Essays in Econometrics and Macroeconomics

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von
Jörn Tenhofen
aus Rhede (Westf.)

Bonn 2011

Dekan: Prof. Dr. Klaus Sandmann
Erstreferent: Prof. Dr. Jörg Breitung
Zweitreferent: Prof. Monika Merz, Ph.D.

Tag der mündlichen Prüfung: 15. Februar 2011

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn
(http://hss.ulb.uni-bonn.de/diss_online) elektronisch publiziert.

to my parents

Acknowledgments

This dissertation would not have been possible without the help and support of many people. I owe all of them a large debt of gratitude.

First and foremost, I am deeply grateful to my two main advisors, Jörg Breitung and Monika Merz, for their continuing support, guidance, and advice throughout the dissertation process. I would especially like to thank them for tolerating my "disparate" research interests. Jörg Breitung gave me the freedom and support to develop and follow my own research agenda. He was always ready to discuss my work and other related issues and gave invaluable advice whenever I needed it. I learned a lot from him about doing research – in econometrics in particular –, creative work, and beyond. Monika Merz made countless suggestions and comments to improve my work at all stages of the dissertation project. She forced me to really think all issues through and alerted me to the important role played by the presentation of a research idea. I am particularly indebted to her for her tireless efforts to make my unique and extremely stimulating research visit at Columbia University possible.

I would also like to thank my co-authors, Jörg Breitung and Guntram Wolff, for their excellent cooperation and challenging as well as stimulating discussions. I learned a lot during our collaboration, about the respective topics and conducting research more generally.

During the years, the Institute of Econometrics has almost become a second home for me. Its pleasant and stimulating atmosphere provided the perfect basis for my work, as well as a starting point for some recreational activities. Many thanks to all present and former members of the Institute, especially Heide Baumung, Benjamin Born, Jörg Breitung, Norbert Christopeit, Matei Demetrescu, Uli Homm, Michael Massmann, as well as Christian and Uta Pigorsch, who made this possible!

Parts of the dissertation were written while I was a visiting scholar at Columbia University during the 2008-9 academic year. This was a fascinating and extremely stimulating experience and I would like to thank the entire Department of Economics at Columbia University for their help, support, and hospitality during my visit. In particular, I am grateful to Mike Woodford, who was so kind to sponsor my stay. I benefited a lot from his support, the open discussions with him, and his numerous helpful and constructive comments on my work. Many thanks also go to Jushan Bai, John Leahy, Stephanie Schmitt-Grohé, and Martín Uribe for insightful conversations and so many important suggestions related to my research. Financial support from the Heinrich Hertz foundation of the state of North Rhine-Westphalia and from the German Academic Exchange Service (DAAD) is gratefully acknowledged.

I am also indebted to Urs Schweizer, Jürgen von Hagen, as well as Silke Kinzig for the tireless efforts in managing the Bonn Graduate School of Economics (BGSE) and the research training group "Quantitative Economics" (Graduiertenkolleg). These institutions provided an invaluable platform for my doctoral studies and my research in general. In this regard, I would like to thank the German Research Foundation (DFG) for financial support.

Many thanks also go to my fellow graduate students as well as various members of the Department of Economics at the University of Bonn, in particular Almut Balleer, Zeno Enders, Michael Evers, Jürgen Gaul, Stefan Koch, Matthias Lux, Julian Merschen, Daniel Müller, Gernot Müller, Johannes Pfeifer, Thomas Rieck, and Philipp Wichardt, for many insightful discussions and a lot of fun.

Finally, and most importantly, I owe my family, my parents in particular, an immeasurable debt of gratitude. Without their unconditional and continuing support, I would not have gone so far. Steffi's patience, encouragement, and enduring belief in me deserves more than just gratitude.

Contents

List of Figures

List of Tables

Make everything as simple as possible,
but not simpler.

– Albert Einstein –

Introduction

When studying real-world phenomena employing models, scientists in general and economists as well as econometricians, in particular, are almost always confronted with the trade-off between completeness and manageability. On the one hand, the respective model should be as complete as possible in the sense of capturing all those aspects of reality which are considered to be relevant to the particular investigation. On the other hand, it is, however, necessary to abstract from certain features of the phenomenon, which are regarded as being unimportant, to keep the model manageable and to be able to carry out a fruitful analysis at all. While recently, driven by the increase in computing power, the trade-off tilted more towards completeness, it is nevertheless true that simplification and abstraction is still a necessary and important ingredient of model building. It would be impossible to work with a model that tries to capture *all* aspects of a real-world phenomenon up to the smallest detail, let alone, would it be possible to comprehend such a model. Furthermore, as reality already provides us with such a complete "model," the actual phenomenon itself, we would not be able to gain any insights or conclusions over and above those we could obtain from looking at the phenomenon directly. Going too far in the other direction, however, by abstracting from essential features of reality, which are relevant to the questions at hand, is similarly problematic. The structural analysis in general and the policy implications, in particular, could depend to a large extend on the chosen degree of abstraction. The crucial aspect of model building then is to strike the right balance between completeness and manageability.

In choosing this balance, the close interaction of macroeconomic theory and empirical analysis as well as econometric techniques is of particular importance. When studying empirically a certain economic relationship, a natural starting point of the econometric model is, of course, the theory underlying this relationship. It offers, for example, guidance on the relevant variables to include and on the functional relation between those variables. As a more specific example, the principle of optimization underlying macroeconomic models in conjunction with the information structure, i.e., what is known to the different agents at the respective point in time, could have implications for the data-generating process of the resulting economic series. This, in turn, could indicate the appropriate econometric model and technique to use in the empirical study. Similarly, the findings of empirical investigations could help to decide how to further refine, or even set up, theoretical models. For instance, what are the dynamic relationships between different economic variables found in the data and what are the consequences for theoretical modeling, such that it is possible to match those dynamics? Alternatively, is there empirically sufficient heterogeneity in certain subaggregates, which would warrant considering those in the theoretical model? What are the (policy) implications of choosing such a different level of abstraction? All this is affected by the level of sophistication of the econometric techniques available. New methodological developments, for example, could facilitate more precise estimates leading to stronger results and implications for model building. Furthermore, more advanced econometric methods could allow to take into account certain features of the data-generating process implied by economic theory, as indicated above.

The overall contribution of this dissertation is to illustrate the importance of choosing an appropriate balance between completeness and manageability, both in the field of macroeconomics and econometrics. In this regard, it focuses strongly on the close interaction of macroeconomic theory and empirical analysis supported by novel econometric techniques. On a more general level, this interplay is not least reflected in the fact that each of the three chapters of this dissertation has a focus on one of these three aspects. While Chapter 1 centers on econometric theory and suggests a new econometric technique, Chapter 2 is a contribution to macroeconomic theory in the area of monetary economics. Finally, Chapter 3 is a combination of the two fields in the sense of featuring an empirical analysis of a macroeconomic question using a newly suggested econometric method. Each individual chapter, however, also draws on the other aspects, thereby highlighting the aforementioned interaction. As a unifying theme, each of the chapters shows that there are potentially dramatic consequences of taking into account additional layers of reality. Those added aspects pertain to the core of the respective investigation, so that the models exhibit an increased level of completeness. This is achieved, however, without forfeiting manageability, not least due to the development and application of novel econometric methods. The particular models and techniques employed are still easy to handle and understand.

In terms of econometric theory, Chapter 1 shows that considerably more precise estimates within a so-called dynamic factor model are attainable by taking into account additional features of the data-generating process. In particular, we suggest a simple two-step estimation procedure to obtain efficient estimates in the presence of both autocorrelation and heteroskedasticity. We demonstrate that those features are in fact present in a widely used macroeconomic data set and illustrate the superior performance of our estimator via a simulation exercise based on this set of time series. Moreover, the dynamic factor model itself is a nice example of how to reconcile the, in most cases, conflicting goals of completeness and manageability. Factor models are based on the idea that a potentially very large set of time series can be represented as the sum of two (unobservable) parts: first, the common component, which is ultimately driven by a small number of common factors shared by the entire panel, and second, the idiosyncratic component, driven by shocks only relevant to the specific series. Due to this separation, the information contained in an exhaustive set of time series can be easily summarized by this small set of factors and utilized for structural analysis and forecasting.

With respect to macroeconomic theory, Chapter 2 considers a New-Keynesian dynamic stochastic general equilibrium (DSGE) model featuring labor market frictions. Within this setup, Chapter 2 highlights the dramatic consequences for equilibrium allocations and optimal monetary policy when replacing the standard approach of a uniformly rigid real wage by heterogeneous wage setting with different degrees of rigidity. The introduction of the latter is motivated by empirical evidence and correspondingly implemented by distinguishing new hires and ongoing workers. This emphasizes once more the close interplay of empirical analysis and macroeconomic model building. With only these minor changes compared to the standard setup and despite an economy-wide average sticky wage, the sizable short-run inflation unemployment trade-off which is obtained in the original model with a uniformly sticky wage disappears. This profoundly affects the optimal conduct of monetary policy. It leaves the monetary authority with a single target so that it can solely focus on inflation with no concern for employment stabilization. The costs of this increase in the level of completeness with respect to such an important aspect of the model, i.e., the wage setting mechanism, are small as manageability basically does not change compared to the original setup. Overall, this chapter illustrates that policy implications derived in a particular model might depend to a large extend on the chosen degree of abstraction.

Finally, the empirical investigation presented in Chapter 3 highlights the importance of taking into account particularities of the information structure as well as of focusing on subcomponents of certain fiscal aggregates when estimating the effects of fiscal policy on the macroeconomy. In particular, we suggest a new empirical approach based on a structural vector autoregression (SVAR), which explicitly allows for the fact that major fiscal policy measures are typically anticipated. Moreover, our investigation indicates that it is crucial to distinguish those subcomponents of total government spending, which might have different effects on the macroeconomy as implied by economic theory. Those two ingredients allow us to reconcile the conflicting results obtained in the literature based on the narrative and standard SVAR approaches, in particular with respect to the consumption response to an increase in government spending. These approaches just take into account either anticipation issues or disaggregate variables but not both. Thus, our findings again illustrate the important role played by the chosen level of abstraction. While at a certain level, the findings of the different approaches seem to be in conflict with each other, at another level, i.e., when allowing for fiscal policy anticipation and considering subcomponents of government spending, the antagonism vanishes. Within this chapter, the interaction of macroeconomic theory, empirical analysis, and novel econometric techniques is particularly rich. The necessity to augment the econometric model in order to account for fiscal policy anticipation results from economic and institutional considerations. In particular, fiscal policy actions are usually known before they are actually implemented. This is mostly due to the extensive public debate typically preceding political decisions, but it also results from the fact that many measure are usually introduced at a certain date, e.g., next January 1st. Optimizing agents, in turn, adjust their plans as soon as they learn about the respective measure and do not wait until implementation. This special information structure must be taken into account when estimating the dynamic relationships between the relevant macroeconomic variables. Intuitively, the econometrician needs the same amount of information as the private agents in order to be able to uncover the dynamics correctly. Using a standard VAR might not be sufficient in this regard. Moreover, distinguishing different subcomponents of total government spending is motivated by theoretical findings of the macroeconomic literature and our empirical results indeed correspond to those findings. The empirical results, in turn, also have implications for macroeconomic modeling. Since the results of the standard fiscal VAR literature are difficult to reconcile with benchmark macroeconomic models, the literature recently increased efforts to align those models with the aforementioned empirical results. Our findings at least raise the question, whether this is a promising way to proceed.

After having discussed the contributions of the different chapters from a global perspective, i.e., with respect to the unifying theme of the dissertation, the remainder of the introduction focuses on each of the chapters individually and summarizes the respective contributions and main findings.

Chapter 1.[1] Dynamic factor models can be traced back to the work of Sargent and Sims (1977) and Geweke (1977), where only systems with a small number of time series are considered. Recent work by Forni, Hallin, Lippi, and Reichlin (2000) and Stock and Watson (2002a, 2002b) extends the setup to large dimensional panels, so that both the time series dimension as well as the number of cross section units are potentially large. Important contributions such as Bai and Ng (2002) and Bai (2003) consolidate this development and lay the foundations for the success of this class of models in areas such as macroeconomic forecasting and structural analysis.

Consistent estimates of the parameters of the model under the weak assumptions of an approximate factor model (Chamberlain and Rothschild 1983) can be ob-

[1]This chapter is based on a joint paper with Jörg Breitung (Breitung and Tenhofen 2010).

tained by employing either the standard principal component (PC) estimator (Stock and Watson 2002a, Bai 2003) or Forni, Hallin, Lippi, and Reichlin's (2000) dynamic principal component estimator. The situation concerning the *efficient* estimation of those parameters, however, is not as clear-cut. This is particularly true when moving away from the rather strong assumption of Gaussian i.i.d. errors. While there are some suggestions for the cases when the errors are *either* heteroskedastic (Boivin and Ng 2006, Doz, Giannone, and Reichlin 2006, Choi 2008) *or* autocorrelated (Stock and Watson 2005), simple approaches that allow for both of those features are nonexistent. This is all the more important since, as we show in Chapter 1, the idiosyncratic components obtained from typical data sets, such as the one of Stock and Watson (2005), indeed feature a considerable amount of heterogeneity with respect to their (sample) variances and first order autocorrelations.

In order to obtain efficient parameter estimates of the dynamic factor model in the presence of both autocorrelation and heteroskedasticity, in Chapter 1 a simple two-step estimation procedure is suggested. We derive the asymptotic distribution of the resulting estimators, investigate the asymptotic efficiency relative to standard PC, and study the performance of the different estimators in small samples via Monte Carlo simulations.

The two-step estimator is derived from an approximate Gaussian log-likelihood function. In particular, the approximating model features mutually uncorrelated idiosyncratic components, but it allows for both individual specific autocorrelations and variances. The resulting estimator employs standard PC in the first stage in order to obtain preliminary estimates of the common factors and factor loadings. Intuitively, PC can be considered as an ordinary least squares (OLS)-like estimator, as it does not take into account the covariance structure of the errors. In the second stage, generalized least squares (GLS)-type transformations are applied, yielding the ultimate two-step PC-GLS estimates of the common factors and factor loadings. Interestingly, when estimating the factors, it is only necessary to take into account possible heteroskedasticity of the errors, whereas the loadings are estimated using just the traditional GLS transformation for autocorrelated errors. Not having to compute the full two-way GLS transformations with respect to both autocorrelation and heteroskedasticity for the respective estimator highlights the simplicity of our approach, which furthermore enables fast computation.

In contrast to the assumptions underlying the aforementioned approximating model, which is employed to derive the estimator, our main results concerning the asymptotic distribution are obtained under much weaker assumptions. The idiosyncratic components, for example, are allowed to be weakly correlated in the sense of Bai and Ng (2002) and Stock and Watson (2002a). With respect to the asymptotic distribution of the two-step estimator, we show that it is not affected by the estimation error in the regressors, i.e., in the estimated covariance parameters and the PC estimates of the factors or loadings. Thus, the feasible two-step PC-GLS estimator is asymptotically as efficient as the estimator that (locally) maximizes the full approximate likelihood function. To obtain small sample gains in efficiency, the two-step estimator can be iterated, using the second stage estimates in future steps as well as improved estimates of the covariance parameters based on the second step residuals.

With respect to the relative asymptotic efficiency, it can only be shown that the PC-GLS estimators are at least as efficient as (and generally more efficient than) the standard PC estimators if the temporal and contemporaneous variance and covariance functions of the errors are correctly specified. In order to obtain an estimator which is always at least as efficient as standard PC and two-step PC-GLS individually, we suggest a generalized method of moments (GMM) estimator based on the two sets of moment conditions corresponding to the aforementioned estimators.

While it is a theoretical possibility that there are situations when standard PC is asymptotically more efficient than two-step PC-GLS, the extensive Monte Carlo simulations presented in Chapter 1 indicate that this is unlikely to occur in practice, even if the covariance functions are misspecified. We compare the performance of the various estimators in different scenarios featuring autocorrelation, heteroskedasticity, as well as cross-sectional correlation. As a final simulation experiment, we generate data based on the widely used set of time series provided by Stock and Watson (2005). This allows to examine the performance of the respective estimators when applied to more “realistic” data sets. In all those simulations, we document the superior performance of the two-step PC-GLS estimator and particularly its iterated version compared to standard PC estimation.

Chapter 2.[2] A striking aspect of the standard version of the New-Keynesian DSGE model, the workhorse model of monetary policy analysis, is the absence of interesting dynamics and even concepts with respect to the labor market. In its basic version, it employs a Walrasian labor market and thus lacks equilibrium unemployment and dynamics in related variables, even though these features are an important aspect of the business cycle. Recent research has begun to address this shortcoming by integrating labor market frictions into the model (e.g., Krause and Lubik 2007, Trigari 2009, Blanchard and Galí 2010, Christoffel and Linzert 2010), where the wage-determination mechanism is a particularly important ingredient. Motivated by the so-called "unemployment volatility puzzle," the standard approach in this regard is to employ an overall rigid real wage. The aforementioned puzzle, most visibly documented by Shimer (2005), describes the difficulty of the baseline Diamond-Mortensen-Pissarides style search and matching model to generate fluctuations in unemployment and vacancies which are consistent with the data. Recent contributions by Haefke, Sonntag, and van Rens (2008) and Pissarides (2009), however, argue that it is inconsistent with empirical evidence to use such a uniformly rigid real wage. They find that the wages for workers in ongoing job relationships are indeed rigid, but those of new hires are highly cyclical. As the latter kind of wages are the relevant ones for search and matching models, they conclude that wage rigidity cannot be the answer to the unemployment volatility puzzle. Instead, Pissarides (2007) recommends augmenting the model by additional driving forces.

Correspondingly, Chapter 2 investigates optimal monetary policy in an environment characterized by labor market frictions, heterogeneous wage setting, as well as markup shocks. In particular, this chapter features two main contributions. First, it studies the consequences for equilibrium allocations, particularly for labor market dynamics and optimal monetary policy, of employing heterogeneous wage setting which is consistent with the empirical findings of the aforementioned authors. Second, this chapter investigates the implications for the dynamic responses of inflation and unemployment under different monetary policy regimes of adding markup shocks as

[2]This chapter is based on Tenhofen (2010).

additional driving forces to the model.

In the first part of the chapter, I introduce heterogeneous wage setting into the New-Keynesian DSGE model of Blanchard and Galí (2010), which features labor market frictions in terms of hiring costs. While Blanchard and Galí (2010) follow the traditional approach in the literature of employing a uniformly rigid wage, I distinguish two kinds of workers in order to introduce some degree of wage heterogeneity. In particular, I distinguish between workers in ongoing job relationships and newly hired workers. Consistent with the empirical studies of Haefke, Sonntag, and van Rens (2008) as well as Pissarides (2009), the former earn a rigid wage in the spirit of Hall (2005), whereas the latter bargain over the wage for the current period, modeled by employing the generalized Nash solution. As the main finding it emerges that the sizable short-run inflation unemployment trade-off, which is obtained in the original setting with a uniformly rigid wage, disappears. This results even though I change the setup of Blanchard and Galí (2010) only to a small extend and despite an economy-wide average wage which is still sticky. As a result, employing a form of wage rigidity consistent with empirical findings has profound effects on the policy implications of this model, in particular, with respect to the optimal conduct of monetary policy. It is left with a single target, so that it can concentrate exclusively on inflation with no concern for employment stabilization.

Nevertheless, the question remains, how to address the unemployment volatility puzzle and what are the consequences of a corresponding mechanism for monetary policy. Hence, in the second part of the chapter, I follow the suggestion of Pissarides (2007) and add markup shocks as additional driving forces to the aforementioned DSGE model with heterogeneous wage setting. This is achieved by assuming a stochastic elasticity of substitution in the Dixit-Stiglitz constant-elasticity-of-substitution consumption aggregator à la Steinsson (2003) and Rotemberg (2008). The resulting markup fluctuations are consistent with empirical evidence as documented by Rotemberg and Woodford (1991, 1999) and Galí, Gertler, and López-Salido (2007). A short-run inflation unemployment trade-off arises and I investigate the dynamics of the model under different monetary policy regimes. The main finding of this part of Chapter 2 is that within this model featuring labor market frictions, heterogeneous wage setting, and markup shocks, optimal policy is characterized by an overriding focus on inflation stabilization. This result is in line with much of the recent literature on optimal monetary policy (e.g., Woodford 2003), but contrasts with the findings of Blanchard and Galí (2010). Furthermore, markup shocks are not able to generate an extensive amount of fluctuations in unemployment within the setup considered in this chapter.

Chapter 3.[3] When considering the empirical literature on the effects of fiscal policy on the macroeconomy, rather conflicting results emerge. On the one hand, the narrative approach typically finds that GDP increases while private consumption and real wages fall in response to shocks to government expenditure (Ramey and Shapiro 1998, Edelberg, Eichenbaum, and Fisher 1999, Burnside, Eichenbaum, and Fisher 2004). This approach uses dummy variables that indicate large (exogenous) increases in government spending related to wars. On the other hand, the findings of the SVAR literature are that GDP as well as private consumption usually increase in response to a shock to government spending (Blanchard and Perotti 2002, Perotti 2005, 2008). The SVARs are typically identified by assuming that government expenditure are predetermined within the quarter. In sum, the main difference concerns the consumption response to a shock to government spending. While the findings of the narrative approach are readily aligned with theoretical predictions of standard macroeconomic models, both of the neoclassical (Baxter and King 1993) and most New-Keynesian (Linnemann and Schabert 2003) variants, this is not so easy when considering the SVAR results. Recently, however, there have been efforts to reconcile current business cycle models with the latter strand of the empirical literature (Galí, López-Salido, and Vallés 2007). The crucial aspect is to generate a positive consumption response to an increase in government spending. All in all, these conflicting results of the empirical literature constitute a rather unfortunate situation, as empirical findings shape our modeling efforts and understanding of the economy.

The starting point of the investigation presented in Chapter 3 is the contribution of Ramey (2009). Her explanation for the different results is that VARs miss the fact that major changes in government expenditure are typically anticipated. On a more

[3] This chapter is based on joint work with Guntram Wolff (Tenhofen and Wolff 2010).

general level, this corresponds to the fundamental problem that in certain settings a misalignment of the information sets of private agents and the econometrician arises. With respect to the anticipation of fiscal policy, this means that private agents not only know the variables observed by the econometrician, but in addition have information on the fiscal shocks occurring in future periods. This misalignment of information sets could impair the ability of standard VARs to recover the actual economic shocks, so that tools based on those econometric models may yield incorrect inferences (e.g., Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson 2007, Leeper, Walker, and Yang 2009).

Chapter 3 investigates the response of private consumption to fiscal shocks within an SVAR framework, taking into account anticipation of fiscal policy actions. In order to avoid the problems of standard VARs, we suggest a new empirical approach which is designed to align the information sets of the private agents and the econometrician. A simulation study based on a theoretical model featuring (imperfect) fiscal foresight is performed, in order to illustrate the ability of this method to correctly capture macroeconomic dynamics. Finally, we present an application to real life data, with a particular focus on the response of private consumption to shocks to different subcomponents of government spending.

With respect to the empirical approach, we start out from the well established SVAR setup suggested by Blanchard and Perotti (2002), but augment it to explicitly take into account perfectly anticipated fiscal policy one period in advance. This is achieved, in particular, by adding expectation terms for next period's fiscal variables as well as equations modeling the formation of those expectation to a standard AB-model SVAR. The crucial point is that the aforementioned expectation are formed with respect to an information set which not only includes current and past endogenous variables of the system but also next period's fiscal shocks. This reflects the special information structure due to fiscal policy anticipation.

As the information structure is generally unknown in practice, it is interesting to investigate the robustness of our methodology to possible deviations from the assumption of one period perfect foresight. Consequently, we simulate data based on a standard neoclassical growth model featuring both anticipated and unanticipated fiscal shocks, so that private agents only have imperfect foresight. By comparing the theoretical impulse responses of the model to the estimated ones, it is possible to check whether the respective approach is able to capture anticipation effects. In particular, the estimated impulse responses are obtained by applying both a VAR à la Blanchard and Perotti (2002) as well as our expectation augmented VAR to the simulated data. The main finding of this part of the chapter is that the new approach, in contrast to the standard VAR, delivers impulse responses which are very close to the theoretical ones. This holds not just in a setup when there are only anticipated fiscal shocks, perfectly corresponding to the underlying assumptions of the expectation augmented VAR, but also in the case of imperfect foresight. Thus, this exercise indicates that the approach is robust to situations with a potentially different information structure.

In the subsequent application to real life data, we distinguish in particular government defense and non-defense expenditure. This is motivated by economic theory, as we expect the response of private consumption to be different to rather wasteful defense and potentially productive non-defense expenditure. Indeed, the results indicate that it is important in empirical work to allow for anticipation of fiscal policy and, particularly, it is crucial to distinguish subcomponents of total government spending. As expected from economic theory, we find private consumption to decrease significantly in an expectation augmented VAR in response to a shock to defense expenditure, whereas it increases significantly to non-defense expenditure shocks. Both findings are in line with Ramey's (2009) general argument and the former result corresponds to her findings for the narrative approach. Consequently, by distinguishing defense and non-defense expenditure it is possible to reconcile the results of the narrative and SVAR approaches to the study of fiscal policy effects. On the other hand, when considering total government spending, the resulting impulse responses are not as clear-cut, since we lump together items with different macroeconomic effects. Moreover, the responses resulting from a VAR à la Blanchard and Perotti (2002) are all rather weak and mostly insignificant, highlighting the importance of anticipation issues.

While this introduction summarizes the key contributions and main findings of this dissertation and highlights its unifying theme, the subsequent three chapters are each developed in an independent and self-contained way.

Chapter 1

GLS estimation of dynamic factor models

1.1 Introduction

Since the influential work of Forni, Hallin, Lippi, and Reichlin (2000), Stock and Watson (2002a, 2002b), Bai and Ng (2002), and Bai (2003), dynamic factor models have become an important tool in macroeconomic forecasting (e.g., Watson 2003, Eickmeier and Ziegler 2008) and structural analysis (e.g., Giannone, Reichlin, and Sala 2002, Bernanke, Boivin, and Eliasz 2005, Eickmeier 2007). Under the weak assumptions of an approximate factor model (Chamberlain and Rothschild 1983), the parameters of the models can be consistently estimated by applying the traditional principal component (PC) estimator (Stock and Watson 2002a, Bai 2003) or – in the frequency domain – by using the dynamic principal component estimator (Forni, Hallin, Lippi, and Reichlin 2000). Assuming Gaussian i.i.d. errors, the PC estimator is equivalent to the ML estimator and, therefore, the PC estimator is expected to share its asymptotic properties. It is well known that a generalized least squares (GLS)-type criterion function yields a more efficient estimator than the OLS based PC estimator if the errors are heteroskedastic (e.g., Boivin and Ng 2006, Doz, Giannone, and Reichlin 2006, Choi 2008). It is less clear how the estimator can be improved in the case of serially correlated errors. Stock and Watson (2005) suggest a GLS transformation similar to the one that is used to correct for autocorrelation in the linear regression model. However, as we will argue below, this transformation affects the static representation of the factor model.

In this chapter, we consider the Gaussian (pseudo) ML estimator in models, where the errors are assumed to be heteroskedastic and autocorrelated. We derive the first order conditions for a local maximum of the (approximate) log-likelihood function and show that the resulting system of equations can be solved by running a sequence of GLS regressions. Specifically, the factors can be estimated by taking into account possible heteroskedasticity of the errors, whereas the factor loadings are estimated by using the usual GLS transformation for autocorrelated errors. We show that the feasible two-step GLS estimation procedure is asymptotically equivalent to the estimator that locally maximizes the approximate likelihood function. In small samples, however, our Monte Carlo simulations suggest that the iterated PC-GLS estimator can be substantially more efficient than the simpler two-step estimator. In a related paper, Jungbacker and Koopman (2008) consider the state space representation of the factor model, where the number of variables (N) is fixed and the vector of common factors has a VARMA representation. As we will argue below, as $N \to \infty$ their (exact) ML estimator converges to the approximate ML estimator suggested in this chapter. Thus, the two-step GLS estimator can be seen as a simplification of the exact ML approach proposed by Jungbacker and Koopman (2008) as N gets large. Furthermore, we do not specify a particular parametric model for the vector of common factors, as the data generating process of the factors becomes irrelevant as $N \to \infty$. Accordingly, our approach sidesteps the problem of choosing an appropriate lag length for the VARMA representation of the factors.

It may be argued that in practice the efficiency gain from taking into account serial correlation and heteroskedasticity may be small if the variances of the idiosyncratic components are similar and their autocorrelations are small. To assess the potential of the suggested estimator, we therefore consider the distribution of the variances and first order autocorrelations estimated from the widely used data set provided by Stock and Watson (2005). This data set contains 132 monthly US series including measures of real economic activity, prices, interest rates, money and credit aggregates, stock prices, and exchange rates. The sampling period runs from 1960 to 2003.[1] As

[1]The original data set is provided for the years 1959 to 2003. Some observations are, however,

usual, the time series are differenced if unit root tests are not able to reject the null hypothesis of nonstationarity. Applying the information criteria of Bai and Ng (2002) suggests that the number of common factors is $r = 7$. The idiosyncratic component is obtained by subtracting the estimated common component from the standardized series. The resulting histograms with respect to sample variances and first order autocorrelations of the idiosyncratic components are presented in Figures 1.1 and 1.2. Since the variables are standardized, the variances of the idiosyncratic components are identical to $1 - c_i$, where c_i is the "commonality" of variable i. A value of c_i close to zero implies that the factors do not contribute to the variance of the time series. In our example, 13 percent of the variables have a commonality less than 0.05. Furthermore, the variances do not seem to be concentrated around some common value. Accordingly, ignoring the heteroskedasticity in the data will lead to a severe loss of efficiency. A similar picture emerges for the autocorrelations of the idiosyncratic errors. Most of the estimates are far away from zero. Moreover, there is substantial heterogeneity among the estimates, suggesting that the model should allow for individual specific autocorrelations. In order to investigate the impact of those features of the data and to illustrate the potential of our suggested estimators, one of the Monte Carlo experiments presented in the latter part of this chapter is based on Stock and Watson's (2005) data set.

The rest of Chapter 1 is organized as follows. In Section 1.2, we consider some prerequisites of the dynamic factor model. Section 1.3 introduces the PC-GLS estimator and Section 1.4 studies the asymptotic distribution of the two-step estimator. The relative asymptotic efficiency of the standard PC and PC-GLS estimators is investigated in Section 1.5. The small sample properties of alternative estimators are compared by means of Monte Carlo simulations in Section 1.6. Finally, Section 1.7 concludes.

---

missing in 1959. We therefore decided to use a balanced data set starting in 1960.

Figure 1.1: Histogram of the sample variances

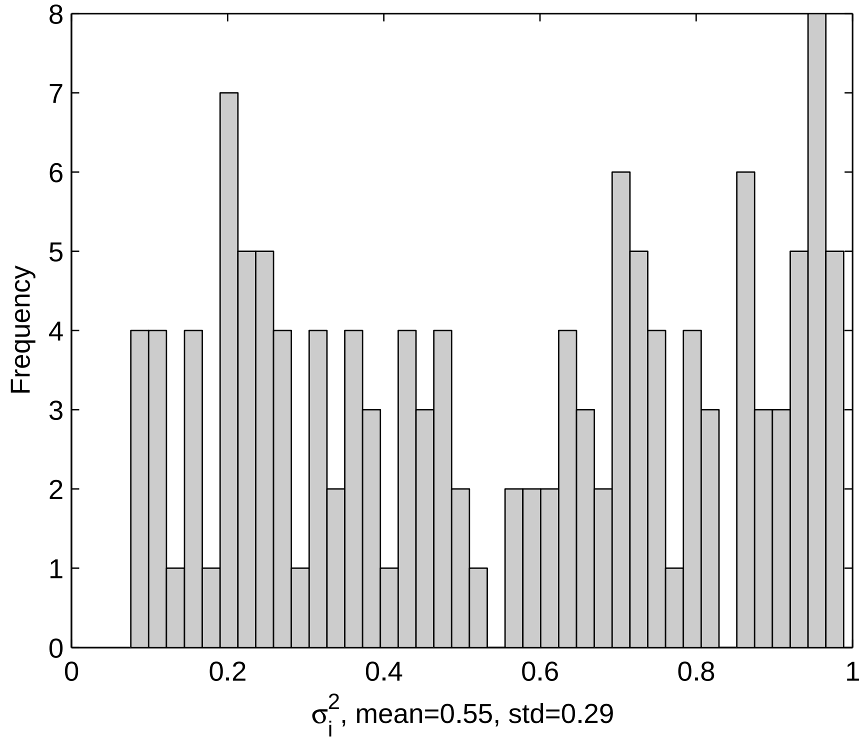
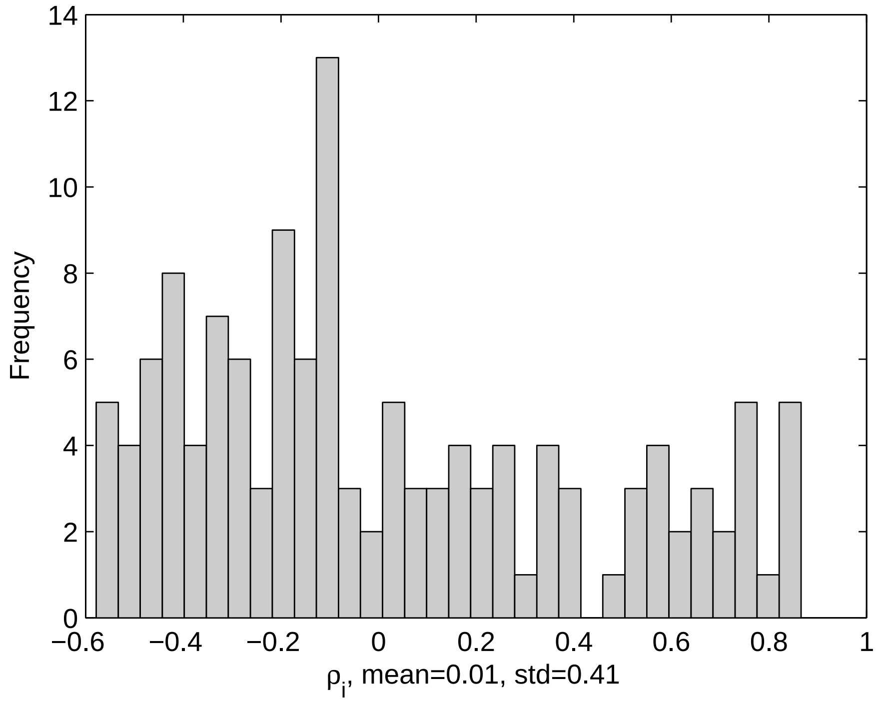


Figure 1.2: Histogram of the sample autocorrelations



1.2 The dynamic factor model

Following Stock and Watson (2002a, 2002b) and Bai and Ng (2002), we consider the dynamic factor model

$$x_{it} \quad = \quad \theta_i(L)'g_t + e_{it} \ , \tag{1.1}$$

where x_{it} is the i'th variable ($i = 1, \dots, N$) observed in period t ($t = 1, \dots, T$), g_t is a $k \times 1$ vector of dynamic factors, and $\theta_i(L) = \theta_{0i} + \theta_{1i}L + \cdots + \theta_{mi}L^m$ is a $k \times 1$ polynomial of factor loadings. As usual in this literature, we ignore possible deterministic terms and assume $E(x_{it}) = E(e_{it}) = 0$.

Let $\Theta(L) = \Theta_0 + \Theta_1 L + \cdots + \Theta_m L^m$ and $\Theta_j = [\theta_{j1}, \dots, \theta_{jN}]'$ ($j = 0, \dots, m$) and define $G_t = [g_t', \dots, g_{t-m}']'$. The static factor representation results as

$$X_t = \Theta G_t + e_t \ ,$$

where $X_t = [x_{1t}, \dots, x_{Nt}]'$, $\Theta = [\Theta_0, \dots, \Theta_m]$, and $e_t = [e_{1t}, \dots, e_{Nt}]'$. It is important to note that Θ need not have full column rank. For example, a subset of the factors may not enter with all lags. In this case the respective columns of Θ are zero. Let $r \leq (m+1)k$ be the rank of the matrix Θ. Then there exists an $N \times r$ matrix Λ such that $\Theta G_t = \Lambda F_t$, where $F_t = RG_t$ and R is a nonsingular $r \times (m+1)k$ matrix. F_t is called the vector of *static factors*.

Finally, in full matrix notation the model is written as

$$X = F\Lambda' + e, \tag{1.2}$$

where $X = [X_1, \dots, X_T]'$ and $e = [e_1, \dots, e_T]'$ are $T \times N$ matrices. The columns of the $T \times r$ matrix $F = [F_1, \dots, F_T]'$ collect the observations of the r static factors.

Under fairly weak assumptions, the factors and factor loadings can be estimated consistently as $N \to \infty$ and $T \to \infty$ by the PC estimator that minimizes the total sum of squares

$$S(F, \Lambda) = tr\left[(X - F\Lambda')'(X - F\Lambda')\right],$$

subject to the constraint $T^{-1}F'F = I_r$ (Stock and Watson 2002b, Bai and Ng 2002). The estimators of F and Λ result as $\widehat{F} = \sqrt{T}\,\widehat{V}_r$ and $\widehat{\Lambda} = T^{-1/2}X'\widehat{V}_r$, respectively, where $\widehat{V}_r$ is the matrix of the r orthonormal eigenvectors corresponding to the r

largest eigenvalues of the matrix XX' (e.g., Stock and Watson 2002b). The resulting estimators for Λ and F will be called the PC-OLS estimators.

If the idiosyncratic errors are heteroskedastic or autocorrelated, the PC-OLS estimator is not efficient. Before introducing our estimator, we briefly discuss existing proposals for efficient estimation in the presence of either heteroskedastic or autocorrelated errors. First, for the heteroskedastic case, Doz, Giannone, and Reichlin (2006) and Choi (2008) suggest GLS-type estimators that minimize the weighted sum of squares

$$S(F, \Lambda, \Omega) = tr\left[\Omega^{-1}(X - F\Lambda')'(X - F\Lambda')\right] ,$$

where Ω =diag$[E(e_{1t}^2), \ldots, E(e_{Nt}^2)]$ for all t. Forni, Hallin, Lippi, and Reichlin (2005) and Choi (2008) consider the case of an arbitrary covariance matrix Ω. It should be noted, however, that the factors are not identified without additional assumptions on the matrix Ω. To see this, consider the spectral decomposition $\Omega = \sum_{i=1}^{N} \mu_i v_i v_i'$, where μ_i and v_i denote the i'th eigenvalue and corresponding eigenvector, respectively. The matrix Ω may be decomposed in form of a factor model yielding $\Omega = \Gamma\Gamma' + \Omega^*$ where, for example, $\Gamma\Gamma' = \sum_{i=1}^{k} w_i \mu_i v_i v_i'$, $k \leq N$, $0 < w_i < 1$ for all i, and

$$\Omega^* = \sum_{i=1}^{k}(1 - w_i)\mu_i v_i v_i' + \sum_{i=k+1}^{N} \mu_i v_i v_i'$$

is a symmetric positive definite matrix. Thus, $E(X_t X_t') = \Lambda\Lambda' + \Omega = \Lambda^*\Lambda^{*'} + \Omega^*$, where $\Lambda^* = [\Lambda, \Gamma]$. In order to distinguish the common factors from the idiosyncratic components, the covariance matrix Ω has to be restricted in such a way that the idiosyncratic errors cannot mimic the pervasive correlation due to the common factors. This is usually ensured by assuming that all eigenvalues of Ω are bounded as $N \to \infty$. One possibility in this regard is to specify Ω as a diagonal matrix, which is what we do in our approach. Another possibility is to allow for some spatial correlation of the form $\Omega = \sigma^2(I_N - \varrho W_N)(I - \varrho W_N')$, where all eigenvalues of the spatial weight matrix W_N are smaller than one and $0 \leq \varrho \leq 1$ (e.g., Chudik, Pesaran, and Tosetti 2010). An additional problem is that in a model with arbitrary covariance matrix Ω, the number of parameters increases with the square of N, i.e., the model implies a large number of additional parameters that may even exceed the number of observations. Finally, the estimator $\widehat{\Omega} = T^{-1}(X - \widehat{F}\widehat{\Lambda}')'(X - \widehat{F}\widehat{\Lambda}')$ is singular and, hence, the inverse does not exist (see also Boivin and Ng 2006). As a result, when deriving our estimator we start from an approximate likelihood function featuring mutually uncorrelated idiosyncratic components, thereby following the traditional factor framework. Our main results concerning this estimator, however, are obtained under much weaker assumptions. In particular, the idiosyncratic components are assumed to be weakly correlated in the sense of Bai and Ng (2002) and Stock and Watson (2002a).

Second, to account for autocorrelated errors, Stock and Watson (2005) consider the model $\rho_i(L)x_{it} = \lambda_i'\widetilde{F}_t + \widetilde{\varepsilon}_{it}$, where $\rho_i(L) = 1 - \rho_{1,i}L - \cdots - \rho_{p_i,i}L^{p_i}$, which implies that $\widetilde{F}_t$ and $\widetilde{\varepsilon}_{it}$ enter the time series in a similar way. However, if a more general dynamic structure as in (1.1) is assumed, the approach suggested by Stock and Watson (2005) yields

$$\rho_i(L)x_{it} = \theta_i^*(L)'g_t + \varepsilon_{it}, \tag{1.3}$$

where $\theta_i^*(L) = \rho_i(L)\theta_i(L)$ and $\varepsilon_{it} = \rho_i(L)e_{it}$ is white noise. In general, this transformation increases the number of static factors. As an example, assume that the scalar factor g_t enters with a single lag (i.e., $m = 1$) and the autoregressive lag order is $p_i = 1$ for all i. Since $\theta_i^*(L)'g_t = [\theta_{0i}, \theta_{1i} - \theta_{0i}\rho_i, -\theta_{1i}\rho_i][g_t, g_{t-1}, g_{t-2}]'$, the number of static factors is $r = 3$, whereas the original model implies only $r = 2$ static factors.[2]

In the following section, we propose a GLS-type estimator which in contrast to earlier work focusing on either heteroskedastic (Forni, Hallin, Lippi, and Reichlin 2005, Doz, Giannone, and Reichlin 2006, Choi 2008) *or* autocorrelated errors (Stock and Watson 2005) accommodates both features.

1.3 The PC-GLS estimator

In this section, we follow Stock and Watson (2005) and assume that the idiosyncratic components have a stationary heterogeneous autoregressive representation of the form

$$\rho_i(L)e_{it} = \varepsilon_{it}, \tag{1.4}$$

where $\rho_i(L)$ is defined above. It is important to note that (1.4) is employed as an auxiliary model to capture the main features of the idiosyncratic dynamics. Our

[2] Only if all autoregressive coefficients are the same, i.e., $\rho_i = \rho$ for all i, then this representation implies two static factors $F_t^* = [g_t - \rho g_{t-1}, g_{t-1} - \rho g_{t-2}]'$.

asymptotic analysis allows for misspecification of the dynamic process that gives rise to some remaining autocorrelation in ε_{it}.

The autoregressive structure of the idiosyncratic component can be represented in matrix format by defining the $(T - p_i) \times T$ matrix

$$R(\rho^{(i)}) = \begin{bmatrix} -\rho_{p_i,i} & -\rho_{p_i-1,i} & -\rho_{p_i-2,i} & \cdots & 1 & 0 & 0 & \cdots \\ 0 & -\rho_{p_i,i} & -\rho_{p_i-1,i} & \cdots & -\rho_{1,i} & 1 & 0 & \cdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \ddots & \cdots \end{bmatrix}.$$

Thus, the autoregressive representation (1.4) is written in matrix form as

$$R(\rho^{(i)})e_i = \varepsilon_i \ , \tag{1.5}$$

where $\varepsilon_i = [\varepsilon_{i,p_i+1}, \ldots, \varepsilon_{iT}]'$ and $e_i = [e_{i1}, \ldots, e_{iT}]'$. Furthermore, we do not impose the assumption that the idiosyncratic errors have the same variances across i and t, but assume that $\sigma_i^2 = E(\varepsilon_{it}^2)$ may be different across i.

We do not need to make specific assumptions about the dynamic properties of the vector of common factors, F_t. Apart from some minor regularity conditions the only consequential assumption that we have to impose on the factors is that they are weakly serially correlated (Assumption 1 in Section 1.4).

Consider the approximate Gaussian log-likelihood function:

$$S(F, \Lambda, \rho, \Sigma) = -\sum_{i=1}^{N} \frac{T - p_i}{2} \log \sigma_i^2 - \sum_{i=1}^{N} \sum_{t=p_i+1}^{T} \frac{(e_{it} - \rho_{1,i}e_{i,t-1} - \ldots - \rho_{p_i,i}e_{i,t-p_i})^2}{2\sigma_i^2}, \tag{1.6}$$

where $\Sigma = \text{diag}(\sigma_1^2, \ldots, \sigma_N^2)$. Note that this likelihood function results from conditioning on the p_i initial values. If x_{it} is normally distributed and $N \to \infty$, then the PC-GLS estimator is asymptotically equivalent to the ML estimator. This can be seen by writing the log-likelihood function as $\mathcal{L}(X) = \mathcal{L}(X|F) + \mathcal{L}(F)$, where $\mathcal{L}(X|F)$ denotes the logarithm of the density function of $x_{11}, \ldots, x_{NT}$ conditional on the factors F and $\mathcal{L}(F)$ is the log-density of $(F_1', \ldots, F_T')$. Since $\mathcal{L}(X|F)$ is $O_p(NT)$ and $\mathcal{L}(F)$ is $O_p(T)$, it follows that as $N \to \infty$ maximizing $L(X|F)$ is equivalent to maximizing the full likelihood function $\mathcal{L}(X)$.

An important challenge for the maximization of this likelihood function is that the likelihood function is unbounded in general (see e.g., Anderson 1984, p. 570). To see this, consider a factor model with a single factor (i.e., $r = 1$). If $\widehat{F}_t = y_{it}/(T^{-1} \sum_{t=1}^{T} y_{it}^2)$

and $\widehat{\lambda}_i = 1$ for some i and $t = 1, \ldots, T$, then $\widehat{\sigma}_i^2 = 0$ and, therefore, the likelihood tends to infinity. This problem is well-known also in other fields of statistics (e.g., the estimation of mixture densities) and we adapt techniques for obtaining the maximum likelihood estimator that were developed to cope with this problem. Specifically, we are focusing on the estimator $\widehat{\theta} = (\widehat{F}_t', \widehat{\lambda}_i')'$ that attains a *local* maximum of the likelihood function. Redner and Walker (1984) provide two conditions under which the local maximum in a neighborhood of the true values θ^0 yields a consistent and asymptotically normally distributed estimator. These two conditions ensure that the likelihood function is concave in a neighborhood of the true values.

Consider the derivatives of the likelihood function:

$$g_{\lambda_i}(\cdot) = \frac{\partial S(\cdot)}{\partial \lambda_i} = \frac{1}{\sigma_i^2}\left\{\sum_{t=p_i+1}^{T} \varepsilon_{it}[\rho_i(L)F_t]\right\} \tag{1.7}$$

$$\begin{aligned} g_{F_t}(\cdot) = \frac{\partial S(\cdot)}{\partial F_t} &= \sum_{i=1}^{N} \frac{1}{\sigma_i^2}\left(\varepsilon_{it}\lambda_i - \rho_{1,i}\varepsilon_{i,t+1}\lambda_i - \cdots - \rho_{p_i,i}\varepsilon_{i,t+p_i}\lambda_i\right) \\ &= \sum_{i=1}^{N} \frac{1}{\sigma_i^2}[\rho_i(L^{-1})\varepsilon_{it}]\lambda_i \end{aligned} \tag{1.8}$$

$$g_{\rho_{k,i}}(\cdot) = \frac{\partial S(\cdot)}{\partial \rho_{k,i}} = \frac{1}{\sigma_i^2}\sum_{t=p_i+1}^{T} \varepsilon_{it}(x_{i,t-k} - \lambda_i' F_{t-k}) \tag{1.9}$$

$$g_{\sigma_i^2}(\cdot) = \frac{\partial S(\cdot)}{\partial \sigma_i^2} = \frac{\sum_{t=p_i+1}^{T} \varepsilon_{it}^2}{2\sigma_i^4} - \frac{T - p_i}{2\sigma_i^2}, \tag{1.10}$$

where $\varepsilon_{is} = 0$ for $s > T$. It is not difficult to verify that *Condition 1* of Redner and Walker (1984) related to the derivatives of the likelihood function is satisfied. Furthermore, the Fisher information matrix is well defined and positive definite at θ^0 (*Condition 2* of Redner and Walker 1984). It follows that the ML estimator that locally maximizes the log-likelihood function is consistent and asymptotically normally distributed. Our proposed estimator maximizes the likelihood in the neighborhood of the PC estimator. Since this estimator is consistent for a particular normalization of the parameters, the local maximizer of the log-likelihood function in the neighborhood of the PC estimator is consistent and asymptotically normally distributed.

A practical problem is the large dimension of the system consisting of $2Nr + N + \sum p_i$ equations. Accordingly, in many practical situations it is very demanding to compute the inverse of the Hessian matrix that is required to construct an iterative minimization algorithm. We therefore suggest a simple two-step estimator that is asymptotically equivalent to locally maximizing the Gaussian likelihood function.

Let us first assume that the covariance parameters ρ and Σ are known. The (infeasible) two-step estimators $\widetilde{F}_t$ $(t = 1, \ldots, T)$ and $\widetilde{\lambda}_i$ $(i = 1, \ldots, N)$ that result from using PC in the first stage, are obtained by solving the following sets of equations:

$$\begin{aligned} g_{F_t}(\widehat{\Lambda},\ \widetilde{F}_t\ , \rho, \Sigma) &= 0 &(1.11)\\ g_{\lambda_i}(\ \widetilde{\lambda}_i\ , \widehat{F}, \rho, \Sigma) &= 0, &(1.12)\end{aligned}$$

where $\widehat{F} = [\widehat{F}_1, \ldots, \widehat{F}_T]'$ and $\widehat{\Lambda} = [\widehat{\lambda}_1, \ldots, \widehat{\lambda}_N]'$ are the ordinary PC-OLS estimators of F and Λ.

It is not difficult to see that the two-step estimator of λ_i is equivalent to the least-squares estimator of λ_i in the regression:

$$\Big(\rho_i(L)x_{it}\Big) = \Big(\rho_i(L)\widehat{F}_t\Big)'\lambda_i + \varepsilon^*_{it} \quad (t = p_i + 1, \ldots, T), \tag{1.13}$$

where $\varepsilon^*_{it} = \varepsilon_{it} + \rho_i(L)(F_t - \widehat{F}_t)'\lambda_i$.

The two-step estimator of F_t (given $\widehat{\Lambda}$) is more difficult to understand. Consider the two-way GLS transformation that accounts for both serial correlation and heteroskedasticity:

$$\frac{1}{\sigma_i}\rho_i(L)x_{it} = \frac{1}{\sigma_i}\widehat{\lambda}_i'[\rho_i(L)F_t] + \frac{1}{\sigma_i}\varepsilon_{it}, \tag{1.14}$$

where for notational convenience we assume $p_i = p$ for all i. Furthermore, our notation ignores the estimation error that results from replacing λ_i by $\widehat{\lambda}_i$.[3]

We will argue below that in order to estimate F_t we can ignore the GLS transformation that is due to serial correlation. But let us first consider the full two-step GLS estimator of F_t that corresponds to condition (1.8). Collecting the equations for $t = p + 1, \ldots, T$, the model can be re-written in matrix notation as

$$\widetilde{X}_i = \widetilde{Z}_i f + \widetilde{\varepsilon}_i, \tag{1.15}$$

where $\widetilde{X}_i = \sigma_i^{-1}[\rho_i(L)x_{i,p+1}, \ldots, \rho_i(L)x_{iT}]'$, $\widetilde{\varepsilon}_i = \sigma_i^{-1}[\varepsilon_{i,p+1}, \ldots, \varepsilon_{iT}]'$, $\widetilde{Z}_i = \sigma_i^{-1}[\widehat{\lambda}_i{}' \otimes R(\rho^{(i)})]$, and $f = vec(F)$. The complete system can be written as

$$\widetilde{x} = \widetilde{Z}f + \widetilde{\varepsilon}, \tag{1.16}$$

[3]The complete error term is given by $\sigma_i^{-1}[\varepsilon_{it} + (\lambda_i - \widehat{\lambda}_i)'\rho_i(L)F_t]$. However, as will be shown below, the estimation error in $\widehat{\lambda}_i$ does not affect the asymptotic properties of the estimator.

where $\widetilde{x} = [\widetilde{X}_1', \ldots, \widetilde{X}_N']'$, $\widetilde{Z} = [\widetilde{Z}_1', \ldots, \widetilde{Z}_N']'$, and $\widetilde{\varepsilon} = [\widetilde{\varepsilon}_1', \ldots, \widetilde{\varepsilon}_N']'$. To see that the least-squares estimator of f is equivalent to a two-step estimator setting the gradient (1.8) equal to zero (given some initial estimator of λ_i), consider the model with only one factor (i.e., $f = F$) and $\rho_i(L) = 1 - \rho_i L$. Since

$$\sum_{i=1}^{N} \widetilde{Z}_i' \widetilde{\varepsilon}_i = \sum_{i=1}^{N} \frac{\widehat{\lambda}_i}{\sigma_i^2} \begin{bmatrix} -\rho_i & 0 & 0 & \cdots & 0 \\ 1 & -\rho_i & 0 & \cdots & 0 \\ 0 & 1 & -\rho_i & \cdots & 0 \\ \vdots & & \ddots & & \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{i2} \\ \varepsilon_{i3} \\ \varepsilon_{i4} \\ \vdots \end{bmatrix} = \sum_{i=1}^{N} \frac{\widehat{\lambda}_i}{\sigma_i^2} \begin{bmatrix} -\rho_i \varepsilon_{i2} \\ \varepsilon_{i2} - \rho_i \varepsilon_{i3} \\ \varepsilon_{i3} - \rho_i \varepsilon_{i4} \\ \vdots \\ \varepsilon_{iT} \end{bmatrix}$$

it follows that the system estimator based on (1.16) solves the first order condition (1.8). Note that the resulting estimator involves the inversion of the $T \times T$ matrix $\widetilde{Z}'\widetilde{Z}$, which is computationally demanding if T is large.

Fortunately, this estimator can be simplified, since the GLS transformation due to the serial correlation of the errors is irrelevant. The GLS transformation resulting from heteroskedastic errors yields $X_t^* = \Lambda^* F_t + u_t$, where $X_t^* = \Sigma^{-1/2} X_t$, $\Lambda^* = \Sigma^{-1/2}\Lambda$, and $u_t = \Sigma^{-1/2} e_t$. Replacing Λ^* by $\widehat{\Lambda}^* = \Sigma^{-1/2}\widehat{\Lambda}$, two-step estimation implies estimating $F_1, \ldots, F_T$ from the system

$$\begin{aligned} X_1^* &= \widehat{\Lambda}^* F_1 + u_1^* \\ \vdots \quad & \quad \vdots \\ X_T^* &= \widehat{\Lambda}^* F_T + u_T^*, \end{aligned}$$

where $u_t^* = u_t + (\Lambda^* - \widehat{\Lambda}^*)F_t$. Note that the vectors u_t^* and u_s^* are correlated, which suggests to estimate the system by using a GLS estimator. However, it is well known that the GLS estimator of a seemingly unrelated regressions (SUR) system is identical to (equation-wise) OLS estimation, if the regressor matrix is identical in all equations. Indeed, since in the present setup the regressor matrix is $\widehat{\Lambda}^*$ for all equations, it follows that single-equation OLS estimation is as efficient as estimating the whole system by using a GLS approach. Thus, the estimation procedure for F_t can be simplified by ignoring the serial correlation of the errors. This suggests to estimate F_t from the cross-section regression

$$\frac{1}{\omega_i} x_{it} = \left(\frac{1}{\omega_i} \widehat{\lambda}_i' \right) F_t + u_{it}^* \quad (i = 1, \ldots, N), \tag{1.17}$$

where $u^*_{it} = \omega_i^{-1}\left[e_{it} + (\lambda_i - \widehat{\lambda}_i)'F_t\right]$ and $\omega_i^2 = E(e_{it}^2)$, i.e., ignoring the GLS transformation with respect to autocorrelation. In what follows, we focus on this simplified version of the two-step estimation approach as its properties are equivalent to those of the full two-way GLS estimation procedure.

1.4 Asymptotic distribution of the two-step PC-GLS estimator

Our analysis is based on a similar set of assumptions as in Bai (2003), which is restated here for completeness.

Assumption 1: *There exists a positive constant $M < \infty$ such that for all N and T: (i) $E||F_t||^4 \leq M$ for all t and $T^{-1}\sum_{t=1}^T F_tF_t' \xrightarrow{p} \Psi_F$ (p.d). (ii) $||\lambda_i|| \leq \overline{\lambda} < \infty$ for all i and $N^{-1}\Lambda'\Lambda \to \Psi_\Lambda$ (p.d.). (iii) For the idiosyncratic components it is assumed that $E(e_{it}) = 0$, $E|e_{it}|^8 \leq M$, $0 < |\gamma_N(s,s)| \leq M$, $T^{-1}\sum_{s=1}^T\sum_{t=1}^T |\gamma_N(s,t)| \leq M$, where $\gamma_N(s,t) = E(N^{-1}\sum_{i=1}^N e_{is}e_{it})$. Furthermore, $N^{-1}\sum_{i=1}^N\sum_{j=1}^N \tau_{ij} \leq M$, $\sum_{i=1}^N \tau_{ij} \leq M$, where $\tau_{ij} = \sup_t\{|E(e_{it}e_{jt})|\}$,*

$$\frac{1}{NT}\sum_{i=1}^N\sum_{j=1}^N\sum_{t=1}^T\sum_{s=1}^T |E(e_{it}e_{js})| \leq M$$

$$E\left|\frac{1}{\sqrt{N}}\sum_{i=1}^N [e_{is}e_{it} - E(e_{is}e_{it})]\right|^4 \leq M.$$

(iv) $E(N^{-1}\sum_{i=1}^N ||T^{-1/2}\sum_{t=1}^T F_{t-k}e_{it}||^2) \leq M$ for all N, T and k.
(v) For all t, k, N and T:

$$E\left|\left|\frac{1}{\sqrt{NT}}\sum_{s=1}^T\sum_{i=1}^N F_{s-k}[e_{is}e_{it} - E(e_{is}e_{it})]\right|\right|^2 \leq M$$

$$E\left|\left|\frac{1}{\sqrt{NT}}\sum_{s=1}^T\sum_{i=1}^N F_{s-k}\lambda_i' e_{is}\right|\right|^2 \leq M$$

$$\frac{1}{\sqrt{N}}\sum_{i=1}^N \lambda_i e_{it} \xrightarrow{d} \mathcal{N}(0, V_{\lambda e}^{(t)}),$$

where $V_{\lambda e}^{(t)} = \lim_{N\to\infty} N^{-1}\sum_{i=1}^N\sum_{j=1}^N \lambda_i\lambda_j' E(e_{it}e_{jt})$ and for each i

$$\frac{1}{\sqrt{T}}\sum_{t=p_i+1}^{T-p_i} F_t e_{i,t+k} \xrightarrow{d} \mathcal{N}(0, V_{Fe}^{(i)}) \quad \textit{for } -p_i \leq k \leq p_i,$$

where $V_{Fe}^{(i)} = \lim_{T\to\infty} T^{-1} \sum_{s=1}^{T}\sum_{t=1}^{T} E(F_t F_s' e_{i,s-k} e_{i,t-k})$.

For a thorough discussion of these assumptions, see Bai and Ng (2002) and Bai (2003). It is well known (e.g., Bai and Ng 2002) that for the asymptotic analysis of the estimators, the factors have to be normalized such that in the limit the common factors obey the same normalization as the estimated factors. Following Bai and Ng (2002), this is achieved by normalizing the factors as

$$\begin{aligned} F\Lambda' &= (FH)(H^{-1}\Lambda') \\ &= F_*\Lambda_*', \end{aligned}$$

where

$$H = T\Lambda'\Lambda F'\widehat{F}(\widehat{F}'XX'\widehat{F})^{-1}.$$

It can be shown that $T^{-1}F_*{}'F_* \xrightarrow{p} I_r$ and, therefore, F_* has asymptotically the same normalization as $\widehat{F}$.

As we do not impose the assumptions of a strict factor model with stationary idiosyncratic errors, we define the following "pseudo-true" values of the autoregressive and variance parameters:

$$\begin{aligned} \omega_i^2 &= \lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} E(e_{it}^2) \\ [\rho_{1,i},\ldots,\rho_{p_i,i}]' &= \Gamma_{i,11}^{-1}\Gamma_{i,10}, \end{aligned}$$

where

$$\Gamma_i = \lim_{T\to\infty} E\left(\frac{1}{T}\sum_{t=p_i+1}^{T}\begin{bmatrix} e_{i,t-1} \\ \vdots \\ e_{i,t-p_i}\end{bmatrix}\begin{bmatrix} e_{it} & \cdots & e_{i,t-p_i}\end{bmatrix}\right) = \begin{bmatrix}\Gamma_{i,10} & \Gamma_{i,11}\end{bmatrix},$$

$\Gamma_{i,10}$ is a $p_i \times 1$ vector, and $\Gamma_{i,11}$ is a $p_i \times p_i$ matrix.

For the asymptotic analysis, we need to impose the following assumption.

Assumption 2: *(i) There exists a positive constant* $C < \infty$*, such that for all* i*:* $\frac{1}{C} < \omega_i^2 < C$*. (ii) The matrix* $\Gamma_{i,11}$ *is positive definite.*

In practice, the covariance parameters are usually unknown and must be replaced by consistent estimates. The feasible two-step PC-GLS estimators $\widetilde{\lambda}_{i,\widehat{\rho}}$ and $\widetilde{F}_{t,\widehat{\omega}}$ solve the first order conditions

$$\widetilde{g}_{\lambda_i}(\widetilde{\lambda}_{i,\widehat{\rho}}, \widehat{F}, \widehat{\rho}^{(i)}) = \sum_{t=p_i+1}^{T} [\widehat{\rho}_i(L)(x_{it} - \widetilde{\lambda}'_{i,\widehat{\rho}}\widehat{F}_t)][\widehat{\rho}_i(L)\widehat{F}_t] \;=\; 0 \tag{1.18}$$

$$\widetilde{g}_{F_t}(\widehat{\Lambda}, \widetilde{F}_{t,\widehat{\omega}}, \widehat{\Omega}) = \sum_{i=1}^{N} \frac{1}{\widehat{\omega}_i^2}(x_{it} - \widehat{\lambda}'_i\widetilde{F}_{t,\widehat{\omega}})\widehat{\lambda}_i \;=\; 0, \tag{1.19}$$

where

$$\widehat{\omega}_i^2 = \frac{1}{T}\sum_{t=1}^{T} \widehat{e}_{it}^2 \tag{1.20}$$

and $\widehat{e}_{it} = x_{it} - \widehat{\lambda}'_i\widehat{F}_t$. Furthermore, $\widehat{\rho}^{(i)} = [\widehat{\rho}_{1,i}, \ldots, \widehat{\rho}_{p_i,i}]'$ is the least-squares estimator from the regression

$$\widehat{e}_{it} = \widehat{\rho}_{1,i}\widehat{e}_{i,t-1} + \cdots + \widehat{\rho}_{p_i,i}\widehat{e}_{i,t-p_i} + \widehat{\varepsilon}_{it}. \tag{1.21}$$

To study the limiting distribution of the feasible two-step PC-GLS estimator, the following Lemma is used.

Lemma 1: *Let* $\widehat{\rho}^{(i)} = [\widehat{\rho}_{1,i}, \ldots, \widehat{\rho}_{p_i,i}]'$ *denote the least-squares estimates from (1.21) and* $\widehat{\omega}_i^2$ *is the estimator defined in (1.20). Under Assumption 1 we have as* $(N, T \to \infty)$

$$\widehat{\rho}^{(i)} = \rho^{(i)} + O_p(T^{-1/2}) + O_p(\delta_{NT}^{-2}) \quad \textit{and} \quad \widehat{\omega}_i^2 = \omega_i^2 + O_p(T^{-1/2}) + O_p(\delta_{NT}^{-2}),$$

where $\delta_{NT} = \min(\sqrt{N}, \sqrt{T})$*.*

The following theorem presents the limiting distribution of the feasible two-step PC-GLS estimator.

Theorem 1: (i) *Under Assumptions 1-2 and if* $(N, T \to \infty)$ *and* $\sqrt{T}/N \to 0$*, then for each* i*,*

$$\sqrt{T}(\widetilde{\lambda}_{i,\widehat{\rho}} - H^{-1}\lambda_i) \;\xrightarrow{d}\; \mathcal{N}(0, \widetilde{\Psi}_F^{(i)^{-1}}\widetilde{V}_{Fe}^{(i)}\widetilde{\Psi}_F^{(i)^{-1}}),$$

where

$$\begin{aligned}
\widetilde{V}_{Fe}^{(i)} &= \lim_{T\to\infty}\frac{1}{T}\sum_{s=p_i+1}^{T}\sum_{t=p_i+1}^{T} E[\rho_i(L)H'F_t\,\rho_i(L)F'_sH\,\varepsilon_{is}\varepsilon_{it}] \\
\varepsilon_{it} &= \rho_i(L)e_{it} \\
\widetilde{\Psi}_F^{(i)} &= \lim_{T\to\infty}\frac{1}{T}\sum_{t=p_i+1}^{T} E\left\{[\rho_i(L)H'F_t][\rho_i(L)H'F_t]'\right\}.
\end{aligned}$$

(ii) *If* $(N, T \to \infty)$ *and* $\sqrt{N}/T \to 0$, *then for each* t,

$$\sqrt{N}(\widetilde{F}_{t,\widehat{\omega}} - H'F_t) \ \xrightarrow{d}\ \mathcal{N}(0, \widetilde{\Psi}_\Lambda^{-1}\widetilde{V}_{\lambda e}^{(t)}\widetilde{\Psi}_\Lambda^{-1}),$$

where

$$\widetilde{V}_{\lambda e}^{(t)} = \lim_{N\to\infty} N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\omega_i^2\omega_j^2}H^{-1}\lambda_i\lambda_j'H'^{-1}E(e_{it}e_{jt})\,,$$

$\widetilde{\Psi}_\Lambda = \lim_{N\to\infty} N^{-1}H^{-1}\Lambda'\Omega^{-1}\Lambda H'^{-1}$, *and* $\Omega = diag(\omega_1^2, \ldots, \omega_N^2)$.

Remark 1: From (i) it follows that the asymptotic distribution remains the same if the estimate $\widehat{\rho}_i(L)\widehat{F}_t$ in (1.13) is replaced by $\rho_i(L)H'F_t$. This suggests that the estimation error in $\widehat{F}_t$ and $\widehat{\rho}_i(L)$ does not affect the asymptotic properties of the estimator $\widetilde{\lambda}_{i,\widehat{\rho}}$. A similar result holds for the regressor $\widehat{\omega}_i^{-1}\widehat{\lambda}_i$. In other words, the additional assumptions on the relative rates of N and T ensure that the estimates of the regressors in equations (1.13) and (1.17) can be treated as "super-consistent".

Remark 2: The assumptions on the relative rates of N and T may appear to be in conflict with each other. However, the two parts of Theorem 1 are fulfilled if $N = cT^\delta$ where $1/2 < \delta < 2$. Therefore, the limiting distribution should give a reliable guidance if both dimensions N and T are of comparable magnitude.

Remark 3: It is interesting to compare the result of Theorem 1 with the asymptotic distribution obtained by Choi (2008). In the latter paper, it is assumed that $E(e_te_t') = \Omega$ for all t, where $e_t = [e_{1t}, \ldots, e_{Nt}]'$, i.e., the idiosyncratic components are assumed to be stationary. In this case, the model can be transformed as $X^* = F^*\Lambda^{*'} + e^*$, where $X^* = X\Omega^{-1/2}$, $F^* = FJ$, $\Lambda^* = \Omega^{-1/2}\Lambda(J')^{-1}$, $e^* = e\Omega^{-1/2}$ and

$$J = T\Lambda'\Omega^{-1}\Lambda F'\widetilde{F}(\widetilde{F}'X\Omega^{-1}X'\widetilde{F})^{-1},$$

such that the covariance matrix of e^* is identical the identity matrix. Note that the matrix normalizing the factors in Choi (2008), J, is different from the one employed for the PC-OLS estimator (and for our PC-GLS estimator), H. Imposing the former normalization, the asymptotic covariance matrix of Choi's (2008) GLS estimator $\widetilde{F}$ reduces to a diagonal matrix.

Remark 4: If the errors are serially uncorrelated, our PC-GLS estimator of F_t is related to the estimator suggested by Forni, Hallin, Lippi, and Reichlin (2005). Analogous to Choi (2008), an important distinguishing feature is the different normalization of the factors. Forni, Hallin, Lippi, and Reichlin (2005) propose an estimator of Ω that is obtained from integrating the estimated spectral density matrix of the idiosyncratic errors. The factors are obtained from solving the generalized eigenvalue problem $|\nu\widetilde{\Omega} - T^{-1}X'X| = 0$, where $\widetilde{\Omega}$ denotes the estimated covariance matrix of the idiosyncratic errors. Note that the matrix of eigenvectors, $\widetilde{V}$, of the generalized eigenvalue problem obeys the normalization $\widetilde{V}'\widetilde{\Omega}\widetilde{V} = I$.

Remark 5: The two-step approach can also be employed to an unbalanced data set with different numbers of time periods for the variables. Stock and Watson (2002b) suggest an EM algorithm, where the missing values are replaced by an estimate of the common component. Let $\hat{x}_{it} = \widehat{\lambda}_i'\widehat{F}_t$ denote the estimated observation based on the balanced data set ignoring all time periods with missing observations. The updated estimates of the common factors and factor loadings are obtained by applying the PC-OLS estimator to the enhanced data set, where the missing values are replaced by the estimates $\hat{x}_{it}$. Employing the updated estimates of F_t and λ_i, results in improved estimates of the missing values that can in turn be used to yield new PC-OLS estimates of the common factors and factor loadings. This estimation procedure can be iterated until convergence. Similarly, the two-step estimation procedure can be initialized by using the reduced (balanced) data set to obtain the PC-OLS estimates $\widehat{F}_t$ and $\widehat{\lambda}_i$. In the second step, the vector of common factors is estimated from regression (1.17). As the T cross-section regressions may employ different numbers of observations, missing values do not raise any problems. Similarly, the N time series regressions (1.13) may be based on different sample sizes. As in the EM algorithm, this estimation procedure can be iterated until convergence.

Remark 6: The two-step estimators can be iterated using the resulting estimates $\widetilde{\lambda}_{i,\widehat{\rho}}$ and $\widetilde{F}_{t,\widehat{\omega}}$ instead of the estimates $\widehat{F}_t$ and $\widehat{\lambda}_i$ in regressions (1.13) and (1.17). Similarly, improved estimators of the covariance parameters can be obtained from the second step residuals, $\widetilde{e}_{it} = x_{it} - \widetilde{\lambda}_{i,\widehat{\rho}}'\widetilde{F}_{t,\widehat{\omega}}$. However, since the estimation error of the factors, factor loadings, and covariance parameters does not affect the limiting distribution of the two-step estimators, additional iteration steps do not improve their *asymptotic* properties. Nevertheless, further iterations may improve the performance in *small samples.*

1.5 Asymptotic efficiency

In this section, we study the relative asymptotic efficiency of the estimators. The following theorem shows that the PC-GLS estimator is generally more efficient than the PC-OLS estimator if the temporal and contemporaneous variance and covariance functions of the error, e_{it}, are correctly specified.

Theorem 2: *Under Assumptions 1–2 and if* $E(e_t e_t') = \Omega = diag\{\omega_1, \ldots, \omega_N\}$, $E(\varepsilon_i \varepsilon_i') = \sigma_i^2 I_{T-p_i}$, *and* F_t *is independent of* e_{it} *for all* i *and* t, *then the PC-GLS estimators of* λ_i *and* F_t *are asymptotically more efficient than the respective PC-OLS estimator in the sense that the difference of the respective asymptotic covariance matrices is positive semidefinite.*

Admittedly, this result is of limited practical use as we typically cannot assume that all variance and covariance functions are correctly specified. In Section 1.3 we have argued that the PC-GLS estimator can be seen as a pseudo ML estimator, which provides us with simple and generally more efficient estimators even if the variance and covariance functions are misspecified. Indeed, our Monte Carlo experiments (some of which are presented in Section 1.6) indicate that the PC-GLS estimator tends to outperform the PC-OLS estimator even if the covariance functions are misspecified. This finding may suggest that the PC-GLS estimator is always more efficient, as it takes into account at least some form of heteroskedasticity and autocorrelation, whereas the PC-OLS estimator simply ignores the possibility of individual specific variances and serial correlation. Unfortunately, this is not true as it is possible to construct special cases characterized by misspecification of the variance or covariance function, where the PC-OLS estimator is asymptotically more efficient than the PC-GLS estimator.[4]

[4]For example, the PC-GLS estimator of λ_i is less efficient than the PC-OLS estimator if we fit an AR(1) model to the idiosyncratic errors, which are in fact generated by the MA(2) model $e_{it} = \varepsilon_{it} + 0.7\varepsilon_{i,t-1} - 0.7\varepsilon_{i,t-2}$. Note, that in this case the fitted AR(1) model implies a positive second

One possibility to cope with this problem is to minimize the misspecification by carefully determining the autoregressive lag order employing, for instance, the Akaike or Schwarz criterion.

In what follows, we propose an alternative approach that combines the two aforementioned estimators yielding a “hybrid estimator” which is at least as efficient as each of the two estimators. To construct such an estimator, we combine the moment conditions of PC-OLS and PC-GLS such that the respective generalized method of moments (GMM) estimator based on two sets of moment conditions is more efficient than any estimator based on a subset of moment conditions. The respective moments for estimating the common factors are:

$$\text{PC-OLS:} \quad m_1(F_t) = \sum_{i=1}^{N} \widehat{\lambda}_i (x_{it} - \widehat{\lambda}_i' F_t) \tag{1.22}$$

$$\text{PC-GLS:} \quad m_2(F_t) = \sum_{i=1}^{N} \frac{1}{\widehat{\omega}_i^2} \widehat{\lambda}_i (x_{it} - \widehat{\lambda}_i' F_t). \tag{1.23}$$

Define $z_i = [\widehat{\lambda}_i', \widehat{\lambda}_i'/\widehat{\omega}_i^2]'$ as the vector of instruments and $Z = [z_1, \ldots, z_N]'$ is an $N \times (2r)$ matrix. The GMM estimator $\widehat{F}_t^{gmm}$ is given by (e.g., Hayashi 2000, chap. 3)

$$\widehat{F}_t^{gmm} = (\widehat{\Lambda}' Z W_t Z' \widehat{\Lambda})^{-1} \widehat{\Lambda}' Z W_t Z' X_t. \tag{1.24}$$

The optimal weight matrix W_t results as

$$W_t = \left[E\left(Z' e_t e_t' Z\right)\right]^{-1}.$$

If e_t is independent of Z and $E(e_t e_t') = \Omega$ for all $t = 1, \ldots, T$, we can invoke the law of iterated expectations yielding $W_t = W = [E(Z'\Omega Z)]^{-1}$. This suggests to estimate the weight matrix as $\widehat{W} = (Z'\widehat{\Omega}Z)^{-1}$, where $\widehat{\Omega} = T^{-1} \sum_{t=1}^{T} \widehat{e}_t \widehat{e}_t'$ and $\widehat{e}_t$ denotes the residual vector from the PC-OLS or PC-GLS estimator, respectively. Unfortunately, this yields a singular weight matrix since the residual vector is orthogonal to the respective columns of the instrument matrix Z. We therefore employ the estimator for the covariance matrix of the idiosyncratic components suggested by Forni, Hallin, Lippi, and Reichlin (2000, 2005). This estimator (denoted as $\widetilde{\Omega}$) is obtained from the dynamic principal component method by integrating the estimated spectral density

order autocorrelation, whereas the actual second order autocorrelation of the errors is negative.

matrix of the vector of idiosyncratic components. Estimating the weight matrix as $\widetilde{W} = (Z'\widetilde{\Omega}Z)^{-1}$ yields the GMM estimator

$$\widehat{F}^{gmm} = [\widehat{F}_1^{gmm}, \ldots, \widehat{F}_T^{gmm}]' = XZ\widetilde{W}Z'\widehat{\Lambda}(\widehat{\Lambda}'Z\widetilde{W}Z'\widehat{\Lambda})^{-1}. \qquad (1.25)$$

A similar approach can be employed to combine the moment conditions of the PC-OLS and PC-GLS estimators of the factor loadings. First, consider the moments of the PC-GLS estimator based on AR(1) errors:

$$\begin{aligned} m^*(\lambda_i) &= \sum_{t=2}^{T}(\widehat{F}_t - \widehat{\rho}_i\widehat{F}_{t-1})[x_{it} - \widehat{\rho}_i x_{i,t-1} - \lambda_i'(\widehat{F}_t - \widehat{\rho}_i\widehat{F}_{t-1})] \\ &= (1+\widehat{\rho}_i^2)\sum_{t=1}^{T}\widehat{F}_t(x_{it} - \lambda_i'\widehat{F}_t) - \widehat{\rho}_i\sum_{t=2}^{T-1}(\widehat{F}_{t+1} + \widehat{F}_{t-1})(x_{it} - \lambda_i'\widehat{F}_t) + O_p(1). \end{aligned}$$

Therefore, if T is large, these moments are equivalent to a linear combination of the moments

$$m_3(\lambda_i) = \sum_{t=1}^{T}\widehat{F}_t(x_{it} - \lambda_i'\widehat{F}_t) \qquad (1.26)$$

$$m_4(\lambda_i) = \sum_{t=2}^{T-1}(\widehat{F}_{t+1} + \widehat{F}_{t-1})(x_{it} - \lambda_i'\widehat{F}_t). \qquad (1.27)$$

Since $m_3(\lambda_i)$ is the moment of the PC-OLS estimator, the "hybrid estimator" is obtained by employing the vector of instruments $\xi_t = [\widehat{F}_t', \widehat{F}_{t+1}' + \widehat{F}_{t-1}']'$ for $t = 2, \ldots, T-1$, $\xi_1 = [\widehat{F}_1, 0]$, and $\xi_T = [\widehat{F}_T, 0]$. Furthermore, define the matrix $\Xi = [\xi_1, \ldots, \xi_T]'$. The GMM estimator for λ_i results as

$$\widehat{\lambda}_i^{gmm} = (\widehat{F}'\Xi W_i \Xi'\widehat{F})^{-1}\widehat{F}'\Xi W_i \Xi' X_i. \qquad (1.28)$$

To estimate the weight matrix W_i, we employ the usual heteroskedasticity and autocorrelation consistent covariance (HAC) estimator suggested by Hansen (1982). Using $\widehat{e}_{it} = x_{it} - \widehat{\lambda}_i'\widehat{F}_t$ and

$$\widehat{\Gamma}_i(k) = \frac{1}{T}\sum_{t=k+1}^{T}\widehat{e}_{it}\widehat{e}_{i,t-k}\xi_{it}\xi_{i,t-k}',$$

the weight matrix is estimated as

$$\widehat{W}_i = \widehat{\Gamma}_i(0) + \sum_{j=1}^{\ell}\tau_j(\widehat{\Gamma}_i(j) + \widehat{\Gamma}_i(j)'),$$

where $\tau_j = (\ell - j + 1)/(\ell + 1)$ is the weight function and ℓ denotes the truncation lag obeying $\ell/T \to 0$.

Since the hybrid estimators for λ_i and F_t employ r additional moments, the small sample properties may deteriorate if r is large relative to T. Thus, although the GMM estimators are asymptotically more efficient than the PC-OLS or PC-GLS estimators, the hybrid estimators may perform worse in small samples, in particular, when the weight matrices are poorly estimated. Some improvement of the small sample properties may be achieved by using the continuously updated GMM estimator proposed by Hansen, Heaton, and Yaron (1996).

1.6 Small sample properties

In order to investigate the small sample properties of the proposed estimators, we perform a Monte Carlo study. In particular, we calculate a measure of efficiency and compare the performance of five different approaches: the standard PC estimator, the two-step and iterated PC-GLS estimators as described above, the hybrid estimator introduced in the previous section, and the quasi maximum likelihood (QML) estimator of Doz, Giannone, and Reichlin (2006). The latter authors suggest maximum likelihood estimation of the approximate dynamic factor model via the Kalman filter employing the EM algorithm. In order to make the standard estimation approach of traditional factor analysis feasible in the large approximate dynamic factor environment, Doz, Giannone, and Reichlin (2006, 2007) abstract from possible cross-sectional correlation of the idiosyncratic component. However, their estimation procedure does take into account factor dynamics as well as heteroskedasticity of the idiosyncratic errors.[5]

Furthermore, we consider two sets of simulation experiments. First, we study a setup featuring a single factor, where the parameters governing the data-generating

[5] Even though their actual implementation of the estimator does not allow for serial correlation of the idiosyncratic components, Doz, Giannone, and Reichlin (2006) point out that, in principle, it is possible to take into account this feature in the estimation approach. However, the resulting estimator is computationally demanding as it implies N additional transition equations for the idiosyncratic components.

process only exhibit a relatively loose relation to those obtained from real-life data. This allows us to get a transparent overview of the relative merits of the respective estimator in various – perfectly controlled – environments concerning autocorrelation, heteroskedasticity, as well as cross-sectional correlation. In our second experiment, on the other hand, we generate data based on the widely used data set of Stock and Watson (2005), where we consider the case of multiple factors. This puts us in a position to study the performance of the various estimators when applied to "more realistic" data sets, i.e., data sets representative for the ones typically encountered in practice.

1.6.1 Simulation in a controlled environment

The data-generating process of our first Monte Carlo experiment is the following:

$$x_{it} = \lambda_i F_t + e_{it},$$

where

$$\begin{aligned}
F_t &= \gamma F_{t-1} + u_t \ , \quad u_t \overset{iid}{\sim} \mathcal{N}(0, (1-\gamma^2)) \\
e_{it} &= \rho_i e_{i,t-1} + \varepsilon_{it} \ , \quad \varepsilon_t \overset{iid}{\sim} \mathcal{N}(0, R\Gamma R) \ , \quad \varepsilon_t = [\varepsilon_{1t} \ \dots \ \varepsilon_{Nt}]' \\
\Gamma &= \Sigma\Omega\Sigma \\
R &= diag\left(\sqrt{1-\rho_1^2}, \dots, \sqrt{1-\rho_N^2}\right), \quad \Sigma = diag(\sigma_1, \dots, \sigma_N) \\
\rho_i &\overset{iid}{\sim} \mathcal{U}[a, b] \\
\lambda_i &\overset{iid}{\sim} \mathcal{U}[0, 1],
\end{aligned} \tag{1.29}$$

where $\mathcal{U}[a, b]$ denotes a random variable uniformly distributed on the interval $[a, b]$. As mentioned above, in these baseline simulations, we set the number of static and dynamic factors equal to one and, therefore, F_t is a scalar.

Four different scenarios are considered. In the first two, we abstract from cross-sectional correlation, i.e., $\Omega = I$, and concentrate on either autocorrelation or heteroskedasticity. In the autocorrelation case, we focus on the dynamic aspects and set $\gamma = 0.7$, $\rho_i \overset{iid}{\sim} \mathcal{U}[0.5, 0.9]$, as well as $\sigma_i^2 = 2$ for all i. In the case focusing on heteroskedasticity, we set $\gamma = 0$, $\rho_i = 0$ for all i, and $\sigma_i \overset{iid}{\sim} |\mathcal{N}(\sqrt{2}, 0.25)|$.

In the other two scenarios, we allow for non-zero cross-sectional correlation, so that Ω is not the identity matrix. In constructing Ω we follow Chang (2002), in order to ensure that only weak cross-sectional correlation is present. In particular, the covariance matrix is generated using the spectral decomposition with $\Omega = HVH'$, where H is a matrix consisting of orthonormal column vectors and V is a diagonal matrix. H is constructed as $H = M(M'M)^{-1/2}$, where the elements of M are drawn from $\mathcal{U}[0,1]$. To obtain V, a set of eigenvalues ν_i, $i = 2, \ldots, N-1$, is generated by drawing from $\mathcal{U}[w,1]$, where $w > 0$. Furthermore, to control the ratio of the minimum to the maximum eigenvalue via w, $\nu_1 = w$ and $\nu_N = 1$. In line with Chang (2002), we choose $w = 0.1$ in all simulations. By construction, in the scenarios featuring cross-sectional correlation, heteroskedasticity is always present, where we also set $\sigma_i \overset{iid}{\sim} |\mathcal{N}(\sqrt{2}, 0.25)|$. As a result, we distinguish only the cases when autocorrelation is present or not. In the former scenario we set $\gamma = 0.7$ and $\rho_i \overset{iid}{\sim} \mathcal{U}[0.5, 0.9]$, whereas in the latter $\gamma = 0$ and $\rho_i = 0$ for all i.

We generate 1000 replications for different sample sizes. In particular, we set $N = 50, 100, 200, 300$ and $T = 50, 100, 200$. In order to assess how precisely we can estimate the true factors or loadings, we compute a measures of efficiency. Concerning the common factors, it is simply $R^2(F, \widehat{F})$, i.e., the coefficient of determination of a regression of F (the true factor) on $\widehat{F}$ (the estimator under consideration) and a constant. Obviously, an analogous measure can also be defined for the factor loadings. Employing the coefficient of determination has the advantage that our measure of efficiency is invariant to the normalization of the factors (or loadings).[6]

1.6.1.1 Autocorrelation and heteroskedasticity

Table 1.1 reports the results for the case of autocorrelated errors when abstracting from cross-sectional correlation. Apparently, the PC and QML estimators perform poorly, in particular with respect to the factor loadings, where the R^2s are of comparable magnitude. The low accuracy can be explained by the fact that both estimators fail to take into account serial correlation of the idiosyncratic component. The QML procedure takes into account the dynamics of the common factors. However, as has

[6] Doz, Giannone, and Reichlin (2006) also employ the (trace) R^2 as their performance measure.

Table 1.1: Efficiency: one factor, autocorrelated errors

	loadings (λ_i)				factors (F_t)			
	PC	two-step	iterated	QML	PC	two-step	iterated	QML
T=50								
N=50	0.287	0.525	0.622	0.267	0.735	0.730	0.848	0.640
N=100	0.294	0.559	0.646	0.287	0.812	0.809	0.924	0.752
N=200	0.298	0.576	0.653	0.300	0.858	0.855	0.959	0.809
N=300	0.304	0.591	0.660	0.313	0.884	0.884	0.974	0.846
T=100								
N=50	0.487	0.756	0.774	0.438	0.837	0.833	0.871	0.761
N=100	0.511	0.781	0.793	0.492	0.908	0.906	0.935	0.875
N=200	0.525	0.792	0.801	0.519	0.945	0.943	0.967	0.928
N=300	0.525	0.794	0.802	0.522	0.957	0.956	0.978	0.945
T=200								
N=50	0.685	0.875	0.876	0.645	0.872	0.870	0.881	0.830
N=100	0.700	0.886	0.888	0.683	0.932	0.930	0.939	0.915
N=200	0.708	0.890	0.891	0.701	0.963	0.962	0.969	0.956
N=300	0.711	0.892	0.893	0.707	0.973	0.973	0.979	0.969

Notes: Entries are the R^2 of a regression of the true factors or loadings on the corresponding estimate and a constant. *PC* is the ordinary principal component estimator, *two-step* and *iterated* indicate the two-step PC-GLS and iterated PC-GLS estimators, respectively, introduced in Section 1.3, and *QML* is the quasi maximum likelihood estimator of Doz, Giannone, and Reichlin (2006). The following parameter values are used: $\gamma = 0.7$, $\rho_i \overset{iid}{\sim} \mathcal{U}[0.5, 0.9]$, $\sigma_i^2 = 2$ for all i.

been argued in Section 1.3, the dynamic properties of the factors are irrelevant for the asymptotic properties as $N \to \infty$. In contrast, for the factor loadings the R^2s for both the two-step and the iterated PC-GLS estimator are considerably larger than the ones for the PC and QML estimators. In particular, with larger T the two PC-GLS estimators become increasingly accurate and show a similar performance as expected from Theorem 1. This picture changes somewhat when examining the results for the factors. Using the two-step estimator basically leads to the same R^2s as using standard PC. In this respect, note that the two-step regression for the common factors is not affected by possible autocorrelation of the errors but exploits possible heteroskedasticity. Interestingly, iterating the PC-GLS estimator until convergence, on the other hand, leads to a substantial increase in accuracy.[7] This is due to the fact that the loadings are estimated more precisely by taking into account the autocorrelation of the errors. Thus, in the second step, the regressors have a smaller error and this improves the efficiency of the factor estimates. This increase in accuracy is mainly noticeable for small T. For larger sample sizes, iteration still leads to more precise estimates, but in absolute terms the improvement is marginal and all estimators show a similar performance characterized by high accuracy.

The results for heteroskedastic errors and without cross-sectional correlation are presented in Table 1.2. Considering the results for the factor loadings, standard PC and two-step PC-GLS estimation show a similar performance, where both estimators become increasingly accurate with sample size. These findings are not surprising, since the two-step PC-GLS estimator of the factor loadings has the same asymptotic properties as the ordinary PC estimator if the errors are serially uncorrelated. A slight efficiency improvement with respect to the loadings is attainable by employing the iterated PC-GLS estimator, in particular if N is small compared to T. Analogous to the case with autocorrelated errors, the efficiency gain is due to the fact that by estimating the factors more precisely via incorporating heteroskedasticity, in the second step, the regressors have a smaller error, thus improving the accuracy of the estimated factor loadings. However, in line with Theorem 1, for larger samples the two PC-GLS estimators perform similarly. There are two reasons for the limited gain in efficiency via iteration compared to standard PC and two-step PC-GLS estimation. First, as explained in Section 1.3, relevant for the efficient estimation of the factor loadings is to allow for autocorrelation of the errors. Not surprisingly, as there is no serial correlation in this scenario, the accuracy of PC (and two-step PC-GLS) is relatively high. Second, and in contrast to the autocorrelation case, the reduction in the error of the regressors is not that large as indicated by the rather small absolute gain in efficiency with respect to the common factors by taking into account heteroskedasticity. Overall, the results concerning the common factors imply that heteroskedasticity

[7]The number of iterations is limited to a maximum of 5. First, this reduces the computational burden and we find no further improvement if the number of iterations is increased.

Table 1.2: Efficiency: one factor, heteroskedastic errors

	loadings (λ_i)				factors (F_t)			
	PC	two-step	iterated	QML	PC	two-step	iterated	QML
T=50								
N=50	0.569	0.559	0.618	0.629	0.833	0.917	0.929	0.932
N=100	0.605	0.596	0.623	0.632	0.915	0.964	0.969	0.970
N=200	0.630	0.618	0.630	0.640	0.958	0.984	0.987	0.987
N=300	0.630	0.618	0.625	0.635	0.972	0.991	0.992	0.992
T=100								
N=50	0.714	0.710	0.770	0.774	0.849	0.929	0.935	0.939
N=100	0.756	0.751	0.774	0.778	0.924	0.968	0.970	0.972
N=200	0.772	0.767	0.777	0.781	0.962	0.986	0.988	0.988
N=300	0.777	0.772	0.778	0.782	0.975	0.992	0.993	0.993
T=200								
N=50	0.821	0.820	0.871	0.872	0.857	0.931	0.934	0.938
N=100	0.858	0.857	0.875	0.876	0.929	0.970	0.972	0.973
N=200	0.872	0.870	0.878	0.878	0.964	0.987	0.988	0.988
N=300	0.875	0.874	0.878	0.879	0.976	0.992	0.993	0.993

Notes: Entries are the R^2 of a regression of the true factors or loadings on the corresponding estimate and a constant. The following parameter values are used: $\gamma = 0$, $\rho_i = 0$ for all i, $\sigma_i \overset{iid}{\sim} |\mathcal{N}(\sqrt{2}, 0.25)|$. For further information see Table 1.1.

of the errors does not seem to be that severe of a problem, as the R^2s of the four estimators under consideration basically all indicate high accuracy. While taking into account heteroskedasticity of the errors does indeed lead to an increase in the R^2s compared to standard PC, the difference is really noticeable only in small samples. Finally, the QML estimates of the factors as well as the factor loadings show a strong performance, even slightly better than the iterated PC-GLS estimator. This is due to the fact, that in this scenario the approximating model coincides with the true model and the QML estimator is equivalent to the exact ML estimator.

1.6.1.2 Introducing cross-sectional correlation

Another typical feature of many data sets is non-zero cross-sectional correlation. Consequently, in the next simulations, we check whether the superior performance of the two PC-GLS estimators still holds under such a correlation structure. First, consider the case of autocorrelated idiosyncratic errors as presented in Table 1.3.[8] The general conclusions from the autocorrelation case presented above carry over to this scenario with added cross-correlation, even though with some modifications due to the presence of heteroskedasticity. With respect to the loadings, the gain in efficiency of using the two-step and iterated PC-GLS estimators compared to standard PC is considerable, where the two PC-GLS estimators show a similar performance in large samples. The improvement when iterating the PC-GLS estimator is even more noticeable. This is due to the presence of heteroskedasticity in addition to autocorrelation. Consequently, there is not only a direct beneficial effect in terms of taking into account dynamic aspects with respect to the factor loadings, but also an indirect effect in terms of a reduction of the error in the regressors, i.e., the common factors, by taking into account heteroskedasticity. Concerning the common factors, standard PC, two-step PC-GLS, and iterated PC-GLS all show a strong performance. Still, employing the two PC-GLS estimators leads to more efficient estimates than PC, where the gain is most noticeable in small samples. Due to the presence of heteroskedasticity, the performance of two-step PC-GLS relative to standard PC is more comparable to the heteroskedasticity case presented above. Again, the benefit from iteration is even more noticeable than in the autocorrelation scenario abstracting from cross-sectional correlation. Analogous to the factor loadings, this is the result of the presence of both autocorrelation and heteroskedasticity. As a result, not only the factors are estimated more precisely by taking into account heteroskedasticity, but there is also the indirect effect of reducing the error in the regressors by taking into account autocorrelation in the idiosyncratic component. This leads to R^2s very close to one.

[8] As mentioned above, heteroskedasticity is always present by construction in this set of simulations. Moreover, due to the computational burden, we do not present results for the QML estimator for these simulations. We checked, however, whether the overall findings of the previous scenarios carry over to the cross-correlation case by running a subset of the simulations including the QML estimator. Indeed, we do not find a substantial change in results.

Table 1.3: Efficiency: one factor, cross-sectional correlation, autocorrelated errors

	loadings (λ_i)					factors (F_t)				
	PC	two-step	iterated	GMM$_e$	GMM$_{\hat{e}}$	PC	two-step	iterated	GMM$_W$	GMM$_{\widehat{W}}$
T=50										
N=50	0.419	0.649	0.751	0.700	0.507	0.828	0.893	0.955	0.894	0.829
N=100	0.442	0.702	0.762	0.753	0.541	0.903	0.949	0.982	0.948	0.907
N=200	0.459	0.721	0.761	0.770	0.560	0.941	0.973	0.993	0.973	0.950
N=300	0.464	0.727	0.761	0.776	0.568	0.952	0.980	0.995	0.979	0.962
T=100										
N=50	0.611	0.816	0.862	0.835	0.740	0.890	0.944	0.960	0.947	0.889
N=100	0.644	0.851	0.869	0.867	0.781	0.944	0.976	0.983	0.977	0.944
N=200	0.654	0.860	0.870	0.876	0.793	0.970	0.989	0.993	0.990	0.971
N=300	0.659	0.863	0.870	0.878	0.797	0.977	0.993	0.996	0.993	0.979
T=200										
N=50	0.767	0.901	0.927	0.906	0.877	0.912	0.957	0.962	0.960	0.909
N=100	0.794	0.922	0.931	0.927	0.902	0.957	0.982	0.984	0.982	0.956
N=200	0.801	0.928	0.932	0.932	0.909	0.978	0.992	0.993	0.992	0.978
N=300	0.803	0.929	0.932	0.933	0.910	0.984	0.995	0.996	0.995	0.984

Notes: Entries are the R^2 of a regression of the true factors or loadings on the corresponding estimate and a constant. GMM_e is the hybrid estimator for the factor loadings as suggested in Section 1.5 using the true idiosyncratic errors to compute the weighting matrix, whereas $GMM_{\hat{e}}$ is the corresponding estimator using the *estimated* idiosyncratic components. GMM_W is the hybrid estimator for the common factors using the true covariance matrix of the idiosyncratic components to compute the optimal weighting matrix, whereas $GMM_{\widehat{W}}$ employs the estimator for the covariance matrix suggested by Forni, Hallin, Lippi, and Reichlin (2000, 2005). The following parameter values are used: $\nu_i \overset{iid}{\sim} \mathcal{U}[0.1, 1]$, $\gamma = 0.7$, $\rho_i \overset{iid}{\sim} \mathcal{U}[0.5, 0.9]$, $\sigma_i \overset{iid}{\sim} |\mathcal{N}(\sqrt{2}, 0.25)|$. For further information see Table 1.1.

Second, consider the scenario without autocorrelation presented in Table 1.4. The overall findings are very similar to the ones with heteroskedasticity but abstracting from cross-sectional correlation, so that the introduction of the latter does not really seem to affect the performance of the suggested estimators. With respect to the factor loadings, the precision of standard PC and two-step PC-GLS is again very similar due to the absence of autocorrelation of the idiosyncratic errors. Iterating the PC-GLS estimator leads to slight efficiency improvements for the reasons stated above, while for large samples the two PC-GLS estimators perform similarly, in line with Theorem 1. Moreover, the results concerning the common factors again indicate high accuracy for all estimators, i.e., the R^2s tend to be very close to one, where the increase in the R^2s of using the two PC-GLS estimators is really noticeable only in small samples.

1.6.1.3 The hybrid estimator

As a final investigation within our first Monte Carlo setup, we consider the hybrid estimator suggested in Section 1.5. As it is constructed to circumvent the problem that in some special cases, characterized by misspecification of the variance or covariance function, standard PC is asymptotically more efficient than the PC-GLS estimator, we compare our new estimator in particular to PC-OLS and the two-step PC-GLS estimator.[9] We consider the data-generating process with non-zero cross-sectional correlation and heteroskedasticity as presented in Tables 1.3 and 1.4, where setups with and without autocorrelation are distinguished.

Moreover, while asymptotically the GMM estimator is more efficient than both the PC-OLS and PC-GLS estimators, this does not need to hold in finite samples. In particular, it turns out that having a reliable estimate of the optimal weighting matrix greatly affects the performance of the hybrid estimator. Consequently, we present results for two versions of this estimator. First, a variant which actually estimates the optimal weighting matrix as suggested in Section 1.5. In the second variant, we employ improved estimates of this matrix. For the estimator of the common factors, it

[9] It is not possible to actually construct such a pathological case in our standard simulation setup. But to be consistent with our results discussed above, we stay within this setting. The simulation experiment nevertheless allows to draw comparisons between the different estimators and gives a comprehensive picture of the hybrid estimator's performance in finite samples.

Table 1.4: Efficiency: one factor, cross-sectional correlation, no autocorrelation in the errors

	loadings (λ_i)					factors (F_t)				
	PC	two-step	iterated	GMM$_e$	GMM$_{\hat{e}}$	PC	two-step	iterated	GMM$_W$	GMM$_{\widehat{W}}$
T=50										
N=50	0.726	0.719	0.749	0.772	0.712	0.910	0.956	0.960	0.958	0.909
N=100	0.748	0.740	0.754	0.795	0.734	0.956	0.982	0.984	0.983	0.956
N=200	0.757	0.748	0.753	0.802	0.740	0.979	0.992	0.993	0.992	0.978
N=300	0.759	0.750	0.753	0.804	0.742	0.986	0.995	0.996	0.995	0.985
T=100										
N=50	0.836	0.833	0.859	0.854	0.828	0.917	0.960	0.962	0.962	0.917
N=100	0.851	0.848	0.860	0.869	0.843	0.959	0.983	0.983	0.983	0.958
N=200	0.863	0.860	0.864	0.879	0.855	0.980	0.993	0.993	0.993	0.980
N=300	0.865	0.862	0.864	0.881	0.856	0.987	0.996	0.996	0.996	0.987
T=200										
N=50	0.903	0.902	0.924	0.910	0.900	0.919	0.961	0.962	0.963	0.919
N=100	0.922	0.921	0.928	0.928	0.919	0.961	0.984	0.985	0.984	0.961
N=200	0.926	0.926	0.929	0.931	0.923	0.981	0.993	0.993	0.993	0.981
N=300	0.928	0.927	0.929	0.933	0.925	0.987	0.996	0.996	0.996	0.987

Notes: Entries are the R^2 of a regression of the true factors or loadings on the corresponding estimate and a constant. GMM_e is the hybrid estimator for the factor loadings as suggested in Section 1.5 using the true idiosyncratic errors to compute the weighting matrix, whereas $GMM_{\hat{e}}$ is the corresponding estimator using the *estimated* idiosyncratic components. GMM_W is the hybrid estimator for the common factors using the true covariance matrix of the idiosyncratic components to compute the optimal weighting matrix, whereas $GMM_{\widehat{W}}$ employs the estimator for the covariance matrix suggested by Forni, Hallin, Lippi, and Reichlin (2000, 2005). The following parameter values are used: $\nu_i \overset{iid}{\sim} \mathcal{U}[0.1, 1]$, $\gamma = 0$, $\rho_i = 0$ for all i, $\sigma_i \overset{iid}{\sim} |\mathcal{N}(\sqrt{2}, 0.25)|$. For further information see Table 1.1.

is computed using the true covariance matrix of the idiosyncratic components resulting from our simulation setup. Concerning the estimator of the factor loadings, we replace the estimated idiosyncratic components by the true errors generated in each simulation run to obtain a better estimate of the corresponding weighting matrix. The second variant, while not feasible in practice, clearly illustrates the effect of using a (possibly poor) estimator of the optimal weighting matrix.

Consider first the case featuring autocorrelated errors as presented in Table 1.3. The hybrid estimator with an improved estimate of the optimal weighting matrix already delivers the results in finite samples, which we expect the standard version to attain asymptotically. It is at least as efficient as both the standard PC and two-step PC-GLS estimator for the common factors and factor loadings. In most cases, in particular for the loadings, this variant of the hybrid estimator yields considerably more efficient estimates than both reference estimators. The gain in efficiency, however, vanishes in larger samples. The hybrid estimator employing an estimated weight matrix, on the other hand, does not yield as strong a performance. With respect to the factor loadings, the corresponding R^2s lie in between those of the standard PC and two-step PC-GLS estimators. The loss in efficiency in finite samples compared to the (more efficient) two-step estimator, however, becomes less severe as N and T get larger. This does not result for the common factors, where the efficiency of the hybrid estimator with estimated weight matrix is similar to that of the less efficient PC-OLS estimator for all sample sizes considered.

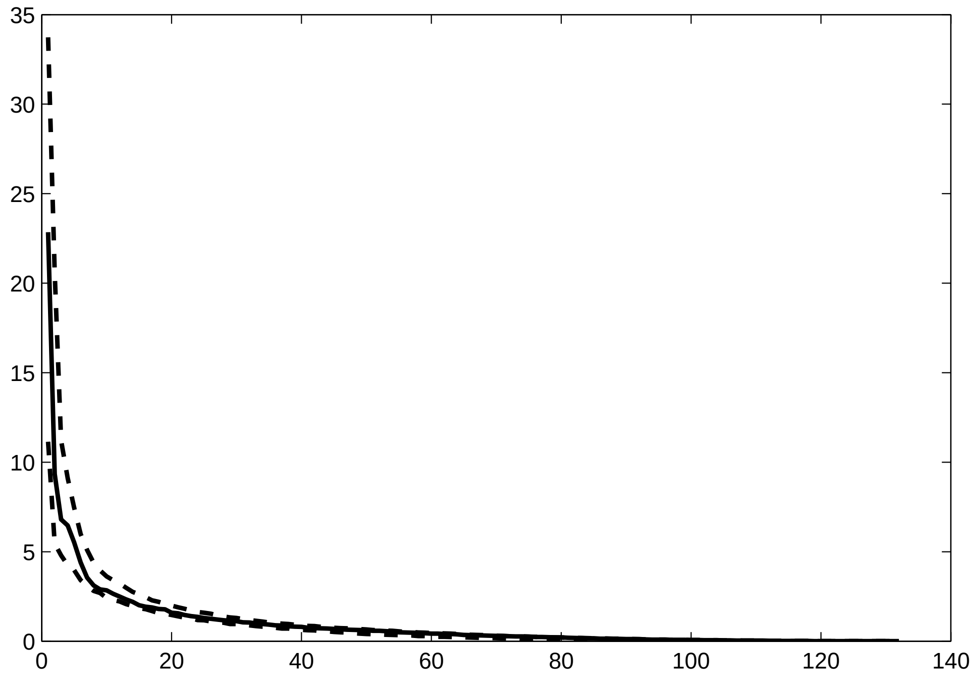
Analogous findings are obtained in the setup without autocorrelation in the idiosyncratic components (Table 1.4). The only difference is that the hybrid estimator for the factor loadings with estimated weight matrix performs about as well as the two reference estimators, where only a slight loss in efficiency can be observed in small samples. This is, of course, due to the absence of autocorrelation in the errors so that there is basically no difference in the performance of the standard PC and two-step PC-GLS estimators. The hybrid estimator with an improved estimate of the optimal weighting matrix is again at least as efficient as the two reference estimators, with a particularly strong performance for the factor loadings. The version with estimated weight matrix, moreover, estimates the common factors about as well as the less efficient standard PC estimator. In sum, these results clearly illustrate the important role played by the estimation of the optimal weighting matrix.

1.6.2 Simulation based on Stock and Watson's (2005) data set

In our second Monte Carlo experiment, we study the performance of the different estimators when applied to more realistic data sets. The starting point is the well known set of time series provided by Stock and Watson (2005). Since we do not want to impose a specific ARMA-structure on the simulated common factors or idiosyncratic components, which would constitute an advantage for the ARMA-based PC-GLS estimators, we employ the circular block bootstrap of Politis and Romano (1992). In particular, motivated by its superior performance in the first simulation experiment, we apply the iterated PC-GLS estimator to the aforementioned data set. In line with Stock and Watson's (2005) findings, we set the number of factors equal to seven. Thus, we obtain estimates for the common factors, F, the factor loadings, Λ, and thus for the common component, $F\Lambda'$, as well as for the idiosyncratic errors, e. In each simulation run, we resample overlapping blocks of a given length from the estimated factors and idiosyncratic errors to obtain a new set of factors and idiosyncratic components. Those series are then combined with the estimated factor loadings to obtain an individual sample for our simulation experiment. The block length differs between the factors and idiosyncratic errors, but is the same over all individual factors and errors, respectively. It is chosen optimally, based on Politis and White (2004) and including the correction of Patton, Politis, and White (2009).[10] In order to preserve the structure of the cross-sectional correlation, we choose the same permutation for all individual series, i.e., factors and errors, respectively, within one simulation run. The scree plots presented in Figure 1.3 suggest that our simulated series are in fact representative for the real-life data set of Stock and Watson (2005). This figure shows

[10]We use the MATLAB-implementation provided by Andrew Patton on his web page (http://econ.duke.edu/ ap172/code.html). As it only delivers the optimal block length of the individual series, we use the mean of the individual block lengths as our optimal value. As a robustness check, we also used larger and smaller block lengths than suggested by the procedure. This does not, however, change our results.

Figure 1.3: Scree plots



Notes: This figure shows the scree plot of Stock and Watson's (2005) original data set (solid line) as well as the respective smallest and largest eigenvalue over 1000 simulation runs from the simulated series (dashed lines). (N=132, T=526)

the scree plot of Stock and Watson's (2005) original data set as well as the respective smallest and largest eigenvalue over 1000 simulation runs from the simulated series. As can be seen from the graph, while there is of course some variation over the different simulations, the scree plots are close so that the basic dependence structure is preserved in the simulations.

Again, we generate 1000 replications for different sample sizes. Since this set of simulations is based on Stock and Watson's (2005) data set, we use their number of variables in all simulations, i.e., $N = 132$. We consider different time series dimensions, however, and set $T = 100, 200, 400, 526, 800$, where $T = 526$ is the number of time periods in Stock and Watson's (2005) data set. Varying the times series dimension is achieved by resampling the respective T observations *with replacement* from the estimated factors and idiosyncratic components. Such a bootstrap resampling scheme

Table 1.5: Efficiency using Stock and Watson's (2005) data set: circular block bootstrap

	loadings (λ_i)			factors (F_t)		
	PC	two-step	iterated	PC	two-step	iterated
N=132						
T=100	0.579	0.602	0.679	0.818	0.856	0.884
T=200	0.683	0.706	0.770	0.829	0.870	0.892
T=400	0.758	0.777	0.826	0.836	0.878	0.897
T=526	0.782	0.798	0.837	0.839	0.881	0.896
T=800	0.810	0.823	0.857	0.843	0.884	0.897

Notes: Entries are the trace R^2 of a regression of the true factors or loadings on the corresponding estimate and a constant. Simulation based on circular block bootstrap of factors and idiosyncratic component as estimated from Stock and Watson's (2005) data set, where $r = 7$. For further information see Table 1.1.

allows us, in particular, to obtain the necessary time series with a dimension that is larger than that of Stock and Watson's (2005) original data set. To assess the performance of the different estimators, the same measure of efficiency as introduced above is used. It has to be generalized, however, to make it applicable to the multi-factor case. In particular, now the *trace* R^2 of a regression of the true factors or loadings on the respective estimated factors or loadings and a constant is used.[11]

The results of this final Monte Carlo experiment are presented in Table 1.5. With respect to the factor loadings, an efficiency gain compared to standard PC is observed for both the two-step and iterated PC-GLS estimators. While the increase in accuracy of the two-step estimator is not that large, it is more pronounced for the iterated version. The latter again stems from the reduction in the error of the regressors, which is supported by the particular relation between N and T observed in this simulation. An increasing T relative to N positively affects the accurate estimation of the common factors and adversely affects the precision of the estimated loadings for

[11]We do not present results for the QML estimator in the second simulation experiment. A setup with seven factors increases the computational burden considerably, making it infeasible in practice to compute the required 1000 estimations in a reasonable amount of time.

the PC-OLS and two-step PC-GLS estimators. Consequently, this leaves more room for improvement with respect to the factor loadings and, furthermore, the regressors for the additional steps of the iterated estimator are estimated quite precisely, which additionally boosts the performance of that estimator.

Similar findings arise for the estimated common factors, where the efficiency of standard PC is larger than for the factor loadings, however. Nevertheless, employing two-step PC-GLS and, in particular, the iterated PC-GLS estimator leads to more precise estimates compared to PC-OLS. The gain in efficiency of using two-step PC-GLS compared to standard PC is already quite noticeable, whereas the subsequent increase of using iterated PC-GLS is not as large as for the factor loadings. This also stems from the particular relation between N and T present in this simulation. As noted above, the increasing T relative to N positively affects the two-step estimates of the common factors, so that there is less room for improvement via iterating the PC-GLS estimator. Furthermore, since the relation between N and T adversely affects the accuracy of the estimated factor loadings, the reduction in the error of the regressors in the subsequent iteration steps is not that large. Overall, this table clearly illustrates the advantages of employing the two-step PC-GLS estimator or its iterated version, when confronted with a real-life data set.

1.7 Conclusion

In this chapter, we propose a GLS-type estimation procedure that allows for heteroskedastic and autocorrelated errors. Since the estimation of the covariance parameters does not affect the limiting distribution of the estimators, the feasible two-step PC-GLS estimator is asymptotically as efficient as the infeasible GLS-estimator (assuming that the covariance parameters are known) and the iterated version that solves the first order conditions of the (approximate) ML estimator. Furthermore, we show that the PC-GLS estimator is generally more efficient than the PC-OLS estimator provided the variance and covariance functions are correctly specified. We also propose a GMM estimator that combines the moments of PC-OLS and PC-GLS to yield an estimator that is asymptotically at least as efficient as each of the two estimators if the second moments are misspecified. Notwithstanding these asymptotic results, our

Monte Carlo experiments suggest that in small samples the hybrid estimator suffers from the poor properties of the estimated weight matrix. We therefore recommend the (iterated) PC-GLS estimator for datasets of moderate sample size.

If one is willing to accept the framework of a strict factor model (that is a model with cross-sectionally uncorrelated factors and idiosyncratic errors), then our approach can also be employed for inference. For example, recent work by Breitung and Eickmeier (2011) shows that a Chow-type test for structural breaks can be derived using the iterated PC-GLS estimator. Other possible applications are LR tests for the number of common factors or tests of hypotheses on the factor space.

Appendix to Chapter 1

The following lemma plays a central role in the proofs of the following theorems:

Lemma A.1: *It holds for all $k \leq p_i$ that*

$$(i) \qquad T^{-1}\sum_{t=p_i+1}^{T}(\widehat{F}_t - H'F_t)F'_{t-k} = O_p(\delta_{NT}^{-2}), \quad T^{-1}\sum_{t=p_i+1}^{T}(\widehat{F}_t - H'F_t)\widehat{F}'_{t-k} = O_p(\delta_{NT}^{-2})$$

$$(ii) \qquad T^{-1}\sum_{t=p_i+1}^{T}\widehat{F}_t\widehat{F}'_{t-k} = T^{-1}\sum_{t=p_i+1}^{T}H'F_tF'_{t-k}H + O_p(\delta_{NT}^{-2})$$

$$(iii) \qquad T^{-1}\sum_{t=p_i+1}^{T}(\widehat{F}_t - H'F_t)e_{i,t-k} = O_p(\delta_{NT}^{-2})$$

$$(iv) \qquad N^{-1}\sum_{i=1}^{N}\frac{1}{\omega_i^2}(\widehat{\lambda}_i - H^{-1}\lambda_i)\lambda_i' = O_p(\delta_{NT}^{-2}), \quad N^{-1}\sum_{i=1}^{N}\frac{1}{\omega_i^2}(\widehat{\lambda}_i - H^{-1}\lambda_i)\widehat{\lambda}_i' = O_p(\delta_{NT}^{-2})$$

$$(v) \qquad N^{-1}\sum_{i=1}^{N}\frac{1}{\omega_i^2}(\widehat{\lambda}_i - H^{-1}\lambda_i)e_{it} = O_p(\delta_{NT}^{-2}) \ .$$

Proof: (i) The proof follows closely the proof for $k = 0$ provided by Bai (2003, Lemmas B.2 and B.3). We therefore present only the main steps.

We start from the representation

$$\widehat{F}_t - H'F_t = \frac{1}{NT}V_{NT}^{-1}\left(\widehat{F}'F\Lambda'e_t + \widehat{F}'e\Lambda F_t + \widehat{F}'ee_t\right),$$

where $e_t = [e_{1t}, \ldots, e_{Nt}]'$, $e = [e_1, \ldots, e_T]'$, and V_{NT} is an $r \times r$ diagonal matrix of the r largest eigenvalues of $(NT)^{-1}XX'$ (Bai 2003, Theorem 1). Consider

$$\begin{aligned}\frac{1}{T}\sum_{t=p_i+1}^{T}(\widehat{F}_t - H'F_t)F'_{t-k} &= \frac{1}{NT^2}V_{NT}^{-1}\Bigg(\widehat{F}'F\Lambda'\sum_{t=p_i+1}^{T}e_tF'_{t-k} + \widehat{F}'e\Lambda\sum_{t=p_i+1}^{T}F_tF'_{t-k} \\ &\qquad + \widehat{F}'e\sum_{t=p_i+1}^{T}e_tF'_{t-k}\Bigg) \\ &= I + II + III.\end{aligned}$$

From Assumption 1 (v) it follows that

$$\Lambda' \sum_{t=p_i+1}^{T} e_t F'_{t-k} = \sum_{i=1}^{N} \sum_{t=p_i+1}^{T} e_{it}\lambda_i F'_{t-k} = O_p(\sqrt{NT}),$$

and using Lemma B.2 of Bai (2003) it follows that $T^{-1}\widehat{F}'F = T^{-1}H'F'F + T^{-1}(\widehat{F} - FH)'F = T^{-1}H'F'F + O_p(\delta_{NT}^{-2})$. Thus, we obtain

$$I = V_{NT}^{-1}\left(T^{-1}\widehat{F}'F\right)\left(\frac{1}{\sqrt{NT}}\Lambda' \sum_{t=p_i+1}^{T} e_t F'_{t-k}\right)\frac{1}{\sqrt{NT}} = O_p\left(\frac{1}{\sqrt{NT}}\right).$$

Next, we consider

$$\Lambda' e' \widehat{F} = \Lambda' \sum_{t=1}^{T} e_t F'_t H + \Lambda' \sum_{t=1}^{T} e_t(\widehat{F}_t - H'F_t)'.$$

Following Bai (2003, p. 160), we have

$$\begin{aligned} \frac{1}{NT}\Lambda' \sum_{t=1}^{T} e_t F'_t H &= O_p\left(\frac{1}{\sqrt{NT}}\right) \\ \frac{1}{NT}\Lambda' \sum_{t=1}^{T} e_t(\widehat{F}_t - H'F_t)' &= O_p\left(\frac{1}{\delta_{NT}\sqrt{N}}\right). \end{aligned}$$

Using $T^{-1}\sum_{t=p_i+1}^{T} F'_t F_{t-k} = O_p(1)$, we obtain

$$II = V_{NT}^{-1}\left(\frac{1}{NT}\widehat{F}'e\Lambda\right)\left(\frac{1}{T}\sum_{t=p_i+1}^{T} F_t F'_{t-k}\right) = \left[O_p\left(\frac{1}{\sqrt{NT}}\right) + O_p\left(\frac{1}{\delta_{NT}\sqrt{N}}\right)\right] O_p(1).$$

For the remaining term, we obtain

$$\begin{aligned} & \frac{1}{NT^2}\widehat{F}'e \sum_{t=p_i+1}^{T} e_t F'_{t-k} = \frac{1}{NT^2}\sum_{s=1}^{T}\sum_{t=p_i+1}^{T} e'_s e_t \widehat{F}_s F'_{t-k} \\ = & \frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=p_i+1}^{T} \widehat{F}_s F'_{t-k}\zeta_{NT}(s,t) + \frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=p_i+1}^{T} \widehat{F}_s F'_{t-k}\gamma_N(s,t), \end{aligned}$$

where

$$\begin{aligned} \zeta_{NT}(s,t) &= e'_s e_t/N - \gamma_N(s,t) \\ \gamma_N(s,t) &= E(e'_s e_t/N). \end{aligned}$$

As in Bai (2003, p. 164f), we obtain

$$III = V_{NT}^{-1}\left[O_p\left(\frac{1}{\delta_{NT}\sqrt{T}}\right) + O_p\left(\frac{1}{\delta_{NT}\sqrt{N}}\right)\right].$$

Collecting these results, we obtain

$$I + II + III = O_p\left(\frac{1}{\sqrt{NT}}\right) + O_p\left(\frac{1}{\sqrt{T}\delta_{NT}}\right) + O_p\left(\frac{1}{\sqrt{N}\delta_{NT}}\right) = O_p\left(\frac{1}{\delta_{NT}^2}\right).$$

The proof of the second result in (i) is a similar modification of Lemma A.1 in Bai (2003) and is therefore omitted.

(ii) Consider

$$\begin{aligned}
T^{-1}\sum_{t=p_i+1}^{T}\widehat{F}_t\widehat{F}'_{t-k} &= T^{-1}\sum_{t=p_i+1}^{T}[H'F_t + (\widehat{F}_t - H'F_t)][H'F_{t-k} + (\widehat{F}_{t-k} - H'F_{t-k})]' \\
&= T^{-1}\left(\sum_{t=p_i+1}^{T} H'F_tF'_{t-k}H + (\widehat{F}_t - H'F_t)F'_{t-k}H + H'F_t(\widehat{F}'_{t-k} - F'_{t-k}H) + (\widehat{F}_t - H'F_t)(\widehat{F}'_{t-k} - F'_{t-k}H)\right) \\
&= T^{-1}\left(\sum_{t=p_i+1}^{T} H'F_tF'_{t-k}H + \underbrace{H'F_t(\widehat{F}'_{t-k} - F'_{t-k}H)}_{Ta} + \underbrace{(\widehat{F}_t - H'F_t)\widehat{F}'_{t-k}}_{Tb}\right) \\
&= \left(T^{-1}\sum_{t=p_i+1}^{T} H'F_tF'_{t-k}H\right) + a + b.
\end{aligned}$$

Using (i) the terms a and b can be shown to be $O_p(\delta_{NT}^{-2})$.

(iii) The proof for $k = 0$ is given in Bai (2003, Lemma B.1). It is not difficult to see that the result remains unchanged if $k \neq 0$.

(iv) Following Bai (2003, p. 165) we have

$$\widehat{\lambda}_i - H^{-1}\lambda_i = T^{-1}H'F'e_i + T^{-1}\widehat{F}'(F - \widehat{F}H^{-1})\lambda_i + T^{-1}(\widehat{F} - FH)'e_i, \tag{1.30}$$

where $e_i = [e_{i1}, \ldots, e_{iT}]'$. Post-multiplying by $\omega_i^{-2}\lambda_i'$ and averaging yields

$$\begin{aligned}
N^{-1}\sum_{i=1}^{N}\frac{1}{\omega_i^2}(\widehat{\lambda}_i - H^{-1}\lambda_i)\lambda_i' &= T^{-1}H'F'\left(N^{-1}\sum_{i=1}^{N}\frac{1}{\omega_i^2}e_i\lambda_i'\right) \\
&+ T^{-1}\widehat{F}'(F - \widehat{F}H^{-1})\left(N^{-1}\sum_{i=1}^{N}\frac{1}{\omega_i^2}\lambda_i\lambda_i'\right) + T^{-1}(\widehat{F} - FH)'\left(N^{-1}\sum_{i=1}^{N}\frac{1}{\omega_i^2}e_i\lambda_i'\right).
\end{aligned}$$

From Bai (2003, p. 165) it follows that the last two terms are $O_p(\delta_{NT}^{-2})$. From Assumption 1 (v) and Assumption 2 (i) it follows that

$$\left\|\sum_{t=1}^{T}\frac{1}{\omega_i^2}H'F_t\lambda_i'e_{it}\right\| \leq \frac{1}{\omega_{\min}^2}\left\|\sum_{t=1}^{T}H'F_t\lambda_i'e_{it}\right\| = O_p(1/\sqrt{T}),$$

where $\omega_{\min} = \min(\omega_1, \ldots, \omega_N)$. Thus, the first part of (iv) is $O_p(\delta_{NT}^{-2})$. The second equation can be shown by using the first part and Lemma A.1 (v).

(v) From (1.30) it follows that

$$\begin{aligned} N^{-1}\sum_{i=1}^{N}(\widehat{\lambda}_i - H^{-1}\lambda_i)e_{it} &= N^{-1}T^{-1}\sum_{s=1}^{T}\sum_{i=1}^{N}\widehat{F}_s e_{is}e_{it} + N^{-1}T^{-1}\sum_{s=1}^{T}\sum_{i=1}^{N}\widehat{F}_s(F_s - H'^{-1}\widehat{F}_s)'\lambda_i e_{it} \\ &= a + b. \end{aligned}$$

For expression a we write

$$N^{-1}T^{-1}\sum_{s=1}^{T}\widehat{F}_s\sum_{i=1}^{N}e_{is}e_{it} = T^{-1}\sum_{s=1}^{T}\widehat{F}_s\left[N^{-1}\sum_{i=1}^{N}e_{is}e_{it} - E(e_{is}e_{it})\right] + T^{-1}\sum_{s=1}^{T}\widehat{F}_s\gamma_N(s,t).$$

From Lemma A.2 (a) and (b) of Bai (2003), it follows that the first term on the r.h.s. is $O_p(N^{-1/2}\delta_{NT}^{-1})$, whereas the second term is $O_p(T^{-1/2}\delta_{NT}^{-1})$.

To analyze b we note that by Lemma A.1 (i) and Assumption 1 (v)

$$\left[T^{-1}\sum_{s=1}^{T}\widehat{F}_s(F_s - H'^{-1}\widehat{F}_s)'\right]\left[N^{-1}\sum_{i=1}^{N}\lambda_i e_{it}\right] = O_p(\delta_{NT}^{-2})O_p(N^{-1/2}).$$

Collecting these results, it follows that

$$\begin{aligned} \left\|N^{-1}\sum_{i=1}^{N}\frac{1}{\omega_i^2}(\widehat{\lambda}_i - H^{-1}\lambda_i)e_{it}\right\| &\le \frac{1}{\omega_{\min}^2}\left\|N^{-1}\sum_{i=1}^{N}(\widehat{\lambda}_i - H^{-1}\lambda_i)e_{it}\right\| \\ &= O_p(T^{-1/2}\delta_{NT}^{-1}) + O_p(N^{-1/2}\delta_{NT}^{-1}) + O_p(N^{-1/2}\delta_{NT}^{-2}) = O_p(\delta_{NT}^{-2}). \end{aligned}$$

Proof of Lemma 1:

Let

$$z_t = \begin{bmatrix} e_{it} \\ \vdots \\ e_{i,t-p_i+1} \end{bmatrix} \text{ and } \widehat{z}_t = \begin{bmatrix} x_{it} - \widehat{\lambda}_i'\widehat{F}_t \\ \vdots \\ x_{i,t-p_i+1} - \widehat{\lambda}_i'\widehat{F}_{t-p_i+1} \end{bmatrix}.$$

Using the same arguments as in Lemma 4 of Bai and Ng (2002), it can be shown that

$$T^{-1}\sum_{t=p_i+1}^{T}\widehat{e}_{it}\widehat{z}_{t-1} - T^{-1}\sum_{t=p_i+1}^{T}e_{it}z_{t-1} = O_p(\delta_{NT}^{-2})$$

and $T^{-1}\sum_{t=p_i+1}^{T}(\widehat{z}_{t-1}\widehat{z}_{t-1}' - z_{t-1}z_{t-1}') = O_p(\delta_{NT}^{-2})$. Therefore, we obtain for the least-squares estimator of $\rho^{(i)}$

$$
\begin{aligned}
\widehat{\rho}^{(i)} &= \rho^{(i)} + \left(\sum_{t=p_i+1}^{T} z_{t-1}z_{t-1}'\right)^{-1} \sum_{t=p_i+1}^{T} z_{t-1}\varepsilon_{it} + O_p(\delta_{NT}^{-2}) \\
&= \rho^{(i)} + O_p(T^{-1/2}) + O_p(\delta_{NT}^{-2})
\end{aligned}
$$

and, similarly, for the least-squares estimator of ω_i^2:

$$
\begin{aligned}
\widehat{\omega}_i^2 &= \omega_i^2 + \left(T^{-1}\sum_{t=p_i+1}^{T} e_{it}^2 - \omega_i^2\right) + \left(T^{-1}\sum_{t=p_i+1}^{T}\left(\widehat{e}_{it}^2 - e_{it}^2\right)\right) \\
&= \omega_i^2 + O_p(T^{-1/2}) + O_p(\delta_{NT}^{-2}).
\end{aligned}
$$

Proof of Theorem 1:

The feasible two-step estimator of λ_i is obtained as

$$
\begin{aligned}
\widetilde{\lambda}_{i,\widehat{\rho}} &= [\widehat{F}'R(\widehat{\rho}^{(i)})'R(\widehat{\rho}^{(i)})\widehat{F}]^{-1}\widehat{F}'R(\widehat{\rho}^{(i)})'R(\widehat{\rho}^{(i)})X_i \\
&= [\widehat{F}'R(\widehat{\rho}^{(i)})'R(\widehat{\rho}^{(i)})\widehat{F}]^{-1}\widehat{F}'R(\widehat{\rho}^{(i)})'R(\widehat{\rho}^{(i)})(F\lambda_i + e_i) \\
&= [\widehat{F}'R(\widehat{\rho}^{(i)})'R(\widehat{\rho}^{(i)})\widehat{F}]^{-1}\widehat{F}'R(\widehat{\rho}^{(i)})'R(\widehat{\rho}^{(i)})\{[\widehat{F} + (FH - \widehat{F})]H^{-1}\lambda_i + e_i)\} \\
\widetilde{\lambda}_{i,\widehat{\rho}} - H^{-1}\lambda_i &= [\widehat{F}'R(\widehat{\rho}^{(i)})'R(\widehat{\rho}^{(i)})\widehat{F}]^{-1}\widehat{F}'R(\widehat{\rho}^{(i)})'R(\widehat{\rho}^{(i)})[(FH - \widehat{F})H^{-1}\lambda_i + e_i],
\end{aligned}
$$

where $e_i = [e_{i1}, \ldots, e_{iT}]'$.

Using Lemma A.1 (ii) and Lemma 1, we obtain

$$
\begin{aligned}
& \frac{1}{T}\widehat{F}'R(\widehat{\rho}^{(i)})'R(\widehat{\rho}^{(i)})\widehat{F} \\
= \; & \frac{1}{T}\sum_{t=p_i+1}^{T}(\widehat{F}_t - \widehat{\rho}_{1,i}\widehat{F}_{t-1} - \ldots - \widehat{\rho}_{p_i,i}\widehat{F}_{t-p_i})(\widehat{F}_t - \widehat{\rho}_{1,i}\widehat{F}_{t-1} - \ldots - \widehat{\rho}_{p_i,i}\widehat{F}_{t-p_i})' \\
= \; & \frac{1}{T}\widehat{F}'R(\rho^{(i)})'R(\rho^{(i)})\widehat{F} + O_p(T^{-1/2}) + O_p(\delta_{NT}^{-2}) \quad \text{[by Lemma 1]} \\
= \; & \frac{1}{T}H'F'R(\rho^{(i)})'R(\rho^{(i)})FH + O_p(T^{-1/2}) + O_p(\delta_{NT}^{-2}) \quad \text{[by Lemma A.1 (ii)]}.
\end{aligned}
$$

Lemma A.1 (i) yields $T^{-1}\sum_{t=p_i+1}^{T}\widehat{F}_{t-k}(\widehat{F}_{t-k}' - F_{t-k}'H) = O_p(\delta_{NT}^{-2})$ and by using Lemma 1

$$
T^{-1}\widehat{F}'R(\widehat{\rho}^{(i)})'R(\widehat{\rho}^{(i)})(\widehat{F} - FH)H^{-1}\lambda_i = O_p(\delta_{NT}^{-2}) + O_p(\delta_{NT}^{-2}T^{-1/2}).
$$

Next, we consider

$$\begin{aligned}
& T^{-1/2}\sum_{t=p_i+1}^{T}[\widehat{\rho}_i(L)\widehat{F}_t]\,[\widehat{\rho}_i(L)e_{it}] \\
= \;& T^{-1/2}\sum_{t=p_i+1}^{T}\widehat{\rho}_i(L)[H'F_t+(\widehat{F}_t-H'F_t)]\widehat{\rho}_i(L)e_{it} \\
= \;& T^{-1/2}\sum_{t=p_i+1}^{T}\rho_i(L)H'F_t\,[\rho_i(L)e_{it}]+O_p(\sqrt{T}/\delta_{NT}^2)+O_p(T^{-1/2}),
\end{aligned}$$

where Lemma A.1 (iii) and Lemma 1 are invoked. Hence, we find

$$\begin{aligned}
\sqrt{T}(\widetilde{\lambda}_{i,\widehat{\rho}}-H^{-1}\lambda_i) \;=\;& [T^{-1}H'F'R(\rho^{(i)})'R(\rho^{(i)})FH]^{-1}T^{-1/2}H'F'R(\rho^{(i)})'R(\rho^{(i)})e_i \\
& +O_p(\sqrt{T}/\delta_{NT}^2)+O_p(T^{-1/2}),
\end{aligned}$$

where $\sqrt{T}/\delta_{NT}^2 \to 0$ if $\sqrt{T}/N \to 0$. Finally, Assumption 1 (v) implies

$$T^{-1/2}H'F'R(\rho^{(i)})'R(\rho^{(i)})e_i \;\xrightarrow{d}\; \mathcal{N}(0,\widetilde{V}_{Fe}^{(i)}),$$

where $\widetilde{V}_{Fe}^{(i)}$ is defined in Theorem 1. With these results, part (i) of the theorem follows.

The proof of part (ii) is similar. We therefore present the main steps only. The feasible two-step estimator of the common factors is given by

$$\begin{aligned}
\widetilde{F}_{t,\widehat{\omega}} &= (\widehat{\Lambda}'\widehat{\Omega}^{-1}\widehat{\Lambda})^{-1}\widehat{\Lambda}'\widehat{\Omega}^{-1}X_t \\
&= (\widehat{\Lambda}'\widehat{\Omega}^{-1}\widehat{\Lambda})^{-1}\widehat{\Lambda}'\widehat{\Omega}^{-1}[(\widehat{\Lambda}-\widehat{\Lambda}+\Lambda H'^{-1})H'F_t+e_t] \\
\widetilde{F}_{t,\widehat{\omega}}-H'F_t &= (\widehat{\Lambda}'\widehat{\Omega}^{-1}\widehat{\Lambda})^{-1}\widehat{\Lambda}'\widehat{\Omega}^{-1}[(\Lambda H'^{-1}-\widehat{\Lambda})H'F_t+e_t],
\end{aligned}$$

where $e_t=[e_{1t},\ldots,e_{Nt}]'$. Under Lemma 1, the (diagonal) elements of $(\widehat{\Omega}^{-1}-\Omega^{-1})$ are $O_p(T^{-1/2})+O_p(\delta_{NT}^{-2})$. Following Bai (2003) and using Lemma A.1 (iv) and (v), we obtain

$$\begin{aligned}
N^{-1}\widehat{\Lambda}'\widehat{\Omega}^{-1}\widehat{\Lambda} &= N^{-1}H^{-1}\Lambda'\Omega^{-1}\Lambda H^{'-1} + O_p(T^{-1/2}) + O_p(\delta_{NT}^{-2}) \xrightarrow{p} \widetilde{\Psi}_\Lambda \\
N^{-1}\widehat{\Lambda}'\widehat{\Omega}^{-1}(\widehat{\Lambda} - \Lambda H^{'-1}) &\leq N^{-1}\widehat{\Lambda}'\Omega^{-1}(\widehat{\Lambda} - \Lambda H^{'-1}) \\
&\quad + \left(\frac{1}{N}\sum_{i=1}^N |\widehat{\omega}_i^{-2} - \omega_i^{-2}|^2\right)^{1/2} \left(\frac{1}{N}||(\widehat{\Lambda}' - H^{-1}\Lambda')\widehat{\Lambda}||^2\right)^{1/2} \\
&= O_p(\delta_{NT}^{-2}) + O_p(T^{-1/2}/\delta_{NT}^2) \\
N^{-1}(\widehat{\Lambda} - \Lambda H'^{-1})'\widehat{\Omega}^{-1}e_t &= N^{-1}(\widehat{\Lambda} - \Lambda H'^{-1})'\Omega^{-1}e_t + N^{-1}\sum_{i=1}^N \left(\frac{1}{\widehat{\omega}_i^2} - \frac{1}{\omega_i^2}\right) e_{it}(\widehat{\lambda}_i - H^{-1}\lambda_i) \\
&= O_p(\delta_{NT}^{-2}) + O_p(T^{-1/2}/\delta_{NT}^2) \\
N^{-1/2}H^{-1}\Lambda'\widehat{\Omega}^{-1}e_t &= N^{-1/2}H^{-1}\Lambda'\Omega^{-1}e_t + O_p(T^{-1/2}) + O_p(\delta_{NT}^{-2}) \xrightarrow{d} \mathcal{N}(0, \widetilde{V}_{\lambda e}^{(t)}) \\
\widetilde{V}_{\lambda e}^{(t)} &= E\left(\lim_{N\to\infty} N^{-1}H^{-1}\Lambda'\Omega^{-1}e_t e_t'\Omega^{-1}\Lambda H^{'-1}\right) \\
&= \lim_{N\to\infty} N^{-1}\sum_{i=1}^N\sum_{j=1}^N \frac{1}{\omega_i^2\omega_j^2} H^{-1}\lambda_i\lambda_j' H^{'-1}E(e_{it}e_{jt}).
\end{aligned}$$

From these results the limit distribution stated in Theorem 1 (ii) follows.

Proof of Theorem 2:

First, we compare the asymptotic covariance matrices of the PC-OLS estimator $\widehat{F}_t$ and the PC-GLS estimator $\widetilde{F}_t$ (where for notational convenience the dependence on $\widehat{\omega}$ is suppressed). Using the results presented in Theorem 1 (ii), the asymptotic covariance matrix of $\widetilde{F}_t$ can be written as

$$\lim_{N,T\to\infty} N\ (\Lambda_0'\Omega^{-1}\Lambda_0)^{-1}\Lambda_0'\Omega^{-1}E(e_te_t')\Omega^{-1}\Lambda_0(\Lambda_0'\Omega^{-1}\Lambda_0)^{-1},$$

where $\Lambda_0 = \Lambda H'^{-1}$. If the covariance structure is correctly specified, then $E(e_te_t') = \Omega$ and the asymptotic covariance matrix reduces to

$$\left(\lim_{N,T\to\infty} \frac{1}{N}\Lambda_0'\Omega^{-1}\Lambda_0\right)^{-1}.$$

The asymptotic covariance matrix of the PC-OLS estimator is (Bai 2003)

$$\begin{aligned}
&\lim_{N,T\to\infty} N\ (\Lambda_0'\Lambda_0)^{-1}\Lambda_0'E(e_te_t')\Lambda_0(\Lambda_0'\Lambda_0)^{-1} \\
= &\lim_{N,T\to\infty} N\ (\Lambda_0'\Lambda_0)^{-1}\Lambda_0'\Omega\Lambda_0(\Lambda_0'\Lambda_0)^{-1},
\end{aligned}$$

if the covariance matrix Ω is correctly specified.

Let $\widehat{F}_t = \widetilde{F}_t + \delta_t$ where $\delta_t = \widehat{F}_t - \widetilde{F}_t$. From

$$var(\widehat{F}_t) = var(\widetilde{F}_t) + var(\delta_t) + cov(\widetilde{F}_t, \delta_t) + cov(\delta_t, \widetilde{F}_t)$$

it follows that $\widetilde{F}_t$ is asymptotically more efficient than $\widehat{F}_t$ if $N\, cov(\widetilde{F}_t, \delta_t) \to 0$ or

$$\lim_{N,T\to\infty} N\, E[(\widetilde{F}_t - H'F_t)(\widehat{F}_t - H'F_t)'] = \lim_{N,T\to\infty} N\, E[(\widetilde{F}_t - H'F_t)(\widetilde{F}_t - H'F_t)'].$$

Since

$$\begin{aligned} & \lim_{N,T\to\infty} N\, E[(\widetilde{F}_t - H'F_t)(\widehat{F}_t - H'F_t)'] \\ = & \lim_{N,T\to\infty} N\, (\Lambda_0'\Omega^{-1}\Lambda_0)^{-1}\Lambda_0'\Omega^{-1}E(e_te_t')\Lambda_0(\Lambda_0'\Lambda_0)^{-1} \\ = & \lim_{N,T\to\infty} \left(\frac{1}{N}\Lambda_0'\Omega^{-1}\Lambda_0\right)^{-1}, \end{aligned}$$

it follows that the difference of the asymptotic covariance matrices of the PC-GLS and PC-OLS estimators is positive semidefinite.

In a similar manner, it can be shown that the PC-GLS estimator of λ_i is asymptotically more efficient than the PC-OLS estimator. Let $F_0 = FH$ and $R(\rho^{(i)})$ as defined in Section 1.3. The asymptotic distribution of the PC-GLS estimator $\widetilde{\lambda}_i$ presented in Theorem 1 (i) can be written as

$$\begin{aligned} & \left[\lim_{N,T\to\infty} E\left(\frac{1}{T}F_0'R(\rho^{(i)})'R(\rho^{(i)})F_0\right)\right]^{-1} \\ \times & \lim_{N,T\to\infty} E\left(\frac{1}{T}F_0'R(\rho^{(i)})'R(\rho^{(i)})e_ie_i'R(\rho^{(i)})'R(\rho^{(i)})F_0\right) \\ \times & \left[\lim_{N,T\to\infty} E\left(\frac{1}{T}F_0'R(\rho^{(i)})'R(\rho^{(i)})F_0\right)\right]^{-1}. \end{aligned}$$

If the autoregressive model for the idiosyncratic errors is correctly specified, we have $E(\varepsilon_i\varepsilon_i') = \sigma_i^2 I_{T-p_i}$, where $\varepsilon_i = R(\rho^{(i)})e_i$. If F_0 is independent of ε_i, it follows from the law of iterated expectations that the asymptotic covariance matrix of the PC-GLS estimator $\widetilde{\lambda}_i$ reduces to

$$\sigma_i^2 \left[\lim_{N,T\to\infty} E\left(\frac{1}{T}F_0'R(\rho^{(i)})'R(\rho^{(i)})F_0\right)\right]^{-1}.$$

Consider

$$
\begin{aligned}
& \lim_{N,T\to\infty} T\, E[(\widetilde{\lambda}_i - H^{-1}\lambda_i)(\widehat{\lambda}_i - H^{-1}\lambda_i)'] \\
= \ & \left[\lim_{N,T\to\infty} E\left(\frac{1}{T}F_0' R(\rho^{(i)})' R(\rho^{(i)}) F_0\right)\right]^{-1} \\
& \times \lim_{N,T\to\infty} E\left(\frac{1}{T}F_0' R(\rho^{(i)})' R(\rho^{(i)}) e_i e_i' F_0\right) \left[\lim_{N,T\to\infty} E\left(\frac{1}{T}F_0' F_0\right)\right]^{-1} \\
= \ & \left[\lim_{N,T\to\infty} E\left(\frac{1}{T}F_0' R(\rho^{(i)})' R(\rho^{(i)}) F_0\right)\right]^{-1} \\
& \times \lim_{N,T\to\infty} E_{F_0}\left\{E\left[\frac{1}{T}F_0' R(\rho^{(i)})' R(\rho^{(i)}) e_i e_i' R(\rho^{(i)})' R(\rho^{(i)}) [R(\rho^{(i)})' R(\rho^{(i)})]^{-1} F_0 \middle| F_0\right]\right\} \\
& \times \left[\lim_{N,T\to\infty} E\left(\frac{1}{T}F_0' F_0\right)\right]^{-1} \\
= \ & \sigma_i^2 \left[\lim_{N,T\to\infty} E\left(\frac{1}{T}F_0' R(\rho^{(i)})' R(\rho^{(i)}) F_0\right)\right]^{-1}
\end{aligned}
$$

and, therefore, the asymptotic covariance between $\widetilde{\lambda}_i$ and $\widehat{\lambda}_i - \widetilde{\lambda}_i$ tends to zero. It follows that the PC-GLS estimator $\widetilde{\lambda}_i$ is asymptotically more efficient than the PC-OLS estimator $\widehat{\lambda}_i$.

Chapter 2

Optimal monetary policy under labor market frictions: the role of wage rigidity and markup shocks

2.1 Introduction

Considering the importance of unemployment fluctuations with respect to the business cycle, recent research has started to integrate labor market frictions into the workhorse model of monetary policy analysis, the New-Keynesian model.[1] The latter, in its standard specification featuring a Walrasian labor market, lacks equilibrium unemployment and related labor market dynamics. Thus, incorporating the aforementioned type of rigidities addresses a salient shortcoming of this framework with respect to the labor market dimension. A central feature of models with labor market frictions is the wage-determination mechanism. Employing rigid real wages has become standard in this respect, in order to address the so-called "unemployment volatility puzzle." The latter describes the difficulty of the standard Diamond-Mortensen-Pissarides style search and matching model employing Nash bargaining to determine wages, to generate empirically plausible fluctuations in unemployment and vacancies in response to

[1] Starting with Merz (1995) and Andolfatto (1996), who introduce rigid labor markets into a standard real-business-cycle model, recent contributions with respect to a New-Keynesian setting include Krause and Lubik (2007), Trigari (2009), Blanchard and Galí (2010), and Christoffel and Linzert (2010).

shocks.[2]

In this chapter, I take up the recent criticism and suggestions by Haefke, Sonntag, and van Rens (2008) and Pissarides (2009) with respect to such an approach. They argue that using a uniformly rigid real wage is not consistent with empirical evidence. As an alternative mechanism to address the unemployment volatility puzzle, the latter author suggests extending the models to include additional driving forces. Accordingly, I investigate optimal monetary policy in an environment with labor market frictions, heterogeneous wage setting, and markup shocks. My contribution is twofold. First, I investigate the implications of introducing heterogeneous wage setting which is consistent with the aforementioned authors' empirical findings for equilibrium allocations, and specifically, labor market dynamics and optimal monetary policy. Second, I examine the consequences of introducing additional driving forces in the form of markup shocks for the dynamic responses of inflation and unemployment to those shocks under different monetary policy regimes.

The two independent studies by Haefke, Sonntag, and van Rens (2008) and Pissarides (2009) challenge the empirical relevance of a uniformly rigid real wage, for example, in the spirit of Hall's (2005) "wage-norm" idea. They show by either performing their own empirical investigation or surveying empirical evidence on wage rigidity, that the wages which are rigid, are those of workers in ongoing job relationships, whereas wages for new hires are highly cyclical. Moreover, as these authors argue, the relevant wage series for search and matching models is wages for new hires. Consequently, since empirically the latter move one-for-one with labor productivity and Nash bargaining implies wages which are highly responsive to changes in productivity, it is consistent with the data to employ this standard mechanism to determine wages. Thus, the authors conclude that wage rigidity cannot be the answer to the unemployment volatility puzzle.

Accordingly, in the first part of this chapter, I introduce heterogeneous wage setting into a New-Keynesian dynamic stochastic general equilibrium (DSGE) model featuring labor market frictions in terms of hiring costs, following Blanchard and Galí (2010). Their model constitutes a particularly convenient benchmark and starting

[2]See, for instance, Shimer (2005).

point due to its transparency as well as sharp results concerning the efficient allocation. The latter makes it very easy to trace the effects of introducing features like heterogeneous wage setting and additional driving forces. At the same time, it gives an analytical relation between the dynamics of the economy and the underlying characteristics of the labor market. Blanchard and Galí (2010) also take the standard route of the literature of introducing an overall rigid wage, which constitutes the starting point of my investigation. In order to introduce heterogeneous wage setting into their setup, I distinguish between two kinds of workers: those in ongoing job relationships and newly hired workers. Opposed to Blanchard and Galí (2010), who introduce a rigid real wage for every worker, and consistent with the empirical studies mentioned above, I assume that only ongoing workers earn a rigid real wage in the spirit of Hall (2005). New hires, on the other hand, bargain over the wage for the current period, modeled by employing the generalized Nash solution. The main finding of this section is that with only these minor changes to the Blanchard and Galí (2010) setup, and despite an economy-wide *average* sticky wage, the inflation unemployment trade-off which they obtain in their model with an overall sticky wage disappears. This is because the expected wage sum and thus the expected labor costs for an individual worker over the course of her tenure at an individual firm moves one for one with labor productivity. This, in turn, eliminates potential hiring incentives, leading to unchanged employment and unemployment levels in response to technology shocks, which corresponds to the constrained efficient allocation. Consequently, introducing a form of wage rigidity which is consistent with empirical evidence leaves the monetary authority with a single target. It can solely focus on inflation with no concern for employment stabilization.

However, this still leaves open the question of what other mechanisms can account for the observed fluctuations in unemployment and what are the implications for monetary policy. In the second part of this chapter, I therefore examine the consequences of introducing an alternative approach to address the unemployment volatility puzzle. In particular, following the suggestion of Pissarides (2007),[3] I incorporate additional driving forces in the form of markup shocks into the New-Keynesian DSGE

[3] This is the more extensive working paper version of Pissarides (2009).

model with heterogeneous wage setting described above.[4] Following Steinsson (2003) and Rotemberg (2008), the elasticity of substitution in the Dixit-Stiglitz constant-elasticity-of-substitution (CES) consumption aggregator is assumed to be stochastic. As a result, the elasticity of demand and thus the desired markup are also stochastic. This can be interpreted as a constantly changing market power of firms due to changes in substitutability of the different varieties of goods. The consistency of markup fluctuations with empirical evidence can be seen from, for example, Rotemberg and Woodford (1991, 1999) and, more recently, Galí, Gertler, and López-Salido (2007). In my setup, shocks to the market power of firms and consequently movements in the desired markup feed via markup pricing into price dynamics and via resulting shifts in the labor demand schedule into employment and unemployment dynamics. Furthermore, a short-run inflation unemployment trade-off emerges, which I study by calibrating the system and simulating the movements of the endogenous variables in response to shocks under different monetary policy regimes. In this regard, I consider three different policies: first, completely stabilizing unemployment, which brings about the allocation of the latter variable in the constrained efficient allocation. Second, I investigate a policy of perfect inflation stabilization. In standard New-Keynesian models, which do not feature a trade-off, such a policy would be optimal. Optimality is used here in the sense of the utility-based approach to welfare analysis, as extensively described in, for example, Woodford (2003). Ultimately it means minimizing a loss function derived from the preferences of the private agents and the equilibrium conditions of the model. Finally, I consider optimal monetary policy in the preceding sense. The investigation is rounded off by deriving the efficient policy frontier, i.e., the plot of the standard deviations of unemployment and inflation under a policy of optimal commitment while varying the relative weight on

[4] Both Mortensen and Nagypál (2007) as well as Hall and Milgrom (2008) also point out by running simple regressions of labor market variables on productivity measures that one cannot expect productivity shocks to be the only source of fluctuations in unemployment, as implicitly assumed by Shimer (2005). Similarly, Balleer (2009) shows by employing a structural vector autoregression (SVAR) that the standard deviations of different labor market variables *conditional* on identified technology shocks are a lot smaller than the corresponding unconditional quantities. Consequently, all those authors emphasize the importance of alternative driving forces with respect to the labor market to match the observed (unconditional) moments.

unemployment stabilization in the monetary authority's loss function from zero to one. The main finding of this exercise in a setup with labor market frictions, heterogeneous wage setting, and markup shocks again supports the result of the literature concerning the importance of price stability.[5] Moreover, using markup shocks within the framework employed in this chapter, it is difficult to generate a significant amount of volatility in unemployment.

Concerning the related literature, not much work has been done on optimal monetary policy in an environment exhibiting labor market frictions. The main contributions in this area are Arseneau and Chugh (2008), Thomas (2008), Faia (2008, 2009), and Blanchard and Galí (2010). The two articles by Ester Faia study optimal monetary policy in an environment with monopolistic competition, price adjustment costs, and matching frictions in the labor market. Faia (2008), in addition, introduces wage rigidity in the spirit of Hall (2005). In her two articles, she considers optimal policy in the sense of both a globally optimal as well as constrained Ramsey approach and in terms of simple interest rate reaction functions. However, she neither looks at heterogeneous wage setting nor at alternative approaches to address the unemployment volatility puzzle. Thomas (2008) incorporates matching frictions into a New-Keynesian model and studies the effects of staggered nominal wage bargaining à la Gertler and Trigari (2009). The latter leads to a setup where inflation stabilization is not optimal. While he considers some type of wage heterogeneity, he does not distinguish between wages for ongoing workers and new hires, even though empirical evidence suggests the importance of such a differentiation.[6] Furthermore, he only looks at productivity shocks. Arseneau and Chugh (2008) consider a DSGE model with search and matching frictions in the labor market and costs of adjusting nominal wages to study optimal monetary and fiscal policy via the Ramsey approach. They do not focus, however, on tackling the unemployment volatility puzzle, and abstract from heterogeneous wage setting. Blanchard and Galí (2010), finally, start out from a simple New-Keynesian model with labor market frictions, where they introduce an ad-hoc rigid real wage in the sense of Hall (2005). This leads to a sizable

[5] See, for example, Woodford (2003).

[6] In particular, he assumes "that workers hired in between contracting periods receive the same wage as continuing workers" (p. 943).

short-run inflation unemployment trade-off. They only consider a uniformly rigid real wage, however, and do not take into account other possible approaches to generate data-consistent fluctuations in labor market variables.

The remainder of Chapter 2 is organized as follows. Section 2.2 presents a simple New-Keynesian model featuring labor market frictions following Blanchard and Galí (2010). I incorporate heterogeneous wage setting into this framework and study the implications for the resulting equilibrium allocation. In Section 2.3, I introduce markup shocks into the New-Keynesian model with heterogeneous wage setting derived in the preceding section. Different monetary policy regimes in this environment are studied in Section 2.4. In particular, I calibrate the model and simulate the dynamic responses of the endogenous variables to shocks to the elasticity of substitution under those monetary policy regimes and calculate the efficient policy frontier. Finally, Section 2.5 concludes.

2.2 A simple New-Keynesian model with labor market frictions

2.2.1 Economic environment

With respect to the basic setup, I follow Blanchard and Galí (2010), who present a simple New-Keynesian model where labor market frictions are introduced via hiring costs. The latter are increasing in labor market tightness, i.e., the ratio of hirings to unemployed. Even though this is a different formalism than in a standard search and matching model, it has similar implications for hiring decisions and unemployment dynamics.[7] In general, the economic environment is specified as follows:[8]

---

[7]Galí (2010), in fact, shows the equivalence of the hiring cost approach employed here and the more traditional matching function setup.

[8]Even though, in the following, I am primarily concerned with *monetary* policy analysis, I follow Woodford (2003) by considering a "cashless" economy and abstract from monetary frictions, which would give rise to a demand for money. Alternatively, the model presented here could be thought of as a "cashless limiting economy." This simplification of the analysis is mainly motivated by the fact that the consequences of incorporating monetary frictions are negligible from a quantitative point of view, as shown by the aforementioned author. Furthermore, since the short-term nominal interest rate is considered as the monetary policy instrument, which follows the current practice of major

- **Preferences:**

In order to avoid distributional issues potentially originating from the absence of perfect income insurance, I follow Merz (1995) and Andolfatto (1996) and assume income- and consumption-pooling of the households. Consequently, the representative household consists of a continuum of members normalized to measure 1 with preferences

$$U_0 = E_0 \left[\sum_{t=0}^{\infty} \beta^t \left(\log C_t - \chi \frac{N_t^{1+\phi}}{1+\phi} \right) \right], \tag{2.1}$$

where $\beta \in (0,1)$ is the household's discount factor, C_t is a CES aggregator over a continuum of goods (Dixit and Stiglitz 1977)

$$C_t = \left[\int_0^1 c_t(i)^{\frac{\varepsilon-1}{\varepsilon}} di \right]^{\frac{\varepsilon}{\varepsilon-1}}, \tag{2.2}$$

with elasticity of substitution $\varepsilon > 1$. $N_t \in [0,1]$ denotes the fraction of household members employed and ϕ indicates the inverse of the Frisch labor supply elasticity.

- **Technology:**

There is a continuum of firms indexed by i on the unit interval. Using an identical technology, which is a constant returns to scale (CRS) production function,

$$Y_t(i) = A_t N_t(i), \tag{2.3}$$

each firm produces a differentiated good. A_t indicates the common state of technology, which moves exogenously over time. Labor is the sole input to production and evolves according to

$$N_t(i) = (1-\delta) N_{t-1}(i) + H_t(i), \tag{2.4}$$

where separation is modeled as an exogenous process governed by the rate $\delta \in (0,1)$,[9] and the measure of workers hired by the individual firm in period t is indicated by

central banks like the European Central Bank or the Federal Reserve, the absence of an explicit money demand relation is inconsequential.

[9] As suggested by Hall (2004), the flow into unemployment is rather constant over time. Consequently, unemployment dynamics mainly result from changes in the exit rate out of unemployment and not from changes in the entrance rate into unemployment. Hence, I assume a constant separation rate.

$H_t(i)$.

- **Labor market:**

The labor market is characterized by the following timing and laws of motion. Since assumptions will be made to ensure full participation,[10] at the *beginning of period t* there is a measure of U_t unemployed workers, who are available for hire, consequently given by

$$U_t = 1 - (1-\delta)N_{t-1}. \tag{2.5}$$

This is just the difference between the labor force of measure 1 and the non-separated workers of the previous period, where aggregate employment at time t is denoted by $N_t \equiv \int_0^1 N_t(i)di$. Firms hire exclusively from this pool U_t, and the measure of newly hired workers, who start working in the same period, is given by $H_t \equiv \int_0^1 H_t(i)di$ and evolves according to

$$H_t = N_t - (1-\delta)N_{t-1}, \tag{2.6}$$

which is the difference between aggregate employment in the current period and last period's employment after separation.

Furthermore, *end-of-period* unemployment is denoted by u_t, and due to the full-participation assumption, it is given by

$$u_t = 1 - N_t, \tag{2.7}$$

the measure of potential workers left without a job after hiring decisions have been made in the given period.[11]

Following the labor search and matching literature, an index of labor market tightness is defined as

$$x_t \equiv \frac{H_t}{U_t} \in [0,1], \tag{2.8}$$

---

[10] This facilitates a better comparison to the constrained efficient allocation, which also features full participation. See Section 2.2.2.

[11] Later on, I use this unemployment measure when analyzing different monetary policies. Unemployment at the beginning of a period, U_t, is just given to completely specify the hiring process.

the ratio of aggregate hires to unemployment.[12] For the representative unemployed this ratio is just the job-finding rate for period t, i.e., the probability of being hired in the given period.

As a final characteristic of the labor market, the costs creating a friction in that market are defined. In a standard labor search and matching model these frictions are introduced by the cost per period of posting a vacancy multiplied by the expected time to fill it. Consequently, in such a model these costs are stochastic, since the time needed to fill the vacancy is uncertain, depending on the degree of labor market tightness. In the model presented here, the corresponding costs, called hiring costs, are deterministic. It is assumed that vacancies are filled instantly by paying these costs, which are a function of labor market tightness, as well. Hence, even though a different formalism is used to create hiring-cost frictions, the implications for hiring decisions and unemployment dynamics are similar. In particular, firm's hiring costs in terms of the CES aggregator of goods are given by $G_t H_t(i)$. This is just the product of the cost per hire, G_t, and the firm's hires in period t. Since G_t is taken as given by the firm and is independent of $H_t(i)$, but is an increasing function of labor market tightness, an externality arises. Accordingly, it is assumed

$$G_t = A_t B x_t^{\alpha}, \tag{2.9}$$

where $\alpha \geq 0, B \geq 0$, and $\delta B < 1$.

2.2.2 Allocating resources

So far, this setup is basically Blanchard and Galí's (2010) economic environment. Thus, in order to facilitate a transparent comparison of their findings to the ones of this chapter, I just repeat their main results with respect to the different allocations. In particular, this makes obvious the consequences of introducing heterogeneous wage setting and markup shocks.

First, consider the constrained efficient allocation.[13] In this case a social planner

[12] In the standard labor search and matching model the corresponding index is defined as the ratio of vacancies to unemployed, $\theta = \frac{v}{u}$.

[13] I present this allocation in a little bit more detail, since it also constitutes the constrained efficient

maximizes the utility of the representative household subject to the technological constraints and labor market frictions, but internalizes the externality indicated above. Because labor market participation is unambiguously beneficial, the allocation exhibits full participation and due to the symmetric setup production and consumption will be the same for all i. Furthermore, the optimality condition results as

$$\underbrace{\frac{\chi C_t N_t^{\phi}}{A_t}}_{\frac{MRS}{A_t}} \leq \underbrace{1-(1+\alpha)Bx_t^{\alpha} +\beta(1-\delta)E_t\left[\frac{C_t}{C_{t+1}}\frac{A_{t+1}}{A_t}B(x_{t+1}^{\alpha}+\alpha x_{t+1}^{\alpha}(1-x_{t+1}))\right]}_{\frac{MRT}{A_t}}, \qquad (2.10)$$

which holds with strict equality in the case of $N_t < 1$. This is just the standard optimality result that the marginal rate of substitution between labor and consumption should equal the marginal rate of transformation, here both simply divided by A_t. For both with and without hiring costs, i.e., $B > 0$ and $B = 0$, the preceding optimality condition implies a constant level of employment, $N_t = N^* \; \forall t$, which does not move in response to technology shocks.[14] Plugging N^* into the production function delivers the efficient level of output, $Y_t^* = A_t N^*$, and inserting it into the aggregate resource constraint results in the efficient level of consumption, $C_t^* = A_t N^*(1-\delta B x^{*\alpha})$. Consequently, both of these quantities move one for one with productivity, A_t.[15]

As a second way to allocate resources, consider the decentralized economy with monopolistic competition and flexible prices. In this environment, Blanchard and Galí (2010) look at two different wage setting mechanisms: the generalized Nash solution as the standard bargaining setup of the search and matching model and an ad-hoc rigid real wage in the spirit of Hall's (2005) "wage-norm" idea. With respect to Nash

allocation of the model with markup shocks presented in Section 2.3.

[14]This invariance result is not least due to the specification of preferences and the absence of capital accumulation. However, such a setup makes transparent the effects of incorporating labor market frictions, heterogeneous wage setting, and markup shocks.

[15]For more on the constrained efficient allocation, see Blanchard and Galí (2010).

bargained wages, the equilibrium condition results as

$$\frac{\chi C_t N_t^{\phi}}{A_t} = \frac{1}{\mathcal{M}} - (1+\vartheta)Bx_t^{\alpha} + \beta(1-\delta)E_t\left[\frac{C_t}{C_{t+1}}\frac{A_{t+1}}{A_t}B(x_{t+1}^{\alpha} + \vartheta x_{t+1}^{\alpha}(1-x_{t+1}))\right], \tag{2.11}$$

where $\mathcal{M} \equiv \frac{\varepsilon}{\varepsilon-1}$ denotes the optimal markup of the value maximizing, monopolistically competitive firm and ϑ is the worker's relative bargaining weight in the Nash bargain. As in the constrained efficient allocation, this equilibrium with Nash bargaining features a constant employment level, i.e., a level unaffected by technology shocks. Consequently, when employing the generalized Nash solution, there is no inflation unemployment trade-off and the monetary authority can exclusively focus on inflation with no concern whatsoever for employment stabilization. Intuitively, the equilibrium Nash bargained wage moves one for one with productivity, thus eliminating all potential hiring incentives, leading to an unchanged employment level in response to technology shocks.

In a next step, Blanchard and Galí (2010) follow Hall (2005) when introducing an ad-hoc rigid real wage of the form $W_t = \Theta A_t^{1-\gamma}, \gamma \in [0,1], \Theta > 0$, which breaks the one-for-one relation between productivity and the wage if $\gamma > 0$. With this kind of wage setting mechanism, productivity shocks affect firms' hiring incentives, leading in turn to fluctuations in (un)employment. Moreover, these fluctuations are inefficient since the constrained efficient employment level is constant, implying that the monetary authority is faced with a short-run inflation unemployment trade-off. Provided the monetary authority assigns at least some weight to unemployment stabilization, the optimal monetary policy problem becomes nontrivial.[16]

2.2.3 Heterogeneous wage setting

Even though a standard approach in the recent literature, introducing wage rigidity for each and every worker is not supported by the facts, as pointed out by Haefke, Sonntag, and van Rens (2008) and Pissarides (2009). When disaggregating the rigid aggregate wage series, the authors find that this rigidity is mostly due to noncyclical wages for

[16]Of course, to be able to continue the analysis, Blanchard and Galí (2010) introduce sticky prices such that monetary policy is effective in this setting.

workers in ongoing job relationships, whereas wages for new hires are highly cyclical. The latter, in turn, is consistent with employing Nash bargaining to determine wages. Accordingly, the authors conclude that wage rigidity cannot be the answer to the unemployment volatility puzzle, which typically serves as a rationale for introducing those rigidities.

Taking these findings as a starting point, in this section, I investigate the implications of introducing a wage setting mechanism into the presented model which is consistent with empirical evidence. In particular, heterogeneous wage setting is employed when considering the decentralized economy. In this setup, I distinguish between two kinds of workers: those who are in ongoing job relationships and those who were hired in the current period, i.e., newly hired workers. Contrary to Blanchard and Galí (2010) who use a rigid wage for every worker, in my setup only workers in ongoing job relationships earn this rigid wage. New hires bargain over the wage for the current period, modeled here by employing the generalized Nash solution. It should be noted that this is just a slight change to the framework of Blanchard and Galí (2010), since I use the same type of rigid wage as in their model. It applies, however, only to workers in ongoing jobs. Nevertheless, since the overwhelming part of the labor force is in ongoing jobs, it is just a very small fraction of workers who are affected by the changes in the wage determination mechanism.[17] Moreover, the economy-wide *average* wage is still rigid.

First, consider the representative firm's problem given the wage, $W_t^a(i)$, where a indicates that this is the average real wage the representative firm pays. It is just the wage sum divided by the number of workers employed by the individual firm. Thus,

$$\begin{aligned} W_t^a(i) &= \frac{(1-\delta)N_{t-1}(i)W_t^o + H_t(i)W_t^n}{N_t(i)} && (2.12) \\ &= \frac{(1-\delta)N_{t-1}(i)}{N_t(i)}\Theta A_t^{1-\gamma} + \frac{H_t(i)}{N_t(i)}W_t^n \,, && (2.13) \end{aligned}$$

where $W_t^o = \Theta A_t^{1-\gamma}$ is the rigid wage for ongoing jobs specified as in Blanchard and Galí (2010), and W_t^n is the wage for new hires, to be determined in a separate step by Nash bargaining. Furthermore, the number of workers in ongoing job relationships is $(1-\delta)N_{t-1}(i)$, i.e., last period's non-separated worker, and new hires are given by

[17] For empirical evidence on the US, see Haefke, Sonntag, and van Rens (2008), for instance.

$H_t(i)$. Since all firms are identical, it is possible to write $W_t^a(i) = W_t^a \; \forall i$.

In this setup with flexible prices and monopolistic competition, the representative firm maximizes its value

$$\max_{\{P_{t+k}(i)\}_{k=0}^{\infty}} E_t \left[\sum_{k=0}^{\infty} Q_{t,t+k} (P_{t+k}(i) Y_{t+k}(i) - P_{t+k} W_{t+k}^a N_{t+k}(i) - P_{t+k} G_{t+k} H_{t+k}(i)) \right], \tag{2.14}$$

by setting the price of its differentiated good, $P_t(i)$, optimally each period, subject to the production function and demand given by

$$Y_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\varepsilon} (C_t + G_t H_t) \quad \forall t. \tag{2.15}$$

In addition, the time path for the aggregate price index $P_t = \left[\int_0^1 P_t(i)^{1-\varepsilon} di \right]^{\frac{1}{1-\varepsilon}}$, the average real wage W_t^a, cost per hire G_t, and the stochastic discount factor for nominal payoffs $Q_{t,t+k} \equiv \beta^k \frac{C_t}{C_{t+k}} \frac{P_t}{P_{t+k}}$ are taken as given.

Solving this problem leads to the usual optimal price setting rule in such an environment, i.e., relative prices are set as a markup over real marginal cost

$$\frac{P_t(i)}{P_t} = \mathcal{M} \, MC_t \quad \forall t, \tag{2.16}$$

where the optimal markup is given by $\mathcal{M} \equiv \frac{\varepsilon}{\varepsilon - 1}$, and real marginal cost are obtained as

$$MC_t = \frac{W_t^n}{A_t} + B x_t^{\alpha} - \beta(1-\delta) E_t \left[\frac{C_t}{C_{t+1}} \frac{A_{t+1}}{A_t} \left(\frac{W_{t+1}^n}{A_{t+1}} - \Theta A_{t+1}^{-\gamma} + B x_{t+1}^{\alpha} \right) \right]. \tag{2.17}$$

This is just the respective costs less expected savings of hiring a worker now instead of next period. The former consist of this period's (Nash) wage and hiring costs, each normalized by productivity. The latter depend on next period's expected hiring costs and the expected difference between the wage for a newly hired worker and the ongoing wage, again normalized by productivity.

Furthermore, symmetry of the equilibrium implies $P_t(i) = P_t \; \forall i$, and thus due to equation (2.16)

$$MC_t = \frac{1}{\mathcal{M}} \quad \forall t. \tag{2.18}$$

Finally, plugging this equilibrium condition for real marginal cost into equation (2.17) leads to

$$\frac{W_t^n}{A_t} = \frac{1}{\mathcal{M}} - Bx_t^\alpha + \beta(1-\delta)E_t\left[\frac{C_t}{C_{t+1}}\frac{A_{t+1}}{A_t}\left(\frac{W_{t+1}^n}{A_{t+1}} - \Theta A_{t+1}^{-\gamma} + Bx_{t+1}^\alpha\right)\right]. \tag{2.19}$$

These conditions are derived under the assumption that wages, and in particular the wage for new hires, are taken as given. In order to specify the equilibrium, I have to assume a wage determination scheme for the newly hired workers, leading to an expression for the process of W_t^n, which can be combined with the preceding equation.

In accordance with the high cyclicality of wages for new hires, I use the generalized Nash solution. To derive the wage schedule, first consider the household side. The (real) value of a newly hired worker to the household at time t is given by

$$\mathcal{V}_t^N = W_t^n - C_t\chi N_t^\phi + \beta E_t\left[\frac{C_t}{C_{t+1}}\left(\delta(1-x_{t+1})\mathcal{V}_{t+1}^U + (1-\delta)\mathcal{V}_{t+1}^O + \delta x_{t+1}\mathcal{V}_{t+1}^N\right)\right]. \tag{2.20}$$

This is just the (Nash) wage minus the marginal rate of substitution plus the discounted expected continuation values. With respect to the latter, conditional on being employed in period t, $\delta(1-x_{t+1})$ is the probability of being separated and not rehired in the next period, thus becoming unemployed. With probability $(1-\delta)$ a worker is not separated, i.e., she is in an ongoing job in the next period, and δx_{t+1} is the probability of being separated but hired again in $t+1$, i.e., being a newly hired worker. Similarly, the value of a worker in an ongoing job to the household at time t results as

$$\mathcal{V}_t^O = \Theta A_t^{1-\gamma} - C_t\chi N_t^\phi + \beta E_t\left[\frac{C_t}{C_{t+1}}\left(\delta(1-x_{t+1})\mathcal{V}_{t+1}^U + (1-\delta)\mathcal{V}_{t+1}^O + \delta x_{t+1}\mathcal{V}_{t+1}^N\right)\right]. \tag{2.21}$$

The preceding expression has the same structure as the one for the value of a newly hired worker except that the Nash wage is replaced by the rigid wage for workers in ongoing jobs. Finally, the value of an unemployed member to the household at time

t is given by

$$\mathcal{V}_t^U = \beta E_t \left[\frac{C_t}{C_{t+1}} \left(x_{t+1} \mathcal{V}_{t+1}^N + (1 - x_{t+1}) \mathcal{V}_{t+1}^U \right) \right], \tag{2.22}$$

where unemployment income is set to zero, and the probability of being employed, and thus newly hired, in the next period conditional on being unemployed in the current period is the job-finding rate next period, x_{t+1}. From these expressions it is possible to calculate the household's surplus from a newly created job, $\mathcal{V}_t^N - \mathcal{V}_t^U$.

Concerning the firm side, and as in Blanchard and Galí (2010), the surplus of that agent from a newly created job is simply $\mathcal{V}_t^J = G_t$. This is due to the fact that the hiring cost are the marginal cost the firm has to pay when it chooses to substitute a newly hired worker for another one.

Employing the usual sharing rule, $\mathcal{V}_t^N - \mathcal{V}_t^U = \vartheta \mathcal{V}_t^J$, where ϑ indicates the worker's *relative* bargaining weight,[18] results in the following expression for the wage:

$$\begin{aligned} &\frac{W_t^n}{A_t} - \beta(1-\delta) E_t \left(\frac{C_t}{C_{t+1}} \frac{A_{t+1}}{A_t} \frac{W_{t+1}^n}{A_{t+1}} \right) = \\ &\quad \vartheta B x_t^\alpha + \frac{C_t \chi N_t^\phi}{A_t} - \beta(1-\delta) E_t \left[\frac{C_t}{C_{t+1}} \frac{A_{t+1}}{A_t} \left(\vartheta (1 - x_{t+1}) B x_{t+1}^\alpha + \Theta A_{t+1}^{-\gamma} \right) \right]. \end{aligned} \tag{2.23}$$

Combining this with the equilibrium condition (2.19), leads to

$$\frac{\chi C_t N_t^\phi}{A_t} = \frac{1}{\mathcal{M}} - (1+\vartheta) B x_t^\alpha + \beta(1-\delta) E_t \left[\frac{C_t}{C_{t+1}} \frac{A_{t+1}}{A_t} B (x_{t+1}^\alpha + \vartheta x_{t+1}^\alpha (1 - x_{t+1})) \right], \tag{2.24}$$

which together with the equation describing the evolution of newly hired workers (2.6), the definition of labor market tightness (2.8), the aggregate resource constraint $C_t = A_t(N_t - B x_t^\alpha H_t)$, and an exogenous process for A_t characterizes the *equilibrium under heterogeneous wage setting.*

The important thing to note here is that this is the *same equilibrium* as in Blanchard and Galí's (2010) setting with Nash bargaining for every worker and not only for

[18] Alternatively, the optimality condition can be written as $(1-\zeta)\left(\mathcal{V}_t^N - \mathcal{V}_t^U\right) = \zeta \mathcal{V}_t^J$, where $\zeta \in (0,1)$ such that $\vartheta = \frac{\zeta}{1-\zeta} \in (0,\infty)$. ζ indicates the share of the joint surplus going to the household.

new hires, i.e., equation (2.11). Consequently, as in their setup and also as in the constrained efficient allocation, the equilibrium features a constant unemployment level. This, in turn, implies that the short-run inflation unemployment *trade-off,* obtained with a rigid wage for every worker, *disappears,* even though the economy-wide average wage is rigid if $\gamma > 0$. The latter can be seen from the expression for the equilibrium average wage, which results as

$$\begin{aligned} W_t^a &= \Theta A_t^{1-\gamma} + \delta A_t \Bigg[\frac{1}{1-\beta(1-\delta)} \left(\frac{1}{\mathcal{M}} - (1-\beta(1-\delta))B(x^*)^\alpha \right) \\ &\quad - \Theta \sum_{i=0}^{\infty} \beta^i (1-\delta)^i E_t(A_{t+i}^{-\gamma}) \Bigg], \end{aligned} \tag{2.25}$$

where x^* is the (constant) equilibrium job-finding rate. For $\gamma = 0$ the average wage moves one for one with productivity. If $\gamma > 0$, however, this one-for-one relation breaks down and the average wage is rigid. The preceding equation (2.25) is obtained by plugging in the equilibrium Nash bargained wage for the new hires into equation (2.13). This equilibrium Nash wage, in turn, results when combining the equilibrium condition (2.24) and the wage schedule (2.23), yielding

$$\frac{W_t^n}{A_t} = \frac{1}{1-\beta(1-\delta)} \left(\frac{1}{\mathcal{M}} - (1-\beta(1-\delta))B(x^*)^\alpha \right) - \Theta \sum_{i=1}^{\infty} \beta^i (1-\delta)^i E_t(A_{t+i}^{-\gamma}). \tag{2.26}$$

The first term of this expression is the expected discounted Nash wage from period t into the infinite future, where the term in parenthesis is the equilibrium Nash bargained wage in a setup where every worker gets the Nash wage, normalized by productivity. The second term just subtracts the expected discounted future rigid wage starting from period $t+1$, normalized by productivity.[19] Consequently, in expected discounted value terms an individual worker gets the same wage sum over the course of her tenure at a firm in this setup as in the framework with Nash bargaining for every worker in every period. Thus, since what matters for hiring incentives and thus employment fluctuations is the *permanent wage* and not how the stream of wage payments is distributed over the duration of the job, it comes as no surprise that the same

[19] Depending on the stochastic process for productivity, even the Nash wage could be rigid to some degree in this setup.

equilibrium arises in the case with heterogeneous wages as in the case with general Nash bargaining.[20] Moreover, as a result, the expected labor costs for an individual worker over the course of her tenure at an individual firm *normalized by productivity* corresponds to the first term of expression (2.26), which is constant. Hence, expected labor costs move one for one with productivity, eliminating all potential hiring incentives, which in turn leads to an unchanged employment level in response to technology shocks.

In sum, introducing a form of wage rigidity which is consistent with empirical evidence leaves the monetary authority with a single target. It can exclusively focus on inflation with no concern whatsoever for employment stabilization. Furthermore, since wage rigidity cannot be a valid answer to the unemployment volatility puzzle, the question remains what other mechanisms can account for the observed fluctuations in unemployment and what are the implications for monetary policy. The following section sheds light on this issue.

2.3 Introducing markup shocks

In this part of Chapter 2, I follow the suggestion put forward by Pissarides (2007) to generate data-consistent employment fluctuations. He recommends introducing additional driving forces, e.g., in the form of *markup shocks* as in Rotemberg (2008). In particular, I investigate the implications for (optimal) monetary policy. I start out from the model presented in the preceding section, i.e., the New-Keynesian model with labor market frictions of Blanchard and Galí (2010), extended by heterogeneous wage setting as described above. Then, in order to introduce markup shocks and following Steinsson (2003) and Rotemberg (2008), the elasticity of substitution in the Dixit-Stiglitz CES aggregator is assumed to be stochastic. As a result, the elasticity of demand and thus the desired markup are also stochastic. This can be interpreted

[20]For related results with respect to a standard search and matching framework, see, for example, Shimer (2004) and Pissarides (2009). They show that labor market dynamics are unaffected by rigidities in wages for ongoing workers compared to a model with period-by-period Nash bargaining, as long as wages in new matches are determined via the generalized Nash solution. Consequently, also this property carries over to the hiring cost setup employed in this chapter.

economically as a permanently changing market power of firms due to changes in substitutability of the different varieties of goods. The consistency of markup fluctuations with empirical evidence can be seen from Rotemberg and Woodford (1991, 1999) and, more recently, Galí, Gertler, and López-Salido (2007).[21]

The *economic environment* as described in Section 2.2.1 is basically unchanged. The only difference is the stochastic elasticity of substitution, ε_t, in the CES aggregator

$$C_t = \left[\int_0^1 c_t(i)^{\frac{\varepsilon_t - 1}{\varepsilon_t}} di\right]^{\frac{\varepsilon_t}{\varepsilon_t - 1}}, \quad \varepsilon_t > 1. \tag{2.27}$$

Since this does not affect the social planner's problem, the *constrained efficient allocation* is the same as depicted in Section 2.2.2. Thus, the constrained efficient employment level is again constant, i.e., it does not move in response to shocks.

2.3.1 Equilibrium in the decentralized economy with flexible prices

As a next step, consider the decentralized economy, first with flexible prices. Again, with respect to the firm side, by setting the price of its differentiated good optimally each period, the monopolistically competitive firm maximizes its value

$$\max_{\{P_{t+k}(i)\}_{k=0}^{\infty}} E_t \left[\sum_{k=0}^{\infty} Q_{t,t+k}(P_{t+k}(i)Y_{t+k}(i) - P_{t+k}W^a_{t+k}N_{t+k}(i) - P_{t+k}G_{t+k}H_{t+k}(i))\right], \tag{2.28}$$

subject to the production function and demand now given by

$$Y_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\varepsilon_t} (C_t + G_t H_t) \quad \forall t. \tag{2.29}$$

Note that due to the stochastic elasticity of substitution, the elasticity of demand is now also time varying. Furthermore, the firm takes as given the time path for the average real wage, W^a_t, as defined in Section 2.2.3, cost per hire, G_t, the stochastic

[21]The highly tractable modeling approach used in this chapter already provides the main insights concerning the effects of introducing this kind of shocks. The "deep habits" model of Ravn, Schmitt-Grohé, and Uribe (2006) represents a possible framework to provide more fundamental microfoundations for markup variation.

discount factor for nominal payoffs, $Q_{t,t+k}$, and the aggregate price index, which is also slightly altered due to the time-varying market power of firms,

$$P_t = \left[\int_0^1 P_t(i)^{1-\varepsilon_t} di \right]^{\frac{1}{1-\varepsilon_t}} . \tag{2.30}$$

Once more, this optimization problem leads to the well known markup pricing rule

$$\frac{P_t(i)}{P_t} = \mathcal{M}_t \, MC_t \quad \forall t. \tag{2.31}$$

However, now the desired markup reflects the time-varying market power of firms and results as $\mathcal{M}_t \equiv \frac{\varepsilon_t}{\varepsilon_t - 1}$, where real marginal costs are unaltered and thus given by equation (2.17). Furthermore, using the symmetry of the equilibrium in conjunction with this equation leads to an equilibrium condition, given the wage for new hires,

$$\frac{W_t^n}{A_t} = \frac{1}{\mathcal{M}_t} - Bx_t^\alpha + \beta(1-\delta)E_t\left[\frac{C_t}{C_{t+1}}\frac{A_{t+1}}{A_t}\left(\frac{W_{t+1}^n}{A_{t+1}} - \Theta A_{t+1}^{-\gamma} + Bx_{t+1}^\alpha\right)\right], \tag{2.32}$$

where the only difference to the analogous expression (2.19) in Section 2.2.3 is the stochastic markup. In accordance with the heterogeneous wage setting framework, the wage for new hires is derived by using Nash bargaining. Since this derivation is not affected by the introduction of markup shocks, the expression for the wage is unchanged and thus given by equation (2.23). Combining the latter with the equilibrium condition above leads to an equation which together with the equation describing the evolution of newly hired workers (2.6), the definition of labor market tightness (2.8), the aggregate resource constraint, and exogenous processes for A_t *and* ε_t describes the *equilibrium under heterogeneous wage setting and markup shocks:*

$$\frac{\chi C_t N_t^\phi}{A_t} = \frac{1}{\mathcal{M}_t} - (1+\vartheta)Bx_t^\alpha + \beta(1-\delta)E_t\left[\frac{C_t}{C_{t+1}}\frac{A_{t+1}}{A_t}B(x_{t+1}^\alpha + \vartheta x_{t+1}^\alpha(1-x_{t+1}))\right]. \tag{2.33}$$

By using the definition of labor market tightness, the equation describing the evolution of aggregate hirings as well as unemployed and *imposing* $\mathcal{M}_t = \overline{\mathcal{M}}$ $\forall t$ leads to the constant equilibrium employment level

$$N^d = N(x^d) = \frac{x^d}{\delta + (1-\delta)x^d} , \tag{2.34}$$

where x^d denotes the (constant) equilibrium level of labor market tightness. The latter is implicitly given by the solution to the following equation, which is obtained by plugging in the resource constraint and the equation describing the evolution of aggregate hirings into the equilibrium condition (2.33), and imposing again a constant markup:

$$\begin{aligned}&\chi\,(1-\delta Bx^{\alpha})\,N(x)^{1+\phi} = \\ &\qquad \frac{1}{\overline{\mathcal{M}}} - (1-\beta(1-\delta))(1+\vartheta)Bx^{\alpha} - \beta(1-\delta)\vartheta Bx^{1+\alpha}.\end{aligned} \tag{2.35}$$

Consequently, equilibrium output and consumption are given by

$$Y_t^d = A_t N^d \tag{2.36}$$

$$C_t^d = A_t N^d(1-\delta B(x^d)^{\alpha}). \tag{2.37}$$

In order to ensure full participation, I only consider equilibria where the respective wage stays above the marginal rate of substitution which results in the case of full employment

$$\frac{W_t^o}{A_t} > \chi(1-\delta B), \quad \frac{W_t^n}{A_t} > \chi(1-\delta B), \quad \forall t. \tag{2.38}$$

Of course, this constancy result with respect to the employment variables only holds when the markup is constant. If it is not, thus reflecting a time-varying market power of firms, it can be seen from equation (2.35) that x_t^d and in turn N_t^d will move in equilibrium. Furthermore, in addition to the movements of A_t this will also be reflected in equilibrium output and consumption. Intuitively, changes in the substitutability of goods implies changes in the market power of firms, which thus leads to movements in the desired markup. In turn, this shifts the labor demand schedule (2.32), leading to fluctuations in employment and unemployment. Consequently, shocks to the elasticity of substitution translate into movements in those labor market variables in this setup, which are *inefficient,* since the constrained efficient allocation exhibits constant employment variables. Provided the monetary authority assigns at least some weight to unemployment stabilization, the optimal monetary policy problem becomes non-trivial. In order for monetary policy to be effective in this environment, the following section introduces sticky prices. Subsequently, I derive the policymaker's loss function to determine optimal policy.

2.3.2 Equilibrium in the decentralized economy with sticky prices

As is standard in this kind of literature, I introduce nominal rigidities in the form of price staggering à la Calvo (1983). Accordingly, there is a constant probability of firms having the opportunity to adjust their prices each period, $1-\theta$, so that the measure of firms which reset and not reset prices in a given period are $1-\theta$ and θ, respectively. Analogous to the derivation in Blanchard and Galí (2010), value maximization by the representative firm leads to the following optimal price setting rule for a firm with the opportunity to reset prices in period t

$$E_t\left[\sum_{k=0}^{\infty}\theta^k Q_{t,t+k}Y_{t+k|t}\frac{1-\varepsilon_{t+k}}{1-\overline{\varepsilon}}(P_t^*-\mathcal{M}_{t+k}P_{t+k}MC_{t+k})\right]=0. \tag{2.39}$$

$Y_{t+k|t}$ is the output level at time $t+k$ for a firm adjusting its price at time t, $\overline{\varepsilon}>1$ is the steady-state value of the elasticity of substitution, P_t^* is the optimal price set in period t by the firm under consideration, and real marginal costs are unchanged compared to the flexible price case and thus given by equation (2.17). Except for the latter, the preceding optimal price setting rule collapses to the corresponding expression obtained by Blanchard and Galí (2010), if the elasticity of substitution is nonstochastic, i.e., $\varepsilon_t=\overline{\varepsilon}=\varepsilon\ \forall t$.

In a next step, I log-linearize the equilibrium relations obtained so far. First, consider the optimal price setting rule (2.39) as well as the expression for the price index in this case

$$P_t=\left(\theta(P_{t-1})^{1-\varepsilon_t}+(1-\theta)(P_t^*)^{1-\varepsilon_t}\right)^{\frac{1}{1-\varepsilon_t}}. \tag{2.40}$$

Log-linearizing those equations around a zero inflation steady state and combination leads to a New-Keynesian Phillips curve of the form

$$\pi_t=\beta E_t(\pi_{t+1})+\lambda(\widehat{mc}_t+\widehat{m}_t), \tag{2.41}$$

where hat variables denote log deviations from their respective steady-state values, i.e., $\widehat{mc}_t = \log\left(\frac{MC_t}{MC}\right)$ as well as $\widehat{m}_t = \log\left(\frac{\mathcal{M}_t}{\overline{\mathcal{M}}}\right)$, and $\lambda \equiv \frac{(1-\theta\beta)(1-\theta)}{\theta}$. The presence of the additional term $\lambda\widehat{m}_t$ indicates that stabilizing marginal cost does not lead to stable inflation and vice versa. Furthermore, as implied by the following derivations, it is possible to establish a relation between marginal cost and unemployment. Thus, the appearance of this extra term highlights the short-run trade-off between inflation and unemployment.

As a first step to derive such a relation between marginal cost and unemployment, I log-linearize equation (2.17), i.e., the expression for real marginal cost, leading to

$$\begin{aligned} \widehat{mc}_t \quad = \quad & \overline{\mathcal{M}}W^n(\widehat{w}^n_t - \widehat{a}_t) + \overline{\mathcal{M}}\alpha g\widehat{x}_t \\ & -\beta(1-\delta)\overline{\mathcal{M}}E_t[(W^n - \Theta + g)[(\widehat{c}_t - \widehat{a}_t) - (\widehat{c}_{t+1} - \widehat{a}_{t+1})] \\ & +W^n(\widehat{w}^n_{t+1} - \widehat{a}_{t+1}) + \Theta\gamma\widehat{a}_{t+1} + \alpha g\widehat{x}_{t+1}]. \end{aligned} \tag{2.42}$$

Variables without a time index denote steady-state values where steady-state productivity is normalized to one, i.e., $A = 1$, $g \equiv Bx^\alpha$, and note that $MC = \frac{1}{\overline{\mathcal{M}}}$.

Moreover, approximation of the wage equation (2.23) results in

$$\begin{aligned} & \overline{\mathcal{M}}W^n\Big[(\widehat{w}^n_t - \widehat{a}_t) - \beta(1-\delta)E_t(\widehat{w}^n_{t+1} - \widehat{a}_{t+1})\Big] = \\ & \quad \overline{\mathcal{M}}\Big[(\vartheta g - W^n)\widehat{a}_t + \alpha\vartheta g\widehat{x}_t + C\chi N^\phi\widehat{c}_t + \phi C\chi N^\phi\widehat{n}_t \\ & \quad -\beta(1-\delta)E_t[(\vartheta(1-x)g + \Theta - W^n)(\widehat{c}_t + (\widehat{a}_{t+1} - \widehat{c}_{t+1})) \\ & \quad -\Theta\gamma\widehat{a}_{t+1} + \vartheta g(\alpha(1-x) - x)\widehat{x}_{t+1}]\Big], \end{aligned} \tag{2.43}$$

which can be combined with the preceding equation to yield the following expression for real marginal cost

$$\begin{aligned} \widehat{mc}_t \quad = \quad & \overline{\mathcal{M}}(\vartheta g - W^n)\widehat{a}_t + \overline{\mathcal{M}}\alpha g(1+\vartheta)\widehat{x}_t + \overline{\mathcal{M}}C\chi N^\phi\widehat{c}_t \\ & +\overline{\mathcal{M}}\phi C\chi N^\phi\widehat{n}_t - \beta(1-\delta)\overline{\mathcal{M}}E_t\Big[(\vartheta(\alpha(1-x) - x) + \alpha)g\widehat{x}_{t+1} \\ & -g(1+\vartheta(1-x))(\widehat{c}_{t+1} - \widehat{a}_{t+1}) + g(1+\vartheta(1-x))\widehat{c}_t \\ & -(W^n - \Theta + g)\widehat{a}_t\Big]. \end{aligned} \tag{2.44}$$

Next, log-linearization of the desired markup, labor market tightness, and consumption leads to

$$
\begin{aligned}
\widehat{m}_t &= \frac{1}{1-\bar{\varepsilon}}\widehat{\varepsilon}_t &(2.45)\\
\delta\widehat{x}_t &= \widehat{n}_t - (1-\delta)(1-x)\widehat{n}_{t-1} &(2.46)\\
\widehat{c}_t &= \widehat{a}_t + \frac{1-g}{1-g\delta}\widehat{n}_t + \frac{g(1-\delta)}{1-g\delta}\widehat{n}_{t-1} - \frac{\alpha g}{1-g\delta}\delta\widehat{x}_t. &(2.47)
\end{aligned}
$$

Then, consider the optimization problem of the representative household, i.e., maximizing (2.1) subject to the following budget constraint

$$\int_0^1 P_t(i)c_t(i)di + Q_tB_t \leq B_{t-1} + W_t^aN_t + T_t \tag{2.48}$$

and the no-Ponzi-game condition

$$\lim_{T\to\infty} E_t(B_T) \geq 0 \quad \forall t, \tag{2.49}$$

where $c_t(i)$ is household's consumption of good i in period t, B_t denotes purchases in period t of one-period nominal riskless discount bonds, where at maturity a single bond pays one unit of money, Q_t is the price of that bond, and T_t indicates the lump-sum part of household income, e.g., dividends due to firm ownership. Using the demand equation $c_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\varepsilon_t} C_t$ and the definition of the aggregate price index, it is possible to rewrite the budget constraint as

$$P_tC_t + Q_tB_t \leq B_{t-1} + W_t^aN_t + T_t. \tag{2.50}$$

Next, solving this optimization problem and approximating the resulting consumption Euler equation yields

$$\widehat{c}_t = E_t(\widehat{c}_{t+1}) - (i_t - E_t(\pi_{t+1}) - \rho), \tag{2.51}$$

where $i_t \equiv -\log Q_t$ is the short-term nominal interest rate and $\rho \equiv -\log\beta$ denotes the household's discount rate.[22]

[22]Note that $Q_t = \frac{1}{1+\bar{\imath}_t}$, where $\bar{\imath}_t$ is the yield of the one-period bond. The definition of the nominal interest rate is motivated by the approximation $\log(1+x) \approx x$, which is accurate for small x. The discount rate is derived analogously.

The preceding equations (2.41) and (2.44) to (2.47) as well as (2.51) together with exogenous processes for technology and the elasticity of substitution, as well as a characterization of monetary policy specify the equilibrium of this model. The equations just given, however, can be significantly simplified by applying the additional approximations suggested by Blanchard and Galí (2010). First, they assume that hiring costs are small relative to output, or more exactly, g as well as δ are considered to be of the same order of magnitude as fluctuations in $\widehat{n}_t$. Accordingly, terms featuring $g\widehat{n}_t$, $\delta\widehat{n}_t$, and δg can be dropped since they are of second order. A second approximation the aforementioned authors consider to be justified is that "fluctuations in $\widehat{x}_t$ are large relative to those in $\widehat{n}_t$" (p. 14). This assumption is motivated by the log-linearization of labor market tightness, i.e., equation (2.46). It implies that $\delta\widehat{x}_t$ and $\widehat{n}_t$ are of the same order of magnitude so that terms featuring the former expression cannot be dropped. Furthermore, since δ and g are assumed to be of the same order, this also holds for $g\widehat{x}_t$. These two assumptions markedly simplify the expressions for consumption and marginal cost[23]

$$\widehat{c}_t = \widehat{a}_t + \widehat{n}_t \tag{2.52}$$

$$\widehat{mc}_t = \overline{\mathcal{M}}\chi N^{1+\phi}(1+\phi)\widehat{n}_t + \overline{\mathcal{M}}\alpha g(1+\vartheta)[\widehat{x}_t - \beta E_t(\widehat{x}_{t+1})]. \tag{2.53}$$

As a subsequent step, I combine the preceding equations in order to obtain a more compact description of the equilibrium. First, using the approximation $\widehat{u}_t = -(1-u)\widehat{n}_t$, where $\widehat{u}_t \equiv u_t - u$ is the deviation of the end-of-period unemployment rate from steady state, in conjunction with the log-linearization of labor market tightness (2.46) leads to

$$\widehat{u}_t = (1-\delta)(1-x)\widehat{u}_{t-1} - (1-u)\delta\widehat{x}_t. \tag{2.54}$$

Second, plugging the expressions for marginal cost (2.53) and the optimal markup (2.45) into the New-Keynesian Phillips curve (2.41), using the aforementioned approximation of unemployment, as well as the preceding equation yields

[23] Note that equation (2.53) relates marginal cost to employment and labor market tightness, which both in turn can be linked to unemployment. Hence, a relation between marginal cost and unemployment results, as mentioned at the beginning of this section.

$$\pi_t = -\eta \sum_{k=1}^{\infty} \beta^k E_t(\widehat{u}_{t+k}) - (\eta + \kappa(1+\vartheta))\widehat{u}_t$$
$$+\kappa(1+\vartheta)(1-\delta)(1-x)\widehat{u}_{t-1} + \frac{\lambda}{1-\overline{\varepsilon}} \sum_{k=0}^{\infty} \beta^k E_t(\widehat{\varepsilon}_{t+k}), \qquad (2.55)$$

where

$$\kappa \equiv \frac{\alpha g \overline{\mathcal{M}} \lambda}{(1-u)\delta}$$
$$\eta \equiv \frac{\lambda \overline{\mathcal{M}} \chi (1-u)^{1+\phi}(1+\phi)}{1-u}.$$

Third, I combine the expressions for unemployment, consumption, as well as the log-linearized consumption Euler equation (2.51) to obtain

$$\widehat{u}_t = E_t(\widehat{u}_{t+1}) + (1-u)(i_t - E_t(\pi_{t+1}) - r_t^*), \qquad (2.56)$$

where $r_t^* = \rho - \widehat{a}_t + E_t(\widehat{a}_{t+1})$ is the efficient real interest rate.

In sum, the New-Keynesian Phillips curve (2.55) together with the expectational IS curve (2.56), exogenous processes for technology and the elasticity of substitution, as well as a description of monetary policy completely specify the equilibrium of this model.

As implied by the constancy of the unemployment level in the constrained efficient allocation, an optimal policy in this environment would aim at achieving both constant unemployment and inflation. This is, however, not feasible in the presence of shocks to the elasticity of substitution, as can be seen from the Phillips curve relation (2.55). For example, stabilizing unemployment by setting the nominal interest rate in such a way that the real interest rate tracks the efficient rate, will not achieve stable inflation, as shocks to the market power of firms, i.e., the last term in equation (2.55), will lead to changes in the inflation rate. Analogously, stabilizing inflation will not achieve constant unemployment, since such a policy basically delivers unemployment equal to its natural rate. The latter, in turn, is the level obtained in a setting without nominal rigidities, which is not constant, as shown in the preceding section. The fundamental reason behind these results is the non-constancy of the gap between the natural rate and the constrained efficient unemployment level. If it was constant then stabilizing the gap between the actual unemployment rate and the natural rate, done via a constant inflation policy, would be equivalent to stabilizing the gap between actual and efficient unemployment, which is the gap relevant from a welfare point of view. In the model described here, however, shocks to the elasticity of substitution will result in a gap between the natural and efficient level which is not constant so that simply implementing a policy of stable inflation will not achieve the optimal outcome. In the words of Blanchard and Galí (2007), there is no “divine coincidence.” I analyze the nature of this short-run inflation unemployment trade-off and its implications for monetary policy in more detail in the next section.

2.4 Monetary policy analysis

In order to continue the investigation, it is necessary to specify the stochastic processes governing technology and, most importantly, the elasticity of substitution. Following Blanchard and Galí (2010), both are assumed to be described by stationary AR(1) processes

$$\begin{aligned} \widehat{a}_{t+1} &= \rho_a \widehat{a}_t + e^a_{t+1}, \quad |\rho_a| < 1, \quad e^a_{t+1} \overset{iid}{\sim} (0, \sigma^2_{e^a}) && (2.57) \\ \widehat{\varepsilon}_{t+1} &= \rho_\varepsilon \widehat{\varepsilon}_t + e^\varepsilon_{t+1}, \quad |\rho_\varepsilon| < 1, \quad e^\varepsilon_{t+1} \overset{iid}{\sim} (0, \sigma^2_{e^\varepsilon}). && (2.58) \end{aligned}$$

Combining this with the Phillips curve (2.55) leads to the following simplification of the latter relation

$$\begin{aligned} \pi_t &= -\eta \sum_{k=1}^{\infty} \beta^k E_t(\widehat{u}_{t+k}) - (\eta + \kappa(1+\vartheta))\widehat{u}_t \\ &\quad + \kappa(1+\vartheta)(1-\delta)(1-x)\widehat{u}_{t-1} + \frac{\lambda}{1-\bar{\varepsilon}} \frac{1}{1-\beta\rho_\varepsilon} \widehat{\varepsilon}_t. && (2.59) \end{aligned}$$

The expectational IS curve (2.56), in turn, is used to determine the respective interest rate rule leading to the allocation characterized by the particular time path for inflation and unemployment under a certain policy regime.

2.4.1 Two polar cases and optimal monetary policy

To obtain a first impression concerning the extent of the short-run trade-off, two polar cases are considered.[24] First, I examine a policy which implements the constrained efficient unemployment level. Consequently, this implies an approach which completely stabilizes unemployment around its efficient level, i.e., $\widehat{u}_t = \widehat{n}_t = \widehat{x}_t = 0 \; \forall t$. Thus, by equation (2.59), inflation in such a regime is described by the following equation

$$\pi_t = \frac{\lambda}{1-\bar{\varepsilon}} \frac{1}{1-\beta\rho_\varepsilon} \widehat{\varepsilon}_t. \tag{2.60}$$

The magnitude of the inflation fluctuations resulting from shocks to the market power of firms depends, of course, on the parameters of the model. The more persistent the process of the elasticity of substitution and the closer its steady-state value is to one, the larger will be the fluctuations in inflation in response to shocks. On the other hand, the higher the degree of nominal rigidities, i.e., the larger θ, leading to a smaller λ, the smaller are the inflation fluctuations in absolute terms. However, in general, it should be noted that the overall coefficient on $\widehat{\varepsilon}_t$ is negative. Thus, a positive deviation of the elasticity of substitution from its steady state will lead to a reduction in the inflation rate. Intuitively, an increase in the elasticity of substitution implies a loss in market power for the individual firm, leading to a reduction of the desired markup. The latter, in turn, due to markup pricing in this environment, brings about the reduction in prices. All this, however, is influenced by price staggering of firms, implying a role for the degree of price rigidity. The higher the rigidity, the larger the incentive for a firm actually being able to adjust prices not to reduce them as much as in the flexible price case, in order to avoid setting prices in deviation from the general price level. This mechanism amplifies the direct effect of a higher degree of rigidity, being that fewer firms are able to change their price in the first place.

As a second polar case, I consider a policy of completely stabilizing inflation, which in the usual setup of a New-Keynesian model featuring a divine coincidence, would also lead to stable unemployment. Here, however, when setting $\pi_t = 0 \; \forall t$ the Phillips curve implies

[24] Note that implicitly it is assumed that monetary policy is completely credible.

$$
\begin{aligned}
\widehat{u}_t &= -\frac{\eta}{\eta+\kappa(1+\vartheta)}\sum_{k=1}^{\infty}\beta^k E_t(\widehat{u}_{t+k}) + \frac{\kappa(1+\vartheta)(1-\delta)(1-x)}{\eta+\kappa(1+\vartheta)}\widehat{u}_{t-1} \\
&\quad + \frac{1}{\eta+\kappa(1+\vartheta)}\frac{\lambda}{1-\overline{\varepsilon}}\frac{1}{1-\beta\rho_\varepsilon}\widehat{\varepsilon}_t. \qquad (2.61)
\end{aligned}
$$

Current unemployment will depend on future expected unemployment and also exhibits serial correlation beyond that of the elasticity of substitution. Moreover, shocks to the latter move unemployment in the same direction as inflation in the preceding case. The magnitude, however, is influenced by an additional factor depending mainly on parameters describing the labor market, e.g., steady-state employment and hiring costs as well as the separation rate, but also on the steady-state markup, the workers' relative bargaining power, and the degree of nominal rigidities. Intuitively, the loss in market power due to the increased substitutability of goods, implies a reduction of the desired markup. This shifts out the labor demand schedule (2.32), leading to an increase in employment and reduced unemployment. Overall, this positive shock to the elasticity of substitution brings the equilibrium closer to the one under perfect competition. The latter, compared to the equilibrium under monopolistic competition, features higher output as well as employment and a lower price level. Furthermore, the movements of unemployment under a flexible price regime will also be described by the equation above, since a policy of inflation stabilization brings about the same allocation as the one of a setup with flexible prices.

Finally, I consider optimal monetary policy where it is possible for the monetary authority to credibly precommit to such a strategy. In this regard, I follow Blanchard and Galí (2010) and assume that unemployment moves around a constrained efficient steady-state value. Furthermore, as is standard in this kind of literature, I base my normative analysis on the preferences of the private agents in the economy. In particular, a second-order Taylor series approximation to the level of expected utility of the representative household in the steady state is performed.[25] An analogous derivation to the one in Blanchard and Galí (2010) leads to the following quadratic

[25] For an extensive overview of this approach to welfare analysis, see Woodford (2003).

loss function:

$$\mathcal{L} = E_0 \left[\sum_{t=0}^{\infty} \beta^t (\pi_t^2 + \alpha_u \widehat{u}_t^2) \right], \tag{2.62}$$

where

$$\alpha_u \equiv \frac{\lambda(1+\phi)\chi(1-u)^{\phi-1}}{\bar{\varepsilon}} > 0. \tag{2.63}$$

Consequently, the *monetary authority's problem* is to minimize the loss function (2.62) subject to the time path of Phillips curve relations (2.59), $\forall t$. Unfortunately, there does not exist an analytical solution so that numerical methods are used. In particular, I calibrate the model and use the approach of Söderlind (1999) to obtain the dynamics of the economy in response to shocks under the optimal commitment policy.

2.4.2 Calibration and dynamics of the economy

In order to get an impression of the dynamic properties of the model in response to shocks to the elasticity of substitution under the aforementioned policy regimes, I first calibrate the equilibrium relations, then numerically solve for the optimal policy if applicable, and finally simulate the evolution of the endogenous variables in response to those shocks.[26]

Concerning the calibration, I take the same values as Blanchard and Galí (2010), and consequently the same time structure, i.e., one time period in the model is chosen to correspond to a quarter. This is done mainly for comparability reasons, but also since the basic modeling structure is similar so that their values also apply to the setup of this chapter. Moreover, to investigate the implications of different degrees of rigidity in the labor market, I also distinguish between two calibrations. The first represents a flexible labor market characterized by a low steady-state unemployment rate and high job-finding and separation rates. The second corresponds to a more sclerotic labor market featuring a higher unemployment rate and lower turnover, i.e., lower

[26] As can be inferred from the preceding equations and also from the discussion in Section 2.2.3, productivity shocks do not lead to a short-run inflation unemployment trade-off in this model. In particular, these shocks do not bring about movements in those endogenous variables. Consequently, in the calibration exercise I only consider shocks to the market power of firms.

Table 2.1: Calibration (common values)

Symbol	Value	Description
β	0.99	household's discount factor
ϕ	1	inverse of Frisch labor supply elasticity
$\bar{\varepsilon}$	6	steady-state elasticity of substitution
$\overline{\mathcal{M}}$	1.2	steady-state markup
θ	0.67	measure of firms not resetting prices in a given period
α	1	elasticity of hiring costs w.r.t. labor market tightness
B	0.11	scale factor of hiring costs
ϑ	1	worker's relative bargaining weight

job-finding and separation rates. I present a first set of parameters, being identical in both calibrations, in Table 2.1. The values given are standard in the literature and consistent with the relevant micro and macro evidence. Fortunately, it is not necessary to calibrate the parameter for which there is the weakest empirical basis in Blanchard and Galí's (2010) calibration: the one describing the degree of real wage rigidity. This is due to the fact that the equilibrium as presented at the beginning of this section does not depend on the latter.

Table 2.2 indicates the calibration for the flexible and sclerotic labor market, respectively. As in the preceding table, the parameters are chosen such that they correspond to the relevant empirical evidence, where the flexible labor market refers to the United States and the sclerotic labor market to continental Europe. This table also includes the parameter governing the relative importance of the disutility of work in total utility, χ. The latter is set to obtain an efficient steady state in the two calibrations, in order to be able to apply the log-linear approximation in both cases.[27]

Figures 2.1 - 2.3 present the results of the simulation exercise, i.e., the dynamic responses of unemployment and inflation to a shock to the elasticity of substitution under the various policy regimes and different degrees of persistence of the shock. The impulse response functions plot the dynamics of the endogenous variables in percent over a horizon of 20 periods in response to a market power shock corresponding to a

[27]For more on the calibration, see Blanchard and Galí (2010).

Table 2.2: Calibration (specific values)

Symbol	Value (flexible)	Value (sclerotic)	Description
x	0.7	0.25	steady-state job-finding rate
u	0.05	0.1	steady-state end-of-period unemployment rate
δ	0.12	0.04	separation rate
g	0.077	0.028	steady-state hiring costs
χ	1.03	1.22	scale factor of disutility of work

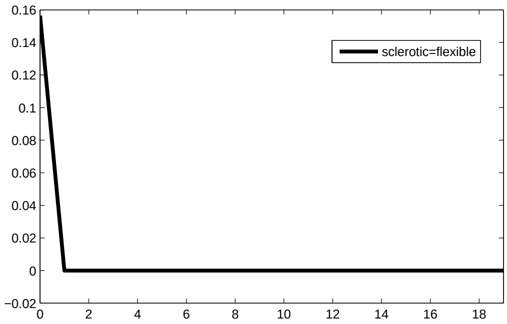
one percent increase in the desired markup.[28]

Consider first a policy of complete unemployment stabilization. Consequently, Figure 2.1 only shows the inflation response to a market power shock. Moreover, as can be seen from equation (2.60), the impulse response function does not depend on the degree of rigidity in the labor market so that the dynamic responses under the "sclerotic" and the "flexible" labor market calibration coincide. As expected, the magnitude and persistence of the inflation response increases with an increasing degree of persistence of the shock. For the case of a purely transitory shock, inflation is only affected on impact and to a relatively small extend, it increases by about 0.16%. Increasing the autoregressive parameter to 0.5 and then 0.9 amplifies the instantaneous impact considerably, being now approximately 0.31% and 1.44%, respectively. This is a result of the forward looking character of the Phillips curve. Moreover, it also increases the persistence of the response, which becomes particularly apparent in the case of $\rho_\varepsilon = 0.9$. Basically, the persistence of the shock carries over to the inflation process. Furthermore, the magnitude of the inflation response is in all cases about as large as the comparable one in Blanchard and Galí (2010), where productivity shocks are considered.

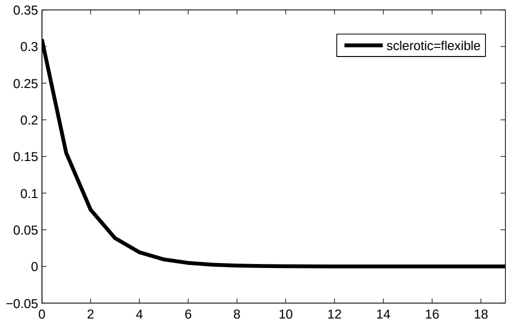
Next, Figure 2.2 depicts the response of unemployment to a shock to the elasticity of substitution under a policy which completely stabilizes inflation. In this case, the responses in a sclerotic and a flexible labor market are different, even though not to a large extent. Analogous to the mechanism indicated above, the inward shift in the

[28]Note that this implies a *decrease* in the elasticity of substitution.

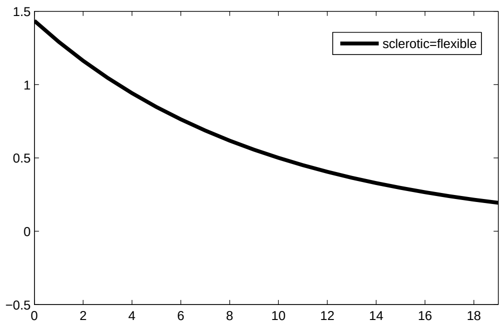
Figure 2.1: Unemployment stabilization regime



(a) $\rho_\varepsilon = 0$



(b) $\rho_\varepsilon = 0.5$

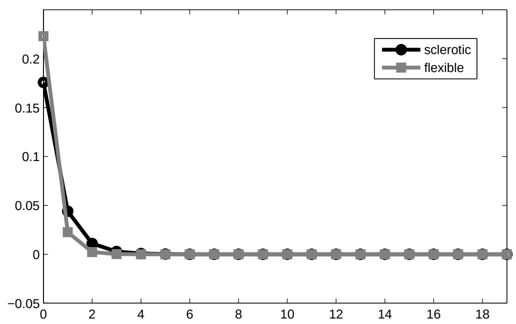


(c) $\rho_\varepsilon = 0.9$

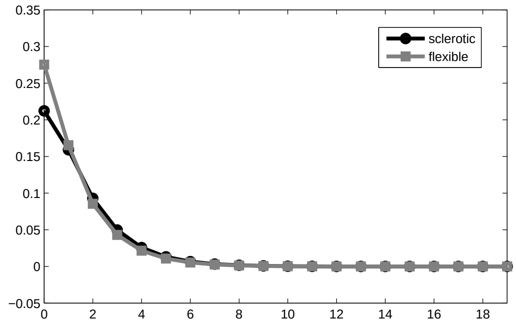
Notes: These figures show the respective response of inflation to shocks to market power for different degrees of persistence of the shock process.

labor demand schedule due to the increase in the desired markup temporarily leads to a higher level of unemployment, which ultimately reverts back to its steady-state level. As expected, the speed of this reversion is inversely related to the degree of persistence of the shock. The latter also influences the magnitude of the unemployment response, however not by as much as in the preceding case. The maximum response for the "sclerotic" calibration, for instance, increases from 0.18% via 0.21% to 0.29%. The last maximum occurs not on impact as in the other cases, but in the subsequent period, thereby indicating the well-known hump-shaped pattern. Overall, the flexible labor market exhibits slightly larger increases in unemployment, which are less persistent than in the "sclerotic" calibration, however. This is basically a consequence of the higher turnover, in particular, a larger sacrifice ratio, under the "flexible" calibration. Due to the smaller responsiveness of inflation to changes in unemployment as indicated

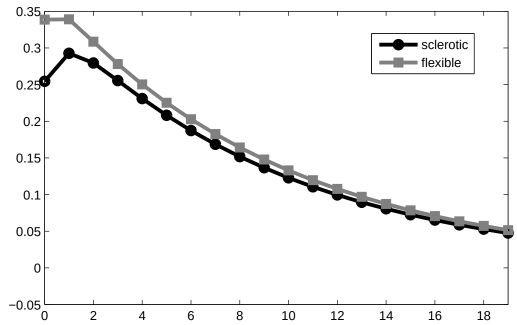
Figure 2.2: Inflation stabilization regime



(a) $\rho_\varepsilon = 0$



(b) $\rho_\varepsilon = 0.5$



(c) $\rho_\varepsilon = 0.9$

Notes: These figures show the respective response of unemployment to shocks to market power for different degrees of persistence of the shock process.

by the coefficients of the Phillips curve, the movements in unemployment which are needed to obtain the same change in inflation are larger in the US-style calibration than in the European one. This stems from two factors: a higher separation rate and a smaller steady-state unemployment rate in the "flexible" specification. With respect to the former, as indicated by equation (2.46), in an environment with a higher separation rate, larger movements in *employment* are needed to obtain a given change in labor market tightness. These feed via changes in marginal cost into changes in inflation.[29] As a second factor, the smaller steady-state unemployment rate leads, in percentage terms, to larger changes in *unemployment* which are necessary to obtain a given change in employment. The general intuition is that a particular change in employment can be digested much easier by a flexible labor market, featuring a higher

[29]See equations (2.53) and (2.41).

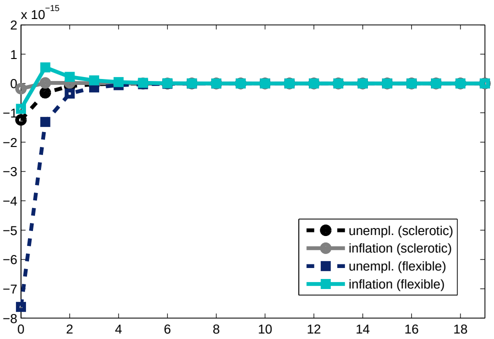
turnover. Furthermore, due to the same channel as for the former factor, higher separation and steady-state job-finding rates lead to a less persistent unemployment response. Again, as indicated by equation (2.46), last period's employment is less of an importance for changes in labor market tightness the larger δ and x.[30]

Overall, the responses of unemployment under this policy are quite small and, in particular, considerably smaller than in Blanchard and Galí (2010). Furthermore, as indicated by those dynamic responses, the model allows for the markup to move countercyclically, i.e., an increase in the desired markup coincides with an increase in unemployment or equivalently a decrease in employment. These countercyclical movements correspond to empirical evidence as presented, for example, in Rotemberg and Woodford (1991, 1999) and Galí, Gertler, and López-Salido (2007).

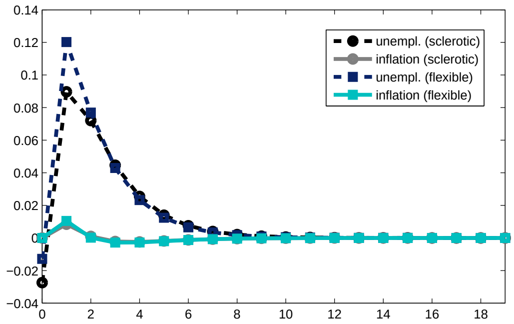
Figure 2.3, finally, presents the dynamics of unemployment and inflation under the optimal monetary policy. Again, the evolution of the endogenous variables differs only slightly between the sclerotic and flexible labor market calibration. The difference in the dynamics, i.e., the marginally larger but less persistent unemployment response in the flexible labor market, can be explained as in the preceding case. In response to a purely transitory shock to market power, both unemployment and inflation practically do not move. Qualitatively, however, they broadly follow the same pattern as in the two cases with a persistent market power shock, which I describe in the following. In those two simulations, it is optimal to almost completely stabilize inflation. Only in the second time period is the inflation rate slightly positive. Unemployment, on the other hand, decreases somewhat on impact and subsequently it follows a path quite close to the one under complete inflation stabilization, i.e., in particular an increase followed by a reversion back to the steady-state unemployment level. The magnitude of the unemployment response, however, is smaller than in the case of complete inflation stabilization. As is characteristic of a policy of optimal commitment, by inducing a certain time path of inflation expectations it is possible to improve the trade-off between inflation and unemployment stabilization faced by the monetary authority in the period of the shock. More specifically, here the expectation of a monetary policy

[30]Ultimately, the effect on unemployment persistence can be seen from the coefficient on $\widehat{u}_{t-1}$ in equation (2.61), which is 0.108 under the US-style calibration and 0.299 under the European specification.

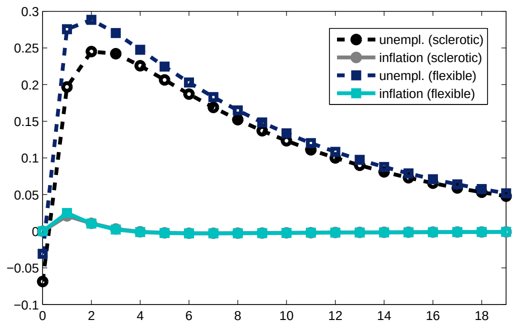
Figure 2.3: Optimal policy regime



(a) $\rho_\varepsilon = 0$



(b) $\rho_\varepsilon = 0.5$



(c) $\rho_\varepsilon = 0.9$

Notes: These figures show the respective responses of unemployment and inflation to shocks to market power for different degrees of persistence of the shock process.

response that leads to an increase in unemployment for a couple of periods following the inflationary market power shock, induces the expectation of countervailing deflationary pressures in the coming periods. This brings about less pronounced price increases in the first place so that the initial unemployment response does not have to be that large to counter the inflationary pressures. In fact, on impact there is a *negative* unemployment response, which is also explained by the optimal stabilization motive and the particular form of the Phillips curve (2.59). Since it features lagged unemployment with a *positive* coefficient, it is optimal to reduce unemployment initially to induce additional *deflationary* pressures in the next period to counter the inflationary shock. The counterproductive *inflationary* effect obtained on impact due to this approach can easily be balanced by offsetting expectations of positive deviations of unemployment from steady state. Moreover, as in the preceding case, the dynamic responses indicate countercyclical movements of the markup in this model. Overall, an optimal policy in this environment with shocks to the elasticity of substitution is one of "price stability" with only minor emphasis on employment stabilization.

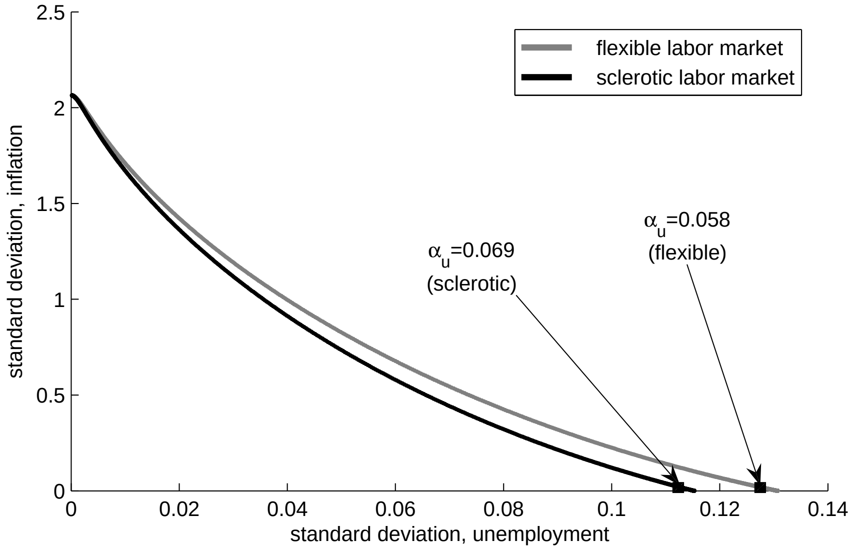
The preceding results can also be illustrated by computing the efficient policy frontier, i.e., the plot of the standard deviations of unemployment and inflation under a policy of optimal commitment while varying the relative weight on unemployment stabilization in the monetary authority's loss function from zero to one. Such a graph highlights, in particular, the trade-off between unemployment and inflation variability faced by the policymaker. Figure 2.4, for instance, presents the results when $\rho_\varepsilon = 0.9$ for both the flexible and sclerotic calibration.[31] As expected from the preceding discussion, the findings for the two specifications are not markedly different. Only the standard deviation of unemployment is slightly larger for the flexible labor market, for the reasons stated above, when the monetary authority assigns at least some weight to inflation stabilization. However, in line with the preceding results, the standard deviation of unemployment is rather small, not exceeding 0.115 and 0.131 under the European- and US-style calibration, respectively. The maximum inflation variability, on the other hand, is considerably larger, increasing to a value of 2.064 under a policy of complete unemployment stabilization. In that case the standard deviations under both specifications coincide in accordance with equation (2.60). Apart from this variability relation in favor of inflation stabilization, the curvature of the efficient policy frontier also points to the direction of the optimality of a strong focus on stabilizing inflation.[32] It is already quite steep in the lower right region, whereas the social marginal rate of substitution between unemployment and inflation variability, i.e., the α_u resulting from the respective calibration, is rather small, being just 0.058 and 0.069 in the flexible and sclerotic calibration, respectively.[33] Consequently, as highlighted in

[31] In particular, each point on the efficient policy frontier depicts the combination of the standard deviation of unemployment in deviation from steady state as well as of the annualized rate of inflation corresponding to a certain value of α_u in the loss function (2.62). Furthermore, a value for $\sigma^2_{e^\varepsilon}$ in the specification of the process of the market power shock is assumed such that a shock to $\widehat{\varepsilon}_t$ in the magnitude of one standard deviation changes annualized inflation, $4\pi_t$, by one percentage point in absolute value.

[32] Note the different scale of the two axes.

[33] This small weight on unemployment variability in the loss function is in line, for instance, with

Figure 2.4: Efficient policy frontier



Notes: This figure plots the standard deviation (in percentage points) of unemployment in deviation from steady state and of the annualized rate of inflation under optimal commitment as the relative weight on unemployment stabilization varies from zero to one ($\rho_\varepsilon = 0.9$).

the graph, the standard deviation combinations referring to optimal policy, indicated by a point of tangency of an indifference curve with slope α_u with the efficient policy frontier, are already in that lower right region. Thus, unemployment variability is only slightly reduced compared to a policy of complete inflation stabilization, whereas inflation variability is still quite small.

In sum, the fluctuations introduced by shocks to the elasticity of substitution are, in general, not large. In particular, at least in the setting presented in this chapter, it does not seem to be a promising approach of using this kind of shocks to generate a significant amount of volatility in unemployment. Moreover, the result of the optimality of stabilizing inflation also in this environment with labor market frictions and markup shocks is consistent with much of the recent literature on optimal monetary policy, represented by Woodford (2003), for example. This finding contrasts, however, with the results of Blanchard and Galí (2010), where only technology shocks are considered. Optimal policy in their model leads to a significant reduction in

results presented in Woodford (2003).

unemployment fluctuations compared to a policy that completely stabilizes inflation. This, in turn, implies some additional inflation fluctuations, thus reflecting a more pronounced focus on unemployment relative to inflation than in the model presented here.[34]

2.5 Conclusion

In this chapter, I investigate the consequences of introducing heterogeneous wage setting and shocks to the market power of firms for the dynamics of inflation and unemployment under different monetary policy regimes and, in particular, optimal monetary policy.

In the first part, I examine the implications for the equilibrium allocation as well as monetary policy of incorporating a form of wage rigidity which is consistent with empirical evidence. Even though widely used in the literature in order to obtain data-consistent fluctuations in labor market variables, uniform wage rigidity is not supported by empirical evidence. I show by using the modeling framework of Blanchard and Galí (2010) that replacing the latter type of rigidity by a form of wage rigidity which is consistent with empirical results can have significant consequences for the equilibrium allocation and conduct of monetary policy. In that environment, the equilibrium unemployment level is again constant, and hence the short-run inflation unemployment trade-off disappears, even though the *average* wage is still sticky. Since expected labor costs are highly responsive to changes in labor productivity, hiring incentives are unaffected by technology shocks, which in turn leads to the absence of unemployment fluctuations.

Still, this leaves the question unanswered what other mechanisms can provide a

[34] In a recent paper, Sala, Söderström, and Trigari (2008) find considerably larger effects of a corresponding type of markup shock. In their setup with various sources of disturbances, the significant trade-off obtained mainly results from price markup shocks. However, the magnitude of this trade-off critically depends on the presence of wage rigidities. Under flexible wages, the standard deviation of the unemployment gap when completely stabilizing inflation, reduces by a factor of eight compared to the case of rigid wages. The latter are introduced into their model via staggered wage bargaining, and thus they do not distinguish between wages for ongoing workers and new hires, which is important for search and matching models, however, as discussed in the introduction of this chapter.

solution to the unemployment volatility puzzle and what are the implications for monetary policy. Consequently, in the second part, I follow the suggestion of Pissarides (2007) and incorporate additional driving forces in the form of markup shocks into the preceding model with heterogeneous wage setting. These shocks lead via markup pricing to movements in prices and via shifts in the labor demand schedule to employment and unemployment dynamics. Even though, as a consequence, a short-run trade-off arises in this setup with markup shocks, optimal policy primarily focuses on inflation stabilization. This finding is in line with the general prescription put forward in the standard literature concerning optimal monetary policy. It is not least due to the other main result of this part of this chapter, i.e., shocks to the market power of firms do not generate significant fluctuations in unemployment in this model.

Nevertheless, since wage rigidity does not provide a data-consistent solution to the unemployment volatility puzzle, a further examination of the empirical performance of alternative approaches to generate fluctuations in labor market variables, and an investigation of the consequences for monetary policy are relevant topics for future research.

Chapter 3

Does anticipation of government spending matter? The role of (non-)defense spending

3.1 Introduction

The empirical literature on the effects of fiscal policy on the macroeconomy is inconclusive. It can broadly be divided into two strands according to the identification approach. On the one hand, fiscal policy events are identified with the narrative approach employing dummy variables that indicate large increases in government expenditure related to wars.[1] These foreign policy events are assumed to be exogenous to the state of the economy and can therefore be used to identify the effects of fiscal policy. This line of research typically finds that in response to such a shock to government spending, GDP increases whereas private consumption and real wages fall (Ramey and Shapiro 1998, Edelberg, Eichenbaum, and Fisher 1999, Burnside, Eichenbaum, and Fisher 2004). On the other hand, structural vector autoregressions (SVARs) usually achieve identification by assuming that government spending is predetermined within the quarter and government revenue does not respond to macroeconomic developments in the same quarter except for exogenous automatic stabilizers (Blanchard and Perotti 2002). This strand of the literature finds that private consumption, similar

[1] The narrative approach goes back to Romer and Romer (1989) in the area of monetary policy. A recent paper by Romer and Romer (2010) employs the narrative approach for tax changes.

to GDP, usually increases after a shock to government spending. Those results have been confirmed and extended in the papers by Perotti (2005, 2008), for example.[2]

These contrasting empirical findings have important implications for our view of the macroeconomy. Standard macroeconomic models focusing on fiscal policy, such as the neoclassical model of Baxter and King (1993) but also most New-Keynesian variants (for example, Linnemann and Schabert 2003), have an unambiguous prediction concerning the response of private consumption to a shock to government spending. Whereas output is expected to increase in response to such a shock, consumption should fall. The central reason for the latter dynamic response in those models is that government expenditure (financed by lump-sum taxes) constitute a withdrawal of resources from the economy, which in turn do not substitute or complement private consumption nor contribute to productivity. The resulting adverse wealth effect drives the negative consumption response. In contrast, Galí, López-Salido, and Vallés (2007) construct a New-Keynesian model with a positive consumption response, in order to reconcile current business cycle models with the empirical findings of the SVAR literature. Galí, López-Salido, and Vallés (2007) make clear, however, that many very special conditions have to be fulfilled for the model to be able to generate a positive response of private consumption. In particular, sticky prices and "rule-of-thumb" consumers drive the result.[3] Empirical findings therefore shape our modeling and understanding of the economy. Unfortunately, however, the different methods employed do not yield consistent results.

In an important contribution, Ramey (2009) aims at explaining the difference between the results of the two empirical approaches. She argues that VAR techniques miss the fact that major changes in government spending, such as expenditure related to wars, are usually anticipated. Within a standard model, it is easy to show that

[2]More empirical evidence with respect to European countries is provided by Biau and Girard (2005) for France, Giordano, Momigliano, Neri, and Perotti (2007) for Italy, de Castro and de Cos (2008) for Spain, and Tenhofen, Wolff, and Heppke-Falk (2010) for Germany. A different identification procedure was proposed by Fatás and Mihov (2001) and Mountford and Uhlig (2009), who also document a positive consumption response.

[3]An earlier contribution featuring a positive consumption response is Devereux, Head, and Lapham (1996), for instance. In this paper, consumption only increases if returns to specialization are sufficiently high.

missing the point of anticipation will result in a positive response of consumption to a shock to government spending, as consumption following the initial drop increases with investment. In support of her hypothesis that shocks are indeed anticipated, Ramey (2009) documents that the war dummy shocks Granger-cause the VAR shocks, but not vice versa.

These problems fit into the more general discussion on when it is possible to relate the innovations recovered by a VAR to the shocks of a particular economic model. Early contributions in this regard are Hansen and Sargent (1980, 1991), Townsend (1983), Quah (1990), and Lippi and Reichlin (1993, 1994), with a recent reminder of these problems to the profession in Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson (2007). An application of these insights to fiscal policy anticipation, in particular concerning tax changes, with a thorough discussion of the related issues can be found in Leeper, Walker, and Yang (2009). This literature centers on the fundamental problem that in certain setups the information sets of the private agents and the econometrician are misaligned. In the case of fiscal policy anticipation, this means that private agents in addition to the variables observed by the econometrician know about the fiscal policy shocks occurring in future periods and act immediately on this information. The econometrician, on the other hand, only observing variables up to the current period, does not possess this information. On a more technical note, (fiscal) foresight in a generic dynamic stochastic general equilibrium (DSGE) model may introduce a non-invertible moving-average (MA) component into the equilibrium process. In this case, the shocks identified by a VAR using only current and past endogenous variables do not match the shocks of the economic model. As a result, standard tools based on VARs, like impulse response functions or variance decompositions, can yield incorrect inferences.

We contribute to the empirical literature on the effects of fiscal policy by explicitly modeling anticipation in an SVAR framework. Our approach is designed to align the information sets of the econometrician and the private agents. Thereby we are able to avoid the problems encountered by standard VARs in settings featuring fiscal policy anticipation. In particular, we are able to exactly capture a situation, where private agents perfectly know fiscal shocks one period in advance. While our method is not general in the sense of being applicable in the presence of all possible (and in practice unknown) kinds of information flows, the findings of a simulation exercise support the approach. In particular, this exercise indicates that the methodology is robust to situations with a potentially different information structure. In order to document the validity of our method, we simulate data from a theoretical model with fiscal foresight, where we demonstrate that the equilibrium process features a non-invertible MA component by using methods recently developed by Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson (2007). Despite having both anticipated and unanticipated fiscal shocks in the model, so that private agents only have *imperfect* foresight, our approach correctly captures the dynamics within a VAR framework, while a standard VAR does not deliver the negative consumption response of the theoretical model.

In a next step, we apply our methodology to real life data to investigate the effects of anticipated fiscal policy on private *consumption*. As Ramey (2009) argues, fiscal policy anticipation could have dramatic consequences by changing the sign of the consumption response. Our findings indeed highlight the importance of taking into account fiscal foresight in empirical work. We show that it is crucial to distinguish those subcomponents of total government spending, which might have different effects on the macroeconomy. In this regard, we take advantage of the flexibility of the econometric approach. Motivated by economic theory and in line with previous studies, we consider government defense and non-defense expenditure.[4] This allows us to reconcile the results of the narrative and SVAR approaches mentioned above and qualify recent findings in the literature.

We find that when taking into account anticipation issues private consumption significantly *decreases* on impact and in subsequent periods in response to a shock to government defense expenditure, exactly in line with Ramey's (2009) findings using the narrative approach. When considering shocks to non-defense spending, on the other hand, consumption *increases* significantly on impact and in the following peri-

[4]While Blanchard and Perotti (2002) have a short subsection where they distinguish defense and non-defense expenditure, they only consider the response of output and do not take into account anticipation issues. Perotti (2008) also distinguishes defense and non-defense spending shocks in one of his SVAR specifications. Again, he does not allow for fiscal policy anticipation, which is the main focus of our investigation, where we show the importance of taking into account those issues.

ods in our expectation augmented VAR. In contrast, the corresponding responses in a standard VAR à la Blanchard and Perotti (2002) are quite weak and mostly insignificant. This highlights the importance of taking into account anticipation issues and is in line with Ramey's (2009) general argument that standard VAR techniques fail to allow for fiscal foresight thereby invalidating the structural analysis.

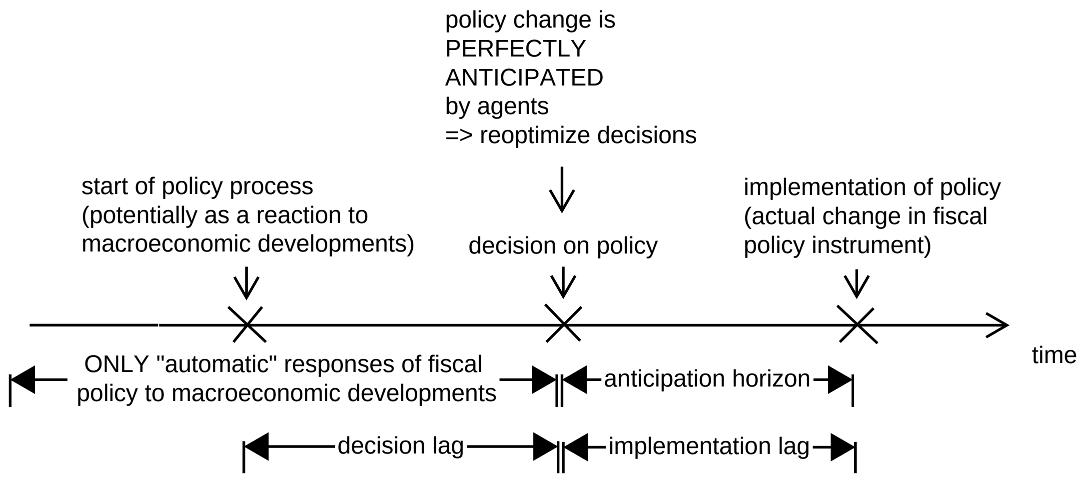
Furthermore, the responses reported for the expectation augmented VAR are in line with central predictions of standard macroeconomic models. In those settings, less productive defense expenditure lead to a decrease in consumption while other, potentially more productive expenditure have the opposite effect. If we do not separate different expenditure components but use total government spending, we do not obtain clear-cut results, as we lump together spending items with different macroeconomic effects. Our findings are robust to adding real GDP and/or a short-term interest rate to the specification as well as to changes in the exogenous elasticities needed to identify the SVAR.

The remainder of Chapter 3 is structured as follows. The next section develops the expectation augmented VAR, while Section 3.3 presents estimation results based on model-generated data. Section 3.4 presents the findings of the empirical investigation with a particular focus on government defense and non-defense expenditure. Section 3.5 checks robustness and, finally, the last section concludes.

3.2 An expectation augmented VAR

In order to explicitly take into account perfectly anticipated fiscal policy, we develop a new empirical approach. It is based on the framework put forward by Blanchard and Perotti (2002), which constitutes a well established SVAR methodology focusing on fiscal policy. Their basic idea is to exploit fiscal policy decision lags to identify structural shocks. In particular, the authors argue that as governments cannot react in the short run, e.g., within the same quarter, to changes in the macroeconomic environment, reactions of fiscal policy to current developments only result from so-called "automatic" responses. However, apart from *decision lags*, policymaking is also characterized by *implementation lags*. After a decision on a spending increase or tax cut, for instance, has been made, it takes time for the public authorities to imple-

Figure 3.1: Sequence of events



ment those measures. As a result, even though there has been no actual adjustment of the respective policy instrument yet, private agents already know that there will be a change in fiscal policy, i.e., they anticipate fiscal policy actions, and act immediately on this information. Not taking account of those implementation lags could invalidate the analysis due to the potential misalignment of the information sets of the private agents and the econometrician. Such a misalignment arises particularly in standard setups, where the econometrician uses data only up to the current period and neglects information on future fiscal shocks. Figure 3.1 summarizes graphically the aforementioned ideas by means of a timeline and illustrates, in particular, the concepts of decision and implementation lags.

Blanchard and Perotti (2002) address anticipation issues by including expectations of future fiscal policy variables in their model. In particular, they assume that agents perfectly know fiscal policy shocks one period in advance and are able to react to it. Thus, the aforementioned expectations are taken with respect to an information set which includes next period's fiscal shocks. Impulse responses to anticipated fiscal shocks are derived by simulating the system under rational expectations. They only consider the response of output, however, which is weaker but still positive. In particular, they do not report consumption responses, where anticipation effects could result in a different sign of the response as argued by Ramey (2009). The weaker output effect, though, might be an indication of a negative consumption response.

To allow for anticipation by the private sector, we go beyond the standard SVAR of Blanchard and Perotti (2002) by explicitly modeling the process describing expectation formation within such a multivariate time series framework. Furthermore, a central contribution of this chapter is to investigate the relevance of anticipation effects for the dynamic response of private consumption to fiscal policy shocks. We emphasize in particular the importance (of the nature) of the particular spending category under consideration, e.g., productive vs. unproductive public expenditure.

We propose the following setup, based on a standard AB-model SVAR, but augmented with expectation terms and equations describing the formation of those expectations:

$$Y_t = G\widehat{F}_{t+1} + C(L)Y_{t-1} + U_t \tag{3.1}$$

$$AU_t = BV_t \tag{3.2}$$

$$\widehat{F}_{t+1} = D(L)Y_t + HV_{t+1}, \tag{3.3}$$

where $Y_t = [c_t \;\; g_t \;\; r_t]'$ is the vector of endogenous variables, $\widehat{F}_{t+1} = [\widehat{g}_{t+1} \;\; \widehat{r}_{t+1}]'$ denotes *next* period's expected fiscal variables which are described in more detail in the next paragraph, U_t is the vector of reduced form residuals, and $V_t = [v_t^c \;\; v_t^g \;\; v_t^r]'$ is the vector of structural shocks to be identified. Here c_t denotes real private consumption, g_t is real government expenditure, r_t denotes real government revenue, and v_t^i is the respective structural shock.

The important difference relative to a standard (S)VAR is the presence of $\widehat{g}_{t+1}$ and $\widehat{r}_{t+1}$ in the preceding equations, both in equation (3.1) and, in particular, via equation (3.3). These expressions, reflecting fiscal policy anticipation, denote the conditional expectation of the respective fiscal variable with respect to current and past endogenous variables as well as next period's fiscal shocks, i.e., $\widehat{g}_{t+1} = E(g_{t+1}|\Upsilon_t, v_{t+1}^g, v_{t+1}^r)$ and $\widehat{r}_{t+1} = E(r_{t+1}|\Upsilon_t, v_{t+1}^g, v_{t+1}^r)$, where $\Upsilon_t = [Y_t, Y_{t-1}, Y_{t-2}, \ldots]$. Accordingly, agents in the economy form expectations about the course of future fiscal policy on the basis of all information available to them. Besides the current and past realizations of the variables in the system, the agents know about the fiscal shocks occurring next period.

These fiscal shocks are known as fiscal policy actions require time to be implemented. Moreover, they are usually subject to a broad public discussion before their actual implementation making the information available to a very broad audience. In order to capture this special information structure, equation (3.3) is added to the system. It describes how model-consistent expectations with respect to future fiscal variables are formed.[5]

Those features of our approach are designed to align the information sets of the private agents and the econometrician. The goal is to avoid the problems encountered by standard VARs, when confronted with data generated from a process featuring a non-invertible moving-average component due to fiscal foresight. The setup is able to exactly capture a situation, where private agents have one-period perfect foresight with respect to fiscal shocks. Even though this is not a general approach applicable in the presence of all possible kinds of information flows, the findings of the subsequent simulation exercise support our method. It indicates that the methodology is robust to situations with a potentially different information structure. Moreover, it is easily applicable to different spending categories. Without much effort and in a readily reproducible way, we can go beyond defense spending, i.e., beyond the point for which studies using the narrative approach exist.

3.2.1 A simplified setting: the general idea of the approach

In order to describe the basic idea of the approach, we first consider a *simplified version* of the aforementioned model, in particular, a setup which does not exhibit lagged dependent variables. This framework, however, easily generalizes to the standard case including lags, which is discussed subsequently. The system can be partitioned into two parts: first, one set of equations representing the basic structure of the economy, and second, the remaining equations modeling the process describing expectation formation. In the general setup given above, the former part is described by equations (3.1) and (3.2), whereas the latter part is modeled by equation (3.3).

More specifically, the basic framework of the economy in the simplified setup is

[5] An extensive discussion of the issues related to those equations can be found in the next subsection.

given by the first three equations of the model, presented here in structural form:

$$c_t = \gamma_1 \widehat{g}_{t+1} + \gamma_2 \widehat{r}_{t+1} + \alpha_g^c g_t + \alpha_r^c r_t + v_t^c \tag{3.4}$$

$$g_t = \alpha_c^g c_t + v_t^g \tag{3.5}$$

$$r_t = \alpha_c^r c_t + \beta_g^r v_t^g + v_t^r. \tag{3.6}$$

In accordance with the idea of fiscal policy anticipation by the private sector and following Blanchard and Perotti (2002), the two expectation terms, $\widehat{g}_{t+1}$ and $\widehat{r}_{t+1}$, appear in the consumption equation. Furthermore, we have to assume a relative ordering of the fiscal variables. Here we act on the assumption that spending decisions come first, i.e., the structural revenue shock, v_t^r, does not enter the expenditure equation, whereas v_t^g enters the revenue equation.[6]

As indicated above, the remaining part of the model consists of equations modeling the process describing expectation formation, in the simple framework given by:[7]

$$\widehat{g}_{t+1} = E(g_{t+1}|\Upsilon_t, v_{t+1}^g, v_{t+1}^r) = \eta_g^{Eg} v_{t+1}^g + \eta_r^{Eg} v_{t+1}^r \tag{3.7}$$

$$\widehat{r}_{t+1} = E(r_{t+1}|\Upsilon_t, v_{t+1}^g, v_{t+1}^r) = \eta_g^{Er} v_{t+1}^g + \eta_r^{Er} v_{t+1}^r. \tag{3.8}$$

Even though a standard VAR also implicitly models expectation formation, here we have to augment the basic VAR equations with the expectation terms and expectational equations, since we have to deal with a special information structure. In particular, not only variables indexed up to time t are part of the information set with respect to time t, but it also contains future variables, i.e., shocks indexed $t + 1$. Accordingly, one-period anticipation of fiscal policy actions is reflected in the presence of v_{t+1}^g and v_{t+1}^r in the preceding equations.

Analogous expectation terms, however, do not appear in the fiscal equations, i.e., equations (3.5) and (3.6) in this simplified setting, and there are no separate expectational equations for the non-fiscal variables. That does not mean that the public sector does not form (rational) expectations about future developments in the economy. It just reflects the fact that the fiscal authority's information set with respect

[6] Note that since the model is presented in structural form, the coefficients $\alpha_g^c, \alpha_r^c, \alpha_c^g, \alpha_c^r$, and β_g^r are elements of the A and B matrices, respectively.

[7] In this simplified setup, due to the absence of lagged dependent variables, Υ_t is not relevant for expectation formation. In the general case, however, Υ_t does play a role.

to the private sector only includes variables indexed up to the current period.[8] It is hard to think of a case of *aggregate* implementation lags for the private sector, which would give rise to the anticipation of future private sector actions by the government, analogous to the setting of fiscal foresight described in this chapter. Consequently, we do not have to augment the fiscal equations by expectation terms and the system by corresponding expectational equations to accommodate such a setup.

Ultimately, we are interested in deriving impulse response functions with respect to perfectly anticipated fiscal policy shocks. Consequently, we have to obtain the corresponding MA-representation of the model. Concerning consumption, which is the main variable of interest, such a representation in this simplified setup results when using equations (3.5) - (3.8) in equation (3.4) and solving for c_t:

$$\begin{aligned} c_t &= \frac{1}{1-\alpha_g^c\alpha_c^g-\alpha_r^c\alpha_c^r}\Big[(\gamma_1\eta_g^{Eg}+\gamma_2\eta_g^{Er})v_{t+1}^g+(\gamma_1\eta_r^{Eg}+\gamma_2\eta_r^{Er})v_{t+1}^r \\ &\quad +(\alpha_g^c+\alpha_r^c\beta_g^r)v_t^g+\alpha_r^c v_t^r+v_t^c\Big]. \end{aligned} \tag{3.9}$$

Consequently, concerning government expenditure for example, the dynamic response of consumption results as

$$\frac{\partial c_t}{\partial v_{t+1}^g} = \frac{\gamma_1\eta_g^{Eg}+\gamma_2\eta_g^{Er}}{1-\alpha_g^c\alpha_c^g-\alpha_r^c\alpha_c^r} \tag{3.10}$$

$$\frac{\partial c_{t+1}}{\partial v_{t+1}^g} = \frac{\alpha_g^c+\alpha_r^c\beta_g^r}{1-\alpha_g^c\alpha_c^g-\alpha_r^c\alpha_c^r} \tag{3.11}$$

$$\frac{\partial c_{t+s}}{\partial v_{t+1}^g} = 0 \quad \forall s\geq 2. \tag{3.12}$$

Note that this is the response to *next period's* fiscal shock, which is, however, perfectly anticipated today. In particular, consumption at time t moves in response to the fiscal shock of period $t+1$.[9]

[8]As the private sector, the government of course does know its own fiscal shocks next period and its effects on *current* non-fiscal variables. This is reflected in the system by equation (3.4) in combination with the fiscal equations.

[9]Due to the absence of lagged dependent variables in this simplified setting, the dynamic response is zero for c_{t+s}, $\forall s\geq 2$. In the general framework, of course, this is typically not the case as indicated in the impulse responses presented below.

We would like to emphasize the rationale of our expectational equations (3.7) and (3.8). The purpose of those equations is to describe how model-consistent expectations with respect to future fiscal variables are formed. In this respect, we are not interested in the structural relations between the different variables and thus the structural coefficients, but rather in the expectation of the respective fiscal variable in the sense of an optimal forecast based on the structure of the economy and all information available to the agent at the respective point in time.

Due to the linear structure of the economy, we consider linear projections as forecasts, which are the (reduced form) conditional expectation in this kind of setting. Consequently, since the conditional expectation leads to the forecast with the smallest mean squared error, linear projections produce optimal forecasts in this sense in such an environment. What remains to be specified are the relevant variables on which to project. In this respect, we consider all information available to the agent, which at time t comprises Υ_t, v^g_{t+1}, and v^r_{t+1}. In particular, both future fiscal shocks are relevant variables to produce a forecast for *both* government expenditure and revenue despite the relative ordering assumption of the structural equations. This can be seen by leading those equations (3.5) and (3.6) by one period and taking expectations. The resulting expressions imply that to obtain a forecast of next period's respective fiscal variable, a forecast of future private consumption is necessary. To this, in turn, *both* expected future government expenditure as well as expected future government revenue and thus the corresponding shocks are relevant as indicated by equation (3.4). Intuitively, the two fiscal shocks are useful for estimating future private consumption, which in turn is relevant to forecasting the fiscal variables.

Moreover, in this simplified setting we can easily solve for $\widehat{g}_{t+1}$ and $\widehat{r}_{t+1}$ by leading equations (3.4) to (3.6) by one period, combination, and taking expectations with respect to the information available at time t, yielding:

$$\widehat{g}_{t+1} = \underbrace{\frac{1-\alpha_r^c\alpha_c^r+\alpha_c^g\alpha_r^c\beta_g^r}{1-\alpha_g^c\alpha_c^g-\alpha_r^c\alpha_c^r}}_{\eta_g^{Eg}} v_{t+1}^g + \underbrace{\frac{\alpha_c^g\alpha_r^c}{1-\alpha_g^c\alpha_c^g-\alpha_r^c\alpha_c^r}}_{\eta_r^{Eg}} v_{t+1}^r \tag{3.13}$$

$$\widehat{r}_{t+1} = \underbrace{\frac{\alpha_c^r\alpha_g^c+\beta_g^r-\alpha_g^c\alpha_c^g\beta_g^r}{1-\alpha_g^c\alpha_c^g-\alpha_r^c\alpha_c^r}}_{\eta_g^{Er}} v_{t+1}^g + \underbrace{\frac{1-\alpha_g^c\alpha_c^g}{1-\alpha_g^c\alpha_c^g-\alpha_r^c\alpha_c^r}}_{\eta_r^{Er}} v_{t+1}^r. \tag{3.14}$$

This demonstrates the consistency of the expectational equations with the equations describing the basic structure of the economy. In particular, the linear projection coefficients of equations (3.7) and (3.8) can be related to the structural coefficients of equations (3.4) to (3.6).

3.2.2 The general setting: estimating an expectation augmented VAR

After having discussed the basic idea of our approach in the simplified setting, we now turn to the *general case* and present the estimation procedure. Taking into account lagged dependent variables, the basic structure of the economy is given by the following set of equations:

$$\begin{aligned} c_t &= C_{11}(L)c_{t-1}+\gamma_1\widehat{g}_{t+1}+\alpha_g^c g_t+C_{12}(L)g_{t-1}+\gamma_2\widehat{r}_{t+1} \\ &\quad +\alpha_r^c r_t+C_{13}(L)r_{t-1}+v_t^c \end{aligned} \tag{3.15}$$

$$g_t = \alpha_{c1}^g c_t+\alpha_{c2}^g c_{t-1}+\widetilde{C}_{21}(L)c_{t-2}+C_{22}(L)g_{t-1}+C_{23}(L)r_{t-1}+v_t^g \tag{3.16}$$

$$\begin{aligned} r_t &= \alpha_{c1}^r c_t+\alpha_{c2}^r c_{t-1}+\widetilde{C}_{31}(L)c_{t-2}+C_{32}(L)g_{t-1}+C_{33}(L)r_{t-1} \\ &\quad +\beta_g^r v_t^g+v_t^r, \end{aligned} \tag{3.17}$$

where we pulled c_{t-1} out of the lagpolynomial, since we have to treat the corresponding coefficients separately due to the identification scheme of Blanchard and Perotti (2002).

The expectational equations in the general setup result as:

$$
\begin{aligned}
\widehat{g}_{t+1} &= E(g_{t+1}|\Upsilon_t, v^g_{t+1}, v^r_{t+1}) \\
&= D_{11}(L)c_t + D_{12}(L)g_t + D_{13}(L)r_t + \eta^{Eg}_g v^g_{t+1} + \eta^{Eg}_r v^r_{t+1} \qquad (3.18) \\
\widehat{r}_{t+1} &= E(r_{t+1}|\Upsilon_t, v^g_{t+1}, v^r_{t+1}) \\
&= D_{21}(L)c_t + D_{22}(L)g_t + D_{23}(L)r_t + \eta^{Er}_g v^g_{t+1} + \eta^{Er}_r v^r_{t+1}. \qquad (3.19)
\end{aligned}
$$

Estimation of this model basically proceeds in three steps.[10] First, we look at the fiscal equations (3.16) and (3.17). Here we start by exploiting the assumption concerning decision lags. In particular, in order to address endogeneity issues, we use exogenous consumption elasticities of government expenditure and revenue to compute adjusted real government direct expenditure and net revenue.[11] Furthermore, we not only have to assume that there is no fiscal policy discretionary response to consumption developments within the quarter but also no response to such developments in the previous quarter. This indicates a tradeoff inherent in this method. On the one hand, we are able to incorporate fiscal foresight in the benchmark fiscal VAR model of Blanchard and Perotti (2002), but on the other we are constrained by the assumptions on which this approach is based. In particular, the maximum anticipation horizon we can implement depends on the number of periods we are willing to assume that fiscal policy is not able to discretionarily respond to macroeconomic developments. This step leads to the following setup:

$$
\begin{aligned}
g^A_t \equiv g_t - \alpha^g_{c1}c_t - \alpha^g_{c2}c_{t-1} &= \widetilde{C}_{21}(L)c_{t-2} + C_{22}(L)g_{t-1} + C_{23}(L)r_{t-1} + v^g_t \qquad (3.20) \\
r^A_t \equiv r_t - \alpha^r_{c1}c_t - \alpha^r_{c2}c_{t-1} &= \widetilde{C}_{31}(L)c_{t-2} + C_{32}(L)g_{t-1} + C_{33}(L)r_{t-1} \\
&\quad + \beta^r_g v^g_t + v^r_t. \qquad (3.21)
\end{aligned}
$$

[10]Here our focus is on the aspect of anticipation. A more detailed description of the general estimation approach can be found in Blanchard and Perotti (2002) and Tenhofen, Wolff, and Heppke-Falk (2010).

[11]Blanchard and Perotti (2002) argue that fiscal policy decision making is a slow process, involving many agents in parliament, government, and civil society. As a result, reactions of fiscal policy to current developments only result from automatic responses. Those are defined by existing laws and regulations and can be taken into account by applying exogenous output or consumption elasticities. Adjusting government expenditure or revenue using these elasticities allows to obtain unbiased estimates of the structural coefficients and thus the structural fiscal policy shocks.

Subsequently, we recursively estimate the resulting equations by OLS to obtain the structural shocks to the respective fiscal variable, i.e., we first estimate equation (3.20) and obtain v_t^g, and then use this shock series as an additional regressor to estimate equation (3.21).

In the second step, we consider the equation modeling private consumption. We begin by rewriting equation (3.15) as follows:

$$\begin{aligned} c_t &= C_{11}(L)c_{t-1} + \gamma_1 g_{t+1} + \alpha_g^c g_t + C_{12}(L)g_{t-1} + \gamma_2 r_{t+1} \\ &\quad + \alpha_r^c r_t + C_{13}(L)r_{t-1} + \tilde{v}_t^c, \end{aligned} \tag{3.22}$$

where

$$g_{t+1} = E(g_{t+1}|\Upsilon_t, v_{t+1}^g, v_{t+1}^r) + u_{t+1}^g \tag{3.23}$$

$$r_{t+1} = E(r_{t+1}|\Upsilon_t, v_{t+1}^g, v_{t+1}^r) + u_{t+1}^r, \tag{3.24}$$

and consequently $\tilde{v}_t^c = v_t^c - \gamma_1 u_{t+1}^g - \gamma_2 u_{t+1}^r$. Subsequently, equation (3.22) is estimated by instrumental variables, in order to account for the correlation of the respective regressors and error term. Since current and next period's *adjusted* government expenditure and revenue are perfectly known at time t,[12] they are uncorrelated with the expectational errors in $\tilde{v}_t^c$. Furthermore, due to the adjustment procedure they are also uncorrelated with v_t^c, so that we can use g_{t+1}^A, g_t^A, r_{t+1}^A, and r_t^A as instruments to estimate γ_1, α_g^c, γ_2, and α_r^c.

Finally, in the third step, we look at the equations modeling expectations. Since, as mentioned above, with respect to these two equations we are only interested in forecasting and not in estimation of the structural parameters, it is sufficient to just plug equations (3.18) and (3.19) into equations (3.23) and (3.24), respectively, and estimate these by OLS, as OLS provides a consistent estimate of the linear projection coefficient.[13]

Following this procedure, we obtain all coefficients necessary to compute the structural impulse response functions. In particular, it is possible to derive the dynamic

[12]To see this, note that due to the informational assumptions, all variables on the utmost right hand side of equations (3.20) and (3.21), and the corresponding ones for g_{t+1}^A and r_{t+1}^A, are known as of time t.

[13]See, for example, Hamilton (1994, p. 76).

response to a perfectly anticipated fiscal policy shock. The actual computation of the impulse responses starts out from the VARMA-representation of the model which is obtained by combining the equations of the general system (3.1) - (3.3):

$$Y_t = R^{-1}K(L)Y_{t-1} + R^{-1}GHV_{t+1} + R^{-1}A^{-1}BV_t \, , \tag{3.25}$$

where $R = (I - GD_0)$ and D_0 is the first coefficient matrix of the lag polynomial $D(L)$, and $K(L) = G\widetilde{D}(L) + C(L)$, where $\widetilde{D}(L) = (D(L) - D_0)\, L^{-1}$. Owing to the assumption of fiscal policy anticipation, the only unusual aspect of this representation are the time indices of the moving average part. By defining $W_t \equiv V_{t+1}$, however, we arrive at a standard VARMA model with corresponding pure MA-representation, which can be used to compute impulse response functions.

3.3 Application to simulated data

In order to illustrate the ability of our approach to capture fiscal policy anticipation, we present an application to model-generated data. We consider a stylized theoretical model featuring fiscal foresight to assess whether the approach is able to address problems related to non-invertibility due to fiscal policy anticipation. In particular, we use a variation of the model of Ramey (2009), which is a standard neoclassical growth model, to simulate time series and subsequently use these artificial data to estimate both a standard VAR and an expectation augmented VAR to derive impulse response functions. A convenient feature of simulating data from a theoretical model is that we know the true impulse response functions in this setup. Consequently, by comparing the estimated impulse responses to the theoretical ones, we can check whether the two aforementioned VAR models are able to address anticipation effects.

Ramey (2009) presents a simple neoclassical growth model featuring government spending financed via nondistortionary taxes, where agents learn about changes in government expenditure before their actual realization. We take her setup as a starting point, but augment it with a few features to be able to apply Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson's (2007) invertibility condition.[14] As mentioned

[14] Our model is still relatively close to Ramey's (2009) original specification. In particular, in the two models the impulse responses which are at the center of our investigation, i.e., the ones with

in the introduction, fiscal foresight in a generic DSGE model may lead to an equilibrium process with a non-invertible MA component, posing substantial problems for standard VAR analysis.[15] These problems can be illustrated as follows: for each non-invertible process there exists an invertible one, featuring the same mean and autocovariance-generating function. This implies that these processes cannot be distinguished based on the first two moments, so that Gaussian likelihood or least-squares procedures, for instance, run into an identification problem. As a result, it is standard in the VAR literature to disregard all non-invertible representations and focus solely on the corresponding invertible process. This means, however, that the econometrician is only able to recover the fundamental innovations corresponding to the invertible representation of the process, whereas the true economic shocks might correspond to the non-fundamental innovations of a non-invertible process.[16] As a result, standard tools based on such VARs, like impulse response functions or variance decompositions, potentially yield incorrect inferences.

In order to detect whether non-invertibility is present in a given DSGE model, Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson (2007) derive a condition based on the state-space representation of the equilibrium process of an economic model:

$$x_{t+1} = Ax_t + Bw_{t+1} \tag{3.26}$$

$$y_{t+1} = Cx_t + Dw_{t+1}, \tag{3.27}$$

where x_t is a vector of (possibly unobserved) state variables, y_t is a vector of variables the econometrician observes, and w_t denotes the vector of economic shocks. If "the eigenvalues of $A - BD^{-1}C$ are strictly less than one in modulus,"[17] a standard VAR will be able to recover the true economic shocks, w_t. Note, however, to be able to

respect to a government spending shock, are quite similar.

[15] An MA process is called invertible, if all the roots of the corresponding characteristic equation are outside the unit circle.

[16] Please note that in this description, we use a relation between (non-)invertibility and (non-)fundamentalness which abstracts from the borderline case, when at least one root of the characteristic equation of the moving-average process is *on* the unit circle (and none inside). Then, the process is non-invertible but the innovations are said to be fundamental.

[17] CONDITION 1 in Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson (2007, p. 1022).

apply this condition, the matrix D has to be nonsingular. In particular, the matrix must be square, i.e., the number of variables observed by the econometrician has to equal the number of economic shocks. For many models, this will not be the case, and this prerequisite is not met in Ramey's (2009) original setup. Consequently, we add investment-specific technology shocks and an error in forecasting government expenditure to the model, to obtain a nonsingular matrix D.[18] The latter feature is particularly interesting for this exercise. It allows to vary the relative importance of anticipated vs. unanticipated shocks to government expenditure. In particular, the model is able to represent a setting where foresight is not perfect.

With respect to the economic environment of the model, preferences and technology are specified as follows: the representative household maximizes

$$U_0 = E_0 \left[\sum_{t=0}^{\infty} \beta^t \left(\log C_t + \psi_t \log L_t \right) \right], \tag{3.28}$$

where β is the household's discount factor, C_t is private consumption, and L_t denotes leisure. The production function of the representative firm is given by

$$Y_t = (Z_t N_t)^{1-\alpha} K_t^{\alpha}, \tag{3.29}$$

where Y_t is output, N_t denotes labor input, and K_t is the capital stock, which evolves according to

$$K_{t+1} = (1-\delta)K_t + X_t I_t. \tag{3.30}$$

In the latter equation, I_t denotes (gross) investment, X_t is the level of investment-specific technology, and δ is the rate of depreciation for capital.[19] The two resource constraints in this economy are given by

$$\begin{aligned} L_t + N_t &\leq 1 && (3.31)\\ C_t + I_t + G_t &\leq Y_t. && (3.32) \end{aligned}$$

[18]Going back to Greenwood, Hercowitz, and Huffman (1988), investment-specific technology shocks are considered to be a major source of economic growth as well as business cycle fluctuations. With respect to the former, see for example Greenwood, Hercowitz, and Krusell (1997), whereas the latter point is made, for instance, by Greenwood, Hercowitz, and Krusell (2000) and Fisher (2006).

[19]This way of introducing investment-specific technological change follows Fisher (2006).

The stochastic processes governing the shocks to technology, the marginal rate of substitution, and investment-specific technology are assumed to evolve according to

$$
\begin{aligned}
\log Z_t &= \rho_1 \log Z_{t-1} + e_t^z, \quad e_t^z \overset{iid}{\sim} (0, \sigma^2_{e^z}) && (3.33)\\
\log \psi_t &= \rho_2 \log \psi_{t-1} + e_t^\psi, \quad e_t^\psi \overset{iid}{\sim} (0, \sigma^2_{e^\psi}) && (3.34)\\
\log X_t &= \rho_3 \log X_{t-1} + e_t^x, \quad e_t^x \overset{iid}{\sim} (0, \sigma^2_{e^x}). && (3.35)
\end{aligned}
$$

Finally, the evolution of government spending, financed via non-distortionary taxes, is specified as follows:

$$
\begin{aligned}
\log G_t &= \log G_{t-j}^{F,j} + \log \mathcal{E}_t^G, \quad j > 0 && (3.36)\\
\log G_t^{F,j} &= d_1 \log G_{t-1}^{F,j} + d_2 \log G_{t-2}^{F,j} + d_3 \log G_{t-3}^{F,j} + e_t^{GF}, \quad e_t^{GF} \overset{iid}{\sim} (0, \sigma^2_{e^{GF}}) && (3.37)\\
\log \mathcal{E}_t^G &= d_1 \log \mathcal{E}_{t-1}^G + d_2 \log \mathcal{E}_{t-2}^G + d_3 \log \mathcal{E}_{t-3}^G + e_t^{\mathcal{E}G}, \quad e_t^{\mathcal{E}G} \overset{iid}{\sim} (0, \sigma^2_{e^{\mathcal{E}G}}), && (3.38)
\end{aligned}
$$

where G_t is *actual* government spending at time t, $G_t^{F,j}$ is the *j-period forecast* of government spending made at time t, and $\mathcal{E}_t^G$ is the error made in forecasting government expenditure. Alternatively and perhaps more intuitively, one can think of government expenditure as following an AR(3) process, where the error consists of an anticipated and an unanticipated part:

$$
\begin{aligned}
\log G_t &= d_1 \log G_{t-1} + d_2 \log G_{t-2} + d_3 \log G_{t-3} + e_t^G && (3.39)\\
e_t^G &= e_{t-j}^{GF} + e_t^{\mathcal{E}G}. && (3.40)
\end{aligned}
$$

Combining such a specification with the forecasting relation (3.36) and the process for the forecast error (3.38) yields equation (3.37). The anticipated part of the error is known j periods in advance. Consequently, the preceding equations imply j-period *imperfect* foresight with respect to government expenditure shocks. In the following exercise, j is set to 1, corresponding to the specification in our empirical application in the next section.[20] This setup is quite convenient in the sense that by varying the variances of the anticipated and unanticipated shock, e_t^{GF} and $e_t^{\mathcal{E}G}$, respectively, it is

[20]This is an additional slight deviation from Ramey's (2009) original model, where she introduces two periods of foresight. Our estimation approach could also accommodate such a setting, but we want to be consistent with the informational assumptions employed in our subsequent empirical investigation.

Table 3.1: Calibration

Symbol	Value	Symbol	Value	Symbol	Value	Symbol	Value
β	0.99	ρ_2	0.95	$\sigma_{e^{\psi}}$	0.008	σ_{e^m}	0.005
α	0.33	ρ_3	0.95	σ_{e^x}	0.012	d_1	1.4
δ	0.023	ρ_4	0.95	$\sigma_{e^{GF}}$	0.0275	d_2	-0.18
ρ_1	0.95	σ_{e^Z}	0.01	$\sigma_{e^{\varepsilon G}}$	0.005	d_3	-0.25

possible to vary the relative importance of the two shocks for government expenditure. As $\sigma^2_{e^{\varepsilon G}}$ tends to zero, we approach a case of j-period *perfect* foresight, whereas when $\sigma^2_{e^{GF}}$ goes to zero, fiscal foresight will vanish. Furthermore, Ramey (2009) introduces measurement error in the logarithm of output, governed by an AR(1) process with autocorrelation coefficient ρ_4 and variance $\sigma^2_{e^m}$.

With respect to the calibration of the model, the same parameters are chosen as in Ramey (2009), where one time period in the model corresponds to a quarter. The calibration of the stochastic process for investment-specific technology, which is not present in Ramey's (2009) original model, is taken from In and Yoon (2007). These authors estimate this process for quarterly data, following an approach introduced by Greenwood, Hercowitz, and Krusell (1997, 2000), where the latter use annual data. Furthermore, we distribute the variance of the government expenditure shock given by Ramey (2009) among the anticipated and unanticipated part. In our benchmark calibration, we choose the same value for the standard deviation of the forecast error with respect to government spending as for the standard deviation of the measurement error in output. All in all, the values chosen are standard and summarized in Table 3.1.

Based on this calibration, we compute the eigenvalues of the matrix mentioned in Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson's (2007) invertibility condition. In this way we can check, whether the equilibrium process of the model just presented features a non-invertible moving-average component. Indeed, two eigenvalues are larger than one in modulus, implying that a standard VAR will not be able to recover the true economic shocks from current and past endogenous variables.[21] Even

[21]For this model, the eigenvalues of the matrix $A - BD^{-1}C$ in modulus are as follows: 1.6245, 1.6245, 0.9977, 0.7442, 0.7442, 0, 0, 0, 0, 0, 0.

though we know that the economic shocks cannot be *exactly* recovered from the observed current and past endogenous variables used in a VAR, it is still possible that (a subset of) those shocks can be reconstructed with relatively high accuracy. This point is made by Sims and Zha (2006) and demonstrated for a particular DSGE model. Since we are primarily interested in impulse response functions, in the following we check the actual severity of the invertibility problem introduced by fiscal foresight by comparing the theoretical impulses responses to the estimated ones obtained from a standard VAR using Blanchard and Perotti's (2002) identification scheme. Furthermore, by computing the corresponding impulse responses using an expectation augmented VAR, we can examine whether our approach is able to align the information sets of the agents and econometrician and can cope with the more demanding informational setup introduced by anticipation of fiscal policy.

Taking the theoretical impulse responses as a reference point, we simulate 1000 sets of time series of 100 observations from the setup described above and subsequently employ these artificial data in the estimation of a standard VAR and an expectation augmented VAR. Since our main focus is on the consumption response to an anticipated government spending shock, we concentrate on bivariate VARs in consumption and *actual* government expenditure while solely plotting the impulse response for consumption with respect to a shock to the latter variable. In the standard VAR, we use a Cholesky decomposition to identify the structural shocks, where government spending is ordered first. In this simplified setting, this amounts to the identification scheme of Blanchard and Perotti (2002), where the consumption elasticity of government spending is assumed to be zero contemporaneously. Concerning the expectation augmented VAR, we proceed as described in the previous section. In both cases, we include a constant and four lags of the endogenous variables in the estimation.[22]

The results are presented in Figure 3.2.[23] Each graph plots the response of con-

[22]This follows the specification of Ramey (2009). In her paper, she performs a similar exercise, in order to stress the importance of timing in a VAR. In particular, she compares two recursively identified VARs, where in the first estimation she uses *actual* government expenditure, G_t, and in the second one the *forecast* of that variable, $G_t^{F,j}$.

[23]The corresponding results when only anticipated shocks are present in the economic model can be found in Figure 3.19 in the appendix. The dynamic responses are almost identical to the ones presented here, highlighting the robustness of the expectation augmented VAR to the joint presence

sumption to a one standard deviation anticipated or unanticipated shock to government expenditure over a horizon of 20 periods. In the theoretical model, the response to both of those shocks is qualitatively the same. Consequently, and since our main focus is on the issue of fiscal foresight, we just show the theoretical impulse response resulting from the model for the anticipated shock to government expenditure, displayed in the first graph of the figure. The remaining plots show the corresponding impulse response function for the standard and expectation augmented VAR, respectively. We present the median dynamic response as well as the 16th and 84th percentile obtained from the 1000 simulation runs, thus also plotting 68% confidence intervals.[24] The timeline is normalized in such a way that period 0 corresponds to the point in time when there is the actual change in government spending, potentially coinciding with an unanticipated shock to government expenditure. The starting point, however, is period -1, when in the theoretical model, which governs the data generating process, the news about an increase in government expenditure arrives. This corresponds to the anticipated government spending shock.[25]

In the theoretical model, even though government spending does not move until period 0, consumption reacts immediately upon arrival of the news, i.e., in period -1. Due to the negative wealth effect, consumption drops on impact followed by a slow increase. Such a response, however, does not result when estimating a standard VAR and employing the well-established identification approach of Blanchard and Perotti (2002). In particular note that this conclusion is unaltered if instead an unanticipated government expenditure shock is considered, since the dynamic response in the theoretical model is qualitatively the same for both of those shocks.[26] The

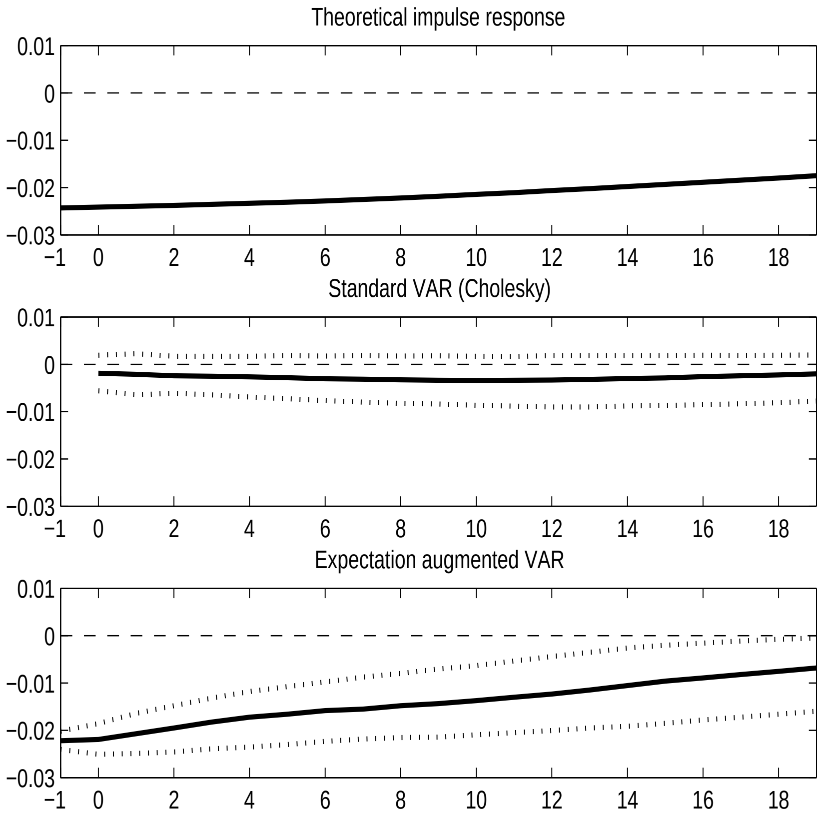
of anticipated and unanticipated shocks.

[24]In this regard, we follow the literature on the effects of fiscal policy shocks. See, for example, Blanchard and Perotti (2002) or Ramey (2009).

[25]The remaining theoretical impulse responses corresponding to an anticipated government expenditure shock are presented in Figure 3.20 in the appendix. Note in particular that all variables except government spending, of course, move immediately when the news about the shock arrives.

[26]The latter comparison might be more appropriate, as a standard VAR is only able to identify a government spending shock which *immediately* leads to a change in government expenditure. The arrival of the news in this setup coincides with the actual change in the fiscal variable. Consequently, the impulse response of consumption in this case starts at period 0.

Figure 3.2: Theoretical and VAR impulse responses



Notes: This figure shows the theoretical and VAR impulse responses of consumption to an anticipated one standard deviation shock to government spending as well as 68% confidence intervals. The economic model features both anticipated and unanticipated shocks.

consumption response for the standard VAR is insignificant over the entire horizon, where the median response is basically zero on impact and then somewhat decreases. Such a result is in line with typical findings of the VAR approach concerning the effects of fiscal policy shocks. In this model, problems related to non-invertibility due to fiscal policy anticipation do not seem to be only a theoretical feature of the data, but have important consequences for empirical research. Reflecting Ramey's (2009) argument, when using standard VAR techniques, structural shocks are not identified correctly, invalidating the structural analysis in a qualitatively and quantitatively important way.[27]

[27]As expected, these problems become less severe when the importance of unanticipated relative

The expectation augmented VAR, on the other hand, seems to be able to align the information sets of the private agents and the econometrician. It correctly captures the response of consumption to the anticipated government spending shock (third graph of Figure 3.2), even in the case when foresight is not perfect but obscured by unanticipated fiscal shocks. Not only the sign and subsequent qualitative movement of consumption corresponds to the true response derived from the model, but also the estimated impulse responses are very close to the theoretical one. The median of the estimated impact responses is -0.022 compared to -0.024 in the theoretical model. Moreover, the 68% confidence band includes the true impulse response up to period 5, and the theoretical response is just marginally outside the confidence interval after that.

Overall, the expectation augmented VAR thus correctly captures the effects of an anticipated fiscal shock. It addresses the more complex information structure of anticipated shocks within a VAR framework and delivers results closely matching the theoretical impulse responses. Opposed to standard approaches, it thus correctly takes into account the informational setup of the underlying data generating process, thereby rendering valid structural analysis feasible. In the next section, we apply our expectation augmented VAR to real-life data in order to investigate the impact of fiscal policy anticipation on the consumption response to a shock to total government expenditure and its subcomponents.

3.4 Empirical investigation

3.4.1 Data and elasticities

With respect to the data of our empirical investigation, real private consumption, real GDP, as well as real government direct expenditure, and real government net revenue for the US are defined as in Blanchard and Perotti (2002).[28] The series are

---

to anticipated government spending shocks is increased. Reducing the importance of fiscal foresight yields impulse responses for a standard VAR which are quite close to the theoretical ones.

[28] Figures 3.21 and 3.22 in the appendix plot the expenditure and tax to GDP ratio, respectively, as shown in Blanchard and Perotti (2002). The data are taken from the Bureau of Economic Analysis website (www.bea.gov).

seasonally adjusted, in per capita terms, and we take logs. The frequency of the employed time series is crucial for the identification approach. In order to exclude the possibility of discretionary fiscal policy actions within one time period, quarterly data are used. The system is estimated in levels including a constant, a time trend, and a dummy to account for the large tax cut in 1975:2. The sample starts in 1947:1 and runs up to 2009:2. The number of lags for the VAR is chosen to be three as suggested by the Akaike information criterion (AIC). With respect to the output and consumption elasticities, we follow Blanchard and Perotti (2002) and assume that there is no automatic response of government spending in the current and the previous quarter, and that the consumption elasticities of net revenue are $2.08 * 0.6468$ and $0.16 * 0.6468$ for time t and $t-1$, respectively, where 2.08 and 0.16 are the output elasticities and 0.6468 is the average share of consumption in GDP over the sample period. We perform various robustness checks concerning these elasticities without any substantial change in results.[29]

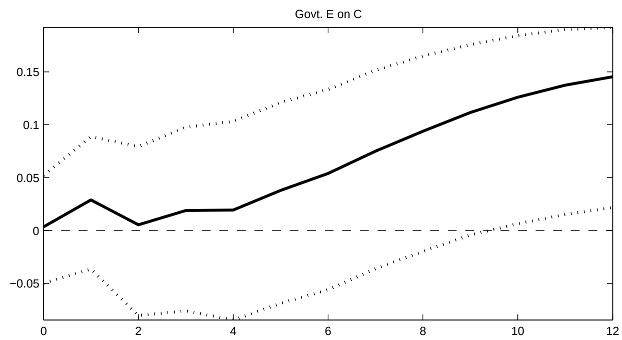
3.4.2 Total government expenditure

The starting point of our empirical investigation is a VAR à la Blanchard and Perotti (2002), featuring highly aggregated fiscal variables. In order to investigate Ramey's (2009) hypothesis that when fiscal policy anticipation is properly taken into account the positive consumption response typically found in VAR studies will turn negative, our VAR models include real private consumption, real direct expenditure, and real net revenue as endogenous variables. In Figures 3.3 and 3.4, we present the responses of private consumption to a shock to government spending derived from a standard VAR and an expectation augmented VAR, respectively.[30] Both of those responses are basically insignificant. In the model which is not taking into account anticipation,

[29] In particular, as do Blanchard and Perotti (2002), we also set the output elasticity of net revenue at $t-1$ to 0 and 0.5, and consequently the consumption elasticity to 0 and $0.5 * 0.6468$; see Section 3.5.

[30] We plot the point estimate of the impulse response function as well as 68% bootstrap confidence bands based on 5000 replications. We show 68% confidence intervals to be comparable to the literature, e.g., Blanchard and Perotti (2002) or Ramey (2009). Moreover, the corresponding impulse response functions with respect to a shock to government revenue for the current and following specifications can be found in the appendix.

Figure 3.3: Standard VAR: government expenditure

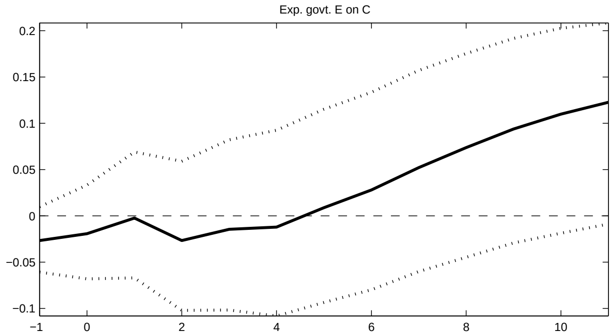


Notes: This figure plots the response of private consumption to a government expenditure shock, employing a standard SVAR model without anticipation. Sample: 1947q1-2009q2.

however, consumption turns significantly positive after the ninth quarter, while in the model with anticipation the consumption response is insignificant over the entire horizon considered. Of course, the insignificant response of the standard VAR stands somewhat in contrast to the paper by Blanchard and Perotti (2002). It should be noted, however, that we show the effect on private consumption, not GDP. Moreover, the respective sample periods under consideration are different. Whereas Blanchard and Perotti (2002) base their results on the sample 1960:1 – 1997:4, we not only use data also from the first decade of the new century but in addition include the 1950s. The latter period might be important, which we will discuss below. The main point, though, to be taken from this first set of results is that at least at this highly aggregated level, taking into account anticipation issues does not overturn the results obtained from a standard VAR.

When considering a variable like real government direct expenditure, however, we are lumping together the different subcomponents of this variable, which could have very different effects on private consumption. For example, expenditure on education might have a different effect on economic activity than defense expenditure. Indeed, the crucial feature of models à la Baxter and King (1993) to generate a negative consumption response to an increase in government expenditure is that the latter represents a withdrawal of resources from the economy, which does not substitute or complement private consumption nor contributes to production. Thus, even though government spending might affect utility, it does not influence private decisions except

Figure 3.4: Expectation augmented VAR: government expenditure



Notes: This figure plots the response of private consumption to an anticipated government expenditure shock, employing an expectation augmented VAR. The shock occurs in period 0 and is anticipated in period -1. Sample: 1947q1-2009q2.

through the budget constraint. However, Baxter and King (1993) show that once government expenditure enter the production function, for example, an increase in this kind of spending can have very expansionary effects depending on the productivity of the good. Consequently, already in the framework of this model, we might expect public expenditure on non-defense items like education, infrastructure, or law enforcement, which probably contribute to aggregate productivity, to induce an increase in private consumption. Public spending on national defense, on the other hand, lacking any complementarity or substitutability with respect to private consumption or any contribution to the private production process, might lead to the opposite response.[31] In fact, a change in defense spending is probably the closest approximation to the standard policy experiment conducted in models like Baxter and King (1993), i.e., a setup where in particular unproductive government expenditure are considered. But when we combine those defense and non-defense items in a single variable and study its dynamic effects on private consumption, the respective individual responses might cancel and lead to such weak results as reported above.

Consequently, in order to avoid this blurring of results, we focus in the following on

[31] Following the same reasoning, Turnovsky and Fisher (1995) in their theoretical investigation of the macroeconomic effects of subcomponents of government spending, distinguish “government consumption expenditure” and “government infrastructure expenditure.” The former includes items like national defense or social programs, whereas the latter consists of spending on roads, education, and job training, for example.

different subcomponents of government spending. In particular, we distinguish defense and non-defense expenditure. Considering defense spending is, of course, similar in spirit to Ramey's (2009) exercise of using dummy variables or other more sophisticated measures to capture large increases in government spending related to wars. Thus, we are able to check whether we can replicate Ramey's (2009) findings in an SVAR-based framework, when taking into account anticipation issues. Our method, however, is not confined to defense spending, so that we can also investigate the role of fiscal foresight when considering non-defense items of government expenditure.[32]

3.4.3 Defense expenditure

First, we look at public expenditure on national defense, which exhibit some noticeable features, particularly compared to non-defense spending. Major movements in total US government expenditure since the 1950s are related to defense spending. Figure 3.5 shows that while real non-defense expenditure per capita have increased substantially, the increase is rather smooth and follows GDP growth. In contrast, defense spending moved considerably and is rather volatile reflecting the different engagements of the USA in international wars. Most notably, the 1950s are characterized by a strong increase in defense expenditure, mainly due to the Korean War build-up. As depicted in Figure 3.6, this military engagement, along with increased defense spending due to the cold war, led to an increase of the ratio of defense expenditure to GDP from less than 7 percent in 1948 to almost 15 percent in 1952.[33] Moreover, the correlation be-

[32]We distinguish defense and non-defense spending and interpret them in terms of their respective degree of substitutability or complementarity or degree of productivity in the private production process in the spirit of Baxter and King (1993) and Turnovsky and Fisher (1995). Another strand of the literature highlights the importance of breaking total government spending down into purchases of goods and services and compensation of public employees (Rotemberg and Woodford 1992, Finn 1998, Forni, Monteforte, and Sessa 2009, Gomes 2009). Our focus, however, is on the different results of the narrative and SVAR approaches concerning the effects of fiscal policy and we therefore highlight defense and non-defense expenditure as subcomponents of total government spending.

[33]Concerning the choice of the sample period, we follow Ramey's (2008) argument and do not disregard the 1950s – including the Korean War – in the subsequent estimations. The Korean War, she forcefully argues, is an important source of variation in the data and should not be ignored. She notes that "[e]liminating the Korean War period from a study of the effects of government spending shocks makes as much sense as eliminating the 1990s from a study of the effects of information

Figure 3.5: Real per capita government spending

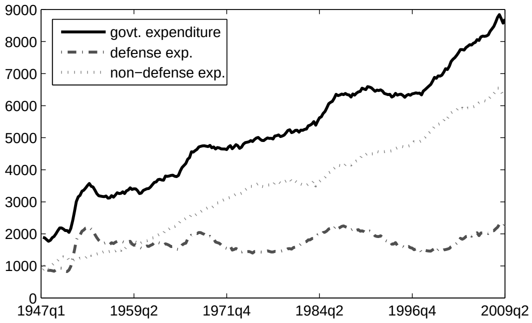
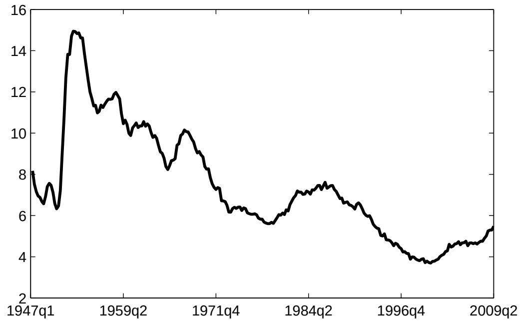


Figure 3.6: Ratio of defense expenditure to GDP (in percent)



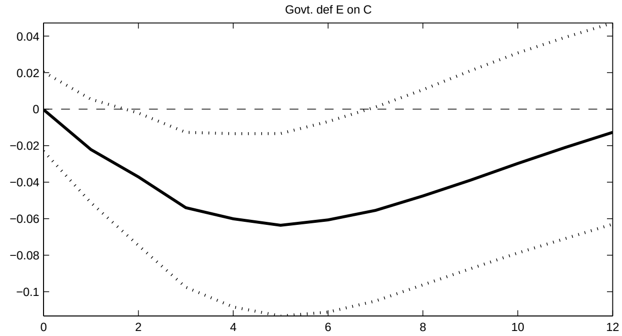
tween the detrended series of total government spending and defense spending is 0.81, whereas it is only 0.39 for total government expenditure and non-defense spending.

Turning to the estimation results, Figure 3.7 shows the response of consumption to a shock to defense spending derived from a standard fiscal VAR in the spirit of Blanchard and Perotti (2002). Compared to the dynamic response to a shock to total government spending, the point estimate shifts markedly downwards, in line with our expectations derived from economic theory. However, it is insignificant except for periods 2-6. In particular, the point estimate on impact is zero and not significant. A very different picture emerges, when the VAR is augmented with our methodology to account for anticipation effects, depicted in Figure 3.8. The dynamic response of consumption is significantly negative up to period 7. In particular, we find that consumption falls on impact and after further decreasing for a couple of quarters it slowly increases again. Consequently, even though defense spending does not move before period 0, the private agents respond immediately when they learn about the shock in period -1.

Thus, we can reconcile the narrative and SVAR approaches by replicating Ramey's (2009) findings in an SVAR-based framework. Our results are furthermore in line with Ramey's (2009) hypothesis that the difference between those two approaches arises because standard VAR techniques fail to allow for anticipation issues. In order

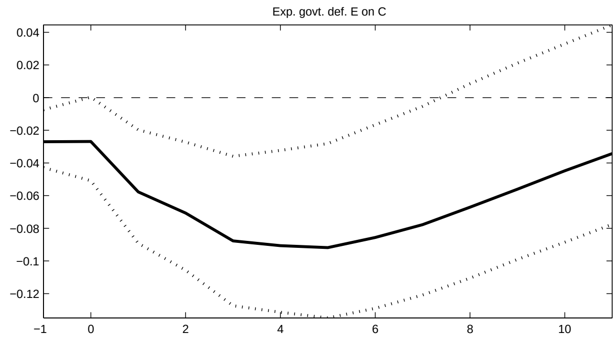
technology." Not surprisingly, when disregarding the important period 1947-1959 in the following estimation, we obtain weaker results (Figures 3.25 and 3.26 in the appendix).

Figure 3.7: Standard VAR: defense expenditure



Notes: This figure plots the response of private consumption to a government defense expenditure shock, employing a standard SVAR model without anticipation. Sample: 1947q1-2009q2.

Figure 3.8: Expectation augmented VAR: defense expenditure



Notes: This figure plots the response of private consumption to an anticipated government defense expenditure shock, employing an expectation augmented VAR. Sample: 1947q1-2009q2.

to see those effects clearly, however, it is necessary to look at more disaggregated variables to avoid interferences due to potentially different dynamic responses to other items of total government expenditure. All in all, our results underscore the need to appropriately take into account fiscal foresight in empirical research.

We can also look at these results from the viewpoint of the problems related to the misalignment of information sets of private agents and the econometrician due to fiscal policy anticipation. In those settings, even though we cannot obtain the true structural shocks from current and *past* endogenous variables, the system is invertible in current and *future* variables. Thus, as pointed out by Leeper, Walker, and Yang (2009), for example, it is possible to understand the two aforementioned approaches within the single framework of finding instruments for future variables. In this regard, it is encouraging that two different approaches of tackling those problems, in particular two different sets of instruments - “war dummies” on the one hand and future identified shocks to defense spending on the other - yield very similar results.

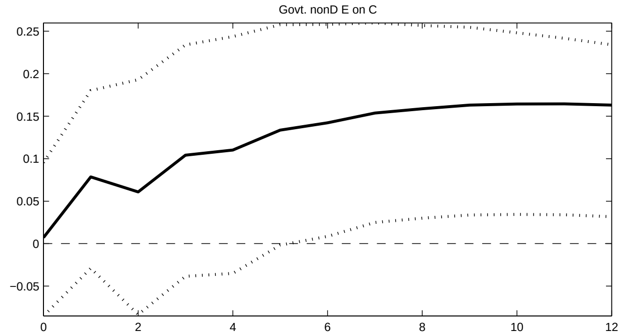
3.4.4 Non-defense expenditure

Next, we move to non-defense spending. As explained at the beginning of this section, we might expect private consumption to react differently to rather wasteful defense and potentially productive non-defense expenditure. Since private agents reoptimize and thus respond to new information as soon as it arrives regardless of whether it concerns defense or non-defense items of government spending, fiscal foresight is not confined to changes in the former variable. Thus, we move beyond Ramey’s (2009) exercise and take advantage of the flexibility of our econometric approach, and investigate the consequences of fiscal policy anticipation for dynamic responses to non-defense expenditure.

In Figure 3.9, we plot the impulse-response function of private consumption to a shock to government expenditure, where the latter do not include defense spending. It is derived from a three variable VAR estimated over the entire sample period without taking into account anticipation. In this standard framework, we find a significantly positive consumption response after 5 quarters. Thus, the dynamics move broadly in the direction implied by economic theory. The point estimate, however, is still basically zero on impact and insignificant, and it takes a couple of quarters for the response to move significantly into positive territory. As Figure 3.10 makes clear, extending the VAR to allow for anticipation of fiscal shocks yields a different picture. We now find a significantly positive consumption response already in period -1, when the increase in non-defense expenditure is anticipated. Furthermore, the response stays significantly positive over the entire horizon under consideration.

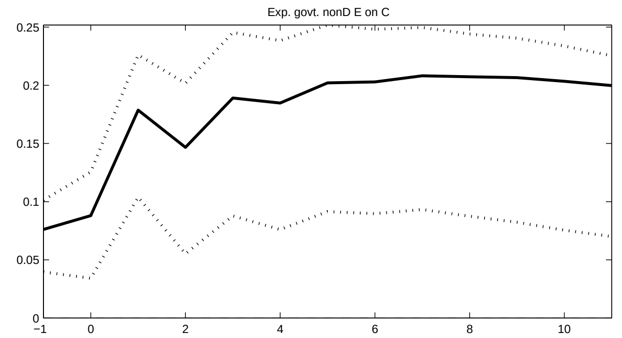
Analogous to the results obtained for defense spending, anticipation effects are also of empirical relevance when considering non-defense expenditure. This finding is in line with Ramey’s (2009) overall argument, even though we obtain a significant *increase* in private consumption. Thus, it is important to distinguish the potentially

Figure 3.9: Standard VAR: non-defense expenditure



Notes: This figure plots the response of private consumption to a government non-defense expenditure shock, employing a standard SVAR model without anticipation. Sample: 1947q1-2009q2.

Figure 3.10: Expectation augmented VAR: non-defense expenditure



Notes: This figure plots the response of private consumption to an anticipated government non-defense expenditure shock, employing an expectation augmented VAR. Sample: 1947q1-2009q2.

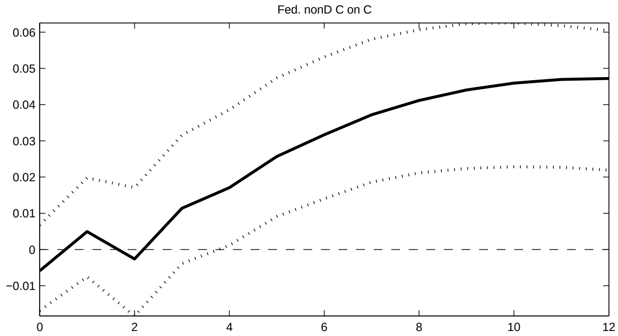
different dynamic responses to the separate subcomponents of total government expenditure.

An unambiguously positive consumption response would be expected when considering the model of Baxter and King (1993) for the case of productive government expenditure, for example.[34] Given the opposite findings for defense and non-defense expenditure, the effects of fiscal policy when lumping together those two items in one fiscal aggregate are likely to be weak.

As a final analysis of this section, we take up another point made by Ramey (2009).

[34]Of course, this result is also in line with the model of Galí, López-Salido, and Vallés (2007).

Figure 3.11: Standard VAR: federal non-defense expenditure



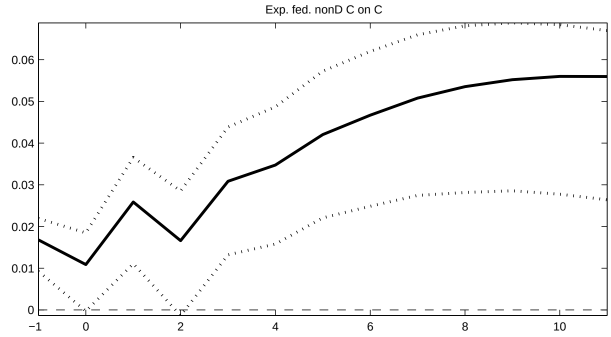
Notes: This figure plots the response of private consumption to a federal non-defense expenditure shock, employing a standard SVAR model without anticipation. Sample: 1947q1-2009q2.

She argues that aggregate VARs are not very good at capturing shocks to spending which is determined locally. Consequently, in order to make sure that our findings are not driven by the fact that large parts of non-defense expenditure are made by states and local authorities, we look at *federal* non-defense consumption spending.[35] As depicted in Figures 3.11 and 3.12, we find our previous results confirmed. In particular, the consumption response derived from an expectation augmented VAR is again significantly positive on impact and over the entire horizon. But also the dynamic response based on a standard VAR is very similar. These results suggest that the difference between defense and non-defense spending is not determined by the fact that large parts of non-defense spending are made by states and local authorities.

All in all, our findings highlight the importance of taking into account fiscal foresight when studying empirically the dynamic effects of changes in fiscal policy on economic activity. Our results are in line with Ramey's (2009) hypothesis that standard VARs fail to take into account anticipation issues and therefore yield incorrect inferences. Motivated by economic theory, we emphasize the need to look at different subcomponents of total government spending and show that they have different effects on the macroeconomy. Lumping together the different items in a single fiscal aggregate blurs the results. For defense spending, we are able to replicate Ramey's (2009) findings of a decrease in private consumption in an SVAR-based framework

[35]Please note that since state and local governments do not have expenditure on national defense, federal defense spending equals total defense spending.

Figure 3.12: Expectation augmented VAR: federal non-defense expenditure



Notes: This figure plots the response of private consumption to an anticipated federal non-defense expenditure shock, employing an expectation augmented VAR. Sample: 1947q1-2009q2.

and can thereby reconcile the narrative and SVAR approaches of studying the effects of fiscal policy. For non-defense spending, we also find an important role for fiscal policy anticipation, but in this case private consumption increases significantly. This result is exactly what would be expected when considering standard neoclassical or New-Keynesian models of fiscal policy for the case of productive public expenditure, for example.

Our findings, moreover, correspond to the results of the recent papers by Kriwoluzky (2009) and Mertens and Ravn (2009). These authors also study the effects of fiscal foresight on the dynamic responses to government expenditure shocks.[36] Neither paper, however, looks at subcomponents of total government spending. By distinguishing defense and non-defense spending, we can put their findings into perspective and also qualify the result in an earlier version of the paper on which this chapter is based, where we find a negative consumption response in an expectation augmented VAR (Tenhofen and Wolff 2007). For instance, similar to our finding for the consumption response to total government expenditure, Kriwoluzky (2009) also obtains a rather weak response in the first couple of quarters. Mertens and Ravn (2009), on the other hand, conclude based on their results that anticipation of fiscal policy does not

[36]The former employs sign restrictions derived from a DSGE model to identify the structural shocks of a vector MA (VMA) model estimated by likelihood methods. The latter consider a vector error-correction model (VECM) and use Blaschke matrices as suggested by Lippi and Reichlin (1993, 1994) to obtain non-fundamental innovations.

alter the positive effects of fiscal policy on consumption and output. Finally, from the viewpoint of the problems related to the misalignment of information sets due to fiscal foresight, we find encouraging that different approaches of tackling these problems, in particular different sets of instruments, yield basically the same results. In the next section, we turn to the robustness of our findings.

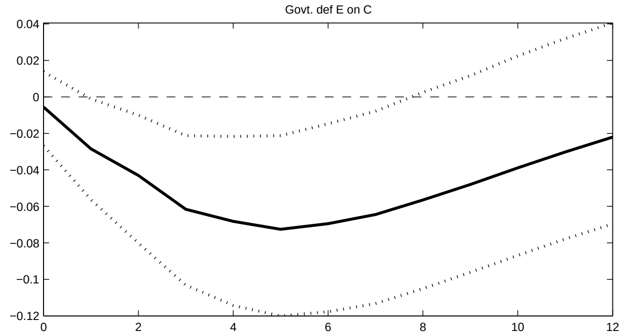
3.5 Robustness checks

First, we want to make sure that our results are not driven by the omission of other, potentially important macroeconomic variables. In particular, we consider adding measures of real output and/or a short-term interest rate to the specifications mentioned above.

With respect to the latter variable, while Blanchard and Perotti (2002) also do not control for short-term interest rates, follow-up papers by Perotti add such a variable to a standard fiscal SVAR. Since monetary policy is not orthogonal to fiscal policy, its inclusion might alter our results. We therefore extend our SVAR approach to also feature a short-term interest rate. In particular, following Giordano, Momigliano, Neri, and Perotti (2007) and Tenhofen, Wolff, and Heppke-Falk (2010), we assume a recursive ordering for the equations of the non-fiscal variables. Accordingly, whereas consumption is assumed not to react to the short-term interest rate contemporaneously, this is not true vice versa. This ordering assumption, reflecting the more sluggish nature of consumption compared to financial variables like interest rates, is common practice in the monetary VAR literature. Furthermore, when estimating the interest-rate equation, we have to add to the set of instruments the structural shock to consumption, v_t^c, obtained from the consumption equation, in order to get unbiased estimates. Apart from that, the additional equation for the interest rate also includes expectation terms of the fiscal variables, in order to be consistent with the assumption of fiscal policy anticipation.[37]

[37] For more details on the estimation when the block of non-fiscal variables includes more than one variable, see Giordano, Momigliano, Neri, and Perotti (2007) and Tenhofen, Wolff, and Heppke-Falk (2010).

Figure 3.13: Standard VAR: defense expenditure (incl. 3-month T-bill rate)



Notes: This figure plots the response of private consumption to a government defense expenditure shock, employing a standard SVAR model without anticipation. The VAR includes the 3-month T-bill rate. Sample: 1947q1-2009q2.

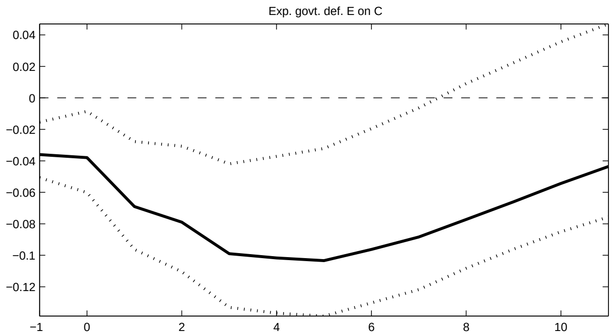
With respect to data, in our estimation we use the 3-month T-bill rate.[38] Concerning the (semi-)elasticities, we follow Perotti (2005) in assuming that government spending does not react to changes in the interest rate in the current and also in the previous quarter. Indeed, the government spending variable does not include interest payments. Regarding the impact on revenue, we also follow Perotti (2005) and assume no contemporaneous response, but also no response to movements in the interest rate in period $t-1$. However, we checked robustness of the results to changes in these elasticities. Our findings are not altered in substance and available from the authors.

Figures 3.13 and 3.14 show the results for a defense expenditure shock once the respective specification is extended to control for the 3-month T-bill rate. As in the benchmark case, we find consumption to fall on impact in the expectation augmented SVAR, while in the case of the standard SVAR it is insignificant on impact. Furthermore, the resulting impulse responses are quite similar to the ones arising in the corresponding three-variable benchmark case. Thus, the inclusion of an interest rate does not significantly alter the effects of government defense spending on private consumption.

Next, we consider the effects of including real GDP in addition to the 3-month T-bill rate and the three variables of our specification focusing on defense spending, i.e., real private consumption, real government defense expenditure, as well as real gov-

[38]The corresponding time series is taken from the FRED database of the Federal Reserve Bank of St. Louis.

Figure 3.14: Expectation augmented VAR: defense expenditure (incl. 3-month T-bill rate)

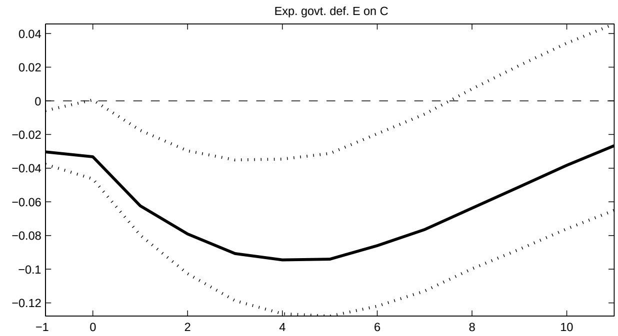


Notes: This figure plots the response of private consumption to an anticipated government defense expenditure shock, employing an expectation augmented VAR. The VAR includes the 3-month T-bill rate. Sample: 1947q1-2009q2.

ernment net revenue. GDP and private consumption are two closely linked variables. The SVAR approach up to now did not control for the developments of the former variable. It is therefore possible that our results are spuriously driven by the omission of this important determinant of private consumption as well as of government activity. We therefore extend the specification of the preceding paragraph to also control for real GDP per capita. This extension is analogous to the one just discussed, where we assume that output does not react contemporaneously to consumption and the short-term interest rate, whereas consumption does react to developments in output within the same period, but not to movements in the interest rate. The latter variable, in turn, is considered to be the least sluggish one among the non-fiscal variables, so that it is assumed to react to both output and consumption contemporaneously.[39] Whereas the assumption with respect to the interest rate is probably uncontroversial, the ordering of the other two variables might be less so. Consequently, in order to check the robustness of our findings, we changed the ordering of output and consumption in our estimation. However, this does not affect our results. As already indicated in Section 3.4, with respect to the output elasticities, we assume the same values as in Blanchard and Perotti (2002), which are furthermore in line with our assumptions

[39]Note that in the estimation of the consumption equation, we have to extend the set of instruments to include the structural shock to output, v_t^y. When estimating the interest-rate equation, we furthermore have to add the structural shock to consumption, v_t^c.

Figure 3.15: Expectation augmented VAR: defense expenditure (incl. GDP and 3-month T-bill rate)



Notes: This figure plots the response of private consumption to an anticipated government defense expenditure shock, employing an expectation augmented VAR. The VAR includes GDP and the 3-month T-bill rate. Sample: 1947q1-2009q2.

concerning the consumption elasticities.

Considering Figure 3.15, we indeed find, in line with standard economic theory as well as our previous results, that shocks to government defense expenditure lead to a decrease in private consumption in the expectation augmented VAR, even when controlling for output per capita, where the consumption response is also quantitatively of similar size. Thus, the inclusion of GDP does not affect our main results.[40]

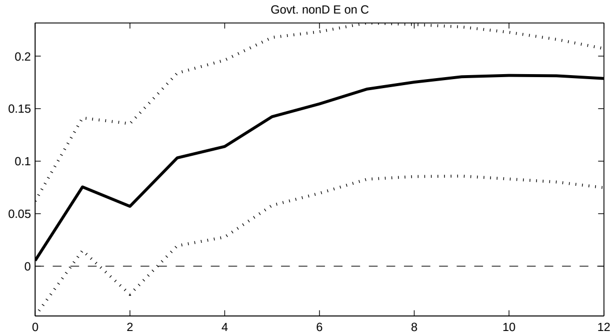
When looking at non-defense expenditure, we also find our main results confirmed (Figures 3.16 to 3.18).[41] The inclusion of a short-term interest rate or GDP does not alter the previous findings. Consumption increases, in particular on impact, in response to a non-defense spending shock in the expectation augmented VAR. In the standard VAR, on the other hand, consumption only increases after a couple of periods and the point estimate is basically zero on impact and insignificant.

Our final robustness check focuses on the elasticities. First, in our specification featuring defense expenditure, we set the elasticity of revenue to private consumption at $t-1$ to zero. Figure 3.37 in the appendix shows that the negative consumption response is unaffected. Increasing this elasticity to $(0.5 * 0.6468)$ yields Figure 3.38,

[40]The corresponding graph for the standard fiscal VAR is also basically unchanged and given in the appendix (Figure 3.35).

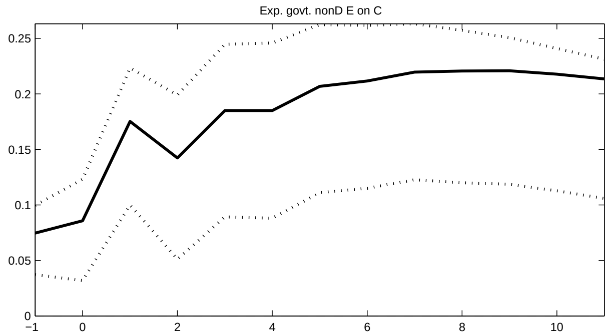
[41]The graph concerning the standard VAR when including real GDP as well as a short-term interest rate is again given in the appendix (Figure 3.36).

Figure 3.16: Standard VAR: non-defense expenditure (incl. 3-month T-bill rate)



Notes: This figure plots the response of private consumption to a government non-defense expenditure shock, employing a standard SVAR model without anticipation. The VAR includes the 3-month T-bill rate. Sample: 1947q1-2009q2.

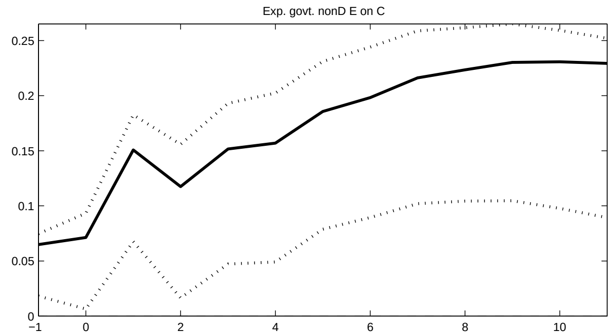
Figure 3.17: Expectation augmented VAR: non-defense expenditure (incl. 3-month T-bill rate)



Notes: This figure plots the response of private consumption to an anticipated government non-defense expenditure shock, employing an expectation augmented VAR. The VAR includes the 3-month T-bill rate. Sample: 1947q1-2009q2.

where the response to a shock to defense spending also remains negative and significant. Next, when doing the same exercise based on our specification featuring non-defense expenditure, we also find our previous results confirmed. Regardless whether we use an elasticity of revenue to private consumption at $t-1$ of zero or $(0.5*0.6468)$, private consumption increases significantly on impact and over the entire horizon considered (Figures 3.39 and 3.40 in the appendix). Furthermore, using the tax revenue elasticity to GDP as the elasticity of tax revenue to consumption does not change the results (Figures 3.41 to 3.44 in the appendix). All in all, even when adding macroeconomic variables to the system or when changing the exogenous elasticities needed

Figure 3.18: Expectation augmented VAR: non-defense expenditure (incl. GDP and 3-month T-bill rate)



Notes: This figure plots the response of private consumption to an anticipated government non-defense expenditure shock, employing an expectation augmented VAR. The VAR includes GDP and the 3-month T-bill rate. Sample: 1947q1-2009q2.

to identify the SVAR, we clearly find our previous findings confirmed.

3.6 Conclusion

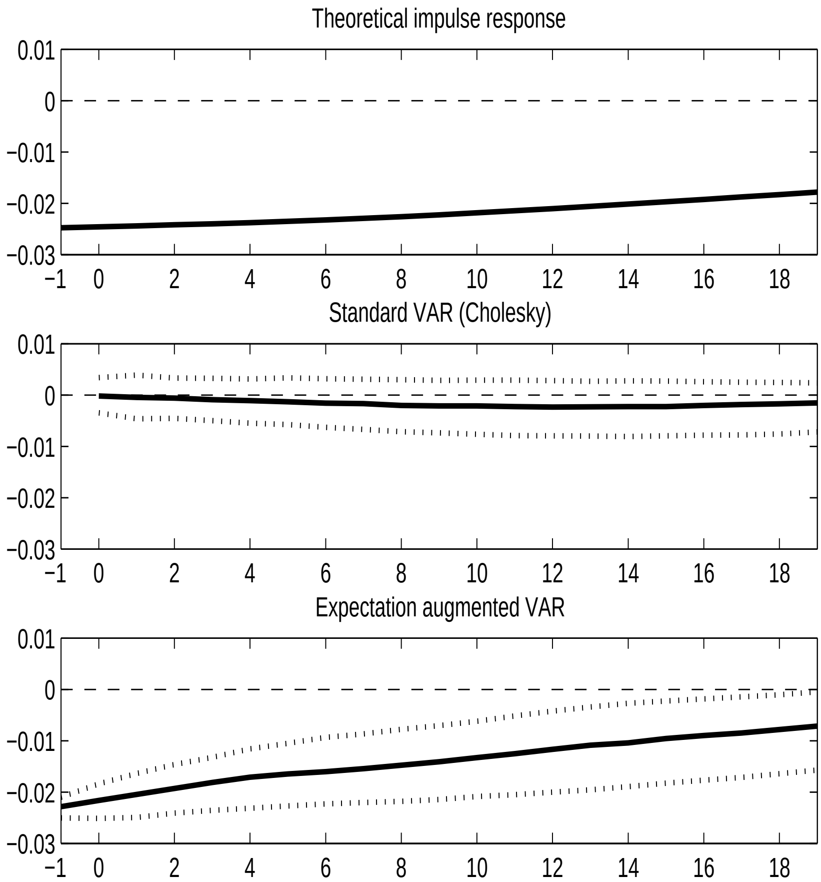
How does private consumption react to public expenditure shocks? In this chapter, we develop an SVAR approach which allows for anticipation of fiscal policy shocks. Our goal is to avoid problems encountered by standard VARs and align the information sets of the private agents and the econometrician, which makes valid structural analysis feasible. We are able to exactly capture a situation, where private agents perfectly know fiscal shocks one period in advance. Even though our method is not general in the sense of being applicable in the presence of all possible kinds of information flows, the findings of a simulation exercise document that our approach is robust to situations with a potentially different information structure. When confronted with data simulated from a model featuring fiscal foresight and an equilibrium process with a non-invertible MA component, our method correctly captures macroeconomic dynamics. In contrast, standard VARs do not capture the dynamics properly. This performance is even more noticeable as the economic model under consideration features both anticipated and unanticipated fiscal shocks, so that private agents only have imperfect foresight. This makes it more difficult for our method to trace out the individual dynamic effects.

The empirical investigation highlights the importance of taking into account anticipation issues in fiscal VAR studies. In contrast to the rather weak and mostly insignificant consumption responses in a standard VAR in the spirit of Blanchard and Perotti (2002), our expectation augmented VAR yields unambiguous responses. In this regard, we show that it is important to distinguish subcomponents of total government spending, which might have different effects on the macroeconomy. This focus on more disaggregated variables is facilitated by the flexibility of our econometric approach and allows us to qualify recent findings in the literature. Considering total government expenditure, on the other hand, does not yield clear-cut results. This is due to the fact that when considering this aggregate, we lump together subcomponents with potentially different effects on the macroeconomy.

The response of private consumption to a shock to defense spending in the expectation augmented VAR corresponds to Ramey's (2009) finding of a negative consumption response. Thus, we are able to reconcile the narrative and SVAR approaches of studying the effects of fiscal policy. Non-defense spending, on the other hand, yields a significantly positive response of private consumption. All in all, our findings are in line with Ramey's (2009) overall argument that standard VAR techniques fail to allow for anticipation issues which invalidates the structural analysis. Moreover, the results reported for the expectation augmented VAR are what would be expected when considering standard macroeconomic models for different degrees of productivity of public expenditure. Defense and non-defense spending are very different in nature, where the latter has a more productive character.

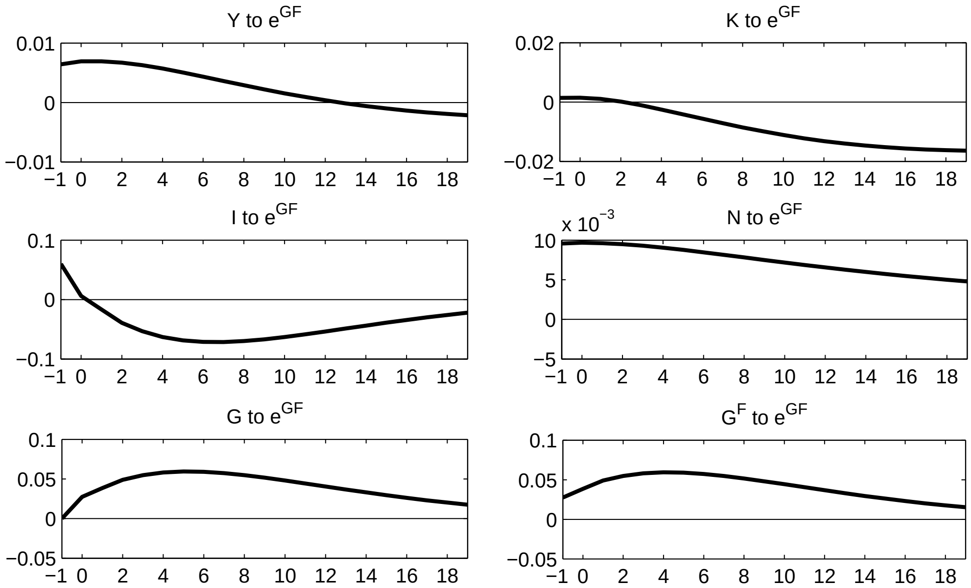
Appendix to Chapter 3

Figure 3.19: Theoretical and VAR impulse responses (only anticipated shocks)



Notes: This figure shows the theoretical and VAR impulse responses of consumption to an anticipated one standard deviation shock to government spending as well as 68% confidence intervals. The economic model features only anticipated shocks.

Figure 3.20: Theoretical impulse responses



Notes: This figure shows the theoretical impulse responses to a one standard deviation anticipated shock to government spending resulting from the economic model.

Figure 3.21: Ratio of government direct expenditure to GDP (in %)

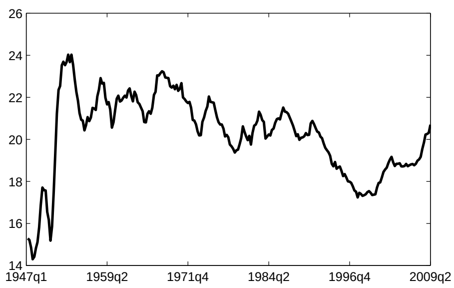


Figure 3.22: Ratio of government net revenue to GDP (in %)

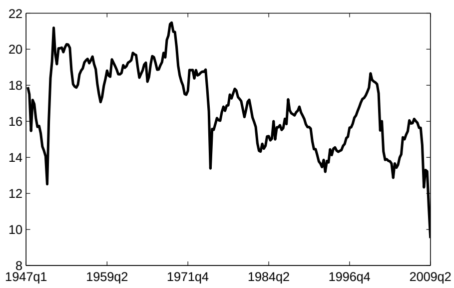
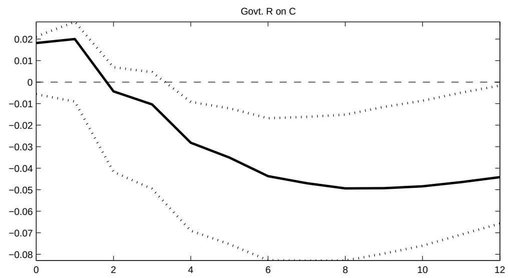
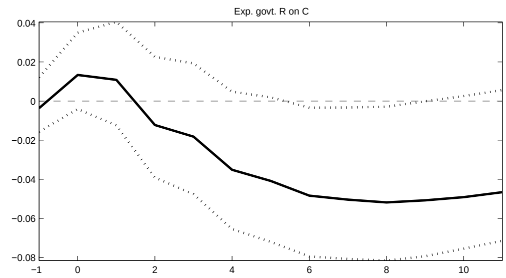


Figure 3.23: Standard VAR: government revenue



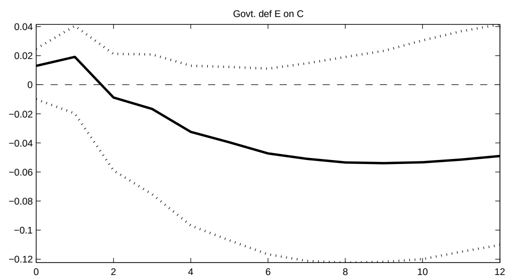
Notes: Response of private consumption to a government revenue shock, employing a standard SVAR model without anticipation. Sample: 1947q1-2009q2.

Figure 3.24: Expectation augmented VAR: government revenue



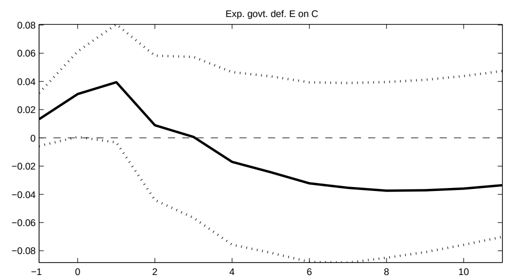
Notes: Response of private consumption to an anticipated government revenue shock, employing an expectation augmented VAR. Sample: 1947q1-2009q2.

Figure 3.25: Standard VAR: defense expenditure (ex 1950s)



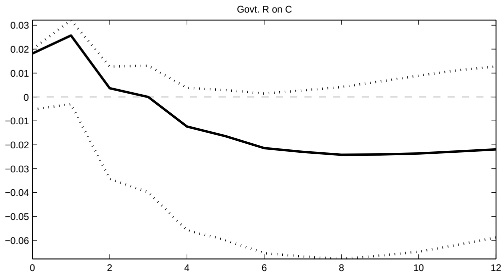
Notes: Response of private consumption to a government defense expenditure shock, employing a standard SVAR model without anticipation. Sample: 1960q1-2009q2.

Figure 3.26: Expectation augmented VAR: defense expenditure (ex 1950s)



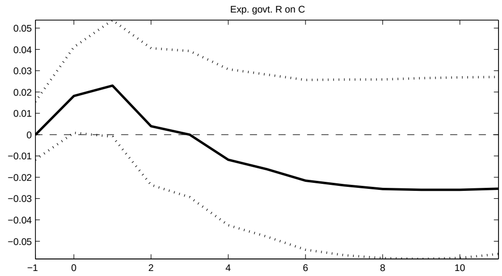
Notes: Response of private consumption to an anticipated government defense expenditure shock, employing an expectation augmented VAR. Sample: 1960q1-2009q2.

Figure 3.27: Standard VAR: government revenue (incl. defense expenditure)



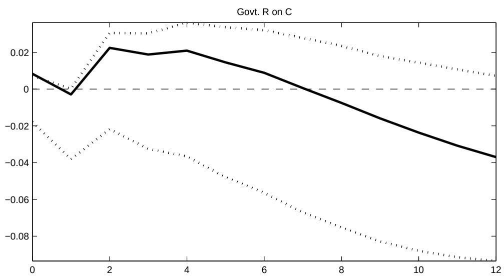
Notes: Response of private consumption to a government revenue shock, employing a standard SVAR model without anticipation featuring defense expenditure. Sample: 1947q1-2009q2.

Figure 3.28: Expectation augmented VAR: government revenue (incl. defense expenditure)



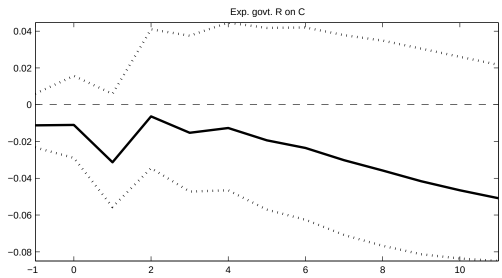
Notes: Response of private consumption to an anticipated government revenue shock, employing an expectation augmented VAR featuring defense expenditure. Sample: 1947q1-2009q2.

Figure 3.29: Standard VAR: government revenue (incl. defense expenditure, ex 1950s)



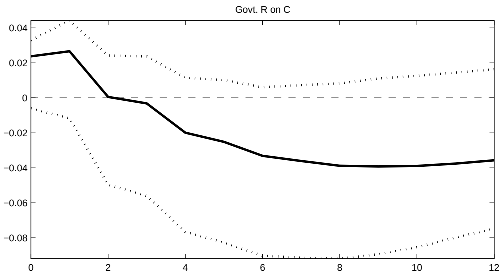
Notes: Response of private consumption to a government revenue shock, employing a standard SVAR model without anticipation featuring defense expenditure. Sample: 1960q1-2009q2.

Figure 3.30: Expectation augmented VAR: government revenue (incl. defense expenditure, ex 1950s)



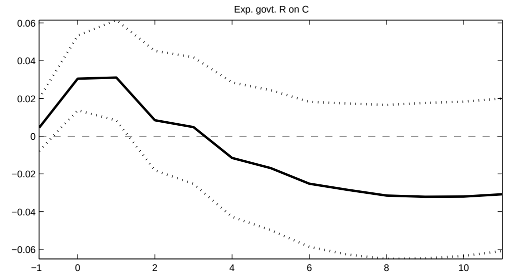
Notes: Response of private consumption to an anticipated government revenue shock, employing an expectation augmented VAR featuring defense expenditure. Sample: 1960q1-2009q2.

Figure 3.31: Standard VAR: government revenue (incl. non-defense expenditure)



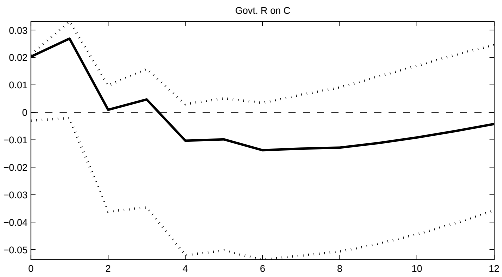
Notes: Response of private consumption to a government revenue shock, employing a standard SVAR model without anticipation featuring non-defense expenditure. Sample: 1947q1-2009q2.

Figure 3.32: Expectation augmented VAR: government revenue (incl. non-defense expenditure)



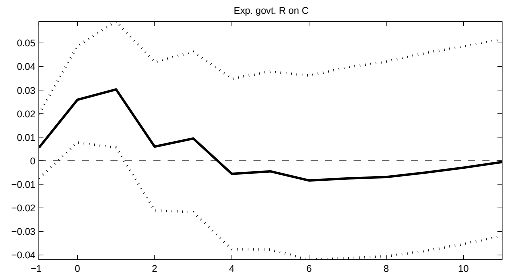
Notes: Response of private consumption to an anticipated government revenue shock, employing an expectation augmented VAR featuring non-defense expenditure. Sample: 1947q1-2009q2.

Figure 3.33: Standard VAR: government revenue (incl. federal non-defense expenditure)



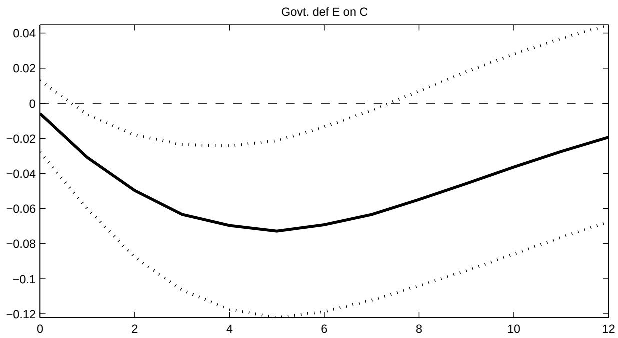
Notes: Response of private consumption to a government revenue shock, employing a standard SVAR model without anticipation featuring federal non-defense expenditure. Sample: 1947q1-2009q2.

Figure 3.34: Expectation augmented VAR: government revenue (incl. federal non-defense expenditure)



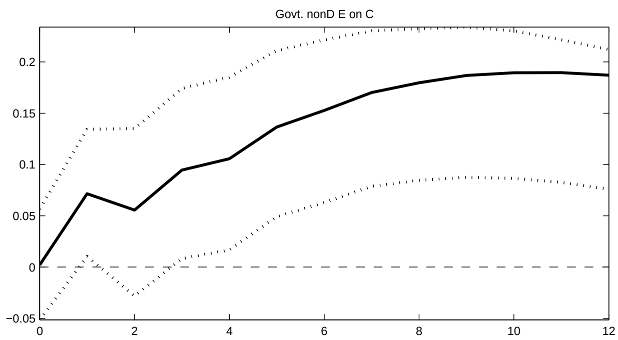
Notes: Response of private consumption to an anticipated government revenue shock, employing an expectation augmented VAR featuring federal non-defense expenditure. Sample: 1947q1-2009q2.

Figure 3.35: Standard VAR: defense expenditure (incl. GDP and 3-month T-bill rate)



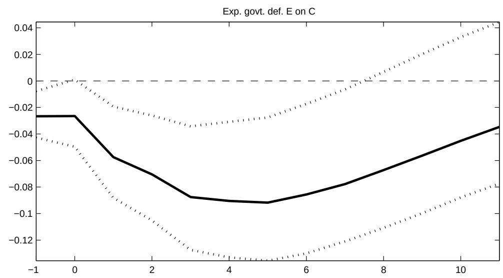
Notes: Response of private consumption to a government defense expenditure shock, employing a standard SVAR model without anticipation. VAR includes GDP and the 3-month T-bill rate. Sample: 1947q1-2009q2.

Figure 3.36: Standard VAR: non-defense expenditure (incl. GDP and 3-month T-bill rate)



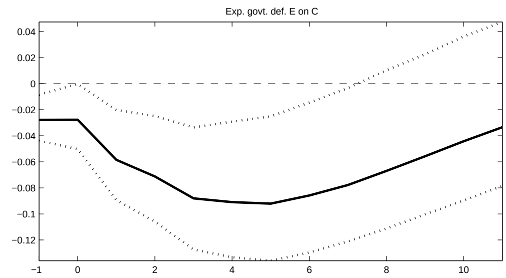
Notes: Response of private consumption to a government non-defense expenditure shock, employing a standard SVAR model without anticipation. VAR includes GDP and the 3-month T-bill rate. Sample: 1947q1-2009q2.

Figure 3.37: Expectation augmented VAR: defense expenditure ($\varepsilon_{c,r}(t-1) = 0$)



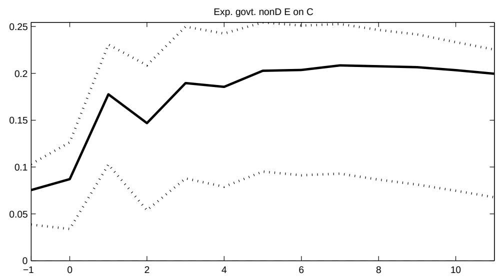
Notes: Response of private consumption to an anticipated government defense expenditure shock, employing an expectation augmented VAR. Elasticity of tax revenue to consumption at $t-1$: 0. Sample: 1947q1-2009q2.

Figure 3.38: Expectation augmented VAR: defense expenditure ($\varepsilon_{c,r}(t-1) = 0.5 * 0.6468$)



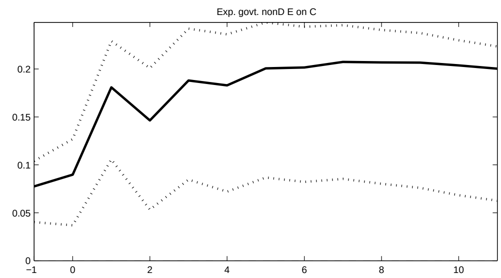
Notes: Response of private consumption to an anticipated government defense expenditure shock, employing an expectation augmented VAR. Elasticity of tax revenue to consumption at $t-1$: 0.5*0.6468. Sample: 1947q1-2009q2.

Figure 3.39: Expectation augmented VAR: non-defense expenditure ($\varepsilon_{c,r}(t-1) = 0$)



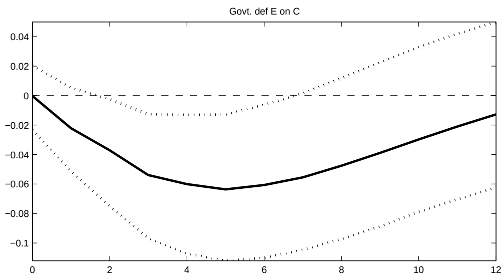
Notes: Response of private consumption to an anticipated government non-defense expenditure shock, employing an expectation augmented VAR. Elasticity of tax revenue to consumption at $t-1$: 0. Sample: 1947q1-2009q2.

Figure 3.40: Expectation augmented VAR: non-defense expenditure ($\varepsilon_{c,r}(t-1) = 0.5 * 0.6468$)



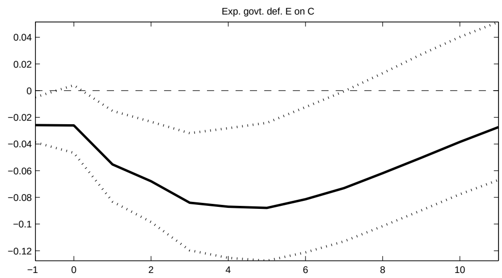
Notes: Response of private consumption to an anticipated government non-defense expenditure shock, employing an expectation augmented VAR. Elasticity of tax revenue to consumption at $t-1$: 0.5*0.6468. Sample: 1947q1-2009q2.

Figure 3.41: Standard VAR: defense expenditure ($\varepsilon_{c,r}(t) = 2.08$)



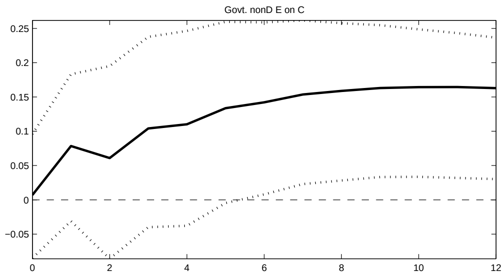
Notes: Response of private consumption to a government defense expenditure shock, employing a standard SVAR model without anticipation. Elasticity of tax revenue to consumption at t: 2.08. Sample: 1947q1-2009q2.

Figure 3.42: Expectation augmented VAR: defense expenditure ($\varepsilon_{c,r}(t) = 2.08$)



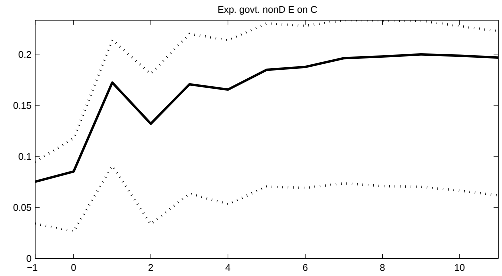
Notes: Response of private consumption to an anticipated government defense expenditure shock, employing an expectation augmented VAR. Elasticity of tax revenue to consumption at t: 2.08. Sample: 1947q1-2009q2.

Figure 3.43: Standard VAR: non-defense expenditure ($\varepsilon_{c,r}(t) = 2.08$)



Notes: Response of private consumption to a government non-defense expenditure shock, employing a standard SVAR model without anticipation. Elasticity of tax revenue to consumption at t: 2.08. Sample: 1947q1-2009q2.

Figure 3.44: Expectation augmented VAR: non-defense expenditure ($\varepsilon_{c,r}(t) = 2.08$)



Notes: Response of private consumption to an anticipated government non-defense expenditure shock, employing an expectation augmented VAR. Elasticity of tax revenue to consumption at t: 2.08. Sample: 1947q1-2009q2.

Concluding remarks

From a global perspective, this dissertation illustrates the consequences of choosing a particular balance between completeness and manageability in terms of model building, both in the field of macroeconomics and econometrics. As a fundamental basis, it emphasizes the close interaction of macroeconomic theory and empirical analysis as well as novel econometric techniques. Each of the three chapters shows that there are potentially dramatic consequences of taking into account, in a manageable way, additional and – with respect to the question at hand – essential layers of reality. In particular, in terms of econometric theory, Chapter 1 demonstrates that considerably more precise estimates within a dynamic factor model are obtainable by employing simple two-step estimators taking into account additional features of the data-generating process, i.e., autocorrelation and heteroskedasticity. Chapter 2, furthermore, considers a macroeconomic model featuring labor market frictions. It highlights the important consequences for equilibrium allocations and optimal monetary policy when altering the central aspect of the wage determination mechanism, so that it is consistent with empirical evidence. This illustrates that the chosen degree of abstraction might determine to a large extend the policy implications of a particular model. Finally, Chapter 3 presents an empirical investigation, studying the effects of fiscal policy on the macroeconomy. In this regard, it demonstrates the importance of allowing for particular features of the information structure as well as of distinguishing certain subcomponents of the fiscal variables, which might have different macroeconomic effects as implied by economic theory. As a result, we can illustrate that while at a certain level of abstraction, the findings of different approaches in the literature seem to be in conflict with each other, at another level the antagonism vanishes.

More specifically, Chapter 1 considers efficient estimation of dynamic factor models, a class of models popular in areas such as, for instance, macroeconomic forecasting and structural analysis. A simple two-step estimation procedure is suggested to obtain efficient estimates in the presence of both heteroskedasticity and autocorrelation. Interestingly, with respect to the factors, it is only potential heteroskedasticity which has to be taken into account, whereas for the loadings the relevant aspect is just autocorrelation. We derive the asymptotic distribution of the estimators and show that it is not affected by the estimation error in the covariance parameters and first stage PC estimates of the factors or loadings. While, as a result, the feasible two-step PC-GLS estimator is asymptotically as efficient as the estimator that (locally) maximizes the full approximate likelihood function, small sample gains may be obtained by iterating the two-step estimator. This is indeed reflected in the results of our extensive Monte Carlo investigation, which includes scenarios featuring autocorrelation, heteroskedasticity, and cross-sectional correlation as well as a setup based on a popular macroeconomic data set. Moreover, we also document the superior performance of the two-step PC-GLS estimator compared to standard PC.

The investigation of Chapter 2 is motivated by recent empirical findings with respect to the structure of wage rigidity. It studies optimal monetary policy using a simple New-Keynesian model featuring labor market frictions, heterogeneous wage setting, as well as markup shocks. Replacing the typically used uniformly rigid wage by a form of wage heterogeneity consistent with the data, has profound effects on the policy implications of this model. In particular, the sizable short-run inflation unemployment trade-off, which is present in the original setup, disappears. This results despite the fact that the original setup is just slightly changed and even though the model features an economy-wide average wage which is still rigid. Consequently, optimal monetary policy can exclusively concentrate on inflation with no concern for employment stabilization. As an overall rigid real wage is typically employed to address the so-called unemployment volatility puzzle, I follow suggestions in the literature with respect to an alternative mechanism and introduce markup shocks as additional driving forces into the model. While a short-run inflation unemployment trade-off indeed arises in this setup, optimal policy is nevertheless characterized by an overriding focus on inflation stabilization. Moreover, markup shocks do not generate a considerable amount of unemployment fluctuations within the model under consideration.

In light of the conflicting empirical results concerning the effects of fiscal policy on the macroeconomy and the potentially important role of fiscal policy anticipation in this regard, Chapter 3 investigates the response of private consumption to fiscal shocks within an SVAR framework, explicitly taking into account fiscal foresight. A new empirical approach is suggested, designed to align the information sets of the private agents and the econometrician, which allows us to avoid the problems of standard VARs. A simulation experiment based on a theoretical model featuring (imperfect) fiscal foresight documents the ability of the approach, in contrast to a standard VAR, to correctly capture macroeconomic dynamics. This result is even robust to deviations from the underlying informational assumptions of the expectation augmented VAR. The subsequent application to real life data indicates that it is indeed important in empirical work to allow for anticipation of fiscal policy. Moreover, it shows that it is crucial to distinguish subcomponents of total government expenditure which might have different macroeconomic effects according to economic theory. By distinguishing government defense and non-defense spending, it is possible to reconcile the results of the narrative and SVAR approaches to the study of fiscal policy effects.

In addition to the more abstract unifying theme indicated above, when considering future work it is possible to draw a more direct line between the three chapters of this dissertation. It would be a potentially fruitful avenue for further research to bring together the different aspects of the respective parts of this thesis. Once more, this would reflect the point stressed above of the importance of a close interaction of macroeconomic theory and empirical analysis as well as novel econometric techniques.

Considering Chapters 1 and 2, it would be interesting to employ dynamic factor models and particularly the suggested estimators to establish stylized facts and additional empirical regularities, which could help in guiding future macroeconomic modeling efforts. This would take the analysis presented in Chapter 2, which focuses on the aspect of the structure of wage rigidity found in the data, one step further. As this chapter illustrates the potentially crucial role played by aspects of the labor market for policy implications, it would be interesting to extend the set of stylized facts in this regard. Dynamic factor models in general and factor-augmented VARs (FAVARs) in the spirit of Bernanke and Boivin (2003) and Bernanke, Boivin, and Eliasz (2005), in particular, could be especially helpful in this context. In order to avoid degrees-of-freedom problems, standard (and also Bayesian) VARs are restricted in the number of variables which can be included. As a result, labor market variables are typically not considered in a monetary VAR. Hence, stylized facts with respect to the dynamic responses of the various labor market variables to monetary policy shocks are not well established. Since FAVARs do not have this limitation, it could be a potentially fruitful investigation to estimate those models with a particular focus on labor market aspects. The corresponding results, in turn, could help to further refine macroeconomic models with respect to the labor market dimension, potentially yielding new insights concerning the policy implications of those models. Employing the estimators presented in Chapter 1 could be of particular importance in this regard, as this would lead to more precise estimates for the impulse response functions, for instance. This could potentially increase the range of variables for which we could make statements with a certain degree of confidence.

Bringing together Chapters 2 and 3, it would be an interesting topic for further research to investigate the effects of fiscal policy on various labor market variables, taking into account fiscal policy anticipation. Macroeconomic models in the spirit of Chapter 2, but extended to include an interesting fiscal dimension, could help to decide which labor market variables are important to consider in the VAR and which subcomponents should be distinguished. The empirical findings, in turn, could give guidance on how to further refine those macroeconomic models. Furthermore, refining current models to take into account the empirical regularities concerning fiscal policy and the labor market might have important consequences with respect to the policy implications of the different models. Analogous to the investigation of Chapter 3, it would also be interesting to examine, whether the empirical results concerning the labor market are indeed affected by the presence of fiscal foresight and what is the importance of distinguishing different fiscal variables.

Finally, the methods developed in Chapter 1 could also be brought to bear on the problems related to fiscal policy anticipation as presented in Chapter 3. The extensive amount of information captured by a dynamic factor model could help to address the fundamental difficulty that the information set typically used by an econometrician is strictly smaller than the information set of the private agents.[42] As a result, it would be possible to recover the actual economic shocks and perform valid structural analysis. Thus, as an alternative to the approach presented in Chapter 3 and as a cross-check, estimating FAVARs or related models as suggested, for example, by Forni, Giannone, Lippi, and Reichlin (2009) could be an interesting topic for future research. Indeed, one motivation for estimating FAVARs when studying the effects of *monetary* policy is the so-called "price puzzle" found in standard monetary VARs, which can also be explained by a misalignment of information sets. The price puzzle describes a situation where following a positive shock in the interest rate the price level increases rather than decreases, as implied by standard economic theory. A possible explanation for this dynamic response is given by Sims (1992). He argues that the central bank possesses information about future inflation developments that is not included in the VAR. A typical "solution" to this problem is to enhance the information of the VAR by adding a commodity price index to the variables already present. However, this is quite arbitrary so that FAVARs have been employed (successfully) to address this problem. With respect to fiscal policy, a recent paper by Forni and Gambetti (2010) in fact uses the approach of Forni, Giannone, Lippi, and Reichlin (2009) to study the effects of government expenditure in the presence of fiscal policy anticipation. An interesting extension of that investigation, which would be in line with the analysis presented in Chapter 3, would consider shocks to different subcomponents of government spending. Moreover, applying the estimators suggested in Chapter 1 could address a shortcoming pervading almost the entire fiscal VAR literature and also the paper by Forni and Gambetti (2010). When presenting impulse response functions, what is typically plotted in conjunction with the point estimate are just 68% confidence bands. Using the more efficient estimators of Chapter 1 could help to raise the standard in this regard.

[42] See, for instance, Giannone and Reichlin (2006).

Bibliography

Anderson, T. W. (1984): *An Introduction to Multivariate Statistical Analysis*. Wiley, Hoboken, New Jersey, 2nd edn.

Andolfatto, D. (1996): "Business Cycles and Labor-Market Search," *The American Economic Review*, 86(1), 112–132.

Arseneau, D. M., and S. K. Chugh (2008): "Optimal Fiscal and Monetary Policy with Costly Wage Bargaining," *Journal of Monetary Economics*, 55(8), 1401–1414.

Bai, J. (2003): "Inferential Theory for Factor Models of Large Dimensions," *Econometrica*, 71(1), 135–172.

Bai, J., and S. Ng (2002): "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70(1), 191–221.

Balleer, A. (2009): "New Evidence, Old Puzzles: Technology Shocks and Labor Market Dynamics," *Institute of International Economic Studies, Stockholm University, mimeo.*

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

Bernanke, B. S., and J. Boivin (2003): "Monetary Policy in a Data-Rich Environment," *Journal of Monetary Economics*, 50(3), 525–546.

Bernanke, B. S., J. Boivin, and P. Eliasz (2005): "Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach," *The Quarterly Journal of Economics*, 120(1), 387–422.

BIAU, O., AND E. GIRARD (2005): “Politique Budgétaire et Dynamique Économique En France: L’approche VAR Structurelle,” *Economie et Prévision*, 169-71, 1–24.

BLANCHARD, O. J., AND J. GALÍ (2007): “Real Wage Rigidities and the New Keynesian Model,” *Journal of Money, Credit and Banking*, 39(1), 35–66.

——— (2010): “Labor Markets and Monetary Policy: A New Keynesian Model with Unemployment,” *American Economic Journal: Macroeconomics*, 2(2), 1–30.

BLANCHARD, O. J., AND R. PEROTTI (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.

BOIVIN, J., AND S. NG (2006): “Are more data always better for factor analysis?,” *Journal of Econometrics*, 132(1), 169–194.

BREITUNG, J., AND S. EICKMEIER (2011): “Testing for structural breaks in dynamic factor models,” *Journal of Econometrics, forthcoming*.

BREITUNG, J., AND J. TENHOFEN (2010): “GLS estimation of dynamic factor models,” *University of Bonn, mimeo*.

BURNSIDE, C., M. EICHENBAUM, AND J. D. M. FISHER (2004): “Fiscal Shocks and their Consequences,” *Journal of Economic Theory*, 115(1), 89–117.

CALVO, G. (1983): “Staggered Prices in a Utility-Maximizing Framework,” *Journal of Monetary Economics*, 12(3), 383–398.

CHAMBERLAIN, G., AND M. ROTHSCHILD (1983): “Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets,” *Econometrica*, 51(5), 1281–1304.

CHANG, Y. (2002): “Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency,” *Journal of Econometrics*, 110(2), 261–292.

CHOI, I. (2008): “Efficient Estimation of Factor Models,” *Sogang University, mimeo*.

CHRISTOFFEL, K., AND T. LINZERT (2010): “The Role of Real Wage Rigidity and Labor Market Frictions for Inflation Persistence,” *Journal of Money, Credit and Banking*, 42(7), 1435–1446.

Chudik, A., M. H. Pesaran, and E. Tosetti (2010): "Weak and strong cross section dependence and estimation of large panels," *The Econometrics Journal, forthcoming*.

de Castro, F., and P. H. de Cos (2008): "The Economic Effects of Fiscal Policy: The Case of Spain," *Journal of Macroeconomics*, 30(3), 1005–1028.

Devereux, M. B., A. C. Head, and B. J. Lapham (1996): "Monopolistic Competition, Increasing Returns, and the Effects of Government Spending," *Journal of Money, Credit and Banking*, 28(2), 233–254.

Dixit, A. K., and J. E. Stiglitz (1977): "Monopolistic Competition and Optimum Product Diversity," *The American Economic Review*, 67(3), 297–308.

Doz, C., D. Giannone, and L. Reichlin (2006): "A quasi maximum likelihood approach for large approximate dynamic factor models," *Working Paper No. 674, European Central Bank, Frankfurt am Main*.

——— (2007): "A Two-Step Estimator for Large Approximate Dynamic Factor Models Based on Kalman Filtering," *Discussion Paper No. 6043, Centre for Economic Policy Research, London, UK*.

Edelberg, W., M. Eichenbaum, and J. D. M. Fisher (1999): "Understanding the Effects of a Shock to Government Purchases," *Review of Economic Dynamics*, 2(1), 166–206.

Eickmeier, S. (2007): "Business cycle transmission from the US to Germany - A structural factor approach," *European Economic Review*, 51(3), 521–551.

Eickmeier, S., and C. Ziegler (2008): "How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach," *Journal of Forecasting*, 27(3), 237–265.

Faia, E. (2008): "Optimal Monetary Policy Rules with Labor Market Frictions," *Journal of Economic Dynamics & Control*, 32(5), 1600–1621.

——— (2009): "Ramsey Monetary Policy with Labour Market Frictions," *Journal of Monetary Economics*, 56(4), 570–581.

Fatás, A., and I. Mihov (2001): "The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence," *Discussion Paper No. 2760, Centre for Economic Policy Research, London, UK.*

Fernández-Villaverde, J., J. F. Rubio-Ramírez, T. J. Sargent, and M. W. Watson (2007): "ABCs (and Ds) of Understanding VARs," *The American Economic Review*, 97(3), 1021–1026.

Finn, M. G. (1998): "Cyclical Effects of Government's Employment and Goods Purchases," *International Economic Review*, 39(3), 635–657.

Fisher, J. D. M. (2006): "The Dynamic Effects of Neutral and Investment-Specific Technology Shocks," *The Journal of Political Economy*, 114(3), 413–451.

Forni, L., L. Monteforte, and L. Sessa (2009): "The general equilibrium effects of fiscal policy: Estimates for the Euro area," *Journal of Public Economics*, 93(3-4), 559–585.

Forni, M., and L. Gambetti (2010): "Fiscal foresight and the effects of government spending," *Università di Modena e Reggio Emilia, mimeo.*

Forni, M., D. Giannone, M. Lippi, and L. Reichlin (2009): "Opening the black box: structural factor models with large cross sections," *Econometric Theory*, 25(5), 1319–1347.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000): "The Generalized Dynamic Factor Model: Identification and Estimation," *The Review of Economics and Statistics*, 82(4), 540–554.

——— (2005): "The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting," *Journal of the American Statistical Association*, 100(471), 830–840.

Galí, J. (2010): "Monetary Policy and Unemployment," *Universitat Pompeu Fabra, mimeo.*

Galí, J., M. Gertler, and J. D. López-Salido (2007): "Markups, Gaps, and the Welfare Costs of Business Fluctuations," *The Review of Economics and Statistics*, 89(1), 44–59.

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

Gertler, M., and A. Trigari (2009): "Unemployment Fluctuations with Staggered Nash Wage Bargaining," *Journal of Political Economy*, 117(1), 38–86.

Geweke, J. F. (1977): "The Dynamic Factor Analysis of Economic Time Series," in *Latent Variables in Socio-Economic Models*, ed. by D. J. Aigner, and A. S. Goldberger, Amsterdam. North-Holland.

Giannone, D., and L. Reichlin (2006): "Does information help recovering structural shocks from past observations?," *Journal of the European Economic Association*, 4(2-3), 455–465.

Giannone, D., L. Reichlin, and L. Sala (2002): "Tracking Greenspan: Systematic and Unsystematic Monetary Policy Revisited," *Discussion Paper No. 3550, Centre for Economic Policy Research, London, UK.*

Giordano, R., S. Momigliano, S. Neri, and R. Perotti (2007): "The Effects of Fiscal Policy in Italy: Estimates from a VAR Model," *European Journal of Political Economy*, 23(3), 707–733.

Gomes, P. (2009): "Fiscal Policy and the Labour Market: the Effects of Public Sector Employment and Wages," *London School of Economics, mimeo.*

Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988): "Investment, Capacity Utilization, and the Real Business Cycle," *The American Economic Review*, 78(3), 402–417.

Greenwood, J., Z. Hercowitz, and P. Krusell (1997): "Long-Run Implications of Investment-Specific Technological Change," *The American Economic Review*, 87(3), 342–362.

——— (2000): "The role of investment-specific technological change in the business cycle," *European Economic Review*, 44(1), 91–115.

Haefke, C., M. Sonntag, and T. van Rens (2008): "Wage Rigidity and Job Creation," *Discussion Paper No. 3714, Institute for the Study of Labor (IZA), Bonn.*

Hall, R. E. (2004): "The Labor Market is the Key to Understanding the Business Cycle," *Stanford University, mimeo.*

——— (2005): "Employment Fluctuations with Equilibrium Wage Stickiness," *The American Economic Review*, 95(1), 50–65.

Hall, R. E., and P. R. Milgrom (2008): "The Limited Influence of Unemployment on the Wage Bargain," *The American Economic Review*, 98(4), 1653–1674.

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

Hansen, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50(4), 1029–1054.

Hansen, L. P., J. Heaton, and A. Yaron (1996): "Finite sample properties of some alternative GMM estimators," *Journal of Business & Economic Statistics*, 14(3), 262–280.

Hansen, L. P., and T. J. Sargent (1980): "Formulating and Estimating Dynamic Linear Rational Expectations Models," *Journal of Economic Dynamics & Control*, 2, 7–46.

——— (1991): "Two Difficulties in Interpreting Vector Autoregressions," in *Rational Expectations Econometrics*, ed. by L. P. Hansen, and T. J. Sargent, pp. 77–119. Westview Press.

Hayashi, F. (2000): *Econometrics*. Princeton University Press, Princeton, New Jersey.

IN, F., AND J. H. YOON (2007): "Determination of asset prices with an investment-specific technology model: Implications for the equity premium puzzle," *Journal of Economic Dynamics & Control*, 31(8), 2637–2658.

JUNGBACKER, B., AND S. J. KOOPMAN (2008): "Likelihood-based analysis for dynamic factor models," *VU University Amsterdam, mimeo.*

KRAUSE, M. U., AND T. A. LUBIK (2007): "The (Ir)Relevance of Real Wage Rigidity in the New Keynesian Model with Search Frictions," *Journal of Monetary Economics*, 54(3), 706–727.

KRIWOLUZKY, A. (2009): "Pre-Announcement and Timing – the Effects of a Government Expenditure Shock," *European University Institute, mimeo.*

LEEPER, E. M., T. B. WALKER, AND S.-C. S. YANG (2009): "Fiscal Foresight and Information Flows," *Indiana University, mimeo.*

LINNEMANN, L., AND A. SCHABERT (2003): "Fiscal Policy in the New Neoclassical Synthesis," *Journal of Money, Credit and Banking*, 35(6), 911–929.

LIPPI, M., AND L. REICHLIN (1993): "The Dynamic Effects of Aggregate Demand and Supply Disturbances: Comment," *The American Economic Review*, 83(3), 644–652.

——— (1994): "VAR analysis, nonfundamental representations, Blaschke matrices," *Journal of Econometrics*, 63(1), 307–325.

MERTENS, K., AND M. O. RAVN (2009): "Measuring the Impact of Fiscal Policy in the Face of Anticipation: a Structural VAR Approach," *Discussion Paper No. 7423, Centre for Economic Policy Research, London, UK.*

MERZ, M. (1995): "Search in the Labor Market and the Real Business Cycle," *Journal of Monetary Economics*, 36(2), 269–300.

MORTENSEN, D. T., AND É. NAGYPÁL (2007): "More on Unemployment and Vacancy Fluctuations," *Review of Economic Dynamics*, 10(3), 327–347.

Mountford, A., and H. Uhlig (2009): "What are the Effects of Fiscal Policy Shocks?," *Journal of Applied Econometrics*, 24(6), 960–992.

Patton, A., D. N. Politis, and H. White (2009): "Correction to "Automatic block-length selection for the dependent bootstrap" by D. Politis and H. White," *Econometric Reviews*, 28(4), 372–375.

Perotti, R. (2005): "Estimating the Effects of Fiscal Policy in OECD Countries," *Discussion Paper No. 4842, Centre for Economic Policy Research, London, UK.*

——— (2008): "In Search of the Transmission Mechanism of Fiscal Policy," in *NBER Macroeconomics Annual 2007*, ed. by D. Acemoglu, K. Rogoff, and M. Woodford, vol. 22, pp. 169–226. University of Chicago Press.

Pissarides, C. A. (2007): "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?," *Discussion Paper No. 839, Centre for Economic Performance, London School of Economics, London.*

——— (2009): "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?," *Econometrica*, 77(5), 1339–1369.

Politis, D. N., and J. P. Romano (1992): "A circular block-resampling procedure for stationary data," in *Exploring the limits of bootstrap*, ed. by R. LePage, and L. Billard, pp. 263–270, New York, NY. John Wiley & Sons, Inc.

Politis, D. N., and H. White (2004): "Automatic block-length selection for the dependent bootstrap," *Econometric Reviews*, 23(1), 53–70.

Quah, D. (1990): "Permanent and Transitory Movements in Labor Income: An Explanation for "Excess Smoothness" in Consumption," *The Journal of Political Economy*, 98(3), 449–475.

Ramey, V. A. (2008): "Comment," in *NBER Macroeconomics Annual 2007*, ed. by D. Acemoglu, K. Rogoff, and M. Woodford, vol. 22, pp. 237–246. University of Chicago Press.

——— (2009): “Identifying Government Spending Shocks: It’s All in the Timing,” *University of California, San Diego, mimeo.*

Ramey, V. A., and M. D. Shapiro (1998): “Costly Capital Reallocation and the Effects of Government Spending,” *Carnegie Rochester Conference Series on Public Policy*, 48, 145–194.

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

Redner, R. A., and H. F. Walker (1984): “Mixture Densities, Maximum Likelihood and the EM Algorithm,” *SIAM Review*, 26(2), 195–239.

Romer, C. D., and D. H. Romer (1989): “Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz,” in *NBER Macroeconomics Annual 1989*, ed. by O. J. Blanchard, and S. Fischer, vol. 4, pp. 121–170. University of Chicago Press.

——— (2010): “The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks,” *The American Economic Review*, 100(3), 763–801.

Rotemberg, J. J. (2008): “Cyclical Wages in a Search-and-Bargaining Model with Large Firms,” *NBER International Seminar on Macroeconomics 2006*, pp. 65–114.

Rotemberg, J. J., and M. Woodford (1991): “Markups and the Business Cycle,” in *NBER Macroeconomics Annual*, ed. by O. J. Blanchard, and S. Fischer, vol. 6, pp. 63–140, Cambridge, Mass. MIT Press.

——— (1992): “Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity,” *The Journal of Political Economy*, 100(6), 1153–1207.

——— (1999): “The Cyclical Behavior of Prices and Costs,” in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1B, pp. 1051–1135, Amsterdam. Elsevier, North Holland.

SALA, L., U. SÖDERSTRÖM, AND A. TRIGARI (2008): "Monetary Policy under Uncertainty in an Estimated Model with Labor Market Frictions," *Journal of Monetary Economics*, 55(5), 983–1006.

SARGENT, T. J., AND C. A. SIMS (1977): "Business Cycle Modeling Without Pretending to Have Too Much a Priori Economic Theory," in *New Methods in Business Cycle Research*, ed. by C. A. Sims, and C. W. J. Granger, Minneapolis. Federal Reserve Bank of Minneapolis.

SHIMER, R. (2004): "The Consequences of Rigid Wages in Search Models," *Journal of the European Economic Association*, 2(2-3), 469–479.

——— (2005): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *The American Economic Review*, 95(1), 25–49.

SIMS, C. A. (1992): "Interpreting the macroeconomic time series facts: the effects of monetary policy," *European Economic Review*, 36(5), 975–1000.

SIMS, C. A., AND T. ZHA (2006): "Does Monetary Policy Generate Recessions?," *Macroeconomic Dynamics*, 10(2), 231–272.

SÖDERLIND, P. (1999): "Solution and Estimation of RE Macromodels with Optimal Policy," *European Economic Review*, 43(4-6), 813–823.

STEINSSON, J. (2003): "Optimal Monetary Policy in an Economy with Inflation Persistence," *Journal of Monetary Economics*, 50(7), 1425–1456.

STOCK, J. H., AND M. W. WATSON (2002a): "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business & Economic Statistics*, 20(2), 147–162.

——— (2002b): "Forecasting Using Principal Components from a Large Number of Predictors," *Journal of the American Statistical Association*, 97(460), 1167–1179.

——— (2005): "Implications of Dynamic Factor Models for VAR Analysis," *Working paper no. 11467, National Bureau of Economic Research, Cambridge, Massachusetts.*

Tenhofen, J. (2010): “Optimal monetary policy under labor market frictions: the role of wage rigidity and markup shocks,” *University of Bonn, mimeo.*

Tenhofen, J., and G. B. Wolff (2007): “Does Anticipation of Government Spending Matter? Evidence from an Expectation Augmented VAR,” *Deutsche Bundesbank Discussion Paper*, No. 14/2007, Deutsche Bundesbank, Frankfurt am Main.

——— (2010): “Does anticipation of government spending matter? The role of (non-)defense spending,” *University of Bonn, mimeo.*

Tenhofen, J., G. B. Wolff, and K. H. Heppke-Falk (2010): “The Macroeconomic Effects of Exogenous Fiscal Policy Shocks in Germany: A Disaggregated SVAR Analysis,” *Journal of Economics and Statistics*, 230(3), 328–355.

Thomas, C. (2008): “Search and Matching Frictions and Optimal Monetary Policy,” *Journal of Monetary Economics*, 55(5), 936–956.

Townsend, R. M. (1983): “Forecasting the Forecasts of Others,” *The Journal of Political Economy*, 91(4), 546–588.

Trigari, A. (2009): “Equilibrium Unemployment, Job Flows and Inflation Dynamics,” *Journal of Money, Credit and Banking*, 41(1), 1–33.

Turnovsky, S. J., and W. H. Fisher (1995): “The composition of government expenditure and its consequences for macroeconomic performance,” *Journal of Economic Dynamics & Control*, 19(4), 747–786.

Watson, M. W. (2003): “Macroeconomic Forecasting Using Many Predictors,” in *Advances in Economics and Econometrics: Theory and Applications*, ed. by M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, vol. III, pp. 87–114, Cambridge, UK. Cambridge University Press.

Woodford, M. (2003): *Interest and Prices.* Princeton University Press, Princeton, New Jersey.