according to the indexation rule (10) is given by

$$C_{H,t+\tau|t} = \left(\frac{P'_{H,t}}{P_{H,t+\tau}}\left(\frac{P_{H,t+\tau-1}}{P_{H,t-1}}\right)^{\delta_H}\right)^{-\varepsilon} (C_{H,t+\tau} + C^*_{H,t+\tau}) \tag{12}$$

The expected discounted profit for a firm that can re-optimize its price in period t is given by

$$E_t \sum_{\tau=0}^{\infty} \theta_H^\tau Q_{t,t+\tau} C_{H,t+\tau|t}\left[P'_{H,t}\left(\frac{P_{H,t+\tau-1}}{P_{H,t-1}}\right)^{\delta_H} - P_{H,t+\tau} MC_{t+\tau}\right]$$

where $MC_t = W_t/(P_{H,t}\,\tilde{\varepsilon}_{a,t})$ are real marginal costs and $Q_{t,t+\tau}$ is a time-dependent stochastic discount factor. Under the assumption that households have access to a complete set of state-contingent claims, $Q_{t,t+\tau}$ is the pricing kernel of such a security maturing in $t+\tau$ and is given by $Q_{t,t+\tau} = \beta^{\tau}\Lambda_{t+\tau}/\Lambda_t$ with $\Lambda_t = \lambda_t/P_t$. Since the household is the owner of the firms and receives the profits, it directs firms to make their decisions based on the households intertemporal rate of substitution. The optimal price resulting from the firm's maximization problem is given by

$$P'_{H,t} = \frac{\varepsilon}{\varepsilon-1}\frac{\sum_{\tau=0}^{\infty}\theta_H^{\tau}E_t\left[Q_{t,t+\tau}C_{H,t+\tau|t}P_{H,t+\tau}MC_{t+\tau}\right]}{\sum_{\tau=0}^{\infty}\theta_H^{\tau}E_t\left[Q_{t,t+\tau}C_{H,t+\tau|t}(P_{H,t+\tau-1}/P_{H,t-1})^{\delta_H}\right]} \tag{13}$$

2.3 Domestic Retailers

The domestic retailers import foreign differentiated goods for which the law of one price holds at the docks. However, we assume that also the retail sector is characterized by monopolistic competition so each retailer has some degree of price setting power. In other words, imports are subject to local currency pricing which gives rise to deviations from the law of one price in the short run. Analogous to the domestic goods producing firms, retailers face a staggered price setting problem with indexation. The price stickiness parameter for this sector is denoted by θ_F and the indexation parameter by δ_F. The indexation rule and the price index for imports can then be expressed along the lines of (10) and (11). When focussing again only on the symmetric equilibrium in which all retailers behave identically, the demand faced by a retailer in period $t+\tau$ conditional on having last re-optimized its price in period t is given by

$$C_{F,t+\tau|t} = \left(\frac{P'_{F,t}}{P_{F,t+\tau}}\left(\frac{P_{F,t+\tau-1}}{P_{F,t-1}}\right)^{\delta_F}\right)^{-\varepsilon} C_{F,t+\tau} \tag{14}$$

The expected discounted profit for a retailer that can re-set its price in period t is given by

$$E_t\sum_{\tau=0}^{\infty}\theta_F^{\tau}Q_{t,t+\tau}C_{F,t+\tau|t}\left[P'_{F,t}\left(\frac{P_{F,t+\tau-1}}{P_{F,t-1}}\right)^{\delta_F} - \tilde{e}_{t+\tau}P^{*}_{F,t+\tau}\right]$$

The optimal price results from maximizing this expression with respect to $P'_{F,t}$, subject to the demand (14) and is given by

$$P'_{F,t} = \frac{\varepsilon}{\varepsilon-1}\frac{\sum_{\tau=0}^{\infty}\theta_F^{\tau}E_t\left[Q_{t,t+\tau}C_{F,t+\tau|t}\tilde{e}_{t+\tau}P^{*}_{F,t+\tau}\right]}{\sum_{\tau=0}^{\infty}\theta_F^{\tau}E_t\left[Q_{t,t+\tau}C_{F,t+\tau|t}(P_{F,t+\tau-1}/P_{F,t-1})^{\delta_F}\right]} \tag{15}$$

2.4 Monetary Authority

As already mentioned, we assume that the monetary policy instrument is the short term interest rate i_t. We consider a policy rule for the central bank in which it takes information on current inflation, output, output growth, changes in the nominal exchange rate and past interest rates into account. We present the explicit specification in the summary of the linearized model equations below.

2.5 Market Clearing

The market clearing condition for the domestic economy requires

$$Y_{H,t} = C_{H,t} + C^*_{H,t} \tag{16}$$

where the left-hand-side is the supply of domestic goods and the right-hand-side is composed of domestic demand and export demand from the rest of the world. Following Kollmann (2002) we assume that export demand resembles domestic demand given by (3) and is given by

$$C^*_{H,t} = \varsigma \left(\frac{P^*_{H,t}}{P^*_t} \right)^{-\eta^*} Y^*_t \tag{17}$$

where we allow the foreign elasticity of substitution η^* to potentially differ from the domestic counterpart η.[9]

We restrict foreign investors from holding domestic bonds. Domestic bond market clearing therefore requires that net supply $D_t = 0$ for all t. The net supply of foreign bonds is zero as well. The reason for this is that number of domestic investors holding foreign debt is negligible relative to foreign investors.

2.6 The Foreign Economy

Instead of just assuming exogenous driving processes for all foreign variables, we want to express also the foreign variables as driven by fundamental economic innovations. We therefore take the model for the small open economy and - loosely speaking - solve it for the closed economy limiting case. In what follows, we point out the main differences. All foreign variables and parameters will be denoted by a $*$-superscript. Since the foreign economy is very large, trade flows to and from the domestic economy

[9]The functional form of (17) is assumed to be the same as in (3). The parameter ς represents the share of foreign imports to total foreign output. Although it is going to zero, we technically need it to have a well-defined steady state for C^*_H.

are negligible compared to total foreign economic activity. Technically, this means that α^* tends to zero. As a consequence, foreign consumption is given by $C_t^* = C_{F,t}^*$ which implies that the foreign consumers price index is entirely determined by foreign goods prices $P_t^* = P_{F,t}^*$. Foreign investors do not face a risk premium so the return on foreign bonds for them is simply $1 + i_t^*$. The price setting problem for foreign goods producing firms is exactly the same as for domestic firms. Since imports from the small open economy constitute a negligible share of total income, there is no foreign import sector. We assume that the foreign monetary authority also follows a Taylor-type rule without the inclusion of the exchange rate.

2.7 Equilibrium

The equilibrium is defined as a vector of prices $\{P_t^\star, P_{H,t}^\star, P_{F,t}^\star, W_t^\star, \tilde{e}_t^\star, i_t^\star, P_t^{*\star}, i_t^{*\star}, W_t^{*\star}\}$ together with a vector of allocations $\{C_t^\star, C_{H,t}^\star, C_{F,t}^\star, Y_{H,t}^\star, N_t^\star, B_t^\star, D_t^\star, Y_t^{*\star}\}$ such that equations (1)-(17) are satisfied. Before proceeding to the log linear approximation of the equilibrium conditions, let us define the real exchange rate as $\tilde{q}_t = \tilde{e}_t P_t^*/P_t$ and the terms of trade for the domestic economy as $S_t = P_{F,t}/P_{H,t}$. Furthermore, it will be convenient to define $\Psi_{F,t} = \tilde{e}_t P_t^*/P_{F,t}$ which according to Monacelli (2005) represents short run deviations from the law of one price. It will subsequently be referred to as the *law of one price gap.*

2.8 Log-Linear Approximation to the Model

For the estimation of the model we derive a log-linear approximation of the equilibrium conditions around a deterministic steady state. The steady state of the model is characterized by zero inflation and balanced trade. All variables are to be interpreted as log-deviations from the steady state. We use small letters unless otherwise noted, i.e. $x_t = \log(X_t/X)$.[10] We list the equations of the linear approximation. For the derivation we refer to the appendix.

The Euler equation (6) can be linearized straightforward to

$$c_t - hc_{t-1} = E_t(c_{t+1} - hc_t) - \frac{1-h}{\sigma}(i_t - E_t\pi_{t+1}) + \frac{1-h}{\sigma}(\varepsilon_{g,t} - E_t\varepsilon_{g,t+1}) \qquad (18)$$

The linearized domestic goods market clearing condition is given by

$$y_t = (1-\alpha)c_t + \alpha\eta(2-\alpha)s_t + \alpha\eta\psi_{F,t} + \alpha y_t^* \qquad (19)$$

[10]Exceptions are the nominal and real exchange rate which are defined as $e_t = \log(\tilde{e}_t/\tilde{e})$ and $q_t = \log(\tilde{q}_t/\tilde{q})$ as well as the shocks $\varepsilon_{x,t} = \log(\tilde{\varepsilon}_{x,t}/\tilde{\varepsilon}_x)$.

where the log of the law of one price gap is $\psi_{F,t} = e_t + p_t^* - p_{F,t}$ and the log terms of trade are $s_t = p_{F,t} - p_{H,t}$. Since for the estimation we use data on the real exchange rate, we must relate the former variable to the real rather than the nominal exchange rate. The terms of trade and the real exchange rate are related according to

$$q_t = e_t + p_t^* - p_t = \psi_{F,t} + (1-\alpha)s_t \tag{20}$$

Since we also use data on the import and export price deflators rather than the indices themselves, we take first differences of the terms of trade and obtain

$$\Delta s_t = \pi_{F,t} - \pi_{H,t} \tag{21}$$

A log-linear approximation to the domestic firms' optimality condition (13) and the price index for domestic goods (11) implies the following Phillips curve

$$\pi_{H,t} - \delta_H \pi_{H,t-1} = \beta E_t(\pi_{H,t+1} - \delta_H \pi_{H,t}) + \kappa_H mc_t \tag{22}$$

where $\kappa_H = \frac{(1-\theta_H)(1-\theta_H\beta)}{\theta_H}$ and the real marginal costs are given by

$$mc_t = \varphi y_t - (1+\varphi)\varepsilon_{a,t} + \alpha s_t + \frac{\sigma}{1-h}(c_t - hc_{t-1})$$

Similarly, a log-linear approximation to the retailers' optimality condition (15) yields another Phillips curve

$$\pi_{F,t} - \delta_F \pi_{F,t-1} = \beta E_t(\pi_{F,t+1} - \delta_F \pi_{F,t}) + \kappa_F \psi_{F,t} + \varepsilon_{cp,t} \tag{23}$$

where $\kappa_F = \frac{(1-\theta_F)(1-\theta_F\beta)}{\theta_F}$. We have augmented the Phillips curve with a cost-push shock $\varepsilon_{cp,t}$ to capture inefficient variations in mark-ups. Domestic CPI inflation and the domestic goods price deflator are related according to

$$\pi_t = \pi_{H,t} + \alpha \Delta s_t \tag{24}$$

Linearizing the uncovered interest rate parity condition (7) is straightforward and together with the definition of the real exchange rate yields

$$i_t - i_t^* = E_t\pi_{t+1} - E_t\pi_{t+1}^* + E_t\Delta q_{t+1} - \chi a_t - \varepsilon_{s,t} \tag{25}$$

The parameter χ captures the elasticity of the risk premium with respect to net foreign debt. A log-linear approximation to the budget constraint yields

$$c_t + a_t = \frac{1}{\beta} a_{t-1} - \alpha(s_t + \psi_{F,t}) + y_t \tag{26}$$

where $a_t = log(\tilde{e}_t B_t) - log(P_t \bar{Y})$ is the log real net foreign asset position as a fraction of steady state domestic income. Monetary policy enters in the model by assuming that the central bank follows a generalized Taylor rule of the form

$$i_t = \rho_i i_{t-1} + \psi_\pi \pi_t + \psi_y y_t + \psi_{\Delta y} \Delta y_t + \psi_e \Delta e_t + \varepsilon_{M,t} \tag{27}$$

$\varepsilon_{M,t}$ is a monetary policy shock. The foreign economy is described by an Euler equation

$$y_t^* - h^* y_{t-1}^* = E_t(y_{t+1}^* - h^* y_t^*) - \frac{1-h^*}{\sigma^*}(i_t^* - E_t \pi_{t+1}^* - \varepsilon_{g,t}^* + E_t \varepsilon_{g,t+1}^*), \tag{28}$$

a Phillips curve

$$\pi_t^* - \delta^* \pi_{t-1}^* = \beta^* E_t(\pi_{t+1}^* - \delta^* \pi_t^*) + \kappa^* mc_t^*, \tag{29}$$

with $\kappa^* = (1-\theta^*)(1-\theta^*\beta^*)/\theta^*$ and the real marginal costs given by

$$mc_t^* = \varphi^* y_t^* - (1+\varphi^*)\varepsilon_{a,t}^* + \frac{\sigma^*}{1-h^*}(y_t^* - h^* y_{t-1}),$$

and a Taylor rule

$$i_t^* = \rho_i^* i_{t-1}^* + \psi_\pi^* \pi_t^* + \psi_y^* y_t^* + \psi_{\Delta y}^* \Delta y_t^* + \varepsilon_{M,t}^* \tag{30}$$

Equations (18)-(30) constitute a system of linear rational expectations difference equations in the 13 variables $\{c_t, y_t, i_t, q_t, s_t, \pi_t, \pi_{H,t}, \pi_{F,t}, \psi_{F,t}, a_t, y_t^*, i_t^*, \pi_t^*\}$. When augmented by driving processes for the exogenous shocks, this system can be solved by means of a numerical routine. We will assume that the monetary policy shocks $\{\varepsilon_{M,t}, \varepsilon_{M,t}^*\}$ are distributed *iid* and that the remaining shocks follow univariate AR(1) processes given by

$$\varepsilon_{x,t} = \rho_x \varepsilon_{x,t-1} + \epsilon_{x,t} \quad \text{with} \quad E[\epsilon_{x,t}\epsilon_{x,t}'] = \sigma_x$$

When solving the system of difference equations, we seek for a representation of the endogenous variables in term of the exogenous shocks. We use the method of Sims (2002) and therefore define a N-dimensional state vector of endogenous variables

S_t[11], a vector of fundamental shocks $\mathcal{E}_t$, and a M-dimensional vector θ that contains all model parameters. The solution of the linearized system takes the form

$$S_t = G(\theta)S_{t-1} + H(\theta)\mathcal{E}_t$$

where the matrices $G(\theta)$ and $H(\theta)$ are complicated non-linear functions of the parameter vector θ.[12] It is important to see that although these matrices are large in dimension - $G(\theta)$ has the dimension $N \times N$ - they are restricted through the model such that the unknown number of parameters M is small.

[11]In our case, $N = 26$. The state vector S_t is composed of the 13 endogenous variables, the 6 exogenous shocks following an AR(1) process, and, because we are using the Sims-algorithm, 7 expectational variables.

[12]We will restrict our analysis to the parameter space which implies a unique stable solution. For a discussion of indeterminate stable solutions in linear rational expectations models, see Lubik and Schorfheide (2004).

3 Estimation Method: The DSGE-VAR Approach

In this section, we introduce the estimation method that will be used for our empirical analysis. It is referred to as the DSGE-VAR approach elaborated by DelNegro and Schorfheide (2004). First, we describe the general idea and discuss the main concepts intuitively. Then we show how to formally implement the idea. This involves the definition of an empirical model in order to derive the likelihood function and the formal definition of the prior distribution. We proceed by showing how a Markov-Chain Monte-Carlo (MCMC) algorithm can be used to simulate the posterior distribution of the DSGE-VAR and discuss how to choose the optimal weight of information from the DSGE model in the estimation. Finally, we describe how structural shocks can be identified using DSGE model restrictions and discuss how to do policy experiments with the estimated model. Much of the subsequent discussion follows closely DelNegro and Schorfheide (2004), DelNegro and Schorfheide (2006), and An and Schorfheide (2007).

3.1 General Idea

There is a close relationship between DSGE models and VARs: The DSGE model described in Section 2 is equivalent to a VAR(1) in the variables S_t. The parameters are constrained by the restriction functions $G(\theta)$ and $H(\theta)$. Let X_t be a subset of S_t that contains variables that can be observed. A general result is that X_t still has a VAR representation. However, the number of lags to be included is possibly infinite. The VAR(∞) can be approximated by a VAR of finite order p. The p first autocovariances of this approximation are equal to the first p autocovariances of the VAR(∞). It follows that the first p autocovariances of the DSGE model are equal to those of a restricted VAR(p). Hence, up to an approximation error in high order autocovariances, estimating the DSGE model is equivalent to estimating a restricted VAR(p). By increasing p, one can successively match autocovariances of higher orders and reduce the approximation error. Importantly, the restricted VAR inherits the properties of the DSGE model: Both are highly stylized versions of the real world by construction. Therefore, from a quantitative point of view, unrestricted VARs often have superior properties, suggesting that the restrictions implied by DSGE model are at odds with the data. On the other hand, unrestricted, non-structural VARs are not directly useful for policy analysis.

The DSGE-VAR approach of DelNegro and Schorfheide (2004) provides a solution to this dilemma. The key idea is the following: The researcher is aware of the fact that some simplifications in the modeling strategy potentially induce deviations of the model implied moments from the moments of the data. Still, she beliefs that the DSGE model contains the mechanisms which are important for policy analysis. Hence, it is useful to imbed the restrictions into the empirical model, but at the same time allow for deviations in order match the properties of the data more closely. The idea can be implemented using a Bayesian approach. In Bayesian analysis, the goal is to derive the distribution of the parameters using information contained in the data and prior beliefs about the parameters. The information contained in the data is described by the likelihood function of a model which is supposed to fit the data well. DelNegro and Schorfheide (2004) choose a VAR for this purpose and exploite the similarities to DSGE models described above. The information contained in the likelihood function is augmented with information from prior beliefs. Hence, it is natural to interpret the DSGE model restrictions as prior beliefs about the relationship of parameters in the VAR. Scepticism about the concrete implications of these beliefs can be introduced using a prior distribution that also assigns mass to VAR parameter combinations which are not exactly satisfying the restrictions. The degree of confidence in the model implication is higher if most of the prior mass is very close to the DSGE restrictions. If the prior distribution is degenerate in the sense the there is mass only at the exact restrictions, the fully restricted VAR approximation to the DSGE model is estimated. If the mass is rather spread, the researcher does not strongly belief in the restrictions implied by the DSGE model. In the limiting case of equally spread mass, an unrestricted VAR is estimated.

Dummy observation interpretation: The way DelNegro and Schorfheide (2004) set out the prior distribution has the following intuitive interpretation: The DSGE model is used to generate a sample of 'dummy observations'. These simulated artificial observations are added to the sample of observed data and the VAR is estimated on the augmented sample. How much the estimates are influenced by the DSGE model restrictions depends on the relative size of the simulated and the actual sample. When the sample of simulated observations is small relative to the actual sample, the estimates are not heavily restricted by the DSGE model. Increasing the size of the artificial sample imposes the restrictions implied by the DSGE model more tightly. For an artificial sample of infinite size, the restrictions of the DSGE model are fully imposed. We will parametrize the tightness of the prior as the ratio of the sample size of artificial data relative to the actual sample size and denote it by λ. This allows to choose how much the estimates are influenced by the restrictions implied by economic theory in a gradual way.

Updating the DSGE model parameters: One advantage of the DSGE-VAR approach is that the deep DSGE model parameters are estimated jointly with the VAR parameters. The intuitive description above is based on a given set of deep parameters in the DSGE model. However, the DSGE model parameters are rarely known with certainty. The researcher may have some uncertain a priori beliefs about their values. Hence, a prior distribution can be placed on the parameters to model the uncertainty. Then, a hierarchical prior can be constructed as follows: The prior for the VAR coefficients given a set of deep structural parameters is multiplied by a prior for these 'auxiliary' parameters. Now, since the model parameters appear in the prior distribution, they are updated with sample information via Bayes' theorem. The two limiting cases are illustrative for this (following Propositions 1 and 2 in DelNegro and Schorfheide, 2004): If the number of dummy observations goes to infinity, that is, if all the prior mass is located at the DSGE restrictions, only the deep DSGE model parameters are free to adjust. As already mentioned, this amounts to estimating a fully restricted VAR approximation to the DSGE model. If the number of dummy observations is small relative to the actual sample size, but both samples are large, the deep parameters are chosen such that a weighted discrepancy between moments implied by the unrestricted VAR and the DSGE model is minimized.[13] That is, the deep parameters are updated in such a way that the unrestricted moments of the data are matched as close as possible.

Optimal weight of the prior: So far, we have left open the question of how the tightness of the DSGE prior, that is, the optimal size of the sample of dummy observations, should be chosen. A natural way in our Bayesian framework is to derive the posterior probabilities of the DSGE-VAR model for different values of the tightness parameter λ. The posterior model probabilities essentially tell what the probability of a certain tightness is. Assuming a quadratic loss function of the decision maker, one would use a weighted average over the set of considered weights to calculate the statistics of interest. In practice, one often uses only the model with the highest posterior probability. There is a decision-based justification of this, assuming a more peculiar loss function. However, if the grid of weights for which the DSGE-VAR is estimated is not very fine, the weight of DSGE-VAR with the highest probability will be almost one anyway. That is, there is no relevant difference between the theoretically optimal choice and the practice of choosing the model with

[13] The weight depends on the relative size of artificial to actual data.

the highest posterior probability when the underlying loss function is quadratic. Based on posterior model probabilities, it is also possible to evaluate the degree of misspecifiation in the DSGE model. If the optimal weight is very low, the restrictions in the DSGE model are shown to be at odds with the data. If the optimal weight is high, theory helps improving the accuracy of the estimates. However, this evaluation will remain vague in the sense that there is nothing like a 'critical value' above which one can say that the DSGE model is particularly good. The fundamental reason for this is that it is not clear what the optimal non-structural reference model would be. A classical VAR(p), corresponding to zero weight of the DSGE prior hardly is the optimal model, even if the DSGE model is meaningless. A non-structural Minnesota style prior would help to get better model probabilities as well as a superior forecast performance. A further symptom of the problem will become apparent in our application. Depending on the choice of the number of lags p, the optimal weight varies: the higher p, the higher is the optimal weight. That is, the more restrictive the reference model is, the less prior information has to be induced.

Identification of shocks based on DSGE model restrictions: The mapping between structural shocks in the economy and the dynamic responses of the VAR variables to these shocks is not identified only by using information contained in the moments of the data. One has to use additional restrictions to relate the covariance matrix of the residuals to the structural shocks. A standard choice is to use a Cholesky decomposition of the covariance matrix of the residuals. This corresponds to timing restrictions: Some variables are assumed not to respond contemporaneously to certain structural shocks. The ordering of variables in the decomposition is of fundamental importance for the results, but usually ad-hoc and not justified by economic theory. In contrast, the DSGE-VAR allows to use economic theory to identify the shocks. Key is to recognize that in the DSGE model, the shocks are exactly identified. That is, there is a unique transformation of the Cholesky decomposition of the covariance matrix of the residuals in the DSGE model to the true mapping. The idea to identify the shocks in the DSGE-VAR is to transform the Cholesky decomposition of covariance matrix of the DSGE-VAR residuals with the unique transformation inferred from the DSGE model. The mapping from the shocks to the DSGE-VAR residuals will deviate from the mapping in the DSGE model if the implications of the DSGE model for the covariance matrix of the residuals are different from the actual covariance matrix in the DSGE-VAR.

Policy analysis in the DSGE-VAR framework: A central issue is how the DSGE-VAR model can be used for policy analysis. Obviously, as there are also non-structural elements involved, the Lucas critique (see Lucas, 1976) potentially applies. A useful result to study the properties is that the posterior mean of the VAR parameters can be decomposed into a 'structural component' and a 'correction' to match the data moments more closely. Depending on the tightness of the DSGE prior, the correction receives more or less weight. The critical assumption making policy analysis with the DSGE-VAR meaningful is that the correction is itself policy invariant. That is, switching to a different policy specification does not influence the way the original structural component should be corrected. This parallels the idea that the main policy dependent mechanisms are build into the model, but there are additional things going on in the economy which are relevant for the concrete quantitative properties of the data, but not the policy question at hand.

3.2 Formal Description

A fundamental result used in Bayesian analysis is that the posterior distribution is proportional to the likelihood function multiplied by the prior distribution. This is referred to as Bayes' theorem and can be stated formally as

$$p(\Theta|X) \propto p(X|\Theta)p(\Theta)$$

where X represents observed data, Θ are the unknown parameters of any model under consideration, and $p(\cdot)$ are generic density functions. In other words, given the likelihood function and the prior distribution, the (unnormalized) posterior distribution is identified. Hence, in a first step we seek to define the former two ingredients. As the mapping from the DSGE model parameters to the moments of the data is highly non-linear, we are not able to analytically derive the posterior distribution of the parameters. In a second step, we therefore describe a numerical method that generates draws from the posterior distribution. Subsequently, we justify the use of the marginal data densities as a measure of the posterior model probabilities and describe the so-called 'modified harmonic mean' estimator that is used to calculate the marginal data density. Finally, we formally describe the implementation of a generic policy experiment in the DSGE-VAR framework.

Likelihood function: To derive the likelihood function, we have to define how the data evolve given a particular set of parameters. We assume that the observable data vector X_t follows a vector autoregressive process of order p:

$$\Phi(L)X_t = e_t$$

where

$$\Phi(L) = 1 - \Phi_1 L - \ldots - \Phi_p L^p$$

and e_t is the vector of one step ahead forecast errors. The dimension of e_t as well as X_t is $N \times 1$. We assume that e_t is normally distributed with mean zero. Formally, this can be written as

$$e_t = H_{VAR}\varepsilon_t, \quad \varepsilon_t \sim iiN(0, I)$$

The normality assumption is made in order to simplify the derivation of the posterior density of the parameters. We will estimate the reduced form of this model: H_{VAR} is not identified without imposing additional restrictions. Only $\Phi(L)$ and $\Sigma = H_{VAR}H'_{VAR}$ can be estimated. It follows that the likelihood function is

$$p(X \mid \Phi, \Sigma) \propto |\Sigma|^{-T/2} \exp\left(-\frac{1}{2}tr(\Sigma^{-1}(X'X - \Phi' X'_P X - X' X_P \Phi + \Phi' X'_P X_P \Phi))\right)$$

where

$$X_P = \begin{pmatrix} X'_p & X'_{p-1} & \ldots & X'_1 \\ X'_{p+1} & X'_p & \ldots & X'_2 \\ \vdots & \vdots & \ddots & \vdots \\ X'_{T-1} & X'_{T-2} & \ldots & X'_{T-p} \end{pmatrix} \quad \text{and} \quad X = \begin{pmatrix} X'_{p+1} \\ X'_{p+2} \\ \vdots \\ X'_T \end{pmatrix}$$

Note that this likelihood function is conditional on the first p observations due to the lag-structure of the VAR.

Prior distribution: The prior distribution is conveniently written in hierarchical form:

$$p(\Sigma, \Phi, \theta) = p(\Sigma, \Phi|\theta)p(\theta)$$

where we denote the vector containing the DSGE model parameters by θ. The decomposition is useful because given θ, the prior for Σ and Φ has a standard form. In order to describe this conditional prior distribution, we define a vector of variables $X_t^*, \quad t = 1, \ldots, T^*$ as follows:

$$X_t^* = ZS_t$$

The selection matrix Z is chosen such that the variables in the model correspond to the observed data. X_t^* has a VAR representation of infinite order. This system is approximated by including only a finite number of lags:

$$X_t^* \approx \sum_{j=1}^{p} \Phi(\theta)_j^* X_{t-j}^* + e_t^*$$

with $e_t^* \sim iiN(0, \Sigma^*(\theta))$. Now define implied coefficient matrices $\Phi^*(\theta)$ and $\Sigma^*(\theta)$ as the maximum likelihood estimates (MLE) of Φ and Σ for a VAR(p) on an infinitely large sample of artificial observations. They are given by the Yule-Walker equations:

$$\Phi^*(\theta) = E(X_P^{*'} X_P^*)^{-1} E(X_P^{*'} X^*)'$$

$$\Sigma^*(\theta) = E(X^{*'} X^*) - E(X_P^{*'} X^*) E(X_P^{*'} X_P^*)^{-1} E(X_P^{*'} X^*)'$$

where

$$X_P^* = \begin{pmatrix} X_p^{*'} & X_{p-1}^{*'} & \dots & X_1^{*'} \\ X_{p+1}^{*'} & X_p^{*'} & \dots & X_2^{*'} \\ \vdots & \vdots & \ddots & \vdots \\ X_{T^*-1}^{*'} & X_{T^*-2}^{*'} & \dots & X_{T^*-p}^{*'} \end{pmatrix} \quad \text{and} \quad X^* = \begin{pmatrix} X_{p+1}^{*'} \\ X_{p+2}^{*'} \\ \vdots \\ X_{T^*}^{*'} \end{pmatrix}$$

The moments $E(X^{*'} X^*)$, $E(X_P^{*'} X^*)$ and $E(X_P^{*'} X_P^*)$ can easily be calculated given the solution to the DSGE model (for details we refer to the appendix of DelNegro and Schorfheide, 2004). Now, the prior distribution of Φ and Σ given θ is chosen to be of the following Inverted-Wishart-Normal form:

$$\Sigma|\theta \sim IW(\Sigma^*(\theta), T^* - Np - 1)$$
$$\Phi|\Sigma, \theta \sim N\left(\Phi^*(\theta), \Sigma \otimes \left(T^* E(X_P^{*'} X_P^*)\right)^{-1}\right)$$

The distributions are centered at Φ^* and Σ^*. One can immediately see that the prior distribution for Φ given Σ gets tighter around the MLE of Φ the larger the size of the artificial sample is. To justify the 'dummy observation' interpretation of the DSGE prior, it is illustrative to look more closely at the functional form of the density of a random variable with an Inverted-Wishart-Normal distribution. In our case

$$p(\Phi, \Sigma \mid \theta) \propto |\Sigma|^{-\frac{\lambda T^* + n + 1}{2}}$$
$$\exp\big\{ -\tfrac{1}{2} tr[\Sigma^{-1}(E(X^{*'}X^*) - \Phi' E(X_P^{*'}X^*) - E(X^{*'}X_P^*)'\Phi + \Phi' E(X^{*'}X^*)'\Phi)] \big\}$$

Comparing this density to the likelihood $p(X \mid \Phi, \Sigma, \theta)$, we see that it resembles the (quasi-) likelihood function[14] of dummy observations X^*

$$p(X^* \mid \theta) \propto |\Sigma|^{-\frac{T^*}{2}}$$
$$\exp\big\{ -\tfrac{1}{2} tr(\Sigma^{-1}(X^{*'}X^* - \Phi' X_P^{*'}X^* - X^{*'}X_P^{*'}\Phi + \Phi' X^{*'}X^{*'}\Phi)) \big\}$$

multiplied with an (improper) prior

$$p(\Phi, \Sigma) = \Sigma^{-\frac{n+1}{2}}$$

The prior density of Φ and Σ given θ only differs because, in order to avoid stochastic variation in the moments of the dummy observation, the simulated sample moments are replaced with their expectations. That is, our prior is chosen as if we estimated the parameters Φ and Σ based on the sample of observed and simulated data using only an improper prior. The sample size of the artificial sample is $T^* = \lambda T$, therefore λ is a parameter which reflects the 'tightness' of the DSGE model prior. The larger λ, the larger the sample compared to the actual sample. If λ is large, the estimates of Φ and Σ will concentrate on the restrictions implied by the DSGE model.
We have defined the prior density of Σ and Φ given θ so far. The prior density $p(\theta)$ remains to be determined. This is done following the standard strategies used in the literature on estimating DSGE models.

Posterior distribution: The posterior distribution can be factorized in the same way as the prior distribution:

$$p(\Sigma, \Phi, \theta | X) = p(\Sigma, \Phi | \theta, X) p(\theta | X)$$

Because of the choice of a conjugate prior for the VAR parameters given θ, the posterior of the same parameters given θ is of the same form as the prior:

[14] For two reasons, this can be the quasi likelihood function: First, the DSGE model does not necessarily have a VAR(p) representation. Second, the likelihood function corresponds to normally distributed shocks, which has not been assumed so far.

$$\Sigma \mid \theta, X \sim IW\left(\tilde{\Sigma}(\theta), (1+\lambda)T - Np - 1\right) \tag{31}$$

$$\Phi \mid \Sigma, \theta, X \sim N\left(\tilde{\Phi}(\theta), \Sigma \otimes (\lambda T E(X_P^{*\prime} X_P^*) + X_P' X_P)^{-1}\right) \tag{32}$$

It is centered at the MLE on both actual and artificial data, $\tilde{\Phi}(\theta)$ and $\tilde{\Phi}(\theta)$. Usually, the prior for θ does not have a simple standard form. Moreover, the mapping from θ to the moments of the artificial data is highly non-linear in most models. Hence, the marginal posterior distribution for θ will not have a standard form. To calculate the distribution, one has to use numerical methods, as described in the next section.

MCMC algorithm: The estimation proceeds in two steps, following the decomposition of the posterior distribution above. Step 1 is to produce J draws from $p(\theta|X)$. In Step 2, we draw from the posterior distribution of the VAR parameters given the draws from Step 1.

Step 1: Drawing from $\mathbf{p}(\theta \mid \mathbf{X})$: The distribution depends on prior knowledge about specific parameters in the model. Usually, there is no way to obtain a standard posterior distribution for θ. A possible way to draw from a non-standard distribution is a Random Walk Metropolis-Hasting (MH) Algorithm. Given an initial value θ_0, a candidate θ^* is drawn from a proposal distribution:

$$\theta^* = \theta^{j-1} + \epsilon_{prop}$$

where a standard choice for the distribution of ϵ_{prop} is a multivariate t-distribution. Then, the following ratio is calculated:

$$r = \frac{p(X \mid \theta^*)p(\theta^*)}{p(X \mid \theta^{j-1})p(\theta^{j-1})}$$

The proposal θ^* is accepted, that is we set $\theta^j = \theta^*$, with probability $r = \min(1, r)$. If the θ^* is rejected, we set $\theta^j = \theta^{j-1}$. These steps are repeated for $j = 1, \ldots, J$. For an exposition of MH algorithms and MH within Gibbs algorithms see Geweke (2005). The key result is that the sequence of draws forms a Markov Chain with unique stationary distribution $p(\theta \mid X)$. In order to mitigate the effects of θ_0, which is chosen arbitrarily, we discard a certain number of initial draws. A prerequisite for using the algorithm is that the likelihood can be evaluated for a given θ. The relevant likelihood is

$$p(X \mid \theta) \propto \frac{|E(X_p^{*\prime} X_p^*) + X_p' X_p|^{-\frac{N}{2}} |(\lambda+1)T\tilde{\Sigma}(\theta)|^{-\frac{(\lambda+1)T-k}{2}}}{|X_p^{*\prime} X_p^*|^{-\frac{N}{2}} |\lambda T \Sigma^*(\theta)|^{-\frac{\lambda T-k}{2}}}$$

Step 2: Drawing from $\mathbf{p}(\boldsymbol{\Phi}, \boldsymbol{\Sigma} \mid \theta, \mathbf{X})$: For $j = 1, \ldots, J$, we first draw Σ from the inverse Wishart distribution (31) and then draw Φ from the multivariate normal distribution (32). In each case, θ is replaced by θ_j from the first step.

Posterior model probabilities: To choose the optimal tightness of the DSGE prior, we will use a measure of in-sample fit. The measure is based on posterior model probabilities for a grid of weights $\lambda_1, \ldots, \lambda_I$. We index each model by its values for the weight of the DSGE prior λ_i and denote the respective models by $\mathcal{M}_{\lambda_i}$. We then calculate the posterior probabilities of each model:

$$p(\mathcal{M}_{\lambda_i}|X) = \frac{p(X|\mathcal{M}_{\lambda_i})p(\mathcal{M}_{\lambda_i})}{p(X)} = \frac{p(X|\mathcal{M}_{\lambda_i})p(\mathcal{M}_{\lambda_i})}{\sum_{j=1}^{I} p(X|\mathcal{M}_{\lambda_j})p(\mathcal{M}_{\lambda_j})}$$

To compare the different models, we put equal prior weight for the each model:

$$p(\mathcal{M}_{\lambda_i}) = \frac{1}{I}, \quad \forall i$$

Hence, in relative terms, only the posterior marginal data density is used as a measure of fit:

$$p(X|\mathcal{M}_{\lambda}) = \frac{p(X|\Theta, \mathcal{M}_{\lambda})p(\Theta|\mathcal{M}_{\lambda})}{p(\Theta|X, \mathcal{M}_{\lambda})}$$

where Θ are all the unknown parameters of the model. The selection criterion mechanically favors models with a high likelihood, but imposes a penalty on too loose priors. The density $p(X|\mathcal{M}_{\lambda})$ cannot be calculated analytically. However, the harmonic mean estimator proposed by Geweke (1999) can readily be applied. This estimator is based on the following identity:

$$\frac{1}{p(X)} = \int \frac{f(\Theta)}{p(X|\Theta)p(\Theta)} p(\Theta|X) d\Theta$$

where $\int f(\Theta) d\Theta = 1$. This expression is estimated with

$$\frac{1}{\widehat{p(X)}} = \frac{1}{J} \sum_{j=1}^{J} \frac{f(\Theta_j)}{p(X|\Theta_j)p(\Theta_j)}$$

In principle, any function $f(\Theta)$ which integrates to one can be used. A standard choice is

$$f(\Theta) = q^{-1}(2\pi)^{-d/2}|V_\Theta|^{-\frac{1}{2}}e^{-\frac{1}{2}(\Theta-\bar{\Theta})V_\Theta^{-1}(\Theta-\bar{\Theta})} \times I\left[(\Theta-\bar{\Theta})V_\Theta^{-1}(\Theta-\bar{\Theta}) < F^{-1}_{\chi^2_d(q)}\right]$$

$\bar{\Theta}$ refers to the posterior mean and V_Θ is the posterior variance of the draws. The parameter q is deliberately chosen to dampen the effect of extreme draws out of the posterior density. One word of caution may be necessary at this point: In theory, the value of q has no influence on the estimated value of the marginal data density. In practice, the estimation depends to some extent on the value of q due to the finite number of draws. It is therefore recommended to calculate $p(X)$ for various values of q.

Identification of shocks: The residuals in the DSGE-VAR relate to structural shocks ε_t as

$$e_t = H_{VAR}\varepsilon_t$$

with $E(\varepsilon_t\varepsilon_t') = \mathcal{I}_M$. We assume that H_{VAR} is invertible, which means that there are as many shocks as observed series. The goal is to estimate the reaction of the series X_t to the shocks ε_t:

$$H_{VAR} = \frac{\partial X_t}{\partial \varepsilon_t'}$$

Given the responses on impact $\frac{\partial X_t}{\partial \varepsilon_t'}$, one can use $\Phi(L)$ to calculate the responses of X_{t+h} for $h > 0$. The problem of identification arises because H_{VAR} can not be uniquely determined using only information from the reduced form estimation of the DSGE-VAR. H_{VAR} is only restricted by its relationship to the covariance matrix of the reduced form residuals:

$$\Sigma = H_{VAR}E(\varepsilon_t\varepsilon_t')H_{VAR}' = H_{VAR}H_{VAR}'$$

It is always possible to plug an orthonormal matrix Ω into the above equation:

$$\Sigma = H_{VAR}\Omega\Omega' H_{VAR}'$$

and define $\tilde{H} = H_{VAR}\Omega$. This matrix also satisfies the restrictions implied by the reduced form estimation. However, it implies potentially very different reactions of X_t to the shocks. Hence, given an arbitrary $\tilde{H}$, there have to be further restrictions on Ω in order to determine the responses H_{VAR}. The identification schemes used in the literature differ in the way Ω is chosen. DelNegro and Schorfheide (2004) propose an approach which relies on the fact that in the DSGE model the shocks are exactly identified. That is, the matrix

$$\frac{\partial Z S_t}{\partial \varepsilon_t'} = H_{DSGE}(\theta)$$

is uniquely determined. Recall that $H(\theta)$ can be calculated using standard methods to solve linear(ized) DSGE models. Furthermore, there is a unique decomposition of this matrix into the product of a triangular matrix $H_{tr,DSGE}(\theta)$ and an orthonormal matrix $\Omega(\theta)$:

$$H(\theta) = H_{tr,DSGE}(\theta)\Omega(\theta)$$

The idea is to set $\tilde{H}$ to $H_{tr,VAR}$, the Cholesky decomposition[15] of Σ, and then to use $\Omega(\theta)$ as a rotation:

$$H_{VAR}(\theta) = H_{tr,VAR}\Omega(\theta)$$

On impact, the responses differ to the extent that $H_{tr,DSGE}(\theta)$ and $H_{tr,VAR}$ differ. That is, if the covariance matrix of residuals is similar to its counterpart in the DSGE model, then the responses on impact will be close. For horizons bigger than zero, there is the influence of $\Phi(L)$ which allows for further deviations of the DSGE-VAR responses from the DSGE model implications.

Policy experiments: The ultimate goal is to use the estimated model for policy experiments. Usually, a subset of parameters in the DSGE model can be interpreted as 'policy parameters' in the sense that some authority can choose their values. In our case, these may be the parameters in the Taylor rule. Denote those parameters by θ_p and the non-policy parameters by θ_s. We want to calculate what the effects of changing θ_p is, given the estimated marginal distribution of θ_s. The post-invervention distribution of the DSGE model parameters is

$$\tilde{p}(\theta|Y) = p(\theta_s|Y,\theta_p)p(\theta_p) = p(\theta_s|Y)p(\theta_p)$$

The second equality follows from the assumption that θ_s are deep parameters, that is, they are not influenced by the policy change. We then calculate the post intervention

[15]There is a subtlety here: The Cholesky decomposition as well as the LR decomposition above is only unique up to the sign of each row. Intuitively, changing the sign of a shock and at the same time changing the sign of the response does not alter the results. Hence, one has to confirm that the 'same' sign convention is applied for both decompositions.

distribution of our VAR model parameters:

$$\tilde{p}(\Phi, \Sigma|Y) = \int p(\Phi, \Sigma|Y, \theta)\tilde{p}(\theta|Y)d\theta$$

Concretely, we use the posterior draws of θ and replace the policy parameters in this distribution with the desired values. We then apply Step 2 of the MCMC algorithm again for each draw of the adjusted distribution to draw from the post-intervention distribution of Φ and Σ. The new VAR parameters can then be used to calculate statistics of interest, such as the variability of output or inflation for example. In this way, it is possible use the DSGE-VAR model to evaluate the effects of policy changes. Note that the second step contains the 'correction mechanism': Instead of using directly the VAR implications of the DSGE model for a particular set of θ, that is replacing Φ and Σ in the integral above with the functions $\Phi^*(\theta)$ and $\Sigma^*(\theta)$, we use the conditional posterior distribution of the DSGE-VAR parameters. This conditional posterior distribution consists in part of this function, but it is augmented with the correction mechanism which is not influenced by the DSGE model parameters by assumption.

4 Empirical Implementation

4.1 Data

Following Justiniano and Preston (2008), we use data on output, inflation, interest rates, the terms of trade and the real exchange rate. A central issue are how data on the 'world' economy is obtained. Approximately 90% of the Swiss imports and almost 80% of all exports are to, respectively from the OECD countries. We therefore approximate output and prices of the world economy with data from these countries. For output we use real GDP of OECD countries available from the OECD.stat. Inflation is measured by the first difference of the implicit GDP deflator. For interest rates, we use the US short-term interest rate as this is presumably the most important interest rate for the world economy. Swiss data is obtain from the Swiss National Bank database and the Swiss State Secretariat for Economic Affairs (SECO). Output is measured by real GDP taken from the SECO database. Again, we use first differences of the implicit GDP deflator as a measure for inflation. The interest rate is the 3-month Libor, converted to quarterly data by averaging over the monthly series. The terms of trade are constructed from the implicit price deflators for imports and exports available from the SECO database. The real exchange rate is the index constructed by Swiss National Bank. As our theoretical model does not explicitly model a trend, we use output growth instead of output in levels. This would be consistent with a linear trend in output. Following Justiniano and Preston (2008) we also use first differences of the the terms of trade and the real exchange rate. The mapping Z is chosen such that the first differences in the data are directly linked to the first differences of the same variables in the model. We therefore augment the state vector S_t with lagged values of output, exchange rates and the terms of trade.

4.2 Prior Distribution and Model Specification

The prior distribution for the DSGE model parameters θ is obviously model dependent. Our model has already been estimated by Justiniano and Preston (2008) using Bayesian methods. We therefore closely follow their specification of the prior distribution. The differences are due to country specific implications of the data for steady state values. We decided to use a higher discount factor for Switzerland than the standard value of 0.99: β is calibrated to be 0.997 corresponding to an average annualized real interest rate of 1.3%. The prior mean of the openness parameter α is 0.43 according to the average of export and import shares (calculated as the sum of exports and imports divided by twice the output, averaged over all time periods).

We decided to estimate the parameters with a rather tight prior distribution.
With respect to the tightness of the DSGE prior, we estimate the model for a grid of values, $\lambda \in \{\frac{2}{3}, 0.8, 1, 1.5, 2, 3, 5, 10, 100\}$. We estimate the model for lag lengths 1 up to 4. In addition, we estimate the model without accounting for misspecification. Based on posterior model probabilities, we select one model out of these 37 estimated models for the policy evaluation.

4.3 MCMC Algorithm

The distribution of the innovations of the proposal draw in the MH-algorithm is a standard multivariate t-distribution with 40 degrees of freedom. To specify its covariance matrix, we proceed as follows: We search for the posterior mode of the DSGE model parameters using the standard likelihood function without adjusting for misspecification. The scaled inverse Hessian at the mode is used as covariance matrix. It is scaled in order to get an acceptance rate between 0.2 and 0.3. We have drawn 50 different sets of initial values out of the prior distribution for θ to initialize the search for the mode. For all the sets, the algorithm converged to the same posterior mode (up to a clearly numerical error). We therefore conclude that posterior distribution is reasonably well behaved. We iterate 1'000'000 times over Step 1 and 2 of the MCMC algorithm described in Section 3 for the grid of λ defined above.[16] As initial value for the MCMC algorithm, we also use the posterior mode based on the likelihood function of the DSGE model. To mitigate the effect of the initial values, we discard the first 20% of the draws. Convergence is checked by graphically verifying that the recursive means remain stable after removing the discarded draws. For computational reasons we evaluate only every 16th draw, such that we are left with 50'000 draws to calculate the distribution of the parameters. In order to select the optimal tightness of the prior, we calculate the harmonic mean estimator of $p(X|\lambda)$ for $q = \{0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95\}$ for each λ in the grid. The results differ across q's only to a small extent. Moreover, the ordering of the model is robust to the choice of q. For the model with the highest marginal data density, we additionally run a chain with 5'000'000 draws as well as 10 chains with 1'000'000 draws starting with initial values drawn out of the prior distribution to check the dependence on initial values. The differences across chains with respect to the posterior distribution of θ are very small. Furthermore, combining these chains resulted in virtually no difference to the posterior distribution obtained from the chain with 5'000'000 draws. Therefore, when evaluating the model with the optimal λ, we

[16] It is actually not necessary to do Step 2 if the aim is only to calculate the marginal data density.

only report results based on the latter. Again, we discard the first 20% of the draws. We calculate the statistics of interest using every 40th draw, such that we are left with 100'000 draws. For all the calculations, we use our own MATLAB programs, except for the posterior mode search where we rely on Christopher Sims' `csminwel` routines available on `http://sims.princeton.edu/yftp/optimize`. Whenever possible, we debugged our code by comparing its results to the Frank Schorfheide's GAUSS code available on `http://www.econ.upenn.edu/~schorf/research.htm`.

5 Results

In this section we present the results from our empirical analysis. First, we elaborate on the selection of the model which will be used for the subsequent analysis. Second, we discuss its prominent properties. Third, we focus on the coefficients in the Taylor rule for different specifications. Finally, we show the results from our policy experiments.

5.1 Model Selection

As described in section 3, the posterior marginal data density can be used as a selection criterion for choosing the optimal tightness of the DSGE prior. Figure 1 shows the marginal data densities for the estimations of DSGE-VAR models with lags 1 up to 4 over a finite grid of the tightness parameter λ. For our estimations the optimal model choice results in $p = 4$ lags and $\lambda = 2.5$. All subsequent analyses will be based on this choice.

Before we proceed, it is worthwhile noting a number of interesting features that emerge from this graphical representation. An outstanding characteristic is the inverse U-shape of the posterior marginal data densities as a function of the tightness parameter λ, irrespective of the number of lags included. The low data densities for high values of λ suggest that DSGE model induced restrictions are not supported by the data. On the other extreme, the low data densities for a low λ indicates that prior information is beneficial because it helps to reduce the parameter space. In sum, the qualitative shape of the data densities can be interpreted as evidence for the usefulness of the DSGE-VAR approach to improve the fit of data. Another striking feature is the fact that optimal weight λ varies with different lags. The reason for this can be attributed to the dual nature of the data density as selection criterion. On the one hand, in-sample fit is rewarded and on the other hand, model complexity is penalized. For small values of λ, the criterion favors a lower lag structure, meaning that more complex models are penalized heavily when prior information is diffuse. Finally, it is interesting to mention that the estimation of the "pure" DSGE model yields a lower data density

5.2 Model Properties

Information about the DSGE parameter estimates for the model chosen in the previous section is contained in Table 2. The means and 80% highest posterior distribution intervals (HPDI80) are reported for the prior and posterior distributions for all parameters. The first conclusion is that the data contains information on the parameters, though not on all of them. This can be seen in that prior and posterior means and HDPI80 do not coincide for most parameters. Turning to the estimates of some selected parameters, it is striking that the inverse intertemporal elasticity of substitution takes a very low value with the posterior mean being at 0.20. Furthermore, we find that habit formation plays essentially no role and also inflation indexation is not as prominent as claimed for instance by Christiano, Eichenbaum, and Evans (2005) for closed economies. Previous studies like Adolfson, Laseen, Linde, and Villani (2007) or Justiniano and Preston (2008) have found the parameter for the trade share α to be particularly difficult to identify, so it was calibrated. We find that although it is updated slightly towards zero compared to the prior, estimation yields a reasonable result with the posterior mean lying at 0.37. The parameters for price stickiness in the domestic and the import sector both take rather high values. The results suggest that prices in the import sector adjust more sluggishly than in the domestic sector. The implied price durations based on the posterior means are 3.1 and 2.4 quarters, respectively.

We do not report the implied VAR coefficients because of their excessive dimension.[17] Instead, we perform impulse-response analysis based on the identification described in section 3 to evaluate the dynamic properties of the model. Figure 2 shows the responses of the domestic output growth, inflation, interest rates as well as the terms of trade and the real exchange rate to a contractionary monetary policy shock. Consider first the impulse response functions of the purely estimated DSGE model which are plotted in red. Contractionary monetary policy appreciates domestic currency and lowers inflation and output. Because prices are sticky, the nominal appreciation entails a drop in the real exchange rate. The terms of trade go up at impact. This result is not evident a priori and can be explained by looking at equation (21). The change in the terms of trade is the difference of import and domestic price inflation. Depending on which one reacts more strongly to the monetary policy shock, the terms of trade can move in either direction. In our case, we have found the Calvo price stickiness parameter in the import sector to be higher than in the domestic sector. This implies a more gentle reaction of import price inflation to the monetary intervention resulting in the increase in the terms of trade. Turning to the comparison of the impulse responses of the plain model to those of the DSGE-VAR, we find that the broad picture is confirmed. However, there are certain discrepancies, e.g. the

[17] Note that we have estimates for 4 coefficient matrices and the covariance matrix. Each of them is of dimension (8×8). Even when accounting for the symmetric structure of the covariance matrix this would amount to reporting information on 292 VAR parameters.

reaction of the terms of trade are nearly zero. This can once again be interpreted as evidence that the DSGE-VAR approach is taking the misspecification of the DSGE model into account and matches the data more closely.

5.3 The Taylor Rule

The main question we want to address in this paper is whether the monetary authority should take movements in the nominal exchange rate into account when conducting monetary policy. The estimates of the coefficients in the policy rule are thus of special interest. They are already reported in Table 2. We additionally plot the prior and posterior densities in Figure 4. To begin with, it stands out that the data seems to contain information on all parameters except the coefficient on output. Interest rate smoothing seems to be an important objective in Swiss monetary policy. The coefficient on the lagged interest rate is updated distinctively towards one and achieves a mean value of 0.88. The coefficients on inflation and output are in the order of magnitude which is in line with the literature.

We now turn to the parameter which is in the center of our focus: the coefficient on the change in the nominal exchange rate $\psi_{\Delta e}$. Although the posterior distribution is pulled towards zero compared to the rather diffuse prior, a lot of its mass is clearly centered around a non-zero value. The posterior mean take a value of 0.12. A justified critique can be pointed at the fact that zero has only been included as boundary in the prior, therefore inherently favoring a non-zero support for the posterior distribution. To dispel this scepticism, we re-estimate a DSGE-VAR for the same specifications but restrict the exchange rate coefficient to be zero. Table 3 contains the log marginal data densities as well as the posterior odds ratios for different truncation values q. The posterior odds ratio can be taken as a criterion in favor or against the hypothesis $\psi_{\Delta e} = 0$ versus $\psi_{\Delta e} > 0$. The results clearly support a non-zero exchange rate response of Swiss policy.

5.4 Policy Experiments

In the previous section we have already presented evidence that favors the inclusion of the nominal exchange rate into the policy rule over the standard specification à la Taylor. In this section we present further findings that support this conclusion. The results of the policy experiments described in section 3 are summarized in Figures 5 and 6. It is assumed that the monetary authority sets the policy parameters but does not influence the others. In this exercise, we let $\psi_{\Delta e}$ vary over a grid and re-draw the

DSGE-VAR parameters from the post-intervention distribution. For each of the new post-intervention estimates, we simulate the model and evaluate the macroeconomic performance on the basis of both the variability of inflation and the variability of output. In both graphs, the blue line represents the distribution of inflation and output volatility, respectively, for the estimated benchmark model. The red line shows the post-intervention distributions. The central finding is that no reaction to the nominal exchange rate deteriorates macroeconomic performance - albeit only slightly - both in terms of inflation and output volatility. A small positive value for the coefficient around 0.1 seems to be optimal. As the coefficient grows too large, macroeconomic performance worsens. Note that the HPDI80 of $\psi_{\Delta e}$ corresponds closely to the region with the lowest volatilities of output and inflation. That is, the Swiss National Bank did react optimally to the exchange rate movements in that respect.

6 Conclusion

The main contribution of this paper is to assess the question of optimal monetary policy in the open economy. Specifically, we ask whether it is optimal for the central bank to react to movements in the nominal exchange rate when macroeconomic performance is evaluated by means of inflation and output variability. We estimate a structural model that is suitable for addressing this question for Swiss data. In addition to only estimating the model, we use the approach proposed by DelNegro and Schorfheide (2004) to account for possible misspecification of the underlying model. Under some assumptions, this approach allows us to perform some counterfactual policy experiments in which the central bank influences the policy parameters.
The key findings of this study is that the estimated coefficient on the exchange rate reaction in the policy rule is non-zero. Moreover, the posterior odds ratio between the basic version of the model and one in which the exchange rate coefficient is restricted to be zero clearly favors the the former. Our policy experiments point towards the same direction. Macroeconomic performance, measured both by the variability of inflation and output, deteriorates for a zero coefficient. The exercises suggests that a small, but positive reaction coefficient is optimal for the Swiss monetary policy conduct.
We conclude by pointing out some issues that are left open for future research. The robustness of our findings can be assessed along various aspects. The restrictions imposed by the DSGE model may be more questionable for the foreign model economy. Therefore, one could alternatively specify a non-structural prior for the foreign block. Another topic that potentially adds further insight could be a different specification of the Taylor-type rule. Specifically, a measure of output gap rather than output could be included which would require the solution of the model for the flexible price equilibrium. Furthermore, we could use alternative data: Prices could be measured by the Consumer Price Index instead of the GDP deflator and the foreign block could be constructed using trade-weighted averages, for example. Finally, our findings could be assessed in the light of a different model such as the standard cash-in-advance model.

References

Adolfson, M., S. Laseen, J. Linde, and M. Villani (2007): "Bayesian Estimation of an Open Economy DSGE Model with Incomplete Pass-Through," *Journal of International Economics*, 72(2), 481–511.

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

Ball, L. (1999): "Policy Rules for Open Economies," in *Monetary Policy Rules*, ed. by J. B. Taylor, pp. 127–144. University of Chicago Press.

Benigno, P. (2001): "Price Stability with Imperfect Financial Integration," Working Paper 2854, CEPR.

Blanchard, O. J., and C. M. Kahn (1980): "The Solution of Linear Difference Models under Rational Expectations," *Econometrica*, 48(5), 1305–1311.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45.

Collard, F., and H. Dellas (2002): "Monetary Policy Rules and Exchange Rate Targeting," Unpublished Manuscript, University of Bern.

DelNegro, M., and F. Schorfheide (2004): "Priors from General Equilibrium Models for VARs," *International Economic Review*, 45(2), 643–673.

——— (2006): "How Good Is What You've Got? DGSE-VAR as a Toolkit for Evaluating DSGE Models," *Federal Bank of Atlanta Economic Review*, (Q2), 21–37.

Garber, P. M., and L. E. Svensson (1995): "The Operation and Collapse of Fixed Exchange Rate Regimes," in *Handbook of International Economics*, ed. by G. M. Grossman, and K. Rogoff, vol. 3, chap. 36, pp. 1865–1911. Elsevier.

Geweke, J. (1999): "Using simulation methods for bayesian econometric models: inference, development,and communication," *Econometric Reviews*, 18(1), 1–73.

——— (2005): *Contemporary Bayesian Econometrics and Statistics*. John Wiley & Sons, Inc.

Justiniano, A., and B. Preston (2008): "Monetary Policy and Uncertainty in an Empirical Small Open Economy Model," *Journal of Applied Econometrics*, forthcoming.

King, R. G., and M. W. Watson (1998): "The Solution of Singular Linear Difference Systems under Rational Expectations," *International Economic Review*, 39(4), 1015–1025.

Klein, P. (2000): "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model," *Journal of Economic Dynamics & Control*, 24, 1405–1423.

Kollmann, R. (2002): "Monetary Policy Rules in the Open Economy: Effects on Welfare and Business Cycles," *Journal of Monetary Economics*, 49(5), 989–1015.

Lees, K., T. Matheson, and C. Smith (2007): "Open Economy DSGE-VAR Forecasting And Policy Analysis: Head To Head With The RBNZ Published Forecasts," CAMA Working Paper 2007-05, Australian National University, Centre for Applied Macroeconomic Analysis.

Lubik, T., and F. Schorfheide (2007): "Do Central Banks Respond to Exchange Rate Movements? A Structural Investigation," *Journal of Monetary Economics*, 54(4), 1069–1087.

Lubik, T. A., and F. Schorfheide (2004): "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic Review*, 94(1), 190–217.

Lucas, R. (1976): "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy*, (1), 19–46.

Monacelli, T. (2005): "Monetary Policy in a Low Pass-Through Environment," *Journal of Money, Credit, and Banking*, 37(6), 1047–1066.

Obstfeld, M., and K. Rogoff (1995): "The Mirage of Fixed Exchange Rates," *Journal of Economic Perspectives*, 9(4), 73–96.

Schmitt-Grohé, S., and M. Uribe (2003): "Closing Small Open Economy Models," *Journal of International Economics*, 61(1), 163–185.

Sims, C. A. (2002): "Solving Linear Rational Expectations Models," *Computational Economics*, 20, 1–20.

Svensson, L. E. O. (1999): "Inflation Targeting as a Monetary Policy Rule," *Journal of Monetary Economics*, 43(3), 607–654.

——— (2000): "Open-Economy Inflation Targeting," *Journal of International Economics*, 50(1), 155–183.

Taylor, J. B. (1993): "Discretion versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy*, 39(1), 195–214.

——— (1999): "The Robustness and Efficiency of Monetary Policy Rules as Guidelines for Interest Rate Setting by the European Central Bank," *Journal of Monetary Economics*, 43(3), 655–679.

——— (2001): "The Role of the Exchange Rate in Monetary-Policy Rules," *American Economic Review*, 91(2), 263–267.

Uhlig, H. (1998): "A Toolkit for Analysing Nonlinear Dynamic Stochastic Models Easily," QM&RBC Codes, Quantitative Macroeconomics & Real Business Cycles.

A Tables and Graphs

Table 1: Prior distribution for DSGE model parameters

parameter	distribution[a]	mean	standard deviation
β	calibrated	0.977	
χ	calibrated	0.01	
α	Normal	0.43	0.05
σ	Gamma	1.20	0.40
φ	Gamma	1.50	0.75
θ_H	Beta	0.50	0.10
θ_F	Beta	0.50	0.10
η	Gamma	1.50	0.75
h	Beta	0.50	0.25
δ_H	Beta	0.50	0.25
δ_F	Beta	0.50	0.25
ρ_i	Beta	0.50	0.25
ψ_π	Gamma	1.50	0.30
ψ_y	Gamma	0.25	0.13
$\psi_{\Delta e}$	Gamma	0.25	0.13
$\psi_{\Delta y}$	Gamma	0.25	0.13
β^*	calibrated	0.99	
σ^*	Gamma	1.20	0.40
φ^*	Gamma	1.50	0.75
θ^*	Beta	0.50	0.10
h^*	Beta	0.50	0.25
δ^*	Beta	0.50	0.25
ρ_i^*	Beta	0.50	0.25
ψ_π^*	Gamma	1.50	0.30
ψ_y^*	Gamma	0.25	0.13
$\psi_{\Delta y}^*$	Gamma	0.25	0.13
ρ_a	Beta	0.50	0.20
ρ_g	Beta	0.50	0.20
ρ_{cp}	Beta	0.50	0.20
ρ_{rp}	Beta	0.50	0.20
ρ_a^*	Beta	0.50	0.20
ρ_g^*	Beta	0.50	0.20
σ_m	Inv Gamma	0.38	0.20
σ_a	Inv Gamma	0.38	0.20
σ_g	Inv Gamma	0.38	0.20
σ_{cp}	Inv Gamma	0.38	0.20
σ_{rp}	Inv Gamma	0.38	0.20
σ_m^*	Inv Gamma	0.38	0.20
σ_a^*	Inv Gamma	0.38	0.20
σ_g^*	Inv Gamma	0.38	0.20

[a] *Note:* The density function of the Inverse Gamma distribution is of the following form $p(x) \propto x^{-\nu-1}e^{-\nu s^2/2x^2}$. In our specification, we use $s = 0.3$ and $\nu = 4$.

Table 2: 80% highest prior and posterior density intervals of DSGE model parameters

	Prior		Posterior	
	Mean	80% Interval	Mean	80% Interval
α	0.43	[0.37 , 0.50]	0.37	[0.31 , 0.43]
σ	1.2	[0.65 , 1.63]	0.20	[0.13 , 0.27]
φ	1.5	[0.47 , 2.20]	1.13	[0.30 , 1.67]
θ_H	0.5	[0.37 , 0.63]	0.58	[0.48 , 0.70]
θ_F	0.5	[0.37 , 0.63]	0.68	[0.62 , 0.75]
η	1.5	[0.47 , 2.20]	1.26	[0.64 , 1.74]
h	0.5	[0.15 , 0.84]	0.08	[0.00 , 0.12]
δ_H	0.5	[0.15 , 0.84]	0.17	[0.00 , 0.26]
δ_F	0.5	[0.15 , 0.84]	0.13	[0.00 , 0.19]
ρ_i	0.5	[0.15 , 0.84]	0.88	[0.80 , 1.00]
ψ_π	1.5	[1.09 , 1.85]	1.47	[1.11 , 1.78]
ψ_y	0.25	[0.07 , 0.37]	0.11	[0.04 , 0.16]
$\psi_{\Delta e}$	0.25	[0.07 , 0.37]	0.16	[0.08 , 0.22]
$\psi_{\Delta y}$	0.25	[0.07 , 0.37]	0.42	[0.27 , 0.55]
σ^*	1.2	[0.65 , 1.63]	0.73	[0.38 , 0.97]
φ^*	1.5	[0.47 , 2.20]	1.51	[0.47 , 2.23]
θ^*	0.5	[0.37 , 0.63]	0.50	[0.37 , 0.62]
h^*	0.5	[0.20 , 0.88]	0.13	[0.01 , 0.21]
δ^*	0.5	[0.20 , 0.88]	0.25	[0.00 , 0.40]
ρ_i^*	0.5	[0.20 , 0.88]	0.78	[0.71 , 0.87]
ψ_π^*	1.5	[1.09 , 1.85]	1.77	[1.40 , 2.11]
ψ_y^*	0.25	[0.07 , 0.37]	0.08	[0.02 , 0.12]
$\psi_{\Delta y}^*$	0.25	[0.07 , 0.37]	0.42	[0.18 , 0.61]
ρ_a	0.50	[0.23 , 0.77]	0.31	[0.08 , 0.48]
ρ_g	0.50	[0.23 , 0.77]	0.79	[0.72 , 0.88]
ρ_s	0.50	[0.23 , 0.77]	0.71	[0.61 , 0.83]
ρ_{cp}	0.5	[0.23 , 0.77]	0.37	[0.13 , 0.54]
ρ_a^*	0.5	[0.23 , 0.77]	0.81	[0.69 , 0.98]
ρ_g^*	0.5	[0.23 , 0.77]	0.78	[0.71 , 0.87]
σ_i	0.38	[0.17 , 0.48]	0.26	[0.20 , 0.32]
σ_a	0.38	[0.17 , 0.48]	0.49	[0.25 , 0.63]
σ_g	0.38	[0.17 , 0.48]	0.33	[0.21 , 0.41]
σ_s	0.38	[0.17 , 0.48]	0.19	[0.14 , 0.22]
σ_{cp}	0.38	[0.17 , 0.48]	0.21	[0.18 , 0.24]
σ_i^*	0.38	[0.17 , 0.48]	0.45	[0.19 , 0.61]
σ_a^*	0.38	[0.17 , 0.48]	0.21	[0.15 , 0.24]
σ_g^*	0.38	[0.17 , 0.48]	0.48	[0.22 , 0.64]

Table 3: Posterior odds

q	Log marginal data densities $\psi_{\Delta e} > 0$	$\psi_{\Delta e} = 0$	Posterior odds[a]
0.05	184.14	168.16	0.11×10^{-6}
0.10	184.30	166.09	0.01×10^{-6}
0.25	184.44	166.98	0.03×10^{-6}
0.50	184.09	167.62	0.07×10^{-6}
0.75	184.23	167.92	0.08×10^{-6}
0.90	184.26	168.09	0.10×10^{-6}
0.95	184.28	168.14	0.10×10^{-6}

[a] *Note:* Posterior odds of the hypothesis $\psi_{\Delta e} = 0$ vs. $\psi_{\Delta e} > 0$.

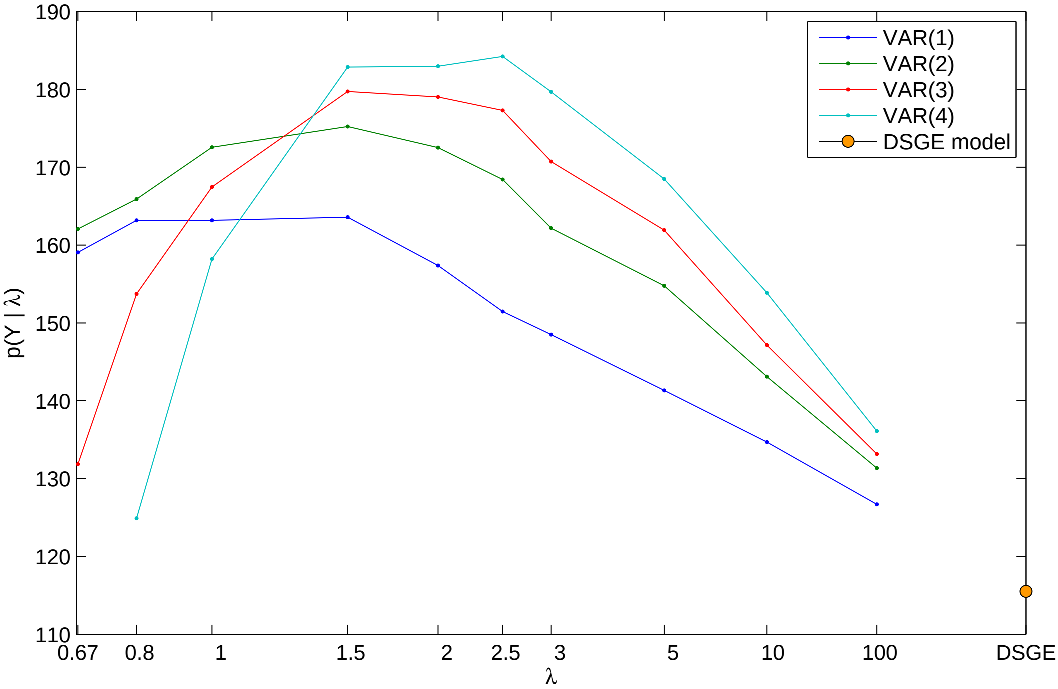

Figure 1: Posterior marginal data densities over a grid for λ

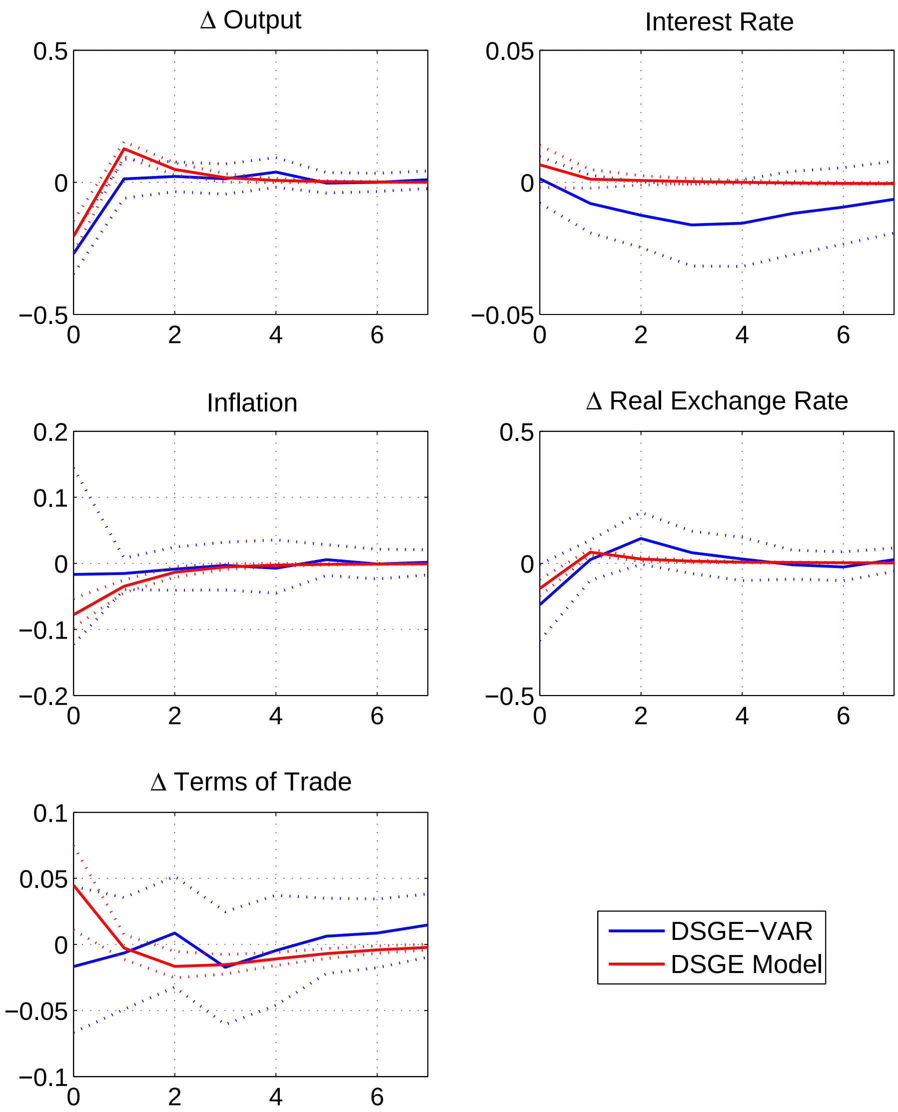

Figure 2: Impulse responses to a contractionary domestic monetary policy shock

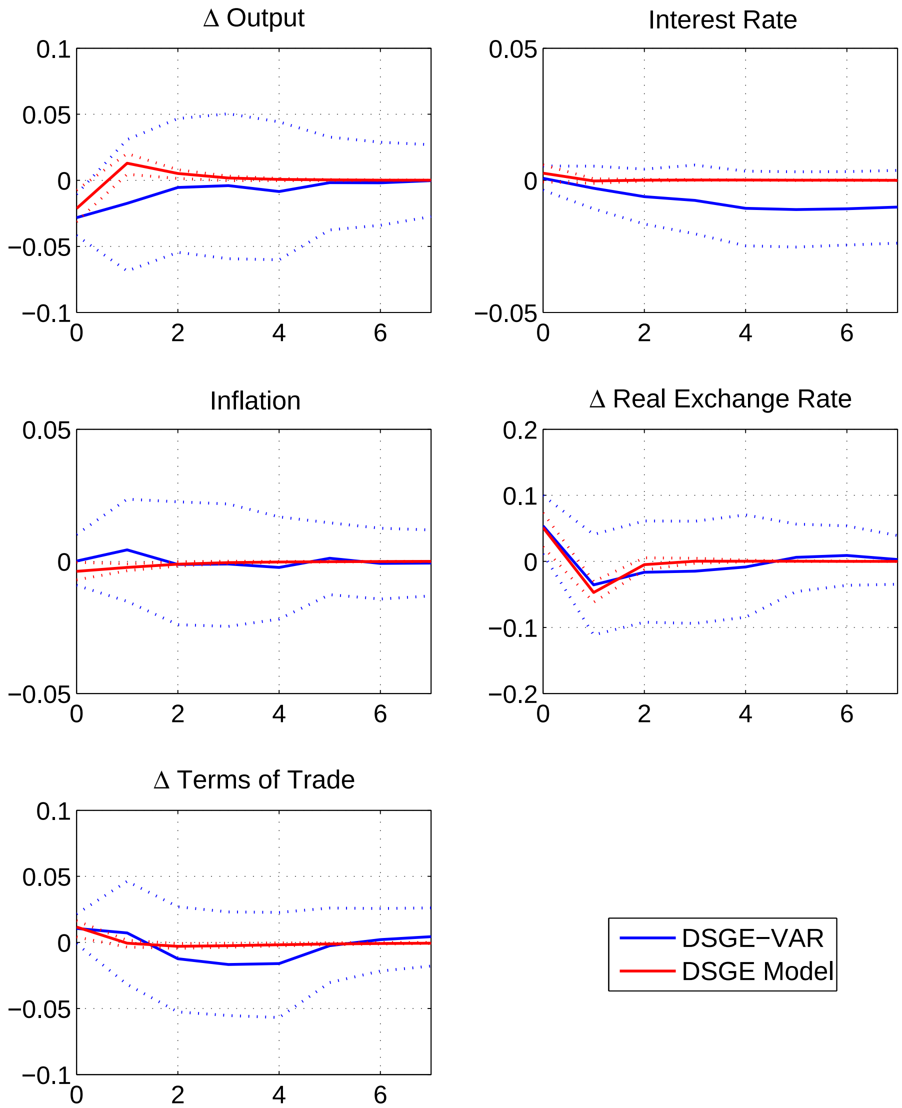

Figure 3: Impulse responses to a contractionary foreign monetary policy shock

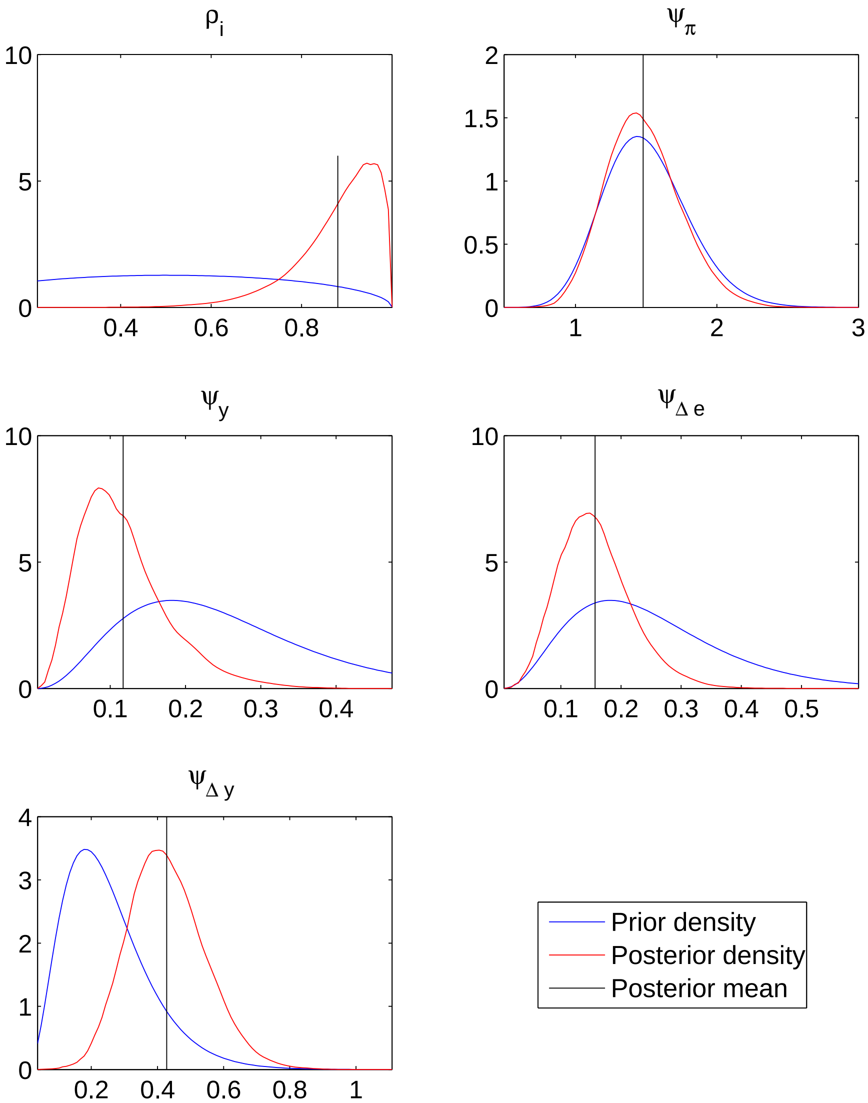

Figure 4: Prior and posterior distributions of the Taylor rule coefficients

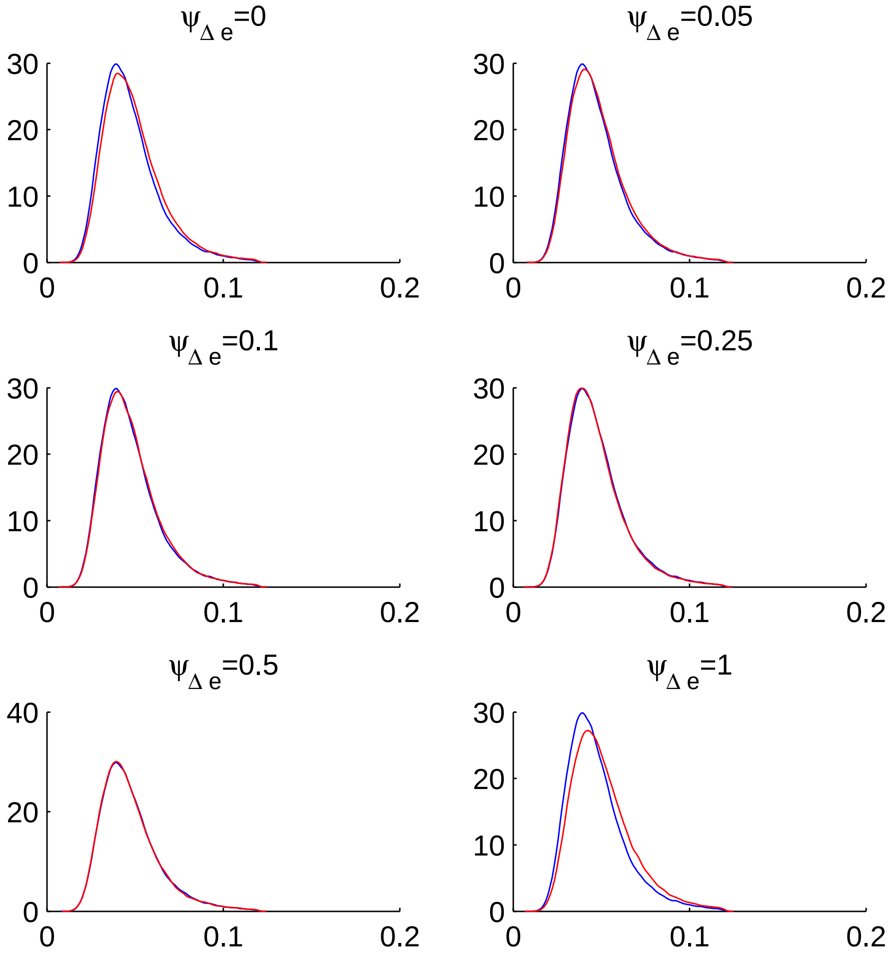

Figure 5: Variance of inflation for different values of $\psi_{\Delta e}$

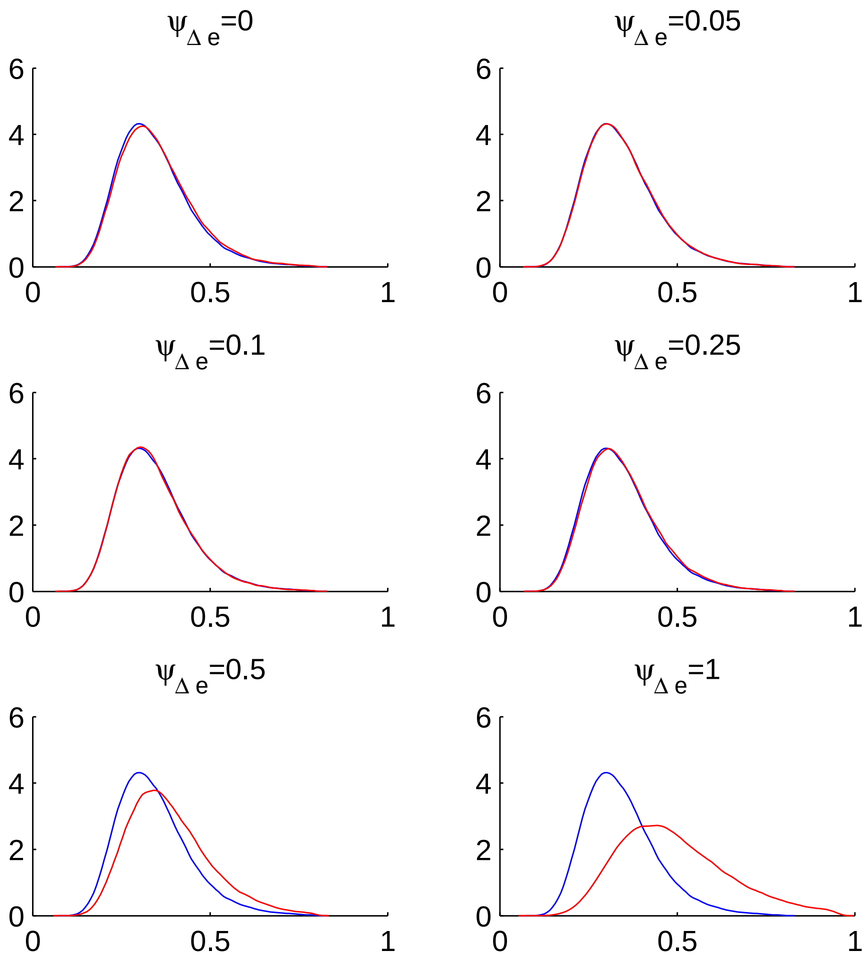

Figure 6: Variance of output for different values of $\psi_{\Delta e}$

B Derivations of the Model Equations

B.1 The Domestic Household

The household problem is solved in two stages. In a first step, we solve for the combination of domestic and foreign goods bundles that minimize costs for any given level of aggregate consumption. Then, given the costs for any level of consumption, the household chooses C_t and N_t so as to maximize lifetime utility.
The cost minimization problem is given by

$$\min_{C_{H,t},C_{F,t}} P_{H,t}C_{H,t} + P_{F,t}C_{F,t} \quad \text{s.t.} \quad C_t = \left[(1-\alpha)^{\frac{1}{\eta}} C_{H,t}^{\frac{\eta-1}{\eta}} + \alpha^{\frac{1}{\eta}} C_{F,t}^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}} \tag{B-1}$$

Attaching the multiplier P_t to the constraint in (B-1) and maximizing with respect to $C_{H,t}$ and $C_{F,t}$ yields the demand functions

$$C_{H,t} = (1-\alpha)\left(\frac{P_{H,t}}{P_t}\right)^{-\eta} C_t \qquad C_{F,t} = \alpha\left(\frac{P_{F,t}}{P_t}\right)^{-\eta} \tag{B-2}$$

When plugging the demand function back into the CES-aggregator we find the theoretically correct consumer price index

$$P_t = \left[(1-\alpha)P_{H,t}^{1-\eta} + \alpha P_{F,t}^{1-\eta}\right]^{\frac{1}{1-\eta}} \tag{B-3}$$

Optimality also requires that the expenditures on all varieties are cost minimizing for any level of the respective consumption bundles. Since we have introduced Calvo-pricing with indexation, the cost minimization complicates a little bit compared to the standard problem. Let $P_{H,t}^1(i)$ and $C_{H,t}^1(i)$ denote the price and quantity demanded of variety i from a firm that can re-optimize its price in period t and $P_{H,t}^2(i)$ and $C_{H,t}^2(i)$ the price and quantity demanded from a firm that cannot re-optimize its price. Then the cost minimization problem of the household for the varieties of domestically produced goods is given by

$$\min_{C_{H,t}^1(i),C_{H,t}^2(i)} \int_0^{\theta_H} P_{H,t}^1(i)C_{H,t}^1(i)di + \int_{\theta_H}^1 P_{H,t}^2(i)C_{H,t}^2(i) \tag{B-4}$$

subject to the following Dixit-Stiglitz aggregator

$$C_{H,t} = \left[\int_0^{\theta_H} C_{H,t}^1(i)^{\frac{\varepsilon-1}{\varepsilon}}di + \int_{\theta_H}^1 C_{H,t}^2(i)^{\frac{\varepsilon-1}{\varepsilon}}di\right]^{\frac{\varepsilon}{\varepsilon-1}} \tag{B-5}$$

Attaching the multiplier $P_{H,t}$ to constraint (B-5) and differentiating the Lagrangian with respect to $C^1_{H,t}(i)$ and $C^2_{H,t}(i)$, the following variety demand functions can be derived

$$C^1_{H,t}(i) = \left(\frac{P^1_{H,t}(i)}{P_{H,t}}\right)^{-\varepsilon} C_{H,t} \tag{B-6}$$

$$C^2_{H,t}(i) = \left(\frac{P^2_{H,t}(i)}{P_{H,t}}\right)^{-\varepsilon} C_{H,t} \tag{B-7}$$

Re-inserting (B-6) and (B-7) into (B-5) yields the theoretically correct price index for the domestic consumption bundle

$$P_{H,t} = \left[\int_0^{\theta_H} P^1_{H,t}(i)^{1-\varepsilon} di + \int_{\theta_H}^1 P^2_{H,t}(i)^{1-\varepsilon} di\right]^{\frac{1}{1-\varepsilon}} \tag{B-8}$$

Since we also impose Calvo-pricing on the retailers that import foreign varieties, we obtain demand functions and a price index for imported goods that are analogous to (B-6), (B-7) and (B-8). What is left to be shown is what prices $P^1_{H,t}(i)$ and $P^2_{H,t}(i)$ are.

B.2 Pricing Decision of the Domestic Firm

Domestic firms face a Calvo-style staggered price setting problem with indexation. That is, with probability θ_H firms cannot re-optimize their price and just follow the indexation rule given by (10). Despite of ex ante heterogeneity of firms we only consider the symmetric equilibrium in which all firms react identically and set the same price. That is, firms that cannot re-optimize set

$$P^1_{H,t} = P_{H,t-1}\left(\frac{P_{H,t-1}}{P_{H,t-2}}\right)^{\delta_H} \tag{B-9}$$

and firms that can re-optimize all set the same optimal price which we denote by $P^2_{H,t} = P'_{H,t}$. In that case, the domestic goods price index (B-8) becomes

$$P_{H,t} = \left[\theta_H \left(P_{H,t-1}\left(\frac{P_{H,t-1}}{P_{H,t-2}}\right)^{\delta_H}\right)^{1-\varepsilon} + (1-\theta_H)\, {P'_{H,t}}^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}} \tag{B-10}$$

The staggered price setting renders the maximization problem of the firms dynamic. We must find an expression for the demand faced by a firm for any period $t+\tau$

when it has last set its price optimally in period t. Any firm sells its output both domestically and abroad. The relevant domestic demand curve for this case is given by (B-6). When assuming that the demand from abroad resembles the domestic demand in a way described by equation (17), the total demand faced by a firm in period $t+\tau$ that has last re-set its price optimally in period t is given by

$$C_{H,t+\tau|t} = \left(\frac{P^1_{H,t+\tau|t}}{P_{H,t+\tau}} \right)^{-\varepsilon} (C_{H,t+\tau} + C^*_{H,t+\tau})$$

Applying the indexation rule (B-9) we can replace $P^1_{H,t+\tau|t}$ and obtain

$$C_{H,t+\tau|t} = \left(\frac{P'_{H,t}}{P_{H,t+\tau}} \left(\frac{P_{H,t+\tau-1}}{P_{H,t-1}} \right)^{\delta_H} \right)^{-\varepsilon} (C_{H,t+\tau} + C^*_{H,t+\tau}) \tag{B-11}$$

A firm that can re-set its price in period t chooses $P'_{H,t}$ such that it maximizes the present discounted value of profits taking into account the probability of not being able to re-set prices in the future:

$$\max_{P'_{H,t}} E_t \sum_{\tau=0}^{\infty} \theta_H^\tau Q_{t,t+\tau} y_{t+\tau|t}(i) \left[P'_{H,t} \left(\frac{P_{H,t+\tau-1}}{P_{H,t-1}} \right)^{\delta_H} - P_{H,t+\tau} MC_{t+\tau} \right]$$

subject to (B-11). The first order condition of the maximization problem can be written as

$$E_t \sum_{\tau=0}^{\infty} \theta_H^\tau Q_{t,t+\tau} y_{t+\tau|t}(i) \left[P'_{H,t} \left(\frac{P_{H,t+\tau-1}}{P_{H,t-1}} \right)^{\delta_H} - \frac{\varepsilon}{\varepsilon - 1} P_{H,t+\tau} MC_{t+\tau} \right] \overset{!}{=} 0 \tag{B-12}$$

The derivation of the optimal price setting condition for the retailers is analogous to the one of the domestic firms.

B.3 Log-Linearizing the Equilibrium Conditions

Euler equation: the log-linearization of (6) is straightforward.

Resource constraint: to linearize (16) we need some preliminary steps. Linearizing the CPI (4) around a zero inflation steady state in which $P = P_H = P_F$ yields

$$p_t = (1-\alpha) p_{H,t} + \alpha p_{F,t} \tag{B-13}$$

The linearized terms of trade are $s_t = p_{F,t} - p_{H,t}$ and the law of one price gap is $\psi_{F,t} = e_t + p_t^* - p_{F,t}$. First differences of the terms of trade imply

$$\Delta s_t = \pi_{F,t} - \pi_{H,t} \tag{B-14}$$

From the definition of the real exchange rate we can infer that it depends on the law of one price gap and the terms of trade according to

$$q_t = e_t + p_t^* - p_t = \psi_{F,t} + (1-\alpha)s_t \tag{B-15}$$

From (B-13) and (B-14) we find that $p_t - p_{H,t} = \alpha s_t$, so that domestic CPI inflation and domestic goods price inflation are related according to

$$\pi_t = \pi_{H,t} + \alpha \Delta s_t \tag{B-16}$$

We now go about linearizing the resource constraint (16). In contrast to the domestic economy, we assume that the the law of one price holds for imports of domestic goods to the foreign economy, i.e. $P_{H,t} = \tilde{e}_t P_{H,t}^* = \tilde{e}_t P_t^*$. Therefore, the resource constraint $Y_t = C_{H,t} + C_{H,t}^*$ can be written as

$$Y_t = (1-\alpha)\left(\frac{P_{H,t}}{P_t}\right)^{-\eta} C_t + \varsigma(\Psi_{F,t} S_t)^{\eta^*} Y_t^*$$

The parameter ς represents the steady state ratio of exports to GDP in the foreign economy. Although due to the large open economy assumption it tends to zero, we still need it for technical reasons, specifically to have a well defined steady state level of C_H^*. Taking a log-linear approximation to the above equation around the zero inflation steady state yields in a first step

$$Y y_t = C_H[\eta(p_t - p_{H,t}) + c_t] + C_H^*[\eta^*(\psi_{F,t} + s_t) + y_t^*]$$

From the steady state version of equation (3) we find that $C_H = (1-\alpha)C$ and $C_F = \alpha C$ where C_F is the steady state level of exports. Assuming balanced trade in the steady state we get $C_H^* = C_F = \alpha C$. Plugging this into the resource constraint we get $Y = (1-\alpha)C + \alpha C = C$. By dividing both sides of the above equation by Y and using equation (B-16) we find

$$y_t = (1-\alpha)c_t + \alpha[(1-\alpha)\eta + \lambda]s_t + \alpha\lambda\psi_{F,t} + \alpha y_t^*$$

When further assuming that the elasticities of substitution across goods coincide in both economies, i.e. $\eta = \eta^*$, we get the linearized resource constraint

$$y_t = (1-\alpha)c_t + \alpha\eta(2-\alpha)s_t + \alpha\eta\psi_{F,t} + \alpha y_t^* \tag{B-17}$$

Optimal price setting for domestic firms: taking a log-linear approximation to (13) and (11) yields a hybrid Phillips curve. Let $\Pi_{H,t}$ denote the gross inflation of domestic goods between $t-1$ and t and $\Pi_{H,t,t+\tau}$ the one between t and $t+\tau$. As a preliminary step, we divide both sides of the domestic price index (11) by $P_{H,t-1}$ and rearrange to obtain

$$\Pi_{H,t}^{1-\varepsilon} = \left[(1-\theta_H)\left(\frac{P'_{H,t}}{P_{H,t-1}}\right)^{1-\varepsilon} + \theta_H \Pi_{H,t-1}^{\delta_H(1-\varepsilon)}\right]$$

So in a zero inflation steady state in which $\Pi = \Pi_H = 1$ it must be that $P'_H = P_H$ and therefore the steady state real marginal cost are $MC = \frac{\varepsilon-1}{\varepsilon}$. We proceed by dividing the firms' optimality condition (13) by $P_{H,t-1}$ (which is in the information set of t) and using the above definition we obtain

$$E_t \sum_{\tau=0}^{\infty} \theta_H^{\tau} Q_{t,t+\tau} C_{H,t+\tau|t} \left[\frac{P'_{H,t}}{P_{H,t-1}} \Pi_{H,t-1,t+\tau-1}^{\delta_H} - \frac{\varepsilon}{\varepsilon-1} \Pi_{H,t-1,t+\tau} MC_{t+\tau}\right] = 0$$

From (12) we can see that in a zero-inflation steady state $y_H(i)$ is well defined and independent of time. Moreover, the steady state stochastic discount factor is just $Q_{t,t+\tau} = \beta^\tau$. Hence, first order Taylor expansion to the above condition around the steady state yields

$$0 = E_t \sum_{\tau=0}^{\infty} (\theta_H\beta)^\tau \left[p'_{H,t} - p_{H,t-1} + \delta_H(p_{H,t+\tau-1} - p_{H,t-1}) - (mc_{t+\tau} + p_{H,t+\tau} - p_{H,t-1})\right]$$

By rearranging we can express the gap between the optimal and last period's price as the sum of expected future marginal costs and a weighted average of domestic goods price inflation and general CPI inflation

$$p'_{H,t} - p_{H,t-1} = (1-\theta_H\beta) E_t \sum_{\tau=0}^{\infty} (\theta_H\beta)^\tau \left[mc_{t+\tau} + \delta_H \pi_{H,t+\tau} + (1-\delta_H)\sum_{k=0}^{\tau} \pi_{H,t+k}\right]$$

After some tedious manipulations the above expression can be written as

$$p'_{H,t} - p_{H,t-1} = E_t \sum_{\tau=0}^{\infty} (\theta_H \beta)^\tau \left[(1-\theta_H\beta) mc_{t+\tau} + (1-\delta_H\theta_H\beta)\pi_{H,t+\tau}\right]$$

or as a first order difference equation

$$p'_{H,t} - p_{H,t-1} = (1-\theta_H\beta)mc_t + (1-\delta_H\theta_H\beta)\pi_{H,t} + (\theta_H\beta)E_t[p'_{H,t+1} - p_{H,t}]$$

By taking a log-linear approximation to the domestic goods price index (11) we find that

$$p_{H,t} = (1-\theta_H)p'_{H,t} + \theta_H((1+\delta_H)p_{H,t-1} - \delta_H p_{H,t-2})$$

Subtracting $p_{H,t-1}$ from both sides and rearranging yields

$$p'_{H,t} - p_{H,t-1} = \frac{\pi_{H,t}}{1-\theta_H} - \frac{\theta_H\delta_H}{1-\theta_H}\pi_{H,t-1}$$

When inserting this into the above difference equation and collecting terms we get the hybrid Phillips curve

$$\pi_{H,t} - \delta_H\pi_{H,t-1} = \frac{(1-\theta_H)(1-\theta_H\beta)}{\theta_H} mc_t + \beta E_t[\pi_{H,t+1} - \delta_H\pi_{H,t}] \quad \text{(B-18)}$$

Marginal costs: To find an expression for the real marginal costs in terms of the other variables we assume that the labor market is in equilibrium and plug the households optimal labor condition into the expression for marginal costs. A log-linear approximation is then

$$mc_t = \varphi n_t + p_t - p_{H,t} - \varepsilon_{a,t} + \sigma(1-h)^{-1}(c_t - hc_{t-1})$$

Substituting for n_t from the production function and using equation (B-16) real marginal cost are given by

$$mc_t = \varphi y_t - (1+\varphi)\varepsilon_{a,t} + \alpha s_t + \sigma(1-h)^{-1}(c_t - hc_{t-1}) \quad \text{(B-19)}$$

Optimal price setting for domestic retailers: Again, we can proceed exactly analogous as when linearizing the optimal price setting condition for domestic firms. Because of zero-inflation in the steady state it must be that the steady state law of one price gap is $\Psi_F = (\varepsilon - 1)/\varepsilon$. A log-linear approximation to the optimality condition (15) then yields the difference equation

$$p'_{F,t} - p_{F,t-1} = (1-\theta_F\beta)\psi_{F,t} + (1-\delta_F\theta_F\beta)\pi_{F,t} + \theta_F\beta E_t[p'_{F,t+1} - p_{F,t}]$$

From the log-linear approximation to the import price index one can derive

$$p'_{F,t} - p_{F,t-1} = \frac{\pi_{F,t}}{1-\theta_F} - \frac{\theta_F\delta_F}{1-\theta_F}\pi_{F,t-1}$$

Combining the last two equations then yields a hybrid Phillips Curve for import price inflation of the form

$$\pi_{F,t} - \delta_F\pi_{F,t-1} = \frac{(1-\theta_F)(1-\theta_F\beta)}{\theta_F}\psi_{F,t} + \beta E_t[\pi_{F,t+1} - \delta_F\pi_{F,t}] \quad \text{(B-20)}$$

Uncovered interest rate parity: When assuming that the function governing the risk-premium is of the form $\phi_{t+1} = \exp(-\chi A_t - \varepsilon_{s,t})$, the log-linearization of (7) is straightforward and directly yields (25).

Budget constraint: Recall that the domestic bond market clearing requires $D_t = 0\ \forall t$. Moreover, we assume that the government transfers T_t are chosen exactly so as to neutralize the distortions stemming from monopolistic competition. It proves convenient to introduce the profits from the final goods producing firm (this does not change anything because it is perfectly competitive and thus makes zero profit): $\Pi_{fg,t} = P_tC_t - P_{F,t}C_{F,t} - P_{H,t}C_{H,t}$. Settling profits with the lump sum transfer yields

$$\Pi_{fg,t} + \Pi_{H,t} + \Pi_{F,t} + T_t = P_tC_t - W_tN_t + \tilde{e}_tP^*_{H,t}C^*_{H,t} - \tilde{e}_tP^*_tC_{F,t}$$

Plugging this into the budget constraint yields

$$\tilde{e}_tB_t = \tilde{e}_tB_{t-1}(1+i^*_{t-1})\phi_t + \tilde{e}_tP^*_{H,t}C^*_{H,t} - \tilde{e}_tP^*_tC_{F,t}$$

Since domestic goods are sold at the law of one price abroad we have $P_{H,t} = \tilde{e}_tP^*_{H,t}$. When dividing the whole equation by P_tY, making use of the definition $A_t = \tilde{e}_tB_t/(P_tY)$ as well as the demand functions (3) and (17), and applying the definition of the real exchange rate we obtain

$$A_t = A_{t-1}\frac{\tilde{q}_t}{\tilde{q}_{t-1}}\frac{(1+i^*_{t-1})}{\Pi^*_t}\phi_t + \frac{1}{Y}\left(\left(\frac{P_{H,t}}{P_t}\right)\varsigma\left(\Psi_{F,t}S_t\right)^{\eta^*}Y^*_t - \tilde{q}_t\alpha\left(\frac{P_{F,t}}{P_t}\right)^{-\eta}C_t\right)$$

where Π^*_t is the gross foreign inflation rate. Since net-foreign debt can take on negative values, we will linearize it (instead of log-linearizing). All other variables are log-linearized and the conventional notation introduced above is used. Combining the Euler equation and uncovered interest parity condition in the steady state yields $1/\beta = (1 - i^*)/\Pi^*$. Because we further assume that net-foreign debt in the steady state is zero, all log-linearized variables of the first term of the right-hand-side in the above equation drop out. We can therefore concentrate on the log-linearization of the second term. The real interest rate in the steady state is equal to unity. Imposing balanced trade we have found in the derivation of the resource constraint that $C_H^* = \varsigma Y^* = \alpha C = \alpha Y$. The log-linear form of the budget constraint can therefore be written as

$$a_t = \tfrac{1}{\beta} a_{t-1} + \alpha \left(p_{H,t} - p_t + \eta(\psi_{F,t} + s_t) + y_t^* - q_t + \eta(p_{F,t} - p_t) - c_t \right)$$

Note that we imposed $\eta^* = \eta$. From the resource constraint we know that

$$y_t^* - c_t = \frac{y_t - c_t}{\alpha} - \eta(2-\alpha)s_t - \eta\psi_{F,t}$$

In addition, since $p_t = (1-\alpha)p_{H,t} + \alpha p_{F,t}$ we can write $p_{H,t} - p_t = -\alpha s_t$ and $p_{F,t} - p_t = (1-\alpha)s_t$. Recalling that the real exchange rate can be expressed as $q_t = \psi_{F,t} + (1-\alpha)s_t$, we can substitute this into the budget constraint and obtain

$$a_t = \tfrac{1}{\beta} a_{t-1} + \alpha \left(\frac{y_t - c_t}{\alpha} + \eta s_t - \psi_{F,t} - s_t + \eta(1-\alpha)s_t - \eta(2-\alpha)s_t \right)$$

Collecting terms and rearranging yields the log-linearized budget constraint (26).