Applying Modern Econometric Methods to Swiss Unemployment and Consumption

Applying Modern Econometric Methods to Swiss Unemployment and Consumption

Inaugural dissertation submitted by Martin Straub in fulfillment of the requirements for the degree of

DOCTOR RERUM OECONOMICARUM

at the Faculty of Business, Economics and Social Sciences of the University of Bern

Submitted by
MARTIN STRAUB
from Egnach TG

2012

The faculty accepted this work as dissertation on April 26, 2012 at the request of the two advisors Prof. Dr. Klaus Neusser and Prof. Dr. Jan-Egbert Sturm, without wishing to take a position on the view presented therein.

Acknowledgements

First of all, I would like to thank my supervisor Prof. Klaus Neusser for his advice and support. His guidance made this thesis possible. Moreover, his lectures have increased my interest in macroeconomics and time series econometrics in the first place.

Furthermore, I would like to thank Prof. Jan-Egbert Sturm, head of KOF Swiss Economic Institute. This thesis was written while I was researcher at KOF. I am thankfull for this opportunity and for the valuable feedback Prof. Sturm gave me on earlier versions of the papers.

Special thanks goes to Rolf Schenker for the joint work on regional unemployment. Moreover, I would like to thank him, Andrea Schnell, Simone Elmer, David Iselin and Anne Stücker for interesting discussions and tea breaks. The ideas and the motivation gained from these discussions were a central ingredient for the successfull conclusion of my thesis.

I am also very thankful to my partner Sabine and to my family for their support. Without them, this thesis would not have been possible.

To Sabine

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Unemployment and consumption are central variables of macroeconomics. Unemployment is a key indicator for the utilization of production factors and inflationary pressure in macro-models and beyond that a variable of major public interest. Until 20 years ago, empirical research on unemployment in Switzerland was difficult because the official unemployment rates permanently lay close to zero. In the meantime, however, the literature has grown considerably. Private Consumption on the other hand is the largest component of Swiss GDP and has delivered growth contributions of 0.9 percentage points on average to the annual real GDP between 1980 and 2010. During the recession caused by the financial crisis, Swiss consumption proved robust and prevented GDP from dropping even deeper. Although private consumption causes less fluctuations in overall GDP than investment, an appropriate modeling of consumption is an important ingredient of macroeconomic analysis.

This thesis presents the application of two modern econometric methods to Swiss data. The first paper estimates a spatial model for the cantonal unemployment rates, the second paper examines the modeling of private consumption in a small DSGE model estimated using Bayesian Methods. The methods used in this thesis require computational effort which would have been very time consuming until some years ago. Nowadays, the inversion of numerous large-scale matrices which is required for the estimation of the spatial model for unemployment takes computers only a few minutes. On the other hand, the Bayesian estimation of the model used in the chapter on consumption by Monte Carlo simulations still takes between several hours and one day, depending on the number of generated draws.

The first paper of the thesis is joint work with Rolf Schenker. In this paper, we apply a spatial time series model to the Swiss cantonal unemployment rates. We start by setting up a non-spatial model and test whether the variables explaining unemployment in this standard setting retain their explanatory power once spatial elements are added. For the spatial model, we use the framework proposed by Cliff and Ord (1973, 1981)[1,2]. Under this setting, the unemployment rate in a specific canton depends not only on exogenous explanatory variables, but also on the unemplyoment rates in the neighboring cantons as well as on the shocks hitting the unemployment rates in the neighboring cantons. Our model thus contains a so-called spatial lag both in the level and in the error term. The intuition behind this modeling strategy is that unemployment rates in neighboring labor markets tend to be equalized as unemployed workers also look for jobs in nearby regions. Moreover, shocks such as the closing of a firm affect unemployment not only in the respective location, but also in adjacent areas. In order to estimate the spatial model, we use the estimation technique proposed by Kelejian and Prucha (1998)[3]. Alhough Kelejian and Prucha originally proposed their setting for cross-sectional data, it can easily be extended to a time series framework. In our estimation results, both spatial lags turn out to be significant. Furthermore, the spatial dependency seems to be stronger in the shocks than in the levels. We contribute to the literature by explicitly modeling spatial spillovers in regional labor markets. Although the existence of such spillovers is generally accepted, only few studies include spatial elements in their estimations. Moreover, we estimate our model using quarterly data while most existing studies analyze regional unemployment in Switzerland on the basis of annual data. Finally, the Cliff-Ord-framework has, to our knowledge, not been used for the Swiss labor market so far.

The second paper investigates the modeling of Swiss private consumption in a small DSGE model. With habit formation and rule-of-thumb consumers, I examine two extensions of the standard new Keynesian model with respect to consumption. Both extensions are motivated by empirical deviations from the permanent income hypothesis. Habit formation affects the speed by which consumers react to shocks and thereby the smoothness of the intertemporal consumption path. The permanent income hypothesis predits that consumers re-optimize their optimal intertemporal consumption path in each period. Accordingly, shocks such as an unexpected change in the interest rate cause immediate adjustments of consumption. In contrast, empirical impulse response functions show gradual responses

[1] Cliff, A. D., and Ord, J. K. *Spatial Autocorrelation.* Monographs in Spatial and Environmental Systems Analysis, 5. Pion, London, 1973.

[2] Cliff, A. D., and Ord, J. K. *Spatial Processes: Models & Applications.* Pion, London, 1981.

[3] Kelejian, H. H., and Prucha, I. R. Generalized Spatial Two-stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. *Journal of Real Estate Finance and Economics 17*, 1 (1998), 99–121.

of consumption to shocks: consumers successively converge to new consumption levels. By allowing for habit formation, these gradual adjustments can be reproduced in theoretical models. Under habit formation, consumers try to smooth not only the level of consumption over time, but also the change in consumption. As a consequence, they react gradually to shocks. The concept of rule-of-thumb consumers is mainly used to explain the sensitivity of consumption to changes in current income, an empirical finding wich is otherwise inconsistent with the permanent income hypothesis. Rule-of-thumb consumers are frequently used in the literature on fiscal multipliers in order to reproduce the empirical impulse response functions of consumption with respect to government spending shocks. In neoclassical and new Keynesian models, expansionary fiscal policy has by default no positive effect on consumption as consumers anticipate the government to eventually finance the expenses by raising taxes. Empirical analyses, on the other hand, suggest that consumption reacts positively to increased government spending. Adding rule-of-thumb consumers can reproduce this result in theoretical models under certain conditions. In my study, I subsequently introduce habit formation and rule-of-thumb consumers in a small DSGE model. The model is a simplified version of the small open economy model presented in Galì and Monacelli (2005)[4]. The goal of the paper is twofold: first, I estimate the degree of habit persistence and the share of rule-of-thumb consumers among all consumers. Second, I investigate the effect of habit formation and rule-of-thumb consumers on the likelihood and posterior probability of the model. I estimate the various model versions using Bayesian methods on Swiss data. According to my estimation results, the data support the concept of habit formation. With respect to rule-of-thumb consumers, the results depend on the imposed labor market structure. While the concept is supported under perfectly competitive labor markets, the results are less clear-cut and not completely robust under imperfectly competitive labor markets. Although several studies on estimated DSGE models for Switzerland have been published, there is so far, to my knowledge, no systematic analysis of the usefulness of habit formation or rule-of-thumb consumers. Moreover, I use consumption data in the estimation process, while the existing publications only use GDP with respect to national accounts.

---

[4] Galí, J., and Monacelli, T. Monetary policy and exchange rate volatility in a small open economy. *Review of Economic Studies 72*, 3 (2005), 707–734.

Chapter 2

Spatio-Temporal Dynamics in Swiss Regional Unemployment[1]

Contents

[1]This chapter is joint work with Rolf Schenker.

2.1 Introduction

Until 20 years ago, the Swiss labor market was characterized by very low unemployment rates. Up to the 1980s, Switzerland experienced a permanent state of nearly full employment, but unemployment rates began to rise in the period of economic contraction of 1991/92. After a long time of low variation both over time and across cantons, unemployment rates began to fluctuate as well as to differ between cantons (see, e.g., Feld and Savioz, 2000). This is why the number of studies about this topic has increased notably since 1990.

By now, the body of literature on regional unemployment in Switzerland is quite large (see, e.g., Flückiger et al., 2007a, Steffen, 2005, Parnisari, 2003) and it is a common finding of most studies that the French and Italian speaking cantons of Switzerland are confronted with higher unemployment rates than the German speaking parts (see, e.g., Filippini and Rossi, 1992). However, only a few studies include spatial elements to their estimations, although it is generally accepted that regional labor markets are characterized by strong interdependencies. In this study, we estimate a spatial time series model for the Swiss cantonal unemployment rates on a quarterly level. We determine the variables which explain the levels and fluctuations in the regional unemployment rates and investigate whether these variables retain their explanatory power once spatial elements are added to the model.

In order to incorporate spatial elements into our model, we use the spatial framework proposed by Cliff and Ord (1973, 1981). In this framework, spatial interaction is modeled such that the dependent variable in one regional entity is influenced by a weighted average of the dependent variables in its neighboring entities. This weighted average is constructed using a weighting matrix W which represents the distances between the entities. The analog dependency can be introduced for the model's disturbances. Such models are referred to as spatial autoregressive models and spatial autoregressive error models, respectively. Using both a spatial lag in the dependent variable and in the disturbances, we model the regional unemployment rates in Switzerland between 1998 and 2007. We use the procedure proposed by Kelejian and Prucha (1998) to estimate the model. We find the coefficients of both spatial lags to be significant, with the coefficient of the spatial error lag being higher than the coefficient of the spatial lag in the dependent variable. Moreover, we find the unemployment rate to be increased by the population share of women, the population density, the population share of people aged between 20 and 24 as well as between 25 and 64, the population share of cross-border commuters and the employment share of the third sector and of modern industries. By contrast, we find decreasing effects on the unemployment rate for the employment share in the public sector. In addition to the mentioned variables, we include national GDP and national wages in the model, allowing for cross-section-specific elasticities.

We contribute to the existing literature in two ways: Firstly, we set up a time series model for the cantonal unemployment rates on a quarterly basis. This allows us to analyze the quarterly regional unemployment rates while most existing studies use annual data or mid-term averages. Secondly, we adapt the Cliff-Ord-framework to the Swiss cantonal unemployment rates. To our knowledge, this framework has not yet been used for the Swiss labor market.

The remainder of the paper is structured as follows: We begin by giving an overview of the existing literature on regional unemployment in Switzerland. Next, we outline the data we use, the structure of our model and the estimation technique. Finally, we discuss the estimation results and summarize our findings in the conclusions.

2.2 Unemployment Patterns in Switzerland

Until the end of the eighties of the last century, unemployment has not been a very interesting phenomenon in Switzerland, as Feld and Savioz (2000) note. In the seventies, the Swiss labor market was basically characterized by full employment. Even in the severe recession following the first oil price shock, unemployment did not raise notably. This was only possible because the foreign work force was considerably reduced in this period. Even in the eighties, unemployment remained quite low, compared to other European countries.[2]

This situation changed notably in the nineties, when the unemployment rate increased from less than 1% in 1990 to more than 5% in 1997. Together with the strong increase of the national unemployment rate, the disparity of unemployment between the cantons rose considerably. In 1997, the canton of Appenzell Inner Rhodes showed an unemployment rate of 1.9%, while Geneva faced 7.8%. This is quite remarkable given the small size of the country (see Feld and Savioz, 2000).

From 1997 to 2000 GDP returned to its growth path and unemployment decreased under 2% again. From 2001 to 2003 Switzerland suffered a period of recession and stagnation that led to an increase of the unemployment rate to more than 4%. From 2003 on, the inverse

[2] The official unemployment rates tend to underestimate the true unemployment, as in this time the unemployment insurance only covered small parts of the economy.

relationship between GDP growth and the evolution of unemployment seemed to become weaker. A notable GDP growth does not necessarily lead to a decrease of unemployment any more. The strong GDP growth (especially in 2005–2007/8) was only followed by a moderate reduction in the unemployment rate. In the same period, Switzerland and the European Union agreed on the free movement of persons. For a discussion of the effects of this agreement on unemployment, see Stalder (2008) and KOF Swiss Economic Institute (2008).

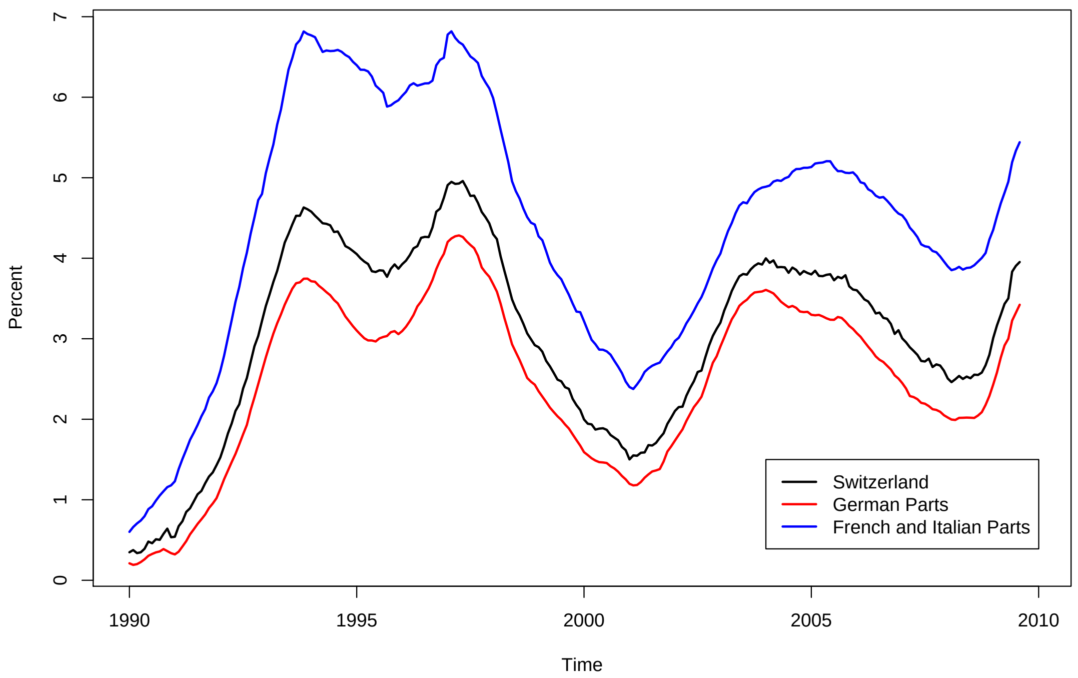
Many studies (see, e.g., Filippini and Rossi, 1992, 1993, Filippini, 1998, Flückiger and Morales, 1998, De Coulon, 1999, Feld and Savioz, 2000, Kleinewefers Lehner, 2001, Flückiger and Vassiliev, 2002, Flückiger et al., 2002, Parnisari, 2003, Steffen, 2005, Flückiger et al., 2007b,a, Brügger et al., 2007) have shown that unemployment is not equally distributed over the Swiss cantons. As can be seen in Figure 2.1, the unemployment rate in the French and Italian speaking cantons has been permanently higher than in the German cantons since 1990. In some periods, the unemployment rates of Latin cantons are even twice as high as in the German cantons.

At first glance, one might think the differences in the unemployment rates to be higher in periods of high unemployment (like around 1995). However, there are different possibilities to measure the disparity of unemployment rates which yield different results. The most commonly used measure is the Gini coefficient (see, e.g., De Coulon, 1999, Parnisari, 2003, Flückiger et al., 2007a). It measures the concentration of unemployment across the cantons. The higher the coefficient, the stronger is the disparity of unemployment in the cantons. The Gini coefficient is plotted together with the Swiss unemployment rate in Figure 2.2. In periods of high unemployment, the Gini coefficient shows smaller values than in periods of low Swiss unemployment. This means that there is a clear tendency towards an equal distribution in periods of high unemployment.[3]

In contrast, Filippini and Rossi (1992, 1993) observe a rising disparity in periods of high unemployment. These studies analyze the standard deviation of the cantonal unemployment instead of the Gini coefficient. There are two problems with this approach: Firstly, Filippini and Rossi (1992) just compare the values for 1976, 1984 and 1991. Secondly, the standard deviation highly depends on the level. The Gini coefficient also depends on the level, but not as heavily as the standard-deviation: If all data values are doubled, the Gini coefficient remains unchanged while the standard-deviation is doubled as well. When looking at the

[3]The same can be shown using the Theil index instead of the Gini coefficient.

Figure 2.1: *Swiss Unemployment Rate, Total, German- and French/Italian-Speaking Part.*



NOTE: The unemployment rate in the French and Italian speaking cantons has been permanently higher than in the German speaking cantons since 1990.

coefficient of variation, the mean-corrected standard-deviation, (see Figure 2.3) we see that the variation is smaller in times of high unemployment rates. This is perfectly in line with what we found analyzing the Gini coefficient.

Besides the existence and extent of regional disparities, also the determinants of the cantons' individual levels of unemployment have been analyzed. Starting from the works of Lewin (1983), Filippini and Rossi (1992, 1993) and Projer (1993) many studies about Swiss regional unemployment rates have been published in the last 10 years. While some of them (see, e.g., De Coulon, 1999, Feld and Savioz, 2000) examine regional Beveridge Curves, others look at in- and outflows of unemployment rather than unemployment rates (like Flückiger et al., 2002, Flückiger and Vassiliev, 2002) and estimate duration models.

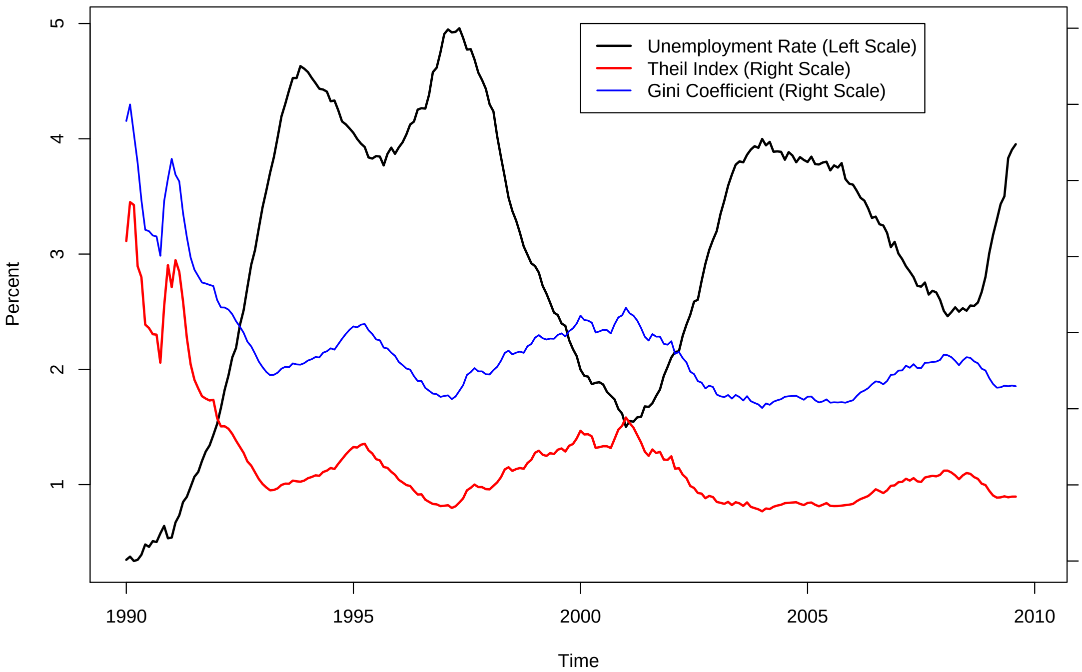
De Coulon (1999) and Flückiger et al. (2002) calculate Beveridge Curves for the Swiss Cantons. Furthermore, De Coulon (1999) analyzes the link between foreign population and regional unemployment and shows that the regional differences in unemployment in Switzerland are higher than in other European countries. The share of foreign residents can partly explain these differences. While, in his estimations, the number of seasonal workers, cross-border commuters and persons with an annual permit are not able to explain the differences, it's the share of foreigners with a permanent residence permit that helps to explain the regional differences.

Flückiger et al. (2002), Flückiger and Vassiliev (2002) focus on the differences between the Canton of Geneva and the other cantons. They divide the unemployment into structural, frictional and cyclical unemployment. The analysis of micro data enables them to calculate in- and outflows of unemployment. They hence find the smaller outflow from unemployment to be one of the reasons for the higher unemployment rates in Geneva. By estimating duration models, they find the duration of unemployment to be higher in Geneva than in other cantons.

Feld and Savioz (2000) estimate a dynamic panel model. They use a model similar to De Coulon (1999), but add different categories of explanatory variables to it. They criticize that variables related to the cantonal economic policy or the skills of the active population have not been analyzed in Swiss studies. In their estimations, they find high tax burdens and the number of foreign workers to increase cantonal unemployment, while human capital reduces unemployment.

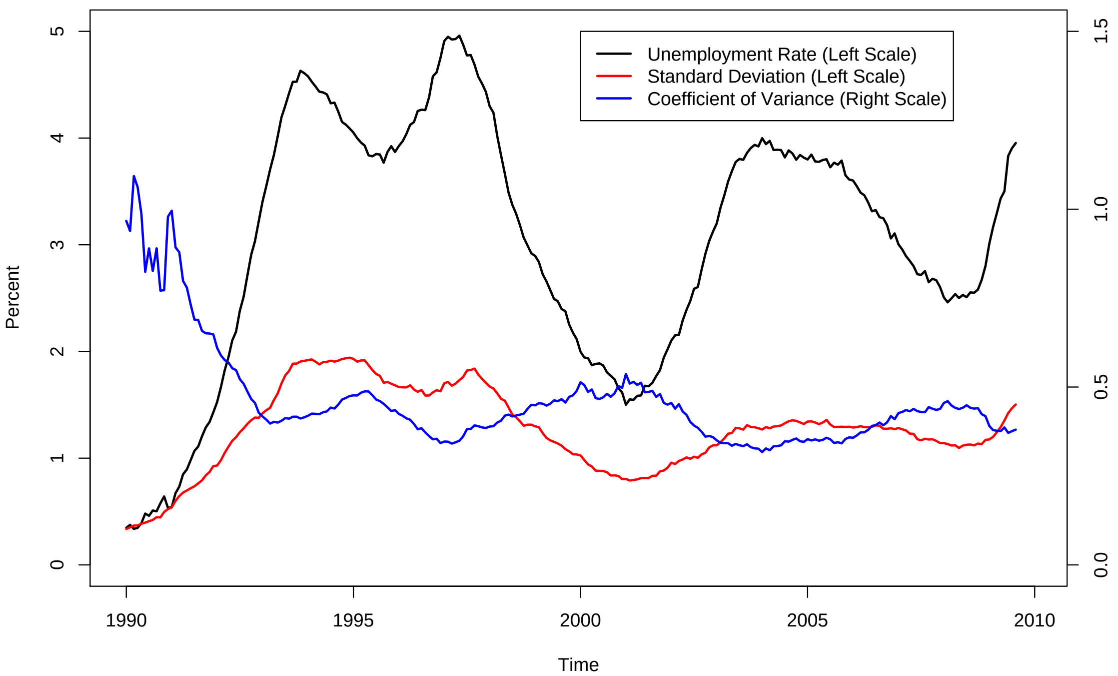
Parnisari (2003) examines the dispersion of unemployment over time. He calculates the Gini coefficient over time and finds the cantonal disparities to augment in times of booms. For the 1980–1990 period, this behavior of cantonal differences is explained by a strong

Figure 2.2: *Swiss Unemployment Rate, Gini Coefficient and Theil Index.*



NOTE: In periods of high unemployment, the Gini coefficient shows smaller values than in periods of low unemployment. This means that there is a clear tendency to an equal distribution in periods of high unemployment.

Figure 2.3: *Swiss Unemployment Rate, Standard Deviation, and Coefficient of Variation.*



NOTE: The variation is smaller in times of high unemployment. This is perfectly in line with what we found analyzing the Gini coefficient.

cyclical component and different structural components in the different cantons. For the second phase, 1990–2002, the reason given is the rise of structural unemployment in large parts of the cantons.

Steffen (2005) focuses more on the institutional settings for explaining the differences in the cantonal unemployment rates. She shows that cyclical variables can explain the national level of unemployment, but are not able to explain the different values in the cantons. Political-institutional variables on the other hand can explain the different responses of the cantons to the cyclical framework. Therefore, she concludes that the macroeconomic framework does not have a direct influence on unemployment.

Brügger et al. (2007) examine how unemployment rates behave at borders. They analyze language borders as well as country borders and are thus able to distinguish institutional from cultural differences. Using micro data they are able to estimate entering and quitting probabilities of unemployment. They find differences in the unemployment rate at two types of borders: At the Swiss language borders, the differences are due to disparities for in- and outflow of unemployed people, while at the national border between Switzerland and Austria, differences occur in the inflows only.

Flückiger et al. (2007a, the extended version of Flückiger et al., 2007b) is a comprehensive study discussing a wide range of approaches and combining macro- and microeconomic approaches. Similar to Parnisari (2003) the authors calculate the Gini coefficient and find cantonal differences to widen in economic boom phases. By calculating in- and outflows as well as duration models, they discover that younger people have a higher probability of becoming unemployed, but their duration in unemployment is shorter, while older people's probability is smaller, but the average duration is longer.

In this article, we assume the unemployment rates of the cantons to be spatially (and temporally) correlated. We investigate which variables determine the cantonal levels of unemployment and whether they also help to explain the differences in the spatial model.

2.3 The Data

2.3.1 The Dependent Variable

We analyze the cantonal unemployment rates. However, cantonal borders are not necessarily congruent with the borders of regional labor markets. Cantons as Basel-Stadt and Basel-Land are probably one labor market, as many people working in Basel-Stadt live in Basel-Land. A drawback of this rather politically motivated partitioning is the fact that cantons are often heterogenic entities. The economic, geographic and demographic characteristics can vary notably inside a canton. By analyzing cantons as entities, these variations get lost, leading to a distortion of the estimation results as actually separated labor markets are merged.[4] An important reason in favor of analyzing cantonal data is the fact that the Swiss cantons have relatively large competences in economic and job market policy and thus, the political and institutional environment mainly depends on the cantons. Accordingly, regional data is available for the cantons (the NUTS-3 regions) or groups of cantons ("statistical regions", the NUTS-2 regions). The statistical regions do not have any political autonomy; therefore the analysis of cantonal data is the best choice.

As unemployment rates show strong seasonal characteristics we use the seasonally adjusted series.[5]

2.3.2 The Independent Variables

The explanatory variables used in the relevant literature can be parted in four groups. The first group covers information about the population structure; the second group contains variables on the labor force; the third group covers institutional and structural information about the regional entities; the fourth group contains macroeconomic variables.

For the population structure, we include the population density (Feld and Savioz, 2000, Steffen, 2005), variables on the age structure (Elhorst, 2003, Feld and Savioz, 2000) and the share of women (Flückiger et al., 2007a, Filippini, 1998) and foreigners in the population in our model. For the foreigners, different variables have been proposed in the literature, namely the number of persons staying for one year, seasonal workers, foreign resident population and cross-border commuters. All of the mentioned studies use at least one of these categories of foreign workers. Data is available for all these categories, but as we need regional time series, we use the share of cross-border commuters in the working population, as proposed by Flückiger et al. (2007b), Parnisari (2003), Feld and Savioz (2000), Steffen (2005), De Coulon (1999), Flückiger and Vassiliev (2002).

From the second group which contains information about the labor force, we cannot include any variables due to data availability. Variables used in the literature include the share of unemployed persons that are in so called labor market procedures ("arbeitsmarktlichen Massnahmen", Flückiger et al., 2007b), the share of unemployed people that are

[4]For a deeper discussion of this, see Spiezia (2003).

[5]We use an additive X12 procedure with the seasonal filter option X11.

under sanctions of the unemployment insurance (Flückiger et al., 2007b, Steffen, 2005), the share of unemployed people that are registered (Flückiger et al., 2007b, Parnisari, 2003, De Coulon, 1999), labor market participation (Parnisari, 2003, Elhorst, 2003, Steffen, 2005, De Coulon, 1999, Flückiger and Vassiliev, 2002), the share of part time workers (Steffen, 2005, Flückiger and Vassiliev, 2002) or the share of temporary workers (Parnisari, 2003). All these variables are not available as regional time series and thus not included in this analysis.

Institutional and structural information about the cantons is contained in the third group of variables. We include in this context the employment shares of the three sectors (Steffen, 2005), of the public sector (Steffen, 2005) and of the traditional, modern and high tech industries as well as the dispersion of employment over industries (Elhorst, 2003, Steffen, 2005, Parnisari, 2003, Filippini, 1998). Moreover, we considered the overall tax burden (Feld and Savioz, 2000, Steffen, 2005) but abstained from including it as it did not show significant explanatory power. The literature has furthermore proposed the strength of the unions (Elhorst, 2003, Steffen, 2005) or variables on the regional administration (public earnings, total expenditures, expenditures for education, expenditures for interests). As there are no sufficiently long and disaggregated time series available for these variables, they are not included in our estimations.

With regard to macroeconomic variables, we use GDP and wages. GDP is an important determinant of unemployment as it measures the overall economy's need for labor. But as regional time series of economic prosperity or income which are usually proposed in the literature (Elhorst, 2003, Feld and Savioz, 2000, Steffen, 2005, Filippini, 1998) are too short, we use the national GDP instead. The wage level represents the price of labor and has been used in many studies (Elhorst, 2003, Steffen, 2005, Filippini, 1998). As there are no cantonal time series on the evolution of wages, we use national time series as well. In turn, we allow the cantons to have individual elasticities to GDP and wages.

2.3.3 The Weighting Matrix

The weighting matrix W specifies the structure and intensity of the spatial effects. For a set of N regions, it is an $N \times N$ matrix whose diagonal elements are set to zero. Hence, the element w_{ij} represents the intensity of effects between two regions i and j (see, e.g., Anselin and Bera, 1998). The literature knows different approaches for specifying these so called spatial weights. The most frequently used weight specifications are the binary and the distance decay weights.

In the binary weighting matrix, $w_{ij} = 1$ if the regions i and j have a common border, and $w_{ij} = 0$ otherwise (see, e.g., Schanne et al., 2008, Kosfeld and Dreger, 2006). The distance decay function is based on the distance between the centers of the regions and takes the inverse or applies a negative exponential function to it (see, e.g., Schanne et al., 2008, Brügger et al., 2007). Simulation studies by Florax and de Graaff (2004), have shown that a combination of these specifications can be promising: When two regions, i and j have a common border, w_{ij} is the inverse or a negative exponential function of the distance between their capitals, otherwise, $w_{ij} = 0$ (see, e.g., Büttner, 1999, Longhi and Nijkamp, 2007).

However, all these metrics are unable to represent the complex geographic structure of the Swiss cantons. For modeling the relatedness of Swiss cantons and their labor markets, we need more than pure geographical information. Two cantons may be close to each other, but because of geographical obstacles – such as mountains or lakes – their labor markets may be quite separated. This is why the binary and the distance decay specification are not suitable in modeling Swiss cantonal data.

Instead, we use the travel times between the canton's capitals by public transport to construct the W-matrix. We therefore set

$$w^*_{ij} = \frac{1}{tt_{ij}}$$

where tt_{ij} is the travel time from region i's capital to region j's capital. One could argue that the travel times can only be calculated for these specific cities and not for the cantons themselves. Regional capitals can be very far from each other, but due to a large common border, the interaction between the cantons can be quite intensive. We therefore also construct an alternative weighting matrix using all regional capitals and all cities with more than 30'000 inhabitants. The travel times between the cantons is then calculated as weighted means of the travel times of all included cities of the cantons. This alternative W-matrix is then used to check the robustness of our estimation results. The resulting estimation results are very similar to those resulting from the original W-matrix.

To facilitate the interpretation and computation of the spatial autocorrelation, the weighting matrix is row-normalized (see, e.g., Kelejian and Robinson, 1993, Anselin and Rey, 1991):

$$w_{ij} = \frac{w^*_{ij}}{\sum_{j=1}^{N} w^*_{ij}}$$

2.4 Modeling Regional Unemployment in a Spatial Framework

The standard test for spatial dependency in the literature is the Moran-I Test. However, it is designed for data without a time dimension. We therefore use Lagrange Multiplier (LM) tests as proposed by Burridge (1980) or Anselin (1988). As we are testing for spatial dependency in the endogenous variable as well as in the error term, we perform two separate LM tests. They find our data to show a strong spatial dependence in the levels as well as in the errors (p-values far below 0.1%). We therefore estimate a model with two spatial components: a spatial lag and a spatial error lag, a so-called SARAR(1,1) model.

The explanatory variables consist mostly of regional time series. Additionally, we use national time series for GDP and wages (which are not available on a regional basis in quarterly frequency) and canton-specific constants. N denotes the number of cross-sections, which is 26 in our case as there are 26 cantons in Switzerland. We thus model the unemployment rate as dependent on its lagged value, the spatial lag and the described set of independent variables.

For each $t \in \mathbf{T}$,

$$\begin{aligned} y_t &= \alpha + \gamma y_{t-1} + A_t\eta + \Delta gdp_{t-1}\theta + \Delta wage_{t-1}\zeta + \rho_1 W y_t + u_t, \qquad |\rho_1| < 1 \quad (2.1) \\ u_t &= \rho_2 W u_t + \epsilon_t, \qquad |\rho_2| < 1 \quad (2.2) \end{aligned}$$

where y_t and y_{t-1} are the $N \times 1$ vectors of unemployment rates in time periods t and $t-1$, respectively, A_t is the $N \times r$ matrix of observations on r exogenous regional time series variables and Δgdp_{t-1} and $\Delta wage_{t-1}$ are the quarter-on-quarter annualized growth rates of seasonally adjusted national real GDP and national nominal wages in time period $t-1$. W is an $N \times N$ spatial weighting matrix of known constants, α is the $N \times 1$ vector of cross-section dummies, η is the $r \times 1$ vector of regression parameters for regional time series variables, θ is the $N \times 1$ vector of GDP elasticities, ζ is the $N \times 1$ vector of wage elasticities, and γ is the temporal autoregressive coefficient. ρ_1 and ρ_2 are scalar spatial autoregressive parameters, u_t is the $N \times 1$ vector of regression disturbances, and ϵ_t is an $N \times 1$ vector of innovations which are assumed to be identically and independently distributed (iid):

$$\epsilon_t \sim iid(0, \sigma_\epsilon^2) \quad (2.3)$$

The variable Wy_t is typically referred to as spatial lag of y_t.

We next define $X_t = (I_N, y_{t-1}, A_t, \Delta gdp_{t-1} I_N, \Delta wage_{t-1} I_N)$ and $\beta = (\alpha', \gamma, \eta', \theta', \zeta')'$, where I_N denotes the identity matrix of dimension N. The model then reads as follows for each time period t:

$$y_t = X_t\beta + \rho_1 W y_t + u_t \tag{2.4}$$

$$u_t = \rho_2 W u_t + \epsilon_t \tag{2.5}$$

X_t has dimension $N \times k$ and β has dimension $k \times 1$, where $k = 3N + r + 1$. We now use stacked matrix notation to write our model in a more compact form, i.e. we stack the cross-section data for all time periods T in matrices or vectors. For example, the $N \times 1$ vectors y_t which contain the cross-sectional data for each time period are stacked into the $NT \times 1$ vector y which contains the cross-sectional data for all time periods. In order to create an $NT \times NT$ matrix containing the W matrix for all time periods, we use the Kronecker product to define $W_{NT} = I_T \otimes W$, where I_T denotes an identity matrix of dimension T. The model then reads as follows:

$$y = X\beta + \rho_1 W_{NT} y + u \tag{2.6}$$

$$u = \rho_2 W_{NT} u + \epsilon \tag{2.7}$$

Some spatial studies assume that the dependent variable in a cross-section depends on the values of the dependent variable in other cross-sections at time $t-1$ (and not in time t, see, e.g., Giacomini and Granger, 2004). In our model, it is more reasonable to assume that unemployment in one canton depends on unemployment in other cantons at the same time. When unemployment arises from a firm closing in region i, people commuting from other regions get unemployed in the same time period as people living in region i.

Several techniques for estimating models including spatial lags and spatial error lags have been proposed. We use the three-step procedure described by Kelejian and Prucha (1998) as it requires relatively weak assumptions and involves lower computational complexity. Although Kelejian and Prucha presented this model for cross-sectional data, it can easily be adapted for time series (see, e.g., Anselin et al., 2008).

The first estimation step consists of a regression according to equation (2.6). As $W_{NT}y$ is endogenous in this setting, the estimation is performed by two-stage least squares. As instruments, denoted by H, the set $(X, W_{NT}X, W_{NT}^2 X)$ is used. This regression produces the estimators $\widetilde{\beta}$ and $\widetilde{\rho_1}$, which are consistent but not efficient, as they do not take into account the dependencies in the errors. The residuals $\widetilde{u}$ are used as estimators for the disturbances u.

In the second step, we use $\tilde{u}$ to estimate the autoregressive parameter ρ_2 by a generalized moments procedure, which is outlined in appendix 2.A.3. The underlying idea is to transform equation (2.7) repeatedly in order to create a system of equations and to use assumption (2.3) to substitute certain terms. By solving the system of equations, the estimators $\widehat{\rho_2}$ and $\widehat{\sigma_\epsilon}$ can be attained.

For the third step, equation (2.6) is premultiplied by $\widehat{\rho_2}W_{NT}$ and then subtracted from its initial version:

$$y - \widehat{\rho_2}W_{NT}y = X\beta - \widehat{\rho_2}W_{NT}X\beta + \rho_1 W_{NT}y - \rho_1\widehat{\rho_2}W_{NT}^2 y + u - \widehat{\rho_2}W_{NT}u \tag{2.8}$$

setting $y^* = y - \widehat{\rho_2}W_{NT}Y$, $X^* = X - \widehat{\rho_2}W_{NT}X$ and substituting ϵ for $u - \widehat{\rho_2}W_{NT}u$, we get

$$y^* = X^*\beta + \rho_1 W_{NT}y^* + \epsilon \tag{2.9}$$

We now have a model with iid disturbances which we can efficiently estimate by two-stage least squares. The regression produces the final estimators $\widehat{\beta}$ and $\widehat{\rho_1}$.

Kelejian and Prucha (1998) call this step feasible spatial generalized two-stage least squares as the theoretical value of ρ_2, which is needed to calculate y^* and X^*, is not known but estimated.

2.5 Empirical Application to Swiss Data

This sections presents the estimation results (section 2.5.1), compares the fit of the spatial model to alternative models (section 2.5.2) and discusses the spatial dispersion of a shock (section 2.5.3).

2.5.1 Estimation Results

In this section, we will present the estimation results and show how they fit into the literature and the relevant theory. The estimation was performed using quarterly data from 1998 to 2007. The detailed estimation results can be found in Table 2.4 in the appendix. This table contains the results for four alternative models. The "Non-Spatial Model" is the starting point of our analysis as it shows which variables explain the cantonal unemployment rates without the inclusion of spatial elements. The variables showing significance are used in the initial spatial model ("Full Model"). The final spatial model ("Selected Model") results by removing the insignificant variables from the full spatial model. Furthermore, the table shows the estimated parameters of the final spatial model when excluding Basel-Stadt and Geneva from the sample ("Robustness Check").

Non-Spatial Model

The non-spatial model neglects the spatial dependency between the cantons and can thus be estimated by OLS. Nevertheless, the estimated coefficients are quite similar to those of the spatial model. In general, the coefficients are larger than in the spatial model. This is what we would expect as the information contained in the spatial lag (which has a positive mean) is distributed on the other explanatory variables.

Lags in Space and Time

Most studies about (regional) unemployment include a temporal lag of unemployment to their set of explanatory variables and a small number of studies furthermore includes a spatial lag. According to Elhorst (2003) there is mainly a statistical matter to do so. Usually, unemployment rates are highly correlated across time and space and normally change by relatively small amounts from period to period. As the economic situation in adjacent regions is similar, unemployment rates tend to be correlated in space.

But besides the statistical matter, there is also an economic motivation for including the spatial lag into the estimation system: In a small economy as Switzerland, many employees do not work in the same canton as they live: 57% of the employees do not work in the same community, 12% not in the same canton as they live (results from the population census 2000). When a firm closes, employees from this region, but also from other regions lose their jobs. Furthermore, an unemployed person living in a region also looks for a job in other regions. These two channels cause further spatial interdependencies between the unemployment rates of Swiss cantons. We assume the amount by which the unemployment rate in region i is affected by the unemployment rate in region j to be proportional to the travel times between these two regions and form our W-matrix accordingly.

In our estimations, the coefficients of the temporal and the spatial lags are all highly significant and positive. The coefficient of the temporal lag is 0.81 and the coefficients for the spatial lags in the dependent variable and in the error term are 0.18 and 0.64, respectively.

While both spatial lags turn out to be significant, the dependency in the error term seems to be even stronger than the one in the level. This is in line with the finding of Parnisari (2003) stating that spatial correlation is stronger for the cyclical than for the structural component of unemployment.

Women

In their study about Swiss regional unemployment, Flückiger et al. (2007a) find the share of women in the population to be positively correlated with the unemployment rate. Their explanation is that after maternity, women have trouble in getting back to working life and stay unemployed for a longer period than men. Filippini (1998) supposes that another reason for the higher unemployment risk of women lies in the fact that the average qualification is lower for women than for men. We can see from national statistics that the average educational attainment is indeed lower for women than for men. But as we will see for other variables, we do not have this data for Switzerland on cantonal level and therefore do not include it into our estimations.

In line with to Flückiger et al. (2007a), we find the share of women in the population to increase cantonal unemployment.

Cross-Border Commuters

The expected effect of immigration on regional unemployment is not straightforward, as it can increase both labor supply and demand. The effect on unemployment is zero, if migrants fill vacancies for which no one in the home region is qualified. If migrants fill jobs for which also domestic people are qualified, unemployment increases. The effects on labor demand work indirectly via an increased demand for goods and services.

Cross border commuters come to Switzerland only for work and live abroad. They tend to spend more income in their home regions than in their work regions. Therefore, the effects on labor demand are weaker than for migrants. For migrants, the labor supply side is usually assumed to dominate the labor demand side (see, e.g., Oud et al., 2008). For cross-border commuters the supply side should consequently dominate the demand side even more clearly.

For Switzerland, there are no regional quarterly time series of migration. What we do have at hand are quarterly time series of cross-border commuters on a cantonal level. Cross-border commuters are an important phenomenon in Switzerland. All in all, they only account for about 2% of the working population. But when analyzing cantonal data, we find large variance in the share of cross-border commuters. Not surprisingly, border cantons are confronted with more cross-border commuters, while cantons without borders hardly receive any of them. In Basel-Stadt and Geneva, the number of cross-border commuters accounts for more than 10% (Geneva) or even more than 15% (Basel-Stadt) of the canton's working population. These two cantons are furthermore very small (Basel is the smallest, Geneva the fifth-smallest) and the most urban ones. Their population density is more than 25 (Geneva) and 8 (Basel) times the national population density of Switzerland. This all makes these two cantons particular and different from other cantons. Therefore, Flückiger and Vassiliev (2002) analyze the differences in unemployment in these two cantons compared to the rest of Switzerland. In our model, these two cantons are modeled in the same way as all other cantons; their particularity is mainly captured by their high number of cross-border commuters and population density. Additionally, we run a robustness check by estimating the model without Basel-Stadt and Geneva.

In a very simple model, Filippini and Rossi (1992) have shown that the number of cross-border commuters has a significant positive effect on the unemployment rate. In a later study, Flückiger et al. (2007b) have concluded that unemployment is higher in border regions. They explain this fact by the higher competition for vacancies in these regions, which decreases the probability to become re-employed and thus increases the unemployment rate. They emphasize that cross-border commuters do not cause unemployment, but make re-entrance to working life more difficult. Of course, this effect does not work in one direction only. The working potential of foreigners increases the labor supply in Swiss border regions. But the labor supply in foreign border regions is also affected by the Swiss working potential. Similarly to the emergence of domestic spatial effects, this leads to an equalization of the unemployment rates in neighboring regions on an international level. As the unemployment rates are generally higher in Switzerland's neighboring regions than in the Swiss border regions, the number of cross-border commuters should increase unemployment in Switzerland.

In our estimations we indeed find a positive effect of cross-border commuters on the regional unemployment rates. This is perfectly in line with the above mentioned literature. Furthermore, our robustness check shows that this result is not dominated by Basel-Stadt and Geneva as the coefficient is virtually unchanged when these two cantons are excluded from the sample.

Age Structure

Several studies add variables about the age structure of the population to their set of explanatory variables. Most of them find that regions with a relatively young population are confronted with a more severe unemployment problem than regions with a relatively old population (Hofler and Murphy, 1989, Johnson and Kneebone, 1991, Elhorst, 1995, Molho, 1995a,b, Partridge and Rickman, 1995).

As has been shown by Oud et al. (2008) with German data, changes in the age structure towards a younger population lead to higher, and more persistent unemployment rates. Looking at Spanish data, Lopez-Bazo et al. (2002) find unemployment rates of people aged 16–25 to be notably higher than for the total population. And consequently, the unemployment rates of regions with a large share of people in this cohort tend to be higher.

For Switzerland, things are different. From 1990 to 2004, the unemployment rate of the cohort 15–19 years has always been lower than for any other cohort (see Figure 2.4). Compared to other countries, the unemployment rate for this cohort is very low. In the European Union, unemployment among young people is measured in the cohort until 25 years and is thus not entirely comparable to Switzerland. Due to the different systems of industrial training in the different countries, it is furthermore difficult to find the correct age categories for comparison. Nonetheless, it can be said that in the European Union unemployment among young people is much higher than for total population. In the last years, the unemployment rate of people below 25 years was twice the rate for the total population and usually between 15% and 20% (cf. Eurostat, 2009).[6]

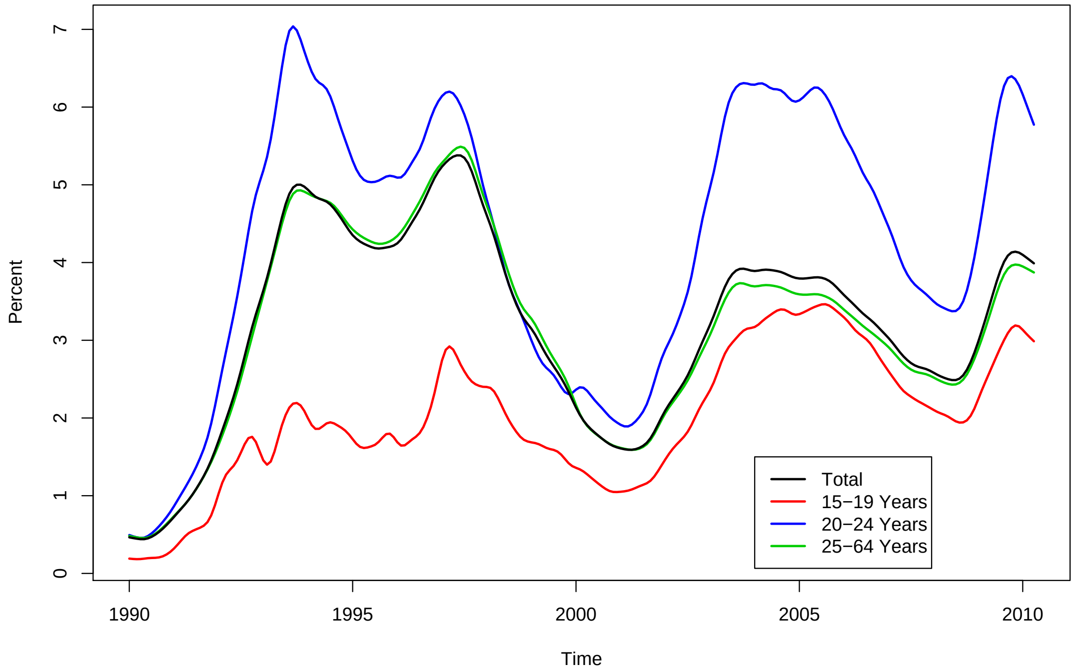
In their study about Switzerland, Flückiger and Vassiliev (2002) have shown that the probability of losing the job is higher for younger people compared to the total population. On the other hand, the average duration of unemployment is much shorter for younger people.

As we would expect from the above argumentation, Feld and Savioz (2000) showed that the higher the number of young people in a canton, the lower is unemployment. They suppose that this could be a human capital effect.

We include the share of people in the following age categories into our estimations: 20–24 years and 25–64 years. The base line category is therefore people aged under 20 years and people aged over 64 years. The shares of 20–24 years and 25–64 years old people both have

[6]In Germany and Austria – where roughly comparable industrial training programs exist – the differences between the age categories were not as high, while in France and Italy this ratio was even larger than 2.

Figure 2.4: *Swiss Unemployment Rates in Different Age Cohorts.*



NOTE: From 1990 to 2004, the unemployment rate of the cohort 15–19 years has always been lower than for any other cohort. On the other hand, the unemployment rate of the 20–24 years old people was higher than in the other cohorts. Due to unstable seasonal patterns, the trend cycles are shown.

a significant positive effect on regional unemployment. We see two reasons for the fact that both coefficients are positive: firstly, with the exception of the 15–19 years old people, the former two age categories cover those people who are actually able to register as unemployed. Thus, it makes sense that a higher population share of these cohorts should in principle increase the unemployment rate. Secondly, we already mentioned that the only age category which can furthermore register as unemployed, namely the 15–19 years old, experience a lower-than-average unemployment rate. This age category was included in earlier versions of our model and its coefficient was actually negative, although not significantly. Nevertheless, compared to the 15–19 years old, the 20–24 and 15–64 years old people experience higher unemployment rates, and thus unemployment should rise with their population share. When comparing the coefficients of the included age categories, we see that the coefficient of the 20–24 years old is higher than the one for the 25–64 years old. This reflects the fact that the 20–24 years old people experience a clearly higher-than-average unemployment rate.

Sectoral Employment

The effects of the sectoral employment mix on unemployment are not a priori clear, as different argumentations are possible. This is reflected in the mixed results of including the sectoral employment shares (Elhorst, 1995, Partridge and Rickman, 1995, 1997, Taylor and Bradley, 1997). Armstrong and Taylor (1993) argue that the unemployment rate is not specific to regions, but to industries. Therefore, they explain the unemployment rate in a specific region by its industry mix and the national rate of unemployment in each industry. But as Martin (1997) shows, the unemployment rate of a specific industry can be different in different regions.

An important phenomenon in the context of sectoral employment and unemployment is the tertiarization process, which is supposed to cause unemployment to be higher in regions specialized in agriculture and manufacturing than in regions specialized in services. As Sheldon (1999) points out, Switzerland was confronted with a tertiarization of the working environment since the 1960s. Employment in production and manufacturing decreased, while the services sector showed an increase in employment and added value. Steffen (2005) thus supposes that cantons with a large manufacturing sector have a larger number of endangered jobs. In cantons with a large service sector, tertiarization is more advanced and the number of endangered jobs is smaller. However, the estimations conducted by Steffen (2005) indicate that unemployment rises with the employment share of the services sector.

As it seems, tertiarization is not the only channel through which the sectoral mix affects unemployment.

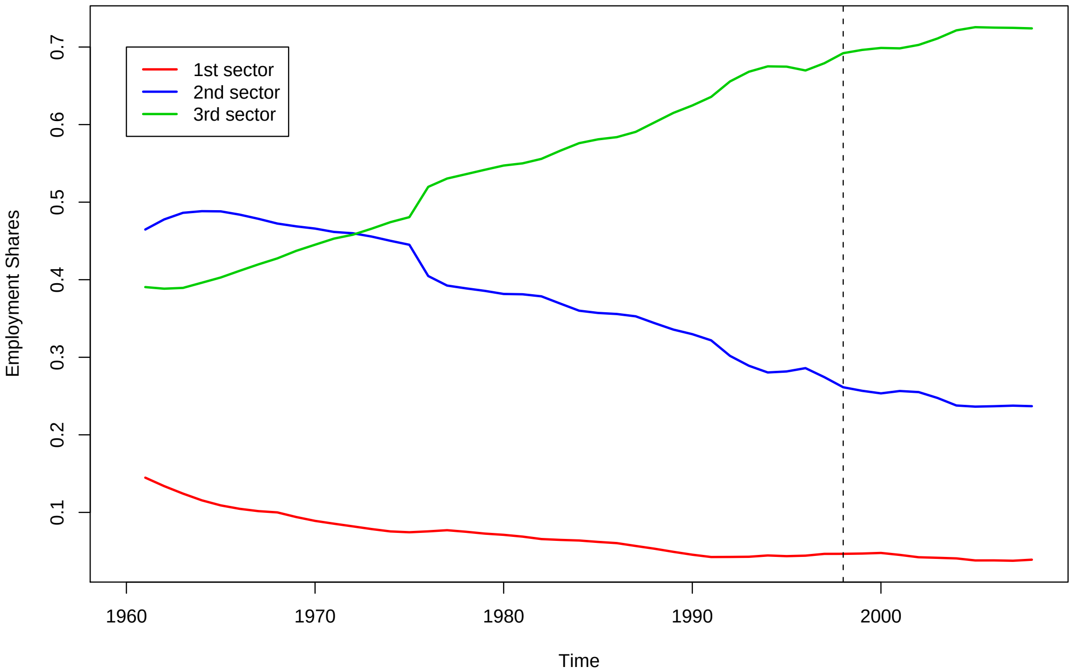
An argument in the opposite direction can be made using employment multipliers. The former are generally higher in agriculture and manufacturing than in the services sector, given that the first two sectors – at least partly – create the demand for some services (Elhorst, 2003). Consequently, unemployment could fall with the employment shares of the first two sectors. With regard to the first sector, the following argumentation can additionally be made: enterprises in the agriculture sector are frequently family enterprises. If a family member loses his/her job outside the farm, he/she often can work on the farm and does not register as unemployed. Furthermore, social control is rather strong in regions with a high employment share of the first sector, i.e. people who lost their job may register less frequently as unemployed than in other regions because they feel ashamed. A high employment share of the first sector could thus dampen unemployment not through economic, but through psychological mechanisms.

In our study, we would not expect tertiarization to play a dominant role as we use data from 1998 to 2007. Figure 2.5 shows that the process of tertiarization was quite advanced in 1998. Although the employment shares changed further during our sample, the main part of the evolution was already completed. With respect to the second sector, we can furthermore say that those industries that still produced in Switzerland in 1998 were mainly high tech industries. They may have a rather high employment multiplier and be less affected by tertiarization.

Similar to Oud et al. (2008), we measure the economic structure as the proportion of employment in agriculture, manufacturing and services. In our estimations, we include the shares in employment of the second and the third sector in the explanatory variables. While the coefficient of the second sector is not significant (and thus not reported), the coefficient of the third sector is positive. A higher employment share of the third sector therefore increases unemployment. We thus conclude that tertiarization does indeed not dominate our result. Instead, higher employment multipliers, robust industries, familial job opportunities and social control seem to decrease unemployment in the first and the sector sector.

Besides the breakdown into the three sectors, we also analyze the internal structure of the modernity of the industries and discriminate three stages: traditional, modern and high-tech industries. The share of employment in traditional industries is supposed to increase unemployment as the former are generally declining. On the other hand, the employment

Figure 2.5: *Employment Shares of the Sectors.*



NOTE: The dotted line shows 1998, the begin of our sample.

share of high-tech industries should decrease unemployment as these are only little affected by structural change. The same argument can be mode for modern industries. Surprisingly, we only find a significant effect for the modern industries: the higher the share of employment in modern industries, the higher is unemployment. We explain this result with the bursting of the "dot-com-bubble", which falls into our estimation sample and caused unemployment among bankers and IT engineers to increase considerably. As modern industries mainly include the financial sector and IT, this one-time shock could explain our result.

Industrial Concentration

Industrial diversity – the opposite of industrial concentration – is the dispersion of employment over industries. Several studies suggest that it is negatively related to unemployment. The idea is that regions with multiple sources of employment provide more chances to become re-employed. Elhorst (2001) cites five studies investigating the relationship between unemployment and industrial diversity (Taylor and Bradley, 1983, Neumann and Topel, 1991, Malizia and Ke, 1993, Simon, 1988, Partridge and Rickman, 1995). All of them find the expected effect. However, it is significant in the first three studies only. On the other hand, one might also argue that concentration in employment is a sign for industrial clustering, which leads to network effects. These can lower unemployment.

The dispersion of employment over the industries is measured by the Herfindahl index. Originally developed to measure the amount of competition among firms, it has undergone several adaptions for other purposes. The idea is to calculate an index of the concentration of output over firms, employment over industries etc. It is calculated as $H = \sum_{i=1}^{N} s_i^2$ with N: number of industries, s_i: share of employment in industry i. It reaches a maximum of 1 in the case of total concentration of employment in one industry. Smaller values indicate more equally distributed employment.

For Switzerland, Filippini (1998) assumes economies with a larger dispersion of employment to show more stable GDP growth and lower unemployment rates than economies specialized in fewer industries. This is confirmed by the findings of Steffen (2005). In contrast, we find a concentration of employment in few industries to lower unemployment. The benefits of clustering seem to dominate the gains of diversification. The differences to the results of the above cited studies may be due by the varying samples used.

This effect only exists in the non-spatial setting. When including spatial elements, the effect of industrial concentration is not significant any more.

Public Sector

As stated by Steffen (2005), the share of employment of the public sector is assumed to have a dampening effect on unemployment. The argument for this is twofold: Firstly, the public sector is less affected by economic crises (see Forni, 2002), and secondly, the enlargement of public employment in times of economic crises can dampen the raise in unemployment (Armingeon, 1999). Other studies (see, e.g., Parnisari, 2003) put it another way and state that a large share of employment in industries with a low responsiveness to the business cycle lower the variance of unemployment.

Steffen (2005) includes the share of employment of the public sector in her models. She cannot find a significant effect in any of the estimated models. In contrast to Steffen (2005), but in line with the above mentioned theory, we find a slight, but negative effect.

Population Density

Population density is often taken as a proxy for the market potential of a region. It is sometimes called potential demand or agglomeration potential. The idea is that through the mechanisms of economies of scale and transport costs, firms tend to produce – and thus demand labor – in regions with high population density. Elhorst (1995) and Molho (1995a,b) model the regional unemployment rate as dependent on the market potential. They all find a negative and significant effect of the market potential on the regional unemployment rate. On the other hand, there is also an intuitive argument for the opposite direction: unemployed people may rather move into the cities as they can benefit from the nearby services and vacancies.

In Switzerland, population density is very different across cantons. In the two urban cantons Basel-Stadt and Geneva it is very high while in more rural cantons as the Grisons or Uri it is quite low. The maximum is reached by Basel-Stadt with more than 5000 people per km^2, in Geneva it is about 1500 people per km^2. As we have discussed above, these two border cantons are confronted with a large number of cross-border commuters. The nation-wide mean population density is 187 people per km^2 and the minimum lies at 25 persons per km^2 in the Grisons. In general, population density is negatively correlated with the share of persons occupied in the first sector (cf. Figure 2.6).

In our estimations, we find a significant positive effect of the population density on unemployment. This contradicts the presented theory, but is in line with the intuition.

As mentioned, there is a high correlation with cross-border commuters in the cantons Basel-Stadt and Geneva. As these two cantons could influence the estimates for population density much, we checked the robustness by estimating our model without these two cantons. The result is given in Table 2.4. We see that the coefficient for population density becomes insignificant when Basel Stadt and Geneva are excluded from the sample while all other coefficients are stable. This result is more in line with the theory than the result when considering all cantons.

GDP

One of the most widely used variables for labor demand is the regional GDP in real terms. The relation is assumed to be positive, i.e. we expect a negative relation to unemployment.

Feld and Savioz (2000) use cantonal GDP for explaining cantonal unemployment rates. In their estimations, they find a negative – but insignificant – effect. Only when aggregating the 26 cantons to seven NUTS-2 regions, the effect becomes significant. Steffen (2005) finds a negative – and for some models even significant – effect of GDP growth on regional unemployment. She reasons that cantonal economic growth is important for describing the different levels of unemployment, but it is not their main determinant.

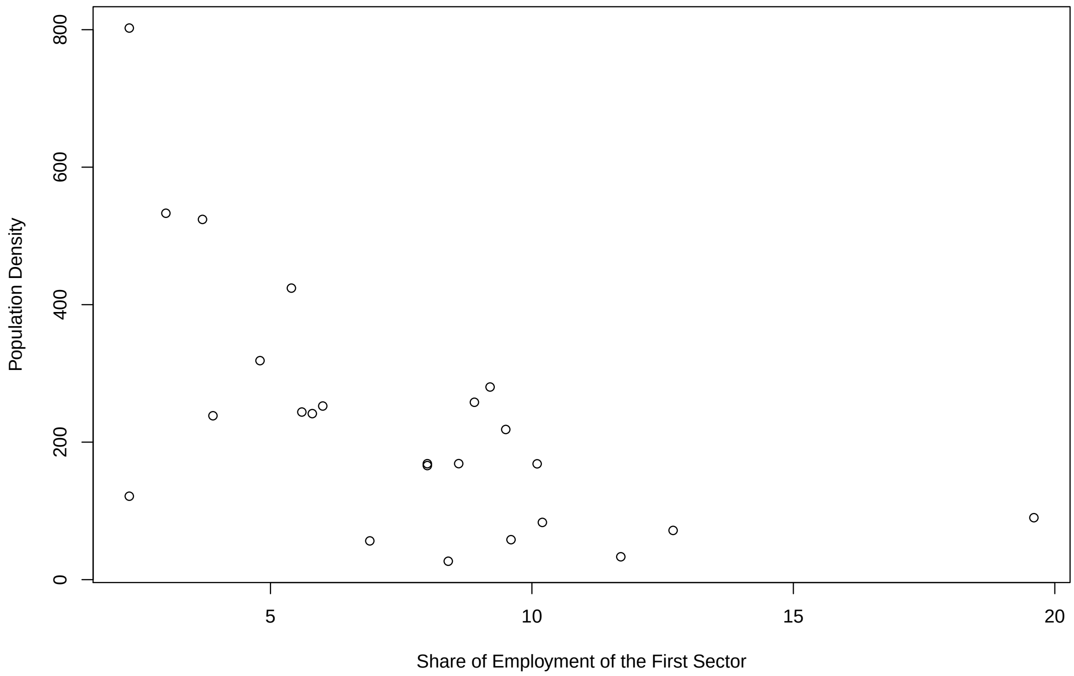
As in Switzerland there is no quarterly data for the regional GDP, we use the national GDP instead. We include annualized quarter-on-quarter GDP growth with a lead of one quarter and individual coefficients to our model. The coefficients can then be interpreted as elasticities. Figure 2.7 illustrates the cantonal elasticities of the unemployment rate to GDP. The strongest reactions are registered in Geneva, Zurich, Jura and Neuchâtel. On the other hand, only small elasticities are registered in Appenzell Inner Rhodes and Outer Rhodes, Ticino, Obwalden and Nidwalden.

As the coefficients for the elasticity to GDP are quite similar across the cantons, we test them on equality. The corresponding F-Test rejects the equality of equal elasticities. We therefore need individual coefficients for GDP.

Wages

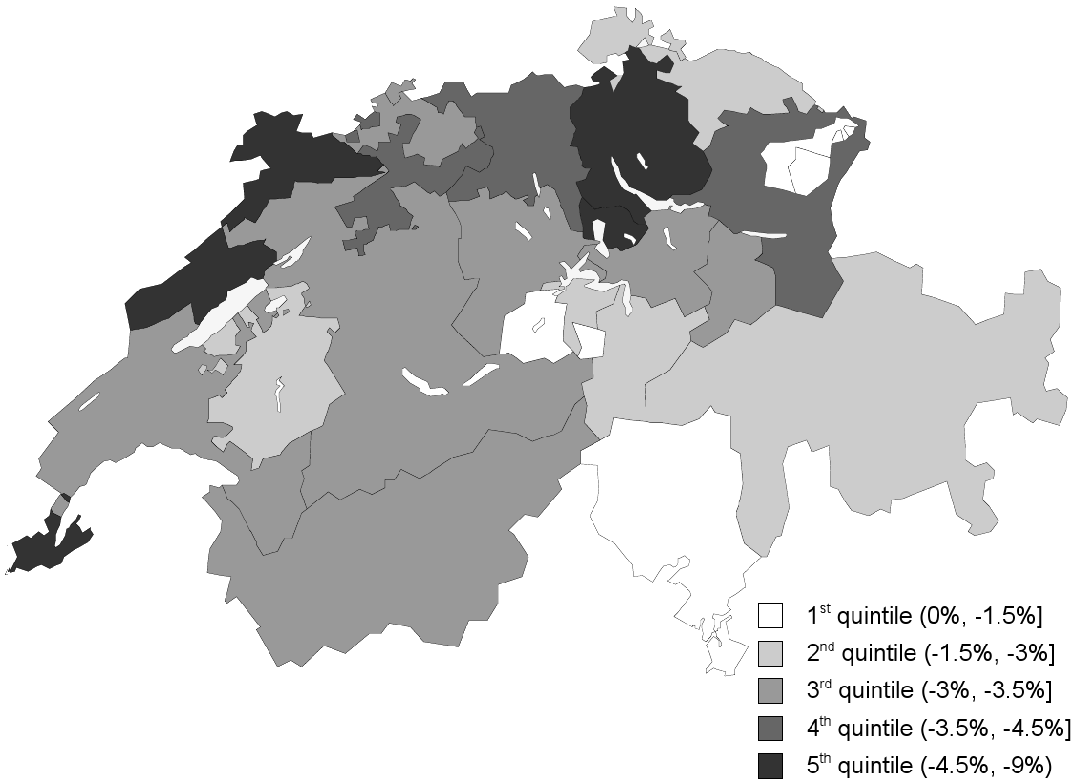
The wage level represents the price of labor and has been used in many studies (Steffen, 2005, Filippini, 1998, Burridge and Gordon, 1981, Hofler and Murphy, 1989, Partridge and Rickman, 1995, Molho, 1995b,a, Partridge and Rickman, 1997). Wages are usually believed to have a positive effect on labor supply and a negative effect on labor demand, hence unemployment should increase if wages rise. Indeed, most Swiss and international studies analyzing the effect of wages on unemployment find a positive relationship (cf., e.g., Steffen,

Figure 2.6: *Employment Share of the First Sector and Population Density.*



NOTE: The values for Basel Stand and Geneva are outside the panel. They both have a small share of employment in the first sector and a very large population density.

Figure 2.7: *Elasticities of Swiss Cantons to a Change in National GDP.*



NOTE: The strongest reactions are registered in Geneva, Zurich, Jura and Neuchâtel. On the other hand, only small elasticities are registered in Appenzell Inner Rhodes and Outer Rhodes, Ticino, Obwalden and Nidwalden.

2005, Filippini, 1998, Burridge and Gordon, 1981, Hofler and Murphy, 1989, Partridge and Rickman, 1995, 1997).

As a measure for wages, we use the wage index of the Swiss Federal Statistical Office, which records the evolution of the national mean wage. This index excludes bonuses and measures the change in the fees paid for given tasks. In other words, it is a good index for the evolution of unit labor costs, rather than a general income measure. We include annualized quarter-on-quarter growth rates of the wage index in our model. As we have no regional data for the wage index, we allow the cantons to have individual elasticities. In line with the above mentioned literature, we find a positive effect of the wages on unemployment in all cantons. Figure 2.8 illustrates the cantonal elasticities of the unemployment rate to the wages. With the exception of Jura, the Latin speaking cantons show lower elasticities than the German ones.

Again, we conduct an F-Test to check whether we really need individual wage elasticities. The null hypothesis of one equal coefficient is clearly rejected.

Constants

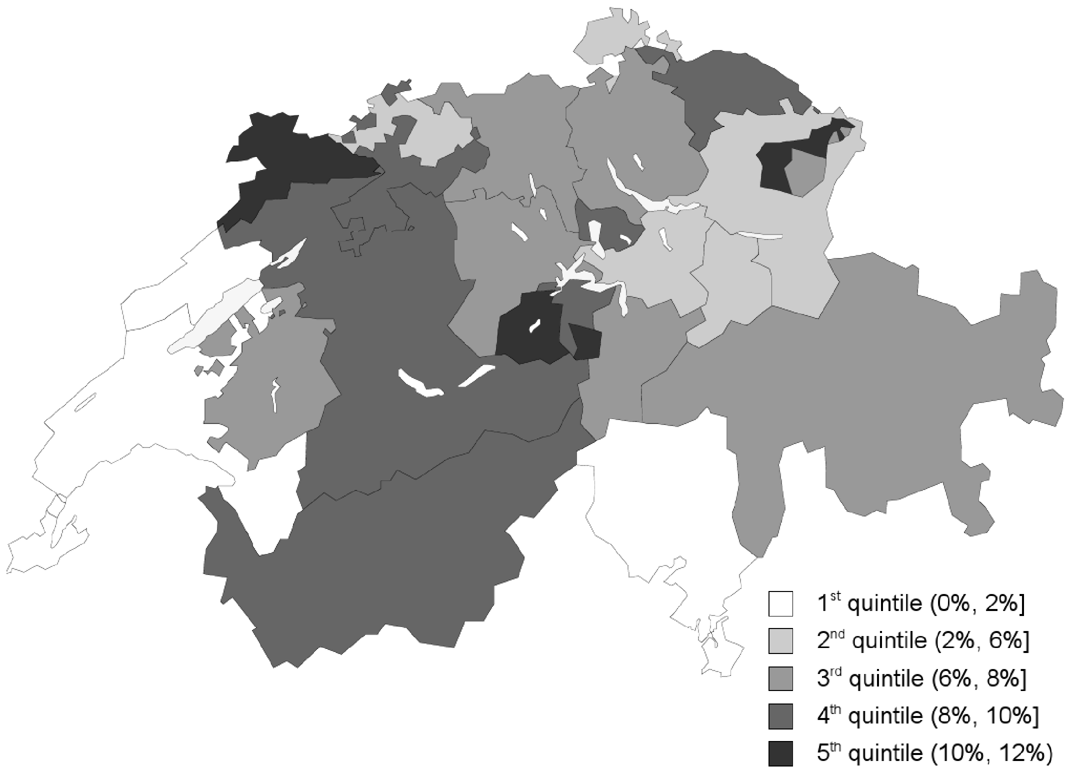
Studies analyzing the different levels in the unemployment rates across regions often use individual constants to summarize all unobserved cross-section specific factors. An F-test clearly indicates that these individual constants should not be summarized into one common intercept in our model.

2.5.2 Model Fit

To evaluate how well our model fits the data, we use two benchmark models (results on the coefficients are not reported): A model with the same exogenous explanatory variables but without spatial components ("OLS model") and a naive autoregressive model, containing individual constants and a common AR-coefficient for all cross-sections ("AR model"). In both models, the coefficient of the time lag of the dependent variable is considerably higher than in the spatial model (OLS model: 0.90; AR model: 0.92). This is due to the high persistence of the unemployment rates: a part of the information that was contained in the spatial lags is now captured by the time lag.

To compare the results of the spatial model to those of the benchmark models, we measure the in-sample fit of the three models. However, there are several possibilities of how to fit the endogenous variable: While for the OLS model and the AR model, the

Figure 2.8: *Elasticities of Swiss Cantons to a Change in National Wages.*



NOTE: With the exception of Jura, the Latin speaking cantons show lower elasticities than the German ones.

information to use for the fit clearly consists of the actual values of the exogenous variables and the temporally lagged dependent variable, for the spatial model, we may furthermore include the spatial lag (i.e., the contemporaneous values of the dependent variable). Kelejian and Prucha (2007) present four predictors for spatial models, based on different information sets:[7]

$$\begin{aligned} \Lambda_1 &= \{X_t, W\} \\ \Lambda_2 &= \{X_t, W, w_{i.}y_t\} \\ \Lambda_3 &= \{X_t, W, y_{-i,t}\}, \end{aligned} \tag{2.10}$$

where $w_{i.}$ denotes the i-th row of W and $y_{-1,t}$ denotes the y_t vector from which the i-th element has been removed. Information set Λ_1 is thus a subset of Λ_2 which again is a subset of the full information set Λ_3. The first three predictors which Kelejian and Prucha (2007) present are the conditional means corresponding to the information sets Λ_p, $p = 1, 2, 3$, in (2.10):

$$\begin{aligned} y_{it}^{(1)} &= E(y_{it}|\Lambda_1) \\ &= (I - \rho_1 W)_{i.}^{-1} x_t \beta \end{aligned} \tag{2.11}$$

$$\begin{aligned} y_{it}^{(2)} &= E(y_{it}|\Lambda_2) \\ &= \rho_1 w_{i.}y_t + x_{ti.}\beta + \frac{cov(u_{it}, w_{i.}y_t)}{var(w_{i.}y_t)}[w_{i.}y_t - E(w_{i.}y_t)] \end{aligned} \tag{2.12}$$

$$\begin{aligned} y_{it}^{(3)} &= E(y_{it}|\Lambda_3) \\ &= \rho_1 w_{i.}y_t + x_{ti.}\beta + cov(u_{it}, y_{-i,t}) \left[VC(y_{-i,t})\right]^{-1} \left[y_{-i,t} - E(y_{-i,t})\right] \end{aligned} \tag{2.13}$$

where VC denotes the variance-covariance matrix. Calculational details are given in Kelejian and Prucha (2007). The fourth predictor they consider is given by

$$y_{it}^{(4)} = \rho_1 w_{i.}y_t + x_{i.t}\beta \tag{2.14}$$

[7] Kelejian and Prucha (2007) define their predictors for a cross-sectional model. We adapt their predictors for a time series model. A fifth predictor which is presented in their paper refers to the spatial error model without spatial lag in the dependent variable and is thus not adequate for our model.

This predictor may be seen as a restricted version of predictor $y_{it}^{(2)}$, implicitly assuming that $cov(u_{it}, w_{i.}y_t) = 0$. This does, however, not hold in general. Thus, conditional on the information set Λ_2, predictor 4 is biased. Kelejian and Prucha (2007) find that predictor 3, which uses the full information set Λ_3, empirically performs best.

To compare the predictors, we measure the goodness of fit by the root mean squared error (RMSE) and the adjusted R^2. Tables 2.1 and 2.2 show the ranks which the predictors of the spatial model and the benchmark models reach as measured by these criteria[8]. Furthermore, figures 2.9 and 2.10 show boxplots of the distribution of the model fit criteria through the cross-sections. We see that predictors 2 and 3 of the spatial model together reach by far the best ranks. Measured by the RMSE ranks, predictor 2 even seems to outperform predictor 3. However, the corresponding boxplots show that the RMSEs of the two predictors are distributed similarly. The adjusted R^2 indicates that the two predictors perform roughly the same as well. This result indicates that in our context, there is no clear difference between the information sets Λ_2 and Λ_3, or rather between the information contained in $w_{i.}y_t$ and $y_{-i,t}$. In words, knowing the weighted mean of the unemployment rates in all other cantons is roughly the same as knowing the actual unemployment rates in the other cantons. This implies that our W matrix is a good measure for the interdependencies between the cantonal labour markets. Behind predictors 2 and 3, predictors 4 and 1 and the OLS model follow. The AR model, which uses the smallest information set of all predictors, clearly performs the worst.

2.5.3 Emanating effects

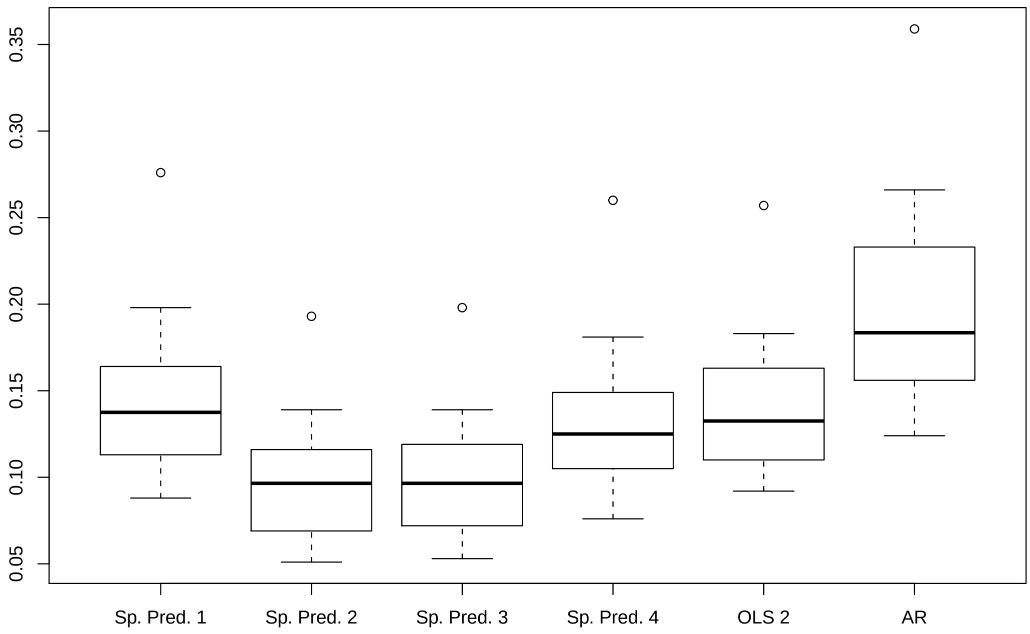
We now use our model to calculate the emanating effects of a change in the explanatory variables. We run an alternative scenario assuming a reduction of cross-border commuters by 10% from the first quarter 2003 onwards in all cantons as compared to the actual values. We then compare the resulting unemployment rates with the fitted unemployment rates in our model which result using the actual values. A change as assumed in our scenario might be caused by a more restrictive handling of border commuting. Such policies are sometimes called for as a measure against unemployment in border cantons. A reduction of cross-border commuters would not only influence unemployment directly in all cantons which record cross-border commuters, but by spatial effects also indirectly in all other cantons.

[8]For the RMSE themselves, see Table 2.5

Table 2.1: *Root Mean Squared Error of the Different Predictors, Ranks.*

	Sp. Pred. 1	Sp. Pred. 2	Sp. Pred. 3	Sp. Pred. 4	OLS	AR
1st	0	15	9	2	0	0
2nd	1	8	16	1	0	0
3rd	0	2	0	17	7	0
4th	9	0	1	6	10	0
5th	16	0	0	0	9	1
6th	0	1	0	0	0	25

Figure 2.9: *Root Mean Squared Error of the Different Predictors.*



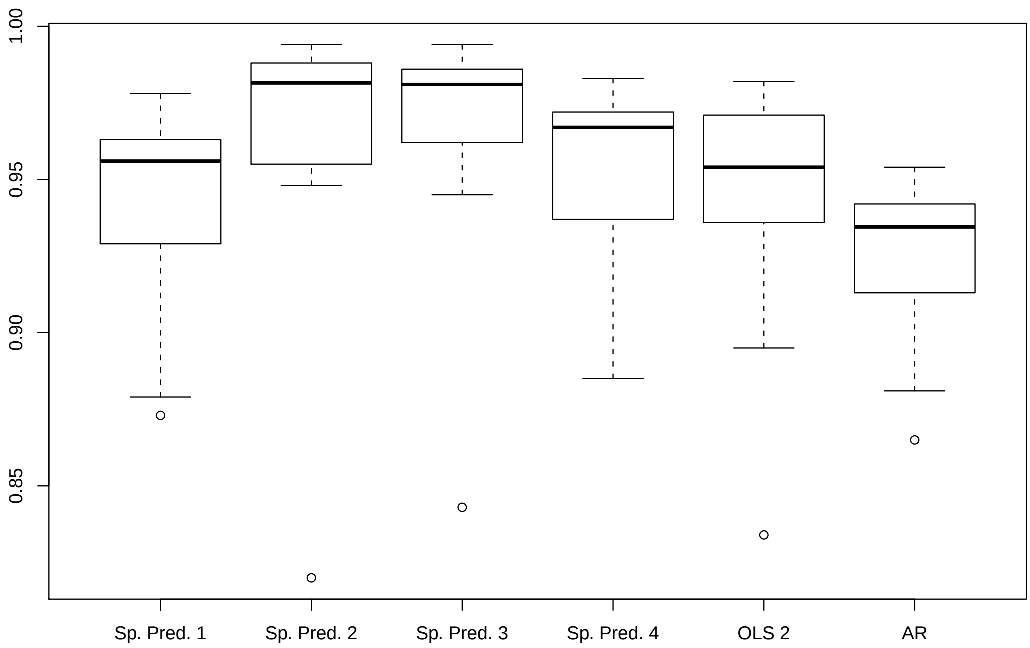
NOTE: Predictor 2 reaches the best performance for some cross-sections, but performs worse than predictor 3 for some others.

Table 2.2: *Adjusted R squared of the Different Predictors, Ranks.*

	Sp. Pred. 1	Sp. Pred. 2	Sp. Pred. 3	Sp. Pred. 4	OLS 2	AR
1st	0	12	11	2	1	0
2nd	1	12	13	0	0	0
3rd	0	1	1	14	8	2
4th	10	0	1	9	5	1
5th	13	0	0	1	12	0
6th	2	1	0	0	0	23

NOTE: Predictors 2 and 3 of the spatial model together reach by far the best ranks. Predictor 2 even obtains the best fit more often than Predictor 3. Predictors 4 and 1 and the OLS model follow, while the AR model performs clearly the worst.

Figure 2.10: *Adjusted R squared of the Different Predictors.*



NOTE: Predictor 2 reaches the best performance for some cross-sections, but performs clearly worse than predictor 3 for some others. Not surprisingly, the models using less information perform worse.

As the coefficient of the share of cross-border commuters is negative in our model, reducing the number of cross-border commuters will ceteris paribus lower unemployment.

The emanating effects of a change in one or more explanatory variables can be calculated by rewriting the model (2.4) as follows (see, e.g., Kelejian et al., 2008):

$$y_t = (I - \rho_1 W)^{-1} X_t \beta + (I - \rho_1 W)^{-1} u_t \tag{2.15}$$

Let $(I - \rho_1 W)^{-1} X_t \beta$ be the $N \times 1$ vector of values resulting from the actual values of X_t. Furthermore, let $(I - \rho_1 W)^{-1} X_t \beta|_{alt}$ denote the corresponding values for our alternative scenario. Assuming that the error terms remain unchanged, the emanating effects can then be calculated as the difference between these values:

$$(I - \rho_1 W)^{-1} X_t \beta|_{alt} - (I - \rho_1 W)^{-1} X_t \beta \tag{2.16}$$

Table 2.3 shows the detailed emanating effects, i.e. the differences which result from the reduction in cross-border commuting. The effects are listed for the yearly average unemployment rate. The emanating effects grow over time, but the increases diminish towards the end of the horizon. This increase over time is due to the auto-regressive structure of our model, where the differences in one period have an impact on the differences in the following periods. The new equilibrum lies up to 1.16 percentage point below the initial values for Basel-Stadt, whereas almost no effect at all is visible in Uri, Schwyz and Zug. Figure 2.11 shows the geographical dispersion of the emanating effects four years after the change in commuting. Not surprisingly, the strongest results are registered in the border cantons in general, and especially in Basel-Stadt, Basel-Land, Jura, Geneva and Ticino. These five cantons record the highest share of cross-border commuters and thus experience the highest reduction of commuters in our alternative scenario. The more centrally situated cantons on the other hand experience small effects only. Some central cantons record no cross-border commuters at all (and therefore no reduction in the alternative scenario) and only experience indirect effects.

A word of caution about the results discussed here is in order. We have run an alternative scenario assuming that all explanatory variables except the number of cross-border commuters remain unchanged. We would, however, expect a forced reduction in the number of cross-border commuters to have an impact on some of the other explanatory variables.

Especially GDP could be affected negatively, which in turn would raise unemployment. In order to determine the dynamic effects of our scenario, we would need to determine the affected explanatory variables endogenously in the model. Therefore, our static analysis cannot determine whether a reduction in the number of cross-border commuters would raise or lower unemployment once all direct and indirect effects are taken into account. It should merely be seen as an illustration of the dispersion of the immediate effects. Because of the spatial lag in the model, the cantons that do not record cross-border commuters at all are hit indirectly via the changed unemployment rate of other cantons.

2.6 Conclusions

In this paper, we estimate a spatial time series model for the Swiss cantonal unemployment rates. This allows us to determine the variables which explain the levels and fluctuations in the regional unemployment rates and investigate whether these variables retain their explanatory power once spatial elements are added to the model. As we find spatial dependency in the levels as well as in the errors, we estimate a SARAR(1,1) model. Our estimation results indicate that spatial spillovers have additional explanatory power to our other model variables.

Furthermore, we find that regional unemployment is raised by the population shares of women, of cross-border commuters and of people aged between 20 and 24 as well as between 25 and 64, by the population density, by the employment share of the third sector and of modern industries and by wage growth. On the other hand, we find decreasing effects by the employment share of the public sector and by GDP growth. Most of these findings are consistent with the existing literature on regional unemployment levels. We then compare the predictions of the spatial model to some alternative models. We see that our spatial model reaches a better fit than a benchmark model which uses the same explanatory variables but no spatial lags.

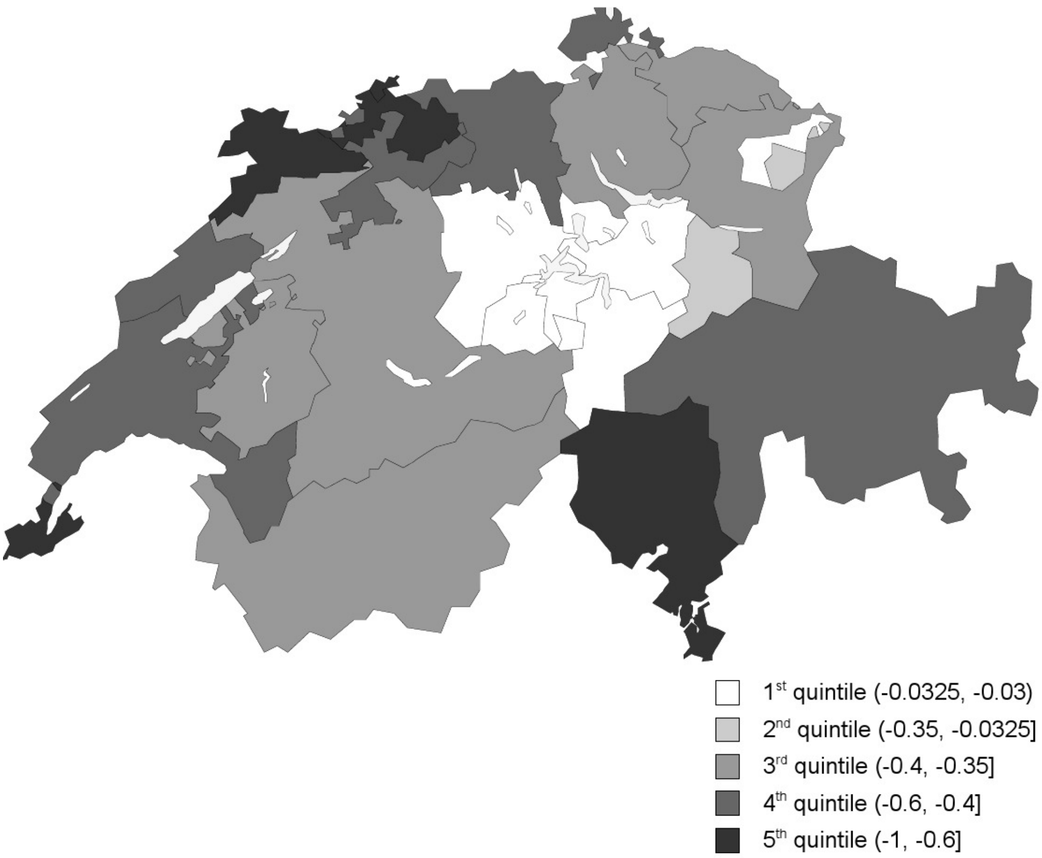
The possibilities of hypothesis testing with our model are restricted by the availability of quarterly regional data. Furthermore, the data we have at hand is often disposable for short time periods only. As data availability improves over time, expansions of our model might well be possible in the future.

Table 2.3: *Emanating Effects for the Alternative Scenario of a Reduction in Cross-Border Commuters.*

	2003	2004	2005	2006	2007
AG	−0.05	−0.14	−0.24	−0.34	−0.44
AI	−0.02	−0.08	−0.16	−0.24	−0.34
AR	−0.02	−0.08	−0.15	−0.23	−0.32
BE	−0.02	−0.09	−0.17	−0.27	−0.37
BL	−0.15	−0.35	−0.51	−0.65	−0.78
BS	−0.32	−0.66	−0.86	−1.02	−1.16
FR	−0.02	−0.08	−0.17	−0.26	−0.36
GE	−0.17	−0.38	−0.54	−0.69	−0.83
GL	−0.02	−0.08	−0.15	−0.24	−0.33
GR	−0.05	−0.14	−0.22	−0.32	−0.41
JU	−0.13	−0.29	−0.43	−0.56	−0.7
LU	−0.02	−0.07	−0.14	−0.23	−0.32
NE	−0.07	−0.19	−0.3	−0.41	−0.54
NW	−0.02	−0.07	−0.14	−0.22	−0.32
OW	−0.02	−0.07	−0.14	−0.22	−0.32
SG	−0.04	−0.11	−0.18	−0.26	−0.36
SH	−0.11	−0.24	−0.35	−0.46	−0.58
SO	−0.03	−0.11	−0.2	−0.3	−0.4
SZ	−0.02	−0.07	−0.14	−0.22	−0.31
TG	−0.04	−0.12	−0.21	−0.3	−0.4
TI	−0.21	−0.44	−0.59	−0.72	−0.85
UR	−0.02	−0.07	−0.14	−0.22	−0.31
VD	−0.06	−0.15	−0.26	−0.36	−0.47
VS	−0.03	−0.11	−0.2	−0.3	−0.4
ZG	−0.02	−0.07	−0.14	−0.22	−0.31
ZH	−0.02	−0.09	−0.17	−0.26	−0.36

NOTE: Annual averages in official unemployment rates.

Figure 2.11: *Emanating Effects in the Year 2007 for the Alternative Scenario of a Reduction in Cross-Border Commuters.*



NOTE: Annual average in official unemployment rates.

2.A Appendix

2.A.1 Estimated Coefficients

Table 2.4: *Estimation Results.*

	Selected Model	Full Model	Robustness Check	Non-Spatial Model
Temporal Lag	0.81***	0.81***	0.82***	0.90***
	(0.02)	(0.02)	(0.02)	(0.01)
Spatial Lag (ρ_1)	0.18***	0.18***	0.17***	
	(0.02)	(0.05)	(0.02)	
Spatial Error Lag (ρ_2)	0.64***	0.63***	0.64***	
	(0.05)	(0.02)	(0.00)	
Cross border commuters	0.09***	0.09***	0.10***	0.12***
(% share in total population)	(0.01)	(0.01)	(0.02)	(0.01)
Women	0.21***	0.21***	0.21***	0.40***
(% share in total population)	(0.04)	(0.04)	(0.04)	(0.05)
Population Density	0.97***	0.96***	1.43	0.68*
(population in 1000 per km^2)	(0.30)	(0.30)	(1.20)	(0.38)
Population aged 20-24	0.06**	0.06**	0.06**	0.21***
(% share in total population)	(0.03)	(0.03)	(0.03)	(0.03)
Population aged 25-64	0.04***	0.04***	0.05***	0.11***
(% share in total population)	(0.02)	(0.02)	(0.02)	(0.02)
Size of Public Sector	−0.02**	−0.02**	−0.02**	−0.05***
(% share in FTE employment)	(0.01)	(0.01)	(0.01)	(0.01)
Industrial Concentration		−3.94		−14.56*
		(4.15)		(6.12)
Size of 3rd Sector	0.01***	0.01***	0.02***	0.04***
(% share in FTE employment)	(0.01)	(0.01)	(0.01)	(0.01)
Size of Modern Industries	0.02**	0.02***	0.01*	0.08***
(% share in FTE employment)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	1040	1040	1040	1040

NOTE: Omitted variables are constants as well as GDP and wage elasticities. Standard errors in parentheses; * significant at the 90% confidence level; ** significant at the 95% confidence level; *** significant at the 99% confidence level; FTE: full-time equivalent

2.A.2 Root Mean Sqared Error of the Different Predictors

Table 2.5: *Root Mean Sqared Error of the Different Predictors.*

	Sp. Pred. 1	Sp. Pred. 2	Sp. Pred. 3	Sp. Pred. 4	OLS	AR
AG	0.14	0.06	0.06	0.12	0.12	0.19
AI	0.15	0.10	0.10	0.14	0.14	0.15
AR	0.16	0.11	0.11	0.14	0.12	0.16
BE	0.12	0.05	0.06	0.11	0.10	0.18
BL	0.11	0.07	0.07	0.09	0.10	0.18
BS	0.15	0.09	0.09	0.13	0.15	0.22
FR	0.09	0.09	0.08	0.08	0.10	0.16
GE	0.17	0.12	0.12	0.16	0.15	0.26
GL	0.15	0.09	0.10	0.13	0.14	0.18
GR	0.11	0.09	0.08	0.09	0.10	0.14
JU	0.28	0.19	0.20	0.26	0.26	0.36
LU	0.14	0.07	0.08	0.13	0.14	0.19
NE	0.16	0.12	0.12	0.15	0.18	0.25
NW	0.10	0.07	0.07	0.09	0.13	0.14
OW	0.12	0.11	0.11	0.11	0.13	0.14
SG	0.11	0.06	0.07	0.09	0.09	0.16
SH	0.17	0.12	0.12	0.15	0.18	0.22
SO	0.20	0.12	0.12	0.18	0.18	0.27
SZ	0.11	0.07	0.07	0.10	0.13	0.16
TG	0.12	0.05	0.05	0.11	0.11	0.18
TI	0.14	0.13	0.12	0.13	0.17	0.23
UR	0.11	0.13	0.12	0.11	0.12	0.12
VD	0.13	0.10	0.09	0.12	0.14	0.22
VS	0.17	0.14	0.14	0.16	0.18	0.24
ZG	0.14	0.10	0.10	0.12	0.11	0.21
ZH	0.17	0.11	0.11	0.16	0.16	0.25
Median	0.14	0.10	0.10	0.13	0.13	0.18

2.A.3 Generalized moments procedure

This section outlines the generalized moments procedure proposed by Kelejian and Prucha (1998), which produces an estimator for the parameter ρ_2. The point of departure are the following transformations of equation (2.7):

$$u - \rho_2 W_{NT} u = \epsilon \tag{2.17}$$

$$W_{NT} u - \rho_2 W_{NT}^2 u = W_{NT}\epsilon \tag{2.18}$$

We define u_i, $\overline{u}_i$ and $\overline{\overline{u}}_i$ to be, respectively, the i-th elements of u, $\overline{u} = W_{NT}u$ and $\overline{\overline{u}} = W_{NT}^2 u$. Similarly, we define ϵ_i and $\overline{\epsilon}_i$ to be the i-th elements of ϵ and $\overline{\epsilon} = W_{NT}\epsilon$. We can then write the above equations as follows:

$$u_i - \rho_2 \overline{u}_i = \epsilon_i, \qquad i = 1, ..., NT \tag{2.19}$$

$$\overline{u}_i - \rho_2 \overline{\overline{u}}_i = \overline{\epsilon}_i, \qquad i = 1, ..., NT \tag{2.20}$$

By squaring and then summing both equations, multiplying (2.19) by (2.20) and summing and finally dividing all terms by NT, the following system of equations is constructed:

$$2\rho_2 \frac{1}{NT}\sum u_i\overline{u}_i - \rho_2^2 \frac{1}{NT}\sum \overline{u}_i^2 + \frac{1}{NT}\sum \epsilon_i^2 = \frac{1}{NT}\sum u_i^2 \tag{2.21}$$

$$2\rho_2 \frac{1}{NT}\sum \overline{u}_i\overline{\overline{u}}_i - \rho_2^2 \frac{1}{NT}\sum \overline{\overline{u}}_i^2 + \frac{1}{NT}\sum \overline{\epsilon}_i^2 = \frac{1}{NT}\sum \overline{u}_i^2 \tag{2.22}$$

$$\rho_2 \frac{1}{NT}\sum (u_i\overline{\overline{u}}_i + \overline{u}_i^2) - \rho_2^2 \frac{1}{NT}\sum \overline{u}_i\overline{\overline{u}}_i + \frac{1}{NT}\sum \epsilon_i\overline{\epsilon}_i = \frac{1}{NT}\sum u_i\overline{u}_i \tag{2.23}$$

As a next step, expectations are taken across the system and the terms involving ϵ replaced as follows:

$$E\left(\frac{1}{NT}\sum \epsilon_i^2\right) = \sigma_\epsilon^2 \tag{2.24}$$

$$E\left(\frac{1}{NT}\sum \overline{\epsilon}_i^2\right) = \frac{1}{NT}\sigma_\epsilon^2 tr(W'_{NT}W_{NT}) \tag{2.25}$$

$$E\left(\frac{1}{NT}\sum \epsilon_i\overline{\epsilon}_i\right) = 0, \tag{2.26}$$

where $tr(.)$ denotes the trace operator. While (2.24) and (2.25) follow from assumption (2.3), (2.26) additionally makes use of the fact that all diagonal elements of W_{NT} are zero.

The system of equations can now be written as follows:

$$\Gamma(\rho_2, \rho_2^2, \sigma_\epsilon^2)' - \gamma \quad = \quad 0 \tag{2.27}$$

where

$$\Gamma \quad = \quad \frac{1}{NT}\begin{pmatrix} 2E(u'\overline{u}) & -E(\overline{u}'\overline{u}) & NT \\ 2E(\overline{\overline{u}}'\overline{u}) & -E(\overline{\overline{u}}'\overline{\overline{u}}) & tr(W_{NT}'W_{NT}) \\ 2E(u'\overline{\overline{u}} + \overline{u}'\overline{u}) & -E(\overline{u}'\overline{\overline{u}}) & 0 \end{pmatrix} \tag{2.28}$$

$$\gamma \quad = \quad \frac{1}{NT}\begin{pmatrix} E(u'u) \\ E(\overline{u}'\overline{u}) \\ E(u'\overline{u}) \end{pmatrix} \tag{2.29}$$

The expectations are next replaced by their sample analogues using $\tilde{u}$ from the two-stage least squares regression in the first estimation step. Equation (2.27) can then be solved by nonlinear least squares to get the estimators $\hat{\rho}_2$ and $\hat{\sigma}_\epsilon^2$.

Bibliography

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Studies in Operational Regional Science 4, Kluwer Academic Publishers, Dordrecht; Boston.

Anselin, L. and Bera, A. K. (1998). Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. In *Handbook of Applied Economic Statistics* (U. A. Giles D, ed.). Marcel Dekker, New York, 237–289.

Anselin, L., le Gallo, J. and Jayet, H. (2008). Spatial Panel Econometrics. In *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice* (L. Mátyás and P. Sevestre, eds.), 2008th ed. Advanced Studies in Theoretical and Applied Econometrics, Springer, Berlin.

Anselin, L. and Rey, S. (1991). Properties of Tests for Spatial Dependence in Linear-Regression Models. *Geographical Analysis* **23** 112–131.

Armingeon, K. (1999). Politische Reaktionen auf Steigende Arbeitslosigkeit. In *Nationaler Staat und Internationale Wirtschaft* (A. Busch and T. Plümper, eds.). Nomos, Baden-Baden, 169–196.

Armstrong, H. and Taylor, J. (1993). *Regional Economics and Policy*. Blackwell, Oxford.

Brügger, B., Lalive D'Epinay, R. and Zweimüller, J. (2007). Regionale Disparitäten in der Arbeitslosigkeit: Kulturelle Grenzen und Landesgrenzen. Tech. rep., SECO.

Büttner, T. (1999). *Agglomeration, Growth, and Adjustment a Theoretical and Empirical Study of Regional Labor Markets in Germany*. Physica-Verlag, Heidelberg Etc.

Burridge, P. (1980). Cliff-Ord Test for Spatial Correlation. *Journal of the Royal Statistical Society Series B-Methodological* **42** 107–108.

BURRIDGE, P. and GORDON, I. R. (1981). Unemployment in the British Metropolitan Labour Areas. *Oxford Economic Papers* **33** 274–97.

CLIFF, A. D. and ORD, J. K. (1973). *Spatial Autocorrelation.* Monographs in Spatial and Environmental Systems Analysis, 5, Pion, London.

CLIFF, A. D. and ORD, J. K. (1981). *Spatial Processes: Models & Applications.* Pion, London.

DE COULON, A. (1999). Disparité Régionale du Chômage et Courbe de Beveridge en Suisse. Tech. rep., University of Geneva, Department of Political Economy.

ELHORST, J. P. (1995). Unemployment Disparities Between Regions in the European Union. In *Convergence and Divergence Among Euopean Unions* (H. W. Armstrong and R. Vickerman, eds.). Pion, London.

ELHORST, J. P. (2001). Dynamic Models in Space and Time. *Geographical Analysis* **33** 119–140.

ELHORST, J. P. (2003). The Mystery of Regional Unemployment Differentials: Theoretical and Empirical Explanations. *Journal of Economic Surveys* **17** 709–748.

EUROSTAT (2009). EUROIND Database.

FELD, L. P. and SAVIOZ, M. R. (2000). Cantonal and Regional Unemployment in Switzerland – A Dynamic Macroeconomic Panel Analysis. *Swiss Journal of Economics and Statistics (SJES)* **136** 463–483.

FILIPPINI, P. (1998). *Regional Unemployment Disparities – The Case of the Swiss Cantons.* vdf Hochsch.-Verlag an der ETH, Zürich.

FILIPPINI, P. and ROSSI, A. (1992). Unemployment in the Swiss Economy – A Border Region Phenomenon? *Aussenwirtschaft* **47** 497–513.

FILIPPINI, P. and ROSSI, A. (1993). Ausländische Arbeitnehmer und Regionale Arbeitslosigkeitsdisparitäten. *Schweizerische Zeitschrift für Volkswirtschaft und Statistik* **3** 357–369.

FLÜCKIGER, Y., COULON, A. D. and VASSILIEV, A. (2002). *Les Raisons de la Différence entre les Taux de Chômage Genevois et Suisse.* LEA Université de Genève, Genève.

Flückiger, Y., Kempeneers, P. and Bazen, S. (2007a). Analyse der Regionalen Unterschiede in der Arbeitslosigkeit. *SECO Publikationen* **22** 115.

Flückiger, Y., Kempeneers, P. and Bazen, S. (2007b). Regionale Unterschiede bei der Arbeitslosigkeit: Ein Revidierter Ansatz. *Die Volkswirtschaft* **80** 57–61.

Flückiger, Y. and Morales, D. (1998). *Analyse des Causes de l' augmentation Actuelle du Chômage en Suisse et à Genève*. LEA Université de Genève, Genève.

Flückiger, Y. and Vassiliev, A. (2002). Les Raisons des Différences de Chômage entre Genève et le Reste de la Suisse. *Swiss Journal of Economics and Statistics (SJES)* **138** 387–410.

Florax, R. J. G. M. and de Graaff, T. (2004). The Performance of Diagnostic Tests for Spatial Dependence in Linear Regression Models – A Meta-Analysis of Simulation Studies. In *Advances in Spatial Econometrics: Methodology, Tools and Application* (L. Anselin, R. Florax and S. Rey, eds.). Springer, Heidelberg, 29–65.

Forni, L. (2002). Public Sector Employment and the Unemployment Rate: Evidence from OECD Countries.

Giacomini, R. and Granger, C. W. J. (2004). Aggregation of Space-Time Processes. *Journal of Econometrics* **118** 7–26.

Hofler, R. A. and Murphy, K. J. (1989). Using a Composed Error Model to Estimate the Frictional and Excess-Supply Components of Unemployment. *Journal of Regional Science* **29** 213–228.

Johnson, J. A. and Kneebone, R. D. (1991). Deriving Natural Rates of Unemployment for Sub-National Regions – The Case of Canadian Provinces. *Applied Economics* **23** 1305–1314.

Kelejian, H. H., Murrell, P. and Shepotylo, O. (2008). Spatial Spillovers in the Development of Institutions. *Electronic Working Papers, Department of Economics, University of Maryland* **07-001**.

Kelejian, H. H. and Prucha, I. R. (1998). Generalized Spatial Two-stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. *Journal of Real Estate Finance and Economics* **17** 99–121.

Kelejian, H. H. and Prucha, I. R. (2007). The Relative Efficiencies of Various Predictors in Spatial Econometric Models Containing Spatial Lags. *Regional Science and Urban Economics* **37** 363–374.

Kelejian, H. H. and Robinson, D. P. (1993). A Suggested Method of Estimation for Spatial Interdependent Models. *Papers in Regional Science* **72** 297–312.

Kleinewefers Lehner, A. (2001). Regionale Unterschiede auf dem Schweizer Arbeitsmarkt im Konjunkturzyklus der 90er Jahre.

KOF Swiss Economic Institute (2008). Effects of the Bilateral Agreements on Swiss Potential GDP. *KOF Studies* **2**.

Kosfeld, R. and Dreger, C. (2006). Thresholds for Employment and Unemployment – A Spatial Analysis of German Regional Labour Markets, 1992-2000. *Papers in Regional Science* **85** 523–542.

Lewin, R. (1983). Niveau und Struktur der Arbeitslosigkeit in der Schweiz 1973-83. *Quartalshefte der Schweizerischen Nationalbank* **5**.

Longhi, S. and Nijkamp, P. (2007). Forecasting Regional Labor Market Developments under Spatial Autocorrelation. *International Regional Science Review* **30** 100–119.

Lopez-Bazo, E., Del Barrio, T. and Artis, M. (2002). The Regional Distribution of Spanish Unemployment – A Spatial Analysis. *Papers in Regional Science* **81** 365–389.

Malizia, E. E. and Ke, S. Z. (1993). The Influence of Economic Diversity on Unemployment and Stability. *Journal of Regional Science* **33** 221–235.

Martin, R. (1997). Regional Unemployment Disparities and their Dynamics. *Regional Studies* **31** 237–252.

Molho, I. (1995a). Migrant Inertia, Accessibility and Local-Unemployment. *Economica* **62** 123–132.

Molho, I. (1995b). Spatial Autocorrelation in British Unemployment. *Journal of Regional Science* **35** 641–658.

Neumann, G. R. and Topel, R. H. (1991). Employment Risk, Diversification, and Unemployment. *Quarterly Journal of Economics* **106** 1341–1365.

Oud, H. L., Johan, Folmer, H., Patuelli, R. and Nijkamp, P. (2008). A Spatial-Dependence Continuous-Time Model for Regional Unemployment in Germany. Tech. rep., Rimini Centre for Economic Analysis.

Parnisari, B. (2003). Anhaltende Regionale und Sektorielle Unterschiede bei der Arbeitslosigkeit. *Die Volkswirtschaft* **76** 39–43.

Partridge, M. D. and Rickman, D. S. (1995). Differences in State Unemployment Rates – The Role of Labor and Product Market Structural Shifts. *Southern Economic Journal* **62** 89–106.

Partridge, M. D. and Rickman, D. S. (1997). The Dispersion of US State Unemployment Rates – The Role of Market and Non-market Equilibrium Factors. *Regional Studies* **31** 593–606.

Projer, E. (1993). Die Arbeitslosigkeit in der Schweiz im Intertemporalen Vergleich. *Quartalshefte der Schweizerischen Nationalbank* **1**.

Schanne, N., Wapler, R. and Weyh, A. (2008). Regional Unemployment Forecasts with Spatial Interdependencies. Tech. rep., Institut für Arbeitsmarkt und Berufsforschung (IAB), Nürnberg.

Sheldon, G. (1999). *Beschäftigungsorientierte Optionen der Schweizer Arbeitsmarktpolitik. Schlussbericht zu einer Studie im Auftrag der Jubiläumsstiftung Schweizerischer Bankverein 1997.* Basel.

Simon, C. J. (1988). Frictional Unemployment and the Role of Industrial Diversity. *Quarterly Journal of Economics* **103** 715–728.

Spiezia, V. (2003). Measuring Regional Economies. *Statistics Brief* **6**.

Stalder, P. (2008). Personenfreizügigkeit: Auswirkungen auf den Arbeitsmarkt und das Wirtschaftswachstum. *Die Volkswirtschaft* **11/2008** 7–11.

Steffen, I. (2005). Determinanten der Arbeitslosigkeit in den Schweizer Kantonen – Eine Empirische Untersuchung zu den Disparitäten in den Kantonalen Arbeitslosenquoten. *Swiss Political Science Review* **11** 27–54.

Taylor, J. and Bradley, S. (1983). Spatial Variations in the Unemployment Rate – A Case-Study of Northwest England. *Regional Studies* **17** 113–124.

Taylor, J. and Bradley, S. (1997). Unemployment in Europe – A Comparative Analysis of Regional Disparities in Germany, Italy and the UK. *Kyklos* **50** 221–245.

Chapter 3

Modeling Swiss Consumption in a DSGE Framework[1]

Contents

[1]I am grateful to Gregor Bäurle for letting me use his Matlab Code and to Stefan Leist for introducing me to the code. Furthermore, I thank the participants of the Brown Bag Seminar at the Department of Economics, University of Berne, for helpful comments.

3.1 Introduction

Private consumption expenditure accounts for roughly 60% of Switzerland's gross domestic product (GDP). Although this share is larger in other countries (e.g., the United Kingdom or the United States), consumption is an important part of GDP and has delivered growth contributions to the annual real Swiss GDP of 0.9 percentage points on average between 1980 and 2010. During the recession caused by the financial crisis, consumption proved robust and prevented Swiss GDP from dropping even deeper. Although private consumption causes less fluctuations in GDP than investment, an appropriate modeling of consumption is an important ingredient of macroeconomic analysis.

The dominant framework for macroeconomic analysis are DSGE models. The corresponding literature is still growing fast, and only a small number of studies on DSGE models for Switzerland have been published so far. Three recent studies estimate variants of the small open economy model presented by Galí and Monacelli (2005) using Swiss data: Beltran and Draper (2008) investigate parameter identification, Leist and Neusser (2010) establish estimates of the output gap and Bäurle and Menz (2008) use a DSGE-VAR framework in order to analyse whether the central bank should react to movements in the nominal exchange rate. The Swiss National Bank uses a more complex DSGE model in its monetary policy process which is documented in Cuche-Curti et al. (2009).

This paper examines two extensions of the standard new Keynesian model with respect to consumption: habit formation and rule-of-thumb consumers. The main motivation for using habit formation is the reproduction of the smoothness found in consumption data. This smoothness is a consequence of the gradual response of consumption to shocks. Under habit formation, households compare current consumption to a habit stock which depends on past consumption. As a consequence, households derive utility not only from the consumption level, but also from consumption growth. Maximization of lifetime utility then implies households to smooth both the consumption level and the change in consumption over time. Two types of habit formation, external and internal habit formation, are commonly used in the literature. The two types differ in the way in which households take into account the effect of their decisions on future utility levels: under internal habit formation, households anticipate that a higher consumption level today will increase the habit stock and lower utility tomorrow. Under external habit formation, households ignore this aspect. The mentioned papers on DSGE models for Switzerland all use some type of habit formation: while Beltran and Draper (2008), Bäurle and Menz (2008) and Leist and Neusser (2010) use external habit formation, Cuche-Curti et al. (2009) use internal habit formation.

Rule-of-thumb consumers are mainly used in two macreoeonomic contexts. First, they have been introduced in consumption models in order to remedy for some of the discrepancies between the implications of the permanent income hypothesis and the empirical evidence. In particular, rule-of-thumb consumers can explain the sensitivity of consumption to current income. Second, they have been used in the literature on fiscal multipliers in order to reproduce the empirical impulse response functions of consumption to government spending shocks.

In this paper, I introduce habit formation and rule-of-thumb consumers in a small DSGE model which I estimate using Bayesian methods on Swiss data. The goal of the paper is twofold: first, I try to estimate the degree of habit persistence and the share of rule-of-thumb consumers. Second, I investigate the effect of habit formation and rule-of-thumb consumers on the likelihood and posterior probability of the model. In the section on habit formation, I estimate three versions of the model: a version without habit formation, a version with external habit formation and a version with internal habit formation. In the section on rule-of-thumb consumers, I consider two alternative labor market structures, namely perfectly and imperfectly competitive labor markets. Under both labor market structures, I estimate model versions with and without rule-of-thumb consumers.

Although all of the mentioned DSGE models for Switzerland use some type of habit formation, there is, to my knowledge, no systematic analysis of its usefulness and the optimal type with respect to Swiss data. Moreover, none of the papers use consumption data in the estimation process. With respect to rule-of-thumb consumers, estimates from DSGE models are generally rare and to my knowledge, there are none available for Switzerland.

I find the data to support both external and internal habit formation. The estimated habit persistence parameter lies at 0.54 for external habit formation and at 0.61 for internal habit formation. Both are significantly different from zero. Moreover, allowing for habit formation improves the likelihood as well as the posterior probability of the model, with external habit formation reaching the best performance. With respect to the share of rule-of-thumb consumers, I find a value of 0.65 under perfectly competitive labor markets and a value of 0.53 under imperfectly competitive labor markets. For the variant with perfectly competitive labor markets, the parameter estimates as well as the likelihood and posterior probability support the concept of rule-of-thumb consumers. For the variant with imperfectly competitive labor markets, the results are less clear-cut: the parameter estimates are not perfectly robust and the posterior probability does not allow for a clear acceptance or rejection of rule-of-thumb consumers.

The remainder of the paper is organized as follows: section 3.2 introduces the baseline model and the estimation technique. Sections 3.3 and 3.4 motivate the use of habit formation and rule-of-thumb consumers, respectively, discuss their implementation in the model and present estimation results. Section 3.5 concludes.

3.2 Baseline Model and Estimation Technique

3.2.1 The Baseline Model

This section presents the baseline DSGE model which will be extended in sections 3.3 and 3.4. The baseline model thus contains neither habit formation nor rule-of-thumb consumers. The model is an adapted version of the model presented by Del Negro and Schorfheide (2008), which is itself a simplified version of the small open economy model presented by Galí and Monacelli (2005). In this model, the world economy consists of two countries, namely the small open economy and the foreign economy summarizing the rest of the world. Both economies produce consumption goods. The production sector of the domestic economy is monopolistically competitive and consists of a continuum of firms producing intermediate goods. A perfectly competitive retail sector buys these intermediate goods, packs them into units of the domestic good and sells them on the domestic market as well as abroad. Another perfectly competitive retail sector then buys units of the domestic good and of the foreign good, packs them into units of the aggregate consumption good and sells them to households. These two retail sectors need not be modeled explicitly as they are assumed to make zero profits. Furthermore, I assume the law of one price to hold, i.e. the price which domestic retailers pay for the foreign good is equal to the price which foreign retailers pay once the nominal exchange rate is taken into account. Monetary policy is conducted according to a Taylor rule containing the output cycle, output growth and inflation. The variables describing the state of the foreign economy are modeled exogenously.

The Representative Household

The domestic economy is populated by a representative household which consumes, invests and supplies labor to domestic firms. In the baseline version of the model, the household's preferences are given by

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{\varepsilon}_{p,t} \left(\log\left(C_t\right) - \frac{N_t^{1+\varphi}}{1+\varphi} \right)$$

where $\tilde{\varepsilon}_{p,t}$ is a preference shock and C_t denotes aggregate consumption and N_t hours worked. Furthermore, β denotes the discount factor and φ the inverse elasticity of labor supply. Aggregate consumption is an index of domestic and foreign consumption goods, as discussed below.

The household's budget constraint is given by

$$P_tC_t + \frac{1}{R_t}D_{t+1} + \frac{\varepsilon_t}{R_t^*}D^*_{t+1} = W_tP_tN_t + D_t + \varepsilon_t D^*_t + \Pi_t$$

where P_t denotes the consumer price index (CPI), D_t and D_t^* denote holdings of domestic and foreign bonds, R_t and R_t^* denote the correponding interest rates, ε_t denotes the nominal exchange rate, W_t the real wage level and Π_t profits from domestic producers. The household problem is solved by maximizing the following lagrangian:

$$\mathcal{L} = \mathbb{E}_0 \sum_{t=0}^{\infty} \left[\begin{array}{c} \beta^t \tilde{\varepsilon}_{p,t} \left(\log\left(C_t\right) - \frac{N_t^{1+\varphi}}{1+\varphi} \right) \\ -\lambda_t \left(P_tC_t + \frac{1}{R_t}D_{t+1} + \frac{\varepsilon_t}{R_t^*}D^*_{t+1} - W_tP_tN_t - D_t - \varepsilon_t D^*_t - \Pi_t \right) \end{array} \right]$$

The derivative of the lagrangian with respect to consumption is

$$\frac{\partial \mathcal{L}}{\partial C_t} = \beta^t \tilde{\varepsilon}_{p,t} \frac{1}{C_t} - \lambda_t P_t$$

and the first order conditions for the household problem are

$$\begin{aligned} N_t^{\varphi} C_t &= W_t \\ \tilde{\varepsilon}_{p,t} \frac{1}{C_t} &= \beta \mathbb{E}_t \left[\tilde{\varepsilon}_{p,t+1} \frac{1}{C_{t+1}} R_t \frac{P_t}{P_{t+1}} \right] \\ 0 &= \mathbb{E}_t \left[\left(R_t - R_t^* e_{t+1}\right) \frac{\tilde{\varepsilon}_{p,t+1}}{\tilde{\varepsilon}_{p,t}} \frac{C_t}{C_{t+1}} \frac{P_t}{P_{t+1}} \right] \end{aligned}$$

where $e_t = \varepsilon_t/\varepsilon_{t-1}$ denotes the depreciation rate of the home currency.

Aggregation of Consumption Goods

Aggregate consumption is a composite index built according to

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

where $C_{H,t}$ and $C_{F,t}$ denote units of the domestic and foreign good, α is the weight of foreign goods in total consumption and η is the elasticity of substitution between domestic and foreign goods. Assuming a perfectly competitive retail sector which buys units of $C_{H,t}$ and $C_{F,t}$, packs them into units of C_t and sells these packages to households yields the following demand functions:

$$C_{H,t} = (1-\alpha)\left(\frac{P_{H,t}}{P_t}\right)^{-\eta} C_t, \qquad C_{F,t} = \alpha\left(\frac{P_{F,t}}{P_t}\right)^{-\eta} C_t \tag{3.1}$$

where $P_{H,t}$ and $P_{F,t}$ are the prices of the home and the foreign good. Assuming zero profits for this retail sector yields the price index

$$P_t = \left[(1-\alpha)P_{H,t}^{1-\eta} + \alpha P_{F,t}^{1-\eta}\right]^{\frac{1}{1-\eta}} \tag{3.2}$$

Details on the derivation of the above demand functions and of the price index are given in the appendix.

The overall production level of the domestic consumption good, denoted by $Y_{C,t}$, is itself an index of intermediate goods $Y_{C,t}(i)$ produced by domestic firms:

$$Y_{C,t} = \left[\int_0^1 Y_{C,t}(i)^{\frac{\varepsilon-1}{\varepsilon}} di\right]^{\frac{\varepsilon}{\varepsilon-1}} \tag{3.3}$$

where ε is the elasticity of substitution between the intermediate goods.

Assuming a perfectly competitive retail sector which buys domestic intermediate goods $Y_{C,t}(i)$, packs them into units of $Y_{C,t}$ and sells the packages to the retailers which aggregate $C_{H,t}$ and $C_{F,t}$ results in the demand function

$$Y_{C,t}(i) = \left(\frac{P_{H,t}(i)}{P_{H,t}}\right)^{-\varepsilon} Y_{C,t} \tag{3.4}$$

where $P_{H,t}(i)$ denotes the price charged by firm i. Assuming zero profits for this retail sector yields the price index

$$P_{H,t} = \left[\int_0^1 P_{H,t}(i)^{1-\varepsilon} di\right]^{\frac{1}{1-\varepsilon}}$$

Details on the derivation are given in the appendix as well.

Domestic Firms

The domestic production sector consists of a continuum of firms which produce intermediate goods according to the production function

$$Y_{C,t}(i) = Z_t N_t(i) \tag{3.5}$$

where Z_t denotes the technology level common to all firms and $N_t(i)$ denotes labor input of firm i. The production sector is monopolistically competitive with a price setting mechanism à la Calvo, augmented with partial indexation. In each period, only a fraction $1-\theta$ of firms is allowed to reset their prices. The fraction θ of firms which are not able to reset their prices instead adjust their prices according to last period's inflation rate in domestic prices:

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

where δ denotes the degree to which these firms index their prices to past inflation. A firm resetting its price maximizes the present value of future profits, taking into account the demand function (3.4) as well as the possibility that it might not be able to reset its price for a number of periods. As all firms face an identical problem, they all set the same price. The present value of future profits is given by

$$\mathbb{E}_t\left[\sum_{\tau=0}^{\infty}\theta^{\tau}Q_{t+\tau|t}Y_{C,t+\tau|t}\left(P'_{H,t}\left(\frac{P_{H,t+\tau-1}}{P_{H,t-1}}\right)^{\delta} - P_{H,t+\tau}MC_{t+\tau}\right)\right]$$

where $P'_{H,t}$ denotes the price set by firms which optimize in period t, $Y_{C,t+\tau|t}$ the quantity sold by these firms, $MC_t = (W_t P_t)/(P_{H,t}Z_t)$ are real marginal costs and $Q_{t+\tau|t}$ is a stochastic discount factor. As the representative household owns the firms, the discount factor is equal to the household's intertemporal rate of substitution and can be expressed in terms of the Lagrange multiplier by $Q_{t+\tau|t} = \lambda_{t+\tau}/\lambda_t$. Using equation (3.4), $Y_{C,t+\tau|t}$ can be written as

$$Y_{C,t+\tau|t} = \left(\frac{P'_{H,t}}{P_{H,t+\tau}}\left(\frac{P_{H,t+\tau-1}}{P_{H,t-1}}\right)^{\delta}\right)^{-\varepsilon} Y_{C,t+\tau} \tag{3.6}$$

The first order condition for the firm's maximization problem is then given by

$$\mathbb{E}_t\left[\sum_{\tau=0}^{\infty}\theta^{\tau}Q_{t+\tau|t}Y_{C,t+\tau|t}\left(P'_{H,t}\left(\frac{P_{H,t+\tau-1}}{P_{H,t-1}}\right)^{\delta}-\frac{\varepsilon}{\varepsilon-1}P_{H,t+\tau}MC_{t+\tau}\right)\right]=0 \qquad (3.7)$$

Under this setup, the price index for domestic goods evolves according to

$$P_{H,t}=\left[(1-\theta)\left(P'_{H,t}\right)^{1-\varepsilon}+\theta\left(P_{H,t-1}\left(\frac{P_{H,t-1}}{P_{H,t-2}}\right)^{\delta}\right)^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}} \qquad (3.8)$$

Trade

The terms of trade are defined as

$$Q_t=\frac{P_{H,t}}{P_{F,t}}$$

where $P_{H,t}$ denotes the price of the domestic good and $P_{F,t}$ the price of the foreign good, both measured in units of the domestic currency. The law of one price is assumed to hold, i.e. the price which domestic retailers pay for the foreign good is equal to the price which is paid abroad once the exchange rate is taken into account:

$$P_{F,t}=\varepsilon_t P^*_{F,t}$$

where $P^*_{F,t}$ denotes the price of the foreign good measured in units of the foreign currency.

Foreign demand for the domestic good, denoted by $C^*_{H_t}$, is defined in analogy to domestic demand for the foreign good:

$$C^*_{H_t}=\alpha^*\left(\frac{P^*_{H,t}}{P^*_t}\right)^{-\eta}C^*_t \qquad (3.9)$$

As the domestic economy is small in relation to the foreign economy, home goods have a negligible weight in foreign consumption. If ϑ denotes the relative size of the domestic economy, α^* can be written as $\alpha^*=\vartheta\alpha$, where $\vartheta\to 0$. This means that $P^*_{F,t}$ will be approximately equal to the foreign CPI and the terms of trade can be expressed as

$$Q_t=\frac{P_{H,t}}{\varepsilon_t P^*_t}$$

The real exchange rate is defined as

$$S_t = \varepsilon_t \frac{P_t^*}{P_t}$$

The relative price $P_{H,t}/P_t$, which determines the competitivenes of domestic goods in the domestic market, can be expressed using the terms of trade and the real exchange rate:

$$\frac{P_{H,t}}{P_t} = Q_t S_t \tag{3.10}$$

Furthermore, using equation (3.2) and the law of one price, the real exchange rate can be expressed directly by the terms of trade:

$$S_t = \left[(1-\alpha) Q_t^{1-\eta} + \alpha\right]^{\frac{1}{\eta-1}} \tag{3.11}$$

The Foreign Economy

The foreign economy is not modeled structurally. Instead, the foreign variables needed for the model, the production level and inflation, are modeled as autoregressive processes. The specifications will be given in the section on the linearized version of the model.

Market Clearing

The market for domestic consumption goods clears if production matches domestic and foreign demand:

$$Y_{C,t} = C_{H,t} + C_{H,t}^*$$

The overall output level, Y_t, contains an additional residual:

$$Y_t = Y_{C,t} + Y_{res,t}$$

Adding this residual allows me to use both consumption and GDP data in the estimation of the model. Using consumption data is necessary given the focus of this study. In principle, the overall output level could directly be linked to consumption – this would be consistent as only consumption goods are produced in the model's economy. However, using consumption data for the overall output level would be not be consistent with using data on the 3-month LIBOR for the interest rate: estimating a Taylor rule would not make sense because the

Swiss National Bank (SNB) reacts (apart from inflation) to overall GDP and not just to consumption. Using both consumption and GDP as well as the above residual assures consistency. The residual contains investment, government spending and net exports of investment goods, commodities, intermediate goods and energy sources, such that it adds up to overall GDP with private consumption and exports of consumption goods. The residual is modeled using an autoregressive process. I assume that the residual is produced in analogy to the intermediate consumption goods according to a production function in the fashion of equation (3.5). This is necessary in order to maintain the relationship between overall output, the overall amount of hours worked and the technology level. However, the price setting behaviour of the implicit production sector of the residual is neglected.

Inserting into the market clearing condition yields

$$\begin{aligned} Y_t &= C_{H,t} + C^*_{H,t} + Y_{res,t} \\ &= (1-\alpha)\left(\frac{P_{H,t}}{P_t}\right)^{-\eta} C_t + \alpha^* \left(\frac{P^*_{H,t}}{P^*_t}\right)^{-\eta} C^*_t + Y_{res,t} \\ &= (1-\alpha)\left(Q_t S_t\right)^{-\eta} C_t + \alpha^* \left(Q_t\right)^{-\eta} C^*_t + Y_{res,t} \end{aligned} \tag{3.12}$$

Log-Linear Approximation to the Model

This section presents the log-linear approximation of the model around a zero-inflation steady state. In the following, capital letters without time index denote steady states, whereas small letters denote logarithmic deviations of variables from their steady states, i.e. $x_t = \log(X_t/X)$. Details on the linearization of the model are given in the appendix.

The Euler equation has the following form:

$$c_t = \mathbb{E}_t\left[c_{t+1}\right] - \left(r_t - \mathbb{E}_t\left[\pi_{t+1}\right]\right) + \mathbb{E}_t\left[\varepsilon_{p,t} - \varepsilon_{p,t+1}\right] \tag{3.13}$$

where π_t and $\varepsilon_{p,t}$ denote the logarithmic deviations of overall inflation and the preference shock from their respective steady states. Inflation in domestic prices, $\pi_{H,t}$, evolves according to the following Phillips curve:

$$\pi_{H,t} - \delta\pi_{H,t-1} = \beta\mathbb{E}_t\left[\pi_{H,t+1} - \delta\pi_{H,t}\right] + \kappa mc_t \tag{3.14}$$

with $\kappa = (1-\beta\theta)(1-\theta)/\theta$. Real marginal costs are given by

$$mc_t = c_t + \varphi y_t - (1+\varphi)\, z_t - \alpha q_t \tag{3.15}$$

By approximating foreign consumption, c_t^*, by foreign output ,y_t^*, the market clearing condition linearizes to

$$\begin{aligned} y_t = (1-\alpha)\frac{C}{Y}c_t - \eta\left(\alpha(1-\alpha)\frac{C}{Y} + \frac{C_H^*}{Y}\right)q_t + \frac{C_H^*}{Y}y_t^* \\ + \left(1-(1-\alpha)\frac{C}{Y} - \frac{C_H^*}{Y}\right)y_{res,t} \end{aligned} \tag{3.16}$$

The central bank sets the short-term interest rate according to the following Taylor rule:

$$r_{t+1} = \rho_r r_t + \psi_\pi \pi_{t+1} + \psi_{\Delta y}\Delta y_{t+1} + \psi_y y_{t+1} + \varepsilon_{M,t+1} \tag{3.17}$$

By including both Δy and y_{t+1}, I do not restrict the central bank to react either to the output cycle or to output growth. The real exchange rate can be expressed directly by the terms of trade using equation (3.11). Therefore, the change in the nominal exchange rate can be connected to the terms of trade as follows:

$$e_t = -(1-\alpha)\Delta q_t + \pi_t - \pi_t^* \tag{3.18}$$

Using equation (3.10), overall CPI inflation can be constructed from inflation in domestic prices and the terms of trade:

$$\pi_t = \pi_{H,t} - \alpha\Delta q_t \tag{3.19}$$

The terms of trade are in principle endogenous to the model. However, I follow Del Negro and Schorfheide (2008) and Leist and Neusser (2010) in modeling them exogenously as an autoregressive process in differences because they are difficult to explain by theoretical models. Furthermore, the technology level (z_t), foreign output (y_t^*), foreign inflation (π_t^*), the preference shock ($\varepsilon_{p,t}$) and the GDP residual ($y_{res,t}$) are also modeled as autoregressive processes. The monetary policy shock ($\varepsilon_{M,t}$) is assumed to be *iid*. This results in the following exogenous processes:

$$\Delta q_{t+1} = \Delta\rho_q q_t + \varepsilon_{q,t+1} \tag{3.20}$$

$$z_{t+1} = \rho_z z_t + \varepsilon_{z,t+1} \tag{3.21}$$

$$y_{t+1}^* = \rho_y y_t^* + \varepsilon_{y,t+1} \tag{3.22}$$

$$\pi_{t+1}^* = \rho_\pi \pi_t^* + \varepsilon_{\pi,t+1} \tag{3.23}$$

$$\varepsilon_{p,t+1} = \rho_p \varepsilon_{p,t} + \epsilon_{p,t+1} \tag{3.24}$$

$$y_{res,t+1} = \rho_{res} y_{res,t} + \varepsilon_{res,t+1} \tag{3.25}$$

Equations (3.13) to (3.25) describe the model's state. In order to solve the model, it has to be rewritten in the following form:

$$\Gamma_0 X_{t+1} = \Gamma_1 X_t + \Psi Z_t + \Pi \eta_t$$

where X_t contains the model variables at time t, $Z_t = (\varepsilon_{M,t}, \varepsilon_{q,t}, \varepsilon_{z,t}, \varepsilon_{y,t}, \varepsilon_{\pi,t}, \epsilon_{p,t}, \varepsilon_{res,t})'$ contains the structural shocks and η_t contains expectational errors. The model can then be solved using Sim's method (Sims, 2001). The model solution describes the transition of the model in response to the vector of shocks:

$$X_{t+1} = \Theta_1 X_t + \Theta_z Z_t$$

3.2.2 Estimation Technique

The parameters of the various model versions are estimated using a Markov Chain Monte Carlo method. Bayesian techniques have in general become the standard in the estimation of DSGE models. Apart from the appealing probabilistic interpretation of the model parameters, Bayesian techniques facilitate the use of prior information about the parameters (see, e.g., DeJong and Dave, 2007). A practical reason for their dominance is that they can also cope well with estimations of high-dimensional dynamic models, a task which poses computational problems to maximum likelihood estimation (Fernández-Villaverde, 2009).

The goal of the estimation process is to estimate the posterior distribution of the model parameters. According to Bayes' rule, the posterior distribution $P(\Theta|Y)$ is proportional to the product of the prior distribution, $\pi(\Theta)$, and the likelihood of the data, L $(Y|\Theta)$:

$$P(\Theta|Y) \propto \pi(\Theta) L(Y|\Theta)$$

where Y denotes the data and Θ the model parameters.

I use a random walk Metropolis-Hastings algorithm in order to estimate the model. The algorithm aims at constructing a Markov chain whose distribution converges to the posterior distribution of the parameters. The principle of the algorithm is described for example in An and Schorfheide (2007). I use the specific implementation described in Bäurle and Menz (2008) which consists of the following steps:

Step 1: Maximize the posterior probability using a numerical optimization procedure.[2] The maximizing parameter set Θ_0 serves as initial value for step 2.

Step 2 (iteration): Generate candidate draws Θ_i^* according to $\Theta_i^* = \Theta_{i-1} + \epsilon_i$, where ϵ_i has a multivariate t distribution with the number of degrees of freedom equal to the number of estimated parameters. Use the inverse hessian at the mode from step 1 as initial covariance matrix of ϵ_i. During the iteration, re-scale the covariance matrix on a regular basis such that the acceptance rate of candidate draws is between 0.2 and 0.3. The probability that a specific candidate draw Θ_i^* is accepted is given by the ratio of the posterior probabilities of the model under Θ_i^* and Θ_{i-1}:

$$q\left(\Theta_i^*, \Theta_{i-1}\right) = \min\left(\frac{\pi\left(\Theta_i^*\right)}{\pi\left(\Theta_{i-1}\right)}\frac{L\left(Y|\Theta_i^*\right)}{L\left(Y|\Theta_{i-1}\right)}, 1\right)$$

The likelihood is calculated using the Kalman filter. The choice of accepting or rejecting Θ_i^* is made by comparing $q\left(\Theta_i^*, \Theta_{i-1}\right)$ to a draw ς from a uniform distribution over the interval $[0, 1]$. If $q\left(\Theta_i^*, \Theta_{i-1}\right) > \varsigma$ then $\Theta_i = \Theta_i^*$, otherwise $\Theta_i = \Theta_{i-1}$.

The parameters have converged when their means over the sequence remain approximately constant (some little variation will always remain, however). In order to judge the convergence of the sequence, I use the graphical plots of the means over the sequence. In order to eliminate the influence of the initial parameter set Θ_0 on the results, I delete the first 20 percent of elements from the sequence. Furthermore, Θ_i is saved only after every 16th draw for computational reasons.

3.2.3 Priors

The prior distributions for the parameters of the baseline model are listed in table 3.1. α, the share of foreign goods in the CPI, and C/Y and C_H^*/Y, the shares of consumption and exports in GDP, are fixed as they are known from the data. The mean of the share of foreign goods in the CPI over the estimation sample lies at 0.27. The means of the shares of consumption and exports in GDP lie at 60 percent and 40 percent, respectively. Furthermore, β, the subjective discount rate, and ρ_p, the autoregressive coefficient of the preference shock, are fixed in order to avoid identification problems. The mean interest

[2] I use Sims' line search algorithm for this step. The algorithm is downloadable from Christopher Sims' homepage as Matlab or R code: http://sims.princeton.edu/yftp/optimize

rate lies at 3.6% in the estimation sample and β is fixed at 0.991 according to the formula $\beta = exp(-r/400)$. ρ_p is fixed at 0.9 as it oscillated around this value in earlier model versions but did not converge well.

The choice of the priors of the parameters of the Phillips curve, κ and δ, is oriented on prior studies using Swiss Data (Leist and Neusser, 2010, Bäurle and Menz, 2008). Accordingly, the parameter of price stickiness, θ, is expected approximately at 0.6, which corresponds to prices being fixed on average for 2 to 3 quarters and to a value of about 0.3 for κ. Furthermore, indexation has been found to play no dominant role for domestic inflation. The priors for the parameters of the Taylor rule are chosen such that the central bank reacts with regard to output primarily to output growth and not to the output cycle. This is a pragmatic step taken because the parameter estimates converged faster and the models reached better likelihoods than in versions where the central bank reacted more strongly to the cycle. Furthermore, note that the response of the central bank to inflation must have a certain strength in order to avoid indeterminate model solutions (see section 3.4.3 for a discussion). In order to allow for the possibility that the central bank does not react to the output cycle nor to output growth, or only to one of the two, I assume the parameters ψ_y and $\psi_{\Delta y}$ to be normally distributed but nonnegative.

The prior means of the autoregressive parameters for the interest rate (ρ_r), foreign output (ρ_{y^*}) and foreign inflation (ρ_π) are set to 0.5 as the corresponding time series can be expected to be moderatly autocorrelated, while a higher persistence can be expected for the productivity level. With respect to the standard deviations of the shocks to the exogenous processes, the priors for σ_M, σ_q and σ_{y^*} are set according to results of OLS regressions, whereas for the unobservable productivity and preference shock, the priors for the standard deviations are set to 1. As foreign inflation is not connected to the data, $\varepsilon_{\pi,t}$ serves primarily as a residual for the nominal exchange rate in the model. Accordingly, σ_π can be expected in the magnitude of the standard deviation of the nominal exchange rate rather than in the magnitude of the standard deviation of some measure of foreign inflation. Finally, σ_{res} can be expected in the magnitude of the standard deviations of the cycles of investment, government spending and exports and imports of goods. In general, the standard deviations of the σ parameters were chosen rather high in order to avoid slow convergence due only to these parameters.

Table 3.1: *Prior Distributions of Parameters for the Baseline Model*

Parameter	Distribution	Mean	Standard Deviation
β	calibrated	0.99	
α	calibrated	0.27	
κ	Gamma	0.30	0.30
δ	Normal on [0,1]	0.10	0.05
φ	Gamma	2.50	1.00
η	Gamma	2.50	1.00
ρ_r	Beta	0.50	0.20
ψ_π	Gamma	1.50	0.05
ψ_y	Normal > 0	0.10	0.05
$\psi_{\Delta y}$	Normal > 0	0.40	0.05
ρ_q	Normal	0.00	0.50
ρ_z	Normal	0.80	0.50
ρ_{y^*}	Normal	0.50	0.50
ρ_{π^*}	Normal	0.50	0.50
ρ_p	calibrated	0.90	
ρ_i	Normal	0.80	0.50
C/Y	calibrated	0.60	
C_H^*/Y	calibrated	0.40	
σ_q	Inv.Gamma	1.50	1.50
σ_r	Inv.Gamma	0.50	0.50
σ_z	Inv.Gamma	1.00	1.00
σ_{y^*}	Inv.Gamma	0.50	0.50
σ_{π^*}	Inv.Gamma	5.00	3.50
σ_p	Inv.Gamma	1.00	3.50
σ_{res}	Inv.Gamma	5.00	3.50

3.2.4 Data

The model is estimated using quarterly data from 1980Q2 to 2010Q4. As the variables of the linearized model are measured in logarithmic deviations from their steady states, all time series are demeaned. The variables of the small open economy are connected to Swiss data, the variables of the foreign economy to data from the European Union. For Switzerland, the time series used are GDP, private consumption, the ratio of the export and import price deflators, overall CPI inflation, the 3-month LIBOR and the nominal effective exchange rate calculated against 24 trade partners. For the European Union, GDP is connected to output in the foreign economy. Table 3.2 gives details on the transformations and sources of the data.

Model variables are usually connected to the data either in levels or differences. In terms of the variables of the linearized model, the first difference of a variable is equal to the first difference of the logarithm of the original absolute variable (i.e., $\Delta x_t = \Delta \log(X_t)$). Therefore, the first difference in the variables of the linearized model corresponds approximately to the growth rate of the original variables (i.e., $\Delta x_t \approx \Delta X_t / X_{t-1}$). Accordingly, some studies connect certain model variables in differences to the corresponding empirical growth rates. In this paper's notation, this corresponds for example to connecting Δy_t to GDP growth. In contrast to this approach, I connect all variables except the exchange rate in levels to the data. In order to measure the deviations from steady state in the data, I run an HP-filter on the time series and calculate cycles by taking the percentage deviations of the original series from the filtered series. These cycles are then connected to the variables of the linearized model. Using levels instead of differences makes sense because the variables are modeled in levels. Moreover, some variables were connected in differences to the data in earlier versions of the paper. Although the model fitted the differences of the variables well to the empirical growth rates, it generated implausible paths for the levels of some variables. Therefore, the nominal effective exchange rate is the only variable to be connected in differences to the data as it is modeled directly in differences.

The HP-filter is often critizised because of its instability at ending points. In order to eliminate this problem, I cut the estimation sample at the end of 2010. Consequently, I consider the instability problem to be of minor importance for the estimation of the model. If the model were, however, to be used for forecasting purposes, it would be necessary to pay greater attention to the appropriate measerement of the cycle at the ending points.

Table 3.2: *Data Used for Estimation*

Time Series	Description	Transformation	Relation to model	Source
CPI Inflation	Inflation in the overall consumer price index	100*dlog(CPI)	π	SFSO
LIBOR	3-Month LIBOR		$4r$	SNB
Exports Deflator (PX)				Seco
Imports Deflator (PM)				Seco
Terms of Trade (ToT)	ToT = PX/PM	cycle	q_t	
Nominal Effective Exchange Rate	Nominal exchange rate against 24 countries	100*dlog(neer)	e_t	SNB
Private Consumption	at prices of preceding year, chained values, ref. year 2000	cycle	c_t	Seco
GDP	at prices of preceding year, chained values, ref. year 2000	cycle	y_t	Seco
GDP European Union (EU27)	chained values, ref. year 2000	cycle	y_t^*	Eurostat

NOTE: All Series except the LIBOR and the nominal effective exchange rate are seasonally adjusted.
Cycle: Percentage Deviation of the seasonally adjusted series from the HP-filtered series
SNB: Swiss National Bank
SFSO: Swiss Federal Statistical Office
Seco: Swiss State Secretariat for Economic Affairs

3.2.5 Results

The parameters of the baseline model converged after generating 1'000'000 candidate draws. The posterior means of the parameters for the baseline version of the model are listed in table 3.3. The mean of κ lies at 3.68, corresponding to a price stickiness parameter (θ) of 0.18 and to prices being fixed for timespans between 1 und 2 quarters. The price stickiness thus turns out smaller than expected. Conversely, the degree of indexation of prices to past inflation, δ, turns out very low, confirming findings of earlier studies. The resulting means for φ, the inverse of the elasticity of labor supply, and η, the elasticity of substitution between domestic and foreign goods, both lie clearly below their prior means. The results for the coefficients of the Taylor rule show a strong response of the monetary authority to inflation. With regard to output, the results stay close to the priors and imply a reaction to output growth but not to the output cycle. The results for the autoregressive coefficients of the exogenous processes imply the change in the terms of trade and foreign inflation to be weakly autocorrelated. While the GDP residual shows moderate autocorrelation, the technology level and foreign output show a high persistence.

In order to study the business cycle dynamics which the model creates, it is instructive to calculate the impulse response functions of the endogenous variables. Figure 3.1 shows the impulse responses of consumption, output, the interest rate and inflation to shocks to the interest rate and the productivity level as well as the corresponding 95-percent confidence intervals. The confidence intervals were calculated by drawing parameter sets from the posterior distributions and calculating the responses of the endogenous variables to the shocks under these parameter sets. By repeating this step 25'000 times, the confidence intervals can be obtained by taking the 5-percent and 95-percent quantiles of the resulting 25'000 impulse responses.

An unexpected increase in the interest rate lowers consumption and output as consumers postpone their expenses into future periods. The lower output level then dampens inflation. The confidence intervals of the corresponding impulse response functions are very small as the interest rate enters the Euler equation with a fixed factor in this model version. A positive shock to the productivity level increases productivity for a longer period; due to the high persistence of the productivity level, it takes long for the effect to wear out. Accordingly, the reactions of the other variables also vanish very slowly. The higher productivity level lowers inflationary pressure and allows the central bank to lower the interest rate. As a consequence, consumption and output rise. The confidence intervals of the impulse responses to a productivity shock are rather large. This uncertainity can be explained by the large

Table 3.3: *Posterior Distributions of Parameters for the Baseline Model*

Parameter	Prior Mean	Posterior Mean	S.D.
κ	0.30	3.68	0.44
δ	0.10	0.05	0.03
φ	2.50	0.24	0.08
η	2.50	0.22	0.05
ρ_r	0.50	0.23	0.07
ψ_π	1.50	1.48	0.05
ψ_y	0.10	0.02	0.01
$\psi_{\Delta y}$	0.40	0.29	0.05
ρ_q	0.00	0.12	0.03
ρ_z	0.80	0.99	0.01
ρ_{y^*}	0.50	0.87	0.04
ρ_{π^*}	0.50	0.25	0.10
ρ_{res}	0.80	0.66	0.03
σ_q	1.50	2.07	0.26
σ_r	0.50	0.37	0.06
σ_z	1.00	0.31	0.04
σ_{y^*}	0.50	0.23	0.03
σ_{π^*}	5.00	3.91	0.46
σ_p	1.00	0.94	0.18
σ_{res}	5.00	8.84	1.14
Log Likelihood		–939.20	
Log Posterior Probability		–980.01	
Candidate Draws		1’000’000	

standard deviation of the parameter κ. In other words, the uncertainity arises from the uncertain effect of productivity on domestic inflation.

3.3 Habit Formation

This section motivates the use of habit formation, discusses the changes in the model structure and presents estimation results. Under habit formation, households wish to smooth the changes in their consumption level. As a consequence, they react gradually to shocks and their intertemporal consumption path is smoothed. I estimate a model version with external habit formation and a version with internal habit formation. The resulting parameters of habit persistence are significantly different from zero. While both models reach a better likelihood and posterior probability than the baseline model from section 3.2, the version with external habit formation performs best.

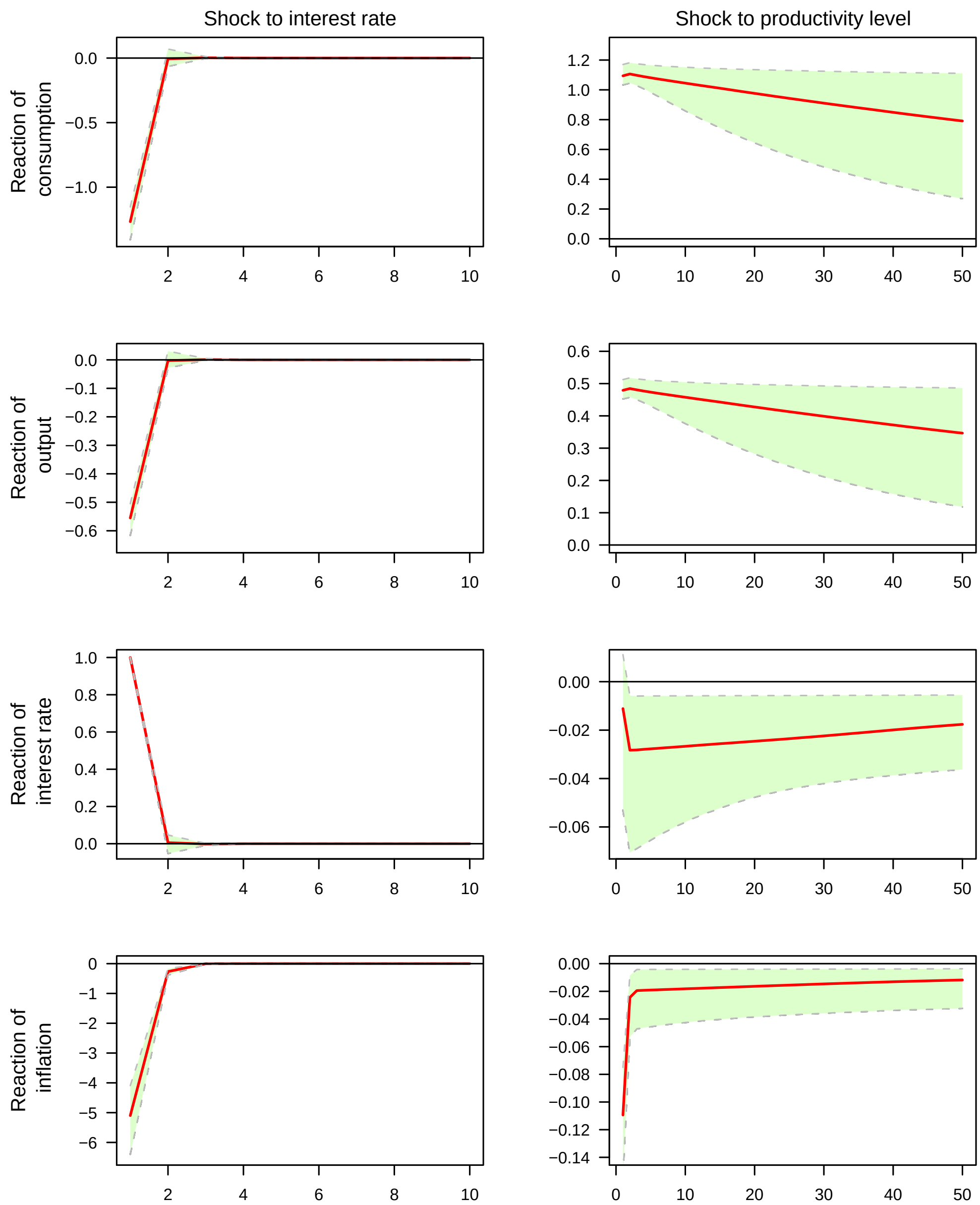
3.3.1 Motivation

From the perspective of macroeconomic modeling, the main reason for introducing habit formation into models is the reproduction of the smoothness found in consumption data. This smoothness is a consequence of the gradual response of consumption to shocks. By default, theory-based macroeconomic models generate immediate reactions to shocks, resulting in volatile consumption paths.

The gradual response of consumption to shocks can be observed in empirical impulse response functions. For example, Leeper et al. (1996) and Christiano et al. (2005) document gradual responses to monetary policy shocks for consumption and a number of other macroeconomic variables using identified VAR approaches on US data. Studies documenting impulse response functions specifically for Swiss consumption are rare. One exception is Kaufmann (1996), who documents the gradual impulse response of consumption to permanent shocks. More recently, the gradual response of Swiss overall GDP to shocks has for example been documented using a structural VAR model in Kugler and Jordan (2004), or using using a structural cointegrated VAR model in Assenmacher-Wesche (2008).

These gradual responses found in the data are not consistent with the permanent income hypothesis (PIH). The latter suggests that households adapt their consumption plan in each period to expected lifetime income such that consumption is smoothed over time. As a consequence, consumption should “jump” immediately in response to news which is relevant for permanent income (Fuhrer, 2000). Hall (1978) pointed out that the PIH

Figure 3.1: *Impulse Responses for the Baseline Model*



Note: Impulse Responses to a one percent interest rate shock (left) and to a productivity shock (right) of magnitude 1 with 95-percent confidence intervals

actually implies that consumption is a random walk: all information that is relevant for future income should be fully included in consumption at any point in time. Therefore, no variable apart from lagged consumption should be useful in forecasting consumption. Under models strictly consistent with the PIH, re-optimization of the consumption path in each period leads to volatile consumption time series, although households try to smooth their future consumption path in each period.

Actual consumption data is thus smoother than the PIH would suggest. Campbell and Deaton (1989) call this the "excess smoothness" of consumption. Moreover, they note a second inconsistency with the PIH, namely that consumption is correlated to lagged income. This "excess sensitivity" is another consequence of the gradual response of consumption to shocks: as the response to shocks takes some time, a fraction of current consumption is determined by past events, as Sommer (2007) notes. Campbell and Deaton (1989) argue that one possible way to model both "excess smoothness" and "excess sensitivity" is habit formation.

Under habit formation, households' current utility depends not only on current consumption, but also on a habit stock, which depends on past consumption. The households compare current consumption to this habit stock. Current utility thus rises with current consumption, but falls with past consumption. In this setting, households derive utility not only from the current consumption level itself, but also from consumption growth. Under utility functions with decreasing marginal utility of consumption, households then wish to smooth not only their consumption level over time, but also the change in consumption. This smoothing produces the gradual impulse responses observed in the data.

Apart from macroeconomics, habit formation is also used in other branches of economics. For example, the asset-pricing literature uses habit formation to explain the equity premium puzzle (see, e.g., Abel, 1990). The latter refers to the fact that the difference between returns on equity and risk-free assets is persistently larger than standard models can explain under plausible extents of risk aversion (Mehra and Prescott, 1985). Under habit formation, the large risk premium can be explained by the fact that consumers dislike strong drops in consumption and therefore require a higher premium for holding risky assets. Moreover, Campbell and Cochrane (1999) show that models featuring habit formation can explain several other empirical features of asset prices.

Further applications of habit formation can be found for example in the growth literature, where habit formation helps to explain the empirical finding that high growth leads to high saving (Carroll et al., 2000). This finding is otherwise inconsisent with the PIH, which suggests that households save less if they expect high income resulting from high growth. In another application, Uribe (2002) shows that under habit formation, exchange-rate-based stabilization programs induce a comovement of consumption and real exchange rates.

3.3.2 Implementation

Throughout the rest of this section, I consider three model versions: the baseline version without habit formation, a version with external habit formation and a version with internal habit formation. In the following, I discuss the implementation of habit formation in the households' preferences as well as the difference between external and internal habit formation.

Habit Formation in the Household's Maximization Problem

Recall the household's preferences in the baseline version of the model:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{\varepsilon}_{p,t} \left(\log C_t - \frac{N_t^{1+\varphi}}{1+\varphi} \right)$$

If we allow for habit formation, the preferences take the following form:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{\varepsilon}_{p,t} \left(\log \left(C_t - hC_{t-1} \right) - \frac{N_t^{1+\varphi}}{1+\varphi} \right)$$

where h is the parameter of habit persistence and the term hC_t is called the habit stock. The degree of habit persistence determines the respective weights of the consumption level and consumption growth for utility. For $h = 0$, only the current consumption level matters to the household, while for $h = 1$, only the change in consumption matters; for $0 < h < 1$, both the level and the change matter. It is always possible to return to the baseline version of the model by setting h to zero.

The lagrangian for the maximization of the household problem then takes the following form:

$$\mathcal{L} = \mathbb{E}_0 \sum_{t=0}^{\infty} \left[\begin{array}{c} \beta^t \tilde{\varepsilon}_{p,t} \left(\log \left(C_t - hC_{t-1} \right) - \frac{N_t^{1+\varphi}}{1+\varphi} \right) \\ -\lambda_t \left(P_t C_t + \frac{1}{R_t} D_{t+1} + \frac{\varepsilon_t}{R_t^*} D_{t+1}^* - W_t P_t N_t - D_t - \varepsilon_t D_t^* - \Pi_t \right) \end{array} \right]$$

Internal and External Habit Formation

The habit stock in the utility function can be interpreted in two ways: it can either be seen as an exogenous reference level of consumption to which the household compares his own consumption level, or as the household's own consumption level from the one period before. These two types of habit formation, generally called external and internal habit formation, generate different model structures.

In the case of external habit formation, the household takes lagged consumption as given. As in "catching up with the Joneses" preferences (see, e.g., Ljungqvist and Uhlig, 2000), lagged consumption is regarded as somebody else's or the economy-wide consumption level, to which the household compares his own consumption level. The habit stock has thus the function of an exogenous reference level of consumption. In the case of internal habit formation, the household takes into account that the lagged consumption term is part of his own maximization problem. Therefore, when deciding about the current level of consumption, the household anticipates that this decision has an impact on future utility. While a higher current consumption level increases today's utility, it lowers tomorrow's utility. As the household anticipates effects on future utility only in the case of internal habit formation, the difference between external and internal habit formation can also be attributed to the household being either myopic or rational (Muellbauer, 1988).

Both external and internal habit formation are widely used in the literature. For example, Fuhrer (2000) and Christiano et al. (2005) use internal habit formation, while Smets and Wouters (2003) and Ravn et al. (2006) use external habit formation. In DSGE models for Switzerland, both types have been used as well: Beltran and Draper (2008), Bäurle and Menz (2008) and Leist and Neusser (2010), which all use adapted versions of the model from Galí and Monacelli (2005), use external habit formation, whereas Cuche-Curti et al. (2009) use internal habit formation. Moreover, studies which directly compare both types reach different conclusions: while Muellbauer (1988) and Korniotis (2010) favor external habit formation, Dennis (2004) finds a slightly better fit for internal habit formation. Furthermore, Dennis (2009) finds the choice between internal and external habit formation to be unimportant for the generated business cycle behavior in new Keynesian models. He sees this as a possible reason for both types being largely used in business cycle models.

Optimality Conditions for the three Model Versions

The two types of habit formation imposed in the utility function lead to different derivatives with respect to consumption and therefore to different first order conditions for the household's maximization problem.

Case 1 – no habit formation: For the baseline version of the model without habit formation, the derivative of the lagrangian with respect to consumption is

$$\frac{\partial \mathcal{L}}{\partial C_t} = \beta^t \tilde{\varepsilon}_{p,t} \frac{1}{C_t} - \lambda_t P_t$$

and the first order conditions for the household problem are

$$\begin{aligned} N_t^{\varphi} C_t &= W_t \\ \tilde{\varepsilon}_{p,t} \frac{1}{C_t} &= \beta \mathbb{E}_t \left[\tilde{\varepsilon}_{p,t+1} \frac{1}{C_{t+1}} R_t \frac{P_t}{P_{t+1}} \right] \\ 0 &= \mathbb{E}_t \left[(R_t - R_t^* e_{t+1}) \frac{\tilde{\varepsilon}_{p,t+1}}{\tilde{\varepsilon}_{p,t}} \frac{C_t}{C_{t+1}} \frac{P_t}{P_{t+1}} \right] \end{aligned}$$

where $e_t = \varepsilon_t / \varepsilon_{t-1}$ denotes the depreciation rate.

Case 2 – external habit formation: For the version with external habit formation, the derivative of the lagrangian is

$$\frac{\partial \mathcal{L}}{\partial C_t} = \beta^t \tilde{\varepsilon}_{p,t} \frac{1}{C_t - hC_{t-1}} - \lambda_t P_t$$

and the first order conditions for the household problem are

$$\begin{aligned} N_t^{\varphi} (C_t - hC_{t-1}) &= W_t \\ \tilde{\varepsilon}_{p,t} \frac{1}{C_t - hC_{t-1}} &= \beta \mathbb{E}_t \left[\tilde{\varepsilon}_{p,t+1} \frac{1}{C_{t+1} - hC_t} R_t \frac{P_t}{P_{t+1}} \right] \\ 0 &= \mathbb{E}_t \left[(R_t - R_t^* e_{t+1}) \frac{\tilde{\varepsilon}_{p,t+1}}{\tilde{\varepsilon}_{p,t}} \frac{C_t - hC_{t-1}}{C_{t+1} - hC_t} \frac{P_t}{P_{t+1}} \right] \end{aligned}$$

Case 3 – internal habit formation: in the case of internal habit formation, the derivative of the lagrangian is

$$\frac{\partial \mathcal{L}}{\partial C_t} = \beta^t \tilde{\varepsilon}_{p,t} \frac{1}{C_t - hC_{t-1}} - h\beta^{t+1} \mathbb{E}_t \left[\tilde{\varepsilon}_{p,t+1} \frac{1}{C_{t+1} - hC_t} \right] - \lambda_t P_t$$

The middle part of the derivative represents the effect of current consumption on next period's utility. The first order conditions for the household problem are

$$N_t^{\varphi}\left(\frac{1}{C_t-hC_{t-1}}-h\beta\mathbb{E}_t\left[\frac{\tilde{\varepsilon}_{p,t+1}}{\tilde{\varepsilon}_{p,t}}\frac{1}{C_{t+1}-hC_t}\right]\right)^{-1}=W_t$$

$$\begin{aligned}\tilde{\varepsilon}_{p,t}\frac{1}{C_t-hC_{t-1}}-h\beta\mathbb{E}_t\left[\tilde{\varepsilon}_{p,t+1}\frac{1}{C_{t+1}-hC_t}\right]&\\ =\beta\mathbb{E}_t\left[\left(\tilde{\varepsilon}_{p,t+1}\frac{1}{C_{t+1}-hC_t}-h\beta\tilde{\varepsilon}_{p,t+2}\frac{1}{C_{t+2}-hC_{t+1}}\right)R_t\frac{P_t}{P_{t+1}}\right]&\\ 0=\mathbb{E}_t\left[\left(R_t-R_t^*e_{t+1}\right)\frac{P_t}{P_{t+1}}\left(\frac{\tilde{\varepsilon}_{p,t+1}}{\tilde{\varepsilon}_{p,t}}\frac{C_t-hC_{t-1}}{C_{t+1}-hC_t}-h\beta\frac{\tilde{\varepsilon}_{p,t+2}}{\tilde{\varepsilon}_{p,t+1}}\frac{C_{t+1}-hC_t}{C_{t+2}-hC_{t+1}}\right)\right]&\end{aligned}$$

Internal and External Habit Formation in the Linearized Model

The three model versions analyzed in this section differ with respect to the Euler equation and the expression for marginal costs. The Euler equation has the following form:

Case 1 – no habit formation:

$$c_t=\mathbb{E}_t\left[c_{t+1}\right]-\left(r_t-\mathbb{E}_t\left[\pi_{t+1}\right]\right)+\mathbb{E}_t\left[\varepsilon_{p,t}-\varepsilon_{p,t+1}\right] \tag{3.26a}$$

Case 2 – external habit formation:

$$\begin{aligned}\left(c_t-hc_{t-1}\right)=\mathbb{E}_t\left[c_{t+1}-hc_t\right]-\left(1-h\right)\left(r_t-\mathbb{E}_t\left[\pi_{t+1}\right]\right)&\\ +\left(1-h\right)\mathbb{E}_t\left[\varepsilon_{p,t}-\varepsilon_{p,t+1}\right]&\end{aligned} \tag{3.26b}$$

Case 3 – internal habit formation:

$$\begin{aligned}\left(c_t-hc_{t-1}\right)-h\beta\mathbb{E}_t\left[c_{t+1}-hc_t\right]=&\ \mathbb{E}_t\left[c_{t+1}-hc_t\right]-h\beta\mathbb{E}_t\left[c_{t+2}-h\beta c_{t+1}\right]\\ &-\left(1-h\right)\left(1-h\beta\right)\left(r_t-\mathbb{E}_t\left[\pi_{t+1}\right]\right)\\ &+\left(1-h\right)\mathbb{E}_t\left[\varepsilon_{p,t}-\varepsilon_{p,t+1}-h\beta\left(\varepsilon_{p,t+1}-\varepsilon_{p,t+2}\right)\right]\end{aligned} \tag{3.26c}$$

The structures of the expression for marginal costs are

Case 1 – no habit formation:

$$mc_t=c_t+\varphi y_t-\left(1+\varphi\right)z_t-\alpha q_t \tag{3.27a}$$

Case 2 – external habit formation:

$$mc_t = \frac{1}{(1-h)}c_t - \frac{h}{(1-h)}c_{t-1} + \varphi y_t - (1+\varphi)\, z_t - \alpha q_t \tag{3.27b}$$

Case 3 – internal habit formation:

$$\begin{aligned} mc_t = & \frac{1}{(1-h)(1-h\beta)}c_t - \frac{h\beta}{(1-h)(1-h\beta)}\mathbb{E}_t\left[c_{t+1}\right] \\ & - \frac{h}{(1-h)(1-h\beta)}c_{t-1} + \varphi y_t - (1+\varphi)\, z_t - \alpha q_t \\ & - \frac{h\beta}{1-h\beta}\left(\varepsilon_{p,t} - \mathbb{E}_t\left[\varepsilon_{p,t+1}\right]\right) \end{aligned} \tag{3.27c}$$

Habit Persistence

The degree of habit persistence, measured by the parameter h, determines the smoothness of the consumption path generated by the model. The closer h is to one, the more delayed is the reaction of consumption to shocks and the smoother is the consumption path. In order to reproduce the empirical impulse response function as close as possible, h should be attributed a high value, e.g. 0.9. However, an exact reproduction of impulse response functions is not the only priority of most empirical studies. While early DSGE models where calibrated in order to reproduce impulse response functions and other stylized facts, models are today normally estimated with the goal of a full characterization of the data generating process (Smets and Wouters, 2003)[3]. Accordingly, models are parametrized in order to reproduce a set of time series as close as possible, but the parameters are not necessarily those which reproduce the stylized facts optimally. Specifically with respect to habit persistence, Sommer (2007) argues that estimators of h are generally biased downwards because consumption data contains measurement errors and other transitory elements. In his study on US data, the habit persistence parameter jumps from 0.35 to 0.89 once measurement error is controlled for.

As it can be expected from the arguments above, estimations of the habit persistence parameter are not always close to 1, but cover nearly the whole range between 0 and 1. For example, using US-data, Christiano et al. (2005) obtain an estimate of 0.65, while Fuhrer (2000) obtains 0.8 and 0.9 using two different estimation techniques. For Switzerland, those

[3]See Geweke (2007) and Smets and Wouters (2003) for further discussions of these different approaches of parametrizing DSGE models.

studies which estimate similar models than the one in this paper reach estimates of 0.08 (Bäurle and Menz, 2008), 0.23 (Leist and Neusser, 2010) and 0.99 (Beltran and Draper, 2008). Cuche-Curti et al. (2009) calibrate the parameter at 0.7. There is thus no standard result for the parameter of habit persistence.

3.3.3 Priors

The prior distributions of the parameters are listed in table 3.4. They are identical to those used for the baseline model, except for the additional parameter h. As there is no standard result for the parameter of habit persistence, I set a relatively neutral Beta prior distribution with a mean of 0.5 and a standard deviation of 0.2.

3.3.4 Results

The means and standard deviations of the posterior distributions of the two habit formation models are listed in table 3.5. The parameters of both models converged after generating 3'000'000 candidate draws.

Habit Persistence

The mean of the habit persistence parameter, h, converges to 0.54 in the model version with external habit formation and to 0.61 in the model with internal habit formation. Accordingly, consumption growth enters the utility function by 54% in the case of external habit formation and by 61% in the case of internal habit formation. In both cases, h has a low standard deviation of 0.05; the parameter is therefore significantly different from zero. This is also visible in the posterior plots of this parameter (see figure 3.2). In order to assess whether the choice of the prior distribution for h affects the posterior mean, I run the estimations also with prior means of 0.25 and 0.75. In the case of external habit formation, h converges to 0.50 (prior mean 0.25) and 0.56 (prior mean 0.75). For internal habit formation, h converges to 0.51 (prior mean 0.25) and 0.69 (prior mean 0.75). While these posterior means are not exactly equal to those obtained with the original prior mean of 0.5, they lie in the same order of magnitude and are still clearly different from zero. The estimation results for h thus support both types of habit formation.

Table 3.4: *Prior Distributions of Parameters for the Habit Formation Models*

Parameter	Distribution	Mean	Standard Deviation
h	Beta	0.50	0.20
β	calibrated	0.99	
α	calibrated	0.27	
κ	Gamma	0.30	0.30
δ	Normal on [0,1]	0.10	0.05
φ	Gamma	2.50	1.00
η	Gamma	2.50	1.00
ρ_r	Beta	0.50	0.20
ψ_π	Gamma	1.50	0.05
ψ_y	Normal > 0	0.10	0.05
$\psi_{\Delta y}$	Normal > 0	0.40	0.05
ρ_q	Normal	0.00	0.50
ρ_z	Normal	0.80	0.50
ρ_{y^*}	Normal	0.50	0.50
ρ_{π^*}	Normal	0.50	0.50
ρ_p	calibrated	0.90	
ρ_i	Normal	0.80	0.50
C/Y	calibrated	0.60	
C_H^*	calibrated	0.40	
σ_q	Inv.Gamma	1.50	1.50
σ_r	Inv.Gamma	0.50	0.50
σ_z	Inv.Gamma	1.00	1.00
σ_{y^*}	Inv.Gamma	0.50	0.50
σ_{π^*}	Inv.Gamma	5.00	3.50
σ_p	Inv.Gamma	1.00	3.50
σ_{res}	Inv.Gamma	5.00	3.50

Figure 3.2: *Posterior Plots for the Habit Persistence Parameter*

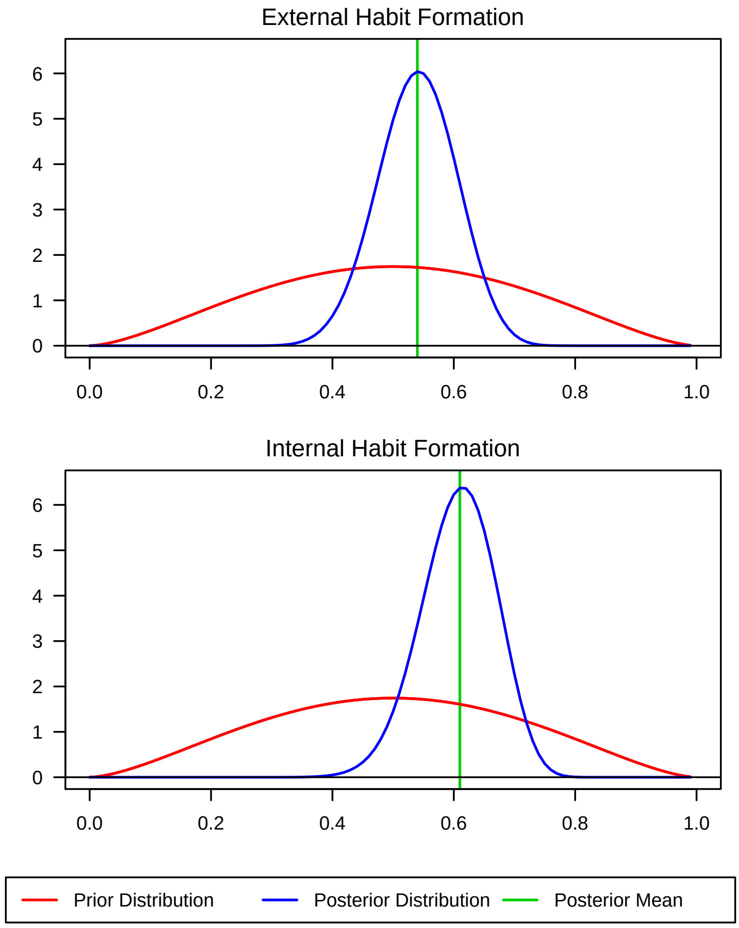


Table 3.5: *Posterior Distributions of Parameters for the Baseline and Habit Formation Models*

		No HF		External HF		Internal HF	
Parameter	Prior Mean	Mean	S.D.	Mean	S.D.	Mean	S.D.
h	0.50	0.00[a]		0.54	0.05	0.61	0.05
κ	0.30	3.68	0.44	2.33	0.31	1.16	0.08
δ	0.10	0.05	0.03	0.03	0.02	0.02	0.02
φ	2.50	0.24	0.08	0.33	0.11	0.46	0.13
η	2.50	0.22	0.05	0.27	0.05	0.26	0.05
ρ_r	0.50	0.23	0.07	0.29	0.07	0.51	0.08
ψ_π	1.50	1.48	0.05	1.46	0.05	1.43	0.05
ψ_y	0.10	0.02	0.01	0.01	0.01	0.01	0.01
$\psi_{\Delta y}$	0.40	0.29	0.05	0.30	0.05	0.31	0.05
ρ_q	0.00	0.12	0.03	0.12	0.03	0.20	0.03
ρ_z	0.80	0.99	0.01	1.00	0.00	0.99	0.01
ρ_{y^*}	0.50	0.87	0.04	0.88	0.04	0.89	0.04
ρ_{π^*}	0.50	0.25	0.10	0.25	0.09	0.25	0.10
ρ_{res}	0.80	0.66	0.03	0.98	0.02	0.92	0.04
σ_q	1.50	2.07	0.26	2.09	0.29	2.07	0.25
σ_r	0.50	0.37	0.06	0.37	0.06	0.40	0.07
σ_z	1.00	0.31	0.04	0.75	0.21	0.86	0.25
σ_{y^*}	0.50	0.23	0.03	0.23	0.03	0.23	0.03
σ_{π^*}	5.00	3.91	0.46	3.60	0.38	3.96	0.53
σ_p	1.00	0.94	0.18	1.20	0.22	3.00	0.58
σ_{res}	5.00	8.84	1.14	6.86	0.68	7.20	0.91
Log Likelihood		–939.20		–917.16		–931.32	
Log Posterior Probability		–980.01		–952.96		–962.24	
Posterior Odds Ratio for h=0				$1.78 * 10^{-12}$		$1.90 * 10^{-8}$	
Candidate Draws		1'000'000		3'000'000		3'000'000	

NOTE: a – calibrated; HF – Habit Formation

Other Model Parameters

The three model versions differ with respect to the Euler equation and the expression for marginal costs. Accordingly, the estimations produce not only different values for the habit persistence parameter, but also for the parameter κ. While the mean of κ lies at 3.68 in the version without habit formation, the versions with habit formation produce smaller values. In the case of external habit formation, κ comes to lie at 2.33, implying a θ of 0.24. The version with internal habit formation produces a κ of 1.45, implying a θ of 0.35. The latter still corresponds to prices being fixed more often than every second quarter on average. The price stickiness thus turns out smaller than expected in all three model versions. On the other hand, the degree of indexation to past inflation, δ, lies below 0.1 for all three versions.

Most of the remaining model parameters turn out similar across the model versions. While, φ, the inverse elasticity of labor supply turns out to be higher for the versions with habit formation, the posterior mean for η, the elasticity of substitution between domestic and foreign goods, as well as the results for the Taylor rule, are very similar for the three versions. On the other hand, there are differences for σ_p, the standard deviation of the innovations to the preference shock. These differences can be explained by the varying model structures: the preference shock enters the Euler equation with different coefficients. Furthermore, the preference shock also enters the Phillips curve in the model with internal habit formation.

Model Comparison

Statistics on the goodness of fit of the three model versions are given at the bottom of table 3.5. According to the log likelihood and the log posterior probability, the models allowing for habit formation are able to fit the data better than the model without habit formation. This is in line with the habit persistence parameter being significantly above zero. Furthermore, the posterior odds ratio between the model without habit formation and those allowing for habit formation can be taken as a criterion for the null hypothesis $h = 0$. The posterior odds ratio clearly favors the models allowing for habit formation. Comparing the log likelihood and log posterior probability of the models with external and internal habit formation shows that the model with external habit formation reaches a better fit of the data. Together with the finding that the habit persistence parameter is significantly different from zero, these results support the concept of habit formation.

Impulse Response Functions

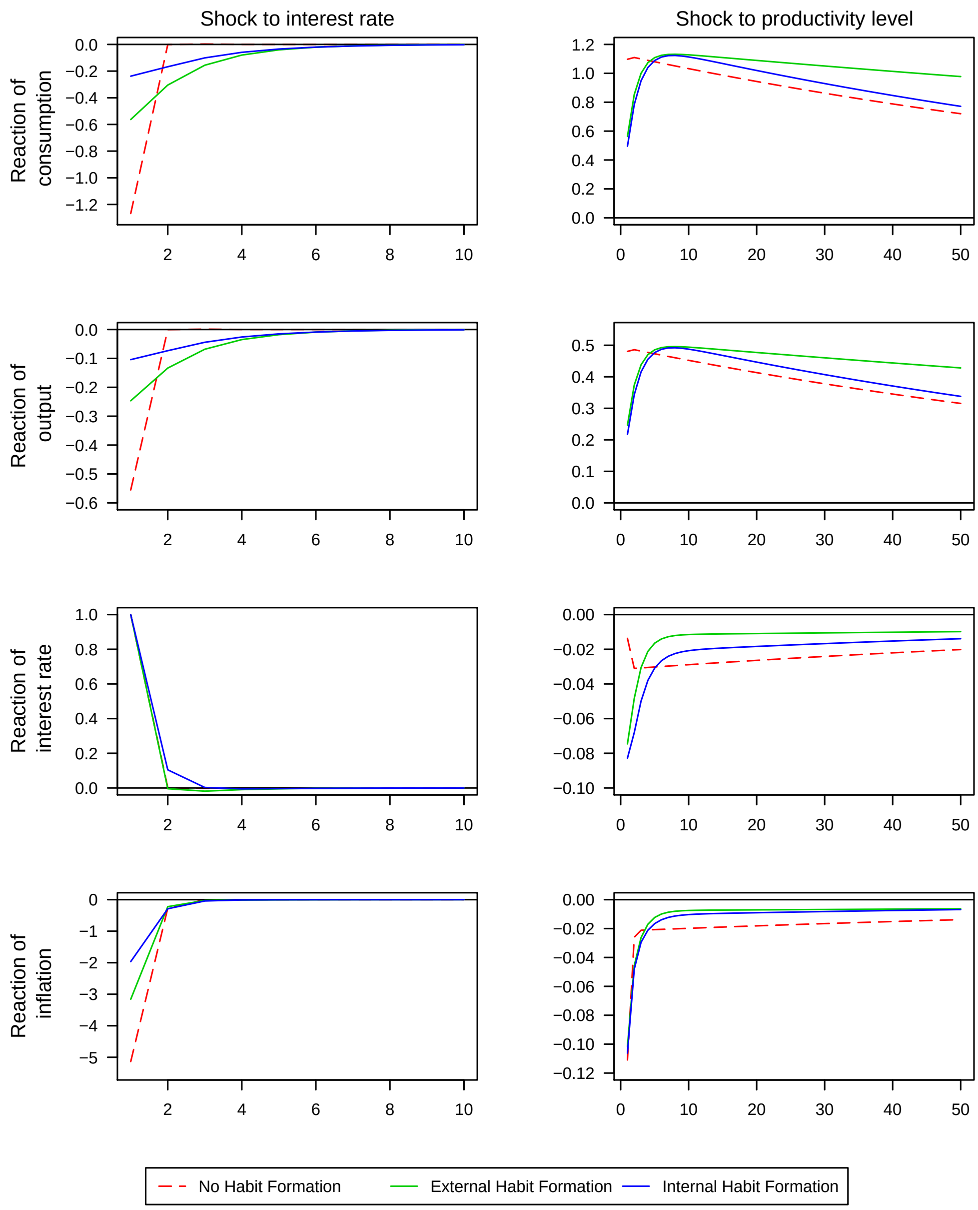
Figure 3.3 shows the impulse responses of consumption, output, the interest rate and inflation to shocks to the interest rate and the productivity level for the three model versions. The impulse response functions show the expected differences between the models: while the version without habit formation produces immediate reactions, the versions allowing for habit formation generate delayed reactions. A positive shock to the interest rate causes a temporary drop in consumption. In the baseline model, consumption reacts strongly in the first period and comes back to the equilibrium level in the second period already. Conversely, cosumption initially falls less in the habit formation models and then converges gradually back to the original level. The effect wears out after about seven quarters. A productivity shock creates persistent responses of consumption in all three model versions. In contrast to the reactions to the interest rate shock, the impulse responses to a productivity shock are hump-shaped. Again, the habit formation models produce more delayed reactions: while the impulse response reaches its peak in the second quarter in the model without habit formation, the peak is not reached until the eighth quarters in the habit formation models.

The confidence intervals of the impulse response functions of the habit formation models are shown in figures 3.7 and 3.8 in the appendix. The confidence interval of the reaction of consumption to an interest rate shock is generally larger in the habit formation models than in the baseline model. This is due to the more complex structure of the Euler equation in these model versions: while the interest rate enters the Euler equation with a factor of one in the baseline model, the factors are $(1-h)$ for external habit formation and $(1-h)(1-h\beta)$ for internal habit formation. The larger confidence intervals can thus be explained by the variance of the parameter h. The confidence intervals in the impulse responses to a productivity shock are generally larger as there are more parameters involved. For example, the reaction of consumption to a productivity shock depends on ρ_z (determines the persistence of the shock), φ (determines the effect of productivity on marginal costs), κ (determines the effect of marginal costs on inflation) and additionally h in the habit formation models (determines the effect of inflation on consumption).

3.4 Rule-of-Thumb Consumers

This section discusses the role of rule-of-thumb consumers in the small open economy model. Rule-of-thumb consumers are a special type of consumers which does not face a saving-spending trade-off. Instead, rule-of-thumb consumers spend all of their labor income on

Figure 3.3: *Comparison of the Impulse Response Functions of the Baseline Model and the Habit Formation Models*



NOTE: Impulse Responses to a one percent interest rate shock (left) and to a productivity shock (right) of magnitude 1 consumption in each time period. As the consumption behaviour of rule-of-thumb households is determined by the real wage and the amount of hours worked, I estimate the model under two alternative labor market structures, namely perfectly and imperfectly competitive labor markets. In most small DSGE models, labor markets are assumed to be perfectly competitive, i.e. individual workers choose the amount of hours they work given the market wage. The market wage in turn comes to lie at the level at which aggregate labor supply by workers matches labor demand by firms. When labor markets are assumed to be imperfectly competitive, some market power is usually given to unions: the latter set wages and workers then meet firms' labor demand.

In order to avoid indeterminacy as well as identification problems, I additionally fix three model parameters when estimating the model variants. The estimated share of rule-of-thumb consumers among all consumers is significantly different from zero for both labor market structures. However, allowing for rule-of-thumb behaviour improves the posterior probability only under perfectly competitive labor markets. Under imperfectly competitive labor markets, on the other hand, the posterior probability does not change noticeably and the estimated share of rule-of-thumb consumers is not completely robust to the specification of the prior distribution.

3.4.1 Motivation

Rule-of-thumb consumers are mainly used in two macreoeconomic contexts. First, they are introduced in consumption models in order to remedy for some of the discrepancies between the implications of the PIH and the empirical evidence. Second, they are used in the literature on fiscal multipliers in order to reproduce the empirical reactions of consumption to government spending shocks in neoclassical and new Keynesian models.

As in the case of habit formation, discrepancies between the PIH and the data are one starting point for the modeling of rule-of-thumb consumers. Empirical evidence suggests that consumption is more strongly correlated with current income than the PIH postulates (see, e.g. Hall and Mishkin, 1982)[4]. Campbell and Mankiw (1989) show that this form of "excess sensitivity" can be explained by a model in which two types of consumers coexist. Half the consumers operate in line with the PIH and make consumption decisions based on

[4]This is no contradiction to the arguments made in section 3.3. The motivation for habit formation is that consumption reacts *slower* to changes in income than the PIH suggests. The motivation for rule-of-thumb consumers is that consumption reacts *stronger* to changes in *current* income than the PIH suggests.

their permanent income. Consumers of this type are also called Ricardian consumers. The remaining consumers, called rule-of-thumb consumers, simply spend their current income in each time period and do not save. As a result, the change in consumption can be expressed as a weighted average of the changes in current and permanent income. While Campbell and Mankiw (1989) report clear evidence for the existence of rule-of-thumb consumers, they nevertheless note that not all consumers seem to consume current income: the fact that periods in which consumption is high relative to income are often followed by rapid income growth suggests that at least some consumers are forward-looking. Therefore, Campbell and Mankiw (1989) argue that the optimal model for consumption contains both Ricardian and rule-of-thumb consumers.

In the context of fiscal multipliers, rule-of-thumb consumers are added to models in order to reproduce impulse responses as they are found in empirical studies. Impulse response functions calculated using VAR models usually suggest a positive response of consumption to expansionary fiscal policy shocks. For US data, impulse responses from identified VAR approaches are documented for example in Fatás and Mihov (2000) or Blanchard and Perotti (2002). While this positive response of consumption can well be reproduced in IS-LM models, where consumption depends on current disposable income, standard neoclassical models generate a different outcome: as consumers anticipate that an increase government spending will eventually be financed by an increase in taxes, they view the increase in spending as a drop in their income. Consequently, they cut consumption and increase labor supply (see, e.g. Baxter and King, 1993, Galí et al., 2007). The higher labor supply then leads to a drop in real wages which lowers income and consumption further. Adding rule-of-thumb consumers can help to reproduce the empirical results if the fiscal shock causes real labor income to rise. While labor demand will always rise in response to this kind of shock, the reaction of the real wage can go in either direction. Galí et al. (2007) show that under sticky prices, the real wage can rise in response to a government spending shock and generate an increase in labor income. A sufficiently large share of rule-of-thumb consumers then assures a positive reaction of overall consumption. In line with this result, Forni et al. (2009) find “mild Keynesian effects” of government spending using an estimated DSGE model for the Euro Area. However, a rise of the real wage in response to a government spending shock is doubted for example by Coenen and Straub (2005). Instead, Coenen and Straub (2005) argue that a positive reaction of consumption to a government spending shock is unlikely to occur in a DSGE model.

The arguments mentioned so far explain why macroeconomists might include rule-of-thumb consumers in their models. However, in order for this extension to be plausible, there also needs to be some a priori reason why consumers might not save. One possible justification is limited asset markets participation (Bilbiie, 2008). Accordingly, rule-of-thumb consumers do not hold assets and do not smooth consumption because they cannot access asset markets. Other reasons brought forward are myopia or fear of saving (Galí et al., 2007). An intuitive reason is mentioned by Mankiw (2000): many households might not have the financial wherewithal to smooth consumption by saving. Instead, their consumption moves with current income.

Estimates of the share of rule-of-thumb consumers typically lie between 0.25 and 0.5. Using time series regressions, Campbell and Mankiw (1989) estimate shares of about 0.5. In studies using DSGE models for the Euro Area, Forni et al. (2009) estimate shares of about 0.35, depending on the labor market structure, while Coenen and Straub (2005) estimate shares of 0.25 to 0.37, depending on the tax system imposed. In their theoretical analysis of government spending effects, Galí et al. (2007) calibrate the share of rule-of-thumb consumers at 0.5. For Switzerland, Madsen and McAleer (2001) find a share of 0.18 and Evans and Karras (1998) a share of 0.37, both using time series estimation. To my knowledge, no estimations from DSGE models are available for Switzerland. In the SNB's DSGE model, Cuche-Curti et al. (2009) calibrate the share of rule-of-thumb consumers at 0.7.

3.4.2 Implementation

Extension of the household sector

The model now contains a continuum of households $i \in [0,1]$ which are divided into two fractions. The first fraction $1-s_r$ of households, called Ricardian or optimizing households, has access to capital markets and chooses the optimal intertemporal consumption path with respect to lifetime utility. The second fraction s_r of households, called rule-of-thumb households, spends all its labor income on consumption in each period. Variables referring to optimizing consumers are denoted by an "o" superscript, variables referring to rule-of-thumb consumers by an "r" superscript. Both consumer types have the familiar utility function:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{\varepsilon}_{p,t} \left(\log\left(C_t^i - hC_{t-1}^i\right) - \frac{\left(N_t^i\right)^{1+\varphi}}{1+\varphi} \right), \quad \text{for} \qquad i = \{o, r\}$$

According to the results from the previous section, I impose external habit formation in the model variant with perfectly competitive labor markets – the variant with imperfectly competitive labor markets cannot be estimated under habit formation due to indeterminacy problems. Furthermore, I assume the preference shock to be equal for the two consumer types. Ricardian consumers face the familiar budget constraint from the baseline model:

$$P_t C_t^o + \frac{1}{R_t} D_{t+1}^o + \frac{\varepsilon_t}{R_t^*} D_{t+1}^{*,o} = W_t P_t N_t^o + D_t^o + \varepsilon_t D_t^{*,o} + \Pi_t^o$$

I assume the real wage to be equal for both consumer types. Rule-of-thumb consumers face a reduced budget constraint:

$$P_t C_t^r = W_t P_t N_t^r$$

As a result, their consumption level is equal to their real labor income:

$$C_t^r = W_t N_t^r$$

Consumption and hours worked can be aggregated using the share of rule-of-thumb consumers, s_r:

$$C_t = s_r C_t^r + (1 - s_r)\, C_t^o$$
$$N_t = s_r N_t^r + (1 - s_r)\, N_t^o$$

Labor Market Structures

I consider perfectly competitive labor markets and imperfectly competitive labor markets. Under perfectly competitive labor markets, individual workers choose the amount of hours they work given the market wage. The market wage in turn comes to lie at the level at which aggregate labor supply by workers matches labor demand by firms. Under imperfectly competitive labor markets, wages are set by unions and workers then meet firms' labor demand. In both cases, the wage level is equal for rule-of-thumb and Ricardian workers. Moreover, an identical relation between the wage level, aggregate hours worked and aggregate consumption holds under both structures. However, while the amount of hours worked by the two household types is allowed to differ under perfectly competitive labor markets, both types work the same amount of hours under imperfectly competitive labor markets.

Under perfectly competitive labor markets, both consumer types face an analog intratemporal optimality condition. This condition determines for both consumer types the relation between hours worked and the marginal utility of consumption, given the real wage:

$$\left(N_t^i\right)^{\varphi}\left(C_t^i - hC_{t-1}^i\right) = W_t, \qquad \text{for} \quad i = \{o, r\}$$

This corresponds to the labor market structure I assumed in section 3.3. Although the wage level is equal for both consumer types, hours worked and consumption will generally differ.

In the model variant with imperfectly competitive labor markets, I follow Galí et al. (2007) and assume a continuum of differentiated labor/worker types $j \in [0,1]$. Effective labor hired by firm k is now an aggregate of the different labor types:

$$N_t(k) = \left(\int_0^1 N_t(j,k)^{\frac{\varepsilon_w - 1}{\varepsilon_w}} dj\right)^{\frac{\varepsilon_w}{\varepsilon_w - 1}} \tag{3.28}$$

where ε_w denotes the elasticity of substitution between the labor types. Each worker type is represented by a specific union. The fraction of Ricardian and rule-of-thumb households is uniformly distributed across the worker types and unions. In hiring labor, firms do not discriminate between Ricardian and rule-of-thumb households. A typical union sets the wage for its worker type z in each period according to the objective function

$$s_r\left[\left(C_t^r(z) - hC_{t-1}^r(z)\right)W_t(z)N_t(z) - \frac{N_t(z)^{1+\varphi}}{1+\varphi}\right] \\ + (1-s_r)\left[\left(C_t^0(z) - hC_{t-1}^0(z)\right)W_t(z)N_t(z) - \frac{N_t(z)^{1+\varphi}}{1+\varphi}\right]$$

and subject to the labor demand by firms

$$N_t(z) = \left(\frac{W_t(z)}{W_t}\right)^{-\varepsilon_w} N_t$$

The labor demand function can be derived from equation (3.28) by assuming cost minimization by firms; details are given in the appendix. In principle, the unions maximize real labor income minus disutility of labor for their worker types. However, as the consumption level will generally differ between Ricardian and rule-of-thumb households, unions weight labor income with the respective marginal utility of consumption. Furthermore, unions take into account that firms do not discriminate between Ricardian and rule-of-thumb households. This implies that in the aggregate, hours worked will be equal for the household types, i.e.

$N_t^r = N_t^0 = N_t$. As the labor types and unions are assumed to be symmetric, the index z can be dropped. The first order condition to the unions' maximization problem is then given by

$$\left(\frac{s_r}{MRS_t^r} + \frac{1-s_r}{MRS_t^0}\right) W_t = \frac{\varepsilon_w}{\varepsilon_w - 1}$$

where $MRS_t^i = (C_t^i - hC_{t-1}^i)N_t^{\varphi}$ for $i = \{o, r\}$.

Changes in the Linearized Model

I assume for simplicity that the steady state consumption levels of both consumer types are equal, i.e. $C^o = C^r = C$. While there are arguments against this assumption, equal steady-state consumption could always be obtained by modeling a government sector which makes transfers to the consumer types in a suitable way (Galí et al., 2007). Steady state hours worked are then also equalized: under perfectly competitive labor markets, an analog intratemporal optimality condition holds for both consumer types and equal steady state consumption thus implies equal steady state hours worked. Under imperfectly competitive labor markets, the amount of hours worked is always equal for the two consumer types. Consumption and hours worked can then be aggregated as follows:

$$c_t = s_r c_t^r + (1 - s_r)\, c_t^o \tag{3.29}$$

$$n_t = s_r n_t^r + (1 - s_r)\, n_t^o \tag{3.30}$$

For Ricardian consumers, the standard Euler equation determines the consumption path:

$$\begin{aligned}\left(c_t^o - hc_{t-1}^o\right) = \mathbb{E}_t\left[c_{t+1}^o - hc_t^o\right] - (1-h)\left(r_t - \mathbb{E}_t\left[\pi_{t+1}\right]\right) \\ + (1-h)\,\mathbb{E}_t\left[\varepsilon_{p,t} - \varepsilon_{p,t+1}\right]\end{aligned} \tag{3.31}$$

For rule-of-thumb consumers, consumption equals labor income:

$$c_t^r = w_t + n_t^r$$

The linearized intratemporal optimality condition for perfectly competitive labor markets is given by

$$\varphi n_t^i = w_t - \frac{1}{(1-h)}\left(c_t^i - hc_{t-1}^i\right), \qquad \text{for} \quad i = \{o, r\} \tag{3.32}$$

For the case of imperfectly competitive labor markets, the linearized first order condition to the unions' maximization problem is given by

$$w_t = \varphi n_t + \frac{1}{(1-h)}(c_t - hc_{t-1}) \tag{3.33}$$

Note that this condition also holds under perfectly competitive labor markets and can be constructed by combining equations (3.29), (3.30) and (3.32). The two labor market structures therefore result in an identical relation between the real wage and aggregate consumption as well as aggregate hours worked. However, hours worked are allowed to differ between Ricardian and rule-of-thumb consumers under perfectly competitive labor markets.

As the real wage is a key variable for the behaviour of rule-of-thumb consumers, it is worth to shortly highlight its determinants. Equation (3.33) states that the real wage is equal to the marginal rate of substitution between consumption and leisure. If the amount of hours worked rises, marginal utility of leisure rises and workers (or unions) ask for higher wages. The size of this effect depends on the elasticity of labor supply and therefore on the parameter φ: a low (high) elasticity of labor supply, corresonding to a high (low) parameter φ, leads to strong (weak) reaction of the real wage. Furthermore, a higher consumption level and/or consumption growth lowers marginal utility of consumption and thus also causes workers (unions) to ask for higher wages. Under habit formation, the size of this effect increases with the habit persistence parameter h. Summarizing the implications of equation (3.33), we note that whenever n_t and c_t move into the same direction, w_t also moves into this direction. On the other hand, if n_t and c_t move into opposite directions, the reaction of w_t depends on the parameters φ and h.

3.4.3 Effects on Determinacy

This section discusses the effect of the introduction of rule-of-thumb consumers on the determinacy of the model solution. A frequent source of indeterminacy in macroeconomic models is a weak reaction of the monetary authority to inflation (this corresponds to the parameter ψ_π being too small). Woodford (2001) shows that for a Taylor rule where the central bank reacts to inflation and the output cycle, the condition for a unique equilibrium in a standard new Keynesian model (i.e., without rule-of-thumb consumers) is given by

$$\psi_\pi > 1 - \frac{(1-\beta)\,\psi_y}{\kappa\,(1+\varphi)} \tag{3.34}$$

This condition is also called the "Taylor principle"[5]. Accordingly, the required reaction to inflation is lower if the central bank also reacts to output (i.e., $\psi_y > 0$). Moreover, a higher degree of price stickiness (corresponding to a lower κ) and/or a lower elasticity of labor supply (corresponding to a higher φ) lower the threshold value for ψ_π. If the central bank does not react to output, the condition reduces to $\psi_\pi > 1$. It is often argued that the fluctuations in output and inflation in the US in the 1960s and 1970s may have been caused by the Federal Reserve responding too little to inflation (see, e.g. Galí et al., 2004).

The presence of rule-of-thumb consumers alters the condition for determinacy. Galí et al. (2004) show that if the share of rule-of-thumb consumers is large enough, the Taylor principle can be an insufficient condition to guarantee determinacy. Instead, a high degree of price stickiness can lead to indeterminacy even if the Taylor principle holds. In order to give intuition on this result, Galí et al. (2004) consider an increase in economic activity which is not caused by fundamental shocks and which agents expect to wear out only gradually. As a consequence, inflation would increase and the central bank would raise the interest rate. Without the existence of rule-of-thumb consumers, such a boom would not be a rational expectations equilibrium because the higher interest rate would cause consumption to decrease and economic activity to decline. However, the presence of rule-of-thumb consumers can change this mechanism: depending on the model parameters, the higher labor demand by firms, together with sticky prices, could raise the real wage level. Rule-of-thumb consumers would then increase their consumption spending and depending on their share, their demand could sustain the boom. In this case, the central bank would not be able to control economic activtiy. Thus, any arbitrary change (i.e., a change not caused by fundamental shocks) in expectations could alter the equilibrium of the economy. Mathematically, the possibility of multiple equilibria leads to indeterminate model solutions.

Depending on the structure of the model, there is thus a threshold share of rule-of-thumb consumers above which model solutions are indeterminate. In addition to the degree of price stickiness, there are several other factors which affect this threshold share. As the key mechanism leading to indeterminacy is the direct reaction of rule-of-thumb consumers to changes in the real wage, any factor which affects the dynamics of the real wage also affects determinacy properties. Galí et al. (2004) discuss the role of the elasticity of labor supply in this context. In the case of an unexpected boom as discussed above, a lower

[5]This condition cannot directly be applied to the setting in the present paper as in the Taylor rule used here, the central bank is allowed to react to output growth.

elasticity of labor supply requires a stronger increase in the real wage in order for labor supply to meet the higher labor demand. But a stronger reaction of the real wage leads to a stronger reaction in rule-of-thumb consumers' spending, which in turn reduces the ability of the central bank to control consumption demand. Therefore, a lower elasticity of labor supply (i.e., a higher parameter φ) lowers the threshold share of rule-of-thumb consumers above which indeterminacy occurs. Bilbiie (2008) examines the interaction of labor and assets markets and finds that a low elasticity of labor supply, together with a sufficiently large share of rule-of-thumb consumers, can actually change the sign of the IS curve (i.e., the relation between the interest rate and output). In this case, an inverted Taylor principle holds: determinacy requires the monetary authority to lower the interest rate in response to inflation. The model used in Bilbiie (2008) captures the effects of hours worked and wages on profits as well as the flow of profits to Ricardian consumers.

General assumptions on the labor market structure also affect determinacy properties. Simulations in Motta and Tirelli (2010) show that indeterminacy problems increase if imperfectly competitive labor markets are assumed instead of perfectly competitive labor markets. The critical share of rule-of-thumb consumers above which model solutions are indeterminate is thus smaller under imperfectly competitive labor markets. Although the model structure in Motta and Tirelli (2010) is slightly different than mine, I found this finding confirmed in the estimation process. A possible explanation for this higher tendency to indeterminacy is the labor supply of the consumer types. Under perfectly competitive labor markets, both consumer types choose their amount of hours worked individually according to the market wage and to their marginal utility of consumption. In the scenario of an unexpected boom discussed above, Ricardian consumers cut their consumption level while rule-of-thumb consumers extend their consumption level. As these reactions affect the respective marginal utilities of consumption, Ricardian consumers will work more and rule-of-thumb consumers will work less. This dampens the increase in rule-of-thumb consumers' labor income and therefore the problematic positive response of rule-of-thumb consumers to the rise in the real wage. However, this effect vanishes once hours worked are equalized in the imperfectly competitive case: now, the difference in hours worked will lie between the optimal reaction for Ricardian and rule-of-thumb consumers and the positive reaction of rule-of-thumb consumers to the increase in the real wage turns out even stronger.

Finally, the effect of habit formation on determinacy is noteworthy. As the impulse response functions in section 3.3 show, habit formation leads to sluggish reactions of Ricardian consumers to monetary policy shocks. Habit formation thus undermines the effectiveness of monetary policy in general and lowers the critical share of rule-of-thumb consumers above which indeterminacy occurs. Indeed, allowing for habit formation created substantial indeterminacy problems under imperfectly competitive labor markets in my estimations.

Introducing rule-of-thumb consumers can thus create various problems with indeterminacy. In the algorithm used in this paper, indeterminacy leads to bounded parameter spaces: as all parameter sets implying indeterminate model solutions are rejected, the share of rule-of-thumb consumers is restricted below its critical value. In order to avoid such boundaries to affect the estimation results, I fix φ, the inverse of the elasticity of labor supply, at a low value. Furthermore, stronger indeterminacy problems require estimating the model variant with imperfectly competitive labor markets without habit formation.

3.4.4 Priors

The prior distributions for this section's estimations are shown in table 3.6. They are identical to those from section 3.3 with the exception of the parameters φ, η, ψ_y and s_r. φ is fixed at 0.1 in order to avoid indeterminacy problems due to combinations of low labor supply elasticites and high shares of rule-of-thumb consumers. I chose a value below the esimation results from sections 3.2 and 3.3 in order to rule out indeterminacy problems as far as possible. Apart from indeterminacy problems, the estimation of the model variants with rule-of-thumb consumers was subject to identification problems and slow convergence in general. In order to avoid these problems, I fix the parameters η and ψ_y. The parameter η, which measures the elasticity of substitution between domestic and foreign goods, was itself subject to identification problems and diverged to zero in some model specifications. According to the estimation results from section 3.3, I therefore fix η at 0.25. The parameter ψ_y, which measures the response of the central bank to the output cycle, was estimated to be approximately zero in the previous sections. With this parameter fixed at zero, the remaining parameter estimates converged faster. Furthermore, the habit persistence parameter h is restricted to 0 for the model version with imperfectly competitive labor markets. For the share of rule-of-thumb consumers (s_r), I set a relatively neutral normal distribution with a mean of 0.5 and standard deviation of 0.25.

3.4.5 Results

The posterior means are shown in table 3.7. Under perfectly competitive labor markets, the estimates converged after generating 5'750'000 candidate draws when the share of rule-

Table 3.6: *Prior Distributions of Parameters for the Models with Rule-of-Thumb Consumers*

Parameter	Distribution	Mean	Standard Deviation
s_r	Normal on [0,1]	0.50	0.25
h	Beta	0.50	0.20
β	calibrated	0.99	
α	calibrated	0.27	
κ	Gamma	0.30	0.30
δ	Normal on [0,1]	0.10	0.05
φ	calibrated	0.10	
η	calibrated	0.25	
ρ_r	Beta	0.50	0.20
ψ_π	Gamma	1.50	0.05
ψ_y	calibrated	0.00	
$\psi_{\Delta y}$	Normal > 0	0.40	0.05
ρ_q	Normal	0.00	0.50
ρ_z	Normal	0.80	0.50
ρ_{y^*}	Normal	0.50	0.50
ρ_{π^*}	Normal	0.50	0.50
ρ_p	calibrated	0.90	
ρ_i	Normal	0.80	0.50
C/Y	calibrated	0.60	
C_H^*	calibrated	0.40	
σ_q	Inv.Gamma	1.50	1.50
σ_r	Inv.Gamma	0.50	0.50
σ_z	Inv.Gamma	1.00	1.00
σ_{y^*}	Inv.Gamma	0.50	0.50
σ_{π^*}	Inv.Gamma	5.00	3.50
σ_p	Inv.Gamma	1.00	3.50
σ_{res}	Inv.Gamma	5.00	3.50

of-thumb consumers was restricted to zero. For the unrestricted version, the estimates converged after generating 4'000'000 candidate draws. Under imperfectly competitive labor markets, the estimates converged after generating 2'000'000 candidate draws for both the restricted and unrestricted version.

Share of Rule-of-Thumb Consumers

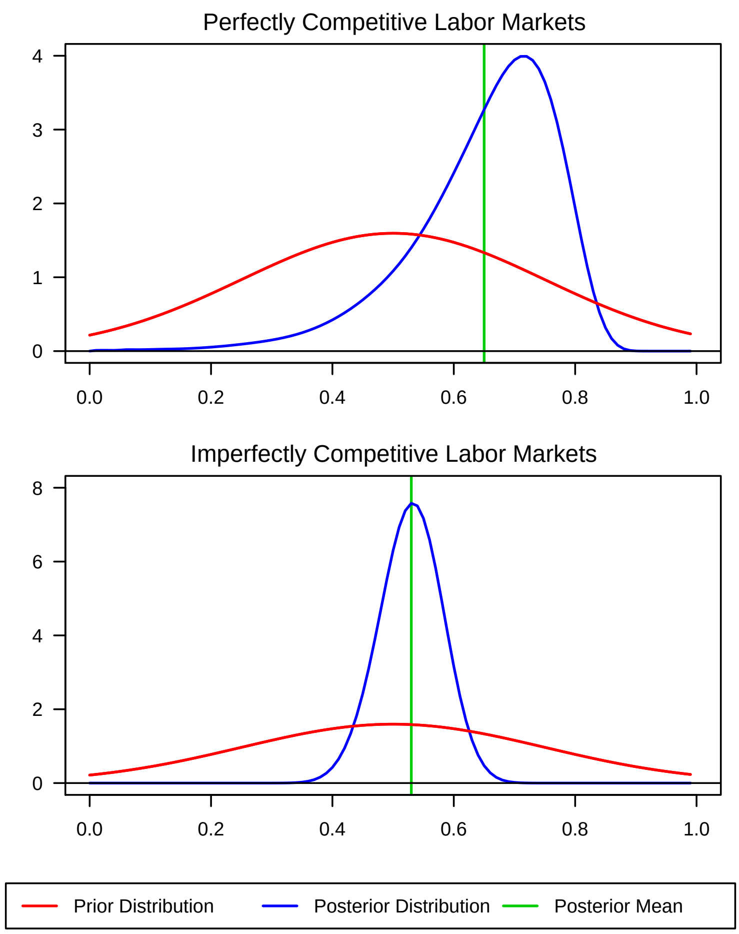
The share of rule-of-thumb consumers is estimated at 0.65 under perfectly competitive labor markets and at 0.53 under imperfectly competitive labor markets. The assumed labor market structure thus affects the estimates. However, the rather large standard deviation under perfectly competitive labor markets indicates that the parameter is not estimated with a high precision. On the other hand, the precision seems to be higher under imperfectly competitive labor markets. The posterior plots are shown in figure 3.4 and reveal that the parameter's posterior distribution is skewed to the left under perfectly competitive labor markets; the mode lies above the mean at approximately 0.75. Under imperfectly competitive labor markets, mode and mean lie close to each other. In both cases, the share of rule-of-thumb consumers is significantly different from zero. In order to check the dependence of these results on the specification of the prior mean, I run the estimations also with prior means of 0.25 and 0.70 for s_r. A prior mean of 0.75 was not possible for the model variant with imperfectly competitive labor markets as it resulted in an indeterminate model solution at the prior means of the parameters and the algorithm was therefore not able to initialize. Under perfectly competitive labor markets, s_r converges to 0.56 (prior mean 0.25) and 0.69 (prior mean 0.70). Under imperfectly competitive labor markets, s_r converges to 0.23 (prior mean 0.25) and 0.50 (prior mean 0.70). For both labor market structures, the lower prior mean of 0.25 alters the posterior mean. While the order of magnitude of the posterior mean stays the same for perfectly competitive labor markets, it changes considerably for imperfectly competitive labor markets. On the other hand, the higher prior mean of 0.70 leads to almost identical results for both labor market structures. Summarizing, the parameter estimation results support the concept of rule-of-thumb consumers under both perfectly and imperfectly competitive labor markets. However, the estimated share of rule-of-thumb consumers is not robust to the specification of the prior mean under imperfectly competitive labor markets.

Table 3.7: *Posterior Distributions of Parameters for the Models with Rule-of-Thumb Consumers*

		Perfectly Competitive LM				Imperfectly Competitive LM			
		Without ROT		With ROT		Without ROT		With ROT	
	Prior Mean	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
s_r	0.50	0.00^a		0.65	0.12	0.00^a		0.53	0.04
h	0.50	0.56	0.05	0.61	0.06	0.00^a		0.00^a	
κ	0.30	2.04	0.22	3.38	0.56	4.51	0.80	2.50	0.23
δ	0.10	0.04	0.03	0.07	0.04	0.06	0.04	0.05	0.03
ρ_r	0.50	0.36	0.07	0.27	0.08	0.25	0.07	0.25	0.07
ψ_π	1.50	1.45	0.05	1.48	0.05	1.48	0.05	1.49	0.05
$\psi_{\Delta y}$	0.40	0.31	0.05	0.30	0.05	0.32	0.05	0.32	0.05
ρ_q	0.00	0.17	0.03	0.13	0.03	0.11	0.03	0.09	0.03
ρ_z	0.80	0.99	0.01	0.99	0.01	0.99	0.01	1.00	0.00
ρ_{y^*}	0.50	0.89	0.04	0.87	0.04	0.88	0.04	0.89	0.03
ρ_{π^*}	0.50	0.25	0.10	0.25	0.10	0.25	0.10	0.25	0.10
ρ_{res}	0.80	0.82	0.05	0.87	0.05	0.82	0.05	0.96	0.02
σ_q	1.50	2.09	0.27	2.06	0.27	2.07	0.27	2.06	0.26
σ_r	0.50	0.38	0.06	0.38	0.06	0.38	0.06	0.39	0.06
σ_z	1.00	1.04	0.26	1.32	0.45	0.33	0.04	0.38	0.05
σ_{y^*}	0.50	0.23	0.03	0.23	0.03	0.23	0.03	0.23	0.03
σ_{π^*}	5.00	3.96	0.53	3.95	0.55	3.97	0.51	3.94	0.52
σ_p	1.00	1.64	0.31	1.06	0.25	1.04	0.24	0.92	0.18
σ_{res}	5.00	6.46	0.84	6.54	0.90	6.65	0.89	6.89	0.87
Log Likelihood		–915.51		–901.74		–925.64		–932.15	
Log Posterior Probability		–934.11		–924.17		–951.59		–950.87	
Posterior Odds Ratio for $s_r = 0$				$4.84 * 10^{-5}$				$4.87 * 10^{-1}$	
Candidate Draws		5'750'000		4'000'000		2'000'000		2'000'000	

NOTE: a – calibrated; ROT – rule-of-thumb consumers; LM – labor markets

Figure 3.4: *Posterior Plots for the Share of Rule-of-Thumb Consumers*



Other Model Parameters

With regard the other model parameters, mainly the parameter κ shows differences between the model versions. These differences can be explained by the differing intratemporal optimality conditions and structures of marginal costs. The implicitly estimated price stickiness parameters θ range between 0.15 (imperfectly competitive labor markets without rule-of-thumb consumers) and 0.26 (perfectly competitive labor markets without rule-of-thumb consumers). In all model versions, the estimation results imply prices to be fixed more often than every second quarter.

The remaining model parameters turn out very similarly across the model versions. Minor differences are visible in the autoregressive parameters of the Taylor rule and of the exogenous process for the terms of trade as well as in the standard deviation of the innovations to the preference shock.

Model Comparison

Statistics on the goodness of fit of the model variants are given at the bottom of table 3.7. Under perfectly competitive labor markets, the model variant allowing for rule-of-thumb consumers reaches a better likelihood as well as a better posterior probability than the variant without rule-of-thumb consumers. Under imperfectly competitive labor markets, however, the restricted version reaches a better log likelihood. The posterior probability is almost identical, indicating bigger deviations between prior distributions and posterior means in the restricted version. Consequently, the posterior odds ratio, which can be taken as a criterion for the null hypothesis $s_r = 0$, clearly supports the existence of rule-of-thumb consumers under perfectly competitive labor markets. Under imperfectly competitive labor markets, the posterior odds ratio marginally favors the version without rule-of-thumb consumers. Comparing the performance of the two labor market structures, the statistics indicate that perfectly competitive labor markets are a more appropriate assumption for Switzerland than imperfectly competitive labor markets.

Impulse Response Functions

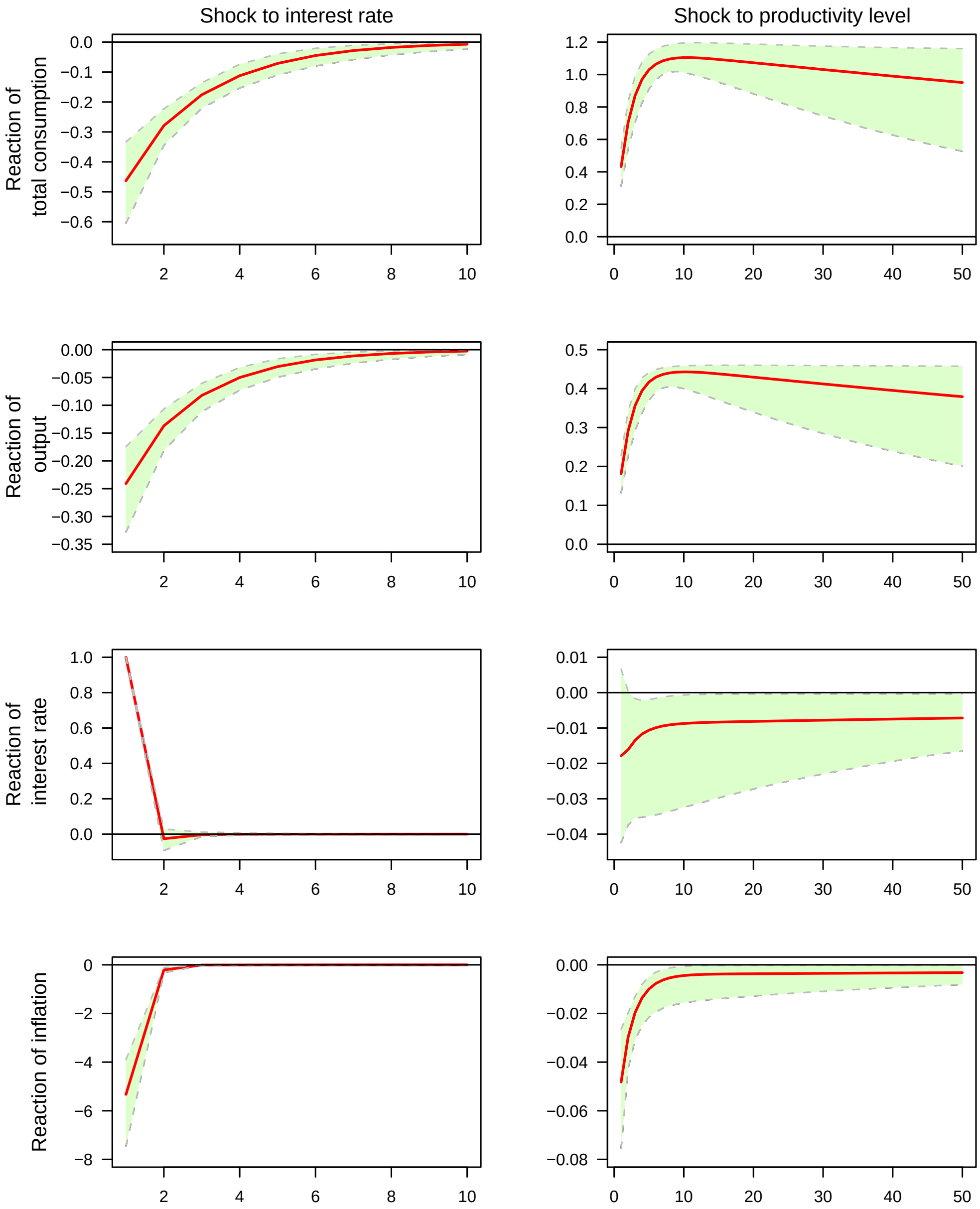
Plots showing the impulse response functions for the model versions with rule-of-thumb consumers are shown in figure 3.5 for perfectly competitive labor markets and in figure 3.6 for imperfectly competitive labor markets. The plots show the reactions of total consumption, output, the interest rate and inflation to shocks to the interest rate and to the technology level. Furthermore, the effects of a shock to the GDP residual are discussed in the text below. The reactions of the model are in principle similar under the two labor market structures. However, the model with imperfectly competitive labor markets produces more sudden reactions because it does not allow for habit formation.

A positive shock to the interest rate lowers consumption of Ricardian consumers and therefore output. The lower overall consumption level and the declining labor demand cause real wages to fall. As a consequence, rule-of-thumb consumers cut their consumption level as well. Both consumer types thus respond to monetary policy. However, it is important to note that the reaction of rule-of-thumb consumers is triggered by the reaction of Ricardian consumers.

A positive shock to the technology level lowers marginal costs as well as labor demand by firms. As a consequence, inflation decreases and the central bank lowers the interest rate. These primary effects cause Ricardian consumers to increase their consumption level. In spite of the corresponding increase in output, the net effect on labor demand is negative. The real wage is subject to two forces in this setting: while the falling labor demand pushes wages down, the higher consumption level of Ricardian consumers pushes them up. As the former effect is very small due to the high elasticity of labor supply, the overall effect on the real wage is positive. The rise in the real wages is stronger than the reduction in hours worked and consumption of rule-of-thumb consumers thus increases. Therefore, a positive technology shock causes an increase in total consumption and output as well as a small decrease in inflation and the interest rate.

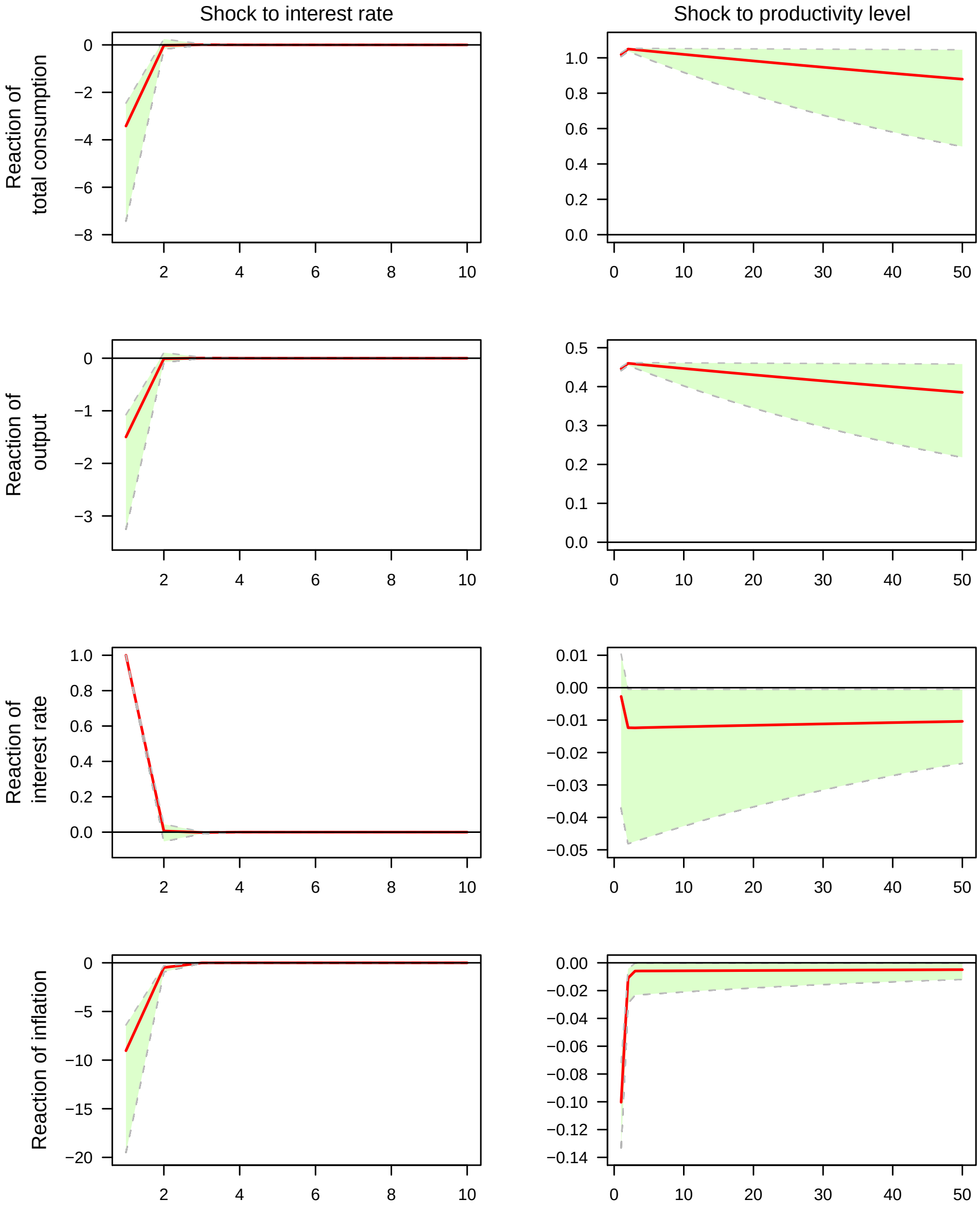
As mentioned, rule-of-thumb consumers are frequently used in order to assess the effects of government spending shocks on consumption. In my model, such a shock corresponds to a shock to the GDP residual. As there is no government sector which eventually raises taxes in order to finance the higher expenses, the full effects of a government spending shock cannot be analyzed. However, the effect on total consumption turns out negative even without accounting for the tax effect. The mechanism is as follows. A shock to the GDP residual directly increases output and labor demand. The central bank then raises the interest rate and Ricardian consumers cut their consumption level. Taken together, these effects lower the real wage: due to the high elasticity of labor supply, the higher labor demand has virtually no effect, but the lower consumption level pulls the wage downwards. As a consequence, rule-of-thumb consumers cut their consumption level as well and total consumption falls in response to the government spending shock.

Figure 3.5: *Impulse Reponses for the Model with Rule-of-Thumb Consumers under Perfectly Competitive Labor Markets*



Note: Impulse Responses to a one percent interest rate shock (left) and to a productivity shock (right) of magnitude 1 with 95-percent confidence intervals

Figure 3.6: *Impulse Reponses for the Model with Rule-of-Thumb Consumers under Imperfectly Competitive Labor Markets*



NOTE: Impulse Responses to a one percent interest rate shock (left) and to a productivity shock (right) of magnitude 1 with 95-percent confidence intervals

3.5 Conclusions

This paper analyzes habit formation and rule-of-thumb consumers in a small open economy model. The model is a simplified version of the model used in Galí and Monacelli (2005). Using a Markov Chain Monte Carlo method on Swiss data, I estimate the degree of habit persistence under external and internal habit formation as well as the share of rule-of-thumb consumers under two labor market structures. Furthermore, I compare the performance of the model versions.

I find habit persistence parameters of 0.54 in the case of external habit formation and 0.61 in the case of internal habit formation. Both are significantly different from zero. Allowing for habit formation improves the likelihood as well as the posterior probability of the model. The direct comparison of the two habit formation models shows that external habit formation reaches a better posterior probability. With respect to the share of rule-of-thumb consumers, I find a value of 0.65 under perfectly competitive labor markets and a value of 0.53 under imperfectly competitive labor markets. Although both estimates are significantly different from zero, the overall performance of the model does not improve in both cases. For the variant with perfectly competitive labor markets, the likelihood and posterior probability improve when rule-of-thumb behaviour is allowed for. For the variant with imperfectly competitive labor markets, the results are less clear-cut: the parameter estimates are not perfectly robust and the posterior probability does not allow for a clear acceptance or rejection of rule-of-thumb consumers.

In future research, the robustness of the results could be checked in more complex models. For example, the law of one price could be relaxed or investment could be modeled structurally. In the case of rule-of-thumb consumers, additional insights on modeling features which reduce indeterminacy problems would be useful.

3.A Appendix

3.A.1 Derivations of Model Equations

Aggregation of Consumption Goods

The perfectly competitive retail sector which buys $C_{H,t}$ and $C_{F,t}$ and sells the aggregated consumption good C_t to consumers faces the following problem:

$$\begin{aligned} \max_{C_{H,t},C_{F,t}} \quad & P_t C_t - P_{H,t} C_{H,t} - P_{F,t} C_{F,t} \\ \text{subject to} \quad & C_t = \left[(1-\alpha)^{\frac{1}{\eta}} C_{H,t}^{\frac{\eta-1}{\eta}} + \alpha^{\frac{1}{\eta}} C_{F,t}^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}} \end{aligned}$$

inserting and taking first order conditions with regard to $C_{H,t}$ and $C_{F,t}$ yields the following demand functions:

$$\begin{aligned} C_{H,t} &= (1-\alpha) \left(\frac{P_{H,t}}{P_t} \right)^{-\eta} C_t \\ C_{F,t} &= \alpha \left(\frac{P_{F,t}}{P_t} \right)^{-\eta} C_t \end{aligned}$$

The overall price index P_t can be calculated from a zero profit condition for the retail sector:

$$\begin{aligned} 0 &= P_t C_t - P_{H,t}(1-\alpha) \left(\frac{P_{H,t}}{P_t} \right)^{-\eta} C_t - P_{F,t} \alpha \left(\frac{P_{F,t}}{P_t} \right)^{-\eta} C_t \\ P_t &= \left[(1-\alpha) P_{H,t}^{1-\eta} + \alpha P_{F,t}^{1-\eta} \right]^{\frac{1}{1-\eta}} \end{aligned}$$

The underlying layer of retailers buys intermediate goods $Y_{C,t}(i)$ and aggregates them to units of the domestic good. This retail sector faces the following problem:

$$\begin{aligned} \max_{Y_t(i)} \quad & P_{H,t} Y_{C,t} - \int_0^1 P_{H,t}(i) Y_{C,t}(i) di \\ \text{subject to} \quad & Y_{C,t} = \left[\int_0^1 Y_{C,t}(i)^{\frac{\varepsilon-1}{\varepsilon}} di \right]^{\frac{\varepsilon}{\varepsilon-1}} \end{aligned}$$

Inserting and taking first order conditions with respect to $Y_t(i)$ yields the demand functions

$$Y_{C,t}(i) = \left(\frac{P_{H,t}(i)}{P_{H,t}}\right)^{-\varepsilon} Y_{C,t}$$

The domestic price index $P_{H,t}$ can be calculated from a zero profit condition for this retail sector:

$$\begin{aligned} 0 &= P_{H,t}Y_{C,t} - \int_0^1 P_{H,t}(i)\left(\frac{P_{H,t}(i)}{P_{H,t}}\right)^{-\varepsilon} Y_{C,t}di \\ P_{H,t} &= \left[\int_0^1 P_{H,t}(i)^{1-\varepsilon}di\right]^{\frac{1}{1-\varepsilon}} \end{aligned}$$

Imperfectly Competitive Labor Markets: Labor Demand

In order to derive labor demand by firms, an additional sector is assumed in analogy to the two retail sectors in the section above (see, e.g., Fernández-Villaverde, 2009). This sector hires the differentiated labor types, aggregates them to homogenous labor units and sells these units to the firms of the production sector. The aggregation of homogenous labor units follows a production function:

$$N_t(k) = \left(\int_0^1 N_t(z,k)^{\frac{\varepsilon_w - 1}{\varepsilon_w}} dz\right)^{\frac{\varepsilon_w}{\varepsilon_w - 1}}$$

where $N_t(k)$ is the labor input sold to firm k, $N_t(z,k)$ is the used amount of hours worked by labor type z and ε_w is the elasiticty of substitution among the labor types. The aggregation sector maximizes profits, taking individual wages $W_t(z)P_t$ and the aggregate wage level W_tP_t as given:

$$\begin{aligned} \max_{N_t(z)} \quad & W_tP_tN_t(k) - \int_0^1 W_t(z)P_tN_t(z,k)dz \\ \text{subject to} \quad & N_t(k) = \left(\int_0^1 N_t(z,k)^{\frac{\varepsilon_w - 1}{\varepsilon_w}} dz\right)^{\frac{\varepsilon_w}{\varepsilon_w - 1}} \end{aligned}$$

Inserting and taking first order conditions with respect to $N_t(z,k)$ yields the demand functions

$$N_t(z,k) = \left(\frac{W_t(z)}{W_t}\right)^{-\varepsilon_w} N_t(k)$$

Furthermore, the overall wage index W_t can be aggregated from the individual wages by assuming zero profits for the aggregation sector:

$$W_t = \left(\int_0^1 W_t(z)^{1-\varepsilon_w} dz \right)^{\frac{1}{1-\varepsilon_w}}$$

Log-Linearization

Prior to the linearization of the model equations, some steady states have to be derived. Following Del Negro and Schorfheide (2008), I set the steady state of the real exchange rate to one, i.e. $S = 1$. From equation (3.11), it then follows that the staedy state of the terms of trade is also equal to one, i.e. $Q = 1$. With regard to prices, equation (3.10) states that $P_H = P$. Together with the price index (3.2), this leads to $P_F = P_H = P$. From equation (3.9), it also follows that $C_H^* = \alpha^* C^*$. Furthermore, I use the approximation $c_t^* = y_t^*$.

Euler equation: the linearization of the various versions of the Euler equation is straightforward.

Real exchange rate: equation (3.11) linearizes to

$$s_t = -(1-\alpha)\, q_t \tag{3.35}$$

Aggregate production function: as mentioned, I assume that the residual $Y_{res,t}$ is produced according to an individual production function analog to equation (3.5). The amount of hours worked can then be aggregated by integrating over the firms of both the production sector of the intermediate consumption good as wall as the implicit production sector of the residual.

$$N_t = \int_0^1 N_t(i)di = \int_0^1 \frac{Y_t(i)}{Z_t} di = \frac{Y_t A_t}{Z_t}$$

where $A_t = \int_0^1 (Y_t(i)/Y_t)di$. Under perfectly competitive labor markets, $N_t(i)$ corresponds to the amount of hours worked which labor i hires. Under imperfectly competitive labor markets on the other hand, $N_t(i)$ corresponds to units of the homogenous labor unit which firm i hires. Galí and Monacelli (2005) show that variations in $\log(A_t)$ around its steady state are of second order. Therefore, the following linearized relationship between aggregate vairables holds up to a first order approximation:

$$n_t = y_t - z_t \tag{3.36}$$

Phillips curve: the derivation of the Phillips curve involves several steps. In equation (3.7), the stochastic discount factor $Q_{t+\tau|t}$ can be substituted by $\lambda_{t+\tau}/\lambda_t$. λ_t can then be canceled and the first order condition becomes

$$\mathbb{E}_t\left[\sum_{\tau=0}^{\infty}\theta^{\tau}\lambda_{t+\tau}Y_{C,t+\tau|t}\left(\left(\frac{P_{H,t+\tau-1}}{P_{H,t-1}}\right)^{\delta}P'_{H,t}-\frac{\varepsilon}{\varepsilon-1}P_{H,t+\tau}MC_{t+\tau}\right)\right]=0$$

By taking the steady state version of this equation, it follows that $MC=(\varepsilon-1)/\varepsilon$. Prior to linearization, the expression is rearranged to

$$\mathbb{E}_t\left[\sum_{\tau=0}^{\infty}\theta^{\tau}\lambda_{t+\tau}Y_{C,t+\tau|t}\left(\frac{P_{H,t+\tau-1}}{P_{H,t-1}}\right)^{\delta}P'_{H,t}\right]=\mathbb{E}_t\left[\sum_{\tau=0}^{\infty}\theta^{\tau}\lambda_{t+\tau}Y_{C,t+\tau|t}\frac{\varepsilon}{\varepsilon-1}P_{H,t+\tau}MC_{t+\tau}\right]$$

With respect to the steady states of the left and right hand sides of the above equation, note that the steady state of $\lambda_{t+\tau}$ cancels out only up to β^{τ}, independently of the type of habit formation imposed in the utility function. Moreover, $MC(\varepsilon/(\varepsilon-1))=1$. Linearization then yields

$$\mathbb{E}_t\left[\sum_{\tau=0}^{\infty}(\beta\theta)^{\tau}\left(\delta\left(p_{H,t+\tau-1}-p_{H,t-1}\right)+p'_{H,t}\right)\right]=\mathbb{E}_t\left[\sum_{\tau=0}^{\infty}(\beta\theta)^{\tau}\left(p_{H,t+\tau}+mc_{t+\tau}\right)\right]$$

Rearranging yields

$$\begin{aligned}
\mathbb{E}_t\left[\sum_{\tau=0}^{\infty}(\beta\theta)^{\tau}\left(p'_{H,t}-p_{H,t-1}\right)\right] &= \mathbb{E}_t\left[\sum_{\tau=0}^{\infty}(\beta\theta)^{\tau}\left((\delta-1)p_{H,t-1}-\delta p_{H,t+\tau-1}+p_{H,t+\tau}+mc_{t+\tau}\right)\right]\\
\frac{1}{1-\beta\theta}\left(p'_{H,t}-p_{H,t-1}\right) &= \mathbb{E}_t\left[\sum_{\tau=0}^{\infty}(\beta\theta)^{\tau}\left(mc_{t+\tau}+\delta\pi_{H,t+\tau}+(\delta-1)\left(p_{H,t-1}-p_{H,t+\tau}\right)\right)\right]\\
p'_{H,t}-p_{H,t-1} &= (1-\beta\theta)\,\mathbb{E}_t\left[\sum_{\tau=0}^{\infty}(\beta\theta)^{\tau}\left(mc_{t+\tau}+\delta\pi_{H,t+\tau}\right)\right]\\
&\quad+(1-\beta\theta)(1-\delta)\,\mathbb{E}_t\left[\sum_{\tau=0}^{\infty}(\beta\theta)^{\tau}\left(\sum_{k=0}^{\tau}\pi_{H,t+k}\right)\right]
\end{aligned}$$

By writing out the sum and regrouping, the last part of the expression can be rewritten as

$$(1-\beta\theta)(1-\delta)\,\mathbb{E}_t\left[\sum_{\tau=0}^{\infty}(\beta\theta)^{\tau}\left(\sum_{k=0}^{\tau}\pi_{H,t+k}\right)\right]=(1-\delta)\sum_{\tau=0}^{\infty}(\beta\theta)^{\tau}\,\pi_{H,t+\tau}$$

and the whole expression becomes

$$\begin{aligned} p'_{H,t} - p_{H,t-1} = (1-\beta\theta)\,\mathbb{E}_t &\left[\sum_{\tau=0}^{\infty} (\beta\theta)^{\tau} \left(mc_{t+\tau} + \delta\pi_{H,t+\tau}\right)\right] \\ + (1-\delta)\,\mathbb{E}_t &\left[\sum_{\tau=0}^{\infty} (\beta\theta)^{\tau}\, \pi_{H,t+\tau}\right] \end{aligned} \tag{3.37}$$

By increasing the time index by one, multiplying the whole expression by $\beta\theta$ and adjusting the starting values of the sums, the following expression for $p'_{H,t+1} - p_{H,t}$ is derived:

$$\begin{aligned} \beta\theta\mathbb{E}_t\left[p'_{H,t+1} - p_{H,t}\right] = (1-\beta\theta)\,\mathbb{E}_t &\left[\sum_{\tau=1}^{\infty} (\beta\theta)^{\tau} \left(mc_{t+\tau} + \delta\pi_{H,t+\tau}\right)\right] \\ + (1-\delta)\,\mathbb{E}_t &\left[\sum_{\tau=1}^{\infty} (\beta\theta)^{\tau}\, \pi_{H,t+\tau}\right] \end{aligned}$$

This expression can then be integrated into equation (3.37):

$$p'_{H,t} - p_{H,t-1} = (1-\beta\theta)\, mc_t + (1-\delta\beta\theta)\,\pi_{H,t} + \beta\theta\mathbb{E}_t\left[p'_{H,t+1} - p_{H,t}\right] \tag{3.38}$$

A log-linear approximation to the domestic price index in equation (3.8) yields

$$p_{H,t} = (1-\theta)\, p'_{H,t} + \theta\left((1+\delta)\, p_{H,t-1} - \delta p_{H,t-2}\right)$$

which can be transformed into

$$p'_{H,t} - p_{H,t-1} = \frac{1}{1-\theta}\pi_{H,t} - \frac{\delta\theta}{1-\theta}\pi_{H,t-1} \tag{3.39}$$

Combining equations (3.38) and (3.39) finally yields the Phillips curve:

$$\pi_{H,t} - \delta\pi_{H,t-1} = \frac{(1-\theta)(1-\beta\theta)}{\theta} mc_t + \beta\mathbb{E}_t\left[\pi_{H,t+1} - \delta\pi_{H,t}\right]$$

Marginal costs: the definition of real marginal costs can be transformed using equation (3.10):

$$MC_t = \frac{W_t P_t}{Z_t P_{H,t}} = \frac{W_t}{Z_t}\frac{1}{Q_t S_t}$$

This expression linearizes to

$$mc_t = w_t - z_t - \alpha q_t$$

Under perfectly competitive labor markets, the real wage can then be substituted from the household's linearized intratemporal optimality condition. Under imperfectly competitive labor markets, the real wage can be substituted from the linearized first order condition to the unions' maximization problem. In other words, the real wage can be substituted using equation (3.33) in both cases. Furthermore, hours worked can be substituted according to equation (3.36). The expression for marginal costs is thus identical under both labor market structures. However, the expression for marginal costs is affected by the type of habit formation imposed in the household's utility function. The resulting expressions are given in equations (3.27a), (3.27b) and (3.27c).

Market clearing: using equation (3.35) as well as the steady states and approximations mentioned above, the market clearing condition in equation (3.12) linearizes to

$$y_t = (1-\alpha)\frac{C}{Y}c_t - \eta\left(\alpha(1-\alpha)\frac{C}{Y} + \frac{C_H^*}{Y}\right)q_t + \frac{C_H^*}{Y}y_t^* + \frac{Y_{res}}{Y}y_{res,t}$$

The steady state version of the market clearing condition is $Y = (1-\alpha)C + C_H^* + Y_{res}$. Therefore, Y_{res} can be expressed in terms of Y, C and C_H^* and the linearized market clearing condition becomes

$$y_t = (1-\alpha)\frac{C}{Y}c_t - \eta\left(\alpha(1-\alpha)\frac{C}{Y} + \frac{C_H^*}{Y}\right)q_t + \frac{C_H^*}{Y}y_t^* + \left(1 - (1-\alpha)\frac{C}{Y} - \frac{C_H^*}{Y}\right)y_{res,t}$$

Change in nominal exchange rate: using the definition of the real exchange rate, the change in the nominal exchange rate can be written as

$$\varepsilon_t = \frac{S_t}{S_{t-1}}\frac{P_t}{P_{t-1}}\frac{P_{t-1}^*}{P_t^*}$$

Using equation (3.35), this expression linearizes to

$$e_t = -(1-\alpha)\Delta q_t + \pi_t - \pi_t^*$$

Overall CPI inflation: according to equation (3.10), CPI inflation can be expressed as

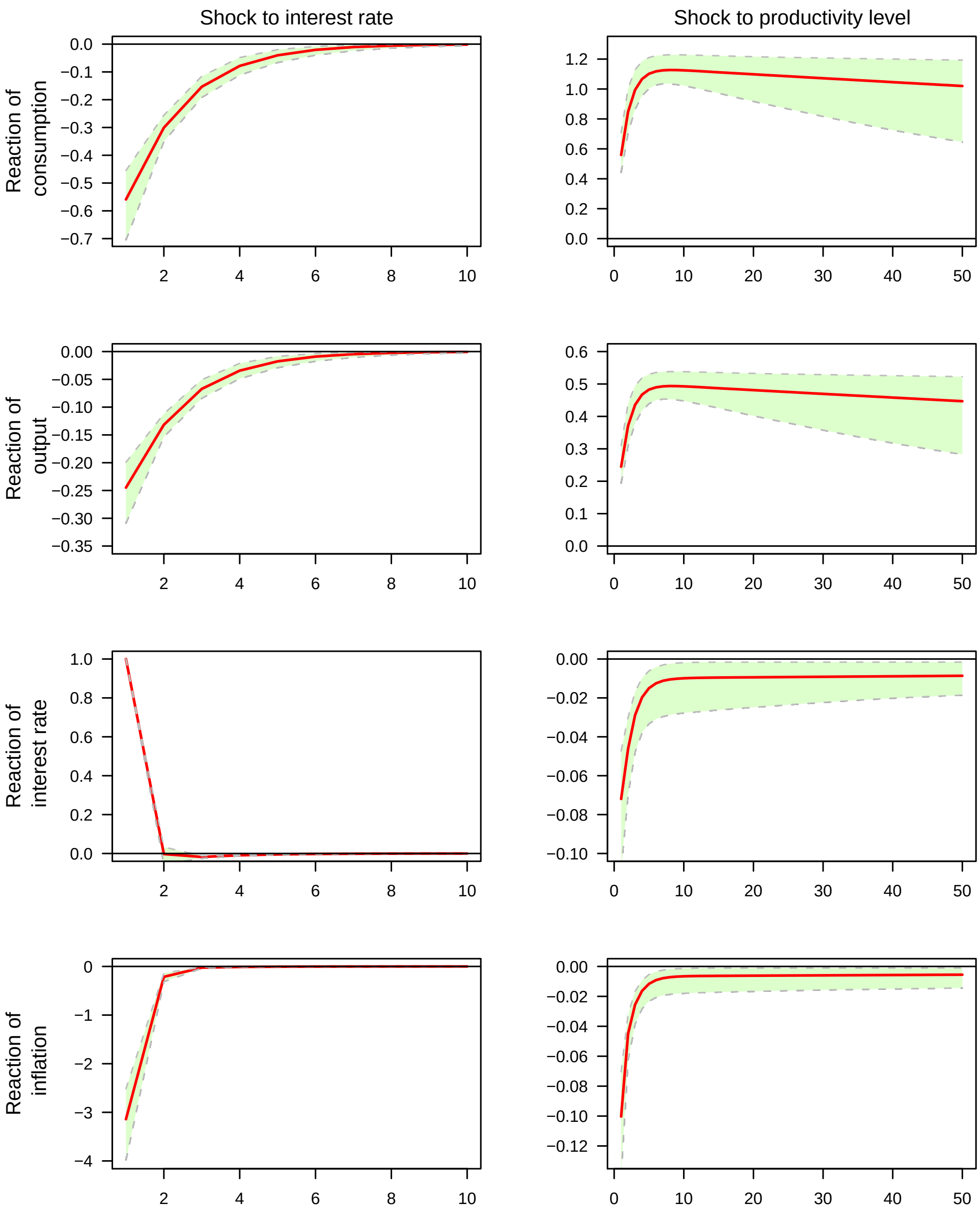
$$\frac{P_t}{P_{t-1}} = \frac{P_{H,t}}{P_{H,t-1}} \frac{Q_{t-1}}{Q_t} \frac{S_{t-1}}{S_t}$$

Using again equation (3.35), this expression linearizes to

$$\pi_t = \pi_{H,t} - \alpha \Delta q_t$$

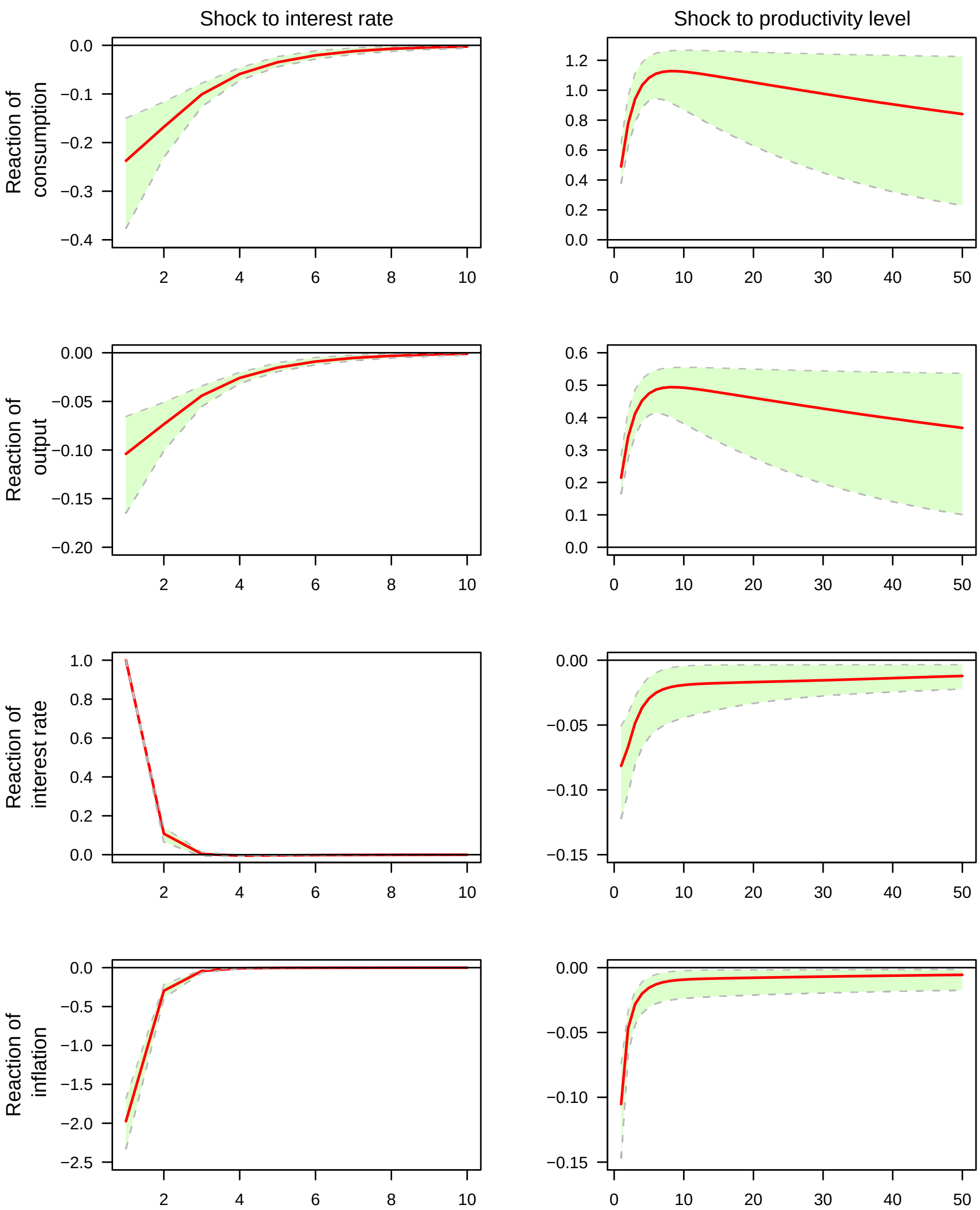
3.A.2 Impulse Response Functions

Figure 3.7: *Impulse Responses for the Model with External Habit Formation*



NOTE: Impulse Responses to a one percent interest rate shock (left) and to a productivity shock (right) of magnitude 1 with 95-percent confidence intervals

Figure 3.8: *Impulse Responses for the Model with Internal Habit Formation*



NOTE: Impulse Responses to a one percent interest rate shock (left) and to a productivity shock (right) of magnitude 1 with 95-percent confidence intervals

Bibliography

Abel, A. B. (1990). Asset Prices under Habit Formation and Catching up with the Joneses. *American Economic Review*, 80(2):38–42.

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

Assenmacher-Wesche, K. (2008). Modeling Monetary Transmission in Switzerland with a Structural Cointegrated VAR Model. *Swiss Journal of Economics and Statistics*, 144(2):197–246.

Baxter, M. and King, R. G. (1993). Fiscal-Policy in General Equilibrium. *American Economic Review*, 83(3):315–334.

Beltran, D. O. and Draper, D. (2008). Estimating the Parameters of a Small Open Economy DSGE Model: Identifiability and Inferential Validity. Technical report, Board of Governers of the Federal Reserve System.

Bilbiie, F. O. (2008). Limited Asset Markets Participation, Monetary Policy and (Inverted) Aggregate Demand Logic. *Journal of Economic Theory*, 140(1):162–196.

Blanchard, O. and Perotti, R. (2002). An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output. *Quarterly Journal of Economics*, 117(4):1329–1368.

Bäurle, G. and Menz, T. (2008). Monetary Policy in a Small Open Economy Model: A DSGE-VAR Approach for Switzerland. *Study Center Gerzensee Working Paper*, 08.03.

Campbell, J. and Deaton, A. (1989). Why Is Consumption So Smooth? *Review of Economic Studies*, 56(3):357–374.

Campbell, J. Y. and Cochrane, J. H. (1999). By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy*, 107(2):205–251.

Campbell, J. Y. and Mankiw, N. G. (1989). *Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence*, volume 1989:4 of *NBER Macroeconomics Annual*. MIT Press, Cambridge.

Carroll, C. D., Overland, J., and Weil, D. N. (2000). Saving and Growth with Habit Formation. *American Economic Review*, 90(3):341–355.

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

Coenen, G. and Straub, R. (2005). Does Government Spending Crowd in Private Consumption? Theory and Empirical Evidence for the Euro Area. *International Finance*, 8(3):435–470.

Cuche-Curti, N. A., Dellas, H., and Natal, J. (2009). DSGE-CH: A Dynamic Stochastic General Equilibrium Model for Switzerland. *Swiss National Bank Economic Studies*, 2009-5.

DeJong, D. N. and Dave, C. (2007). *Structural Macroeconometrics*. Princeton University Press, Princeton, N.J.

Del Negro, M. and Schorfheide, F. (2008). *Inflation Dynamics in a Small Open Economy Model under Inflation Targeting - Some Evidence from Chile*. Monetary Policy under Uncertainty and Learning. Central Bank of Chile, Santiago, Chile.

Dennis, R. (2004). Specifying and Estimating New Keynesian Models with Instrument Rules and Optimal Monetary Policies. *Federal Reserve Bank of San Francisco Working Paper*, 2004-17.

Dennis, R. (2009). Consumption Habits in a New Keynesian Business Cycle Model. *Journal of Money Credit and Banking*, 41(5):1015–1030.

Evans, P. and Karras, G. (1998). Liquidity Constraints and the Substitutability between Private and Government Consumption: The Role of Military and Non-Military Spending. *Economic Inquiry*, 36(2):203–214.

Fatás, A. and Mihov, I. (2000). The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence. *INSEAD Working paper*, 2000/78.

Fernández-Villaverde, J. (2009). The Econometrics of DSGE Models. *NBER Working Paper Series*, 14677.

Forni, L., Monteforte, L., and Sessa, L. (2009). The General Equilibrium Effects of Fiscal Policy: Estimates for the Euro Area. *Journal of Public Economics*, 93(3-4):559–585.

Fuhrer, J. C. (2000). Habit Formation in Consumption and its Implications for Monetary-Policy Models. *American Economic Review*, 90(3):367–390.

Galí, J., López-Salido, J. D., and Vallés, J. (2004). Rule-of-thumb Consumers and the Design of Interest Rate Rules. *Journal of Money Credit and Banking*, 36(4):739–763.

Galí, J. and Monacelli, T. (2005). Monetary Policy and Exchange Rate Volatility in a Small Open Economy. *Review of Economic Studies*, 72(3):707–734.

Galí, J., Vallés, J., and López-Salido, J. D. (2007). Understanding the Effects of Government Spending on Consumption. *Journal of the European Economic Association*, 5(1):227–270.

Geweke, J. (2007). Models, Computational Experiments and Reality. Unpublished manuscript. University of Iowa.

Hall, R. E. (1978). Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence. *Journal of Political Economy*, 86(6):971–87.

Hall, R. E. and Mishkin, F. S. (1982). The Sensitivity of Consumption to Transitory Income - Estimates from Panel Data on Households. *Econometrica*, 50(2):461–481.

Kaufmann, S. (1996). Permanent Components in Swiss Macroeconomic Variables. *Swiss Journal of Economics and Statistics*, 132(4):539–562.

Korniotis, G. M. (2010). Estimating Panel Models With Internal and External Habit Formation. *Journal of Business and Economic Statistics*, 28(1):145–158.

Kugler, P. and Jordan, T. J. (2004). Structural Vector Autoregressions and the Analysis of Monetary Policy Interventions: The Swiss Case. *Swiss Journal of Economics and Statistics*, 140(1):67–87.

Leeper, E. M., Sims, C. A., and Zha, T. (1996). What Does Monetary Policy Do? *Brookings Papers on Economic Activity*, 2:1–63.

Leist, S. and Neusser, K. (2010). Measuring the Natural Output Level by DSGE Models: An Empirical Investigation for Switzerland. *Swiss Journal of Economics and Statistics*, 146(1):275–300.

Ljungqvist, L. and Uhlig, H. (2000). Tax Policy and Aggregate Demand Management Under Catching up with the Joneses. *American Economic Review*, 90(3):356–366.

Madsen, J. B. and McAleer, M. (2001). Consumption, Liquidity Constraints, Uncertainty and Temptation: An International Comparison. *Journal of Economic Psychology*, 22(1):61–89.

Mankiw, N. G. (2000). The Savers-Spenders Theory of Fiscal Policy. *American Economic Review*, 90(2):120–125.

Mehra, R. and Prescott, E. C. (1985). The Equity Premium - a Puzzle. *Journal of Monetary Economics*, 15(2):145–161.

Motta, G. and Tirelli, P. (2010). Rule-of-thumb Consumers, Consumption Habits and the Taylor Principle. *University of Milano-Bicocca, Department of Economics Working Paper Series*, 194.

Muellbauer, J. (1988). Habits, Rationality and Myopia in the Life Cycle Consumption Function. *Annales d'Economie et de Statistique*, 9:47–70.

Ravn, M., Schmitt-Grohe, S., and Uribe, M. (2006). Deep Habits. *Review of Economic Studies*, 73(1):195–218.

Sims, C. A. (2001). Solving Linear Rational Expectations Models. *Computational Economics*, 20:1–20.

Smets, F. and Wouters, R. (2003). An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area. *Journal of the European Economic Association*, 1(5):1123–1175.

Sommer, M. (2007). Habit Formation and Aggregate Consumption Dynamics. *B E Journal of Macroeconomics*, 7(1).

Uribe, M. (2002). The Price-Consumption Puzzle of Currency Pegs. *Journal of Monetary Economics*, 49(3):533–569.

Woodford, M. (2001). The Taylor Rule and Optimal Monetary Policy. *American Economic Review*, 91(2):232–237.

Selbständigkeitserklärung

Ich erkläre hiermit, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle Koautorenschaften sowie alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche gekennzeichnet. Mir ist bekannt, dass andernfalls der Senat gemäss Artikel 36 Absatz 1 Buchstabe o des Gesetzes vom 5. September 1996 über die Universität zum Entzug des aufgrund dieser Arbeit verliehenen Titels berechtigt ist.

Martin Straub