New Issues in Empirical Business Cycle Analysis: News Shocks and Structural Inference

Dissertation
zur Erlangung des wissenschaftlichen Doktorgrades (Dr. rer. pol.)
an der Fakultät Wirtschafts- und Sozialwissenschaften der
Universität Hamburg

vorgelegt von

Thomas Haertel

aus Gifhorn

21. Mai 2016

Prüfungskommission:

Vorsitzender: Prof. Dr. Thomas Straubhaar

Erstgutachter: Prof. Dr. Bernd Lucke

Zweitgutachter: Prof. Dr. Ulrich Fritsche

Datum der Disputation: 18. November 2016

Preface

First of all, I would like to thank my supervisor Prof. Dr. Bernd Lucke for giving me the opportunity to write this thesis. I benefited strongly from his advice, suggestions and guidance, in particular during the time I worked at the Institute for Growth and Fluctuations at the University of Hamburg. Additionally, I would like to thank my second supervisor Prof. Dr. Ulrich Fritsche and Prof. Dr. Thomas Straubhaar as chair of my thesis committee.

I am very indebted to my former colleagues from the institute for their help, moral support and friendship. In particular, I wish to thank Johann Buhné, Angela Ebert, Karin Endrejat, Maximilian Ludwig, Olaf Posch, Pascal Terveer and Jacopo Zotti. I benefited a lot from fruitful discussions and our teamwork but also from the interpersonal relationship not only during the office hours. In this context, I am grateful to have joined the Feierabend crew, which has been a great pleasure for me. Inga Hardeck and Karl Broemel, thank you for your friendship.

Especially, I am very thankful to my friend Atilim Seymen. Atilim has been a great help by lending his ear, being patient and providing valuable comments and feedback. His excellent support improved the thesis substantially.

Last but not least, I am deeply indebted to my family. Over the years, they have had to bear a lot of my mood swings. Their unconditional patience and encouragement helped me complete the thesis.

Frankfurt am Main, November 2016   *Thomas Haertel*

Contents

Contents

Contents

List of Figures

List of Figures

List of Tables

Summary

This thesis deals with the application of structural estimation methods in the context of the so-called news shocks. The main focus is on two estimation approaches for time series, the vector autoregressive model and the subspace algorithm. While the former approach is a standard procedure employed in empirical business cycle analysis, using the subspace algorithm is less common in this field. There exist only a few examples of the usage of the subspace algorithm in the literature, in particular, when taking potential cointegration properties among the underlying time series into account. The thesis addresses this issue and provides a comprehensive comparison of both estimation methods by means of simulation studies and empirical investigations.

Another key aspect of the thesis is its focus on the concept of news-driven business cycles and the challenge of identifying the corresponding shocks in the empirical data. The idea behind news-driven business cycles is that the change in the agents' expectations about the future state of the economy caused by an exogenous event (i.e., the news shock) can generate macroeconomic fluctuations. Considering news shocks as a driving force of business cycle phenomena is controversial in the literature. First, there are several other possible candidates for this role. Second, the identification of news shocks can be more difficult than in the case of the other potential causes of business cycle fluctuations. The intricacy is to uncover the quantitative importance of the agents' forward-looking behavior in macroeconomic data with the aforementioned estimation methods. Since these methods are typically limited to a small number of variables, there might be insufficient information for identifying the actual anticipation effects.

In this thesis, two scenarios are distinguished, each of which considers a particular type of news shock. The first kind of news shock is supposed to reflect market rush behavior of firms which immediately invest in new startups after receiving a signal about future market opportunities in order to secure monopoly rents in newly opened markets. The second type of news shock implies that the economic agents anticipate future productivity growth induced by technical progress. In the first scenario, the above-mentioned difficulty does not exist, meaning that the estimation methods are, in principle, capable of correctly identifying the underlying shocks and the corresponding dynamics in the data. By contrast, the information set in the second scenario is theoretically not sufficient for empirically determining news shocks and their implications in a precise way. Simulation studies are carried out for both scenarios to analyze the goodness of fit of the estimation methods. Finally, I follow two recently published studies and conduct empirical analyses of the US data to explore the quantitative importance of these two types of news shocks. As opposed to these publications, I apply the subspace algorithm in addition to the vector autoregressive model.

The main results of the thesis can be summarized as follows. In the simulation exercises, I

show that the subspace algorithm represents a promising alternative to the vector autoregressive model, even though the findings depend on the parameterization of the underlying data generating process. Moreover, I demonstrate that a suitable modification of both estimation procedures can mitigate the above identification problem in the case of the second scenario. As regards the empirical examinations, my results support the findings in the literature to some extent. While market rush shocks prove to be an important driver of business cycles in the very short-run (i.e., within the one-year-horizon), the role of technological news shocks in explaining macroeconomic fluctuations is slightly ambiguous. At horizons beyond the first year, technological news shocks are a dominant source of US business cycles when a vector autoregressive model is used for estimation, whereas this kind of shock is more a medium- and long-run determinant of fluctuations according to the subspace algorithm analysis.

Zusammenfassung

Die vorliegende Arbeit behandelt die Anwendung von strukturellen Schätzmethoden im Kontext der sogenannten Nachrichtenschocks. Im Mittelpunkt stehen dabei zwei Ansätze zur Schätzung von Zeitreihen, das vektorautoregressive Modell und der Subspace-Algorithmus. Während erstgenanntes zum Standardinstrumentarium in der empirischen Konjunkturanalyse gehört, ist die Anwendung des Subspace-Algorithmus in diesem Bereich weniger verbreitet. Insbesondere, wenn es bei der Analyse um die Berücksichtigung möglicher Kointegrationseigenschaften der zugrunde liegenden makroökonomischen Zeitreihen geht, gibt es in der Literatur bisher wenige Beispiele für den Einsatz des Algorithmus. Diese Arbeit greift diesen Umstand auf und dokumentiert einen umfassenden Vergleich der beiden Schätzmethoden mit Hilfe von Simulationsstudien sowie empirischen Analysen.

Ein weiterer Schwerpunkt der Arbeit ist der besondere Fokus auf das Konzept der nachrichtengetriebenen Konjunkturzyklen und der Herausforderung der Identifikation entsprechender Schocks in den empirischen Daten. Die Idee hinter nachrichtengetriebenen Konjunkturzyklen ist, dass die durch einen exogenen Impuls (in Form des Nachrichtenschocks) ausgelöste Änderung der Erwartungshaltung von Wirtschaftssubjekten über den zukünftigen Zustand der Ökonomie gesamtwirtschaftliche Fluktuationen verursachen kann. Nachrichtenschocks als treibende Kraft hinter konjunkturellen Schwankungen einzuordnen ist in der Literatur umstritten. Erstens gibt es zahlreiche andere Kandidaten, die ebenfalls dafür in Frage kommen. Zweitens ist die Identifizierung der Nachrichtenschocks mitunter schwieriger als im Fall der anderen potentiellen Konjunkturtreiber. Die Schwierigkeit besteht darin, mit Hilfe der zuvor erwähnten Schätzmethoden die quantitative Bedeutung von vorausschauendem Verhalten der Individuen in den makroökonomischen Daten aufzudecken. Da diese Methoden üblicherweise nur eine begrenzte Anzahl an Variablen berücksichtigen, kann dies zur Folge haben, dass nicht ausreichend Informationen genutzt werden können, um die tatsächlichen Antizipationseffekte offenzulegen.

In der Arbeit werden zwei Szenarien unterschieden, in denen jeweils die Rolle einer bestimmten Art von Nachrichtenschocks betrachtet wird. Bei der ersten Form von Nachrichtenschocks wird angenommen, dass Unternehmen, ähnlich wie bei einem Goldrausch, als Reaktion auf eine Ankündigung zukünftiger Markterschließungsoptionen frühzeitig investieren, um entsprechende Monopolrenten auf den neu entstehenden Märkten abschöpfen zu können. Bei der zweiten Art von Nachrichtenschocks wird unterstellt, dass die Wirtschaftsakteure zukünftiges Produktivitätswachstum, welches durch technischen Fortschritt hervorgerufen wird, antizipieren. Im ersten Szenario tritt die oben genannte Schwierigkeit nicht auf. Das bedeutet, dass die vorgestellten Schätzmethoden dabei grundsätzlich in der Lage sind, die wahren zugrunde liegenden Schocks und die damit verbundenen Dynamiken in den Daten korrekt zu identifizieren. Im zweiten Szenario wird hingegen davon ausgegan-

gen, dass die Informationsmenge theoretisch nicht ausreicht, um die Nachrichtenschocks und deren Implikationen empirisch eindeutig bestimmen zu können. Für beide Szenarien werden zunächst Simulationsstudien durchgeführt, um die Güte der zuvor genannten Schätzmethoden zu analysieren. Den Abschluss der jeweiligen Untersuchungen bilden empirische Analysen von US-Daten in Anlehnung an zwei neuere Studien, in denen die quantitative Bedeutung der betrachteten Nachrichtenschocks erforscht wird. Im Gegensatz zu diesen Publikationen, welche nur auf der Anwendung von vektorautoregressiven Modellen basieren, wird dabei auch der Subspace-Algorithmus eingesetzt.

Die wichtigsten Ergebnisse der Arbeit können wie folgt zusammengefasst werden. In den Simulationsübungen zeigt sich, dass der Subspace-Algorithmus im Rahmen einer strukturellen Schätzung eine vielversprechende Alternative zum vektorautoregressiven Modell darstellt, wenngleich die Ergebnisse von der Parametrisierung des zugrunde liegenden datengenerierenden Prozesses abhängen. Zudem wird im Rahmen des zweiten Szenarios verdeutlicht, dass durch eine geeignete Modifikation der beiden Schätzverfahren dem obigen Identifikationsproblem entgegengewirkt werden kann. Im Hinblick auf die empirischen Untersuchungen unterstützen die dazugehörigen Resultate die Befunde in der einschlägigen Literatur in gewissem Maße. Während sich die Goldrauschschocks als bedeutender Auslöser von konjunkturellen Schwankungen in der sehr kurzen Frist (d. h. innerhalb des Ein-Jahres-Horizonts) erweisen, ist die Rolle der technologischen Nachrichtenschocks bei der Erklärung von gesamtwirtschaftlichen Fluktuationen nicht ganz so eindeutig. Sofern die Schätzung auf einem vektorautoregressiven Modell beruht, dominieren die technologischen Nachrichtenschocks die Erklärung von US-Konjunkturzyklen in Bezug auf einen Zeithorizont ab einem Jahr, wohingegen diese Art von Schock im Fall der Analyse mit dem Subspace-Algorithmus eher als Ursache für mittel- bis langfristige Schwankungen angesehen werden kann.

Chapter 1

Introduction

Empirical business cycle analysis covers a broad range of topics. One branch explores the sources of macroeconomic fluctuation and examines the relevant explanatory approaches. These approaches are characterized by considerable variety and reflect a wide spectrum of theories and hypotheses. The concept of news-driven business cycles is one of these issues. Its basic core is that economic agents adjust their expectations or perceptions of future states of the economy such that its actual development is affected substantially. The notion that expectations matter for macroeconomic fluctuations dates back at least to authors, such as Pigou (1927), Keynes (1936) and von Haberler (1937), who assigned an important role to expectations in their explanatory theories of dynamics and interactions within economies. Among these authors, John Maynard Keynes has been the most prominent, and his thoughts and ideas still influence contemporary research in economics.

I refer to the following three quotes of the aforementioned authors to highlight the historical relevance of expectations in the profession. For instance, Keynes describes expectations as important determinants of the production and employment levels in the economy.

> *... a mere change in expectation is capable of producing an oscillation of the same kind of shape as a cyclical movement, in the course of working itself out.* (Keynes (1936), p. 49)

Likewise, Pigou sees variation in agents' expectations as trigger of business cycle movements.

> *Thus, (...) we conclude definitely that they [i.e., varying expectations], and not anything else, constitute the immediate and direct causes or antecedents of industrial fluctuations.* (Pigou (1927), p. 29)

Furthermore, von Haberler hints at the challenge of capturing and quantifying the agents' expectations and perceptions.

1

> *We cannot observe states of mind; but it is possible to make certain observations from which states of mind or changes of mind can be inferred.* (von Haberler (1937), p. 147)

It took almost seven decades for macroeconomists to address expectations-driven business cycles in detail. In this context, Paul Beaudry and Franck Portier have to be mentioned as pioneers who devote themselves to the relevance of expectations-driven business cycles from a theoretical and an empirical perspective. The core of their approach is to model and identify so-called news shocks as potential sources of macroeconomic fluctuation. Beaudry and Portier characterize news shocks as signals about future productivity growth received by the agents at least one period in advance such that the agents adapt their behavior immediately after the arrival of that new information. In theoretical modeling, building a corresponding business cycle model that generates comovement between macroeconomic aggregates, such as output, consumption, investment and hours worked, in response to a news shock is difficult.[1] With respect to empirical business cycle analysis, the challenge, as indicated by von Haberler, is to filter out from the macro data only those elements that reflect agents' expectations using appropriate econometric methods. In their very recent publication, Beaudry and Portier (2014a) provide a comprehensive overview of the concept of the news-driven business cycle and its related issues.

The articles by Beaudry and Portier can be considered the key starting points for extensive subsequent research. On the one hand, the work of Beaudry and Portier serves as inspiration and guide for further studies that emphasize the importance of agents' expectations but may differ in how they model news shocks. On the other hand, the econometric methodology used by Beaudry and Portier is at the heart of a debate about the suitability of their strategy for identifying news shocks. In my thesis, I refer to several previous studies that deal with the aforementioned aspects. A focal point is the examination of the second issue, that is, the appropriateness of the methodological procedures of Beaudry and Portier. Therefore, I concentrate on the main criticism, which is based on the argument that the strategies of Beaudry and Portier are not adequate for studying anticipation behavior of economic agents because they simply rely on the usage of an insufficient information set. Due to the limited number of observation variables, the degree of information is constrained such that the empirical methods applied are not capable of identifying news shocks or their implications in the data. In this thesis, I elaborate this problem and explore whether particular approaches represent possible solutions.

The contribution of this thesis becomes apparent when considering developments over recent years, wherein the role of expectation and anticipation in the economic context has been clearly revealed.[2] In the wake of the financial and debt crises, there has been intense

---

[1]Nonetheless, a few articles published before those of Beaudry and Portier point to this issue, e.g., the contributions by Barro and King (1984), Cochrane (1994b) and Danthine et al. (1998).
[2]See Matheson and Stavrev (2014) and Girardi (2014) for addressing developments in the US and the euro area.

discussion of various related issues, including the monetary policy stance of central banks. In view of the crises and the zero lower bound of interest rates, central banks have been forced to deploy unconventional policies, for instance, changing the design of central bank communication, which has focused more strongly on controlling market expectations.[3] An example is the forward guidance strategy, as recently undertaken by the US Federal Reserve (and Bank of England), wherein the monetary policy measure is directly coupled with the future path of a macroeconomic indicator (in terms of the unemployment rate).[4]

Reappraisal of the financial and debt crises over recent years has included the role of rating agencies, which has also been the focus of public attention. Rating agencies provide information about the current and future conditions and performance of institutions, such as banks or states. Thus, whether these agencies can be held responsible for the occurrence of the financial crisis, e.g., due to the assignment of misleading ratings to collateralized debt obligations, has been discussed. Since the outbreak of the debt crisis, the focus has been directed more toward the agencies' assessment of countries' debt sustainability. In this respect, Durdu et al. (2013) pick up Beaudry and Portier's concept of anticipated future productivity shocks in order to study the impacts of a change in agents' expectations in a model of sovereign debt and default risk, wherein news shocks contain information about the government's ability to repay its debts. Likewise, Gunn and Johri (2013) adopt Beaudry and Portier's idea to present a theoretical explanation of the Great Recession in the US, which they interpret as consequence of revised expectations about efficiency in the financial sector.

Though this thesis does not take up such policy considerations, it highlights the practical relevance of an empirical analysis of the expectations-driven behavior of agents from a macroeconomic perspective and addresses corresponding research questions. If the goal is to analyze whether expectation shocks, especially news shocks, can trigger business cycles, it is necessary to examine the tools that are applied in this context. This investigation is the main contribution of my thesis.

Beaudry and Portier's empirical approach relies on structural vector autoregression (SVAR) models, which explain the dynamics of (or within) an economy with a very limited number of macroeconomic variables that are regressed on their own past values thereby allowing for interdependencies among them. SVARs are used to identify structural shocks in the data

---

[3]Barakchian and Crowe (2013) provide empirical evidence that US monetary policy has tended to become more forward-looking since 1988.

[4]A detailed explanation of forward guidance can be found, e.g., in the article by Justiniano et al. (2012). The communication strategy of the European Central Bank (ECB) also represents a type of forward guidance, although it is not directly linked to a macroeconomic indicator. However, the ECB attempts to strongly influence market expectations by giving concrete statements on the future path of monetary policy. As examples, one could mention the press conferences of Mario Draghi, who announced on 26 July 2012 that "the ECB is ready to do whatever it takes to preserve the euro" (see http://www.ecb.europa.eu/press/key/date/2012/html/sp120726.en.html) and further emphasized on 4 July 2013 that the ECB expects the future path of the ECB's policy interest rates to remain unchanged or even lower (see also ECB Monthly Report, July 2013, pp. 5-10).

as exogenous drivers of the dynamic behavior of economic agents. These shocks are then construed as sources of business cycle movement. Their identification requires appropriate restrictions, which are based on theoretical considerations that permit an economic interpretation of the shocks and their implications. Restrictions on the short-run and/or long-run variable responses to the shocks are imposed. To motivate the selected restrictions, a specific business cycle model can be chosen, but usually, the analysis is not conditioned on a particular model, as is the case with other more theory-oriented methods. The heart of the VAR approach is to let the data speak freely to a large extent rather than putting it into a precise theoretical straightjacket.

Subspace algorithm analysis can be applied in a similar fashion. This method originates in linear systems theory and represents a well-established tool for empirical analysis in engineering. It is based on the state space model representation in which one or more state variables determine the state of the remaining variables in the system. While the latter variables are observable, state variables are generally not. Hence, the challenge of estimating a state space model is to correctly capture the impact of the unobservables. As will be seen in this thesis, subspace algorithm analysis is one possible way to do this. As in the VAR procedure, it does not presume a particular state space model and allows the data to behave more flexibly. Likewise, the implementation of the subspace algorithm analysis depends on the usage of a limited set of available (i.e., observable) data series. The aforementioned problems with the identification of news shocks à la Beaudry and Portier evidently also arise in this context.

To the best of my knowledge, there are only a few empirical studies in business cycle research in which subspace algorithm analysis is used. This is even more the case for situations in which cointegration relationships can be found in the data. Cointegration is a statistical concept describing the property of time series to exhibit comovement that is driven by one (or more) common stochastic trend(s). Macroeconomic time series are often proved to have this feature. Dietmar Bauer and Martin Wagner have developed a modified version of the subspace algorithm accounting for cointegration. This thesis is probably the first empirical business cycles analysis in which Bauer and Wagner's subspace algorithm is applied to structural estimation. The restrictions that are necessary for identification can be employed in the same way as in SVAR, so a comparison of SVAR and structural subspace algorithm analysis is an important goal. Such a comparison involves carrying out Monte Carlo simulations in which artificial time series are generated on the basis of a theoretical business cycle model and, in turn, used for estimation via the above-mentioned methods. The individual performance of a method is evaluated in how well it can reproduce the dynamics of the underlying theoretical model. Thereby, the corresponding impulse responses and the identified structural shocks provide the foundation for the assessment. A subsequent empirical application of the methods completes the comparative study.

Throughout my thesis, I conduct two comparative studies that deal with two different scenarios. In the first, I assume that the available information set coming from the observable data is sufficient to uncover the structure of the underlying data generating process (DGP). In contrast, the second scenario considers the situation in which there is misalignment between the available and the theoretically sufficient information sets. The latter scenario reflects the situation that has been indicated above with respect to the work of Beaudry and Portier. It should be noted that the classification of an information set as sufficient is referred as the asymptotic case here. In this thesis, I designate an information set as insufficient if it cannot disclose the true DGP with the help of the applied methods even if the data used for the analysis contained an infinite number of observations. Of course, there are additional difficulties in practice, such as limited sample sizes, which affect the performance of the estimation methods. Such issues will also be addressed in the Monte Carlo simulations. Differentiating between the two scenarios enables a convenient treatment to systematically identify how various aspects influence the success of the estimation methods and to show how potential problems can be handled appropriately.

My thesis is thus organized as follows. Chapter 2 outlines the theoretical framework, which establishes the basic setting for the subsequent chapters. To classify the suitability of an information set, it is helpful to distinguish between two perspectives. On the one hand, there is the perspective of the economic agents whose actions are represented by a corresponding business cycle model. This model is supposed to be the true DGP from which the observable time series originate. In my thesis, I call the associated representation of the model the fully informed agents (FIA) model. This name is derived from the assumption that all agents in the model economy have access to all relevant information that could influence their decision making. On the other hand, there is the econometrician's perspective. As opposed to the agents, the econometrician is faced with the constraint that her analysis relies only on the results of the agents' decisions that are captured by the available observations in terms of the limited number of macroeconomic time series. In the following, I consider the econometrician someone who applies the previously introduced empirical instruments, i.e., VAR and subspace algorithm analysis. The perspective of the econometrician is described by the limitedly informed econometrician (LIE) model.

The FIA and LIE models are based on their respective state space representations. They are equivalent in the case wherein the information sets of agents and econometrician are the same, so the latter can expose the true DGP. In Chapter 2, I deduce the condition, i.e., the invertibility condition, necessary for this situation to hold. This condition is used to distinguish between the two scenarios I have explained above. Furthermore, I derive the corresponding VAR and moving average (MA) representations of the FIA and LIE models in Chapter 2 to illustrate the connections among the state space, VAR and MA representations. The MA representation describes the observables as an infinite weighted sum of present

and past (structural) shocks driving the system. Impulse responses can directly be derived from the MA representation. The closing section of Chapter 2 is devoted to the prototypical business cycle model, which serves as benchmark for the DGP in the simulation studies conducted in the following chapters. The model is presented in a rather general form but will be specified in the relevant chapters to perform the role of the FIA model.

The subsequent chapters are divided into two parts. In Chapters 3 and 4, I consider the scenario in which the information sets of the econometrician and the agents are equal, whereas Chapters 5 to 8 consider the scenario in which equivalence between both information sets is not satisfied. First, Chapter 3 introduces subspace algorithm analysis as an alternative to the VAR approach for structural estimation. The focus of the chapter is more on explaining its practical application rather than a detailed description of its statistical background. Particular attention is paid to the context of cointegration and the adapted version of the subspace algorithm as designed by Bauer and Wagner. Afterwards, I apply the subspace algorithm cointegration analysis in a Monte Carlo simulation to compare it to its VAR counterpart, namely, the vector error correction model (VECM). In a first experiment, the workhorse model in Chapter 2 is used to build the DGP. In the second exercise, I resort to a more complex model that can be viewed as comprehensive expansion of the workhorse model. The expanded model is attributable to Beaudry et al. (2011) and motivates the empirical examination in Chapter 4.

The idea behind the model by Beaudry et al. (2011) is to picture the market rush behavior of firms, which receive a signal about new market opportunities in the next period. In the model, firms respond to the signal immediately by boosting their investments in new startups to achieve profits in newly opened markets. Beaudry and Portier (2014a) argue that this type of model is an alternative way of interpreting the literature on technological news-driven business cycles. However, as specified by Beaudry et al. (2011), the model does not imply the above-mentioned misalignment between the econometrician and the agents' information sets.[5]

Beaudry et al. (2011) constructed their model to provide a theoretical justification for their examination of US data in which they estimate a structural VECM (SVECM) to identify market rush shocks as triggers of very short-term macroeconomic fluctuations. In Chapter 4, I carry out a comparable investigation using the subspace algorithm cointegration analysis in addition to the SVECM. Moreover, I provide evidence that the market rush interpretation of the structural shocks identified by these methods seems plausible. The results associated with the subspace algorithm method, in particular, appear convincing and constitute a serious alternative to the SVECM, thereby supporting the findings of Beaudry et al. (2011).

Chapter 5 marks the beginning of the second part of my thesis dealing with the situation

---

[5]This is mainly due to the model assumption about the agents' anticipation horizon, which is just supposed to be one period long, i.e., there is only one period between the occurrence of the signal and the materialization.

in which the information sets of the econometrician and agents do not match. Therefore, the econometrician should principally be unable to unveil the structure of the FIA model on the basis of the LIE model. In the literature, this scenario is often called nonfundamentalness (or noninvertibility). In Chapter 5, I offer an explanation of this terminology and show how nonfundamentalness affects the econometrician's analysis and results, e.g., how it can bias the corresponding impulse response functions. To this end, I present analytical and numerical examples based on the workhorse model. Chapter 5 concludes with a discussion of different solutions proposed in the literature to cope with the problem of nonfundamentalness. These approaches aim to expand the econometrician's information set so that there is no disadvantage relative to the information possessed by the agents. Popular examples include adding a large number of further macroeconomic variables or considering specific information variables, such as survey data.

There may not be such a multitude of macroeconomic series or particular information variables available, e.g., because of an insufficient number of observations due to a shorter sample period or lower frequency; thus, I explore other potential solutions in the last three chapters of my thesis. Chapter 6 addresses the question of whether it is possible to extract the necessary information from the results of the structural estimation in order to adequately modify the procedure to correct for bias and to obtain proper findings. Unfortunately, this is not the case, as I will illustrate through analytical and numerical examples.

While the method presented in Chapter 6 can be considered empirically based, Chapter 7 is devoted to the theory-based approach of Mertens and Ravn (2010). In their analysis of US data, the authors modify the standard SVAR/SVECM procedure in order to investigate the impacts of anticipated fiscal shocks. Their approach is characterized by its ability to detect the true impulse responses at least asymptotically. Mertens and Ravn (2010) explain how to adapt the conventional SVAR/SVECM when only a few particular parameters of the underlying FIA model are known. In Chapter 7, I show that this technique can be carried over to subspace algorithm analysis and can be applied to the context of technology-related news shocks. I conduct Monte Carlo simulations on the basis of the workhorse model to compare the performance of the standard with the modified versions of SVAR/SVECM and subspace algorithm analysis. My results indicate that the modified subspace algorithm analysis achieves the best outcome. I also use robustness tests to ascertain whether the improvement due to the modification steps proposed by Mertens and Ravn (2010) depends on certain factors.

The closing chapter, Chapter 8, presents a framework bringing together the techniques of the preceding chapters in order to examine the empirical relevance of news shocks in the US data. I draw on the study by Beaudry and Portier (2014a) and use their original data set. In addition to the SVECM approach of Beaudry and Portier (2014a), I perform a subspace algorithm analysis and apply the theory-based root flipping method, the latter with both the

SVECM and the subspace algorithm. The findings support the hypothesis that technological news shocks can trigger macroeconomic fluctuations in the US to some extent. In particular, the SVECM results provide, in comparison to the subspace algorithm analyses, a more unambiguous picture of the relevance of this type of news shock, including that its dominance is not related to the very short-term horizon. The latter outcome is complementary to the previous finding of Chapter 4, which revealed the superior role of market rushes, as another kind of news shock, in explaining US business cycles over the very short-run. The chapter closes the thesis with a discussion about the challenge to set up a framework in which both types of news shocks coexist and which permits identification of both in empirical applications. Irrespective of this challenge, my thesis contributes to the literature that builds on the view of Pigou (1927) and Keynes (1936) and emphasizes the importance of changes in agents' expectations for business cycle phenomena.

Chapter 2

Theoretical framework

In this chapter, I sketch the general framework that forms the basis for most of the analyses presented in this thesis. In the first section, I present various types of time series representations that follow from an economic model. Following the introductory remarks, I describe the original state space representation of the model. The original state space representation reflects the true dynamics of the agents' behavior in the economy. I label this representation the FIA model. Then, the innovations form, which is an alternative representation associated with the FIA model, is provided in the subsequent subsection. The innovations form is another way of representing a state space system, but in contrast to the original state space representation, it relates to the perspective of an econometrician who seeks to expose the original model when carrying out an empirical analysis. I call this representation the LIE model. The description of the FIA and LIE models is complemented by the MA and VAR representations, which can be derived from them and help establish the link to the empirical tools employed in the thesis. Then, I expound on the condition that ensures that the FIA and LIE models perfectly match. The first section is completed by highlighting important connections between the FIA and LIE models.

The second section of this chapter outlines the workhorse model that will provide the theoretical foundation for several analyses in the subsequent chapters wherein it adopts the role of the FIA model. It can be seen as a prototypical dynamic stochastic general equilibrium (DSGE) model. As I trace different issues and results back to this model throughout the thesis, I introduce the model in a rather simplified version and present its basic components and properties. The workhorse model will be augmented in the following chapters according to the needs of the corresponding subject while keeping its structure as tractable as possible.

2.1   State space models and their representations[1]

2.1.1   Introductory remarks

Formulating a DSGE model usually requires certain elements to be defined. First, one has to characterize the environment in which the agents, i.e., the decision makers in the economy, interact. Second, one has to establish the decision rules that determine the agents' behavior. Last but not least, agents' choices are influenced by the degree of uncertainty they are confronted with. Hence, building a DSGE model needs a specification of the stochastic environment. Combined, these three components finally form a nonlinear system of expectational difference equations.

Following this preparation stage, the next step is to solve the model, i.e., to find the policy functions that describe the dynamic evolution of the economy. As most of these models do not have a closed form solution, numerical methods have to be used to solve them approximately. There are various solution methods available. The well-known procedures by Blanchard and Kahn (1980), Uhlig (1999), Klein (2000) and Sims (2002) rely on the computation of a linear or log-linear approximation around a nonstochastic steady state. Alternatively, linear quadratic approximation methods can be applied (see Anderson et al. (1996) or Díaz-Giménez (1999), among others). In the latter case, optimization problems with quadratic objective functions subject to linear constraints are solved. All these methods yield model solutions in form of linear approximations of the policy functions.[2]

Linear policy functions are highly attractive for the purposes of this thesis, which is why I limit the scope of the thesis to the class of DSGE models for which the solution is mapped onto a linear system of equations. Of course, I thus implicitly impose strong restrictions because I assume that the model class considered here explains economic behavior while ignoring any nonlinearities in the real data. However, focusing on linear dynamic systems allows me to draw direct connections between the theoretical models and the empirical tools I present and develop throughout the thesis. This coherence could not be illustrated in such a clear and demonstrative way using nonlinear systems.

Having obtained a linear(ized) model solution, there are different ways of taking the model to the data. In general, methods for evaluating a DSGE model can be categorized into at least two groups. The first category includes calibration exercises and moment-based estimation

---

[1]In this section, I draw on Fernández-Villaverde et al. (2005, 2007) as well as Hansen and Sargent (2014). See Hannan and Deistler (1988) for a more sophisticated treatment on state space models. Because the latter offer a statistical perspective, my main references are the former authors who provide more economic intuition.

[2]There are also approaches, such as perturbation or projections methods, for obtaining high order approximations. For a detailed description and review of such methods, I refer the reader to, e.g., Aruoba et al. (2006) or DeJong and Dave (2007). In a comparative study of several numerical solution techniques, e.g., second order approximations or value function iterations, Heer and Maußner (2008) find that log-linearization yields the best results in terms of accuracy and adequacy for business cycle analysis when applied to the stochastic growth model with flexible labor, i.e., my workhorse model.

procedures, such as the Generalized Method of Moments (GMM) or Simulated Method of Moments (SMM). These methods involve only a (rather small) subset of observable variables included in the DSGE model. By concentrating on this subset, the aim of the procedures is to match the behavior (in terms of moments) of the selected variables with the behavior of their counterparts in the real data. The second group of methods comprises Maximum Likelihood (ML) and Bayesian estimation. Unlike the first group, these techniques are normally based on a larger set of observables and attempt to appraise a larger set of empirical implications of the chosen DSGE model.

All the preceding methods premise a particular DSGE model upon which they build their analysis. They can be used to test whether the model under investigation is adequately supported by the data. As will be described below, calibration and moment-based procedures aim to match specific empirical objects in order to parameterize the underlying model. ML and Bayesian estimation rest on the assumption that the DSGE model completely characterizes the data from a statistical point of view. Empirical performance is assessed either by defining a second set of empirical targets (as in the case of calibration), by testing hypotheses (as in moment-based and ML approaches) or by comparing conditional probabilities assigned to the model under examination and its alternatives (as in Bayesian estimation).

In a calibration exercise, the model parameters are set to specific values, which can be traced back to micro-level analysis or long-run averages of certain time series (e.g., the labor income share). The empirical target then depends on the question posed in the experiment. For example, one could assess the predictions of the model produced by varying a selected parameter. Matching moments involves estimating the model parameters such that the resulting parameterization leads to the best fit of the underlying economic model regarding the collection of moments chosen (in advance) by the econometrician. An example of a moment-based procedure is the minimization of an objective function that determines the distance between the theoretical moments (as implied by the economic model) and the empirical moments (as estimated by, e.g., a reduced-form VAR) subject to the set of chosen parameters. By testing the statistical significance of the difference between both sets of moments, it can be evaluated whether the model should be rejected as potential DGP.

ML and Bayesian estimation conduct a likelihood analysis. That is, based on the state space representation of the observable data and a distributional assumption for the structural shocks in the model, the likelihood function is initially formed. The likelihood function is the joint probability of observing the data as function of the model parameters. As its name indicates, ML maximizes the likelihood function subject to the parameters, i.e., the objective is to find the parameterization that maximizes the probability that the data could have been generated by the theoretical model. ML estimation represents the classic approach to likelihood analysis – the data are treated as random and the parameters as fixed (but unknown). In particular, the observable data are considered a realization of a random draw

from a sampling distribution that is captured by the likelihood function.

The Bayesian approach presumes the opposite in the sense that the parameters to be estimated are treated as random while the data are fixed. The likelihood function is coupled with a distributional assumption for the parameters to incorporate a priori information about them in order to compute their posterior distribution. From the definition of the posterior distribution, the marginal likelihood associated with the particular model is derived. It can be interpreted as expected value of the likelihood conditioned on the parameters' prior distribution. The empirical plausibility of the model can then be assessed by comparing it with an alternative model in terms of the so-called posterior odds ratio. This ratio relates the marginal likelihood of one model to the marginal likelihood of the alternative, whereby each marginal likelihood is also multiplied by a (prior) probability assigned to the respective model. The evidence provided by the data in favor of one model over another can be evaluated by means of this ratio. The larger the posterior odds ratio, the stronger the evidence against the model represented in the denominator.[3]

A third group of empirical methods, which I treat separately from the previous two categories, is at the core of this thesis. This group contains the estimation methods introduced in the first chapter: (structural) VAR and subspace algorithm analysis. As in the first group, these procedures typically incorporate a small set of (observed) variables. In contrast to all the aforementioned methods, the econometrician who utilizes them does not usually condition on a particular (fully specified) economic model but imposes only a minimum of theoretical structure on her analysis. As a result, her conclusion about which theory gains empirical support from the data might be more general. There is a controversy in the literature about whether this approach is generally capable of exposing the true DGP.[4] This discussion is mainly related to the VAR procedure. A key goal of this thesis is to revisit this debate under various scenarios and to broaden it in the subspace algorithm analysis. Note that, henceforth, whenever I allude to the econometrician, I presume that she applies VAR, subspace algorithm analysis or both.

To acquaint the reader with the debate about whether the true DGP can be adequately captured by VAR (and subspace algorithm) analysis, it is helpful to establish a benchmark representation (i.e., the FIA model) that typifies the theoretical model and to derive the associated representation on which the econometrician's analysis is founded. This allows me to depict the conditions under which the econometrician is, in principle, able to recover the true DGP. My course of action draws on Fernández-Villaverde et al. (2005, 2007), who use state space representations to illustrate the connections between the theoretical model and the econometrician's model. In referring to Fernández-Villaverde et al. (2005), my presentation can be seen as an engineering exercise in which I provide a direct mapping between the parameters

---

[3]A comprehensive overview of the above-mentioned methods and their implementations is given by DeJong and Dave (2007).

[4]See Section 3.1 for more details and references.

of the economic model and those of the econometrician's model.

2.1.2 Fully informed agents model

State space representation

Based on the above explanations, I presume that the solution to a DSGE model can be expressed in terms of a linear policy function. It describes a set of state variables that evolve according to a first order autoregressive process driven by a certain number of economic shocks. Using a definition from dynamic programming, state variables dictate the position of the system before the agents in the economy make their current decisions. I distinguish between endogenous and exogenous state variables. Endogenous states are predetermined variables, i.e., their current value depends on the agents' choices in the previous period. Exogenous states incorporate stochastic processes that are additional sources of model dynamics and cannot be influenced by the agents' decisions. In addition, all other (non-predetermined) endogenous variables can be constructed as linear combination of the states.[5]

A DSGE model can have the following state space representation

$$x_t = \mathcal{A} x_{t-1} + \mathcal{B} w_t \,, \tag{2.1}$$

$$y_t = \Pi x_t \,, \tag{2.2}$$

where the constants are omitted for illustrative purposes. y_t is the $(k \times 1)$ vector of (non-predetermined) endogenous variables, and x_t denotes the $(n \times 1)$ state vector. w_t represents the zero mean $(m \times 1)$ vector of economic shocks satisfying the white noise properties with $E\left(w_t w_t'\right) = I_m$ and $E\left(w_t w_{t-j}\right) = 0$ for integer $j > 0$. The system matrices have the corresponding dimensions in brackets: $\mathcal{A}\left(n \times n\right)$, $\mathcal{B}\left(n \times m\right)$ and $\Pi\left(k \times n\right)$. These matrices are nonlinear functions of the deep model parameters designating, e.g., technology, preferences and endowments. Equation (2.1) is the state equation. Equation (2.2) expresses the linear mapping between (non-predetermined) endogenous and state variables.

Next, I reformulate the state space representation above following the notation of Fernández-Villaverde et al. (2005, 2007). After substituting equation (2.1) into equation (2.2) and setting $\mathcal{C} = \Pi \mathcal{A}$ and $\mathcal{D} = \Pi \mathcal{B}$, I arrive at the system

$$x_t = \mathcal{A} x_{t-1} + \mathcal{B} w_t \,, \tag{2.3}$$

$$y_t = \mathcal{C} x_{t-1} + \mathcal{D} w_t \,, \tag{2.4}$$

---

[5]In this context, Sargent (1989) was one of the first to propose that economic models could be formulated in state space form. He illustrates how an economic model can be mapped onto a setup that enables the application of recursive methods, such as the Kalman filter, to evaluate the model's likelihood function.

where $\mathcal{C}$ and $\mathcal{D}$ represent matrices of dimension $(k \times n)$ and $(k \times m)$, respectively.

The above system of equations represents the original state space. It describes the true dynamics in the underlying economic model. To be precise, the state space system (2.3) and (2.4) can be summarized as a quadruple, $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$, which provides a complete characterization of the agents' behavior in terms of the model parameters. I also adopt the usual assumption in standard DSGE models that all agents know the complete model, including its parameters as well as the current and past realizations of all variables and shocks. Hence, the agents' information set in period t can be defined as $\mathcal{I}_t^A \equiv \{y_\tau, x_\tau, w_\tau \text{ for } \tau \leq t; \mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}\}$. I thus exclude DSGE models that have an imperfect or asymmetric information structure on the part of the agents from my analysis, as these features would complicate my course of action. I call the state space system given in equations (2.3) and (2.4) the FIA model.

The number of variables in y_t is not often equal to the number of economic shocks. Primarily, it holds that $k > m$, implying a singular FIA model. This situation is described in the literature as stochastic singularity. In this case, there are linear combinations of a subset of variables in y_t, which are purely deterministic. If the data were generated by a singular FIA model, they would have a singular variance-covariance matrix, whereas the econometrician's empirical methods suppose a nonsingular variance-covariance matrix. Ingram et al. (1994) illustrate that it is thus impossible for the econometrician to infer the realization of the economic shocks from the realization of the variables in y_t. Note that the problem demonstrated by Ingram et al. (1994) is due to a prediction of a singular FIA model: only a small number of shocks explains a larger number of variables in y_t. This differs from the scenario described in the first chapter, where the existence of a particular type of economic shock, e.g., a news shock, produces misalignment between the information sets of the agents and the econometrician such that the econometrician does not have enough information to identify the shocks. While the latter issue is part of my thesis, I rule out the case of stochastic singularity.[6]

To cope with the problem of stochastic singularity, various proposals exist in the literature. All of these aim to equalize the dimensions of y_t and w_t. The first approach adds error terms to equation (2.2). These error terms can be specified as uncorrelated across variables and therefore constructed as pure measurement errors.[7] Hence, they lack a structural meaning and should simply be viewed as misspecification errors. Ireland (2004) also allows for correlated errors across variables so that they can capture (co)movements in the data, which the underlying theoretical model does not indicate and which do not reflect simple measurement bias.

As an alternative, Ingram et al. (1994) introduce the multiple-shock approach in which the

---

[6]Stochastic singularity does not affect only the empirical methods that I consider in this thesis. Ruge-Murcia (2007) shows that the procedures for matching moments and likelihood analysis are also vulnerable to stochastic singularity.

[7]See, e.g., Altug (1989), Sargent (1989), McGrattan (1994), Hall (1996) or McGrattan et al. (1997).

dimension of w_t is enlarged by augmenting the theoretical model with additional economic shocks until there are as many economic shocks as variables in y_t.[8] This extended model facilitates further sources of stochastic uncertainty, which might help explain variations in the data. A disadvantage of this procedure is that the extended model does not retain the structure of the original model, as appending economic shocks normally increases the state dimension in contrast to the addition of measurement errors.[9]

While the previous approaches expand the vector w_t, the third one decreases the dimension of y_t by dropping some of the variables (see, e.g., Lubik and Schorfheide (2004) or Smets and Wouters (2007)). Thus, it is presumed that there is a one-to-one mapping between the model variables and corresponding (perfectly measured) indicators in real data. In the following, I allude to this approach when I base my analysis on the subsequent assumption in order to exclude stochastic singularity from the forthcoming subjects:

Assumption 1. *The number of variables in y_t equals the number of economic shocks w_t, i.e.,*
$$k = m, \text{ and the matrix } \mathcal{D} \text{ has full rank.}[10]$$

MA representation

For my purposes, I continue with another assumption at first. The assumption is related to the eigenvalues of the system matrix $\mathcal{A}$. These eigenvalues are decisive for the dynamic properties of the state process (as given by x_t) and, as a consequence, for the dynamic properties of the process given by y_t. The following assumption ensures that (x_t and hence) y_t describes a stationary process:

Assumption 2 (temporary). *All eigenvalues of $\mathcal{A}$ are less than one in modulus.*

I consider *Assumption 2* temporary throughout the thesis because I will also address nonstationary processes in the subsequent chapters. The assumption is later relaxed to allow for eigenvalues of $\mathcal{A}$ equal to one such that the process captured by y_t is integrated of order one. In Chapter 3, I show how to derive the corresponding vector MA and VAR representation of (the first difference of) y_t when $\mathcal{A}$ has unit eigenvalues. To ease the explanations in the remaining paragraphs of this subsection, y_t is considered stationary at this point.

The vector MA representation of the FIA model is obtained as follows. I write the state equation as $x_{t-1} = \left(L^{-1} - \mathcal{A}\right)^{-1} \mathcal{B} w_t$ and substitute into equation (2.4), whereby the lag operator L shifts the respective variable(s) one period backward (or forward, in the case of L^{-1}):

$$y_t = \left[\mathcal{D} + \mathcal{C} \left(L^{-1} - \mathcal{A}\right)^{-1} \mathcal{B} \right] w_t . \tag{2.5}$$

---

[8]See Leeper and Sims (1994), Smets and Wouters (2003) or Adolfson et al. (2008), for instance.
[9]Refer to Ruge-Murcia (2007) and Tovar (2009) for a deeper discussion of these issues.
[10]Notice that this assumption is not necessary for the derivation of the LIE model.

This equation can also be formulated in more convenient form

$$y_t = \left[\mathcal{D} + \mathcal{C} \left(I_n - \mathcal{A}L \right)^{-1} \mathcal{B}L \right] w_t. \tag{2.6}$$

Thus, the MA representation describes y_t as a linear combination of all the economic shocks that occurred in the current and (infinite number of) past periods. I use the MA representation to picture the dynamics of the model in terms of impulse response functions. The impulse responses can be extracted from the coefficients in the MA operator $[\mathcal{D} + \mathcal{C} \left(I_n - \mathcal{A}L \right)^{-1} \mathcal{B}L]$.[11]

VAR representation

Contrary to the MA representation, the VAR representation of the FIA model does not always exist. The VAR representation can be derived from the MA representation in equation (2.6) under the condition that the MA lag operator is invertible in nonnegative powers of the lag operator L. In technical terms, this is formulated as the condition that the roots of the characteristic polynomial $det \left[\mathcal{D} + \mathcal{C} \left(zI_n - \mathcal{A} \right)^{-1} \mathcal{B} \right]$ are less than one in absolute value, where $det \left[\bullet \right]$ denotes the determinant, and z stands for a scalar that can take any complex value. I refer the reader to the appendix for more details on the terminology of roots and polynomials. By applying the rules related to the algebra of partitioned matrices, I also show in the appendix that the aforementioned roots equal the eigenvalues of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$.[12] Therefore, I can document the necessary condition for the existence of the VAR representation of the FIA model, whereby I utilize the terminology of Fernández-Villaverde et al. (2007) to state the

(poor man's) invertibility condition. *All eigenvalues of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$ are less than unity in modulus.*

Only if the *invertibility condition* holds, the FIA model has an (infinite) order VAR representation, which can be determined by inverting the MA operator in equation (2.6). This is essential for the econometrician who conducts VAR analysis. Even if the econometrician is not reliant on VAR methods, the *invertibility condition* is crucial. To improve the reader's sense about this, I present the LIE model in the next subsection and discuss the implications of the *invertibility condition* afterwards.

---

[11]Using the power series expansion with respect to equation (2.6) may be helpful, which yields $y_t = \mathcal{D}w_t + \sum_{i=0}^{\infty} \mathcal{C}\mathcal{A}^i \mathcal{B}w_{t-1-i}$. See the appendix for general explanations of the terminology used in this section.

[12]In the appendix, I also differentiate between the characteristic polynomial and the lag polynomial. Using the above MA representation as example, I would write the associated MA lag polynomial as $det \left[\mathcal{D} + \mathcal{C} \left(I_n - \mathcal{A} \right)^{-1} \mathcal{B}L \right]$. Consequently, the roots of the lag polynomial are the reciprocals of the roots corresponding to the characteristic polynomial.

2.1.3 Limitedly informed econometrician model

State space representation

In this subsection, I introduce the LIE model, which is directly associated with the FIA model. The econometrician who seeks to detect the FIA model (or at least its dynamics) is confronted with a limited information set in relation to the agent because the state vector comprises unobservable variables. Thus, the econometrician can only retrieve information from y_t. In that sense, y_t can be interpreted as vector of observables and equations (2.2) and (2.4) as observation equations (of the FIA model).

The LIE model is a way of representing the observable process y_t from the perspective of the econometrician. The derivation of the LIE model requires computations that are inherent in the Kalman filter algorithm, which originates in the work by Kalman (1960). The eponymous Kalman filter is applied to the data for various purposes. For example, it can be used to estimate the parameters of the system matrices of the FIA model when conducting a ML estimation procedure.[13] Alternatively, Hansen and Sargent (2014) explain how to use the computational steps of the Kalman filter to deduce the innovations form from which I derive the LIE model.

I concentrate on a key step of the Kalman filter algorithm: the estimation of the unobserved state by incorporating the underlying FIA model structure and information that can be filtered out from the observations. Thereby, the state is calculated as the linear least squares projection from the available information, i.e., $\widehat{x}_t \equiv \widehat{E}\left[x_t \mid Y^t\right]$, where $\widehat{E}\left[\bullet\right]$ denotes the linear least squares projection operator, and Y^t is the set of observations in the history of y_t, including period t. It is then possible to estimate the state conditional on current and past values of y_t as a weighting average between the prediction of the model and a correction from the observations:

$$\widehat{x}_t = \mathcal{A}\widehat{x}_{t-1} + K_t\left(y_t - \mathcal{C}\widehat{x}_{t-1}\right), \tag{2.7}$$

where K_t represents the so-called Kalman gain, which weights the correction term in this equation. It is calculated as

$$K_t = \left(\mathcal{A}\Sigma_{t-1}\mathcal{C}' + \mathcal{B}\mathcal{D}'\right)\left(\mathcal{C}\Sigma_{t-1}\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}, \tag{2.8}$$

where Σ_t is defined as $E\left\{\left(x_t - \widehat{x}_t\right)\left(x_t - \widehat{x}_t\right)'\right\}$ and can be interpreted as the prediction error variance of the state. It is updated via the recursive formula

$$\begin{aligned}\Sigma_t = {}& \mathcal{A}\Sigma_{t-1}\mathcal{A}' + \mathcal{B}\mathcal{B}' \\ & - \left(\mathcal{A}\Sigma_{t-1}\mathcal{C}' + \mathcal{B}\mathcal{D}'\right)\left(\mathcal{C}\Sigma_{t-1}\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}\left(\mathcal{A}\Sigma_{t-1}\mathcal{C}' + \mathcal{B}\mathcal{D}'\right)'.\end{aligned} \tag{2.9}$$

---

[13]Hamilton (1994) provides a corresponding explanation.

This equation is called a matrix Riccati difference equation.[14]

As for the definition of $\widehat{x}_t$, I define the analogous prediction of y_t conditional on its own history as $\widehat{y}_t \equiv \widehat{E}\left[y_t \mid Y^{t-1}\right]$. Next, defining the one period ahead forecast error in y_t as $a_t \equiv y_t - \widehat{y}_t$, it is straightforward to see that the correction term in equation (2.7) is equal to a_t, i.e., $a_t = y_t - \mathcal{C}\widehat{x}_{t-1}$.[15] In rearranging the latter equation and replacing the correction term in (2.7), I can build the system that Hansen and Sargent (2014) call the innovations representation:

$$\widehat{x}_t = \mathcal{A}\widehat{x}_{t-1} + K_t a_t\,, \tag{2.10}$$

$$y_t = \mathcal{C}\widehat{x}_{t-1} + a_t\,, \tag{2.11}$$

with $E\left(a_t a_t'\right) = \mathcal{C}\Sigma_{t-1}\mathcal{C}' + \mathcal{D}\mathcal{D}'$. The innovations representation results from the recursions given in the equations (2.7)–(2.9), which can be initialized from arbitrary values of $\widehat{x}_0$ and Σ_0. The LIE model represents the time-invariant version of the innovations representation.

To obtain the time-invariant version of the state space system in equations (2.10) and (2.11), I have to set up appropriate conditions under which the recursions converge to a steady state as $t \longrightarrow \infty$. These conditions are common in the linear quadratic Gaussian control literature, where a quadratic return function is optimized subject to a linear transition law. There is a vast literature dealing with the conditions that lead to the existence of solutions to linear regulator problems, including terms such as detectability, stabilizability, reachability, controllability and observability. Anderson and Moore (1979) provide an overview of these attributes.[16] Ljungqvist and Sargent (2004) show that the optimal linear regulator problem and Kalman filtering are equivalent concepts. I therefore adopt certain aspects of control theory in the following.

The Kalman filter is time-invariant and asymptotically stable if a constant or limiting solution to the Riccati difference equation (2.9) exists. The conditions are usually imposed in terms of definitions related to the system matrices. For my purposes, I formulate the following definitions:[17]

Definition 1. *The pair $(\mathcal{A}, \mathcal{C})$ is observable if the rank of $\begin{pmatrix} \mathcal{C}' & \mathcal{A}'\mathcal{C}' & \ldots & (\mathcal{A}')^{n-1}\mathcal{C}' \end{pmatrix}'$ is equal to the number of states n.*

Definition 2. *The pair $(\mathcal{A}, \mathcal{B})$ is reachable if the rank of $\begin{pmatrix} \mathcal{B} & \mathcal{A}\mathcal{B} & \ldots & \mathcal{A}^{n-1}\mathcal{B} \end{pmatrix}$ is equal to the number of states n.*

---

[14]Its derivation is presented in the appendix.

[15]Replace y_t in the definition of $\widehat{y}_t$ by the right-hand side of equation (2.4), i.e., $\widehat{y}_t = \widehat{E}\left[\mathcal{C}x_{t-1} + \mathcal{D}w_t \mid Y^{t-1}\right]$ and recall the definition of $\widehat{x}_t$. It follows that $\widehat{y}_t = \mathcal{C}\widehat{x}_{t-1}$.

[16]See also the discussion in Anderson et al. (1996).

[17]Note that the more common term is controllability rather than reachability. These attributes are equivalent only if the matrix $\mathcal{A}$ is nonsingular (see Aoki (1987), p. 59). Because I also want to allow for singular $\mathcal{A}$, I use the term reachability in the following. Observability means that for a given output sequence $\{y_t\}$, one can uniquely reconstruct the initial state at finite time. Reachability says that the system can be transferred to any desired state target for a given input sequence, i.e., here $\{w_t\}$ for a finite time (see Gu (2012)).

The properties in *Definitions 1* and *2* guarantee a unique limiting solution to the Riccati difference equation (2.9).[18] These attributes also ensure that the dimension of the state vector in the FIA model is minimal, i.e., that no superfluous state variable has an effect on the dynamics of the system.[19] The minimum number of state variables (i.e., n in *Definitions 1* and *2*) is substantial for the development of the subspace algorithm analysis in Chapter 3.

Given the *Definitions 1* and *2*, I make the following assumption:

Assumption 3. *The pair $(\mathcal{A}, \mathcal{C})$ is observable, and the pair $(\mathcal{A}, \mathcal{B})$ is reachable.*

As a result of this assumption, I can determine the unique limiting solution $\overline{\Sigma}$ in terms of the system matrices as the solution to the algebraic Riccati equation

$$\Sigma = \mathcal{A}\Sigma\mathcal{A}' + \mathcal{B}\mathcal{B}' - (\mathcal{A}\Sigma\mathcal{C}' + \mathcal{B}\mathcal{D}')(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}')^{-1}(\mathcal{A}\Sigma\mathcal{C}' + \mathcal{B}\mathcal{D}')' . \tag{2.12}$$

This result is accompanied by the fact that all eigenvalues of $(\mathcal{A} - K\mathcal{C})$ are less than one in modulus,[20] where K is the steady state Kalman gain that fulfills

$$K = \left(\mathcal{A}\overline{\Sigma}\mathcal{C}' + \mathcal{B}\mathcal{D}'\right)\left(\mathcal{C}\overline{\Sigma}\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1} . \tag{2.13}$$

In the literature, the result with respect to the eigenvalues of $(\mathcal{A} - K\mathcal{C})$ is often stated as the following condition:

Minimum phase condition. *All eigenvalues of $(\mathcal{A} - K\mathcal{C})$ are less than unity in modulus.*

Hence, I obtain the time-invariant version of the innovations representation

$$\widehat{x}_t = \mathcal{A}\widehat{x}_{t-1} + Ka_t , \tag{2.14}$$

$$y_t = \mathcal{C}\widehat{x}_{t-1} + a_t , \tag{2.15}$$

with $\Omega_a \equiv E\left(a_t a_t'\right) = \mathcal{C}\overline{\Sigma}\mathcal{C}' + \mathcal{D}\mathcal{D}'$. I denote the state space representation in equations (2.14) and (2.15) as the LIE model (in reduced form) because it reflects the bounded state of knowledge of the econometrician. It can be summarized as the quadruple $(\mathcal{A}, \mathcal{C}, K, \Omega_a)$. In Chapter 3, I show that the elements of this quadruple can be estimated by the subspace algorithm analysis.

The econometrician only observes y_t and does not see the state. She can only forecast the state conditioned on the history of y_t (and an initial state). Thus, her information set is exclusively governed by the observations y_t and the corresponding forecast errors a_t: $\mathcal{I}_t^E \equiv \{y_\tau, a_\tau \text{ for } \tau \le t; \mathcal{A}, \mathcal{C}, K, \Omega_a\}$. Of course, the econometrician's information set is also limited

---

[18]The reader is directed to d'Andréa Novel and Lara (2013) for a corresponding proof in which they presume the initial matrix Σ_0 to be positive semidefinite.

[19]A proof is given by Gu (2012), for instance.

[20]See also the proof by d'Andréa Novel and Lara (2013) for further details.

by a finite sample size in practical applications. However, because I use the LIE model as a theoretical construct to analyze the systematic connections between FIA and LIE model, I take $\mathcal{I}_t^E$ to be the information set that is available to the econometrician on the basis of an infinite set of observations y_t. The reason is that, at this stage, I wish to avoid well-known shortcomings of small samples.

The econometrician who intends to detect the dynamics and sources of the true DGP usually proceeds by imposing a structure on her reduced form model based on theoretically founded restrictions. Structural identification implies that one assumes that the reduced form innovations a_t are linear combinations of the structural innovations $\widehat{\varepsilon}_t$ with the identity covariance matrix. This means that the relationship $a_t = \widehat{\mathcal{D}}\widehat{\varepsilon}_t$ holds, where $\widehat{\mathcal{D}}$ denotes the rotation matrix (assumed to have full rank), which has to be determined by sufficient restrictions on the LIE model. Note that it also follows that $\Omega_a = \widehat{\mathcal{D}}\widehat{\mathcal{D}}'$. At this point, suppose that structural identification has already been conducted, so I can write the LIE model in structural form as

$$\widehat{x}_t = \mathcal{A}\widehat{x}_{t-1} + \widehat{\mathcal{B}}\widehat{\varepsilon}_t\,, \tag{2.16}$$

$$y_t = \mathcal{C}\widehat{x}_{t-1} + \widehat{\mathcal{D}}\widehat{\varepsilon}_t\,, \tag{2.17}$$

where $\widehat{\mathcal{B}} = K\widehat{\mathcal{D}}$ and $E\left(\widehat{\varepsilon}_t\widehat{\varepsilon}_t'\right) = I_k$.

The econometrician is successful if the dynamics and shocks she recovers from the LIE model in structural form are equivalent to those of the FIA model, i.e., if the system matrices $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ and true shocks w_t conform to their counterparts in equations (2.16) and (2.17). Before addressing this issue in detail, I conclude this subsection with the MA and VAR representations of the LIE model.

MA representation

As in the case of the FIA model, I refer to (temporary) *Assumption 2* at this point. I repeat the computations and apply them to the LIE model (in structural form) to obtain

$$y_t = \left[I_k + \mathcal{C}\left(I_n - \mathcal{A}L\right)^{-1}KL\right]\widehat{\mathcal{D}}\widehat{\varepsilon}_t\,. \tag{2.18}$$

Following the same algebraic rules as above, it can be verified that the roots of the characteristic polynomial $det\left[I_k + \mathcal{C}\left(zI_n - \mathcal{A}\right)^{-1}K\right]$ equal the eigenvalues of $(\mathcal{A} - K\mathcal{C})$. As the *minimum phase condition* is satisfied due to *Assumption 3*, the MA operator $\left[I_k + \mathcal{C}\left(I_n - \mathcal{A}L\right)^{-1}K\right]$ is invertible in nonnegative powers of the lag operator L so that the structural innovations $\widehat{\varepsilon}_t$ lie in the space spanned by current and past realizations of y_t, i.e., the series $\{\widehat{\varepsilon}_t\}$ can be constructed using the sequence of observations $\{y_t\}$. Equation (2.18) is known as the Wold representation. Its main properties are that the MA lag operator is invertible (in nonnegative

powers of L) and that $\widehat{\varepsilon}_t$ satisfies white noise characteristics.[21] The latter holds because $\widehat{\varepsilon}_t$ is a linear combination of the one period ahead forecast error in y_t based on a linear least squares projection of y_t on its past.

Structural identification is conducted by imposing appropriate restrictions on the MA operator in order to determine the rotation matrix $\widehat{\mathcal{D}}$. Having found the rotation matrix, the structural innovations $\widehat{\varepsilon}_t$ can be computed as well as the associated impulse responses. In this thesis, I consider short-run and/or long-run restrictions, i.e., the instantaneous and/or total impact of shocks on certain variables are constrained for shock identification. I provide several examples in the subsequent chapters. Furthermore, Chapter 3 contains more detailed explanations of these identification strategies.[22]

VAR representation

Due to the invertibility of the MA lag operator in equation (2.18), the associated VAR representation can be achieved by using the algebraic rule given in the appendix:

$$y_t \;=\; \mathcal{C}\left[I_n - \left(\mathcal{A} - K\mathcal{C}\right)L\right]^{-1} K y_{t-1} + \widehat{\mathcal{D}}\widehat{\varepsilon}_t \,. \tag{2.19}$$

In contrast to the FIA model, the VAR representation of the LIE model always exists because of the fulfillment of the *minimum phase condition* (under *Assumption 3*). The econometrician's success in uncovering the FIA model depends on whether the *invertibility condition* is met. In the next subsections, I consider this aspect and show that it is not related only to the econometrician who uses VAR analysis. Indeed, even when using subspace algorithm analysis, the econometrician's success is connected to the *invertibility condition*.

2.1.4 Perfect match

Note that for the econometrician, I have defined $\widehat{x}_t$ as a linear projection of the state on current and past values of y_t. Using the FIA model equations, I can derive an analogous expression in terms of the quadruple $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ for the agents. Solving the observation equation for w_t yields $w_t = \mathcal{D}^{-1}\left(y_t - \mathcal{C}x_{t-1}\right)$. By substituting this expression into the state equation and solving backward, I obtain

$$x_t \;=\; \left(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C}\right)^t x_0 + \sum_{i=0}^{t-1} \left(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C}\right)^i \mathcal{B}\mathcal{D}^{-1} y_{t-i} \,, \tag{2.20}$$

---

[21] For a more detailed description and discussion of the Wold representation, see Rozanov (1967) and Sargent (1987).

[22] There are also other possibilities for identification, which are not taken into account in this thesis. These strategies employ either medium-term or sign restrictions on the impulse responses (see the references provided in Section 3.4). Another option is the identification procedure undertaken by Barsky and Sims (2011) (in the context of news shocks) to which I allude to in Section 8.1.

where x_0 denotes the initial state vector. Hence, the state x_t lies in the space spanned by current and past realizations of y_t and an initial state. If all eigenvalues of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$ are inside the unit circle, the first term on the right-hand side of equation (2.20) vanishes as $t \to \infty$, so the state is perfectly explained only by the history of observations. This is exactly the case in which $x_t = \widehat{x}_t$ holds such that Σ_t is a zero matrix that, therefore, solves the algebraic Riccati equation (2.12).

The foregoing presentation implies that the FIA and LIE models fully coincide (when the econometrician correctly identifies the structural form of the LIE model, i.e., when she finds the matrix $\widehat{\mathcal{D}}$, which is equal to $\mathcal{D}$ in this case). Correspondingly, the *invertibility condition* represents the criterion for whether the econometrician is able to detect the dynamics and sources of the FIA model irrespective of her estimation technique. This statement is based on the fact that I focus on the impulse responses of the FIA model as the true model dynamics. In the case of the subspace algorithm analysis, the econometrician estimates the system matrices of the LIE model in order to compute the MA representation (2.18). In the other case, the econometrician estimates a VAR model on the basis of equation (2.19) and inverts the VAR operator to produce the same MA representation. Therefore, there is no difference between the estimation approaches. Note that this chapter is a theoretical treatment, so I ignore issues that may arise in practical applications here, such as lag truncation or small sample bias.[23] I broach these issues in the chapters that present the simulation and empirical studies.

The perfect match between the FIA and LIE models designates the scenario in which the *invertibility condition* holds such that the econometrician is able to recover the true dynamics (that follow from the quadruple $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$) and the true shocks w_t. Alternatively, I characterize this scenario as the situation wherein the information set of the agents, $\mathcal{I}_t^A$, is equal to the information set of the econometrician, $\mathcal{I}_t^E$. This scenario is covered by Chapters 3 and 4, whereas Chapters 5 to 8 deal with the opposite scenario whose main implications are elucidated in the next subsection.

2.1.5 Factorization identity

If the *(poor man's) invertibility condition* does not hold, the FIA and LIE model do not match completely. Nevertheless, both systems are alternative representations that describe the same observed process y_t. Hence, the covariance generating functions of y_t associated with both models are equal. I use their state space representations to express the corresponding covariance generating functions of y_t in terms of the system matrices.[24]

According to Sargent (1987), the covariance generating function of y_t is defined as a function

---

[23]For example, equation (2.19) represents an infinite order VAR whose estimation is not practically possible. Thus, it is approximated by truncating high ordered lags.
[24]To ease the explanation in this subsection, *Assumption 2* is supposed to hold.

of a complex scalar z as $g_y(z) = \sum_{\tau=-\infty}^{\infty} Cov_y(\tau) z^\tau$, where $Cov_y(\tau) = E\left(y_t y_{t-\tau}'\right)$ for integer τ. When applying this formula to the FIA model (see equation (2.5)), it can be written as $g_y^{FIA}(z) = M(z) \Omega_w M(z^{-1})'$, where $M(z) = \mathcal{D} + \mathcal{C}(z I_n - \mathcal{A})^{-1} \mathcal{B}$, and Ω_w is the covariance matrix of w_t. Because of $E(w_t w_t') = I_k$, I thus obtain

$$g_y^{FIA}(z) = \left[\mathcal{D} + \mathcal{C}(z I_n - \mathcal{A})^{-1} \mathcal{B}\right] \left[\mathcal{D}' + \mathcal{B}'(z^{-1} I_n - \mathcal{A}')^{-1} \mathcal{C}'\right]. \tag{2.21}$$

Because $E(\widehat{\varepsilon}_t \widehat{\varepsilon}_t') = I_k$ and $\widehat{\mathcal{D}}\widehat{\mathcal{D}}' = \mathcal{C}\overline{\Sigma}\mathcal{C}' + \mathcal{D}\mathcal{D}'$, I can analogously calculate the covariance generating function of y_t in the LIE model as

$$g_y^{LIE}(z) = \left[I_k + \mathcal{C}(z I_n - \mathcal{A})^{-1} K\right] \left(\mathcal{C}\overline{\Sigma}\mathcal{C}' + \mathcal{D}\mathcal{D}'\right) \left[I_k + K'(z^{-1} I_n - \mathcal{A}')^{-1} \mathcal{C}'\right]. \tag{2.22}$$

Equating both expressions yields the so-called spectral factorization identity (see Hansen and Sargent (2014))

$$\left[\mathcal{D} + \mathcal{C}(z I_n - \mathcal{A})^{-1} \mathcal{B}\right] \left[\mathcal{D}' + \mathcal{B}'(z^{-1} I_n - \mathcal{A}')^{-1} \mathcal{C}'\right] = \left[I_k + \mathcal{C}(z I_n - \mathcal{A})^{-1} K\right] \left(\mathcal{C}\overline{\Sigma}\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)$$
$$\times \left[I_k + K'(z^{-1} I_n - \mathcal{A}')^{-1} \mathcal{C}'\right]. \tag{2.23}$$

The main message of the factorization identity is that the FIA and LIE models are observationally equivalent. Recall that the econometrician resorting to the observable process y_t faces the task of finding the true underlying DGP (when there are at least two possible candidates).[25] However, she is confined to the LIE model so that she cannot disclose the true process given that the *invertibility condition* is violated. Consequently, the impulse responses she estimates do not match the impulse responses obtained from the FIA model. The same holds for the identification of the shocks.

Of course, the magnitude of deviation depends on the exact specification of the FIA model and cannot be demonstrated at this stage. Regarding the respective shocks, I can provide more detail by rearranging equations (2.16) and (2.17) as

$$\widehat{x}_t = (\mathcal{A} - K\mathcal{C}) \widehat{x}_{t-1} + K y_t, \tag{2.24}$$

$$\widehat{\mathcal{D}}\widehat{\varepsilon}_t = y_t - \mathcal{C}\widehat{x}_{t-1}. \tag{2.25}$$

Substituting the observation equation of the FIA model into the previous equations, I com-

---

[25]Theoretically, a much larger number of other models could have generated the same observable sequence $\{y_t\}$.

bine them with the state equation of the FIA model to obtain

$$
\begin{pmatrix} x_t \\ \widehat{x}_t \end{pmatrix} = \begin{pmatrix} \mathcal{A} & 0 \\ K\mathcal{C} & \mathcal{A} - K\mathcal{C} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ \widehat{x}_{t-1} \end{pmatrix} + \begin{pmatrix} \mathcal{B} \\ K\mathcal{D} \end{pmatrix} w_t , \tag{2.26}
$$

$$
\widehat{\mathcal{D}}\widehat{\varepsilon}_t = \begin{pmatrix} \mathcal{C} & -\mathcal{C} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ \widehat{x}_{t-1} \end{pmatrix} + \mathcal{D}w_t . \tag{2.27}
$$

Violation of the *invertibility condition* implies that $x_t \neq \widehat{x}_t$. Thus, I can use the above system to calculate $\widehat{\varepsilon}_t$ as a function of current and lagged values of w_t. Recall the stability properties of $\mathcal{A}$ and $(\mathcal{A} - K\mathcal{C})$, which allow me to connect equations (2.26) and (2.27) to

$$
\widehat{\mathcal{D}}\widehat{\varepsilon}_t = \left(\begin{pmatrix} \mathcal{C} & -\mathcal{C} \end{pmatrix} \left(I_k - \begin{pmatrix} \mathcal{A} & 0 \\ K\mathcal{C} & \mathcal{A} - K\mathcal{C} \end{pmatrix} L \right)^{-1} \begin{pmatrix} \mathcal{B} \\ K\mathcal{D} \end{pmatrix} L + \mathcal{D} \right) w_t . \tag{2.28}
$$

This equation contains a central result. The innovations to the econometrician's information set correspond to "old news" for the agents in the model economy as long as the *invertibility condition* does not hold. Equation (2.28) is therefore a way of describing the scenario in which the information set of the agents is not equal to the information set of the econometrician. This scenario is explored in the second part of the thesis in which I present several examples and approaches that are conjectured to help the econometrician solve this problem. To conduct such an examination, I rely on, inter alia, various versions of a business cycle model that serves as the FIA model, as introduced in the following section.

2.2 Workhorse model

2.2.1 Introductory remarks

In this second section of the chapter, I introduce the prototypical DSGE model. This model is a simple real business cycle (RBC) model, such as those found in standard textbooks.[26] Pioneering work has been conducted by Kydland and Prescott (1982) and Long and Plosser (1983), who launched a new category of dynamic equilibrium models. They focus on a stylized economy in which supply-side shocks are the driving forces of macroeconomic fluctuations, and there are no imperfections, asymmetries or sources of friction. The model primarily indicates that business cycles result from optimal market behavior stimulated by exogenous (technological) shocks. In general, a standard RBC model can be seen as merely

---

[26]Usual references for canonical RBC models are the articles by Hansen (1985) and King et al. (1988).

an extension of the neoclassical growth model that is augmented with a household labor-leisure choice.

Over the years, the RBC model has become a matter of considerable debate because it reflects a rather simplified environment, and it is questioned whether such a model is able to explain real world phenomena.[27] Consequently, many modifications and extensions of the standard RBC model as well as alternative models, e.g., New Keynesian models (NKM), which emphasize the role of nominal rigidities, market imperfections and monetary factors, have been studied.[28] The core of most of these models includes the typical elements of an RBC model, which encourages me to use it as the benchmark model.

The model is presented in a rather compact way by emphasizing its basic elements without specifying concrete functional forms of the representative household's utility function or the firm's production function, for instance. I leave the specifications open, as I use particular modifications or extended versions of the workhorse model in the subsequent chapters. Depending on the subject, I also vary the model's complexity. My objective is to elaborate precise interrelationships between estimation methods and the data generating theoretical model. Therefore, I limit the model's framework to the extent that allows me to explain important implications and results that are not obscured by the intricacy of the model.

The agents in the model economy have rational expectations and at point t symmetric information about all events that have occurred up to that date. The agents are characterized by optimal behavior in the sense that they solve optimization problems based on their rational nature in order to make economic choices. All goods and assets are traded on perfectly competitive markets within a closed economy, and all agents are price-takers. The economy is populated by a continuum of identical infinitely lived households and a large number of identical firms. As they are identical, their economic decisions can be assembled into the optimization problems of a representative household and a representative firm.

The household owns capital stock K_t, which it lends at a rental rate to the representative firm. Furthermore, the household's labor force N_t is rented to the firm at the wage rate. The representative firm uses capital and labor to produce a single good Y_t, which is sold in the final goods market. The price of the final good is the numeraire and thus set to unity. In every period, the firm chooses the amount of capital to use and labor to employ subject to a production function $F(\bullet)$ in order to maximize profits. The production function satisfies the neoclassical properties, i.e., it is strictly increasing in both arguments, strictly concave, twice continuously differentiable, homogeneous of degree one, and it fulfills the Inada conditions.

The representative household solves an intertemporal maximization problem. The household gains utility $u(\bullet)$ from consumption C_t and leisure. The usual assumptions on $u(\bullet)$ are

---

[27]See Rebelo (2005) for a review of this discussion.
[28]The articles in Cooley (1995) review the variation within the RBC framework. Galí (2008) provides a textbook insight into the NKM literature.

that it is strictly increasing in both arguments, twice continuously differentiable and strictly concave. As leisure is typically restricted by a given time endowment, i.e., it reflects the share of time (normalized to one) that is not assigned to work, I use labor N_t instead of leisure as the argument in the utility function. Hence, the instantaneous utility can be specified as $u\left(C_t, N_t\right)$ so that it implicitly takes the time restriction into account. Conditional on its information in the present period, the household maximizes its expected lifetime utility subject to its budget constraint. The household receives income from lending capital and labor. Income is used to finance either consumption of the final good and/or investment I_t to build up the capital stock. The capital stock evolves according to a standard capital accumulation equation.

In the following subsection, I establish the social planner program in order to derive the analytical setting and the corresponding equilibrium conditions. As there are no externalities, distortions or imperfections, the first welfare theorem holds, i.e., perfectly competitive markets entail a Pareto optimal equilibrium. According to the second welfare theorem, the Pareto optimal allocation implies a competitive equilibrium. Therefore, the solution to the social planner program associated with the model is equivalent to the competitive equilibrium.

2.2.2   Social planner program

Suppose that there is a benevolent social planner who maximizes the welfare of the representative household subject to the resource constraints of the economy given some initial conditions. I can formulate her optimization problem as follows:

$$\max_{\{C_t, N_t, K_t\}} \quad E_t \sum_{j=0}^{\infty} \beta^j u\left(C_{t+j}, N_{t+j}\right) \tag{2.29}$$

$$s.t. \quad Y_t = A_t F\left(K_{t-1}, N_t\right), \tag{2.30}$$

$$Y_t = C_t + I_t, \tag{2.31}$$

$$K_t = (1-\delta) K_{t-1} + I_t, \tag{2.32}$$

$$K_0 \text{ given}, $$

where β denotes the discount factor, and A_t is a stochastic process that is described by a law of motion for the exogenous technological progress in the economy. Expression (2.29) represents the expected lifetime utility as a discounted sum of all present and future utilities. Equation (2.30) states the neoclassical production function, depending on capital, labor and technological progress. Equation (2.31) is the aggregate resource constraint, and equation (2.32) is the capital accumulation equation, where δ is the depreciation rate.

The associated first order conditions of that problem can be combined into

$$u_C\left(C_t, N_t\right) = \beta E_t\left[u_C(C_{t+1}, N_{t+1})\left(A_{t+1}F_K(K_t, N_{t+1}) + 1 - \delta\right)\right], \tag{2.33}$$

$$u_N(C_t, N_t) = u_C(C_t, N_t)A_t F_N(K_{t-1}, N_t), \tag{2.34}$$

and supplemented with the transversality condition

$$\lim_{s \to \infty} E_t\left[\beta^{t+s}u_C\left(C_{t+s}, N_{t+s}\right)K_{t+s}\right] = 0.$$

Note that $u_i\left(\bullet\right)$ and $F_j\left(\bullet\right)$ with $i = C,\ N$ and $j = K,\ N$ denote the partial derivatives with respect to the functions' arguments. Equation (2.33) is known as the Euler equation, which indicates that the expected marginal benefit of shifting consumption forward by one period is equal to the marginal costs of abstaining from consumption in the present period. Thus, it determines the optimal consumption-savings choice. According to equation (2.34), the marginal disutility of an extra unit of labor equals the marginal utility of additional consumption due to higher labor compensation. The transversality condition guarantees that no resources are left unconsumed (in the infinite future).

Equations (2.30)–(2.34), in conjunction with a specified law of motion for A_t, form a set of equations that determine the equilibrium dynamics of the model economy. Depending on the functional forms of the utility and production functions as well as on the specification of A_t, further steps (such as detrending or log-linearizing) might be necessary to prepare the equilibrium conditions for solving the model and transforming into the (linear) FIA model form given in equations (2.3) and (2.4). These steps are documented to some extent when establishing the specific model environments in the subsequent chapters.

2.2.3 Decentralized economy

Finally, I illustrate a decentralized interpretation of the basic RBC model presented in the preceding section. This interpretation is helpful when studying the asset price implications of the model and the relevance of the forward-looking behavior of asset prices in the context of news shocks. As the latter will be emphasized in Chapters 5 and 7, I provide the theoretical framework at this stage. My exemplification of the decentralized model economy refers to the well-known article by Jermann (1998), but it can also be found more recently in Gershun and Harrison (2008), for instance.[29]

---

[29]In this context, Danthine and Donaldson (2002) provide an overview of alternative interpretations of the decentralized economy.

In this interpretation, there exists a stock market in which equity shares are traded at price P_t. These securities are claims to the net cash flow stream of the representative firm. The representative households sells and purchases fractions S_t of these assets from period to period and receives dividend payments D_t. The number of equity shares is normalized to one. Contrary to the above situation, the household does not own the capital stock directly. Rather, it undertakes its savings decisions based on its shareholder status. The representative household maximizes lifetime utility, i.e.,

$$\max_{\{C_t, N_t, S_t\}} \quad E_t \sum_{j=0}^{\infty} \beta^j u\left(C_{t+j}, N_{t+j}\right) \tag{2.35}$$

subject to its budget constraint

$$W_t N_t + S_{t-1}\left(P_t + D_t\right) \quad = \quad C_t + S_t P_t. \tag{2.36}$$

The representative firm maximizes its stock market value, i.e., the present discounted value of an infinite sequence of net cash flows to the owners, i.e.,

$$\max_{\{K_t, N_t\}} \quad P_t \quad = \quad E_t \sum_{j=0}^{\infty} \beta^j M_{t+j}\left(Y_{t+j} - W_{t+j} N_{t+j} - I_{t+j}\right) \tag{2.37}$$

subject to the production function (2.30), the capital accumulation equation (2.32) and shareholders' preferences in terms of the stochastic discount factor M_t. The latter is given by

$$M_{t+j} \quad = \quad \frac{u_C(C_{t+j}, N_{t+j})}{u_C(C_t, N_t)}$$

for integer $j \geq 0$, i.e., the owners' marginal rate of substitution. The firm finances new capital completely through retained earnings rather than by issuing new shares. The dividends disbursed to shareholders are given by the period net cash flow of the firm, i.e., the residual of the output value after the wage bill has been paid and investments have been financed:

$$D_t \quad = \quad Y_t - W_t N_t - I_t.$$

It is straightforward to derive and combine the first order conditions of the household and firm's optimization problems to obtain the Euler equation (2.33) and the condition for the optimal labor choice (2.34). In equilibrium, it is assumed that the shareholders hold all outstanding equity claims. Along with the goods market clearing equation, the laws of

motion for capital and technology, and the transversality condition

$$\lim_{s \to \infty} E_t \left[\beta^{t+s} M_{t+s} P_{t+s} \right] = 0 \,,$$

I arrive at the same system of equations, which defines the equilibrium as in the previous subsection.

Chapter 3

Using subspace algorithm cointegration analysis for structural estimation

In this chapter, I focus on the application of the econometrician's structural estimation techniques. These methods involve a relatively small set of observable time series to make inferences about the structure of the underlying DGP. In terms of the terminology of the previous chapter, structural estimation means that the econometrician imposes restrictions on her LIE model motivated by economic theory to identify the sources and dynamics of the agents' behavior, as captured by the FIA model. SVAR is a well-known approach in this context. The main object of this chapter is the introduction of an alternative to the SVAR procedure, namely, subspace algorithm analysis and its extension to the particular framework of cointegrated systems. Because cointegration is a common feature among many macroeconomic time series, so-called subspace algorithm cointegration analysis seems to be another appropriate tool for the empirical investigation of business cycle dynamics.

Note that the starting point in this chapter is the LIE model because it reflects the perspective of the econometrician. Throughout the chapter, I assume that the *invertibility condition* of Chapter 2 is satisfied. Hence, the econometrician's information set $\mathcal{I}_t^E$ is theoretically sufficient to uncover the relevant dynamics and shocks of the FIA model. This means that I consider the scenario in which the LIE model coincides with the FIA model if potential problems, such as those caused by small sample uncertainty or incorrect identification schemes, are ignored.

After presenting the motivation for this chapter in the first section, I provide further explanation for the connection between state space models and VAR representations based on the theoretical setup of Chapter 2 in order to expand the scope to the cointegration framework. Furthermore, I give a description in the third section of how to implement the subspace algorithm because it is less commonly employed than VAR estimation. To lay the foundation for its application in the face of cointegration systematically, I begin by emphasizing the sta-

tionary settings and augment the field to allow for cointegrated variables afterwards. In the fourth section, I present the econometrician's identification strategies for structural estimation, which are used in this chapter and in the rest of the thesis. The chapter is concluded by subjecting the VAR and subspace algorithm analysis to two Monte Carlo simulation studies to compare their performance ability for structural estimation.

3.1 Motivation and related literature

SVAR models have become widely used tools in empirical business cycle research. On the one hand, SVAR estimation of macroeconomic time series may reveal stylized facts that economic models, especially DSGE models, should replicate. On the other hand, the SVAR approach is used to test the empirical validation of a theoretical model. Its basic core is the estimation of structural shocks, i.e., to find the economic shocks, which drive the macroeconomic aggregates as linear combinations of the reduced form innovations in an unrestricted VAR. These structural shocks can be identified by appropriate, i.e., theoretically founded, restrictions. Such restrictions are usually imposed on the short-run and/or long-run effects of the shocks. Moreover, SVARs have been applied to discriminate between different economic models that imply the same theoretical restrictions. In recent years, there has been substantial controversy in the literature as to whether SVAR models can serve this purpose.

Galí (1999)'s empirical finding of a negative impact effect of a technology shock on hours worked intensified the debate about the usefulness of SVAR methods. The most prominent contributions are the papers by Chari et al. (2005, 2008) and Christiano et al. (2007). In these studies, the authors simulate artificial time series using a simple RBC model and estimate SVARs using these artificial data sets in order to compare the true with the estimated impulse responses. The authors investigate potential bias caused by lag truncation or small samples.[1] While Chari et al. (2005, 2008) conclude that SVAR methods with long-run restrictions fail to depict the true impulse responses, Christiano et al. (2007) counter that these findings depend on the chosen model parameters. Under a different parameterization, Christiano et al. (2007) achieve results that favor SVARs regardless of the identification scheme (i.e., over either the short-run or long-run).

In a recent paper, Kascha and Mertens (2009) resume this discussion and let the SVAR approach compete with other estimation methods, notably, vector autoregressive moving average (VARMA) and state space models. They motivate their analysis using the frequently stated argument that VARMA and state space estimation methods outperform VARs because they assemble the complete structure of the underlying DSGE model, whereas VARs

---

[1]For a detailed discussion on lag truncation bias, see Ravenna (2007). Lippi and Reichlin (1993) and Cooley and Dwyer (1998) are early contributions that emphasize the difficulties of low order VAR models, whereas Faust and Leeper (1997) and Erceg et al. (2005) address bias due to the sample size. See also Poskitt and Yao (2016) for a very recent article.

are only finite order approximations of the VARMA process indicated by the structural model. Kascha and Mertens (2009) re-examine the simulation studies by Chari et al. (2005, 2008) and Christiano et al. (2007) but focus on the estimated impact and long-run effect of technology shocks instead. Their findings disclose that a certain subspace algorithm for state space model analysis performs substantially better than the SVAR in terms of mean squared error (MSE) between the true and estimated shock effects. This algorithm originates from the article by Larimore (1983) and is also known as the canonical correlation algorithm (CCA). Due to their results, the authors deduce that problems in identifying structural shocks with long-run restrictions are not exclusively attributed to SVAR models but can affect any of the methods considered if specific properties, such as near-noninvertibility or near-nonstationarity, characterize the underlying DGP.

For nonstationary processes, cointegration plays an important role because it takes possible long-run relationships among nonstationary times series into account. In this context, SVECMs are popular instruments in empirical research (see the seminal paper by King et al. (1991) or more recent articles by Gonzalo and Ng (2001), Breitung et al. (2004) and Beaudry et al. (2011)).[2] SVECMs represent SVARs with nonstationary variables that also allow for cointegration. Likewise, there is an analogous counterpart to the CCA that takes account of cointegration. In a series of papers, Bauer and Wagner (2002, 2003, 2009) establish the concept of subspace algorithm cointegration analysis.[3] Bauer and Wagner (2002, 2003) extend the theoretical framework of state space estimation by including unit roots and modify the standard CCA procedure accordingly. Henceforth, I will denote the modified algorithm as adapted CCA (ACCA). Simulation studies by Wagner (2004) and Bauer and Wagner (2009) show that ACCA performs at least as well as the standard methods of cointegration analysis, in particular, the Johansen (1995) VECM approach.

Based on the above discussion, the aim of this chapter is the usage of ACCA for structural estimation to pursue the subspace algorithm cointegration analysis developed by Bauer and Wagner (2002, 2003, 2009). I show that, after estimating the state space system via ACCA, it is straightforward to derive a vector MA representation of the multivariate time series in first differences, which can be used to identify structural shocks and investigate impulse responses by employing the same identification restrictions as for SVECMs.[4] Furthermore, I follow Kascha and Mertens (2009) in comparing the performance of this structural estimation approach to its standard opponent, i.e., SVECM, in a Monte Carlo simulation. I deviate from their procedure by analyzing the estimated impulse responses and structural shocks

---

[2]One also has to cite Beaudry and Portier (2006) and Beaudry and Lucke (2010). However, because they address the role of anticipated shocks, which is the topic of the second part of this thesis, I do not emphasize their prominence here. I highlight their relevance further in the corresponding chapters.

[3]The published article by Bauer and Wagner (2012) is a more recent and modified version of Bauer and Wagner (2003). See also Wagner (2010) in that context.

[4]This vector MA representation does not restrict the analysis to the usage of ACCA. Alternative state space models for cointegration analysis, such as that of Aoki (1987), could be applied in the same way.

and relate them to their true counterpart in terms of correlation and root MSE (RMSE). My results illustrate that structural ACCA is a serious alternative to the standard estimation technique. Yet, it should be noted that the findings depend on the underlying model parameterization.

3.2 Equivalent representations of the state space system

3.2.1 Stationary processes

In this and the following section, I tie in the representations derived in Chapter 2. I work with the representations stemming from the LIE model but keep in mind that, given a successful structural identification, the econometrician is able to discover the FIA model because the *invertibility condition* is supposed to hold. The objective is to depict the equivalence of the VAR, VARMA and state space representations of the model. I also pay some attention to the VARMA representation to refer to the remarks in Section 3.1, but the focus is on the VAR and state space representations.

In this subsection, I rely on *Assumption 2*, i.e., that $\mathcal{A}$ is a stable matrix, and hence consider a stationary state space system. Recall the infinite VAR representation of the LIE model in reduced form (as given in equations (2.14) and (2.15))

$$y_t = \mathcal{C}\left[I_n - (\mathcal{A} - K\mathcal{C})L\right]^{-1} K y_{t-1} + a_t \tag{3.1}$$

or by using power series expansion

$$y_t = \sum_{i=0}^{\infty} \mathcal{C}(\mathcal{A} - K\mathcal{C})^i K y_{t-1-i} + a_t. \tag{3.2}$$

Furthermore, one can find an algebraically equivalent VARMA representation for y_t (see Zellner and Palm (1974)). I rearrange equation (2.14) to obtain

$$(I_n - \mathcal{A}L)\widehat{x}_t = K a_t. \tag{3.3}$$

Multiplying both sides by the adjoint of $(I_n - \mathcal{A}L)$, i.e., $adj\left[I_n - \mathcal{A}L\right]$ yields

$$adj\left[I_n - \mathcal{A}L\right](I_n - \mathcal{A}L)\widehat{x}_t = adj\left[I_n - \mathcal{A}L\right] K a_t. \tag{3.4}$$

Using $adj\left[I_n - \mathcal{A}L\right](I_n - \mathcal{A}L) = det\left[I_n - \mathcal{A}L\right]$ gives

$$det\left[I_n - \mathcal{A}L\right]\widehat{x}_t = adj\left[I_n - \mathcal{A}L\right] K a_t. \tag{3.5}$$

In multiplying both sides of equation (2.15) by $det\,[I_n - \mathcal{A}L]$ and substituting equation (3.5), I achieve a VARMA(n,n) representation[5]

$$det\,[I_n - \mathcal{A}L]\,y_t \;\; = \;\; \mathcal{C}adj\,[I_n - \mathcal{A}L]\,Ka_{t-1} + det\,[I_n - \mathcal{A}L]\,a_t\,. \tag{3.6}$$

It should be clear that there is equivalence among the state space form in equations (2.14) and (2.15), the finite order VARMA model in equation (3.6) and the infinite order VAR representation in equation (3.2). With respect to the introductory discussion in the previous section, it is not possible to exactly maintain the equivalence of the VAR model in practical applications because it can only be estimated by truncating the autoregressive terms. Moreover, the extent of the related bias depends on the eigenvalues of $(\mathcal{A} - K\mathcal{C})$ (or $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$). The rate of decay is determined by the largest eigenvalue of this matrix in modulus (at least asymptotically). Hence, the closer this eigenvalue is to one, the more lagged VAR terms would be necessary to obtain an appropriate approximation of the infinite order VAR representation. Therefore, estimation methods based on the VARMA or state space form may be superior to the VAR model because they aim at providing an exact description of the underlying DGP, while a VAR model is merely a loose approximation of the complete structure of the DGP.

3.2.2   Nonstationary processes

I restrict the class of nonstationary processes that I discuss to series that are integrated of order one, i.e., $I\,(1)$. Many macroeconomic time series show trending behavior that can be described by unit root processes.[6] Moreover, it is well-known that these time series often move together along a common stochastic trend, so the concept of cointegration provides an appropriate basis for the analysis. To introduce cointegrated processes in my framework, I relax *Assumption 2* and allow for eigenvalues of $\mathcal{A}$ (less than or) equal to one in absolute value. Note that this does not affect the other assumptions of Chapter 2 or the overall presumption in this chapter that the *invertibility condition* holds. To illustrate the connection between the eigenvalues of $\mathcal{A}$ and the cointegration property, I reconsider the VAR representation of the previous subsection.[7] Based on the VAR model, I establish the VECM representation and a related state space system that corresponds to the LIE model.[8]

Recall that solving equation (2.15) for a_t and substituting into equation (2.14) yields equation

---

[5]Aoki and Havenner (1991) provide an alternative way to compute the VARMA representation in equation (3.6) by applying the Cayley-Hamilton theorem.
[6]Nelson and Plosser (1982) were the first to detect the presence of unit roots in US macroeconomic time series.
[7]The derivation of the VAR representation is still applicable because I adhere to the *minimum phase condition*.
[8]In Section 3.4, I also derive a VARMA representation in the case of cointegration.

(2.24). I use the lag operator and solve for $\widehat{x}_t$ to obtain

$$\widehat{x}_t \;=\; [I_n - (\mathcal{A} - K\mathcal{C})\,L]^{-1}\,K y_t\,. \tag{3.7}$$

Substituting this expression back into equation (2.15) leads to the infinite VAR representation as in equation (3.1), which can be written as

$$y_t \;=\; \sum_{i=1}^{\infty} A_i y_{t-i} + a_t \tag{3.8}$$

or more compactly as

$$A\,(L)\,y_t = a_t\,, \tag{3.9}$$

where the VAR operator is $A\,(L) = I_k - \sum_{i=1}^{\infty} A_i L^i$. The coefficient matrices A_i for $i = 1,\ldots\infty$ are defined as given in equation (3.2).[9]

If at least some of the variables in y_t are cointegrated, then the matrix $A\,(1)$ is rank deficient. In particular, its rank r is a positive integer smaller than the number of observable variables k. Hence, there are r cointegrating relations, i.e., r linear combinations of variables in y_t that are $I\,(0)$. In other words, $c \equiv k - r$ common stochastic trends drive the system, i.e., the lag polynomial $det\,[A\,(L)]$ has c roots equal to one, whereas all other roots exceed one in modulus. As a consequence, the VAR process in equation (3.8) is unstable. Due to cointegration, simple differencing would imply misspecification because this would ignore the long-run information that is captured in the levels of the variables.[10]

A standard procedure is to formulate the following VECM representation by subtracting y_{t-1} from both sides of equation (3.8) and rearranging the right-hand side:

$$\triangle y_t = \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \triangle y_{t-i} + a_t\,, \tag{3.10}$$

where $p \to \infty$, $\Pi = -\,(I_k - A_1 - \ldots - A_p)$ and $\Gamma_i = -\,(A_{i+1} + \ldots + A_p)$, $i = 1,\ldots,p-1$. Hence, $rank\,(\Pi) = r$. Then, Π can be decomposed as $\Pi = \tilde{\alpha}\tilde{\beta}'$, where $\tilde{\alpha}$ denotes the $(k \times r)$ loading matrix, and $\tilde{\beta}$ is the $(k \times r)$ matrix of cointegrating vectors. As for the VAR in the stationary setting, the estimation of a VECM requires cutting the number of lagged differences to a finite number. The lag length p is usually quite small in practical applications. However, for the remaining discussion in this section, I presume that the lag length is sufficiently large, so the VECM is (more or less) an exact representation of the state space system.

In the following, I develop the particular form of the LIE model starting at the VAR representation in levels so that the direct link to the cointegration properties can be deduced.

---

[9]I draw the reader's attention to the different notation concerning the state space system matrix $\mathcal{A}$ and the VAR operator $A\,(L)$ as well as its related coefficients A_i.

[10]See, e.g., Lütkepohl (2005) for a detailed introduction to cointegration and VECMs.

I thereby draw on Benes and Vavra (2005). Note that Bauer and Wagner (2003) follow a similar but more elaborate approach. First, define

$$
y_t = \begin{pmatrix} I_k & 0_k & \dots & 0_k \end{pmatrix} \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{pmatrix}
$$

$$
\equiv HY_t \tag{3.11}
$$

so that I can write the VAR representation in levels in the so-called companion form:

$$
Y_t = \begin{pmatrix} A_1 & \dots\dots & A_p \\ I_{k(p-1)} & 0_{k(p-1)\times k} \end{pmatrix} Y_{t-1} + \begin{pmatrix} I_k \\ 0_k \\ \vdots \\ 0_k \end{pmatrix} a_t
$$

$$
\equiv TY_{t-1} + H'a_t. \tag{3.12}
$$

Second, I compute a Jordan decomposition of T such that $T = P\Lambda P^{-1}$, where Λ has block-diagonal structure.[11] Next, I premultiply equation (3.12) by P^{-1} and set $\widehat{x}_t = P^{-1}Y_t$ to obtain

$$
\widehat{x}_t = \Lambda\widehat{x}_{t-1} + P^{-1}H'a_t. \tag{3.13}
$$

Finally, I formulate equation (3.11) using $\widehat{x}_t = P^{-1}Y_t$ and substitute equation (3.13):

$$
\begin{aligned}
y_t &= HPP^{-1}Y_t \\
&= HP\widehat{x}_t \\
&= HP\Lambda\widehat{x}_{t-1} + HPP^{-1}H'a_t \\
&= HP\Lambda\widehat{x}_{t-1} + a_t. \tag{3.14}
\end{aligned}
$$

By setting $\mathcal{A} = \Lambda$, $\mathcal{C} = HP\Lambda$ and $K = P^{-1}H'$, I arrive at the LIE model (see equations (2.14) and (2.15)).

To describe the relationship between the structure of $\mathcal{A}$ and the integration orders of the variables, I draw on some results from the literature. For the sake of exposition, I simply point to the references for details and proofs in the following. Given that the state vector has minimal dimension, the unit root structure of y_t conforms to the unit root structure of the state (see Bauer and Wagner (2003)). Thus, in studying matrix $\mathcal{A}$, I uncover the integration properties of y_t.

---

[11] Further description of Jordan decomposition can be found in the appendix.

Suppose that the eigenvalues in $\mathcal{A}$ are organized in decreasing order. According to Archontakis (1998) and Bauer and Wagner (2003), the Jordan blocks associated with the unit roots are all of size one for $I\,(1)$ processes. This means, in technical terms, that in the case of $I\,(1)$ processes, the algebraic multiplicity is equal to the geometric multiplicity of the eigenvalues of $\mathcal{A}$ equal to one.[12] Therefore, the upper left block of $\mathcal{A}$ is an identity matrix of dimension c. The matrix in the lower right block is associated with the eigenvalues smaller than one in modulus. All other entries are zero. Analogously to the structure of $\mathcal{A}$, the other matrices can be partitioned accordingly:

$$\mathcal{A} = \begin{pmatrix} I_c & 0_{c \times (n-c)} \\ 0_{n-c} & \mathcal{A}_{st} \end{pmatrix}, \ K = \begin{pmatrix} K_1 \\ K_{st} \end{pmatrix}, \ \mathcal{C} = \begin{pmatrix} \mathcal{C}_1 & \mathcal{C}_{st} \end{pmatrix}, \tag{3.15}$$

where the indices 1 and st denote the blocks corresponding to the nonstationary and stationary parts of the system, respectively. The submatrices have the following dimensions in brackets: $\mathcal{A}_{st}\,((n-c) \times (n-c))$, $K_1\,(c \times k)$, $K_{st}\,((n-c) \times k)$, $\mathcal{C}_1\,(k \times c)$ and $\mathcal{C}_{st}\,(k \times (n-c))$. Minimality of the system (i.e., the minimal state dimension) implies that $\mathcal{C}_1$ and K_1 have full rank (see Bauer and Wagner (2003)).

Using the partitioned matrices, the LIE model can be solved under the assumption of zero initial condition of the state:[13]

$$y_t = \sum_{i=1}^{t-1} \mathcal{C}_1 K_1 a_i + a_t + \sum_{j=0}^{\infty} \mathcal{C}_{st} \mathcal{A}_{st}^j K_{st} a_{t-j-1}\,. \tag{3.16}$$

Equation (3.16) corresponds to the Granger representation in the case of the VECM. y_t is decomposed into an $I(1)$ part (i.e., the first term of the right-hand side) and an $I(0)$ part (i.e., the remaining terms of the right-hand side). Because the $I(1)$ component consists of k random walks, the rank of $\mathcal{C}_1 K_1$ determines the number of common stochastic trends c. The orthogonal complement to the column space of $\mathcal{C}_1$, $\mathcal{C}_1^{\perp}$, determines the cointegrating space. The cointegration rank is given by the number of columns of the latter, i.e., $r = k - c$. Hence, $\mathcal{C}_1^{\perp}$ is the analogous element to $\tilde{\beta}$ in the VECM representation.

---

[12]Algebraic multiplicity is the multiplicity of an eigenvalue that solves the characteristic polynomial. Geometric multiplicity is the number of Jordan blocks associated to an eigenvalue. See the appendix for further details and some illustrative examples.

[13]This can be done by recursive substitution of the equation (2.14) into equation (2.15), thereby accounting for the structure of the system matrices as presented in (3.15).

3.3 Estimating linear dynamic systems using subspace methods

3.3.1 Stationary processes

In the literature, several subspace methods are proposed, e.g., numerical methods for subspace state space system identification (N4SID) introduced in Overschee and Moor (1994), the multivariate output error state space (MOESP) method by Verhaegen (1994) and the aforementioned CCA.[14] In this context, Bauer (2005b) suggests using CCA because the other methods do not estimate the system matrices related to the error terms. In addition, Bauer (2005b) highlights its success in econometric applications, such as the estimation of multivariate linear systems. As one of my aims is the identification of structural shocks, CCA seems to be a reasonable choice.

Recall that I start my discussion with stationary processes (i.e., the stability assumption on $\mathcal{A}$ is supposed to hold) to illustrate the idea of subspace algorithms in the following. The extension of the concept to cointegrated systems is based on these derivations and described in the subsequent subsection. In both subsections, I present the basic steps for estimating the system matrices of the LIE model in reduced form. The procedure for determining its structural form is explained afterwards in Section 3.4.

I begin with some remarks on the unique identification of the state space system, the LIE model in reduced form. As shown in the previous section, state space and VARMA systems are equivalent representations of the same underlying process. In general, there are different possibilities of representing the underlying process by means of a state space or VARMA model. This aspect is crucial when it comes to the estimation, which needs a unique representation, i.e., a unique set of parameters, to be specified. With the VARMA representation in the previous section, I implicitly achieved identification because I have formulated the VARMA model in final equations form.[15] Concerning the state space estimation, further restrictions on the parameter matrices are required for unique identification of the state space system. To shed more light on this point, I consider some system theory issues.

In system theory, one speaks of a transfer function that encompasses the underlying mechanism behind the state space system. The transfer function maps input signals onto output signals. The state space system is then the operator that describes this transformation process. In terms of system theory, the transfer function is a z-transform of its impulse response.[16] A specific transfer function admits various state space realizations of different

---

[14] A general review of subspace methods can be found in Viberg (1995).

[15] The VARMA representation in equation (3.6) is in final equations form because it has a scalar VAR operator (multiplied by the identity matrix) and an MA operator, which both are equal to the identity matrix when evaluated at $L = 0$.

[16] In general, a z-transform converts discrete signals (i.e., a discrete sequence of complex values) in the time

state dimensions. Therefore, identification places two demands on the corresponding state space system: it has to be minimal, and a particular basis for the state has to be identified. While minimality of the state dimension has already been presumed in the previous section (due to *Assumption 3*), the second aspect needs further explanation.

As the state is unobservable (in general), its basis can be selected arbitrarily. In other words, there can be different realizations of the same transfer function. To clarify this point, I use recursive substitution to write equation (2.14) as

$$\widehat{x}_t = \mathcal{A}^T \widehat{x}_{t-T} + \sum_{i=0}^{T-1} \mathcal{A}^i K a_{t-i} \,. \tag{3.17}$$

Shift it one period backward and substitute it into equation (2.15) to obtain

$$y_t = \mathcal{C}\mathcal{A}^T \widehat{x}_{t-T-1} + \sum_{i=0}^{T-1} \mathcal{C}\mathcal{A}^i K a_{t-i-1} + a_t \,. \tag{3.18}$$

With $T \to \infty$, equation (3.18) reduces to the vector MA representation

$$y_t = \sum_{i=0}^{\infty} \mathcal{C}\mathcal{A}^i K a_{t-i-1} + a_t \,, \tag{3.19}$$

so the transfer function is directly seen to be $k\left(z\right) = I_k + z\mathcal{C}\left(I_n - z\mathcal{A}\right)^{-1}K$. Given the minimal state dimension, a nonsingular transformation of the state vector would lead to a different realization of the transfer function. This aspect can be demonstrated by defining the transformation of the state as $f_t \equiv F\widehat{x}_t$, resulting in another representation form of the LIE model

$$f_t = F\mathcal{A}F^{-1}f_{t-1} + FKa_t \,, \tag{3.20}$$

$$y_t = \mathcal{C}F^{-1}f_{t-1} + a_t \,, \tag{3.21}$$

with the associated transfer function $k\left(z\right) = I_k + z\mathcal{C}F^{-1}\left(I_n - zF\mathcal{A}F^{-1}\right)^{-1}FK$, which implies the same impulse responses as above.

Therefore, a canonical form of the state space system is necessary, which defines a particular basis for the state and guarantees the uniqueness of the model. There are different available approaches for obtaining a canonical form, e.g., echelon parameterization (see Bauer (2005b)) or stochastic balancing (see Bauer (2005a)). I refer the reader to the stated articles for elaborate explanations of these possibilities. With stochastic balancing, the system iden-

---

domain (described by a difference equation) into complex signals in the frequency domain (described by an algebraic equation). The (z-transform) transfer matrix represents the impulse response matrix of the state space system (see Subsection 6.2.6 in Kwakernaak and Sivan (1972)).

tification is achieved via singular value decomposition (SVD) of a particular matrix, which is inherent in the following estimation procedure.

The crucial idea underlying the subspace algorithm is that, if I know the state, the system matrices of the state space model can be estimated using simple regression techniques. Therefore, the predictive properties of the state are decisive for the success of the algorithm in the estimation. I begin by solving the LIE model equations for y_{t+j} with integer $j > 0$, which yields

$$y_{t+j} = \mathcal{C}\mathcal{A}^{j-1}\widehat{x}_t + \sum_{i=1}^{j-1} \mathcal{C}\mathcal{A}^{i-1} K a_{t+j-i} + a_{t+j}. \tag{3.22}$$

Because $\widehat{x}_t$ and the innovations a_{t+l}, where integer $l > 0$, are uncorrelated by assumption, equation (3.22) shows that the best linear predictor of y_{t+j}, based on the knowledge of an initial state and all past realizations of y_t, is simply a function of the current state $\widehat{x}_t$. Solving the state space system for $\widehat{x}_t$ reveals that the state lies in the space spanned by y_{t-i}, where integer $i \geq 0$, and an initial state for some integer $p > 0$:[17]

$$\widehat{x}_t = (\mathcal{A} - K\mathcal{C})^p \, \widehat{x}_{t-p} + \sum_{i=0}^{p-1} (\mathcal{A} - K\mathcal{C})^i \, K y_{t-i}. \tag{3.23}$$

Defining the following vector of variables $\mathcal{Y}_{t,p}^- \equiv \left(y_t', y_{t-1}', \ldots, y_{t-p+1}' \right)'$, equation (3.23) can be written as

$$\widehat{x}_t = (\mathcal{A} - K\mathcal{C})^p \, \widehat{x}_{t-p} + \mathcal{K}_p \mathcal{Y}_{t,p}^-, \tag{3.24}$$

where $\mathcal{K}_p = \left(K, (\mathcal{A} - K\mathcal{C}) K, \ldots, (\mathcal{A} - K\mathcal{C})^{p-1} K \right)$. Under the *minimum phase condition*, $(\mathcal{A} - K\mathcal{C})^p$ vanishes for large p. Thus, the state can be approximated by $\mathcal{K}_p \mathcal{Y}_{t,p}^-$. Moreover, defining the vector of "future" variables $\mathcal{Y}_{t,f}^+ \equiv \left(y_{t+1}', y_{t+2}', \ldots, y_{t+f}' \right)'$ for some integer $f > 0$, equation (3.22) can then be expressed as

$$\mathcal{Y}_{t,f}^+ = \mathcal{O}_f \mathcal{K}_p \mathcal{Y}_{t,p}^- + \mathcal{E}_f E_{t,f}^+ = \beta_{f,p} \mathcal{Y}_{t,p}^- + \mathcal{N}_{t,f}^+, \tag{3.25}$$

where $\beta_{f,p} = \mathcal{O}_f \mathcal{K}_p$, $\mathcal{N}_{t,f}^+ = \mathcal{E}_f E_{t,f}^+$ with $O_f = \left(\mathcal{C}', \mathcal{A}'\mathcal{C}', \ldots, \left(\mathcal{A}^{f-1} \right)' \mathcal{C}' \right)'$,

$$\mathcal{E}_f = \begin{pmatrix} I & 0 & 0 & \ldots & 0 & 0 \\ \mathcal{C}K & I & 0 & \ldots & 0 & 0 \\ \mathcal{C}\mathcal{A}K & \mathcal{C}K & I & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathcal{C}\mathcal{A}^{f-2}K & \mathcal{C}\mathcal{A}^{f-3}K & \ldots & \ldots & \mathcal{C}K & I \end{pmatrix},$$

---

[17]Note that this is the same exercise as in Chapter 2 with respect to equation (2.20).

and $E_{t,f}^+$ denotes the prediction errors.

The application of the subspace algorithm is based upon equation (3.25). This equation is essential and illustrates the eponymous character of the subspace algorithm. The main idea is to estimate the system matrices of the LIE model based on a projection of "future" observations on the subspace spanned by past realizations. Equation (3.25) motivates the following procedure:[18]

Step 1: Select integers f and p.

Step 2: Regress $\mathcal{Y}_{t,f}^+$ on $\mathcal{Y}_{t,p}^-$ to obtain $\widehat{\beta}_{f,p}$ as an estimator of $\mathcal{O}_f \mathcal{K}_p$.

Step 3: Construct a rank n approximation of $\widehat{\beta}_{f,p}$ with decomposition $\widehat{\mathcal{O}}_f \widehat{\mathcal{K}}_p$ to obtain $\widehat{\mathcal{K}}_p$.

Step 4: Estimate the state $\widehat{x}_t$ by $\widehat{\mathcal{K}}_p \mathcal{Y}_{t,p}^-$.

Step 5: Use equation (2.15) and regress y_t on $\widehat{x}_{t-1}$ to obtain $\widehat{\mathcal{C}}$ and residuals $\widehat{a}_t = y_t - \widehat{\mathcal{C}} \widehat{x}_{t-1}$.

Step 6: Use equation (2.14) and regress $\widehat{x}_t$ on $\widehat{x}_{t-1}$ and $\widehat{a}_t$ to obtain $\widehat{\mathcal{A}}$ and $\widehat{K}$.

For the first step, several criteria are suggested in the literature to determine the lags and leads of observations in the stacked vectors $\mathcal{Y}_{t,p}^-$ and $\mathcal{Y}_{t,f}^+$.[19] The fundamental elements of the procedure are the next three steps. On the one hand, because the true order of the state space system (i.e., the state dimension) is n, it follows that $\mathcal{O}_f \mathcal{K}_p$ has rank n for $f, p > n$. However, the estimator $\widehat{\beta}_{f,p}$ of $\mathcal{O}_f \mathcal{K}_p$ has full rank equal to $min \{f, p\} \cdot k$. Hence, the estimation of $\beta_{f,p}$ requires a reduced rank restriction.

The rank approximation is performed on a transformed matrix $\widehat{\mathcal{W}}_f^+ \widehat{\beta}_{f,p} \widehat{\mathcal{W}}_p^-$ instead of $\widehat{\beta}_{f,p}$ itself, where $\widehat{\mathcal{W}}_f^+$ and $\widehat{\mathcal{W}}_p^-$ are weighting matrices that differ among the various subspace algorithms. In the case of CCA, the choice

$$\widehat{\mathcal{W}}_f^+ = \left(\widehat{\Gamma}_f^+\right)^{-\frac{1}{2}} \text{ and } \widehat{\mathcal{W}}_p^- = \left(\widehat{\Gamma}_p^-\right)^{\frac{1}{2}}$$

with

$$\widehat{\Gamma}_f^+ = \frac{1}{T-p-f+1} \sum_{t=p+1}^{T-f+1} \mathcal{Y}_{t,f}^+ \left(\mathcal{Y}_{t,f}^+\right)' \text{ and } \widehat{\Gamma}_p^- = \frac{1}{T-p-f+1} \sum_{t=p+1}^{T-f+1} \mathcal{Y}_{t,p}^- \left(\mathcal{Y}_{t,p}^-\right)'$$

determines its name, where T is the number of observations. Note that $\left(\widehat{\Gamma}_f^+\right)^{-\frac{1}{2}}$ and $\left(\widehat{\Gamma}_p^-\right)^{\frac{1}{2}}$ represent the inverse of the Cholesky factor of $\widehat{\Gamma}_f^+$ and the Cholesky factor of $\widehat{\Gamma}_p^-$, respectively. As $\widehat{\Gamma}_f^+$ and $\widehat{\Gamma}_p^-$ denote the noncentered sample covariances of $\mathcal{Y}_{t,f}^+$ and $\mathcal{Y}_{t,p}^-$, they are estimators of the canonical correlations of $\mathcal{Y}_{t,f}^+$ and $\mathcal{Y}_{t,p}^-$.

---

[18]See Bauer (2005b) for a related discussion on the asymptotic properties.

[19]In the following applications of the subspace algorithm, I use the criterion suggested by Bauer and Wagner (2002), which will be described in Subsection 3.5.1.

The rank n approximation is obtained via SVD of the transformed matrix $\widehat{\mathcal{W}}_f^+ \widehat{\beta}_{f,p} \widehat{\mathcal{W}}_p^- = \widehat{U} \widehat{\Xi} \widehat{V}'$, where $\widehat{U}$ and $\widehat{V}$ contain the left and right singular vectors, respectively, and $\widehat{\Xi}$ is a diagonal matrix with $min\{f,p\} \cdot k$ singular values in decreasing order as corresponding elements. Replacing $\mathcal{O}_f \mathcal{K}_p$ by $\widehat{\beta}_{f,p}$ leads to the fact that all singular values in the SVD of $\widehat{\mathcal{W}}_f^+ \widehat{\beta}_{f,p} \widehat{\mathcal{W}}_p^-$ are nonzero, whereas a SVD of $\widehat{\mathcal{W}}_f^+ \mathcal{O}_f \mathcal{K}_p \widehat{\mathcal{W}}_p^-$ only results in n nonzero singular values because the rank of $\widehat{\mathcal{W}}_f^+ \mathcal{O}_f \mathcal{K}_p \widehat{\mathcal{W}}_p^-$ is n.

Therefore, knowledge of n is essential. To set its value, Bauer and Wagner (2009) propose an order estimation procedure that relies on minimizing a specific criterion function.[20] For a given choice of n, the SVD can be portioned into two parts:

$$\widehat{\mathcal{W}}_f^+ \widehat{\beta}_{f,p} \widehat{\mathcal{W}}_p^- = \widehat{U} \widehat{\Xi} \widehat{V}' = \widehat{U}_n \widehat{\Xi}_n \widehat{V}_n' + \widehat{R}_n \,,$$

where the first part contains the n dominant singular values in decreasing order and its corresponding left and right singular vectors. All remaining singular values are included in $\widehat{R}_n$ and omitted. The rank n approximation and decomposition of $\widehat{\beta}_{f,p}$ yields

$$\widehat{\beta}_{f,p} = \widehat{\mathcal{O}}_f \widehat{\mathcal{K}}_p = \left(\widehat{\mathcal{W}}_f^+\right)^{-1} \widehat{U}_n \widehat{\Xi}_n \widehat{V}_n' \left(\widehat{\mathcal{W}}_p^-\right)^{-1}$$

so that $\widehat{\mathcal{K}}_p = \widehat{\Xi}_n \widehat{V}_n' \left(\widehat{\mathcal{W}}_p^-\right)^{-1}$. Finally, the implementation of Steps 5 and 6 is straightforward by running simple ordinary least square regressions so that the estimation procedure is completed.

3.3.2 Extension to cointegration analysis

My focus is the application of the subspace cointegration approach rather than its theoretical foundation; thus, I refer the reader to Bauer and Wagner (2003) for a detailed description and simply emphasize the necessary modification of the procedure previously presented. Thereby, I draw on Bauer and Wagner (2002), who show how to adapt the CCA algorithm to allow for cointegration properties.

As in the stationary case, I begin with a remark on system identification. Reconsider the particular state space form of the LIE model and the corresponding solution in equation (3.16), where the system matrices are partitioned in block-diagonal structure. This permits decoupling the nonstationary from the stationary system. Concerning the nonstationary part, the solution only provides the product $C_1 K_1$. Thus, C_1 and K_1 cannot be uniquely determined from this product. In applying a nonsingular transformation of the nonstationary

---

[20]The order estimation criterion is given by $SVC(n) = -log\left(1 - \widehat{\sigma}_{n+1}^2\right) + 2nkH_T/T$, where n, k and T are defined as in the main text. $\widehat{\sigma}_{n+1}$ denotes the $(n+1)$th dominant singular value, and H_T/T is a penalty term, where H_T is set equal to $log(T)$. The smallest argument that minimizes the information criterion determines the (estimated) state dimension (see Bauer and Wagner (2002, 2009) for details).

part by the usage of a nonsingular $(c \times c)$ matrix S, it is obvious that one can find another representation of the nonstationary part, which is observationally equivalent, i.e., there is another realization $(I_c, SK_1, C_1 S^{-1})$ associated with the nonstationary part, which also implies a state space system with matrix $\mathcal{A}$ in Jordan normal form. For this reason, additional restrictions have to be imposed on C_1 and K_1 for unique identification. Bauer and Wagner (2003) assume C_1 to satisfy $C_1' C_1 = I_c$. In this case, an orthogonal complement of C_1 exists with $\left(C_1^{\perp}\right)' C_1^{\perp} = I_r$ and $\left(C_1^{\perp}\right)' C_1 = 0$. This restriction guarantees the uniqueness of C_1 and K_1 when determined from the product $C_1 K_1$ (see the proof of Lemma 2 in Bauer and Wagner (2003)).[21]

Based on the previous remarks, Steps 2 to 4 of the CCA algorithm have to be adapted to take the cointegration properties into account. At first, a consistent estimate of C_1 is obtained by regressing y_t on the first c components of the state estimated by the standard algorithm. Then, a new weighting matrix

$$
\widehat{\mathcal{W}}_{f,C_1}^+ = \left[\left(I_f \otimes \widehat{\overline{\mathcal{C}}} \right) \sum_{t=1}^{T} Y_{t,f}^+ \left(Y_{t,f}^+ \right)' \left(I_f \otimes \widehat{\overline{\mathcal{C}}} \right)' \right]^{-1/2} \left(I_f \otimes \widehat{\overline{\mathcal{C}}} \right)
$$

is calculated, where $\widehat{\overline{\mathcal{C}}} = \left(\widehat{\mathcal{C}}_1, \widehat{\mathcal{C}}_1^{\perp} \right)'$. Because of the new weighting matrix, the matrix of left singular values $\widehat{U}_n$ discussed hitherto needs to be replaced by

$$
\widehat{U}_{n,c} = \begin{pmatrix} I_c & 0_{c \times (n-c)} \\ 0_{(fk-c) \times c} & \widehat{U}_n (2,2) \end{pmatrix},
$$

where $\widehat{U}_n (2,2)$ denotes the lower right block of $\widehat{U}_n$, i.e., the matrix obtained after eliminating the first c rows and columns of $\widehat{U}_n$. Because $\widehat{U}_n$ is a unitary matrix, an alternative way to write $\widehat{\mathcal{K}}_p$ is $\widehat{U}_n' \widehat{\mathcal{W}}_f^+ \widehat{\beta}_{f,p}$. The adapted estimate of the state therefore follows as

$$
\hat{x}_t = \widehat{\mathcal{K}}_p^{adap} \mathcal{Y}_{t,p}^- \quad \text{with} \quad \widehat{\mathcal{K}}_p^{adap} = \widehat{U}_{n,c}' \widehat{\mathcal{W}}_{f,C_1}^+ \widehat{\beta}_{f,p} .
$$

3.4 Structural estimation

The basic goal of the structural estimation is to obtain an estimate of the contemporaneous impact matrix of the structural shocks. In terms of the LIE model in structural form, I denote this matrix as $\widehat{\mathcal{D}}$ (see equations (2.16) and (2.17)). Because the *invertibility condition* is assumed to hold, this matrix should be equivalent to the matrix $\mathcal{D}$ of the FIA model representation. Having found $\widehat{\mathcal{D}}$, it is possible to discover the true dynamics and to detect the

---

[21]Note that unique identification of the stationary part of the system is achieved by SVD of the transformed matrix in the estimation procedure, as already explained in Subsection 3.3.1.

structural innovations $\widehat{\varepsilon}_t$, which are equal to the economic shocks w_t in this case.

There are different strategies to identify the parameters of $\widehat{\mathcal{D}}$. The focus of my thesis are restrictions imposed on the short-run and/or long-run responses of the variables to the structural shocks. However, there are also alternative identification schemes that operate on the medium-term (see, e.g., Uhlig (2004) or Francis et al. (2014)) or sign restrictions (see, e.g., Faust (1998), Canova and Nicoló (2002) or Uhlig (2005)).[22] Widely used identification schemes in the literature are the short-run identification introduced by Sims (1980) and the long-run identification by Blanchard and Quah (1989). A common step is the assumption that the structural shocks are orthogonal to each other. Because $a_t = \widehat{\mathcal{D}}\widehat{\varepsilon}_t$ and I normalize the variances of $\widehat{\varepsilon}_t$ to unity, the relationship

$$\Omega_a = \widehat{\mathcal{D}}\widehat{\mathcal{D}}'$$

yields $k(k+1)/2$ restrictions. Therefore, $k(k-1)/2$ additional restrictions are necessary to determine all parameters of $\widehat{\mathcal{D}}$.

Sims (1980) suggests imposing these restrictions on the impact matrix $\widehat{\mathcal{D}}$ itself. In establishing a recursive ordering of the variables in y_t and using a Cholesky decomposition of the variance-covariance matrix of the reduced form residuals Ω_a, $\widehat{\mathcal{D}}$ is identified as a lower triangular matrix. This short-run identification is already possible at this stage in both the stationary and the nonstationary cases of Section 3.3 because I can estimate the state space system by CCA or ACCA, respectively, and decompose the estimated covariance matrix of the residuals.

The long-run identification is more extensive. I begin with the stationary case. Blanchard and Quah (1989) restrict the structural long-run matrix to be lower triangular, i.e., the $k(k-1)/2$ further restrictions are imposed on the cumulative effects of the shocks. From equation (3.19), I write the MA lag operator as $T(L) = I_k + \sum_{i=0}^{\infty} \mathcal{C}\mathcal{A}^i K L^{i+1}$ to express

$$y_t = T(L)a_t. \tag{3.26}$$

Because $a_t = \widehat{\mathcal{D}}\widehat{\varepsilon}_t$, it follows that

$$y_t = \widehat{M}(L)\widehat{\varepsilon}_t, \tag{3.27}$$

where $\widehat{M}(L) = \widehat{\mathcal{D}} + \sum_{i=0}^{\infty} \mathcal{C}\mathcal{A}^i K \widehat{\mathcal{D}} L^{i+1}$. The total impact matrix of the reduced form innovations is $T(1) = I_k + \mathcal{C}(I_n - \mathcal{A})^{-1} K$. Thus, the total impact matrix of the structural shocks

---

[22]See also recent articles by Kilian (2011) and Fry and Pagan (2011). Kilian (2011) provides an overview of various identification procedures. Fry and Pagan (2011) review sign restrictions from a critical perspective.

equals $\widehat{M}(1) = T(1)\widehat{\mathcal{D}}$. Because

$$\widehat{M}(1)\left(\widehat{M}(1)\right)' = T(1)\Omega_a\left(T(1)\right)',$$

$\widehat{M}(1)$ can be calculated as lower triangular matrix by a Cholesky decomposition of the right-hand side. Then, I can obtain $\widehat{\mathcal{D}}$ as $T(1)^{-1}\widehat{M}(1)$.

Unfortunately, in the nonstationary case of Subsection 3.3.2, $T(1)$ is singular because $\mathcal{A}$ has eigenvalues equal to one. Hence, this long-run identification cannot be employed. For this reason, I show an alternative way to calculate the long-run multiplier matrix that can be used for identification. Furthermore, the cointegration property of the system inherently entails restrictions that have to be taken into account when choosing the (long-run) identification scheme.

Recall equation (3.6), which is derived from the LIE model in reduced form. In the non-stationary case, y_t is supposed to be $I(1)$ with $c = k - r$ common stochastic trends (see also Subsection 3.2.2). Hence, $\mathcal{A}$ has c eigenvalues equal to one, i.e., the lag polynomial $det[I_n - \mathcal{A}L]$ has c roots that lie on the unit circle. This implies that $adj[I_n - \mathcal{A}L]$ has $c - 1$ unit root components so that I can eliminate $c - 1$ unit roots on both sides leaving

$$\widetilde{\mathcal{A}}(L)\triangle y_t = \mathcal{C}\breve{\mathcal{A}}(L)Ka_{t-1} + \widetilde{\mathcal{A}}(L)\triangle a_t, \tag{3.28}$$

where $\widetilde{\mathcal{A}}(L)$ denotes the scalar polynomial determined by the eigenvalues of $\mathcal{A}$ less than one and $\breve{\mathcal{A}}(L)$ represents a matrix polynomial that remains after eliminating the unit roots in $adj[I_n - \mathcal{A}L]$. Finally, multiplication of both sides by $\widetilde{\mathcal{A}}(L)^{-1}$ gives the vector MA representation in first differences

$$\triangle y_t = \left[I_k + \left(\widetilde{\mathcal{A}}(L)^{-1}\mathcal{C}\breve{\mathcal{A}}(L)K - I_k\right)L\right]a_t. \tag{3.29}$$

By writing $T(L) = I_k + \left(\widetilde{\mathcal{A}}(L)^{-1}\mathcal{C}\breve{\mathcal{A}}(L)K - I_k\right)L$, the total impact matrix of the structural innovations is again given as $\widehat{M}(1) = T(1)\widehat{\mathcal{D}}$. In the case of cointegration, $\widehat{M}(1)$ has reduced rank of $c = k - r$. In general, this means that there are r transitory shocks and $k - r$ shocks that have a permanent effect. Thus, r columns of $\widehat{M}(1)$ are restricted to zero, which indicates $r(k-r)$ independent restrictions, so $k(k-1)/2 - r(k-r)$ restrictions are still required for full identification. According to Lütkepohl (2005), $r(r-1)/2$ restrictions that have to be imposed on $\widehat{\mathcal{D}}$ are needed to identify the transitory shocks. The remaining $(k-r)((k-r)-1)/2$ restrictions are placed on the nonzero columns of $\widehat{M}(1)$.[23] Finally, after having determined $\widehat{\mathcal{D}}$, impulse responses directly follow from (3.29).

---

[23]In that context, Amisano and Giannini (1997) propose a scoring algorithm for parameter identification. Moreover, Breitung et al. (2004) show how to use the scoring algorithm for simultaneous short- and long-run identification.

3.5 Monte Carlo experiments

3.5.1 General design

In the following subsections, I conduct two simulation studies that differ in the underlying DGP. As mentioned in the introductory section, structural estimation methods are often used to investigate business cycle fluctuations empirically. Because I present the structural ACCA as an alternative to the SVECM approach, it is appropriate to compare these methods in such a framework. I use typical DSGE models as the basis of the experiments.

The first DGP is defined by a simplified version of the RBC model, which has been introduced in Chapter 2. The idea is that I start with a simple model that allows me to show the relationship between the eigenvalues of $\mathcal{A}$ and $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$ and the deep model parameters. I keep the number of parameters as small as possible so that it is easier to control for the properties of the DGP and to evaluate the performance of the estimation methods. The second DGP ought to deepen the analysis. The DGP is based on a DSGE model recently developed by Beaudry et al. (2011). This model has a far more complex structure than the preceding candidate, so it facilitates the comparison of ACCA and VECM in a more elaborate environment.

Regarding the methodology, my course of action is the same in all exercises. In small sample experiments, I generate 1000 samples of the series $\{y_t\}$ with 200 observations. I then apply structural VECM and ACCA estimation to each sample. The lag length of the VECM is chosen by the Akaike information criterion (AIC).[24] In addition, using the subspace algorithm requires specifying values for the integers f, p and the system order n. Bauer and Wagner (2002) propose to set f and p equal to $2 \cdot \hat{p}_{AIC}$, where $\hat{p}_{AIC}$ denotes the optimal lag length for an autoregressive approximation of y_t determined by the AIC. The system order is estimated by minimizing the specific criterion function given in Subsection 3.3.1.[25]

My goal is to detect the true shocks and to recover the dynamics of the DGP in terms of impulse responses. Depending on the underlying model that serves as the DGP, I impose appropriate restrictions for structural identification. Finally, the comparison of the VECM and ACCA outcomes requires certain criteria. With respect to the identified shocks, I calculate correlations of the estimated residuals and true shocks as well as the associated RMSEs. Furthermore, I compare the estimated impulse responses with the true responses and com-

---

[24]It is well-known in the literature that the lag order can be crucial for the performance of VAR models. Because this discussion is not the goal of my thesis, I point the reader to the publications by Braun and Mittnik (1993) and Kilian (2001) for further insights into this topic. I select the lag order according to the AIC, which is common in practical applications. Furthermore, it can be argued that the AIC is preferable to other criteria, such as the Schwarz criterion, because it tends to suggest a longer lag length (see, e.g., Lütkepohl (2005)), which might be more convenient for approximating an infinite order VAR.

[25]I am grateful to Martin Wagner for sharing his program codes for the subspace algorithm analysis and system order estimation.

pute the corresponding RMSEs.

3.5.2   Data generating process 1

Model description

The neoclassical model economy comprises a standard production function, a log utility function and a conventional capital accumulation equation. I simplify the version presented in Section 2.2 by assuming a fixed labor supply. Thus, the representative agent maximizes expected lifetime utility solely over consumption C_t. The production level of output Y_t varies with the capital input and exogenous technical progress. Capital stock K_t grows by investment I_t minus depreciation. Two stochastic shock processes and a resource constraint complete the model, which can be written as a social planner's problem (given some initial conditions):

$$
\begin{aligned}
\max_{\{C_t,\ K_t\}} \quad & E_t \sum_{j=0}^{\infty} \beta^j \left(\psi_{t+j} log \left(C_{t+j} \right) \right) \\
s.t. \quad & Y_t = A_t K_{t-1}^{1-\alpha}, \\
& Y_t = C_t + I_t, \\
& K_t = (1 - \delta) K_{t-1} + I_t, \\
& log \left(A_t \right) = log \left(A_{t-1} \right) + w_t^A, \\
& log \left(\psi_t \right) = (1 - \rho_\psi) log \left(\overline{\psi} \right) + \rho_\psi log \left(\psi_{t-1} \right) + w_t^\psi,
\end{aligned}
\tag{3.30}
$$

where δ denotes the depreciation rate on capital, and $(1 - \alpha)$ is the production elasticity of capital with $0 < \alpha < 1$. A_t represents the exogenous state of technology, which is a driftless random walk with w_t^A as a technology shock. Preferences are described as a stationary AR(1) process for ψ_t with autoregressive parameter $0 < \rho_\psi < 1$ that measures the degree of persistence of the preference shock w_t^ψ. Both shocks are independently and identically distributed random variables with zero mean and unit standard deviation.

After combining the first order conditions, I compute the solution of the corresponding stochastic second order difference equation in the capital stock

$$
E_t \left[\widetilde{K}_{t+1} \right] + \phi_1 \widetilde{K}_t + \phi_2 \widetilde{K}_{t-1} = \Theta_1 log \left(A_t \right) + \Theta_2 E_t \left[log \left(A_{t+1} \right) \right] + \Theta_3 \widetilde{\psi}_t + \Theta_4 E_t \left[\widetilde{\psi}_{t+1} \right], \tag{3.31}
$$

where "~" denotes the log deviation from the steady state value. The coefficients are derived as follows:

$$
\phi_1 \quad = \quad - \left(1 + \frac{1}{\beta} + \alpha \left(1 - \beta \left(1 - \delta \right) \right) \left(\frac{1 - \beta \left(1 - \delta \right)}{\left(1 - \alpha \right) \beta} - \delta \right) \right),
$$

$$\phi_2 = \frac{1}{\beta},$$

$$\Theta_1 = -\left(\frac{1-\beta(1-\delta)}{(1-\alpha)\beta}\right),$$

$$\Theta_2 = \beta(1-\delta)\left(\frac{1-\beta(1-\delta)}{(1-\alpha)\beta}-\delta\right)+\delta,$$

$$\Theta_3 = \frac{1-\beta(1-\delta)}{(1-\alpha)\beta}-\delta,$$

$$\Theta_4 = -\left(\frac{1-\beta(1-\delta)}{(1-\alpha)\beta}-\delta\right).$$

I solve the stable root backward and the unstable root forward. After some rearrangement, I obtain the policy function in terms of the capital stock

$$log(K_t) = \phi^K + \phi_{KK}log(K_{t-1}) + \phi_{KA}log(A_t) + \phi_{K\psi}\widetilde{\psi}_t, \qquad (3.32)$$

where ϕ_{KK} is the stable root of the solution to the difference equation. Like ϕ^K and all the remaining coefficients in equation (3.32), ϕ_{KK} depends on the deep model parameters. The law of motion for any other endogenous variable Z_t has the form

$$log(Z_t) = \phi^Z + \phi_{ZK}log(K_{t-1}) + \phi_{ZA}log(A_t) + \phi_{Z\psi}\widetilde{\psi}_t. \qquad (3.33)$$

The set of observables comprises consumption and output. Using the decision rules for consumption and output, I can derive the vector MA representation in first differences as

$$\left(\begin{array}{c} \triangle log(C_t) \\ \triangle log(Y_t) \end{array}\right) = \left(\begin{array}{cc} \frac{\phi_{CK}\phi_{KA}}{(1-\phi_{KK}L)}L + \phi_{CA} & (1-L)\left(\frac{\phi_{CK}\phi_{K\psi}}{(1-\phi_{KK}L)(1-\rho_\psi L)}L + \frac{\phi_{C\psi}}{(1-\rho_\psi L)}\right) \\ \frac{\phi_{YK}\phi_{KA}}{(1-\phi_{KK}L)}L + \phi_{YA} & (1-L)\left(\frac{\phi_{YK}\phi_{K\psi}}{(1-\phi_{KK}L)(1-\rho_\psi L)}L\right) \end{array}\right)\left(\begin{array}{c} w_t^A \\ w_t^\psi \end{array}\right).$$

$$(3.34)$$

Notice that $\phi_{Y\psi}$ is zero, and denote the MA operator on the right-hand side of equation (3.34) as $M(L)$. Using the definitions of the various coefficients, the roots of the MA lag polynomial $det[M(L)]$ are calculated as $\frac{1}{1-\delta}$ and one. The unit root results from the cointegrating relationship between consumption and output.[26]

Furthermore, I can express the solution as FIA model representation as

$$\left(\begin{array}{c} \widetilde{K}_t \\ \widetilde{\psi}_t \\ log(A_t) \end{array}\right) = \left(\begin{array}{ccc} \phi_{KK} & \phi_{K\psi}\rho_\psi & \phi_{KA} \\ 0 & \rho_\psi & 0 \\ 0 & 0 & 1 \end{array}\right)\left(\begin{array}{c} \widetilde{K}_{t-1} \\ \widetilde{\psi}_{t-1} \\ log(A_{t-1}) \end{array}\right) + \left(\begin{array}{cc} \phi_{KA} & \phi_{K\psi} \\ 0 & 1 \\ 1 & 0 \end{array}\right)\left(\begin{array}{c} w_t^A \\ w_t^\psi \end{array}\right),$$

$$(3.35)$$

---

[26]I present further remarks on MA roots in the nonstationary case in the appendix. Note that a decision rule for investment follows from the resource constraint.

$$
\begin{pmatrix} \widetilde{C}_t \\ \widetilde{Y}_t \end{pmatrix} = \begin{pmatrix} \phi_{CK} & \phi_{C\psi}\rho_\psi & \phi_{CA} \\ \phi_{YK} & \phi_{Y\psi}\rho_\psi & \phi_{YA} \end{pmatrix} \begin{pmatrix} \widetilde{K}_{t-1} \\ \widetilde{\psi}_{t-1} \\ log\,(A_{t-1}) \end{pmatrix} + \begin{pmatrix} \phi_{CA} & \phi_{C\psi} \\ \phi_{YA} & 0 \end{pmatrix} \begin{pmatrix} w_t^A \\ w_t^\psi \end{pmatrix}. \tag{3.36}
$$

Equation (3.35) is the state equation, and equation (3.36) is the observation equation. The system matrices $\mathcal{A}, \mathcal{B}, \mathcal{C}$ and $\mathcal{D}$ are defined accordingly. Note that the system is minimal, i.e., the rank of the observability and reachability matrix is equal to the number of states n, so it is not possible to formulate the system with a smaller state dimension. The eigenvalues of $\mathcal{A}$ are ϕ_{KK}, ρ_ψ and one. The latter implies that there is one common stochastic trend, i.e., that log consumption and log output are cointegrated. There is only one nonzero eigenvalue of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$, which equals $(1 - \delta)$. Consequently, the invertibility condition is satisfied for the plausible space of values for δ, where $0 < \delta \leq 1$. In the following, the FIA model given in equations (3.35) and (3.36) is used to generate the artificial time series. Then, structural estimation is implemented as described above.

Basic exercises

In the basic exercises, I calibrate the model parameters as follows. α is set to $2/3$ so that the production elasticity of capital is $1/3$. Furthermore, I use $\overline{\psi} = 1$, $\beta = 1/1.01$ and fix ρ_ψ at 0.5. For the depreciation rate, I consider two cases so that I can directly control for the eigenvalue of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$. Variation in the depreciation rate aims at demonstrating the sensitivity of the estimation methods to the noninvertibility region. As will be seen below, a higher depreciation rate increases the speed of adjustment of the variables to the shocks. To put it in the words of Erceg et al. (2005), a rise in the depreciation rate decreases the "endogenous persistence" in the FIA model. My particular values for δ are 0.025 and 0.2. The magnitude of δ also affects the value of ϕ_{KK}, which equals 0.9625 and 0.7882, respectively. Note that ϕ_{KK} is determined irrespective of ρ_ψ. Changing ρ_ψ, i.e., varying the exogenous persistence in the FIA model, is part of the robustness check following the basic exercises.

I start with two exercises (Exercises 3.1 and 3.2) in which I use the fact that the preference shock has no contemporaneous effect on output, as is visible from equation (3.36), and impose a corresponding zero short-run restriction for the structural identification. Exercise 3.1 uses the lower depreciation rate of 0.025 while Exercise 3.2 relies on the higher depreciation rate of 0.2. I then repeat these experiments in Exercises 3.3 and 3.4 but focus on a different identification scheme. I thereby refer to the discussion about the performance of structural VAR and state space models identified with long-run restrictions (see the references given in Section 3.1) and analyze the outcomes for the VECM in comparison to the ACCA in my setting. Because the FIA model specified above also implies that the preference shock has no long-run effect on consumption or output, I employ a corresponding identification strategy in Exercise 3.3 (with the lower depreciation rate) as well as in Exercise 3.4 (with the higher

Table 3.1: EXERCISES 3.1 & 3.2 - SIMULATION RESULTS (BASIC, SHORT-RUN RESTRICT.)

	EXERCISE 3.1: $\delta = 0.025$, $\rho_\psi = 0.5$		EXERCISE 3.2: $\delta = 0.2$, $\rho_\psi = 0.5$	
	VECM	ACCA	VECM	ACCA
CORRELATIONS (SHOCKS)	mean (prct.)	mean (prct.)	mean (prct.)	mean (prct.)
w^A	$0.99 \left(\begin{matrix} 0.94 \\ 1.00 \end{matrix}\right)$	$0.97 \left(\begin{matrix} 0.87 \\ 0.99 \end{matrix}\right)$	$0.99 \left(\begin{matrix} 0.95 \\ 1.00 \end{matrix}\right)$	$0.98 \left(\begin{matrix} 0.92 \\ 0.99 \end{matrix}\right)$
w^ψ	$0.93 \left(\begin{matrix} 0.90 \\ 0.96 \end{matrix}\right)$	$0.93 \left(\begin{matrix} 0.88 \\ 0.96 \end{matrix}\right)$	$0.98 \left(\begin{matrix} 0.95 \\ 0.99 \end{matrix}\right)$	$0.97 \left(\begin{matrix} 0.92 \\ 0.99 \end{matrix}\right)$
RMSE (SHOCKS)				
w^A	0.2025	0.2777	0.1895	0.2428
w^ψ	0.3735	0.3841	0.2449	0.2656
RMSE (IMPULSE RESP.)				
$w^A \rightarrow consumption$	0.1887	0.2009	0.2626	0.2450
$w^\psi \rightarrow consumption$	0.1874	0.1750	0.2043	0.1942
$w^A \rightarrow output$	0.2380	0.2536	0.2652	0.2489
$w^\psi \rightarrow output$	0.1946	0.1699	0.2101	0.1990

Note: The table reports the median correlations (and 5th and 95th percentile correlations in parentheses) between the true shocks and the structural shocks (top block row) identified in VECM and ACCA as well as the associated RMSEs (middle block row). The bottom block row contains the corresponding RMSEs w.r.t. the impulse responses.

depreciation rate).

Table 3.1 reports the results of the first two exercises (Exercises 3.1 and 3.2). I document the correlations between the true shocks and the corresponding identified structural innovations as well as the associated RMSEs. For the correlations, I compute the median over all simulation runs for each shock. The 5th and 95th percentile correlations are given in parentheses. Additionally, the RMSEs for the impulse responses are presented. These RMSEs describe the bias between the true and the estimated impulse responses over a time horizon of 32 periods.

In Exercise 3.1, VECM shows better performance than does ACCA because the identification of the structural shocks is more accurate in terms of correlation and RMSE for both shocks. VECM yields a median correlation estimate between the identified technology shocks and the true ones of 0.99, whereas the median correlation in the case of the ACCA is 0.97. Although this difference is not significant when looking at the confidence bounds given by the 5th and 95th percentiles, the ACCA suffers from higher uncertainty, as indicated by the wider confidence intervals. As for the median correlation concerning the preference shocks, both methods achieve a coefficient of 0.93, but ACCA is outperformed by the VECM in terms of the confidence bounds and the shock-related RMSEs. Furthermore, the estimated responses to the technology shock evolve in a closer match to the true theoretical ones for

the VECM relative to the ACCA according to the RMSEs. The latter outcome is reversed for the estimated responses to the preference shock due to more precise estimates by the ACCA.

Figure 3.1 illustrates the theoretical impulse responses and median estimates associated with ACCA and VECM as well as the corresponding confidence bounds. The latter are computed as the 5th and 95th percentiles of the estimated impulse responses. The estimated responses to the technology shock display a higher degree of uncertainty in the case of ACCA, as indicated by wider confidence intervals. The difference between the two methods with respect to the response to the preference shock is not that visible. Notice that the scale of the vertical axis is decreased in the respective lower right graphs because the true response of output to the preference shock is quite small under the chosen parameterization, especially in relation to the distribution of the estimated responses. In fact, ACCA and VECM do not find a statistically significant response to the preference shock at all. In Exercise 3.2, VECM is slightly better than ACCA in the identification of the structural shocks, whereas ACCA dominates in matching the impulse responses with respect to both shocks.

In comparing the results of both exercises, it is not appropriate to contrast the RMSEs of the impulse responses. The increase in the depreciation rate affects the dynamics of the model, particularly the impulse responses. As capital depreciation is higher in Exercise 3.2, investment has a stronger reaction in order to smooth the consumption path, which also impacts the other endogenous variables, e.g., output. Thus, there are sharper movements in the impulse responses, and it would be questionable to compare the RMSE of the impulse responses because they would be based on different patterns. Therefore, I only focus on the identification of the shocks in the following.

Raising the depreciation rate lowers the eigenvalue of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$ from 0.975 to 0.8, i.e., I move further away from the region of noninvertibility in Exercise 3.2 in comparison to Exercise 3.1. My findings confirm that this affects the estimation results of both methods. The performance of both VECM and ACCA improve in this exercise. Both methods gain precision in terms of correlation and RMSE of the shocks; an increase in the correlations is mainly observed with respect to the preference shock. Moreover, the confidence intervals of the correlations narrow for both technology and preference shocks, i.e., estimation uncertainty is reduced. However, one should be cautious in attributing the result of Exercise 3.2 solely to variation in the MA root. As mentioned above, varying the depreciation rate means changing the dynamics of the underlying model, which is essentially reflected by the stable root in the solution to the difference equation in capital. An increase in δ reduces the value of ϕ_{KK}, which is one of the eigenvalues of $\mathcal{A}$. It implies that I cannot isolate the effect of the MA root completely.

It is worth studying the performance of VECM and ACCA under a different identification scheme. Note that a zero long-run restriction on the preference shock is also valid in my setting. Using the same calibrated parameter values, I conduct exercises analogous to the

Figure 3.1: EXERCISES 3.1 & 3.2 - IMPULSE RESPONSES (BASIC, SHORT-RUN RESTRICT.)
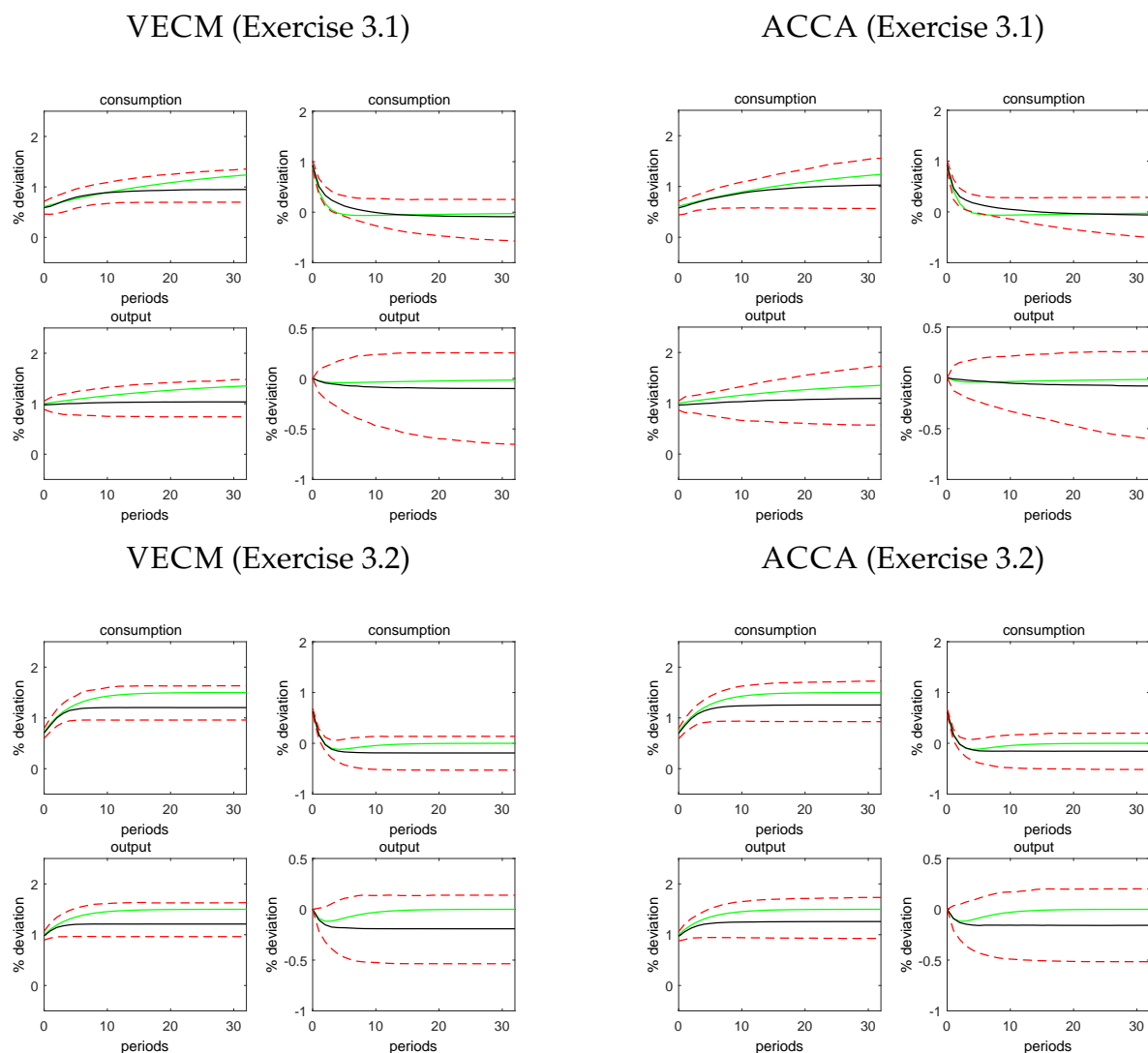
VECM (Exercise 3.1)   ACCA (Exercise 3.1)

VECM (Exercise 3.2)   ACCA (Exercise 3.2)

Note: The figure depicts FIA model impulse responses (in green) and estimated impulse responses (in black). Dashed red lines mark two-sided 90% confidence bounds. Impulse responses to w^A are shown in the left column and impulse responses to w^ψ in the right column of each panel.

Table 3.2: EXERCISES 3.3 & 3.4 - SIMULATION RESULTS (BASIC, LONG-RUN RESTRICT.)

	EXERCISE 3.3: $\delta = 0.025$, $\rho_\psi = 0.5$		EXERCISE 3.4: $\delta = 0.2$, $\rho_\psi = 0.5$	
	VECM	ACCA	VECM	ACCA
CORRELATIONS (SHOCKS)	mean (prct.)	mean (prct.)	mean (prct.)	mean (prct.)
w^A	$0.97 \begin{pmatrix} 0.83 \\ 0.99 \end{pmatrix}$	$0.94 \begin{pmatrix} 0.74 \\ 0.99 \end{pmatrix}$	$0.97 \begin{pmatrix} 0.90 \\ 0.99 \end{pmatrix}$	$0.96 \begin{pmatrix} 0.87 \\ 0.99 \end{pmatrix}$
w^ψ	$0.92 \begin{pmatrix} 0.80 \\ 0.95 \end{pmatrix}$	$0.91 \begin{pmatrix} 0.70 \\ 0.95 \end{pmatrix}$	$0.96 \begin{pmatrix} 0.89 \\ 0.98 \end{pmatrix}$	$0.95 \begin{pmatrix} 0.87 \\ 0.98 \end{pmatrix}$
RMSE (SHOCKS)				
w^A	0.3515	0.4465	0.2744	0.3151
w^ψ	0.4370	0.4748	0.3116	0.3372
RMSE (IMPULSE RESP.)				
$w^A \to consumption$	0.2579	0.2763	0.2509	0.2514
$w^\psi \to consumption$	0.1441	0.1816	0.0735	0.0792
$w^A \to output$	0.3157	0.3546	0.2497	0.2528
$w^\psi \to output$	0.0897	0.1350	0.0695	0.0761

Note: The table reports the median correlations (and 5th and 95th percentile correlations in parentheses) between the true shocks and the structural shocks (top block row) identified in VECM and ACCA as well as the associated RMSEs (middle block row). The bottom block row contains the corresponding RMSEs w.r.t. the impulse responses.

previous ones but implement the zero long-run instead of the zero short-run restriction (see Table 3.2 and Figure 3.2 for the corresponding results). In Exercise 3.3, the VECM approach clearly outperforms ACCA in terms of all categories. A large part of the estimated impulse responses exhibit a high degree of uncertainty in the case of the ACCA resulting in insignificant results with respect to the responses to the technology shock to some extent. As in Exercise 3.2 (compared to Exercise 3.1), the performance of ACCA improves substantially in Exercise 3.4, wherein the depreciation rate is higher, and moves closer to the performance of the VECM, although the VECM is still slightly more accurate.

It can be ascertained that the long-run identification scheme is less successful than the short-run identification scheme, both with ACCA and the VECM, in obtaining the true shocks and impulse responses. This is true without limitation when comparing Exercise 3.1 with Exercise 3.3. A comparison of Exercises 3.2 and 3.4 confirms this outcome with respect to the identification of the shocks. In Exercise 3.4 (compared to Exercise 3.2), the RMSEs of the responses to the technology shock worsen only in the ACCA case, whereas the VECM shows some enhancement. Note that an assessment based on the impulse responses regarding the preference shock is difficult because there is perfect identification only on impact in the one case (Exercise 3.2), whereas there is perfect identification in the long-run in the other case so that uncertainty is much lower in Exercise 3.4 due to the relatively rapid convergence of the system.

Figure 3.2: EXERCISES 3.3 & 3.4: IMPULSE RESPONSES (BASIC, LONG-RUN RESTRICT.)
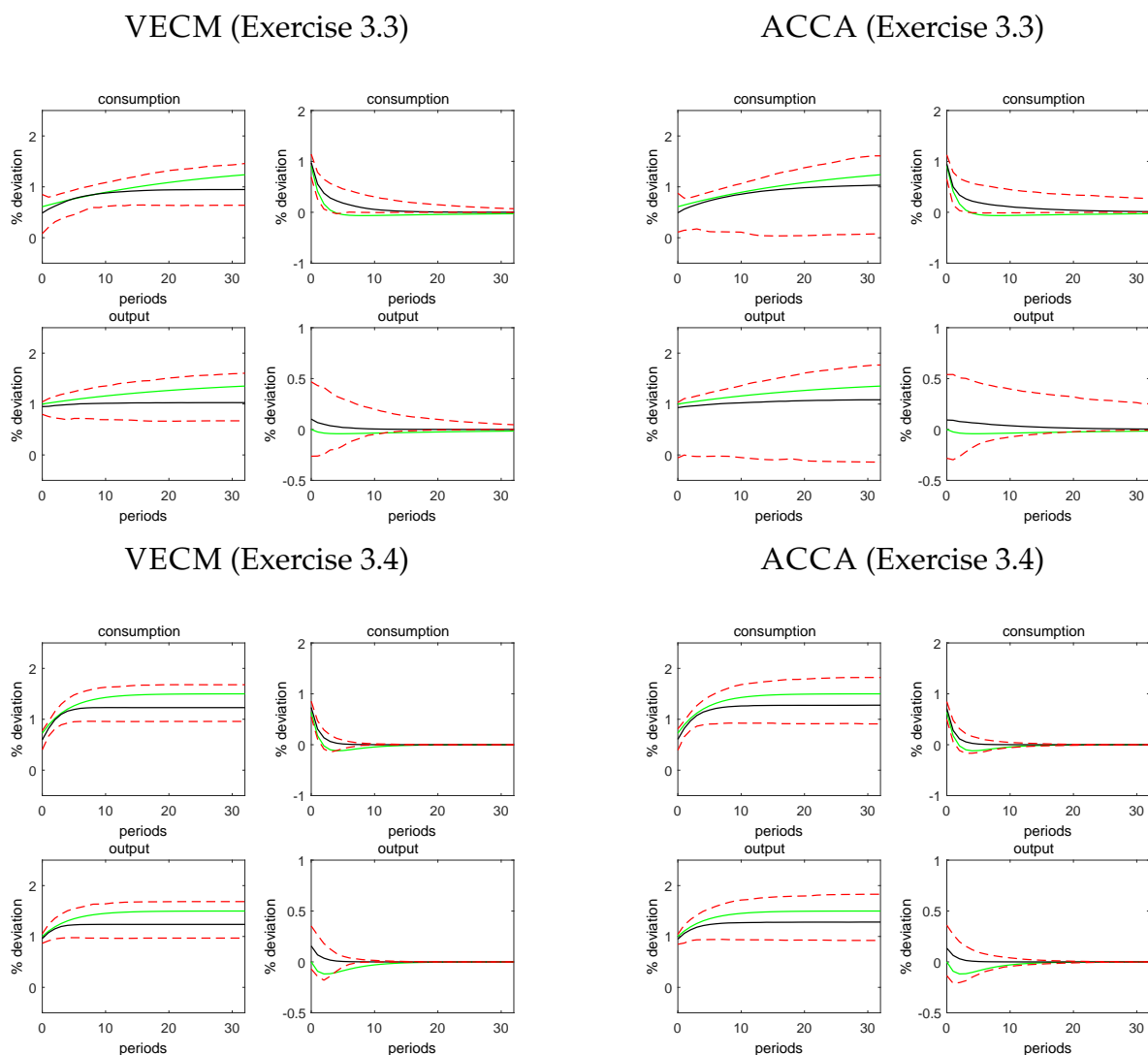
Note: The figure depicts FIA model impulse responses (in green) and estimated impulse responses (in black). Dashed red lines mark two-sided 90% confidence bounds. Impulse responses to w^A are shown in the left column and impulse responses to w^ψ in the right column of each panel.

Robustness exercises

I implement two more exercises, Exercises 3.5 and 3.6, to check the robustness of the above results. The theoretical model I have used so far is highly stylized and exhibits a very simple structure. In the following, I extend the characteristics of the model and allow for variable labor supply. Hence, labor becomes an additional argument of the household and firm's optimization problems. I assume a new functional form of the instantaneous utility function

$$u\left(C_t, N_t\right) \;\; = \;\; log\left(C_t\right) + log\left(1 - \psi_t N_t\right), \tag{3.37}$$

and a production function of the Cobb-Douglas type

$$Y_t \;\; = \;\; K_{t-1}^{1-\alpha}\left(\theta_t N_t\right)^{\alpha}, \tag{3.38}$$

where N_t denotes the new argument. Technological progress θ_t is labor-augmenting and evolves according to a unit root process as modeled previously.[27] The inclusion of variable labor supply augments the household's decision problem. Besides the intertemporal consumption-savings trade-off, the household has to balance consumption and labor within a period. The parameters are calibrated as before, i.e., $\alpha = 2/3$, $\beta = (1.01)^{-1}$ and $\overline{\psi} = 1$. The eigenvalues of $\mathcal{A}$ are the same, i.e., ϕ_{KK}, ρ_ψ and one, as well as the eigenvalue of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$, which equals $(1 - \delta)$. I fix the depreciation rate at $\delta = 0.025$, yielding $\phi_{KK} = 0.953$.

Though I cannot separately control for the eigenvalues of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$, I am able to govern the eigenvalues of $\mathcal{A}$ associated with the exogenous stochastic processes. As I am interested in considering a cointegrating relationship in the model, I keep the unit root in the technology process but use two different persistence levels for the stationary preference shock. While I set ρ_ψ in Exercise 3.5 to the same level (0.5) as above, I increase its value to 0.95 in Exercise 3.6. Thus, I have a quite similar eigenvalue structure in Exercise 3.5 as in Exercise 3.3 (and Exercise 3.1). In Exercise 3.6, I can see the implications when the stationary stochastic process has almost permanent effects. My identification strategy is a zero long-run restriction of the preference shock in both exercises. Note that the extension of the theoretical model no longer permits a zero short-run restriction of the preference shock because output now immediately responds to the preference shock in the FIA model.

The corresponding results of Exercises 3.5 and 3.6 are presented in Table 3.3 and Figure 3.3. In Exercise 3.5, the VECM approach dominates the ACCA with respect to the shock identification, as has already been the case in Exercise 3.3. However, the median correlation and RMSE worsen. Moreover, ACCA provides a better average match of the responses to

---

[27]Note that setting $\theta_t = A_t^{\frac{1}{\alpha}}$ would provide the Hicks-neutral technology as in the simple model with fixed labor.

Table 3.3: EXERCISES 3.5 & 3.6 - SIMULATION RESULTS (EXTENDED, LONG-RUN RESTRICT.)

	EXERCISE 3.5: $\delta = 0.025$, $\rho_\psi = 0.5$		EXERCISE 3.6: $\delta = 0.025$, $\rho_\psi = 0.95$	
	VECM	ACCA	VECM	ACCA
CORRELATIONS (SHOCKS)	mean (prct.)	mean (prct.)	mean (prct.)	mean (prct.)
w^θ	$0.92 \begin{pmatrix} 0.77 \\ 0.98 \end{pmatrix}$	$0.89 \begin{pmatrix} 0.74 \\ 0.97 \end{pmatrix}$	$0.73 \begin{pmatrix} 0.55 \\ 0.88 \end{pmatrix}$	$0.72 \begin{pmatrix} 0.52 \\ 0.87 \end{pmatrix}$
w^ψ	$0.87 \begin{pmatrix} 0.74 \\ 0.94 \end{pmatrix}$	$0.86 \begin{pmatrix} 0.72 \\ 0.94 \end{pmatrix}$	$0.72 \begin{pmatrix} 0.57 \\ 0.88 \end{pmatrix}$	$0.71 \begin{pmatrix} 0.52 \\ 0.88 \end{pmatrix}$
RMSE (SHOCKS)				
w^θ	0.4267	0.4826	0.7308	0.7509
w^ψ	0.5264	0.5433	0.9032	0.9290
RMSE (IMPULSE RESP.)				
$w^\theta \rightarrow consumption$	0.1933	0.1684	0.1546	0.1650
$w^\psi \rightarrow consumption$	0.0850	0.0903	0.3907	0.3839
$w^\theta \rightarrow output$	0.2774	0.2394	0.2688	0.2743
$w^\psi \rightarrow output$	0.0838	0.0945	0.5233	0.5166

Note: The table reports the median correlations (and 5th and 95th percentile correlations in parentheses) between the true shocks and the structural shocks (top block row) identified in VECM and ACCA as well as the associated RMSEs (middle block row). The bottom block row contains the corresponding RMSEs w.r.t. the impulse responses.

the technology shock than the VECM in Exercise 3.5. Neither method estimates a significant effect on consumption of the preference shock at all horizons, which is similar to the result in Exercise 3.3 concerning the response of output to the preference shock. Overall, I assess that both techniques lead to qualitatively similar results as in Exercise 3.3.

When the persistence of the preference shock is increased, both ACCA and the VECM lose precision dramatically in uncovering the true shocks, which becomes apparent in the substantial reduction of the correlation measures and the increasing RMSEs. The sole exception to this exacerbation relates to the responses of consumption to the technology shock, where the RMSEs are smaller compared to Exercise 3.3 but accompanied by higher estimation uncertainty. In contrast, the two estimation methods fail dramatically in detecting the responses to the preference shock. This finding demonstrates that both procedures exhibit severe shortcomings in capturing the dynamics when facing slowly diminishing effects of the preference shock.

Remarks and large sample exercise

At this stage, I can conclude that ACCA can be seen as an alternative to the standard tool, but it cannot surpass the performance of the VECM in the settings hitherto considered. I show that both methods underperform when the underlying parameterization of the DGP

Figure 3.3: EXERCISES 3.5 & 3.6 - IMPULSE RESPONSES (EXTENDED, LONG-RUN RESTRICT.)
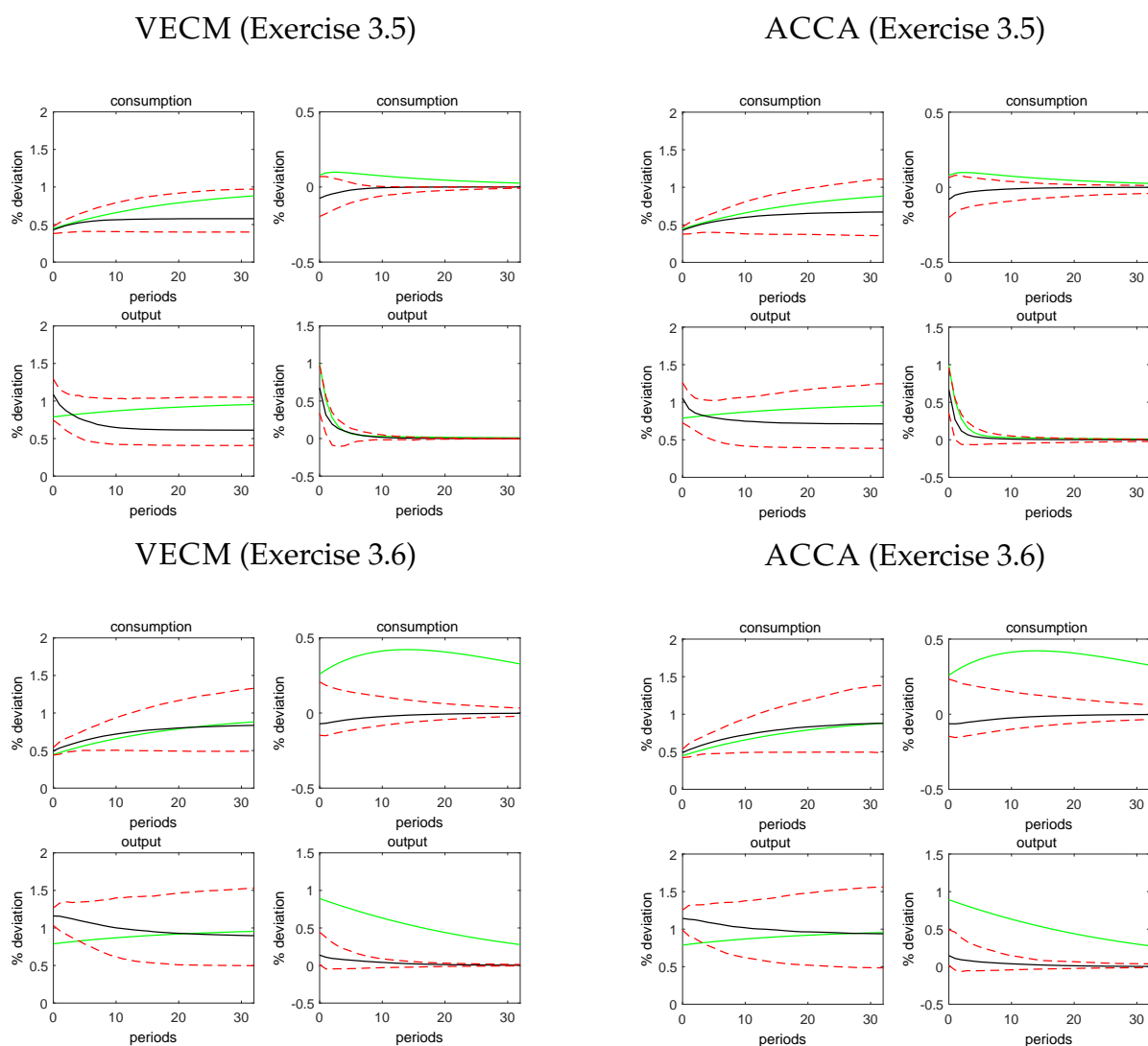
VECM (Exercise 3.5)   ACCA (Exercise 3.5)

VECM (Exercise 3.6)   ACCA (Exercise 3.6)

Note: The figure depicts FIA model impulse responses (in green) and estimated impulse responses (in black). Dashed red lines mark two-sided 90% confidence bounds. Impulse responses to w^θ are shown in the left column and impulse responses to w^ψ in the right column of each panel.

induces an eigenvalue of $(\mathcal{A} - \mathcal{BD}^{-1}\mathcal{C})$ close to one, whereas the ACCA is more sensitive to that situation. In Subsection 3.2.1, I have already stated that the magnitude of this eigenvalue is crucial for the VAR (VECM) identification. My findings seem to reveal the same for the ACCA. An explanation for this phenomenon might be inherent in the derivation of the subspace algorithm.

Due to the coincidence of FIA and LIE model, I can express equation (3.23) as

$$x_t = \left(\mathcal{A} - \mathcal{BD}^{-1}\mathcal{C}\right)^p x_{t-p} + \sum_{i=0}^{p-1} \left(\mathcal{A} - \mathcal{BD}^{-1}\mathcal{C}\right)^i \mathcal{BD}^{-1} y_{t-i} . \tag{3.39}$$

Under the *invertibility condition*, the state can be approximated by the past of the observables, implying that there is the same decisive connection between the eigenvalues of $(\mathcal{A} - \mathcal{BD}^{-1}\mathcal{C})$ and the past values of y_t. The lower the decay rate (in terms of the largest eigenvalue in modulus of $(\mathcal{A} - \mathcal{BD}^{-1}\mathcal{C})$), the greater the number of past observations would be required to estimate the state sufficiently. In practical applications with finite samples, these past values have to be truncated, implying that the integer p is rather small. Therefore, I end up with a problem analogous to the lag truncation bias of the VAR/VECM in the case of ACCA.

This problem becomes more severe in situations when long-run restrictions are used for identification and stationary exogenous processes are highly persistent. Exercises 3.5 and 3.6 illustrate this issue. I verify my explanation from above for this result by conducting two large sample exercises in which I increase the sample size up to 40000 observations. Thus, I eliminate potential errors caused by small sample uncertainty. The two exercises differ in the way the lag length is chosen for both methods. In Exercise 3.7, I apply the same lag order choice as before, i.e., I use the AIC for the VECM and twice the result of the AIC as the default for the lags and leads in the ACCA. In Exercise 3.8, I use a fixed lag length of 200 instead of the AIC suggestion.

The results are presented in Figure 3.4. A comparison of both exercises confirms my expectations. The poor performance of both techniques can mainly be attributed to lag truncation bias and sample size uncertainty. Small sample bias is largely diminished, as can be seen from the relatively tight confidence intervals in all cases. But both methods significantly underestimate the effects of all shocks at almost all horizons when the AIC is selected. Exercise 3.8 then clearly demonstrates that choosing a high lag order of sufficient length helps remove this bias.

3.5.3 Data generating process 2

Model description

In this subsection, I use another model as the DGP, which provides a more complex frame-

Figure 3.4: EXERCISES 3.7 & 3.8 - IMPULSE RESPONSES (EXTENDED, LARGE SAMPLE)
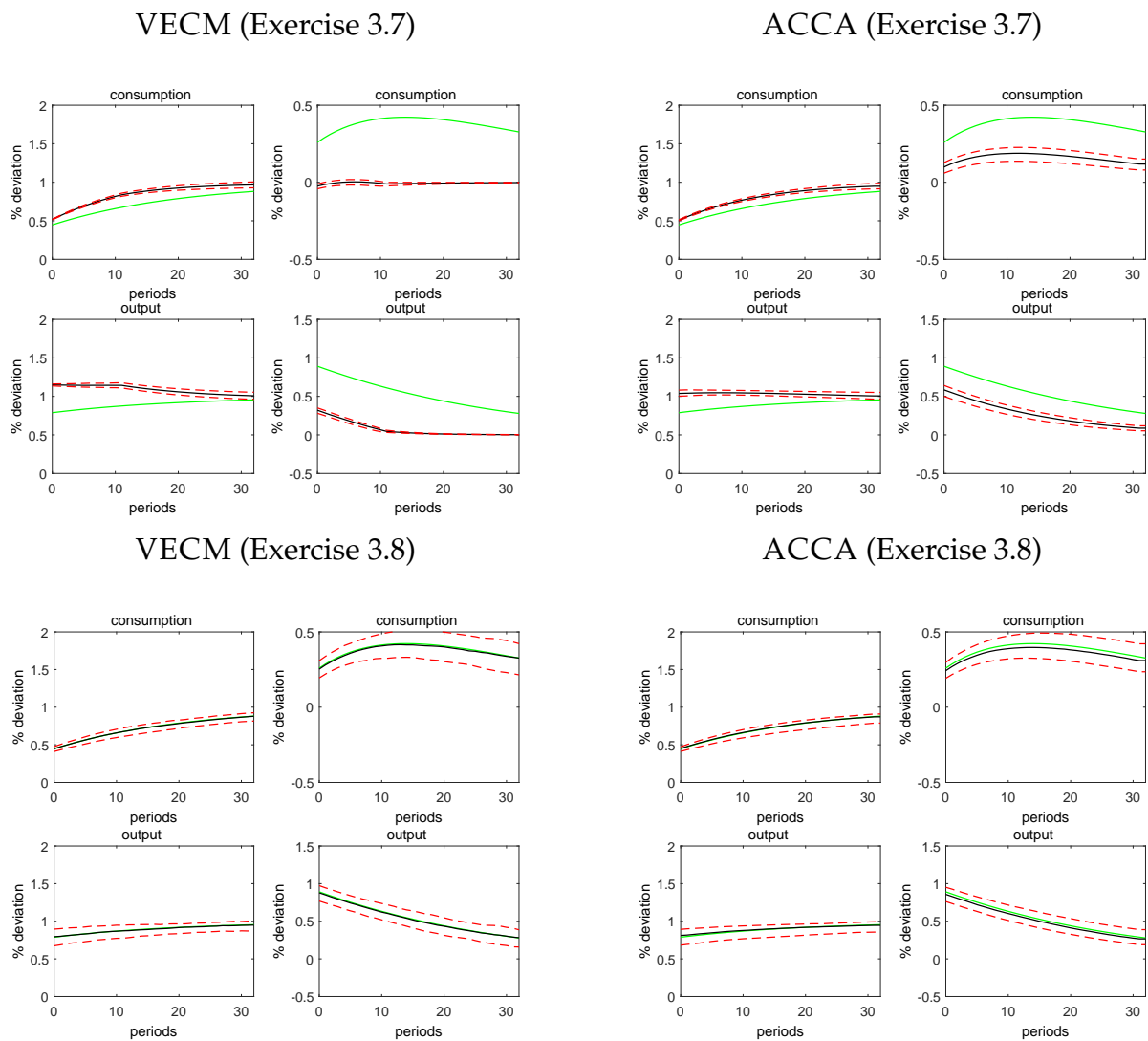
Note: The figure depicts FIA model impulse responses (in green) and estimated impulse responses (in black). Dashed red lines mark two-sided 90% confidence bounds. Impulse responses to w^{θ} are shown in the left column and impulse responses to w^{ψ} in the right column of each panel.

work in which VECM and ACCA can compete. Moreover, the model represents the theoretical basis for the empirical study in the next chapter. It offers several important features that are known to give a suitable description of what can be found in macroeconomic data. The DGP is simulated from a (slightly extended) version of the gold rush fever model by Beaudry et al. (2011) in the following. Beaudry et al. (2011) show that such a model can explain important properties of the US data. The interesting element of the model is the incorporation of market rushes into a standard DSGE model. A market rush occurs because agents receive a signal about a new market opportunity. Expecting future profits, agents found startups to produce new varieties of a good and to compete with each other. The winner of the competition becomes the monopolist in the new market. The perception of such an opportunity initiates a market rush and causes an economic expansion.

Beaudry et al. (2011) demonstrate that such a market expansion shock can be a major trigger of business cycle fluctuation, primarily over the first year. Hence, it provides an alternative and meaningful explanation for short-run movements of macroeconomic aggregates because this role is usually attached to other shocks, e.g., preference shocks as in Beaudry and Lucke (2010). Beaudry et al. (2011) build a theoretical model that can replicate the empirical features of a VECM comprising consumption and output. As highlighted by Cochrane (1994a) using such a VECM, consumption is mainly driven by a permanent component and output by a transitory component. A technology shock is a prominent instance of a permanent shock, whereas a market rush could be considered a transitory shock.

In an extended version of the Beaudry et al. (2011) model, Beaudry et al. (2006) use two types of intermediate firms both of which exhibit the possibility of a variety expansion in their goods. Variety expansion in the first type generates no long-run impact on productivity, whereas it does in the second type. The Beaudry et al. (2006) model version also includes two types of market rush shocks. In their quantitative assessment, Beaudry et al. (2006) ascertain that only market rushes without long-run effects on productivity contribute to the (short-run) dynamics of the economy. Therefore, I omit the second category of intermediate goods and market expansion shocks. Instead, I add the variable capacity utilization, which is not included in the original model but is typical for DSGE models, especially when incorporating investment-specific technology (IST) shocks, as will be the case in Chapter 4.[28]

The gold rush fever model can be characterized as follows.[29] A representative firm produces a raw final good Q_t using capital K_t utilized at rate u_t, labor h_t, and a set of intermediate goods $X_t(j)$ with mass $N_{x,t}$ according to the production function

$$Q_t = (u_t K_t)^{1-\alpha_x-\alpha_h} (\theta_t h_t)^{\alpha_h} N_{x,t}^{\xi} \left(\int_0^{N_{x,t}} X_{it}^{\chi} di \right)^{\frac{\alpha_x}{\chi}}, \qquad (3.40)$$

---

[28]See Smets and Wouters (2007) or Schmitt-Grohé and Uribe (2012), for instance.
[29]A detailed explanation is given in the NBER working paper by Beaudry et al. (2006).

with α_x, $\alpha_h \in (0,1)$, $\alpha_x + \alpha_h < 1$ and $\chi \leq 1$. χ denotes the elasticity of substitution between intermediate goods. A variety expansion has no long-run impact, which is guaranteed by setting $\xi = \frac{-\alpha_x(1-\chi)}{\chi}$. θ_t represents a Harrod-neutral technology process. Its log follows a random walk with drift term γ_θ (see below). $N_{x,t}$ is the number of effectively produced intermediate goods. Each good is produced by a monopolist using the raw final good as the input. The capital accumulation function is given by

$$K_{t+1} = (1 - \delta(u_t)) K_t + \left[1 - \mathcal{S}\left(\frac{I_t}{I_{t-1}}\right)\right] I_t , \tag{3.41}$$

where $\delta(u_t)$ is the depreciation rate depending on the degree of utilization. $\mathcal{S}\left(\frac{I_t}{I_{t-1}}\right)$ represents investment adjustment costs satisfying $\mathcal{S}(\gamma_\theta) = \mathcal{S}'(\gamma_\theta) = 0$ and $\varphi \equiv \mathcal{S}''(\gamma_\theta)\gamma_\theta^2$.

In each period, potential new varieties are produced if they are profitable enough, i.e., if the present value of expected future profits exceeds the setup costs of the startup. Let $N_{s,t}$ denote the total number of startups. The probability of becoming a functioning new firm with a product monopoly is endogenously determined, whereas existing monopolies die out at an exogenous rate μ. Hence, the dynamics of the number of active firms is given by

$$N_{t+1} = (1 - \mu) N_t + \rho_t N_{s,t} . \tag{3.42}$$

In principle, the model distinguishes between the number of active and potential varieties. The law of motion for the latter is described by

$$\mathcal{N}_{t+1} = (1 + \eta_t - \mu) \mathcal{N}_t , \tag{3.43}$$

so the potential varieties vanish at exogenous rate μ. η_t represents the aforementioned market rush shock. It carries information about the number of potential products in the next period. Beaudry et al. (2006, 2011) assume that the optimal behavior of entrepreneurs implies the exploitation of all potential varieties. Therefore, $N_t = \mathcal{N}_t \, \forall t$ in equilibrium.

The representative household maximizes expected lifetime utility

$$E_t \sum_{\tau=0}^{\infty} \beta^\tau \left(log\left(C_{t+\tau} - bC_{t+\tau-1}\right) + \psi\left(\overline{h} - h_{t+\tau}\right)\right) \tag{3.44}$$

subject to the budget constraint

$$C_t + P_t^M \mathcal{E}_t^M + P_t^S \mathcal{E}_t^S = W_t h_t + R_t u_t K_t + \mathcal{E}_t^M \Pi_t + (1 - \mu) P_t^M \mathcal{E}_{t-1}^M + \rho P_t^M \mathcal{E}_{t-1}^S . \tag{3.45}$$

β is the discount factor. The household chooses consumption C_t and labor supply h_t as well as its investment in startups $\mathcal{E}_t^S$ and equity holdings in monopolies $\mathcal{E}_t^M$, where the corresponding prices are P_t^S and P_t^M, respectively. In addition to wages w_t, the household

Table 3.4: CALIBRATED PARAMETERS (GOLD RUSH FEVER MODEL)

β	0.9926	ρ	0.9166
α_x	0.3529	σ_x	0.2865
α_h	0.4235	σ_θ	0.0131
δ	0.0250	b	0.5900
χ	0.8333	φ	0.4376
γ_θ	1.0060	ε_δ	0.42
μ	0.0086		

receives earnings R_t from supplying capital services and dividend payments Π_t. The goods market clearing condition is given by

$$Y_t \;=\; C_t + I_t + S_t \,, \tag{3.46}$$

where Y_t denotes value added, and S_t denotes total startup expenditures.

Finally, there are two stochastic processes, a unit root process for technology and a transitory process for the nonproductive market expansion shock,

$$log\,(\theta_t) \;=\; log\,(\gamma_\theta) + log\,(\theta_{t-1}) + \sigma_\theta w_t^\theta \tag{3.47}$$

and

$$log\,(\eta_t) \;=\; (1-\rho)\,log\,(\mu) + \rho\,log\,(\eta_{t-1}) + \sigma_\eta w_t^\eta \,, \tag{3.48}$$

where w_t^θ and w_t^η are Gaussian white noise with standard deviations σ_θ and σ_η.

I use the same calibrated and estimated parameter values as Beaudry et al. (2006), with one exception. Because I abandon the second set of intermediate goods in my version of the model, the elasticity of output to intermediate goods X equals the value that is distributed to both types in the original model. Additionally, I calibrate the elasticity of marginal depreciation $\varepsilon_\delta \left(\equiv \frac{\delta''(u)}{\delta'(u)} u \right)$ to a value of 0.42 following Greenwood et al. (1988). All parameter values are given in Table 3.4.

As there are two shocks to identify, I construct an artificial data set by generating time series for two variables. To explicitly allow for cointegration and refer to the previous remarks, consumption and output are the chosen observables. Therefore, the following state space representation of the model is my theoretical benchmark. After detrending and log-linearizing around the steady state, I solve the model by applying the method of Uhlig

(1999). Finally, I formulate the state space representation

$$
\begin{pmatrix} \widetilde{k}_{t+1} \\ \widetilde{c}_t \\ \widetilde{i}_t \\ \widetilde{\eta}_t \\ log\left(\theta_t\right) \end{pmatrix} = const_1 + \mathcal{A} \begin{pmatrix} \widetilde{k}_t \\ \widetilde{c}_{t-1} \\ \widetilde{i}_{t-1} \\ \widetilde{\eta}_{t-1} \\ log\left(\theta_{t-1}\right) \end{pmatrix} + \mathcal{B} \begin{pmatrix} w_t^\theta \\ w_t^\eta \end{pmatrix}, \tag{3.49}
$$

$$
\begin{pmatrix} log\left(C_t\right) \\ log\left(Y_t\right) \end{pmatrix} = const_2 + \mathcal{C} \begin{pmatrix} \widetilde{k}_t \\ \widetilde{c}_{t-1} \\ \widetilde{i}_{t-1} \\ \widetilde{\eta}_{t-1} \\ log\left(\theta_{t-1}\right) \end{pmatrix} + \mathcal{D} \begin{pmatrix} w_t^\theta \\ w_t^\eta \end{pmatrix}, \tag{3.50}
$$

where lowercase letters with "~" denote log-deviations of the detrended variables from their steady state level, and $const_i$, $i = 1, 2$, are vectors of constants.

The FIA model is used to generate the artificial time series.[30] Under the chosen parameterization, the eigenvalues of $\mathcal{A}$ are 0.5706, 0.7712, 0.9166, 0.9313 and one. There are two nonzero eigenvalues of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$ with magnitudes of 0.7894 and 0.9116. The selection of lags (and leads) in the VECM (ACCA) is again based on the AIC. The identification scheme is a zero long-run restriction on the second shock, i.e., the innovation to output has no effect on consumption over the long-run. To judge the quality of the estimation methods with respect to the structural shock identification and their dynamics, I calculate the same statistics as in the first simulation study.

Results

Table 3.5 displays the results of both estimation procedures from the simulation exercise. In relation to the results of DGP 1, things have slightly changed, as the ACCA outperforms VECM in terms of correlations when looking at the confidence intervals. In terms of the median correlation, both techniques are very close to each other. The comparison of the corresponding RMSE reveals better performance of ACCA only for the preference shock, whereas the RMSE associated with the technology shock is nearly the same. VECM offers slightly lower RMSEs than the ACCA in three of four estimated impulse responses. In the case of the reaction of output on the technology shock, the ACCA dominates the VECM.

Figure 3.5 illustrates the theoretical impulse responses and median estimates associated with

---

[30]For this purpose, I use the Matlab command ss(sys,'minimal') to compute the minimal realization of the state space system given in equations (3.49) and (3.50).

Table 3.5: SIMULATION RESULTS (GOLD RUSH FEVER MODEL)

	VECM	ACCA
CORRELATIONS (SHOCKS)	mean (prct.)	mean (prct.)
w^θ	$0.95 \begin{pmatrix} 0.53 \\ 1.00 \end{pmatrix}$	$0.94 \begin{pmatrix} 0.65 \\ 0.99 \end{pmatrix}$
w^η	$0.94 \begin{pmatrix} 0.50 \\ 0.99 \end{pmatrix}$	$0.94 \begin{pmatrix} 0.66 \\ 0.99 \end{pmatrix}$
RMSE (SHOCKS)		
w^θ	0.4082	0.4089
w^η	0.4392	0.4203
RMSE (IMPULSE RESP.)		
$w^\theta \rightarrow consumption$	0.2156	0.2273
$w^\eta \rightarrow consumption$	0.1509	0.1571
$w^\theta \rightarrow output$	0.3789	0.3256
$w^\eta \rightarrow output$	0.2067	0.2147

Note: The table reports the median correlations (and 5th and 95th percentile correlations in parentheses) between the true shocks and the structural shocks (top block row) identified in VECM and ACCA as well as the associated RMSEs (middle block row). The bottom block row contains the corresponding RMSEs w.r.t. the impulse responses.

ACCA and VECM as well as the corresponding 90% confidence bounds. The graphical analysis does not give further insights except that, for both methods, the response of consumption to a market expansion shock suffers from high uncertainty, as shown by the wide confidence intervals.

The results of the large sample simulation in Figure 3.6 shed more light on these aspects. The figures disclose that both estimators have large finite sample biases because the asymptotic errors are relatively small. The VECM still induces some deviations from the true impulse responses with respect to the technology shock. These shortcomings can be attributed to lag truncation bias, as the lag length is fixed by the AIC in the large sample simulation. When setting a fixed lag length of 200 and repeating the large sample simulation, the deviation disappears.[31]

3.6 Concluding remarks

In this chapter, I have shown that extending subspace algorithm cointegration analysis to structural estimation can provide an alternative to standard estimation methods. However, its property of capturing the full structure of the underlying DGP does not emerge considerably in a simple simulation study. My explanation is that this procedure might also suffer from lag truncation bias as the VECM does. Nevertheless, when applying this approach

---

[31] A graphical illustration is omitted to save space.

Figure 3.5: IMPULSE RESPONSES (GOLD RUSH FEVER MODEL, SMALL SAMPLE)
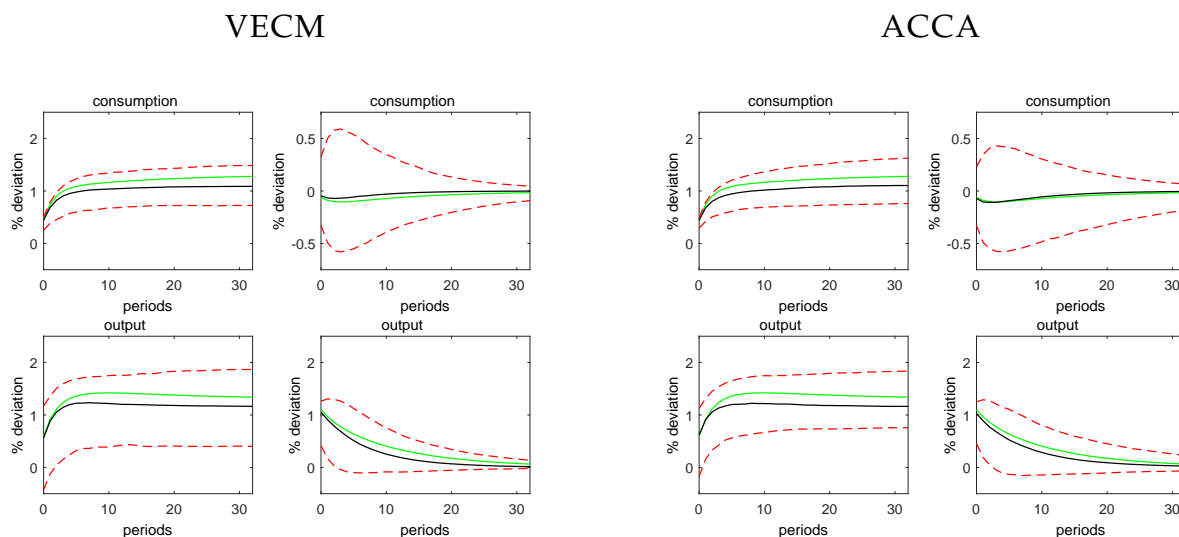
VECM

ACCA

Note: The figure depicts FIA model impulse responses (in green) and estimated impulse responses (in black). Dashed red lines mark two-sided 90% confidence bounds. Impulse responses to w^θ are shown in the left column and impulse responses to w^η in the right column of each panel.

Figure 3.6: IMPULSE RESPONSES (GOLD RUSH FEVER MODEL, LARGE SAMPLE)
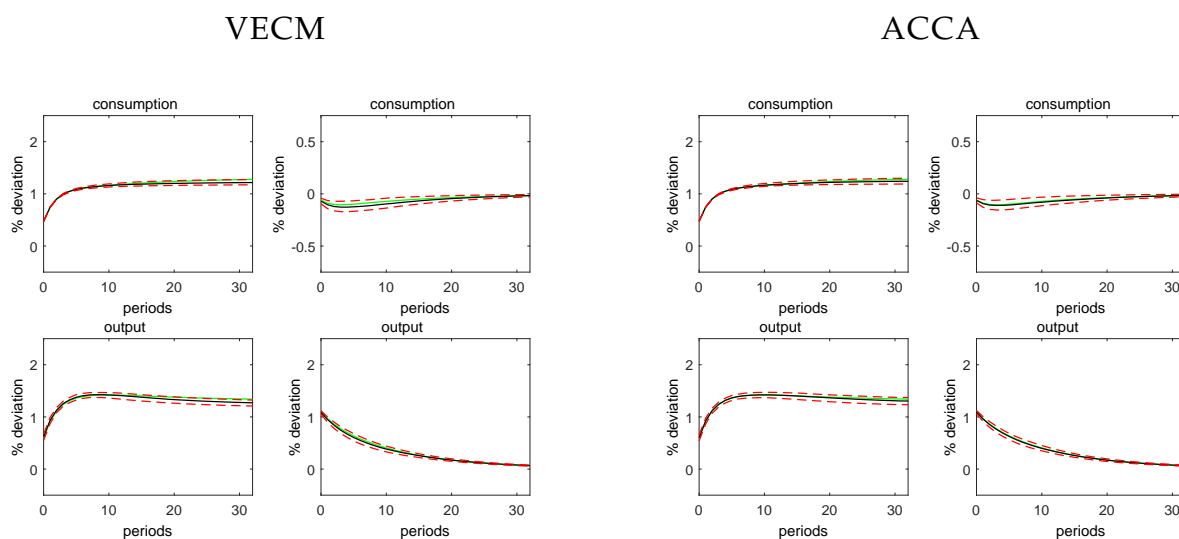
VECM

ACCA

Note: The figure depicts FIA model impulse responses (in green) and estimated impulse responses (in black). Dashed red lines mark two-sided 90% confidence bounds. Impulse responses to w^θ are shown in the left column and impulse responses to w^η in the right column of each panel.

within a more complex setting in terms of the second DGP, my simulations demonstrate that this approach can perform at least comparably to its standard counterpart. The gold rush fever model used as the second DGP lays the foundation for the empirical examination in the subsequent chapter.

Chapter 4

Identifying market rushes in the US data

One goal of empirical business cycle analysis is to uncover the sources of macroeconomic fluctuation. Many studies aim to quantify the contribution of structural shocks to variations in relevant variables. My focus in this chapter is the analysis of very short-run movements. In particular, I am interested in identifying the drivers of economic activity within their first year of unfolding. I thereby pick up on the idea of the market rush shocks introduced in Subsection 3.5.3 and pioneered by Beaudry et al. (2011), who show by examining US data that a market expansion shock accounts for 80% of the variation in output on impact, which remains approximately 50% within the first year. According to the Burns and Mitchell (1946) definition of a business cycle, the typical duration of a business cycle ranges from six to 32 quarters. For that reason, I label shocks that dominate in the forecast error variance decomposition (FEVD) of output (or some other proxy for economic activity) over a maximum horizon of one year as very short-run triggers of macroeconomic fluctuations. In this chapter, I provide further evidence that market rushes can be assigned this role.

4.1 Motivation and related literature

I refer to two bodies of literature. The first category relates to a variety of empirical findings from investigations of driving forces of output fluctuations. As most existing studies focus on explaining longer horizons than I do here, the results reviewed in the following are often merely by-products of the analyses. Thus, I demonstrate by citing the following articles that the diversity of findings does not provide a clear picture of the true triggers of output variation over the very short-run. I find support in this chapter for the market rush interpretation of Beaudry et al. (2011), which helps their approach stand out among disparate results in the literature.

One explanation is a study by Fisher (2006), who utilizes the SVAR method and stresses that nontechnology shocks dominate the first four quarters of output fluctuations by more than

80% when examining postwar US data until 1979. For the post-1982 period, he discovers that neutral technology shocks account for 50–80% of very short-run variation in output.[1] Forni and Gambetti (2010) analyze a much larger data set that consists of 101 quarterly US macroeconomic time series covering the period from 1959 to 2007 and estimate a structural factor model. The authors interpret the shock accounting for 60–70% of the forecast error variance in output in the first year as a supply-side disturbance. Zeev and Khan (2015) are more precise about the nature of the identified shock. According to their SVAR analysis of the US data, a surprise shock in total factor productivity (TFP) boosts output variation on impact by nearly 60%.

There are also contributions in the literature that highlight the role of demand-side disturbances as drivers of output fluctuations over the very short-run. Smets and Wouters (2007) estimate a medium-scale DSGE model of the US economy using Bayesian techniques and decompose the FEVD of output into the contributions of seven shocks. Focusing on the first two quarters, an exogenous spending shock explains the largest part of output variance, with the immediate impact being the highest by approximately 40%.[2] Beaudry and Lucke (2010) apply a cointegrated SVAR model and disclose that a shock that can be interpreted as preference shock is the dominant source of output variation in the first quarters. Milani (2012) builds on the model of Smets and Wouters (2007) and uncovers a risk premium shock as main trigger of output over the very short-run. Last but not least, market rushes à la Beaudry et al. (2011) supplement the list of structural shocks as driving forces of output movements at very high frequencies.

The second body of related literature addresses the role of news shocks in explaining business cycles. A market rush shock can be seen as a specific type of news shock, as it offers the typical news shock characteristic in the sense that agents adjust their behavior in anticipation of future economic outcomes. News shocks have become a popular subject of study in recent years, as evidenced by the wide variety of examples. The most prominent are news about future exogenous events, such as technology growth (as introduced by Beaudry and Portier (2007) and Jaimovich and Rebelo (2009)), tax rates (see Leeper et al. (2008, 2013), Leeper and Walker (2011) and Sirbu (2013)) or government spending (e.g., in Mertens and Ravn (2010), Kriwoluzky (2012), Schmitt-Grohé and Uribe (2012)). Another instance of a news shock is anticipated monetary policy (see Cochrane (1998) or Milani and Treadwell (2012) for a recent publication). In the case of the Beaudry et al. (2011) market rushes, agents anticipate the creation of new markets in the next period and use this (exogenous) information to adapt their decisions in the current period. Therefore, the following empirical investigation of market rushes can be considered a contribution to the news shock literature.

It should be noted that market rushes, as modeled in the previous chapter, do not lead to

---

[1] Note that Fisher (2006) identifies IST shocks as major drivers of output for longer horizons.

[2] Exogenous spending not only reflects government expenditures but also net exports in the Smets and Wouters (2007) model.

the widely debated problems of nonfundamentalness because the theoretical model is not subject to a noninvertible MA representation.[3] As nonfundamentalness is the focus of the upcoming chapters, I concentrate solely on the application of the previously presented estimation methods and the corresponding findings in the remainder of this chapter. The goal is to identify structural shocks in data for the US and to confirm the market rush interpretation of Beaudry et al. (2011). In addition to their SVECM approach, I use structural ACCA for estimation and conduct a Granger causality test to show that the identified shocks exhibit market rush features. My results provide further evidence in support of the idea that market rushes are important drivers of business cycles in the very short-run.

4.2   Empirical examination

4.2.1   Course of action

Having shown that structural ACCA is an appropriate alternative to the SVECM in the previous chapter, I take the next step and apply it to the US data. I begin with a system consisting of two variables, consumption and output, and employ both aforementioned estimation methods as in Subsection 3.5.3. Lag lengths are chosen based on the AIC as before. In a second step, I test for Granger causality between the shock series, which I identify through the zero long-run restriction on consumption and conjecture to reflect the market rush phenomenon, and an additional data series representing the developing number of firms in the economy. I thereby follow the approach of Haertel and Lucke (2008) and Lucke (2013), who find that (neutral) technological news shocks detected through an SVECM are Granger-causal for German and US patent data.

This procedure is repeated for a system of three variables wherein an IST variable is added to the former two aggregates. The reason is that the bivariate system above might be biased due to omitted variables such that the structural shocks uncovered in this system are insufficient to provide a distinct picture of the origins of macroeconomic fluctuations, as they could be a mixture of different sources. One might argue that the expansion to a third dimension is also insufficient, but as my results will show, it seems to be adequate for a nonambiguous finding. Another argument for the inclusion of IST is that it has become a standard element in DSGE models that suggest comprehensive explanations of business cycle dynamics (see Smets and Wouters (2007) or Schmitt-Grohé and Uribe (2012)).

---

[3]The corresponding eigenvalues of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$ are smaller than one in absolute value (see the model description in Subsection 3.5.3).

4.2.2 Data description

The sample for my quarterly data set spans the 1955Q1–2010Q4 period. For the output and consumption variables, I use the time series of gross domestic product (GDP) and personal consumption expenditures, which can be found in the National Income and Product Accounts (NIPA) offered by the U.S. Bureau of Economic Analysis (BEA). To obtain per capita values, I divide them by the number of employed persons measured as the civilian, noninstitutional population aged over 16. The proxy for IST is constructed as the relative price of investment, i.e., as the log-difference of the NIPA deflator for consumption and the NIPA investment price index.

Concerning the number of firms, one could think of appropriate candidates, such as some series provided by the BEA and used in Lewis and Stevens (2012). The BEA's Survey of Current Business publishes data on net business formation and the number of new incorporations. Unfortunately, these are discontinued series that cover data only until 1996. Beaudry et al. (2011) also motivate their idea of gold rush fever using a recent example, the dot-com boom at the end of the 1990s; thus, I consider an alternative measure drawn from the Bureau of Labor Statistics (BLS), which also contains values for the nineties. On its website, the BLS offers data on private sector establishment births. A detailed explanation and treatment of this series is given by Sadeghi (2008). The series starts in the second quarter of 1993 and is updated quarterly by the BLS. Hence, it seems convenient to use the series for my test for market rush interpretation.

4.2.3 Results

Figure 4.1 displays the estimated impulse responses. Comparing the findings of the bivariate system (solid lines), both methods yield similar results in terms of the shape of the impulse responses. The main difference is that there is much more uncertainty with respect to the ACCA indicated by much wider confidence bounds. Another aspect to be mentioned is the fact that the estimated long-run effect of the technology shock is substantially higher in the case of the ACCA, and it evolves more gradually. Turning to the designated market rush shock, both methods induce a positive effect on consumption, which is insignificant for all horizons. This positive (but insignificant) effect contradicts the implications of the theoretical model, but as the gold rush fever model admits only a negligible negative response of consumption, I do not pay closer attention to this aspect. The estimated responses of GDP are quite close to the theoretical responses and emphasize the transitory and positive character of a market rush.

It is interesting to analyze whether this identified shock can definitely be interpreted as a market rush. I therefore test the hypothesis of a potential causal effect of the structural shock on the above described data on establishment births by estimating a bivariate VAR

Figure 4.1: MARKET RUSH ANALYSIS (ESTIMATED IMPULSE RESPONSES)

VECM   ACCA
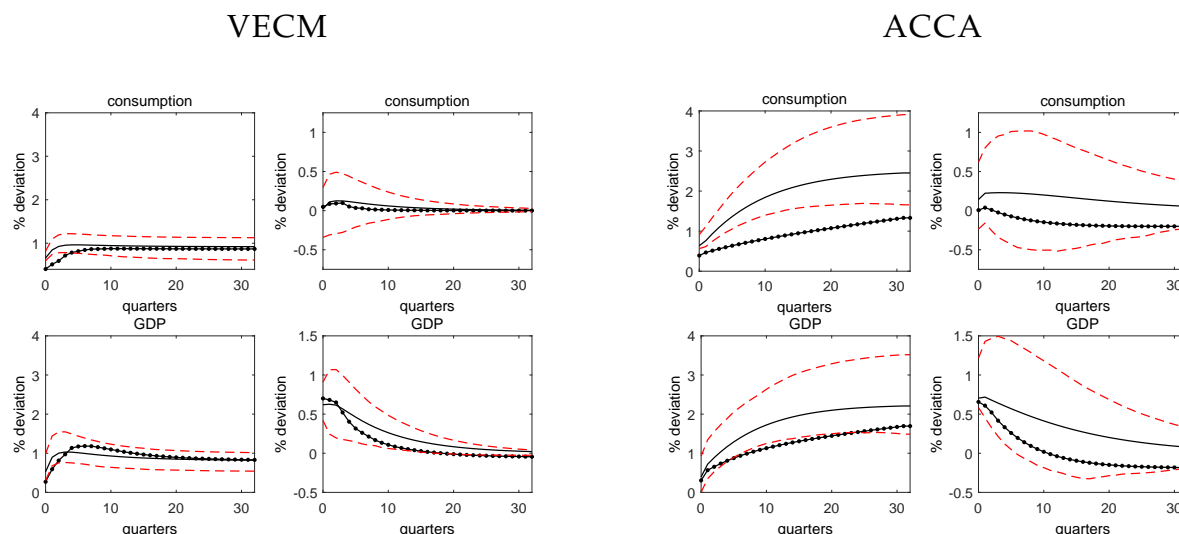
Note: The figure depicts estimated impulse responses of the 2-VAR system (solid black lines) and estimated impulse responses of the 3-VAR system (dotted black lines). Dashed red lines mark the 95% bootstrapped (Hall) confidence intervals w.r.t. the 2-VAR system. Impulse responses to the neutral technology shock are shown in the left column and impulse responses to the market rush shock in the right column of each panel.

model comprising both series and a Granger causality test.[4] The results are presented in (the first block column of) Table 4.1. Unfortunately, the results for the causal relationship are not clear. The p-values for the null hypothesis that the shocks Granger-cause establishment births are relatively low, with values of approximately 7% (ACCA) and 8% (VECM), but testing in the opposite direction reveals p-values that are either slightly lower (6% for the ACCA residuals) or higher (12% for the VECM shocks).

As these findings are not convincing, I modify the analysis by expanding the set of observable variables as explained above, and I use personal consumption expenditures only on nondurables and services as the consumption measure.[5] Adding an IST variable to the system implies a third structural shock, which can be identified as will be described below. Maintaining the gold rush fever model of the previous chapter as the theoretical foundation of the analysis would mean appending another stochastic process to the model, e.g., an exogenous IST process as proposed by Greenwood et al. (1997, 2000).

A Johansen cointegration test to the trivariate system (i.e., consumption, relative price of investment and output) verifies a cointegration rank of one, that is, the system is driven by two common stochastic trends. This result coincides with the theoretical benchmark in which the IST process and the neutral technology process are modeled as unit root processes. Structural estimation thus requires three identifying assumptions. In addition to the zero long-run restriction from above, I impose two zero short-run restrictions: the shocks

---

[4]I include a constant as a deterministic variable and choose the lag length according to the AIC.

[5]The usage of consumption expenditures on nondurables and services is more common in the literature (see also Cochrane (1994a) and Beaudry et al. (2011)). The exclusion of expenses on durable goods is theoretically justified because it seems to not be a good proxy for the corresponding service flow from which consumers derive their utility rather than from the expenditures (see Flavin (1981) or Campbell (1987)).

Table 4.1: MARKET RUSH ANALYSIS (GRANGER CAUSALITY TESTS)

	2-VAR SYSTEM	3-VAR SYSTEM
	ESTABLISHMENT BIRTHS	

Market rush shocks (VECM)	0.0848
	      0.1284
		0.0032
		      0.5099

Market rush shocks (ACCA)	0.0698
	      0.0580
		0.0473
		      0.5508

Note: Upper left corner depicts p-value for null: row variable does not Granger-cause column variable. Lower right corner shows p-value for null: column variable does not Granger-cause row variable.

to consumption and output have no contemporaneous impact on IST. This identification scheme is therefore in line with my theoretical model.

To check whether this procedure works well in distinguishing among the three structural shocks, I examine the FEVD for the modified gold rush fever model and the empirical findings of the ACCA and VECM in Figure 4.2.[6] Consumption apparently captures the neutral technology shock (blue) as predicted by the model. My measure for IST also seems appropriate for conveying the effects of an IST shock (green). Looking at the FEVD of output, market rushes (red) are dominant over the very short-run with an approximately 80% share of the variance on impact. The first ranking position is replaced after three quarters by the neutral technology shock increasing from 15–20% to over 80% of the variation over the medium-run. The contribution of the IST shock is virtually zero in the case of the VECM and closer to the outcome of the theoretical model for the ACCA. I conclude that IST shocks do not play a significant role in explaining output (or consumption) fluctuations over the entire business cycle horizon.

For completeness, I add the associated impulse responses of the trivariate system to Figure 4.1 (dotted black lines). The qualitative pattern is similar to what I have shown before for the bivariate system. The typical feature of the market rush shock is still obvious.[7]

Finally, I owe the reader evidence supporting the market rush interpretation. The corresponding results for the Granger causality tests in the right block column of Table 4.1 provide confirmation of that interpretation. The market rush shocks (identified in the trivariate system) significantly Granger-cause establishment births at the 5% level. This result is not invalidated by the causality test in the opposite direction as in the bivariate case: the null hypothesis that establishment births Granger-cause the structural shocks can be clearly rejected.

---

[6]I use the same calibrated parameter values as in Subsection 3.5.3. The standard deviation of the IST shock is set to one-third of the standard deviation of the neutral technology shock.

[7]The estimated effect of a neutral technology shock in the case of the ACCA is now lower and closer to the VECM and the theoretical model.

Figure 4.2: MARKET RUSH ANALYSIS (FEVD, 3-VAR SYSTEM)

THEORETICAL MODEL

VECM

ACCA
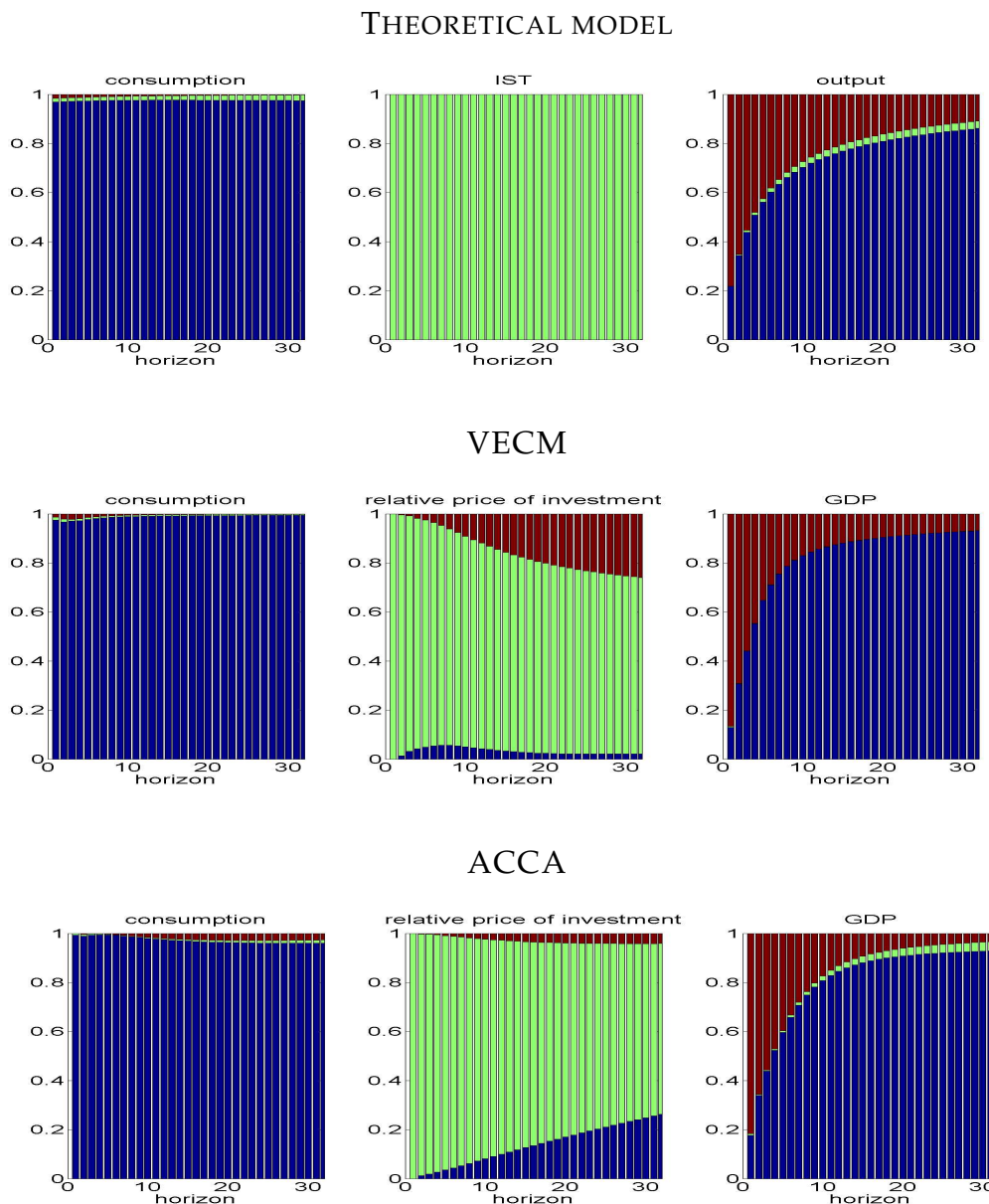
Note: The figure depicts the FEVD shares of consumption (left column), IST (middle column) and output (right column) due to the neutral technology shock (in blue), the IST shock (in green) and the market rush shock (in red) in the theoretical model (top panel), VECM (middle panel) and ACCA (bottom panel).

4.3 Concluding remarks

This chapter concludes the first part of this thesis. I have introduced structural ACCA as an alternative to SVECM and shown that they are on a par with each other. In summarizing this chapter, I draw several conclusions. The modified version of the gold rush fever model by Beaudry et al. (2011) provides a convincing theoretical framework of what is discovered in the US data. Furthermore, my results contradict the findings of Fisher (2003, 2006) in the sense that IST shocks explain only a minor share of output variations. Over the short-run, the market rush shock surpasses its competitors in explaining output fluctuations, whereas the neutral technology shock dominates the medium- and long-run variation in output. By applying a Granger causality test, I find evidence in favor of the market rush interpretation of the structural shock (identified in the trivariate system) through my estimation procedures. Hence, I provide support for the Beaudry et al. (2011) hypothesis and a robustness check that emphasizes the important role of market rushes in stimulating business cycles over the very short-run.

As market rushes represent a particular type of a news shock, I also contribute to the debate about the relevance of news shocks as sources of macroeconomic fluctuation. However, market rushes, as specified in my theoretical model, do not generate a noninvertible MA representation of the model solution. Usually, this characteristic is a central subject in the discussion of news shocks and poses a special challenge for both the VAR and subspace algorithm analysis. In the subsequent chapters, I will address this topic and give detailed insights and examinations of potential solutions.

Chapter 5

Anticipation effects and their consequences for structural estimation

In this chapter, I change the focus and analyze the situation when the *invertibility condition* of Chapter 2 is not satisfied. Hence, the match between the FIA and the LIE model ceases to exist. In this case, the econometrician is not able to uncover the true dynamics of the FIA model. She makes a systematic error due to the fact that her information set $\mathcal{I}_t^E$ is smaller than that of the agents in the FIA model, $\mathcal{I}_t^A$. The size and implication of this error depend on the underlying FIA model structure. This situation is often designated by the term non-fundamentalness. I explain how to quantify this error without implementing a Monte Carlo study. Given an economic model solution in terms of the FIA model representation, it is possible to calculate the structural form of the LIE model, i.e., to compute the representation that would be estimated by the econometrician who applies standard structural estimation techniques. As a consequence, I can isolate the systematic error due to nonfundamentalness from other error sources, such as model misspecification, lag truncation or small sample uncertainty. I provide analytical and numerical examples demonstrating the effects when news shocks are present within the theoretical framework. The chapter concludes with a discussion of the relevance of adding variables to the econometrician's information set in order to reduce her informational disadvantage relative to the agents.

5.1 Motivation and related literature

A recent development in macroeconomics is related to the observation that information can be a decisive factor affecting people's behavior and thus helps explain social and economic phenomena. There are different channels through which information influences individuals and their actions in the economy. Information economics studies these channels. Standard theory postulates that agents' information sets encompass the entire history of realizations

of relevant variables. Information economics moves away from this assumption and investigates the related consequences. What are the implications of imperfect information in the economy? How does asymmetric information arise and how does it impact economic decisions? What are the outcomes when agents have information on future events in addition to their knowledge of past incidents? All these questions include a substantial role of the design of the agents' information sets. Based on this design, theoretical models are built to provide explanations for certain puzzles. Empirical testing of such models poses an essential challenge for the researcher.

In the following, I consider the case where the agent's information set also contains information about future realizations of variables, i.e., I consider situations in which agents anticipate future shocks. Typical examples in the business cycle literature include news about future technology growth, tax foresight or anticipation of government spending shocks. In all these situations, the econometrician who intends to run an empirical test of related theoretical models only has access to current and past observations of variables. Hence, she can be confronted with the problem of having a smaller information set than the agents in the model economy. Consequently, the empirical test may lead to false conclusions.

Specifically, VAR models are a matter of considerable debate in this context. In an environment with anticipation effects, the general claim of VAR econometricians to let the data speak freely becomes a weakness. VAR models are usually limited to a rather small number of variables and therefore constrain the econometrician's information set. Furthermore, structural estimation methods are applied in order to give the estimated shocks an economic interpretation. They require parameter restrictions for identification but try to impose only minimal structure on the data. It can be shown that SVAR models can generally produce systematic errors in the face of anticipation effects. These errors arise because the underlying theoretical model, i.e., the FIA model, exhibits a nonfundamental MA representation. Moreover, the systematic failure is related not only to the VAR models but also to the subspace algorithm analysis, which I have presented in the previous chapter.

This chapter refers to several recent papers that investigate the consequences of nonfundamentalness for the performance of VAR methods in identifying news shocks and their dynamic effects. Leeper and Walker (2011) and Leeper et al. (2013) point to the difficulties an econometrician is faced with when estimating SVAR models that aim at the identification of anticipation effects. In a simple neoclassical model, they show that tax foresight induces model dynamics that cannot be identified correctly by an econometrician using standard SVAR estimation. At worst, she might draw completely incorrect conclusions from her exercise.[1]

As emphasized by Fernández-Villaverde et al. (2007), nonfundamentalness produces a wedge between the true economic shocks and the innovations that result from SVAR estima-

---

[1] I revisit this aspect in Section 7.5 and present an illustrative example in the appendix.

tion. To reduce this wedge, several solutions are proposed and discussed in the literature. A recommendation that is often stated is the incorporation of forward-looking variables, such as stock prices, to resolve this problem. A well-known empirical example is the pioneering work by Beaudry and Portier (2005, 2006), who identify news shocks by estimating an SVECM of TFP and stock prices. Their analysis is theoretically founded on a simple Lucas asset tree model, which has an invertible MA representation. The authors demonstrate that stock prices seem to be an appropriate candidate for identification of news about future technological improvements. Nevertheless, Forni et al. (2014) show that a slight modification of the underlying stochastic process that describes the evolution of technology in the model generates nonfundamentalness. Thus, the true shocks cannot be recovered even when stock prices are included in the VAR.

Fève and Jidoud (2012, 2014) use rather stylistic models to show that the impulse responses to news shocks identified from SVARs imposing either short-run or long-run restrictions are biased. The authors stress that this bias is smaller the more the news shocks account for the variation in the variables included in the VAR. Sims (2012) conducts Monte Carlo simulations using a DSGE model with real and nominal frictions as well as news shocks as the DGP to analyze the effects of nonfundamentalness on the reliability of SVARs in discovering the true dynamic responses to the news shocks. He refers to the terminology of Fernández-Villaverde et al. (2007) and the findings of Fève and Jidoud (2012, 2014) to emphasize that the wedge due to nonfundamentalness decreases with the relative importance of the news shocks. Sims (2012) concludes that the problem of nonfundamentalness is a problem of missing states in the VAR and that the choice of "observable variables (...) included in a VAR might matter - some observables may do a better job of forecasting the missing states, hence leading to (...) a closer mapping between VAR innovations and structural shocks."[2] He also suggests that including as many observable variables as possible represents a remedy for the problem. Because this procedure aims at enlarging the information set of the econometrician, it has its natural consequence in switching from (large scale) VAR to dynamic factor models, e.g., as introduced by Forni et al. (2000), Forni et al. (2009) and implemented by Forni et al. (2014).

In this chapter, I first focus on the situation in which the number of observable variables is limited to the econometrician and consider the expansion of the econometrician's information set at the end of the chapter. In contrast to Sims (2012), I demonstrate that the aforementioned wedge not only affects the estimated responses to the news shocks but also the responses not associated with this type of shock. I argue that the choice of observable variables might matter for this aspect but does not influence the general size of the wedge of nonfundamentalness. Prior to that, I begin with a description of the concept of nonfundamentalness followed by examples on how to quantify the systematic error that arises as a

---

[2]Note that Sims (2012) uses the term structural shocks to label the true economic shocks from the underlying FIA model.

result.

5.2 Nonfundamentalness (and noninvertibility)

5.2.1 The concept of nonfundamentalness

Consider the following covariance stationary zero mean stochastic process with rational spectral density. It has a vector MA representation of the form

$$y_t = M(L) w_t, \tag{5.1}$$

where y_t is the vector of observables, and w_t denotes a white noise vector that satisfies $E(w_t) = 0$, $E(w_t w_t') = I_k$ and $E(w_t w_{t-j}) = 0$ for integer $j > 0$. Suppose that I know the true model describing the economy and I can express the solution to that model as in equation (5.1). In Chapter 2, I have presented the above vector MA representation in terms of the system matrices of the FIA model (see equation (2.6)). w_t represent the innovations to the agents' information set, which I have already labeled as (true) economic shocks.

According to Lippi and Reichlin (1994), a MA representation is called fundamental if

(i) the disturbances are white noise,

(ii) the MA operator is a matrix of rational functions in L with no poles inside the unit circle, and

(iii) the MA lag polynomial has no roots smaller than one in modulus.[3]

The conditions given in (i) and (ii) hold for equation (5.1). If (iii) was also satisfied, I would say that w_t is fundamental for y_t – meaning that w_t lies in the linear space spanned by current and lagged values of y_t.

In the literature, the terms nonfundamentalness and noninvertibility are often used interchangeably. If one would be more precise, both terms have to be distinguished from each other (see Alessi et al. (2011)). In particular, one can differentiate among three cases:

(1) If all roots of $det\,[M(L)]$ are larger than unity, the representation is fundamental and invertible (in the past), i.e., the inverse of $M(L)$ has a representation only in nonnegative powers of L.

---

[3]Note that the above conditions for fundamentalness refer to the square case, i.e., the number of shocks m equals the number of observables k (see *Assumption 1*). If $k \neq m$, the third condition has to be replaced. Instead, one has to state that the MA representation in equation (5.1) is *fundamental* if the rank of $M(L)$ is equal to m (see, e.g., Rozanov (1967) or Proposition F in Forni et al. (2009)).

(2) If at least one root of $det\,[M\,(L)]$ is less than unity while the remaining roots are larger than unity, the representation is nonfundamental and noninvertible (in the past), i.e., the inverse of $M\,(L)$ has a representation not only in nonnegative powers of L but also in negative powers of L.

(3) If at least one root of $det\,[M\,(L)]$ is equal to one in modulus while the remaining roots are larger than unity, the representation is fundamental and noninvertible (in the past and in the future), i.e., the inverse of $M\,(L)$ does not exist.

An example of the third possibility would be if I thought of y_t denoting the first difference of an $I\,(1)$ process. A root of unity would indicate a cointegrating relationship between the observables. It is well known that, in this case, an associated VAR representation in first differences does not exist due to noninvertibility.[4] I keep this possibility in mind but nonetheless abide by the literature and use nonfundamentalness and noninvertibility synonymously, i.e., whenever I use these terms, I refer to the first and second category, respectively. The first case has already been covered in the previous chapters, and I examine the second case in the following.

Suppose that the FIA model exhibits a nonfundamental representation, i.e., that the previous condition (iii) is violated. As I have already shown in Chapter 2, there is no perfect match between the FIA model and the LIE model (in structural form) in this case. Consequently, the true shocks and the structural innovations that would be identified by the econometrician do not coincide because she cannot detect the true shocks by conditioning on the history of the observations y_t. In particular, the structural innovations are an infinite sum of (present and) past realizations of the true shocks (see equation (2.28)). Then, the impulse responses do not accord with each other as a result of the different MA representations.

Throughout the thesis, I exemplify two ways of deriving the Wold MA representation with which the econometrician is faced. In the subsequent sections, I show that applying the time-invariant version of the Kalman filter leads to such a fundamental representation. A second way of converting a nonfundamental into a fundamental representation will be illustrated in Chapter 6. While the first method makes use of the state space representation of the model, the second approach directly relates to the MA representation.

5.2.2 Quantifying the wedge of nonfundamentalness

Recall from Chapter 2 that the FIA model can be expressed as

$$x_t = \mathcal{A}x_{t-1} + \mathcal{B}w_t, \tag{5.2}$$

$$y_t = \mathcal{C}x_{t-1} + \mathcal{D}w_t, \tag{5.3}$$

---

[4]One possible way of proceeding would then be to estimate a VECM.

where w_t are the true shocks that are supposed to be white noise with the corresponding variance-covariance matrix normalized to the identity matrix. The associated LIE model in reduced form is

$$\widehat{x}_t = \mathcal{A}\widehat{x}_{t-1} + Ka_t, \tag{5.4}$$

$$y_t = \mathcal{C}\widehat{x}_{t-1} + a_t, \tag{5.5}$$

where a_t represents the one period ahead forecast errors in y_t. Defining $v_t \equiv \mathcal{D}w_t$ yields the direct counterpart as a linear combination of the true shocks. I can calculate the variance-covariance matrices of the (reduced form) shocks v_t and a_t to obtain $E(v_t v_t') = \mathcal{D}\mathcal{D}'$ and $E(a_t a_t') = \mathcal{C}\overline{\Sigma}\mathcal{C}' + \mathcal{D}\mathcal{D}'$. If there is nonfundamentalness in the FIA model, the state is not completely observable, implying that $x_t \neq \widehat{x}_t$ so that its forecast error variance $\overline{\Sigma} > 0$. Hence, $E(a_t a_t') - E(v_t v_t')$ is a positive definite matrix, i.e., a_t captures less information about y_t than does v_t.

Thus, the ratio of the determinants of $E(a_t a_t')$ and $E(v_t v_t')$ can be used as an indicator of the wedge between (fundamental) innovations and (nonfundamental) economic shocks. Specifically, I construct the indicator as the log of this ratio:[5]

$$\varphi \equiv log\left(det\left[E\left(a_t a_t'\right)\right]\right) - log\left(det\left[E\left(v_t v_t'\right)\right]\right). \tag{5.6}$$

This calculation necessitates finding a solution to the algebraic matrix Riccati equation given in (2.12), which can also be written as

$$\Sigma = \widetilde{\mathcal{A}}\Sigma\widetilde{\mathcal{A}}' + \widetilde{\mathcal{Q}} - \widetilde{\mathcal{A}}\Sigma\mathcal{C}'\left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\Sigma\widetilde{\mathcal{A}}', \tag{5.7}$$

where $\widetilde{\mathcal{A}} = \mathcal{A} - \mathcal{B}\mathcal{D}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}$ and $\widetilde{\mathcal{Q}} = \mathcal{B}\mathcal{B}' - \mathcal{B}\mathcal{D}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{D}\mathcal{B}'$.[6] Normally, there are several (numerical) algorithms at hand for solving this equation.[7] In general, these algorithms hinge on the assumption that $\mathcal{A}$ (or $\widetilde{\mathcal{A}}$) is nonsingular. In the case of anticipation effects in a DSGE model, this assumption can fail (see the benchmark model from Chapter 7 or the simple model in Subsection 5.3.3). When the assumption fails, one can make use of a particular algorithm, e.g., the one proposed by Pappas et al. (1980), which is based on the solution of a generalized eigenvalue problem and does not require the transition matrix (i.e., $\mathcal{A}$ or $\widetilde{\mathcal{A}}$) to be invertible.[8]

---

[5]Note that my way of computing this measure is similar to the determination of standard statistical criteria, such as the AIC or Schwarz information criterion, for lag order selection in VAR models. Sims (2012) uses only the determinant of $\mathcal{C}\overline{\Sigma}\mathcal{C}'$ as an indicator of the wedge. I argue that this is not a convenient choice for the comparison of $E(a_t a_t')$ and $E(v_t v_t')$ because it neglects the information contained in $\mathcal{D}\mathcal{D}'$.

[6]This transformation originates in the typical conversion of an optimal linear regulator problem with cross products in states and controls into one without interdependencies between states and controls. Both problems remain equivalent as do the closed loop transition matrices and the associated matrix Riccati equations (see Ljungqvist and Sargent (2004)). I demonstrate the equivalence of equations (5.7) and (2.12) in the appendix.

[7]See Anderson et al. (1996) for a description and discussion.

[8]A demonstrative calculation is given in the appendix with respect to the multivariate example of Subsection

One way to analyze the implications caused by nonfundamentalness is to contrast the impulse responses of the FIA model with their counterparts from the LIE model. I therefore need a further step that allows the comparison because solving the Riccati difference equation yields only the reduced form representation of the LIE model. Accordingly, I have to impose appropriate restrictions on the reduced form representation to compute the rotation matrix $\widehat{\mathcal{D}}$, which connects the forecast errors in y_t with the structural innovations identified by the econometrician as $a_t = \widehat{\mathcal{D}}\widehat{\varepsilon}_t$. The econometrician would orient herself based on what she conjectured to be the true model, so these restrictions have to be in line with the FIA model. After determining $\widehat{\mathcal{D}}$, the impulse responses can be calculated for the LIE model (see equation (2.18)).

To measure the discrepancy between the impulse responses coming from the (nonfundamental) FIA model and the (fundamental) LIE model representation, I compute an indicator that I call the mean average weighted error (MAWE). It builds the average of all percentage deviations (in absolute terms) between the impulse responses at each point i of response horizon h, where each deviation is separately weighted by the relative share of the response at point i to the sum of all response estimates over the entire horizon h:

$$MAWE_{jk} \;\equiv\; \frac{1}{h}\sum_{i=1}^{h}\left(\frac{|\widehat{m}_{ijk} - m_{ijk}|}{m_{ijk}}\frac{m_{ijk}}{\sum_{l=1}^{h}|m_{ljk}|}\right) = \frac{1}{h}\sum_{i=1}^{h}\left(\frac{|\widehat{m}_{ijk} - m_{ijk}|}{\sum_{l=1}^{h}|m_{ljk}|}\right), \quad (5.8)$$

where m_{ijk} $(\widehat{m}_{ijk})$ is the response of variable j to shock k at point i in the FIA model (LIE model). This statistic is more convenient for my purposes than standard measures because it better accounts for the general shape and magnitude of the impulse responses. As will be seen in the upcoming examples, some of the impulse responses feature a weak reaction of the variables to the news shock during the anticipation horizon followed by a distinct jump when the anticipated shock actually materializes. In this case, a standard measure that computes the error as percentage deviation would punish small absolute deviations between the LIE and FIA impulse responses within the anticipation horizon relatively severely in comparison to deviations between the LIE and FIA impulse responses after materialization of the shock. The MAWE indicator might be more suitable in this context due to its weighting scheme. Moreover, the MAWE also allows me to compare the performance of different compositions of the vector of observables y_t, which would not be appropriate when using a measure whose calculation is based on the deviation in levels, as will be shown in Subsection 5.3.4.

Thus, I use two indicators to quantify the outcomes in cases of nonfundamentalness: a single measure φ, which delivers a general statement about the extent of the systematic error, and the impulse response deviations captured by the MAWEs, which provide more detailed insight into how the systematic error is distributed among the chosen variables. I utilize

___

5.3.3.

both measures in the following examples.

5.3 Implications of nonfundamentalness in the presence of news shocks

5.3.1 General setup

In Chapter 3, I have derived the policy function for capital in the workhorse model introduced in Chapter 2 in which I assumed two stochastic processes for technology and preferences. In the following, I consider only one stochastic process to keep the setting for my discussion simple. The objective is to present analytical examples that disclose the general idea in the upcoming chapters and make connections to Chapter 2. In this respect, the cost of analytical computations is the usage of a simplified setting. Nevertheless, I choose a setting that permits a fairly generalization in terms of model extensions within the theoretical framework.

In the analytical examples, I initially leave the specific type of the exogenous process open, i.e., I use a proxy s_t that can reflect neutral technology, IST, taxes or government spending, for instance. My goal is to illustrate the implications of nonfundamentalness, which arise because of the inclusion of a commonly used kind of news shock in the theoretical framework. Consequently, I specify that particular setup in terms of my workhorse model and designate the stochastic process. The labeling of the process does not generally affect the implications I show, but it sets the direction of the forthcoming sections and topics.

If I assume a standard utility function as in Chapter 3, I obtain the well-known second order difference equation after log-linearizing the model around its nonstochastic steady state,

$$\phi_0 E_t \left[\widetilde{k}_{t+1}\right] + \phi_1 \widetilde{k}_t + \phi_2 \widetilde{k}_{t-1} = \phi_3 \widetilde{s}_t + \phi_4 E_t \left[\widetilde{s}_{t+1}\right] , \qquad (5.9)$$

where $\widetilde{k}_t$ is the log deviation of (the optionally stationarized stock of) capital from its steady state value, and $\widetilde{s}_t$ is the log-linearized stochastic process. The coefficients ϕ_j (with $j = 0, 1, 2, 3, 4$) depend on the deep model parameters. Solving the stable root backward and the unstable root forward unveils how the decision rule for capital depends on future expected shocks:

$$\widetilde{k}_t = \phi_{kk}\widetilde{k}_{t-1} - \frac{\phi_{kk}\phi_3}{\phi_2}\widetilde{s}_t - \frac{\phi_{kk}\left(\phi_4 + \omega\phi_3\right)}{\phi_2}\sum_{j=0}^{\infty} \omega^j E_t \left[\widetilde{s}_{t+j+1}\right] , \qquad (5.10)$$

where ϕ_{kk} is the stable root of the characteristic polynomial associated with the second order difference equation in $\widetilde{k}_t$. ω denotes the inverse of the unstable root to the characteristic

polynomial.[9] Note that I only consider model specifications for which ω has positive values so that I can adhere to the outcome such that $0 < \omega < 1$. According to Mertens and Ravn (2010), ω can be interpreted as an anticipation rate because it measures the magnitude at which agents discount expected future shocks in terms of their effect on the current capital stock.

Incorporating anticipated shocks into the model reveals how nonfundamentalness can emerge. A simple way to do that is to introduce a delay between the announcement of a shock and its materialization. For example, I can write the stochastic process in log-linearized form as

$$\widetilde{s}_t = w^s_{q,t-q}, \tag{5.11}$$

where $w^s_{q,t-q}$ is supposed to be white noise. It is the news that arrives q periods prior to its effect on the exogenous variable and affects the economy. This type of news shock is commonly used in the literature.[10] The policy function of capital is then given by

$$\widetilde{k}_t = \phi_{kk}\widetilde{k}_{t-1} + \phi_{ks}w^s_{q,t-q} + \phi_{ks,1}\sum_{i=0}^{q-1}\omega^{q-1-i}w^s_{q,t-i}, \tag{5.12}$$

where $\phi_{ks} = -\frac{\phi_{kk}\phi_3}{\phi_2}$ and $\phi_{ks,1} = -\frac{\phi_{kk}(\phi_4+\phi_3\omega)}{\phi_2}$. This function is the starting point for the following examples.

5.3.2 Univariate example

To ease the computations in this subsection, I assume that $\phi_3 = 0$, implying that $\phi_{ks} = 0$. This presumption restricts the potential underlying model specifications, but I lift the assumption in the subsequent section in which I find similar outcomes to those observed here.

Consider the last term on the right-hand side of equation (5.12), which reflects the effects of the expected future path of the exogenous variable. Because the agents anticipate future changes, discounted values of the latter impinge on the current level of capital. What might be irritating at first sight is the "pervasive" (Leeper and Walker (2011), Leeper et al. (2013)) result that the older the news, the less it is discounted (i.e., the more it is weighted) by the agents. This makes sense: the further in the past the news is announced, the closer the associated (future) materialization to the current period. Thus, it is discounted less. Notice that this "pervasive" effect leads to the nonfundamental property. As will be seen, the econometrician weights the news conversely. Note that as long as the anticipation horizon q does not exceed one period, nonfundamentalness does not arise because the policy function

---

[9]See Ljungqvist and Sargent (2004) for more details.
[10]Examples can be found in the web appendix for Beaudry and Lucke (2010), in Mertens and Ravn (2010) and in Gunn and Johri (2013). For the investigation of other types of information flows, I refer the reader to the articles by Leeper and Walker (2011), Leeper et al. (2013) and Beaudry and Portier (2014a).

reduces to a simple AR(1) process with $q = 1$.

Supposing $q = 2$ leads to an ARMA(1,1) representation of the policy function, which can be formulated in terms of the FIA model representation. Therefore, write the ARMA process in state space form by setting

$$x_t = \frac{1}{\omega \phi_{ks,1}} \widetilde{k}_t - w_{2,t}^s, \; y_t = \frac{1}{\omega \phi_{ks,1}} \widetilde{k}_{t-1} \text{ and } w_t = w_{2,t-1}^s.$$

The state and observation equations then read

$$x_t = \phi_{kk} x_{t-1} + \left(\frac{1}{\omega} + \phi_{kk} \right) w_t, \tag{5.13}$$

$$y_t = x_{t-1} + w_t. \tag{5.14}$$

When assigning the system matrices $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ to the above representation, one can compute $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$ as $-1/\omega$ to see that the *invertibility condition* is not satisfied.

This simple example easily allows the calculation of the LIE model representation. Because there is only one state variable, and the observability and reachability matrices are trivially given as the coefficients captured by $\mathcal{C}$ and $\mathcal{B}$, *Assumption 3* holds so that the conditions for convergence of the Kalman filter are fulfilled. One can solve the associated Riccati equation straightforwardly to obtain

$$\overline{\Sigma} = \frac{1}{\omega^2} - 1.$$

The corresponding Kalman gain is

$$K = \phi_{kk} + \omega.$$

The LIE model in reduced form is then given by

$$\widehat{x}_t = \phi_{kk} \widehat{x}_{t-1} + (\omega + \phi_{kk}) a_t, \tag{5.15}$$

$$y_t = \widehat{x}_{t-1} + a_t. \tag{5.16}$$

Thus, $\mathcal{A} - K\mathcal{C} = -\omega$, i.e., the *minimum phase condition* is maintained.

The econometrician identifies $a_t = \frac{1}{\omega} \widehat{\varepsilon}_t$, as $E(a_t a_t') = \mathcal{C}\overline{\Sigma}\mathcal{C}' + \mathcal{D}\mathcal{D}' = \frac{1}{\omega^2}$, and thus estimates the LIE model in structural form as

$$\widehat{x}_t = \phi_{kk} \widehat{x}_{t-1} + \left(\frac{\phi_{kk}}{\omega} + 1 \right) \widehat{\varepsilon}_t, \tag{5.17}$$

$$y_t = \widehat{x}_{t-1} + \frac{1}{\omega}\widehat{\varepsilon}_t, \tag{5.18}$$

where $\widehat{\varepsilon}_t$ has unit variance. Solving equation (5.17) for $\widehat{x}_t$, substituting into equation (5.18) and multiplying both sides by $(1 - \phi_{kk}L)$, I obtain the ARMA(1,1) process $y_t - \phi_{kk}y_{t-1} = \frac{1}{\omega}\widehat{\varepsilon}_t + \widehat{\varepsilon}_{t-1}$, in which the lagged error is relatively less weighted (by one) than the error in the current period (by $1/\omega$). The econometrician attaches more importance to more recent news in contrast to the agents. The lower the anticipation rate, the less the agents discount older news relative to more recent news, i.e., the lower the anticipation rate, the more "pervasive" the implications.

The preceding result suggesting that the extent of the difference between the agent and econometrician's model depends on the magnitude of the anticipation rate finds support when considering the two measures of Subsection 5.2.2. Because $v_t \equiv \mathcal{D}w_t$ with $\mathcal{D} = 1$, the wedge of nonfundamentalness is simply $\varphi = -2log(\omega)$, i.e., the lower the anticipation rate, the higher the wedge. The same outcome is expected for the gap between the impulse response of the agents and the one computed by the econometrician. To show this, I derive the MA representation of the FIA model as

$$y_t = \left(1 + \left(\frac{\frac{1}{\omega} + \phi_{kk}}{1 - \phi_{kk}L}\right)L\right)w_t \tag{5.19}$$
$$= \left(\frac{1 + \frac{1}{\omega}L}{1 - \phi_{kk}L}\right)w_t.$$

The MA representation of the LIE model is

$$y_t = \left(\frac{1}{\omega} + \left(\frac{1 + \frac{\phi_{kk}}{\omega}}{1 - \phi_{kk}L}\right)L\right)\widehat{\varepsilon}_t \tag{5.20}$$
$$= \left(\frac{L + \frac{1}{\omega}}{1 - \phi_{kk}L}\right)\widehat{\varepsilon}_t.$$

Evidently, the impact effect on the Wold representation is larger than that in the original model, whereas this result is reversed for the effects at longer horizons. Obviously, the lower the ω, the higher the deviation between the impulse responses. I calculate the total deviation in absolute terms as $2(1/\omega - 1)$. Division by $(1 + 1/\omega)/(1 - \phi_{kk})$ yields the MAWE measure as $2(1 - \phi_{kk})(1 - \omega)/(1 + \omega)$ with the only difference that I do not divide this term by the length of a specific horizon because I deal with the infinite horizon here. It can be confirmed that the MAWE indicator is, as in the description of the "pervasive" implications above, negatively correlated with the anticipation rate.

5.3.3   Extension to the multivariate case

I now relax the assumption that $\phi_3 = 0$ and augment the stochastic process such that it includes another shock that is a standard surprise (white noise) shock $w_{0,t}^s$:

$$\widetilde{s}_t = w_{0,t}^s + \lambda_s w_{q,t-q}^s, \tag{5.21}$$

where λ_s represents the relative standard deviation of the news shock in comparison to the surprise shock. According to equation (5.21), agents no longer have perfect foresight on the level of $\widetilde{s}_t$. For $q = 2$, I can write the policy function in capital as[11]

$$\widetilde{k}_t = \phi_{kk}\widetilde{k}_{t-1} + \phi_{ks}w_{0,t}^s + \omega\phi_{ks,1}\lambda_s w_{2,t}^s + \phi_{ks,1}\lambda_s w_{2,t-1}^s + \phi_{ks}\lambda_s w_{2,t-2}^s, \tag{5.22}$$

where all the coefficients are defined as before. Because there are two shocks, I set up a two-dimensional vector of observables to formulate the state space system in terms of the FIA model representation

$$\begin{pmatrix} \widetilde{k}_t \\ w_{2,t}^s \\ w_{2,t-1}^s \end{pmatrix} = \underbrace{\begin{pmatrix} \phi_{kk} & \phi_{ks,1}\lambda_s & \phi_{ks}\lambda_s \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}}_{=\mathcal{A}} \begin{pmatrix} \widetilde{k}_{t-1} \\ w_{2,t-1}^s \\ w_{2,t-2}^s \end{pmatrix} + \underbrace{\begin{pmatrix} \phi_{ks} & \omega\phi_{ks,1} \\ 0 & 1/\lambda_s \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \lambda_s \end{pmatrix}}_{=\mathcal{B}} \begin{pmatrix} w_{0,t}^s \\ w_{2,t}^s \end{pmatrix},$$

$$\tag{5.23}$$

$$\begin{pmatrix} \widetilde{s}_t \\ \widetilde{k}_t \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 0 & \lambda_s \\ \phi_{kk} & \phi_{ks,1}\lambda_s & \phi_{ks}\lambda_s \end{pmatrix}}_{=\mathcal{C}} \begin{pmatrix} \widetilde{k}_{t-1} \\ w_{2,t-1}^s \\ w_{2,t-2}^s \end{pmatrix} + \underbrace{\begin{pmatrix} 1 & 0 \\ \phi_{ks} & \omega\phi_{ks,1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \lambda_s \end{pmatrix}}_{=\mathcal{D}} \begin{pmatrix} w_{0,t}^s \\ w_{2,t}^s \end{pmatrix}.$$

$$\tag{5.24}$$

Thereby, I take a simplifying step of assigning the endogenous state variable $\widetilde{k}_t$ and the exogenous state variable $\widetilde{s}_t$ to the vector of observables, which contradicts the usual assumption about the unobservability of the states but is convenient for demonstration purposes. Notice that $\widetilde{s}_t$ does not appear in the state vector because I have already documented the minimal state space representation in equations (5.23) and (5.24). To check this, one can verify that the rank of the observability and the reachability matrix is equal to three (see the appendix), which is the dimension of the state vector, i.e., $q + 1$.

If $w_{0,t}^s$ and $w_{2,t}^s$ are interpreted as technology shocks, for instance, one could think of $\widetilde{s}_t$ as an

---

[11]Eliminating the assumption that $\phi_3 = 0$ leads to the inclusion of the news shock that dates back q periods. This shock does not fit in the weighting scheme of all other news shocks of the interval t to $t - q + 1$. Nevertheless, the main conclusion of the previous subsection still holds within this setting.

observable measure of technological progress. In the literature, there are different attempts to quantify technological improvement, which range from a simple Solow residual to more sophisticated approaches, e.g., those by Basu et al. (2006). Capital is chosen as the second observable for ease of computation. In principle, the implications I highlight will not be different if I use any other endogenous variable as the second component in the vector of observables.

Having defined the new state space system, it is now possible to derive the corresponding LIE model representation. First, I ensure that nonfundamentalness is existent by calculating the nonzero eigenvalue of

$$
\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C} = \begin{pmatrix} 0 & 0 & 0 \\ -\frac{\phi_{kk}}{\omega\phi_{ks,1}\lambda_s} & -\frac{1}{\omega} & 0 \\ 0 & 1 & 0 \end{pmatrix}
$$

as $-1/\omega$. Because the conditions with respect to the convergence of the steady state Kalman filter are satisfied, I can find a unique solution to the algebraic matrix Riccati equation. In the appendix, I compute the analytical expressions. Here, I only concentrate on selected results that bear further commonalities with the foregoing univariate example.

The solution to the Riccati equation is given as

$$
\overline{\Sigma} = \left(\frac{1-\omega^2}{1+\lambda_s^2\omega^4}\right) \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -\omega \\ 0 & -\omega & \omega^2 \end{pmatrix}.
$$

Hence, I can calculate the variance-covariance matrix of the forecast errors in y_t, which reads

$$
E\left(a_t a_t'\right) = \begin{pmatrix} \chi_1 & \chi_2 \\ \chi_2 & \frac{\left(\chi_2^2+\lambda_s^2\phi_{ks,1}^2\right)}{\chi_1} \end{pmatrix},
$$

where $\chi_1 = \frac{1+\lambda_s^2\omega^2}{1+\lambda_s^2\omega^4}$ and $\chi_2 = \frac{\phi_{ks}\left(1+\lambda_s^2\omega^2\right)-\lambda_s^2\omega\phi_{ks,1}\left(1-\omega^2\right)}{1+\lambda_s^2\omega^4}$. The determinant of the variance-covariance matrix is $\lambda_s^2\phi_{ks,1}^2$. Because the determinant of

$$
\mathcal{D}\mathcal{D}' = \begin{pmatrix} 1 & \phi_{ks} \\ \phi_{ks} & \lambda_s^2\omega^2\phi_{ks,1}^2 + \phi_{ks}^2 \end{pmatrix}
$$

is $\lambda_s^2\omega^2\phi_{ks,1}^2$, I end up, as in the univariate example, with the wedge of nonfundamentalness $\varphi = -2log\left(\omega\right)$.

In the above examples, I have focused on situations where $q = 2$. In the appendix, I general-

ize the result for the wedge as $-2(q-1)log(\omega)$. However, anticipation horizons longer than two periods lead to analytical expressions that are difficult to handle. Therefore, I switch to numerical calculations, which also facilitate the illustration in the case of further model extensions, in the following subsection.

5.3.4 Numerical simulations

For the numerical exercises, I use a neoclassical production function of a Cobb-Douglas type with constant returns to scale, where α denotes the output elasticity of labor (N_t), and the standard accumulation equation for capital (K_t), as presented in Chapter 2. I specify the other relevant functions in the workhorse model as follows. The household's instantaneous utility at time t is described by

$$u\left(C_t, N_t\right) = \frac{C_t^{1-\sigma}\left(1-\psi N_t\right)^{1-\sigma}-1}{1-\sigma}, \tag{5.25}$$

where C_t is consumption and $\sigma, \psi > 0$. I include one stochastic variable, which is labor-augmenting technology that follows an autoregressive process of order one subject to a surprise ($w_{0,t}^\theta$) and an anticipated shock ($w_{q,t-q}^\theta$):

$$log\left(\theta_t\right) = \rho_\theta log\left(\theta_{t-1}\right) + \sigma_\theta w_{0,t}^\theta + \lambda_\theta \sigma_\theta w_{q,t-q}^\theta, \tag{5.26}$$

where $0 < \rho_\theta < 1$. σ_θ denotes the standard deviation of the technology shocks, which are supposed to be white noise. λ_θ is the relative weight of the standard deviation of the anticipated shock with respect to the surprise shock. The model solution for any endogenous variable z_t (including capital) in terms of log-deviation from its corresponding steady state value (indicated by lowercase letters with "~") is then given by the policy function

$$\widetilde{z}_t = \phi_{zk}\widetilde{k}_{t-1} + \phi_{z\theta}\widetilde{\theta}_t + \phi_{z\theta,1}\sum_{i=0}^{q-1}\omega^{q-1-i}\lambda_\theta\sigma_\theta w_{q,t-i}^\theta. \tag{5.27}$$

Because equation (5.26) represents a zero mean process, $\widetilde{\theta}_t$ is simply $log\left(\theta_t\right)$. Notice that the "pervasive" discounting occurs for all endogenous variables.

Using the policy functions, assuming that technology is observable and defining a two-dimensional vector of observables, I derive the vector MA representation of the model solution as[12]

---

[12]Recall the previous explanations of the possibility of finding a measure of technological progress. I could, e.g., easily define a variable that represents the Solow residual (or TFP) by setting $A_t = \theta_t^\alpha$.

$$
\begin{pmatrix} \widetilde{\theta}_t \\ \widetilde{z}_t \end{pmatrix} =
$$

$$
\underbrace{\begin{pmatrix} \frac{1}{(1-\rho_\theta L)} & \frac{L^q}{(1-\rho_\theta L)} \\ \left(\frac{\phi_{z\theta}}{(1-\rho_\theta L)} + \frac{\phi_{k\theta}\phi_{zk}}{(1-\phi_{kk}L)(1-\rho_\theta L)} L \right) & \begin{matrix} \left(\frac{\phi_{z\theta}}{(1-\rho_\theta L)} + \frac{\phi_{k\theta}\phi_{zk}}{(1-\phi_{kk}L)(1-\rho_\theta L)} L \right) L^q \\ + \left(\phi_{z\theta,1} + \frac{\phi_{k\theta,1}\phi_{zk}}{(1-\phi_{kk}L)} L \right) \Theta(L) \end{matrix} \end{pmatrix} \Sigma_w}_{= M(L)} \begin{pmatrix} w^\theta_{0,t} \\ w^\theta_{q,t} \end{pmatrix},
$$

$$(5.28)$$

where

$$
\Sigma_w = \begin{pmatrix} \sigma_\theta & 0 \\ 0 & \lambda_\theta \sigma_\theta \end{pmatrix}
$$

and

$$
\Theta(L) = \omega^{q-1} + \omega^{q-2}L + \ldots \omega L^{q-2} + L^{q-1} .
$$

The roots of $\Theta(L)$ can be computed as $q-1$ roots of unity multiplied by ω.[13] Hence, these roots lie all on one circle with radius ω in the complex plane.[14] These roots are the zeros of the MA lag polynomial $det[M(L)]$. Because my previous assumption that $0 < \omega < 1$ still holds, there is nonfundamentalness when the anticipation horizon exceeds one period.[15] Notice that $det[M(L)]$ has one additional root that is not directly related to ω or $\Theta(L)$. It can be calculated as $\phi_{z\theta,1}/(\phi_{z\theta,1}\phi_{kk} - \phi_{zk}\phi_{k\theta,1})$.[16]

---

[13]Any complex number $\mathcal{Z}$ that satisfies the equation $\mathcal{Z}^m = 1$, where m is some integer, is called a root of unity.
[14]Notice that one root of $\Theta(L)$ is equal to $-\omega$ when q is even.
[15]See also my explanations in Subsection 5.3.1 with respect to the assumption of possible values for ω.
[16]In the case that capital is the second observable, the additional root would not exist.

Moreover, I set up the FIA model representation[17]

$$
\begin{pmatrix}
\widetilde{k}_t \\
\widetilde{\theta}_t \\
w^\theta_{q,t} \\
w^\theta_{q,t-1} \\
\vdots \\
w^\theta_{q,t-q+2} \\
w^\theta_{q,t-q+1}
\end{pmatrix}
=
\underbrace{
\begin{pmatrix}
\phi_{kk} & \rho_\theta \phi_{k\theta} & \omega^{q-2}\phi_{k\theta,1} & \cdots & \omega\phi_{k\theta,1} & \phi_{k\theta,1} & \phi_{k\theta} \\
0 & \rho_\theta & 0 & \cdots & 0 & 0 & 1 \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & 0 & \frac{1}{\lambda_\theta \sigma_\theta} & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \cdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \frac{1}{\lambda_\theta \sigma_\theta} & 0 & 0 \\
0 & 0 & 0 & \cdots & 0 & \frac{1}{\lambda_\theta \sigma_\theta} & 0
\end{pmatrix}
}_{=\,\mathcal{A}}
\Sigma_1
\begin{pmatrix}
\widetilde{k}_{t-1} \\
\widetilde{\theta}_{t-1} \\
w^\theta_{q,t-1} \\
w^\theta_{q,t-2} \\
\vdots \\
w^\theta_{q,t-q+1} \\
w^\theta_{q,t-q}
\end{pmatrix}
$$

$$
+
\underbrace{
\begin{pmatrix}
\phi_{k\theta} & \omega^{q-1}\phi_{k\theta,1} \\
1 & 0 \\
0 & \frac{1}{\lambda_\theta \sigma_\theta} \\
0 & 0 \\
\vdots & \vdots \\
0 & 0 \\
0 & 0
\end{pmatrix}
}_{=\,\mathcal{B}}
\Sigma_w
\begin{pmatrix}
w^\theta_{0,t} \\
w^\theta_{q,t}
\end{pmatrix},
\tag{5.29}
$$

$$
\begin{pmatrix}
\widetilde{\theta}_t \\
\widetilde{z}_t
\end{pmatrix}
=
\underbrace{
\begin{pmatrix}
0 & \rho_\theta & 0 & \cdots & 0 & 0 & 1 \\
\phi_{zk} & \rho_\theta \phi_{z\theta} & \omega^{q-2}\phi_{z\theta,1} & \cdots & \omega\phi_{z\theta,1} & \phi_{z\theta,1} & \phi_{z\theta}
\end{pmatrix}
}_{=\,\mathcal{C}}
\Sigma_1
\begin{pmatrix}
\widetilde{k}_{t-1} \\
\widetilde{\theta}_{t-1} \\
w^\theta_{q,t-1} \\
w^\theta_{q,t-2} \\
\vdots \\
w^\theta_{q,t-q+1} \\
w^\theta_{q,t-q}
\end{pmatrix}
$$

$$
+
\underbrace{
\begin{pmatrix}
1 & 0 \\
\phi_{z\theta} & \omega^{q-1}\phi_{z\theta,1}
\end{pmatrix}
}_{=\,\mathcal{D}}
\Sigma_w
\begin{pmatrix}
w^\theta_{0,t} \\
w^\theta_{q,t}
\end{pmatrix},
\tag{5.30}
$$

---

[17]Notice that the state space representation presented in equations (5.29) and (5.30) is not minimal. To obtain the minimal realization of the state space representation (for the numerical as well as the Monte Carlo simulations below), I use the Matlab command ss(sys,'minimal'). In this case, the state vector has dimension $q + 1$.

where

$$\Sigma_1 = \begin{pmatrix} I_2 & 0_{2 \times q} \\ 0_{q \times 2} & \lambda_\theta \sigma_\theta \cdot I_q \end{pmatrix},$$

and Σ_w is the same as in (5.28). Recall from Chapter 2 that the roots of $det\,[M\,(L)]$ are the reciprocals of the (nonzero) eigenvalues of $(\mathcal{A} - \mathcal{BD}^{-1}\mathcal{C})$. Hence, the matrix $(\mathcal{A} - \mathcal{BD}^{-1}\mathcal{C})$ has $q - 1$ eigenvalues with modulus $1/\omega$ and one equal to $(\phi_{kk}\phi_{z\theta,1} - \phi_{k\theta,1}\phi_{zk})\,/\phi_{z\theta,1}$.[18]

After deriving the FIA model, I can use the system matrices $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ in combination with the suitable algorithm to solve the associated Riccati equation. This leads to the LIE model in reduced form. To obtain the corresponding structural form, I imitate the econometrician's procedure, i.e., imposing appropriate restrictions on the system to identify the structural innovations. Because I have a bivariate vector of observables, one restriction is necessary (in addition to the conventional assumption of the orthogonality of the structural shocks). I constrain the contemporaneous impact matrix $\widehat{\mathcal{D}}$ to be lower triangular, i.e., the news shock has no immediate impact on technology, which is in line with the FIA model. It thus becomes possible to calculate the impulse responses of the LIE model and to compare them with the true responses.

In the following, I provide some illustrative examples. I choose output as the second observable and calibrate the model parameters as follows: $\alpha = \frac{2}{3}$, $\beta = 1.01^{-1}$, $\sigma = 1$, $\psi = 1$ and $\rho_\theta = 0.95$. Moreover, I set the standard deviation σ_θ and the relative weight of the news shock λ_θ equal to one. Given this calibration, I can calculate the eigenvalues of $(\mathcal{A} - \mathcal{BD}^{-1}\mathcal{C})$. The anticipation rate equals 0.9435, so there are $q - 1$ eigenvalues with modulus 1.06. The additional eigenvalue not related to ω equals 0.9.

I visualize the systematic error made by the econometrician by comparing her impulse responses with the original responses. Figure 5.1 presents the true impulse responses (green lines) and those resulting from the standard identification procedure of the econometrician (blue lines) for the case when $q = 2$. The upper graphs show the responses of technology, the lower graphs display the responses of output to the surprise shock (left column) and the anticipated shock (right column), respectively.

Recall that the model that has served until now as the DGP is a standard RBC model. Thus, I observe the well-known reactions of output on the surprise and the anticipated technology shock. Because I have specified a standard utility function, i.e., without habit formation, for example, a news shock generates a wealth effect such that agents increase consumption and reduce labor supply on impact. Output decreases and is accompanied by a decline in investment due to the resource constraint. During the anticipation phase, the household continues to slowly raise its level of consumption and leisure, whereas output and investment slightly decrease. When the news shock materializes, productivity boosts output to a higher level,

---

[18]In the special case when capital is the second observable, the latter eigenvalue is zero. See also footnote 16.

Figure 5.1: IMPULSE RESPONSES IN WORKHORSE MODEL WITH NEWS SHOCK $(q = 2)$
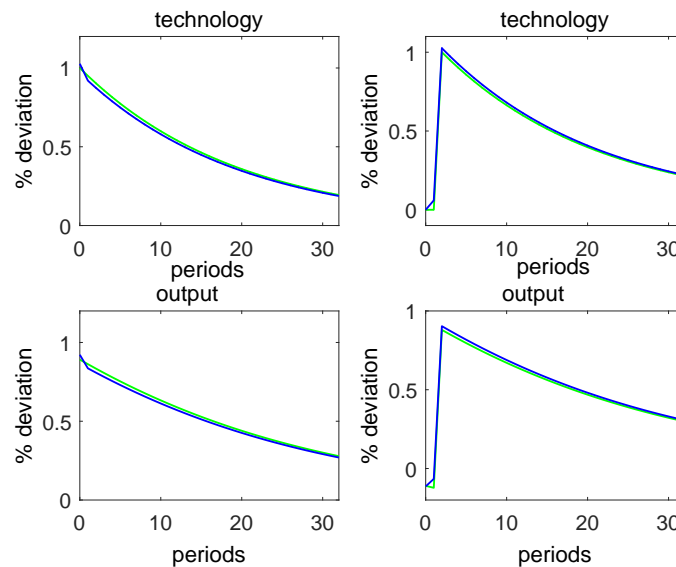
Note: The figure depicts the FIA model impulse responses (in green) and the LIE model impulse responses (in blue). Impulse responses to $w_{0,t}^{\theta}$ are shown in the left column and impulse responses to $w_{2,t}^{\theta}$ in the right column.

and the reaction of all aggregates copy their responses as in the case of the surprise shock (with a delay of q periods).

There is no substantial deviation between the true and the estimated impulse responses in this simulation. This finding changes with a longer anticipation horizon, as will be described below. Nevertheless, the contrary implications are less severe than those found by Leeper et al. (2013), for instance.[19] However, the econometrician might have to care about her conclusions about the impulse responses when addressing anticipated shocks and an anticipation horizon that is relatively long. These findings also suggest that nonfundamentalness affects not only impulse responses corresponding to the news shocks but also to the other conventional (surprise) disturbances associated with the same stochastic process.

Figure 5.2 displays the result when the anticipation horizon is fixed at $q = 8$.[20] The econometrician overestimates the impact effect of the surprise shock. This overshooting ends after three or four periods and transitions to underestimation thereafter until the variables converge back to their steady state values. Concerning the anticipated shock, the first eight periods are characterized by impulse responses in the LIE model, which lie above the true responses. In the original model, there is a distinct jump in both variables when the news shock materializes eight periods after its anticipation. The LIE model responses also show a jump after eights periods but exhibit a gradual increase in the periods before the spike and then overestimate the increase when the spike occurs.

---

[19]Leeper et al. (2013) present an example in which the econometrician would estimate a positive response of capital to an anticipated tax shock, although the opposite holds true for the underlying standard growth model. See also my explanations in Section 7.5 and the appendix to Chapter 7.

[20]See also my remarks on the length of the anticipation horizon in Subsection 7.4.1.

Figure 5.2: IMPULSE RESPONSES IN WORKHORSE MODEL WITH NEWS SHOCK ($q = 8$)

Note: The figure depicts the FIA model impulse responses (in green) and the LIE model impulse responses (in blue). Impulse responses to $w_{0,t}^\theta$ are shown in the left column and impulse responses to $w_{8,t}^\theta$ in the right column.
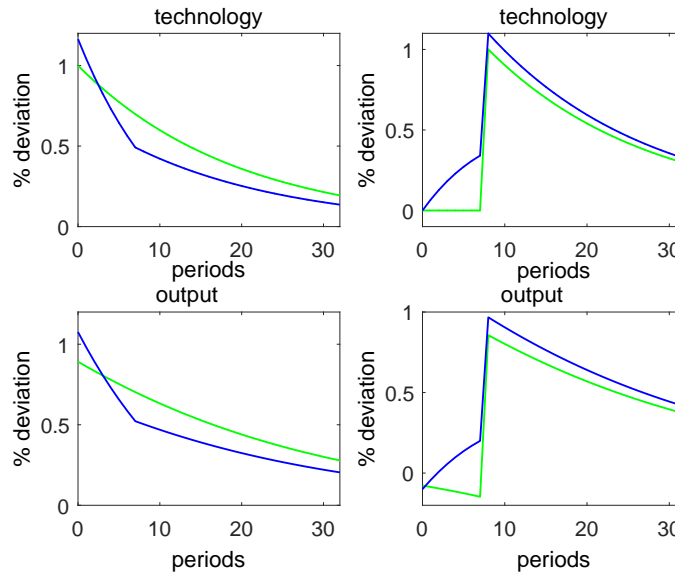
I quantify the deviations between the FIA and the LIE model by computing the corresponding MAWEs, as explained above. Table 5.1 presents the results. Each cell contains 2×2 matrices that include the computed MAWEs associated with the respective impulse responses. For example, the upper left element of a matrix corresponds to the response of technology to the surprise shock, whereas the lower right element is linked to the response of output to the news shock. I include two additional columns in the table, which display the findings for different values of λ_θ. For an illustrative comparison, I also portray the impulse responses in Figure 5.3.

The size of the relative weight λ_θ determines the structural shock on which the differences between the FIA and the LIE model become stronger. The more important the news shock, as indicated by a higher λ_θ, the lower the relative errors on its side but the higher the deviations with respect to the surprise shock. These implications reverse for small values of λ_θ. In the appendix, I show that my measure of the wedge of nonfundamentalness φ equals $-2(q-1)\log(\omega)$ and is therefore not influenced by λ_θ. The magnitude of λ_θ only shifts the systematic error in terms of the wedge to one side, but it does not affect its total level. It is not surprising that the econometrician who tries to disentangle both shocks can identify the dynamics of the shock that has a higher variance more easily. Consequently, discovering the news shock comes at the cost of disregarding the surprise shock and vice versa.

Furthermore, $\varphi = -2(q-1)\log(\omega)$ is also independent of the choice of the second observable in the system because the policy function in equation (5.27) holds for any endogenous variable of the model, i.e., the decision about their current level inherits the same inverse weighting scheme of old news for all variables, which leads to the nonfundamentalness property. As in the previous case with varying values of λ_θ, the choice of the second observ-

Figure 5.3: IMPULSE RESPONSES IN WORKHORSE MODEL WITH NEWS SHOCK (VARYING λ_θ)
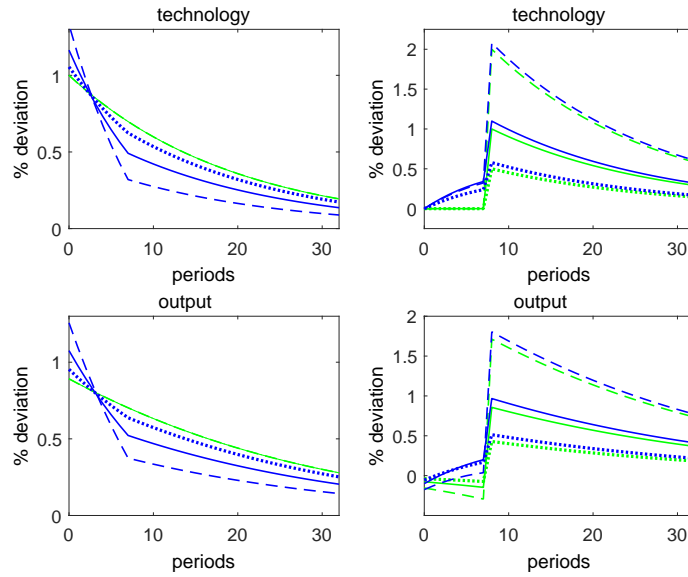
Note: The figure depicts the FIA model impulse responses (in green) and the LIE model impulse responses (in blue) for three cases: $\lambda_\theta = 1$ (solid), $\lambda_\theta = 2$ (dashed) and $\lambda_\theta = 1/2$ (dotted). Impulse responses to $w_{0,t}^\theta$ are shown in the left column and impulse responses to $w_{8,t}^\theta$ in the right column.

Table 5.1: MAWE RESULTS W.R.T. WORKHORSE MODEL WITH NEWS SHOCK ($q = 8$)

	$\lambda_\theta = 1$		$\lambda_\theta = 2$		$\lambda_\theta = 1/2$	
MAWE	0.6951	0.6197	1.2789	0.2690	0.2486	0.9209
	0.6504	0.6433	1.1962	0.2827	0.2331	0.9484

Note: The table reports the computed MAWEs associated with the impulse responses displayed in Figure 5.3. MAWE values are calculated as in equation (5.8) multiplied by 100.

able only matters for how the wedge is divided between the impulse responses to the two structural shocks but not for the general magnitude of φ. To provide support for this point, I compare the dynamic responses of output and investment, respectively, to both shocks depending on whether output or investment is selected as the second observable.

Figure 5.4 presents the responses of output to the surprise shock (left column) and to the news shock (right column) in the top row and the respective impulse responses associated with investment in the bottom row for both the FIA model and the LIE model. Obviously, investment exhibits a much stronger reaction to the shocks, and the deviation between the FIA model and the LIE model becomes more apparent in this case. To enable a fair comparison, I compute the corresponding MAWEs presented in the second column of Table 5.2 to which a third row that contains the ratio of the MAWEs of investment and the MAWEs of output (i.e., the middle row divided by the top row) is added. The findings indicate that, in this setting, investment as the second observable seems to be the better choice if the goal is a closer match for the dynamic responses to the news shock but a worse choice if one wishes to concentrate on the impulse responses to the surprise shock. Note that I add a third col-

Figure 5.4: IMPULSE RESPONSES IN WORKHORSE MODEL WITH NEWS SHOCK (DIFFERENT VARIABLES)
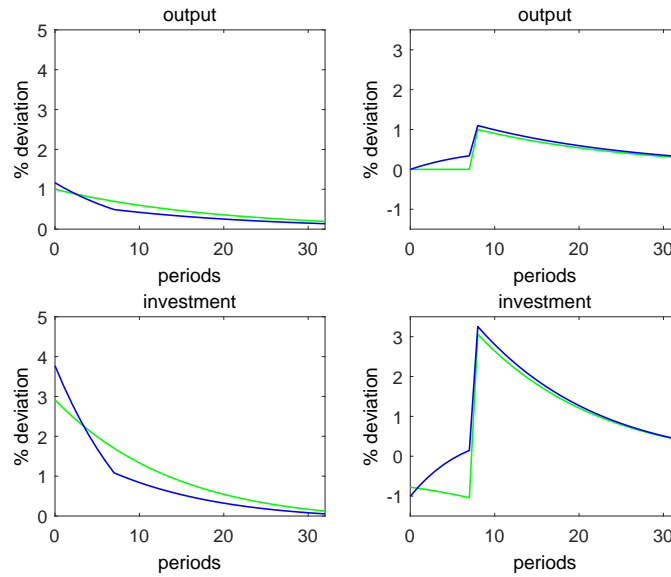
Note: The figure depicts the FIA model impulse responses (in green) and the LIE model impulse responses (in blue) related to the second observable variable in the respective system. The upper row plots the impulse response of output as the second observable, the lower row displays the impulse responses of investment as the second observable. Impulse responses to $w_{0,t}^{\theta}$ are shown in the left column and impulse responses to $w_{8,t}^{\theta}$ in the right column.

Table 5.2: MAWE RESULTS W.R.T. WORKHORSE MODEL WITH NEWS SHOCK (DIFFERENT VARIABLES)

2nd observable	MAWE		RMSE	
Output	0.6504	0.6433	0.1219	0.1244
Investment	0.8621	0.4844	0.3527	0.3646
Ratio	1.3255	0.7530	2.8926	2.9308

Note: In the upper two rows, the table reports the computed MAWEs and RMSEs associated with the impulse responses displayed in Figure 5.4. MAWE values are calculated as in equation (5.8) multiplied by 100. The bottom row shows the ratio between the values in the middle row and the values in the top row.

umn to Table 5.2, which contains the RMSEs with respect to the aforementioned impulse responses, to show that the RMSE is not a convenient indicator in this context. According to the RMSE, one would draw the misleading conclusion that output should be strictly preferred to investment.

I conclude this subsection with a final exercise in which I am concerned with the anticipation rate because this seems to be a crucial element when the implications of nonfundamentalness in the presence of news shocks are studied. I return to the system comprising output as the second observable and present the MAWE of each impulse response depending on the size of the anticipation rate in Figure 5.5. I vary the value of the discount factor β between 0.909 and 0.995 such that I obtain a range of values for ω between 0.78 and 0.96.

The more the agents discount future utility, the lower the anticipation rate and the higher

Figure 5.5: MAWE RESULTS W.R.T. WORKHORSE MODEL WITH NEWS SHOCK (VARYING ω)
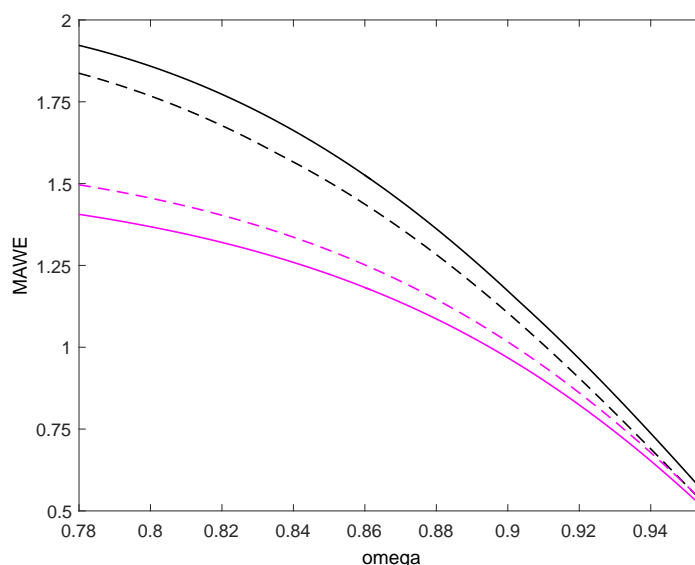
Note: The figure depicts the computed MAWEs w.r.t. the impulse responses of technology (in black) and output (in violet) depending on the value of the anticipation rate ω. MAWE values are calculated as in equation (5.8) multiplied by 100. Solid lines correspond to MAWEs associated with impulse responses to $w_{0,t}^{\theta}$. Dashed lines show the MAWEs associated with the impulse responses to $w_{8,t}^{\theta}$.

the systematic error made by the econometrician. Mertens and Ravn (2010) conduct other experiments in this context and emphasize that, for a reasonable parameter space, standard DSGE models imply an anticipation rate between 0.9 and one. Although this statement might reassure the econometrician that she need not care about nonfundamentalness, it is worth noting that even in the range of values for ω between 0.9 and one, the magnitude of the error more than doubles.

5.4 The role of additional variables in the econometrician's information set

A frequently recited practical recommendation to handle the problem of nonfundamentalness is to include "forward-looking" variables in the set of observables. Because news shocks reflect agents' anticipation of the future, the econometrician faces the challenge of finding data that reveal this information. Popular proxies for this purpose are stock prices because standard theory models these as the expected discounted sum of future profits, dividends or cash flows. Early contributions, e.g., by Fama (1990) and Schwert (1990), find that stock prices mirror changes in agents' expectations about future economic movements. Beaudry and Portier (2006) consider this aspect and show that stock price innovations, which are orthogonal to surprise changes in TFP, comprise anticipated future TFP shocks. Many subsequent studies relate to this result and estimate SVAR models to quantify the contribution of news shocks to macroeconomic fluctuation (see Haertel and Lucke (2008),

Beaudry and Lucke (2010) and Barsky and Sims (2011), for instance). Additionally, Barsky and Sims (2011) use a measure of consumer sentiment to capture the forward-looking characteristics of this variable (see also Barsky and Sims (2012) in that context). Forni et al. (2014) address this issue as well. The authors assert that the bivariate SVAR model by Beaudry and Portier (2006) as well as higher-dimensional models similar to that by Beaudry and Lucke (2010) have nonfundamental representations. Forni et al. (2014) argue that if models similar to that by Beaudry and Lucke (2010) include consumer sentiment as an additional variable, nonfundamentalness does not arise; they corroborate this statement with an empirical test of nonfundamentalness. By contrast, I base my argumentation on a more theoretical perspective in the subsequent paragraphs without leaving the framework developed in Chapter 2.

Recall the policy function in (5.27). It holds for any endogenous variable of the model, i.e., the decision about their current level inherits the inverse weighting scheme of old news for all variables, which leads to nonfundamentalness. Note that it is possible to modify the workhorse model (e.g., by adding capital adjustment costs, as will be conducted in Chapter 7) such that it involves a variable that can be interpreted as stock price and computed as the present value of the future dividend stream. This variable would exhibit the same structure in the policy function as the other endogenous variables. Hence, as long as the set of observables is extended by such endogenous variables, conditioning on more information by including them in the system does not help. A test of this is to compute the wedge of nonfundamentalness for the workhorse model augmented with further standard surprise shocks, such as a surprise preference or government spending shock, so that the vector of observables can be enlarged without facing the problem of stochastic singularity. One would end up with the same value for the wedge of nonfundamentalness as in the lower-dimensional systems.

The "forward-looking" feature of a variable, such as the stock price, is therefore not sufficient for eliminating the nonfundamentalness problem. In fact, the assumption of rational expectations is relevant for the existence of that problem. Under rational expectations, agents determine their actions using all available information about all kind of shocks, whereas the econometrician can only condition on a limited set of observable aggregates. The econometrician cannot condition on the shocks because they are not directly observable, and the objective of the econometrician is to detect them. As a result, there is misalignment between the agent and the econometrician's information sets such that nonfundamentalness manifests.

Alternatively, one could argue more generally and enlarge the econometrician's information set with variables from "outside" the model solution. Giannone and Reichlin (2006) show that this helps solve the problem only under certain assumptions. In principle, if the true shocks are nonfundamental for the small set of variables included in the vector of observ-

ables, then they cannot be fundamental for a larger set of observables. To see that, reconsider the vector MA representation of the general model solution in equation (5.1). For the moment, I do not provide a concrete specification but assume that it describes the true structural model, which has a nonfundamental MA representation. Hence, $det\,[M\,(L)]$ has at least one root inside the unit circle. One could expand the system by augmenting y_t with additional variables y_t^*, which might also involve further economic shocks. Thus, in general, I obtain

$$\begin{pmatrix} y_t \\ y_t^* \end{pmatrix} = \begin{pmatrix} M\,(L) & S\,(L) \\ M^*\,(L) & S^*\,(L) \end{pmatrix} \begin{pmatrix} w_t \\ w_t^* \end{pmatrix}. \tag{5.31}$$

Because equation (5.1) is presumed to be the true model, the additional shocks w_t^* ought to explicitly affect y_t^*. It implies that $S\,(L) = 0$. Then, the roots of the MA lag polynomial in equation (5.31) are the roots of $det\,[M\,(L)] \cdot det\,[S^*\,(L)]$. Unless all the roots of $det\,[M\,(L)]$ inside the unit circle are not canceled out by some of the roots of $det\,[S^*\,(L)]$, the larger system is also nonfundamental.

The question is whether there are possibilities for a remedy? In this context, nonfundamentalness is a problem of missing states in the set of observables, in particular, those states that represent the lagged news shocks. This can be seen in the state space representations in Subsections 5.3.3 and 5.3.4, where the state vector comprises such terms. If the econometrician finds candidates that are directly related to lagged news, she could augment the set of observables and reduce the discrepancy between her own and the agent's information set.

In the literature, there are some attempts that proceed in this direction. One strand of the literature is based on the idea of using instrumental variables as proxy for the news. The studies in that body of literature differ in how they find these proxies. One way is to use identified shocks from conventional SVAR models as instruments. Other approaches incorporate forecasts that are generated by other models. A detailed and comprehensive discussion of these topics can be found in Leeper et al. (2008, 2013), but they mainly concentrate on fiscal foresight.

A different approach is undertaken by researchers who estimate factor models of various types.[21] The basic idea behind those models is to describe the entire economy using a very large set of variables, which are driven by a small number of latent factors. Hence, all the relevant information can be reduced to a small set of factors. In a recent paper, Forni et al. (2014) suggest estimating a dynamic factor model (cf. Forni et al. (2000) and Forni et al. (2009)) or amending the basic VAR model with those factors (i.e., estimation of a factor-augmented VAR (FAVAR) model à la Bernanke et al. (2005)) if it does not contain enough information to cope with nonfundamentalness. Related to the explanations above, this proposition can be a useful way to overcome the dilemma. The idea is tested by Sims (2012), who shows in a sim-

---

[21]For a general overview of factor models and their implications, see Stock and Watson (2005, 2010).

ulation study that adding "information variables" to the conventional VAR helps curtail the wedge of nonfundamentalness in the presence of news shocks. In this study, "information variables" are modeled as noisy signals about the news and are extracted from the factor structure of the generated data.[22]

In the following chapters, my objective is to explore alternatives to the aforementioned procedures and to test their ability to resolve the problem of nonfundamentalness when news shocks are considered. I focus on situations in which the applied researcher estimates a small-scale VAR model (or conducts a corresponding subspace algorithm analysis) because she has only limited access to a small number of observable variables.

---

[22]A very recent example of how to create a direct measure of news is given by Larsen and Thorsrud (2015) and Thorsrud (2016). They use textual data from a Norwegian newspaper to develop a corresponding news index.

Chapter 6

Examining empirically based root flipping

In this chapter, I address the question of whether it is possible to handle the problem of non-fundamentalness without extending the econometrician's information set. What is known in this thesis up to now is that DSGE models with news shocks can have a nonfundamental MA representation, i.e., at least one root of the MA lag polynomial associated with the FIA model lies inside the unit circle. It is also known that, in this case, the econometrician would estimate a corresponding form of the LIE model that has the same autocovariance structure as the FIA model but differs in the location of the MA roots that are originally inside the unit circle. Specifically, estimation of the LIE model flips those roots from the inside to the outside of the unit circle, i.e., these estimated roots represent reciprocals of the originals. Is it then feasible to switch these roots back in order to retrieve the true model?

6.1 Motivation and related literature

Lippi and Reichlin (1993, 1994) give a formal introduction and illustration of the usage of so-called Blaschke matrices to transform fundamental into nonfundamental MA representations. Generally, there exists an infinitely large set of nonfundamental representations having the same autocovariance structure as the fundamental representation. An econometrician who estimates a standard VAR model ignores them a priori. Lippi and Reichlin (1994) show that the roots of the MA polynomial can produce circles of complex roots in the determinant of the VAR operator. This aspect helps limit the space of relevant nonfundamental representations. Furthermore, it allows their construction from the estimated VAR coefficients. As the econometrician uses only information coming from her estimated model, I call this approach empirically based root flipping. The general procedure of Lippi and Reichlin (1993, 1994) can be summarized by the following steps: estimate a VAR model, flip the estimated roots of its vector MA representation and impose standard restrictions based on theory to identify the structural dynamics.

This procedure unfortunately involves some drawbacks: how does the econometrician know which roots have to be flipped? Moreover, the success of the procedure crucially depends on the accuracy of the VAR estimation. Lippi and Reichlin (1993, 1994) apply this approach to pick out specific examples of nonfundamental representations that have either quite similar or different implications in comparison to the corresponding fundamental ones. Hence, if economic theory does not enable the econometrician to be sure about the location of the roots (i.e., whether they are inside or outside the unit circle), the estimation of the standard VAR model can produce misleading conclusions about the true underlying structure.

My goal is to explore whether the method of empirically based root flipping works in principle, i.e., whether it is generally possible to uncover the true dynamics by implementing this procedure. I find that the resulting nonfundamental representations are false, as is the fundamental candidate.

6.2 Background and basic concepts

Flipping the roots of a polynomial can be achieved by means of Blaschke matrices in both directions. To highlight the role of Blaschke matrices in flipping roots, I recapitulate the analytical univariate example of Subsection 5.3.2 in which the nonfundamental MA representation (resulting from the FIA model) is given by

$$y_t = \underbrace{\left(\frac{1 + \frac{1}{\omega} L}{1 - \phi_{kk} L} \right)}_{= \, m(L)} w_t, \tag{6.1}$$

and the fundamental MA process (originating from the LIE model) is

$$y_t = \underbrace{\left(\frac{L + \frac{1}{\omega}}{1 - \phi_{kk} L} \right)}_{= \, \widehat{m}(L)} \widehat{\varepsilon}_t, \tag{6.2}$$

where both shocks are white noise with unit variance and $0 < \phi_{kk}, \omega < 1$. Both representations fulfill the conditions in (i) and (ii), which characterize fundamentalness in conjunction with (iii) (see Subsection 5.2.1). The latter is only satisfied in the case of equation (6.2). As both processes have the same autocovariance structure, they are observationally equivalent, i.e., the econometrician who has access to the sequence of observations $\{y_t\}$ discovers equation (6.2), although the true DGP is equation (6.1). The reason for this is that the MA operator in equation (6.1) is invertible in negative powers of the lag operator L only. In other words, the space spanned by the present and past values of w_t is not contained in the information set with which the econometrician is equipped. That is, the sequence of shocks $\{w_t\}$ cannot

be detected by employing the history of y_t. The econometrician's information set is smaller than the agent's. She would need future realizations of y_t to identify the true shocks w_t. This is (usually) not the case, so she recovers only the sequence of shocks $\{\widehat{\varepsilon}_t\}$.

A so-called Blaschke factor can be applied to transform equation (6.2) into equation (6.1) and vice versa. Formulating the function $B(L) = \frac{L+\omega}{1+\omega L}$ helps move from equation (6.2) to equation (6.1). Setting $m(L) = \widehat{m}(L) B(L)$ and $w_t = B(L)^{-1} \widehat{\varepsilon}_t$ yield the nonfundamental representation in equation (6.1). $\frac{L+\omega}{1+\omega L}$ is called a Blaschke factor, which flips the root from $-1/\omega$ to $-\omega$ by postmultiplying $\widehat{\phi}(L)$ by $B(L)$ in the above example. Alternatively, one can use $B(L^{-1}) = \frac{L^{-1}+\omega}{1+\omega L^{-1}}$ to switch the opposite way, i.e., $\widehat{m}(L) = m(L) B(L^{-1})$ and $\widehat{\varepsilon}_t = B(L) w_t$. Additionally, $\widehat{\varepsilon}_t$ is a linear combination of current and past realizations of w_t in this example.

The Blaschke factor is defined for any complex scalar z and not only for the lag operator. Furthermore, this concept can be extended to the case with k variables. Then, there is a k-dimensional identity matrix with one diagonal element being the Blaschke factor so that an elementary Blaschke matrix has the form

$$R(z, \lambda) = \begin{pmatrix} I_{k-1} & 0 \\ 0 & \frac{z-\lambda}{1-\overline{\lambda}z} \end{pmatrix}$$

with $|\lambda| < 1$, and $\overline{\lambda}$ denotes the complex conjugate of λ (allowing for the case of complex roots). In the case of a single root, a complete Blaschke matrix is the product of the elementary one with an orthogonal matrix G (described below), either by pre- or postmultiplication, e.g., $B(z) = G \cdot R(z, \lambda)$. If there are m roots, $\lambda_1, ..., \lambda_m$, to be flipped, $B(z) = G_1 \cdot R(z, \lambda_1) \cdot G_2 \cdot R(z, \lambda_2) \cdots G_m \cdot R(z, \lambda_m)$.

In general, a Blaschke matrix $B(z)$ satisfies the following two conditions:

(I) $B(z)$ has no poles inside the (closed) unit circle; and

(II) $B(z) B^T(z^{-1}) = I$, i.e., $B(z)^{-1} = B^T(z^{-1})$,

where $B^T(z^{-1})$ is obtained after transposing and taking conjugates of $B(z)$.

To illustrate how to flip roots in a multidimensional case, now suppose that a k-dimensional vector MA process is given by

$$y_t = \widehat{N}(L) \eta_t, \tag{6.3}$$

where η_t is a k-dimensional white noise vector with $E(\eta_t \eta_t') = I_k$. I assume that the MA lag polynomial $det\left[\widehat{N}(L)\right]$ has only one (real) root at λ^{-1} with $|\lambda| < 1$. Using the Blaschke

matrix as introduced above, I can write

$$y_t = N(L)\mu_t, \tag{6.4}$$

where $N(L) = \widehat{N}(L) B(L)$, $\mu_t = B(L)^{-1} \eta_t$ with $E(\mu_t \mu_t') = I_k$ and

$$B(L) = G \cdot \begin{pmatrix} I_{k-1} & 0 \\ 0 & \frac{L-\lambda}{1-\lambda L} \end{pmatrix}.$$

G represents the rotation matrix that determines the last column of $\widehat{N}(L) G$ to include the factor $L - \lambda^{-1}$. The other columns of G are fixed due to the orthogonality condition. Thus, the choice of G is not unique. One way of computing the rotation matrix is to take a SVD of $\widehat{N}(L)$ at $L = \lambda^{-1}$ and use the matrix of right singular vectors as G. Postmultiplication of $\widehat{N}(L) G$ by the elementary Blaschke matrix finally replaces the root λ^{-1} by its reciprocal value.

As in the univariate example, it is also possible to move from the nonfundamental to the fundamental representation. The analytical example in Section 6.4 illustrates this aspect. Prior to that, I describe the procedure of empirically based root flipping in the subsequent section and argue that this procedure fails to discover the true MA representation.

6.3 Empirically based root flipping

Suppose that the FIA model is characterized by a nonfundamental vector MA representation of the form

$$y_t = M(L) w_t, \tag{6.5}$$

where w_t denotes the k-dimensional vector of economic shocks with $E(w_t w_t') = I_k$. By definition, $det[M(L)]$ has at least one root inside the unit circle. A corresponding reduced form of the representation in equation (6.5) is

$$y_t = M(L) M(0)^{-1} v_t, \tag{6.6}$$

where $v_t = M(0) w_t$.

As described above, the representation in equation (6.5) has its fundamental counterpart in form of the Wold decomposition

$$\begin{aligned} y_t &= M(L) B_1(L)^{-1} B_1(L) w_t \\ &= \widehat{M}(L) \varepsilon_t, \end{aligned} \tag{6.7}$$

where $\widehat{M}(L) = M(L) B_1(L)^{-1}$ and $\varepsilon_t = B_1(L) w_t$. $B_1(L)$ now denotes an appropriate Blaschke matrix that flips all roots smaller than unity in absolute value from the inside to the outside of the unit circle, i.e., $det\left[\widehat{M}(L)\right]$ has all roots larger than one in modulus. Notice that $E(\varepsilon_t \varepsilon_t') = I_k$.[1]

The econometrician using standard structural estimation procedures seeks to identify the structural innovations ε_t and their dynamics in reckoning to find the true shocks.[2] In particular, the econometrician does not estimate equation (6.7) directly but its reduced form

$$
\begin{aligned}
y_t &= \widehat{M}(L) \widehat{M}(0)^{-1} \widehat{M}(0) \varepsilon_t \\
&= T(L) a_t,
\end{aligned}
\tag{6.8}
$$

where $T(L) = \widehat{M}(L) \widehat{M}(0)^{-1}$, and $a_t \left(= \widehat{M}(0) \varepsilon_t\right)$ are the reduced form innovations with $E(a_t a_t') = \widehat{M}(0) \widehat{M}(0)'$. Thereby, I assume that the econometrician estimates equation (6.8) correctly. By indirect estimation, I mean that the econometrician applies the estimation procedures, such as VAR or subspace algorithm analysis, and then computes the vector MA representation, which is given in equation (6.8). In that sense, correct estimation indicates that all possible errors due to issues such as small sample uncertainty or lag truncation are neglected here. After the reduced form has been determined, identifying restrictions on $\widehat{M}(L)$ by means of zero short-run and/or zero long-run restrictions are imposed. In the standard procedure, these restrictions are based on (what is reckoned to be) the true model, i.e., on $M(L)$ rather than on $\widehat{M}(L)$. Hence, if the restrictions are not consistent with $\widehat{M}(L)$, the econometrician obtains another structural form, which I write as

$$
y_t = T(L) \widehat{\mathcal{D}} \widehat{\varepsilon}_t,
\tag{6.9}
$$

where $\widehat{\mathcal{D}}$ is the identified impact matrix (known from the LIE model in structural form) and $\widehat{\varepsilon}_t = \widehat{\mathcal{D}}^{-1} \widehat{M}(0) \varepsilon_t$.[3] $\widehat{\mathcal{D}}$ is computed such that $E(\widehat{\varepsilon}_t \widehat{\varepsilon}_t') = I_k$, i.e., the matrix $\widehat{\mathcal{D}}^{-1} \widehat{M}(0)$ has to be orthogonal. Note that if the identifying restrictions are in accordance with $\widehat{M}(L)$, one trivially has $\widehat{\mathcal{D}} = \widehat{M}(0)$, and therefore, $\widehat{\varepsilon}_t$ equals ε_t.

The econometrician is not successful in both steps of structural estimation. Obviously, the econometrician fails in the first step because $M(L) \neq \widehat{M}(L)$. She also fails in the latter step because there can be no matrix $\widehat{\mathcal{D}}$ for which $T(L) \widehat{\mathcal{D}}$ is equal to $M(L)$. Consequently, the econometrician only detects a linear combination of current and past true shocks. Thus, she uncovers dynamics that differ from the true dynamics.

To avoid this shortcoming, one could attempt the approach of Lippi and Reichlin (1993,

---

[1] Additionally, note that ε_t is also white noise due to the particular property of the Blaschke matrix, i.e., the Blaschke matrix transforms a white noise process into another white noise process (see, e.g., Lippi and Reichlin (1994)).

[2] Recall that the terminology distinguishes between the structural innovations ε_t and the true shocks w_t.

[3] This is the case in the analytical example of Subsection 6.4.2.

1994). The application of their procedure flips the MA roots from outside to inside of the unit circle after the reduced form in equation (6.8) has been estimated and before the restrictions are imposed. This step leads to the following representation

$$
\begin{aligned}
y_t &= T(L) B_2(L) B_2(L)^{-1} a_t \\
&= \widetilde{T}(L) \widetilde{a}_t,
\end{aligned} \tag{6.10}
$$

where $\widetilde{T}(L) = T(L) B_2(L)$ and $\widetilde{a}_t = B_2(L)^{-1} a_t.$[4] $B_2(L)$ denotes the Blaschke matrix chosen by the econometrician. In the final step, one would proceed by implementing the identification scheme as in the standard approach above. My claim is that this standard identification procedure is also invalid in the case of empirically based root flipping.

Consider the last representation in equation (6.10), where $\widetilde{T}(L) = M(L)B_1(L)^{-1}\widehat{M}(0)^{-1}B_2(L)$. Standard identification yields a rotation matrix, say $\widetilde{\mathcal{D}}$. To compute the structural form, $\widetilde{T}(L)$ is postmultiplied by $\widetilde{\mathcal{D}}$. Evidently, there can be no such matrix $\widetilde{\mathcal{D}}$ for which $\widetilde{T}(L)\widetilde{\mathcal{D}}$ equals $M(L)$. I conclude that empirically based root flipping collapses as well. A solution to this problem by adjusting the identification scheme appropriately is possible, but only if the econometrician has knowledge of the Blaschke matrix $B_1(L)$. This case is discussed in Chapter 7, but I first present some examples that illustrate the failure of empirically based root flipping.

6.4 The Lippi and Reichlin (1993) model

6.4.1 The model setup

The Lippi and Reichlin (1993) model is a modified version of the one presented by Blanchard and Quah (1989). Blanchard and Quah (1989) provide a theoretical framework for their empirical analysis in which they try to separate transitory from permanent effects of two structural shocks to US output. By estimating a bivariate SVAR model that comprises output growth and the unemployment rate, Blanchard and Quah (1989) introduce a long-run restriction scheme to identify both temporary and permanent sources of output fluctuations. They interpret a permanent innovation as a supply-side shock and a transitory disturbance as a demand shock. One of their findings is that the demand shock substantially contributes to output fluctuations over the short- and medium-term.

---

[4]Note that root flipping as in the representation in equation (6.10) can lead to complex coefficients in $\widetilde{T}(L)$. Furthermore, the MA operator evaluated at $L = 0$ is probably not equal to the identity matrix. It is possible to take an intermediate step that guarantees real coefficients and the identity matrix at $L = 0$ by postmultiplying $\widetilde{T}(L)$ with $\widetilde{T}(0)^{-1}$. The corresponding residuals are then given by $\breve{a}_t = \widetilde{T}(0)\widetilde{a}_t$. Note that this step does not alter the rationale of my arguments in the main text.

Lippi and Reichlin (1993) take up the Blanchard and Quah (1989) study and question their results by alluding to the possible existence of nonfundamental representations. They state that the theoretical model used as motivation by Blanchard and Quah (1989) exhibits poor dynamics in relation to what can be found in the data. Lippi and Reichlin (1993) suggest incorporating a learning-by-doing technology into that theoretical model, which generates a more complex pattern than does the productivity process in the Blanchard and Quah (1989) model. This more complex structure produces a model solution that implies nonfundamentalness and therefore violates the standard assumption of VAR analysis regarding the roots of the corresponding MA lag polynomial. Consequently, Lippi and Reichlin (1993) assess that Blanchard and Quah (1989) simply rule out other possible and economically relevant representations from their analysis by construction and without justification. These alternatives might have different implications compared to those obtained by Blanchard and Quah (1989).

In the following, I use the modified model version of Lippi and Reichlin (1993) to show that the nonfundamental alternatives, as detected by the empirically based root flipping procedure, are as wrong as the fundamental one derived from the standard VAR approach. The theoretical model describes a Keynesian-type economy and consists of the following equations

$$log\left(Y_t\right) \;=\; log\left(M_t\right) - log\left(P_t\right) + a \cdot log\left(\theta_t\right), \tag{6.11}$$

$$log\left(Y_t\right) \;=\; log\left(N_t\right) + log\left(\pi_t\right), \tag{6.12}$$

$$log\left(P_t\right) \;=\; log\left(W_t\right) - log\left(\pi_t\right), \tag{6.13}$$

$$log\left(W_t\right) \;=\; log\left(W\right) \mid \left\{E_{t-1}\left[log\left(N_t\right)\right] = log\left(\overline{N}\right)\right\}, \tag{6.14}$$

where Y_t, N_t and θ_t are the levels of output, employment and productivity, respectively. M_t, P_t and W_t denote the nominal variables of money supply, price level and wages, respectively. $\overline{N}$ reflects full employment, and π_t represents a learning-by-doing technology, which is explained below.

Equation (6.11) describes the aggregate demand as a function of real balances and the term $a \cdot ln\left(\theta_t\right)$, which can be interpreted as the investment demand, where a is the respective elasticity parameter. According to equation (6.12), output is produced by a constant returns to scale production function that only includes labor inputs. Firms set prices as specified in equation (6.13). The fourth equation (6.14) characterizes the wage setting rule in the spirit of Fischer (1977). Nominal wages in period t are set conditional on the expectation in the

previous period that employment in t equals full employment.

The model is closed by specifying the laws of motion for $log(M_t)$, $log(\theta_t)$ and $log(\pi_t)$. The first two variables are supposed to follow driftless random walks:

$$log(M_t) = log(M_{t-1}) + w_t^d, \tag{6.15}$$

$$log(\theta_t) = log(\theta_{t-1}) + w_t^s, \tag{6.16}$$

where w_t^d and w_t^s are the demand and supply shocks, respectively, assumed to be orthogonal and to have unit variance.

In the original model of Blanchard and Quah (1989), equations (6.12) and (6.13) include θ_t instead of π_t. Lippi and Reichlin (1993) integrate the additional variable π_t into the model, which is described by

$$log(\pi_t) = log(\pi_{t-1}) + d(L) w_t^s, \tag{6.17}$$

where $d(L)$ is a lag polynomial used to induce a certain diffusion pattern of the productivity process. The model solution can be expressed by the vector MA representation

$$\begin{pmatrix} \triangle log(Y_t) \\ U_t \end{pmatrix} = \begin{pmatrix} a(1-L) + d(L) & 1-L \\ -a & -1 \end{pmatrix} \begin{pmatrix} w_t^s \\ w_t^d \end{pmatrix}, \tag{6.18}$$

where U_t denotes the unemployment rate defined as $U_t \equiv log(\overline{N}) - log(N_t)$. Note that the solution of the original Blanchard and Quah (1989) model can be reproduced by setting $d(L) = 1$.

It is straightforward to see that the roots of the MA lag polynomial are the roots of $d(L)$. Lippi and Reichlin (1993) emphasize that a specification of $d(L)$, which generates more complex dynamics than the trivial ones of Blanchard and Quah (1989), gives rise to a nonfundamental representation. As an example, they set $d(L) = d_0(1 + 2L + 4L^2 + 4L^3 + L^4 + 0.5L^5)$, where d_0 is a scaling factor. Such a polynomial produces an S-shaped diffusion process and has roots inside the unit circle.

6.4.2 Testing empirically based root flipping

I start my discussion of testing the empirically based root flipping procedure with the representation in equation (6.18) but use a simpler structure of $d(L)$. I assume that $a = 1$ and $d(L) = 1 + \frac{1}{\vartheta}L$, where $0 < \vartheta < 1$. As the MA lag polynomial has one root at $-\vartheta$, the vector MA representation is nonfundamental. Note that the simple process that I assume for $d(L)$

is not decisive for my subsequent arguments. It is chosen merely to ease the computations and without loss of the relevant implications.[5] In fact, a more complex structure of $d(L)$ would highlight the implications even more clearly.

The resulting vector MA representation is given by

$$
\begin{pmatrix} \triangle log\,(Y_t) \\ U_t \end{pmatrix} = \begin{pmatrix} \frac{2\vartheta + (1-\vartheta)L}{\vartheta} & 1-L \\ -1 & -1 \end{pmatrix} w_t \,, \tag{6.19}
$$

where $w_t = \begin{pmatrix} w_t^s & w_t^d \end{pmatrix}'$. I use a Blaschke matrix to flip the root from $-\vartheta$ to $-\vartheta^{-1}$, i.e., I can write equation (6.19) as

$$
\begin{pmatrix} \triangle log\,(Y_t) \\ U_t \end{pmatrix} = M(L)\,B_1(L)^{-1}\,B_1(L)\,w_t
$$

$$
= \underbrace{\frac{\sqrt{2}}{2} \begin{pmatrix} \frac{(1-2\vartheta)L+3\vartheta}{\vartheta} & -\frac{(1+\vartheta L)}{\vartheta} \\ -2 & 0 \end{pmatrix}}_{=\,\widehat{M}(L)} \varepsilon_t \,, \tag{6.20}
$$

where $\varepsilon_t = B_1(L)\,w_t$ with

$$
B_1(L) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{L+\vartheta}{1+\vartheta L} \end{pmatrix} G_1^{-1} \,, \quad G_1 = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}
$$

and $E\,(\varepsilon_t \varepsilon_t') = I_2$.[6]

The econometrician estimates the reduced form

$$
\begin{pmatrix} \triangle log\,(Y_t) \\ U_t \end{pmatrix} = \widehat{M}(L)\,\widehat{M}(0)^{-1}\,\widehat{M}(0)\,\varepsilon_t
$$

$$
= \underbrace{\begin{pmatrix} 1+\vartheta L & \frac{(3\vartheta^2 + 2\vartheta - 1)L}{2\vartheta} \\ 0 & 1 \end{pmatrix}}_{=\,T(L)} a_t \,, \tag{6.21}
$$

---

where $a_t = \widehat{M}(0)\,\varepsilon_t$ are the reduced form innovations with $E(a_t a_t') = \widehat{M}(0)\,\widehat{M}(0)'$ and

$$\widehat{M}(0) = \frac{\sqrt{2}}{2}\begin{pmatrix} 3 & -\frac{1}{\vartheta} \\ -2 & 0 \end{pmatrix}.$$

The root of $det\,[T(L)]$ is then $-\vartheta^{-1}$.

Equation (6.21) is the starting point for standard structural estimation methods. An econometrician estimating a VAR model of $\begin{pmatrix} \triangle log\,(Y_t) & U_t \end{pmatrix}'$ computes the vector MA representation in equation (6.21) and then imposes restrictions to identify the innovations ε_t while supposing to find the true shocks. Recall that her choice of the restriction is contingent upon the true model in equation (6.19) and not upon the representation in equation (6.20).

I imitate the identification procedure of Blanchard and Quah (1989), i.e., a zero long-run restriction of the second shock (which is supposed to be a demand shock) on output growth, which is in line with the theoretical model. As a result, I obtain

$$\widehat{\mathcal{D}} = \frac{1}{2}\begin{pmatrix} 3 + 1/\vartheta & 3 - 1/\vartheta \\ -2 & -2 \end{pmatrix}$$

and the vector MA representation

$$\begin{pmatrix} \triangle log\,(Y_t) \\ U_t \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{1+3\vartheta+L(1-\vartheta)}{2\vartheta} & \frac{(3\vartheta-1)(1-L)}{2\vartheta} \\ -1 & -1 \end{pmatrix}}_{= T(L)\,\widehat{\mathcal{D}}}\widehat{\varepsilon}_t, \qquad (6.22)$$

which differs from equation (6.19) as well as from equation (6.20).[7] Likewise, the identified shocks $\widehat{\varepsilon}_t$ differ from w_t and ε_t, respectively.

Empirically based root flipping is also based on equation (6.21) but flips the root of $det\,[T(L)]$ before imposing the restriction for structural identification. Note that the estimation of $T(L)$ (and its root) crucially depends on the underlying estimation method in practice, and the econometrician does not know the root of $det\,[T(L)]$ a priori. I ignore such problems, as mentioned in the introduction of this chapter. Notice that $T(L)$ is arranged such that the factor $(1 + \vartheta L)$ only appears in the first column.[8] In general, the econometrician is free to choose where to pin down this factor by selecting an appropriate rotation matrix G_2 so that she can control for the column that should include $(1 + \vartheta L)$. Without loss of generality, I

---

[7]$\widehat{\mathcal{D}}$ is computed as $T(1)^{-1}T_{chol}$, where T_{chol} results from a Cholesky decomposition of $T(1)\,\Omega_a T(1)'$ with Ω_a defined as $E(a_t a_t')$.

[8]Due to the postmultiplication of $\widehat{M}(L)$ by $\widehat{M}(0)^{-1}$.

assume that the econometrician selects the first column.[9] Then, the identity matrix is simply a convenient choice for the rotation matrix G_2. Thus, applying the Blaschke matrix to equation (6.21) provides

$$
\begin{pmatrix} \triangle log\left(Y_t\right) \\ U_t \end{pmatrix} = T\left(L\right) B_2\left(L\right) B_2\left(L\right)^{-1} a_t
$$

$$
= \underbrace{\begin{pmatrix} L+\vartheta & \frac{\left(3\vartheta^2+2\vartheta-1\right)L}{2\vartheta} \\ 0 & 1 \end{pmatrix}}_{= \widetilde{T}\left(L\right)} \widetilde{a}_t , \tag{6.23}
$$

where $\widetilde{a}_t = B_2\left(L\right)^{-1} a_t$ with

$$
B_2\left(L\right) = G_2 \cdot \begin{pmatrix} \frac{L+\vartheta}{1+\vartheta L} & 0 \\ 0 & 1 \end{pmatrix}
$$

and $G_2 = I_2$. The variance-covariance matrix of $\widetilde{a}_t$ is $E\left(\widetilde{a}_t\widetilde{a}_t'\right) = G_2^{-1}\widehat{M}\left(0\right)\widehat{M}\left(0\right)'G_2$. This procedure flips the root back to $-\vartheta$.

Finally, the implementation of the standard long-run restriction yields

$$
\begin{pmatrix} \triangle log\left(Y_t\right) \\ U_t \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{3\vartheta^2+\vartheta+\left(2+\vartheta-3\vartheta^2\right)L}{2\vartheta} & \frac{\left(3\vartheta-1\right)\left(1-L\right)}{2} \\ -1 & -1 \end{pmatrix}}_{= \widetilde{T}\left(L\right)\widetilde{T}\left(0\right)^{-1}\widetilde{\mathcal{D}}} \widetilde{\varepsilon}_t , \tag{6.24}
$$

where $\widetilde{\varepsilon}_t = \widetilde{\mathcal{D}}^{-1}\widetilde{T}\left(0\right)\widetilde{a}_t$, and

$$
\widetilde{\mathcal{D}} = \frac{1}{2}\begin{pmatrix} 3\vartheta+1 & 3\vartheta-1 \\ -2 & -2 \end{pmatrix}
$$

is the estimated impact matrix.[10]

To sum up, the analytical example demonstrates that both the standard approach (see equation (6.22)) and the empirically based root flipping procedure (see equation (6.24)) yield a vector MA representation that deviates from the original. The difference from the true responses only refers to the impulse responses of output growth with respect to the impact effect of both shocks. This aspect is due to the simplicity of my example. The resulting rep-

---

[9]If the econometrician selects the second column, setting G_2 equal to $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ would achieve that.

[10]$\widetilde{\mathcal{D}}$ is calculated as $\widetilde{T}\left(1\right)^{-1}\widetilde{T}_{chol}$, where $\widetilde{T}_{chol}$ results from a Cholesky decomposition of $\widetilde{T}\left(1\right)\Omega_{\widetilde{a}}\widetilde{T}\left(1\right)'$ with $\Omega_{\widetilde{a}}$ defined as $E\left(\widetilde{a}_t\widetilde{a}_t'\right)$.

resentations are simple MA(1) processes. If I chose a more complex diffusion process for $d(L)$, e.g., the process suggested by Lippi and Reichlin (1993), the divergence from the true responses would arise at longer horizons as well. Furthermore, the exclusive deviation from the true responses in the case of one variable is not the general finding, as will be shown in the following example.

6.5 A Lucas asset tree model with news shocks

6.5.1 The model setup

In this section, I focus on a setting with news shocks. The relevance of anticipated shocks in explaining macroeconomic fluctuations has become an important topic in empirical business cycle research since the work of Beaudry and Portier (2005, 2006). In a first step, they show that structural innovations to stock prices identified by a short-run identification scheme (i.e., the stock price innovation has no impact effect on TFP) in a bivariate structural VECM comprising TFP and stock prices are (almost) perfectly collinear to structural innovations to TFP, which have been identified by imposing a long-run restriction (i.e., the TFP shock is the only driving force of TFP over the long-run) in the same VECM. Their conclusion is that stock price news, which are orthogonal to surprise changes in TFP, reflect anticipated future TFP shocks. In a second exercise, the authors reveal that such shocks can explain a major part of the forecast error variance of output and consumption, for example. Their results receive confirmation in several papers (see, e.g., Beaudry and Lucke (2010)), but also meet with criticism stressing that the underlying theoretical model features a nonfundamental representation. Forni et al. (2014) address this problem by using a simple Lucas asset tree model including news shocks to demonstrate the critique. I take on their example as second benchmark for the test of empirically based root flipping. I differ from Forni et al. (2014) only in assuming a stationary technology process. My conclusions would not change if I used a unit root in the technology process.

In the model economy, expected lifetime utility of the representative household only depends on consumption as follows:

$$U = E_0 \sum_{t=0}^{\infty} \beta^t C_t, \tag{6.25}$$

where β denotes the discount factor. In every period t, the household consumes C_t and buys tree shares S_{t+1} at the unit price P_t. The household's expenditures are constrained by its wealth that is decomposed into the value of the tree shares held at period t, $P_t S_t$, and the value of the period flow of dividends $D_t S_t$. In equilibrium, there is one unique tree such that the stock market value equals a unit price P_t. Moreover, dividends equal technology θ_t,

which follows a stationary AR(1) process

$$\theta_t = \rho_\theta \theta_{t-1} + w_{0,t}^\theta + w_{q,t-q}^\theta, \tag{6.26}$$

where $0 < \rho_\theta < 1$. $w_{0,t}^\theta$ represents a surprise shock to technology, whereas $w_{q,t-q}^\theta$ displays the anticipated shock to technology, which materializes after q periods. Both shocks are white noise with unit standard deviation.

The representative household maximizes lifetime utility subject to the budget constraint. I obtain

$$P_t = \beta E_t \left(P_{t+1} + D_{t+1} \right), \tag{6.27}$$

which is the first order condition of the optimization problem. Solving forward and imposing the transversality condition yields

$$P_t = \beta E_t \left(\beta \left(P_{t+2} + D_{t+2} \right) D_{t+1} \right) = E_t \sum_{\tau=1}^{\infty} \beta^\tau D_{t+\tau}. \tag{6.28}$$

Finally, taking the technology process into account and using the lag operator L, I determine the stock market value of the tree in terms of current and past shocks:

$$P_t = \frac{\beta \rho_\theta}{(1 - \beta \rho_\theta)(1 - \rho_\theta L)} w_{0,t}^\theta + \left(\frac{\beta^q}{1 - \beta \rho_\theta} \sum_{i=0}^{q-1} \beta^{-i} L^i + \frac{\beta \rho_\theta}{(1 - \beta \rho_\theta)(1 - \rho_\theta L)} L^q \right) w_{q,t}^\theta. \tag{6.29}$$

Defining a bivariate vector of observables with technology and stock prices, I arrive at the following vector MA representation, e.g., for $q = 2$,

$$\begin{pmatrix} \theta_t \\ P_t \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{1}{1-\rho_\theta L} & \frac{L^2}{1-\rho_\theta L} \\ \frac{\beta \rho_\theta}{(1-\beta \rho_\theta)(1-\rho_\theta L)} & \frac{\beta^2}{1-\beta \rho_\theta} + \frac{\beta}{1-\beta \rho_\theta} L + \frac{\beta \rho_\theta}{(1-\beta \rho_\theta)(1-\rho_\theta L)} L^2 \end{pmatrix}}_{= M(L)} \underbrace{\begin{pmatrix} w_{0,t}^\theta \\ w_{2,t}^\theta \end{pmatrix}}_{= w_t}. \tag{6.30}$$

6.5.2 Testing empirically based root flipping

I begin testing empirically based root flipping by setting $q = 2$ and considering the corresponding vector MA representation of $\begin{pmatrix} \theta_t & P_t \end{pmatrix}'$ in equation (6.30). I later extend this example by increasing the anticipation horizon q. The case $q = 2$ allows me to illustrate my main arguments by another analytical example for which I can provide simple expressions for the Blaschke matrices. I refer the reader to the appendix for some of the computational details to save space.

The MA lag polynomial in equation (6.30) has one root at $-\beta$. As β denotes the discounting factor, the root is smaller than one in modulus. I therefore have a nonfundamental

Figure 6.1: IMPULSE RESPONSES IN LUCAS ASSET TREE MODEL WITH NEWS SHOCK ($q = 2$)
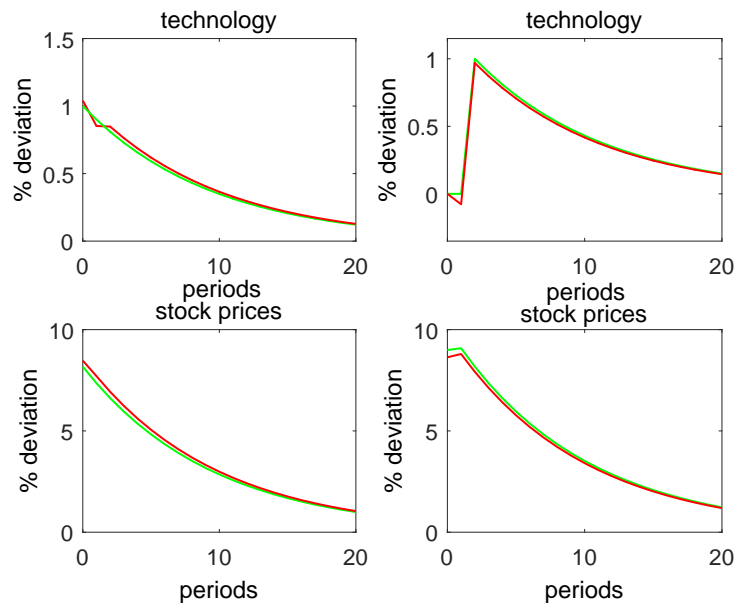
Note: The figure depicts the FIA model impulse responses (in green) and the impulse responses resulting from the empirically based root flipping procedure (in red). Impulse responses to $w_{0,t}^{\theta}$ are shown in the left column and impulse responses to $w_{2,t}^{\theta}$ in the right column.

representation. In Figure 6.1, I depict the impulse responses resulting from the empirically based root flipping procedure (red lines) together with the true impulse responses from the original model (green lines). The chosen identification scheme is a zero restriction on the contemporaneous effect of $w_{2,t}^{\theta}$ on θ_t. The calibrated parameters are $\beta = 0.99$ and $\rho_{\theta} = 0.9$.

Although there is no substantial visible divergence, the impulse responses from empirically based root flipping do not match the true responses in any of the four cases – even at longer horizons. These deviations are far from what an econometrician should be concerned with if one takes other sources of uncertainty and misspecification into account, but the relevance of nonfundamentalness becomes more substantial with an increasing anticipation horizon.

Raising the anticipation horizon to $q = 8$ emphasizes this claim. In this case, $det\,[M\,(L)]$ has seven roots that all have modulus β. When applying the empirically based root flipping procedure, I again assume that the econometrician estimates all roots exactly and knows to flip them all. Figure 6.2 displays the corresponding impulse responses together with the impulse responses produced by the standard approach (blue lines), i.e., the impulse responses detected by the econometrician without empirically based root flipping. As indicated in the figure, the error of empirically based root flipping grows remarkably with the length of the anticipation horizon. Furthermore, an econometrician applying standard techniques generates impulse responses with much less deviation from the true responses than when using empirically based root flipping.
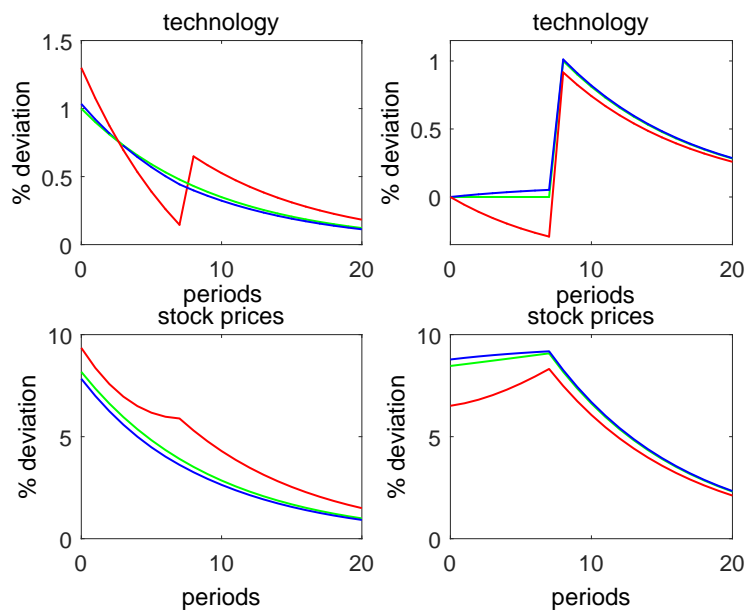
Figure 6.2: IMPULSE RESPONSES IN LUCAS ASSET TREE MODEL WITH NEWS SHOCK ($q = 8$)

Note: The figure depicts the FIA model impulse responses (in green), the impulse responses resulting from the empirically based root flipping procedure (in red) and the impulse responses related to the standard approach (in blue), i.e., the LIE model. Impulse responses to $w_{0,t}^\theta$ are shown in the left column and impulse responses to $w_{8,t}^\theta$ in the right column.

6.6   Remarks

In this chapter, I have demonstrated that empirically based root flipping cannot solve the problem of nonfundamentalness. I presented two examples in which empirically based root flipping produces errors in terms of deviations from the true dynamics. These errors surely depend on the structure of the underlying theoretical model. Testing the approach in a setup that includes news shocks reveals that the length of the anticipation horizon amplifies this bias. Moreover, the findings indicate that empirically based root flipping can produce larger errors than the standard SVAR approach.

Thus, it does not seem feasible to find the correct nonfundamental representation using small-scale VAR without embedding additional information that cannot be filtered out of the estimated VAR coefficients. It is worth exploring alternative approaches that make use of supplemental information that diminish the misalignment between the agent and econometrician's information sets effectively. Successful work in this direction has been conducted, e.g., by Mertens and Ravn (2010), whose method I pursue in the next chapter.

Chapter 7

Testing theory-based root flipping

In the previous chapter, I used simple models to explicate the concept of root flipping through analytical examples. In this chapter, I extend the setting and use the more complex workhorse model for my computations and simulations. Because I need a theoretical foundation for the theory-based root flipping procedure, I revisit the description of the model specified in Chapter 5 and then explain the procedure in this framework.

The theory-based technique applies to a situation in which the econometrician has additional knowledge of the Blaschke matrix that connects the true (nonfundamental) vector MA representation with the estimated (fundamental) vector MA representation. Equipped with this information, the econometrician can compensate for her informational disadvantage relative to the agents. I examine the goodness of fit of the estimation procedure with a Monte Carlo simulation study at the end of the chapter.

7.1   Motivation

The concept of theory-based root flipping originates from Mertens and Ravn (2010). They introduce an estimator adapted from an SVAR model that incorporates auxiliary information drawn from the underlying theoretical model. Based on theory, it is possible to flip the relevant roots and uncover the true impulse responses in the presence of news shocks (at least asymptotically). In a Monte Carlo simulation, Mertens and Ravn (2010) show that this augmented SVAR estimator performs better in comparison to the standard SVAR procedure. When applied to US data, they also report a puzzle, that is, a result that is contrary to the implications of standard theory: consumption rises in response to both surprise and anticipated government spending shocks.

In this chapter, I test the augmented SVAR approach of Mertens and Ravn (2010) in a different environment, a model economy subject to both surprise and anticipated technology shocks. Such a setting is interesting because the relevance of anticipated technology shocks

has become a controversial topic in empirical business cycle research (see also the next chapter). In many studies, technological news shocks are identified through the inclusion of stock prices in the set of observables. Stock prices are considered as an appropriate candidate for identifying news shocks due to their forward-looking character. However, as I have already discussed in Section 5.4 and shown in the Lucas asset tree model in the preceding chapter, the inclusion of stock prices in the set of observables alone may not solve the problem of nonfundamentalness.

I extend the basic concept of the estimator of Mertens and Ravn (2010) and implement it in combination with the subspace algorithm analysis. As shown in Kascha and Mertens (2009) and Chapter 3, subspace algorithm analysis can be a serious alternative to standard SVAR models. Both Chapter 3 and Kascha and Mertens (2009) investigate only the performance of subspace algorithm analysis with respect to models that do not imply nonfundamentalness. In general, subspace algorithm analysis encounters the same problem as SVARs when nonfundamentalness is present, as the derivation of the algorithm rests upon the *minimum phase assumption*. Nonetheless, the nonfundamentalness problem can be eluded as in SVAR models. I compare the performance of the augmented SVAR method and the augmented subspace algorithm analysis in the Monte Carlo simulation that concludes the chapter.[1]

7.2 Specification of the workhorse model

In this section, I use the fully specified workhorse model from Chapter 5, which serves as the theoretical foundation for the computations and simulations as well as for the empirical analysis in the next chapter in which I will examine the role of news shocks using US data. Thus, I closely align my procedure with other studies, such as Beaudry and Portier (2005, 2006), which select stock prices as a relevant variable that captures technological news.

To incorporate stock prices into the workhorse model, I augment it by introducing capital adjustment costs into the model setup. In particular, I redesign the capital accumulation equation as

$$K_t = (1 - \delta) K_{t-1} + \left[1 - \mathcal{S}_K \left(\frac{I_t}{\delta K_{t-1}} \right) \right] I_t , \tag{7.1}$$

where $\mathcal{S}_K (\bullet)$ represents a capital adjustment cost function that satisfies $\mathcal{S}_K (1) = \mathcal{S}_K' (1) = 0$ and $\mathcal{S}_K'' (1) > 0$ in the steady state. Its introduction permits a Tobin's q interpretation within the model, which implies that the shadow price of installed capital is allowed to deviate from the price of an additional unit of capital. I model the stock price as firm value, which I compute as Tobin's q multiplied by the capital stock.[2]

---

[1] See also Sims (2012), Beaudry et al. (2013) and Beaudry and Portier (2014b) for Monte Carlo simulations in the context of news shocks, but these studies only address the performance of the standard SVAR technique.
[2] Note that Tobin's q is not related to the anticipation horizon in the case of news shocks, which I denote q in this thesis.

All other elements of the model as well as the solution technique are as in Chapter 5. The general model solution in terms of the FIA model and the vector MA representation are the same as presented therein. I repeat the vector MA representation of the model solution (equation (5.28)) as a reminder and refer the reader to Chapter 5 for further details and the corresponding FIA model (see equations (5.29) and (5.30)):

$$
\begin{pmatrix} \widetilde{\theta}_t \\ \widetilde{z}_t \end{pmatrix} =
$$

$$
\underbrace{\begin{pmatrix} \frac{1}{(1-\rho_\theta L)} & \frac{L^q}{(1-\rho_\theta L)} \\ \left(\frac{\phi_{z\theta}}{(1-\rho_\theta L)} + \frac{\phi_{zk}\phi_{k\theta}}{(1-\phi_{kk}L)(1-\rho_\theta L)} L \right) & \begin{array}{c} \left(\frac{\phi_{z\theta}}{(1-\rho_\theta L)} + \frac{\phi_{zk}\phi_{k\theta}}{(1-\phi_{kk}L)(1-\rho_\theta L)} L \right) L^q \\ + \left(\phi_{z\theta,1} + \frac{\phi_{zk}\phi_{k\theta,1}}{(1-\phi_{kk}L)} L \right) \Theta(L) \end{array} \end{pmatrix}}_{= M(L)} \Sigma_w \begin{pmatrix} w_{0,t}^\theta \\ w_{q,t}^\theta \end{pmatrix},
$$

$$
(7.2)
$$

where

$$
\Sigma_w = \begin{pmatrix} \sigma_\theta & 0 \\ 0 & \sigma_\theta \lambda_\theta \end{pmatrix}
$$

and

$$
\Theta(L) = \omega^{q-1} + \omega^{q-2} L + \ldots \omega L^{q-2} + L^{q-1}.
$$

As already stated in the discussion in Section 5.4, stock prices are also part of the model solution such that the nonfundamental characteristic in equation (7.2) remains when replacing the proxy for any endogenous variable z_t by the stock price as the second observable.

7.3 The concept of theory-based root flipping

Having established the theoretical framework, I can adopt the approach of Mertens and Ravn (2010). This approach acts on the transformation of the nonfundamental vector MA representation in equation (7.2) into its fundamental counterpart. As shown in the preceding chapter, one has to formulate particular Blaschke matrices to flip the roots of the MA lag polynomial from the inside to the outside of the unit circle. Because the relevant roots of $det\,[M(L)]$ lie on a circle with radius ω separated by the same angle, this task is straightforward because it only requires one specific rotation matrix G. After $M(L)$ has been postmultiplied by G, all the roots associated with $M(L)\,G$ can be flipped successively. Consecutive multiplication by elementary Blaschke matrices provides the fundamental vector MA rep-

resentation.

Setting $y_t = \left(\begin{array}{cc} \widetilde{\theta}_t & \widetilde{z}_t \end{array} \right)'$ and $w_t = \left(\begin{array}{cc} w_{0,t}^\theta & w_{q,t}^\theta \end{array} \right)'$, I formalize these computations as

$$y_t = \widehat{M}(L)\,\varepsilon_t, \tag{7.3}$$

where $\widehat{M}(L) = M(L)\,B(L)^{-1}$ and $\varepsilon_t = B(L)\,w_t$. $B(L)^{-1}$ is formulated as

$$B(L)^{-1} = G \cdot R_1(L)^{-1}\,R_2(L)^{-1}\ldots R_{q-1}(L)^{-1}$$

with

$$R_i(L)^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1-\overline{\omega}_i L}{L-\omega_i} \end{pmatrix},$$

$$G = \frac{1}{\sqrt{1+(\lambda_\theta\omega^q)^2}} \begin{pmatrix} 1 & -\lambda_\theta\omega^q \\ \lambda_\theta\omega^q & 1 \end{pmatrix},$$

where $\overline{\omega}_i$ denotes the complex conjugate of ω_i. The reduced form is given by

$$y_t = T(L)\,a_t, \tag{7.4}$$

where $T(L) = \widehat{M}(L)\,\widehat{M}(0)^{-1}$. Again, $a_t = \widehat{M}(0)\,\varepsilon_t$ are the reduced form innovations with $E(a_t a_t') = \widehat{M}(0)\,\widehat{M}(0)'$.

Following the procedure by Mertens and Ravn (2010) exposes the true impulse responses from the observables.[3] It presupposes the knowledge of the Blaschke matrix $B(L)$, i.e., the knowledge of three additional parameters, ω, λ_θ and q. Under this assumption, it is possible to disclose the model dynamics because the identifying restrictions can be chosen appropriately.

For example, a zero short-run restriction on the upper right element of the impact matrix $M(0)$ would be qualified in the bivariate model. However, an econometrician who uses standard techniques would mistakenly place this restriction on $\widehat{M}(0)$. Therefore, a valid way of proceeding is to impose a proper (but not a zero) restriction on $\widehat{M}(0)$, which implies a zero restriction on the $[1,2]$-element of $M(0)$. Because $\widehat{M}(0) = M(0)\,B(0)$ and $B(L)$ are known, a correct restriction is $\widehat{m}_{12}^0 = -\omega\lambda_\theta\widehat{m}_{11}^0$, where $\widehat{m}_{ij}^0$ denotes the $[i,j]$-element of $\widehat{M}(0)$.[4] Having found $\widetilde{M}(0)$, the true impulse responses captured by $M(L)$ can be computed by

---

[3]Notice that it does not identify the true shocks in practice because premultiplying $\varepsilon_t = B(L)\,w_t$ by $B(L)^{-1}$ and combining with $a_t = \widehat{M}(0)\,\varepsilon_t$ yields w_t as an infinite sum of the reduced form innovations, i.e., $w_t = B(L)^{-1}\,\widehat{M}(0)^{-1}\,a_t$.

[4]Note that this restriction does not yield a unique result. As in the standard case, one has to impose the usual sign restrictions on the diagonal elements of $M(0)$. However, only the diagonal elements of $M(0)$ are uniquely determined. In using elements of $B(0)$, it is possible to identify the $[2,1]$-element of $M(0)$ as well.

postmultiplying $T(L)$ by $\widehat{M}(0) B(L)$.

While Mertens and Ravn (2010) conduct this procedure within a structural VAR framework, I extend it to the subspace algorithm analysis. Recall that the subspace algorithm analysis relies on essentially the same assumption as the VAR estimation, i.e., it results in the same representation in equation (7.4).[5] Thus, the connection between the FIA model and the LIE model is the same as the link between the (true) vector MA representation in equation (7.2) and the corresponding reduced form representation in equation (7.4). In the context of vector MA representations, one speaks of the roots of the MA lag polynomial. The MA lag polynomial in equation (7.2) has roots inside the unit circle, whereas these roots are replaced by their reciprocals in the MA operator in equation (7.4). In the case of the state space systems, one can state this relation in terms of the eigenvalues of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$ and $(\mathcal{A} - K\mathcal{C})$. All eigenvalues of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$, which are larger than one in modulus, have a reciprocal counterpart in $(\mathcal{A} - K\mathcal{C})$. In Chapters 2 and 3, I have demonstrated that the *minimum phase condition* (i.e., all eigenvalues of $(\mathcal{A} - K\mathcal{C})$ are less than unity) is crucial for the derivation of the subspace algorithm and is a result of the assumptions that lead to convergence of the Kalman filter. This duality between VAR and subspace algorithm analysis suggests that the theory-based root flipping procedure can be implemented analogously in both cases.

7.4 Monte Carlo simulation

7.4.1 Monte Carlo design

In the simulation study, I use the FIA model representation, where the stock price is chosen as the second observable, as the DGP. I calibrate the model parameters as before: $\alpha = 2/3$, $\beta = 1.01^{-1}$, $\sigma = 1$, $\psi = 1$ and $\rho_\theta = 0.95$. The remaining two parameters, $\mathcal{S}_K''(1)$ and λ_θ, are chosen such that the variance ratio of investment and stock prices and the variance ratio of consumption and output approximately match their empirical counterparts in the US postwar data. Namely, I set $\mathcal{S}_K''(1) = 0.5$ and the relative weight of the news shock λ_θ equal to two. The anticipation horizon q is fixed at eight periods. Note that the choice of q differs in the literature. On the one hand, there are some conservative candidates, such as Sims (2012) who picks a value of two, for instance. On the other hand, Beaudry and Lucke (2010) use an anticipation horizon of eight periods for the simulation study presented in their web appendix. Their empirical results as well as the findings of Beaudry and Portier (2006) reveal point estimates that legitimate this choice rather than the lower values. Given the calibration, I can calculate the roots of the MA lag polynomial. The anticipation rate equals 0.9675. Because $q = 8$, there are seven roots of the MA lag polynomial with modulus

---

[5]Using the notation of the system matrices in the LIE model representation, I can derive $T(L) = \left(I_k + \mathcal{C}(I_n - \mathcal{A}L)^{-1} KL \right)$.

Figure 7.1: IMPULSE RESPONSES IN WORKHORSE MODEL WITH NEWS SHOCK $(q = 8)$
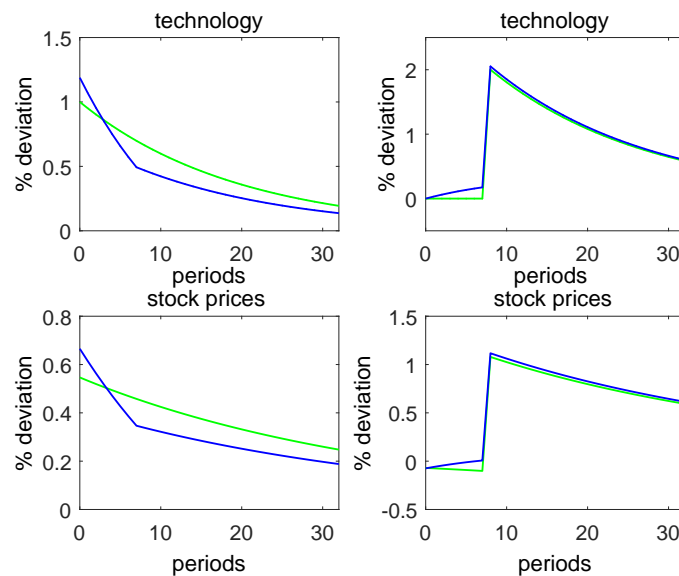
Note: The figure depicts the FIA model impulse responses (in green) and the LIE model impulse responses (in blue). Impulse responses to $w_{0,t}^{\theta}$ are shown in the left column and impulse responses to $w_{8,t}^{\theta}$ in the right column.

0.9675. The additional root not related to ω equals 1.05.

As in Chapter 5, I begin with a graphical presentation in Figure 7.1, which shows the true impulse responses (green lines) and the ones coming from the standard approach for structural identification (blue lines). The picture is similar to the one in Chapter 5, i.e., the systematic error mainly occurs with respect to the responses to the surprise shock.[6]

I can use Figure 7.1 to visually test whether the estimation methods with and without root flipping work correctly. The figure reflects the asymptotic results for the estimation procedures, i.e., if I chose a sufficiently large sample size, the methods should reproduce the blue and green lines, respectively. I start the analysis with a large sample simulation in which I generate 1000 sets of data series each with a sample size of 20000 observations. In the subsequent small sample exercises, I generate 1000 sets of data series each containing 500 observations. The first 300 observations are then discarded to eliminate effects of presample values, so my effective sample size is 200.

The theory-based root flipping procedure is incorporated into two frameworks, the SVAR model and subspace algorithm analysis. In the latter case, I conduct the CCA, which originates from Larimore (1983).[7] The lag length of the VAR model in the large sample exercise is set equal to 200. I perform the large sample experiment only to test the validity of the root flipping procedures quasi-asymptotically. By using such a long lag length, I want to exclude any other potential source of bias. In the small sample exercises, the lag length is chosen

---

[6]Remember that this finding is associated with the relative weight of the news shock. The lower the weight, the smaller the error on the side of the surprise shock and the larger the error with respect to the anticipated shock, and vice versa.

[7]I remind the reader of the description of the subspace algorithm provided in Section 3.3.

according to the AIC, but I also consider fixed lag lengths, as will be described below. In the case of the AIC, I set the maximum number of lags equal to eleven. For the subspace algorithm, I set the number of lead and lagged observations as described in Chapter 3. The system order is determined by the criterion proposed by Bauer and Wagner (2002) (see also Chapter 3).

In each exercise, I implement the standard structural estimation approaches, i.e., I estimate the conventional SVAR model and apply structural CCA, where I erroneously impose a zero short-run restriction on $\widehat{M}(0)$ in both cases. Alternatively, I carry out the theory-based root flipping procedure using both methods. Thus, four methods compete, and I investigate whether the root flipping yields any improvement in comparison to the standard approaches. To compare the performance of these methods, I measure the distance between the estimated and the true impulse responses by means of the MAWE indicator developed in Chapter 5.

7.4.2   Benchmark system

Contrary to the usual course of action, I initially turn to the large sample exercise to convince the reader that theory-based root flipping works. In the following graphs, I present the estimation results for the standard methods, VAR and CCA, as well as the approaches with root flipping. Figure 7.2 is composed of four boxes, each associated with one of the estimation procedures. Each box includes four graphs that illustrate the estimated impulse responses in black and the true impulse responses in green. The upper graphs show the responses of technology. The lower graphs display the responses of the second observable, stock prices, to the respective surprise and anticipated shock. The estimated impulse responses are computed as the median over all simulation runs. The dashed red lines mark the confidence intervals that I calculate as the 5th and 95th percentiles of the estimated impulse responses. It is evident that the estimation methods work as expected. The standard approaches reproduce the shape of the blue lines in Figure 7.1 while the root flipping procedures cover the green lines.

Next, I carry out the small sample experiment and investigate the performance of the methods in face of finite sample uncertainty. Figure 7.3 presents the corresponding estimation results. In the case of the VAR, the visual analysis does not clearly reveal whether there is any improvement due to root flipping. Conspicuously, the impulse responses of the VAR with root flipping exhibit a very choppy pattern. In contrast, the impulse responses of the CCA look much smoother even after the roots have been switched. Moreover, there is some progress visible regarding the responses of both variables to the surprise shock. On average, the estimated impulse responses under theory-based root flipping are closer to the true responses than in the standard case.

Figure 7.2: EXERCISE 7.1 - IMPULSE RESPONSES (LARGE SAMPLE SIMULATION)
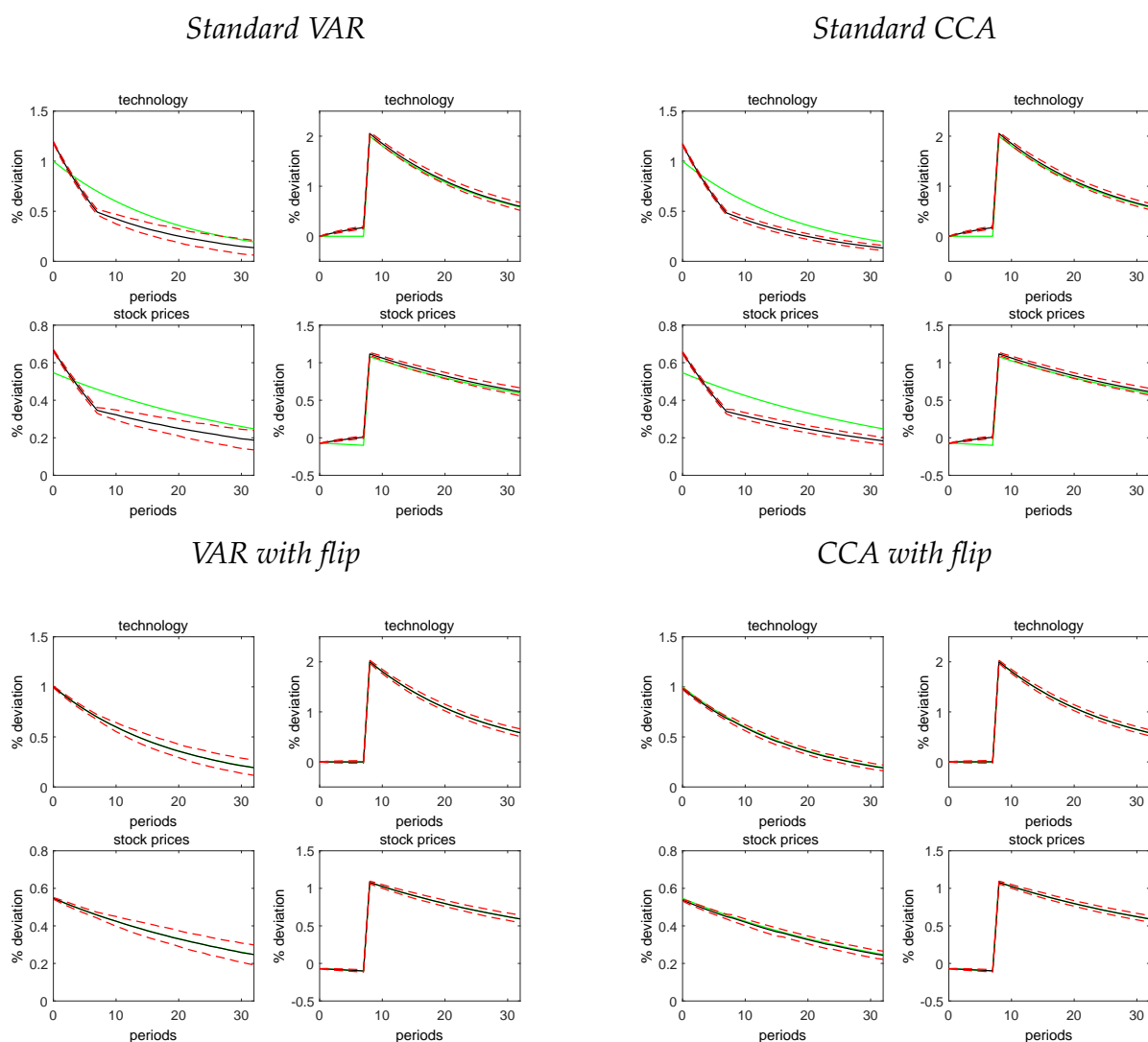
Standard VAR   *Standard CCA*

VAR with flip   *CCA with flip*

Note: The figure depicts FIA model impulse responses (in green) and estimated impulse responses (in black). Dashed red lines mark two-sided 90% confidence bounds. Impulse responses to $w_{0,t}^{\theta}$ are shown in the left column and impulse responses to $w_{8,t}^{\theta}$ in the right column of each panel.

Figure 7.3: EXERCISE 7.2 - IMPULSE RESPONSES (SMALL SAMPLE SIMULATION)
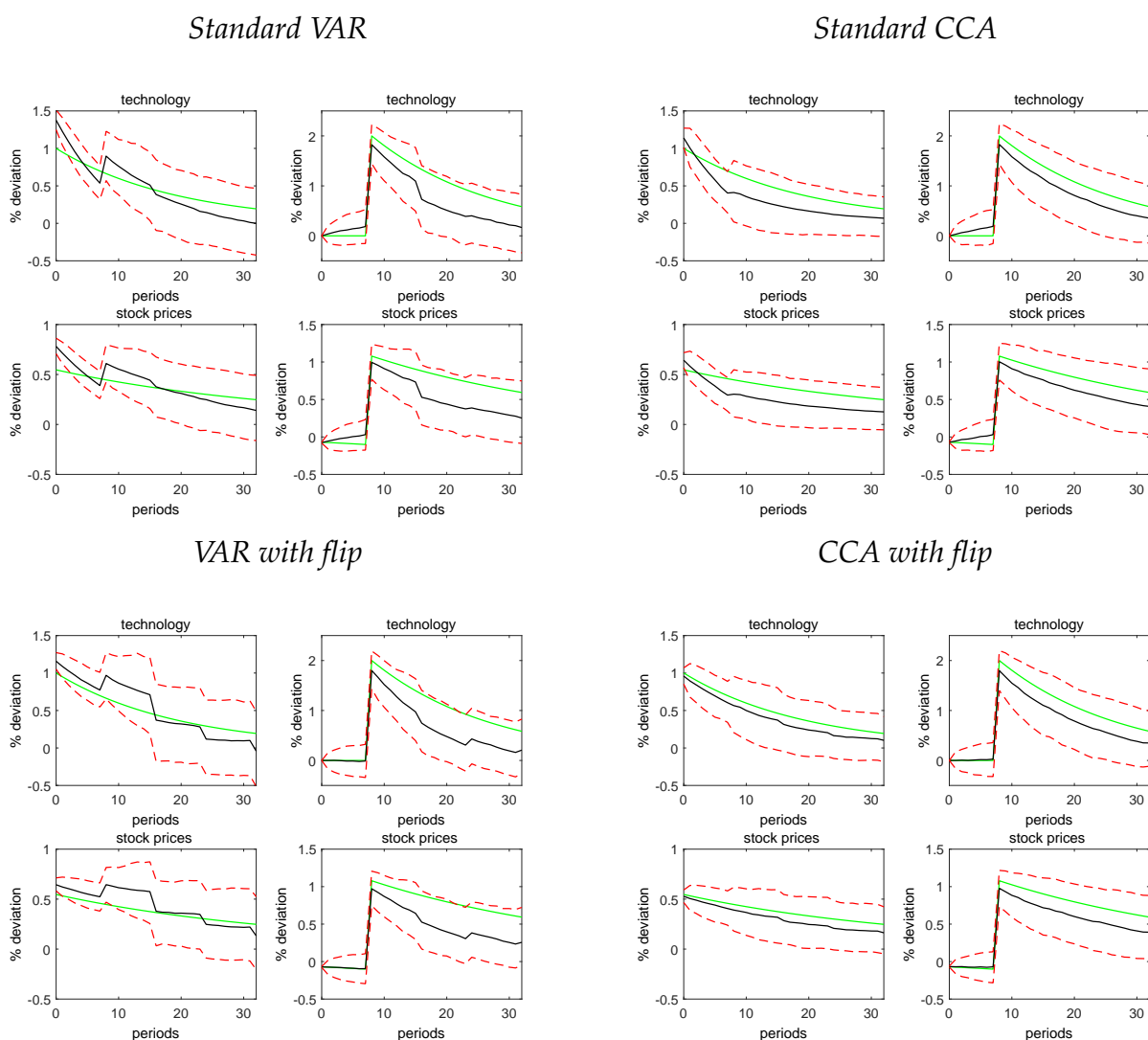
Note: The figure depicts FIA model impulse responses (in green) and estimated impulse responses (in black). Dashed red lines mark two-sided 90% confidence bounds. Impulse responses to $w_{0,t}^{\theta}$ are shown in the left column and impulse responses to $w_{8,t}^{\theta}$ in the right column of each panel.

Table 7.1: EXERCISE 7.2 - MAWE RESULTS (STOCK PRICES)

|	MAWE (VAR)		MAWE (CCA)	
Standard	1.4169	1.3349	1.3437	1.0528
	1.1926	1.1998	1.1987	0.9369
With flip	1.4899	1.3546	1.0949	1.0457
	1.2870	1.2522	0.9638	0.9449
Ratio (flip/standard)	1.0515	1.0147	0.8149	0.9932
	1.0792	1.0437	0.8040	1.0085

Note: In the upper two block rows, the table reports the computed MAWEs associated with the impulse responses displayed in Figure 7.3. MAWE values are calculated as in equation (5.8) multiplied by 100. The bottom block row shows the ratio between the values in the middle block row and the values in the top block row.

I quantify the differences in the impulse responses by computing the corresponding MAWEs. Table 7.1 reports the results. Each cell contains 2×2 matrices that include the MAWEs (or ratios) associated with the respective impulse responses. For instance, the upper left element of each matrix corresponds to the response of technology to a surprise shock, whereas the lower right element is linked to the response of stock prices to a news shock. The composition of the first two block lines (standard, with flip) is the same as in Figure 7.3. Additionally, the lower block line shows the ratio of each pair of the MAWEs so that the relative change due to root flipping can be directly seen. The numbers suggest that the VAR with root flipping yields inferior results, but there is some considerable improvement in the CCA with root flipping. The gain in precision on the side of the surprise shock is approximately 20% for the CCA, whereas there is virtually no change in the responses to the news shock.

The question is whether these findings can be generalized and if there is an explanation for the poor VAR performance. In this context, the choice of the lag length is a key issue. From that reason, I repeat the simulation exercise by estimating the same artificial data sets as above but varying the lag length for the VAR model. I begin with a VAR including one lag and conduct the 1000 replications. I reiterate this procedure and increase the number of included lags stepwise to a maximum of 25 lags. Hence, I can calculate the MAWE depending on the fixed lag length. Figure 7.4 shows the resulting MAWEs for the whole range of fixed lags. The solid line marks the standard VAR estimation, and the dotted line indicates root flipping. It is straightforward to recognize that a lag length that is smaller than (or equal to) the anticipation horizon leads to a relatively large bias.

It is therefore useful to consider the specifications with lag lengths above the anticipation horizon in more detail. This is done in Figure 7.5, where I add the corresponding results for the VAR based on AIC and the MAWEs of the CCA methods. First, the MAWE values of the VAR based on AIC (red lines) appear quite similar to those using a fixed lag length of nine or ten (black lines). This indicates that the AIC mainly chooses lag lengths close to these values. Furthermore, if the lag length exceeds a certain number, root flipping provides

Figure 7.4: EXERCISE 7.2 - MAWE RESULTS (VAR LAG LENGTH VARIATION)
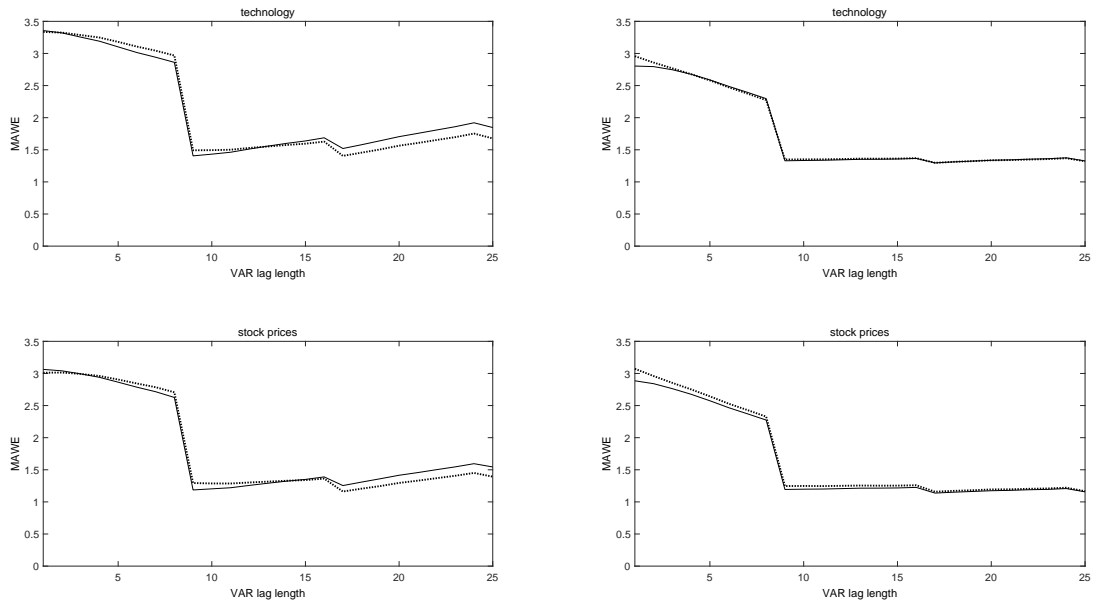
Note: The figure depicts the computed MAWEs associated with the estimated impulse responses related to the standard SVAR model (solid lines) and the SVAR model augmented by the root flipping procedure (dotted lines), depending on selected lag lengths. MAWE values are calculated as in equation (5.8) multiplied by 100. MAWEs corresponding to impulse responses to $w_{0,t}^{\theta}$ are shown in the left column and MAWEs w.r.t. impulse responses to $w_{8,t}^{\theta}$ in the right column.

Figure 7.5: EXERCISE 7.2 - MAWE RESULTS (SELECTED VAR LAG LENGTHS)
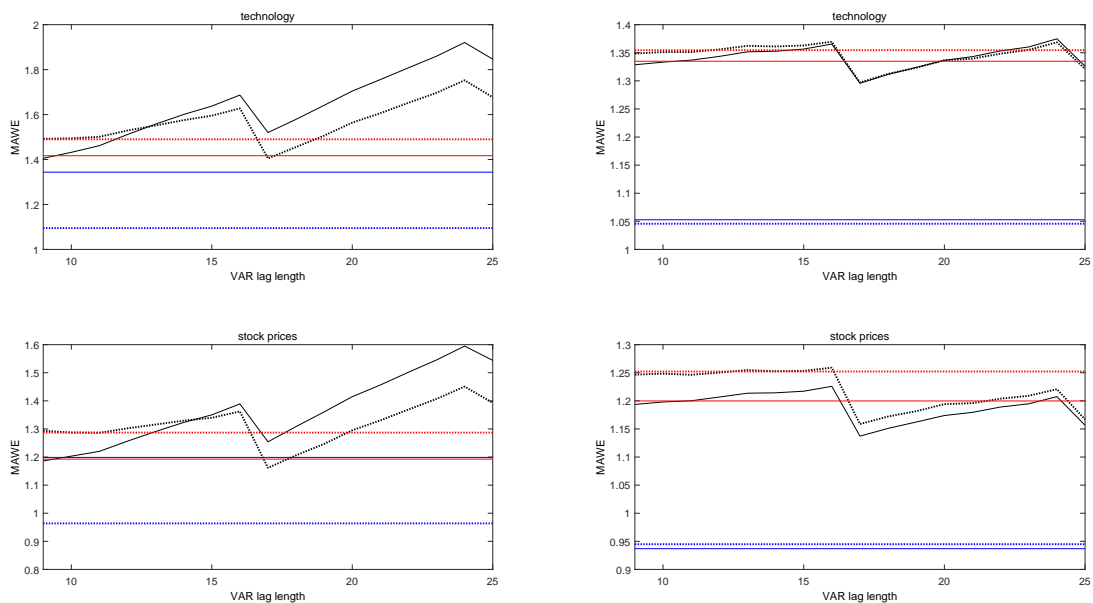
Note: The figure depicts the computed MAWEs associated with the estimated impulse responses resulting from different estimation approaches. These approaches are based on the SVAR model depending on selected lag lengths (in black), SVAR using the lag length suggested by the AIC (in red) and CCA (in blue). Solid lines refer to the standard application of these techniques and dotted lines to the corresponding root flipping procedure. MAWE values are calculated as in equation (5.8) multiplied by 100. MAWEs associated with the impulse responses to $w_{0,t}^{\theta}$ are shown in the left column and MAWEs w.r.t. impulse responses to $w_{8,t}^{\theta}$ in the right column.

Figure 7.6: EXERCISE 7.2 - IMPULSE RESPONSES (VAR WITH FLIP, 17 LAGS)
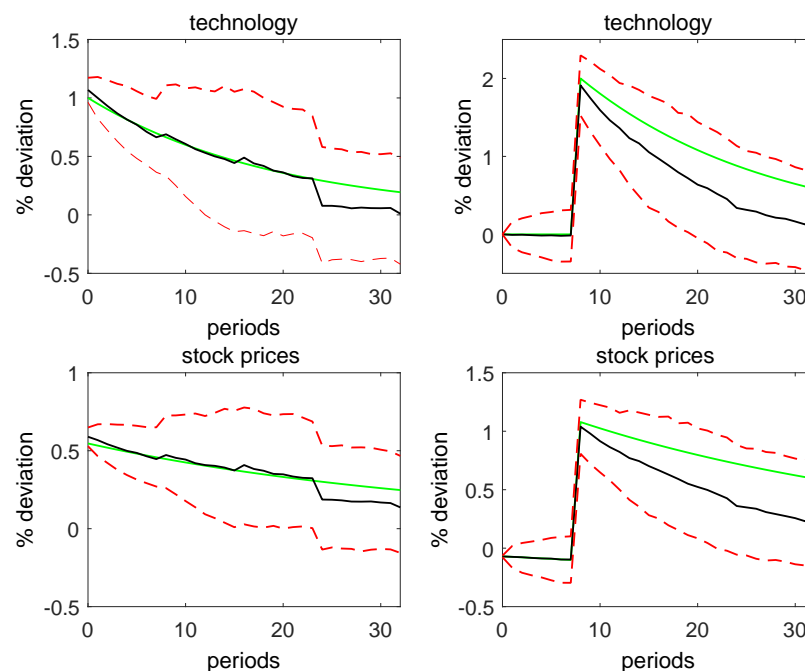
Note: The figure depicts FIA model impulse responses (in green) and estimated impulse responses (in black). Dashed red lines mark two-sided 90% confidence bounds. Impulse responses to $w_{0,t}^{\theta}$ are shown in the left column and impulse responses to $w_{8,t}^{\theta}$ in the right column.

improvement related to the surprise shock responses, whereas this is not the case for the anticipated shock. If one could speak of the optimal VAR lag length with respect to the MAWE, one might suggest a lag length of seventeen. This specification is an ambivalent choice because the minimal MAWE on the news shock side could be achieved without root flipping.

I present the estimation results for the VAR with seventeen lags and root flipping in Figure 7.6, which reveals that increasing the lag length to seventeen helps smooth the estimated impulse responses, primarily with respect to the surprise shock responses. This gain in shapes of the functions is accompanied by only minimal improvements in terms of the MAWE compared to the AIC specification as shown before.

At this stage, it can be stated that the CCA approach clearly outperforms the VAR model. Even CCA without root flipping dominates almost all VAR specifications. This conclusion is disappointing for the VAR econometrician because both methods should provide the same results at least asymptotically.

Hence, it seems reasonable to analyze the finite sample properties. I therefore conduct two more exercises in which I increase the sample size to 300 (Exercise 7.3) and 500 (Exercise 7.4) observations. Of course, such sample sizes are quite unrealistic when quarterly data are present. This aspect originates in the fact that finding a proxy for technology that is more frequent than a quarterly basis is rather difficult, whereas the availability of appropriate stock

Table 7.2: EXERCISES 7.3 & 7.4 - MAWE RESULTS (STOCK PRICES & SAMPLE SIZE)

	MAWE (VAR)				MAWE (CCA)			
Sample size	$t = 300$		$t = 500$		$t = 300$		$t = 500$	
Standard	1.3297	1.0063	1.1038	0.7102	1.0974	0.8031	0.9143	0.6111
	1.1023	0.8725	0.9139	0.6118	0.9652	0.7035	0.7948	0.5342
With flip	1.2078	1.0015	0.9683	0.6901	0.8205	0.7822	0.5923	0.5792
	1.0088	0.8847	0.8148	0.6048	0.7120	0.6967	0.5135	0.5160
Ratio (flip/stand.)	0.9083	0.9952	0.8773	0.9718	0.7477	0.9740	0.6479	0.9479
	0.9152	1.0139	0.8915	0.9885	0.7377	0.9903	0.6460	0.9659

Note: Analogously to Table 7.1, the table reports the computed MAWEs (and ratios) resulting from the simulation exercises in which the sample size is increased up to 300 and 500 observations. MAWE values are calculated as in equation (5.8) multiplied by 100.

price data is given without problems at a monthly or even higher frequency. Nevertheless, the goal of these exercises is to clarify the effects of small sample uncertainty on the results in general. Although I have shown above that the VAR lag length is a crucial determinant of the results, I use the AIC again in the next exercises. With a higher number of observations, I also increase the maximum number of lags that can be suggested by the AIC. In fact, the MAWE results, depending on the fixed lag lengths, indicate that the VAR model based on the AIC seems represent a good compromise in the end (see the appendix).

Table 7.2 displays selected findings, and each cell is organized as in the previous table. In combination with Table 7.1, I interpret the results as follows. Increasing the sample size leads to higher precision of the CCA and VAR methods. The relative gain due to root flipping increases with the sample size. For example, CCA diminishes the MAWE of the response to the surprise shock by approximately 25% (for $t = 300$) and 35% (for $t = 500$) due to root flipping. The VAR with root flipping yields an improvement as well but to a lesser extent. As regards the responses to the news shock, CCA with root flipping raises precision by 1–3% (for $t = 300$) and 3–5% (for $t = 500$), whereas VAR with root flipping appears successful only for $t = 500$. Yet, the CCA approach is strictly superior to its VAR opponent.

The previous findings point to a substantial advantage of the CCA approach when handling nonfundamentalness. In the following robustness check, I investigate several aspects that could be suspected to have an influence on the results. From Chapter 5, I know that the choice of the second observable and the relative standard deviation of the news shock matter for how the wedge of nonfundmentalness is split between the impulse responses to the two shocks. To test whether the choice of the second observable is an issue for the performance of the estimators, I first analyze systems in which stock prices are replaced by other variables (see Exercises 7.5 to 7.8). Afterwards, I alter the relevance of the news shocks by decreasing their relative standard deviation in the stochastic process (see Exercises 7.9 to 7.11). Chapter 5 also documented that the length of the anticipation horizon affects the magnitude of the wedge of nonfundamentalness: the longer the anticipation horizon, the larger the wedge. I

address this issue and investigate the implications of variation in q for the goodness of fit of the estimators in Exercises 7.12 to 7.17. In the final exercise, I lay the foundation for the empirical analysis in the subsequent chapter and extend the framework by adding another stochastic process to the model and allowing for nonstationarity.

7.4.3 Robustness check

Other variables

I repeat the experiments from above by replacing stock prices in the model with either output (Exercises 7.5 & 7.6) or investment (Exercises 7.7 & 7.8), but I report only the results for sample sizes $t = 200$ and $t = 500$ to save space. The findings from these exercises are summarized in Table 7.3.[8]

The experiments with output and investment as the second observables verify the previous outcome that theory-based root flipping can work even with small samples containing 200 observations. For the most part, the improvement is registered on the side of the surprise shock as before. According to the ratio computed as the MAWE of the root flipping technique divided by the MAWE of the standard method, root flipping yields MAWEs that are approximately 4% (for the technology-output system) and 11% (for the technology-investment system) lower in the VAR model and 21% (for the technology-output system) and 15% (for the technology-investment system) lower in the subspace algorithm analysis in comparison to the standard approaches. Increasing the sample size enhances the precision gain. CCA shows a quite robust performance in that context. In the technology-output system, improvement due to root flipping changes from 21% to 40% when increasing the sample size from 200 to 500 observations. In the technology-investment system, it rises from 15% (for $t = 200$) to 30% (for $t = 500$) in this case. VAR with root flipping improves in the technology-output system up to 20% for $t = 500$, whereas the change is negligible in terms of the MAWE when investment is the second observable.

Regarding the responses to the news shock, there is improvement in the experiments even with 200 observations for the subspace algorithm as well as for the VAR model. The improvement turns out to be of smaller magnitude in comparison to what is seen with respect to the surprise shock responses. Again, more observations lead to better performance and to greater benefits of root flipping. In the system including output, the benefit in terms of lower MAWEs increases from approximately 1% (for $t = 200$) to 4% (for $t = 500$) in the VAR model and from approximately 2% to 7% when CCA is applied. With investment as the second observable, the VAR method does not perform much better, but the CCA improves

---

[8]As before, I use the AIC to determine the lag length in the VAR model in all the remaining robustness exercises. In the appendix, I present the MAWE results of the VAR models for Exercises 7.5 to 7.8, depending on the fixed number of lags. The findings support my assessment in the previous exercises that the AIC is an appropriate candidate for selecting the VAR lag length in the simulations rather than fixing it at a particular number.

Table 7.3: EXERCISES 7.5–7.8 - MAWE RESULTS (OUTPUT & INVESTMENT)

		MAWE (VAR)				MAWE (CCA)			
	Sample size	$t = 200$		$t = 500$		$t = 200$		$t = 500$	
Output	Standard	1.3147	1.3009	1.2419	0.7049	1.5161	1.0729	1.0916	0.5995
		1.1849	1.2242	1.1294	0.6554	1.4398	1.0143	1.0298	0.5633
	With flip	1.2586	1.2819	0.9954	0.6737	1.1998	1.0512	0.6572	0.5540
		1.1385	1.2201	0.8994	0.6314	1.1297	1.0013	0.6157	0.5245
	Ratio	0.9574	0.9854	0.8015	0.9557	0.7914	0.9797	0.6021	0.9241
	(flip/std.)	0.9608	0.9967	0.7964	0.9634	0.7846	0.9872	0.5979	0.9312
Investment	Standard	1.9633	1.1076	1.1963	0.6848	1.3677	1.0463	0.9065	0.5922
		1.9083	1.0675	1.1759	0.6586	1.3489	1.0031	0.8822	0.5661
	With flip	1.7345	1.0813	1.0290	0.6677	1.1559	1.0408	0.6303	0.5700
		1.7036	1.0431	1.0253	0.6424	1.1412	0.9957	0.6236	0.5443
	Ratio	0.8835	0.9763	0.8601	0.9752	0.8452	0.9948	0.6953	0.9627
	(flip/std.)	0.8927	0.9772	0.8719	0.9755	0.8460	0.9926	0.7069	0.9614

Note: Analogously to Table 7.1, the table reports the computed MAWEs (and ratios) resulting from the simulation exercises in which the second observable is replaced either by output or investment. The table contains the results of the exercises in which the sample size is set to 200 and 500 observations, respectively. MAWE values are calculated as in equation (5.8) multiplied by 100.

from approximately 1% (for $t = 200$) to 4% (for $t = 500$).

If I consider the MAWE levels rather than the relative change, the subspace algorithm analysis strongly outperforms the VAR estimation. In nearly all cases, CCA is more successful than VAR – both with and without root flipping. The only exceptions are some of the surprise shock responses in the technology-output system.

Relevance of news shocks

As described in Chapter 5, the relative weight of the anticipated shock, λ_θ, is a crucial parameter in the theoretical model. The higher its value, the greater the systematic bias due to nonfundamentalness shifted toward the side of the surprise shock and vice versa. I investigate the implications of reducing the weight of the news shock for the estimation methods in further simulation exercises. Table 7.4 contains the corresponding findings. I present the MAWE results for three bivariate systems that include stock prices (Exercise 7.9), output (Exercise 7.10) or investment (Exercise 7.11) as the second observable. The sample size is again set to 200 observations.

Standard VAR and CCA reduce the MAWE with respect to the surprise shock. The gain due to root flipping is smaller than before. The difference in the precision of both methods is diminished, but CCA still performs somewhat better, except for the technology-investment system. Concerning the news shock, the MAWE substantially increases for the estimation procedures. Root flipping succeeds in decreasing the MAWE in only one case, namely, when the VAR method is applied to the technology-investment system.

I conclude that the relevance of the anticipated shock has strong implications for the estimation results. If the standard deviation of the news shock is small in relation to the surprise shock, the root flipping approach is less successful. Conversely, a more relevant news shock makes theory-based root flipping more effective.

Anticipation horizon

The length of the anticipation horizon is also relevant for the magnitude of the wedge of non-fundamentalness and for the performance of the estimators. To illustrate the latter point, I carry out two more exercises for each of the three bivariate systems, technology and stock prices (Exercises 7.12 and 7.13), technology and output (Exercises 7.14 and 7.15), and technology and investment (Exercises 7.16 and 7.17), in which I fix the anticipation horizon at four and twelve periods. The sample size is kept at 200 observations. Table 7.5 displays the corresponding MAWEs.

In combination with the results for $q = 8$, it can be assessed that the longer the anticipation horizon, the more effective root flipping for CCA and VAR. While there is a relatively small or even no benefit of the root flipping procedure for the short anticipation horizon, its gain

Table 7.4: EXERCISES 7.9–7.11 - MAWE RESULTS (LOW WEIGHT)

2nd observable	MAWE (VAR)			MAWE (CCA)		
	Stock prices	Output	Investment	Stock prices	Output	Investment
Standard	1.0986 2.0217	1.0981 2.0277	1.1734 1.6888	1.0211 2.4052	1.0344 2.7088	1.0804 2.0169
	0.9526 1.8818	1.0236 1.9245	1.1514 1.6436	0.9365 2.3000	0.9894 2.6597	1.0684 1.9864
With flip	1.0840 2.0875	1.0770 2.0274	1.1187 1.6429	0.9909 2.4356	1.0110 2.8026	1.0401 2.0236
	0.9403 2.0344	1.0027 1.9678	1.0996 1.6077	0.9095 2.3826	0.9665 2.7834	1.0319 1.9966
Ratio	0.9867 1.0326	0.9808 0.9999	0.9534 0.9728	0.9704 1.0127	0.9774 1.0346	0.9627 1.0033
(flip/std.)	0.9871 1.0811	0.9796 1.0225	0.9550 0.9782	0.9712 1.0359	0.9768 1.0465	0.9659 1.0051

Note: Analogously to Table 7.1, the table reports the computed MAWEs (and ratios) resulting from the simulation exercises in which the relative standard deviation of the news shock is reduced. The table contains the results of the exercises in which the second observable is either stock prices, output or investment. MAWE values are calculated as in equation (5.8) multiplied by 100.

Table 7.5: Exercises 7.12–7.17 - MAWE results (various q)

		MAWE (VAR)				MAWE (CCA)			
	Anticipation horizon	$q = 4$		$q = 12$		$q = 4$		$q = 12$	
Stock prices	Standard	1.3697	1.1896	1.4292	1.4229	1.1533	0.9566	1.7984	1.4295
		1.1617	1.0836	1.1531	1.2373	1.0456	0.8641	1.5995	1.2692
	With flip	1.3899	1.2042	1.3092	1.3787	1.0870	0.9611	1.3791	1.3592
		1.1846	1.1085	1.1113	1.2362	0.9770	0.8725	1.2076	1.2299
	Ratio	1.0148	1.0123	0.9161	0.9689	0.9425	1.0047	0.7668	0.9508
	(flip/std.)	1.0197	1.0230	0.9637	0.9991	0.9344	1.0097	0.7550	0.9690
Output	Standard	1.2874	1.1479	1.4815	1.3698	1.1999	0.9332	1.9634	1.4637
		1.1834	1.0876	1.3311	1.2820	1.1496	0.8879	1.8532	1.3856
	With flip	1.2767	1.1550	1.1778	1.2714	1.1113	0.9342	1.4885	1.3877
		1.1718	1.0993	1.0515	1.2038	1.0579	0.8913	1.3906	1.3289
	Ratio	0.9917	1.0062	0.7950	0.9282	0.9262	1.0011	0.7581	0.9481
	(flip/std.)	0.9902	1.0107	0.7899	0.9390	0.9202	1.0038	0.7504	0.9591
Investment	Standard	1.8313	1.0136	1.9822	1.2359	1.1316	0.9333	1.7804	1.3959
		1.7825	0.9926	1.9122	1.1738	1.1063	0.9050	1.7652	1.3355
	With flip	1.7807	1.0149	1.5420	1.1504	1.0784	0.9390	1.3834	1.3353
		1.7411	0.9955	1.5060	1.0880	1.0562	0.9105	1.3792	1.2695
	Ratio	0.9724	1.0012	0.7779	0.9308	0.9529	1.0060	0.7770	0.9565
	(flip/std.)	0.9768	1.0030	0.7876	0.9270	0.9547	1.0061	0.7813	0.9506

Note: Analogously to Table 7.1, the table reports the computed MAWEs (and ratios) resulting from the simulation exercises in which the length of the anticipation horizon is set equal to four and twelve, respectively. Note that, in the exercises with $q = 12$, it is also ensured that the maximum number of lags suggested by the AIC exceeds the anticipation horizon. The table contains the results of the exercises in which the second observable is either stock prices, output or investment. MAWE values are calculated as in equation (5.8) multiplied by 100.

in precision becomes obvious for $q = 12$. It should be noticed that for the long anticipation horizon, i.e., $q = 12$, VAR outperforms CCA in some cases, e.g., the MAWE in terms of levels associated with the estimated impulse responses to the surprise shock in the technology-stocks and technology-output systems and the MAWE ratios with respect to the news shock in the systems, where output and investment are selected as the second observables. CCA seems more vulnerable to such a long anticipation horizon because the relative increase in the MAWE levels is much larger in comparison to the VAR when increasing q from four to twelve.

Model extension

If the goal is to employ theory-based root flipping in an empirical investigation, other factors than those thus far considered regarding the performance of the estimation methods have to be taken into account. The preceding theoretical model might not yet be sufficient, as it only features stationary behavior and only two kinds of technology shocks. I therefore augment the model by replacing the stationary stochastic process of technology by a unit root process and adding another stochastic process to the model, which represents a preference shock.[9]

The technology process is now described by

$$log\left(\theta_t\right) \;\; = \;\; log\left(\gamma_\theta\right) + log\left(\theta_{t-1}\right) + \sigma_\theta w_{0,t}^\theta + \sigma_\theta \lambda_\theta w_{q,t-q}^\theta, \tag{7.5}$$

where γ_θ displays the deterministic growth factor of technology. The preference shock is given by a stationary AR(1) process of the form

$$log\left(\psi_t\right) \;\; = \;\; \left(1 - \rho_\psi\right) log\left(\psi\right) + \rho_\psi log\left(\psi_{t-1}\right) + \sigma_\psi w_t^\psi, \tag{7.6}$$

where $0 < \rho_\psi < 1$, and w_t^ψ is white noise with standard deviation σ_ψ.

I arrange the vector of shocks in the empirical model such that the preference shock is in the third position. Because there are three shocks in the setting, I extend the vector of observables to be three-dimensional. The complete Blaschke matrix, which transforms $M\left(L\right)$ into $\widehat{M}\left(L\right)$, is now given by the identity matrix in which the upper left (2×2) block is still represented by the former Blaschke matrix as defined in equation (7.3). Consequently, there is no difference between the dynamics of the preference shock in $M\left(L\right)$ and $\widehat{M}\left(L\right)$.

In the simulation exercise (Exercise 7.18), I use the calibration $\gamma_\theta = 1.004$, $\rho_\psi = 0.5$ and $\sigma_\psi = 1/3$. For all the remaining parameters, I keep the calibrated values of the benchmark system, especially the relative weight of the news shock, which is set to a value of two. As in the bivariate case, I generate 1000 sets of artificial data with an effective sample size of 200 and compute the median over all estimation runs. I choose technology, stock prices and

---

[9]Note that adding the preference shock to the system increases the state dimension by one unit.

consumption as observables for the empirical model. As estimation methods, I use the VAR and CCA counterparts in the framework of cointegrated time series, namely, VECM and ACCA, as introduced in the third chapter.

The higher system dimension also requires the adjustment of the identification scheme. Compared to the bivariate case above, I need two more restrictions for the structural shock identification. Theoretically justified restrictions are a zero short-run and a zero long-run restriction of the preference shock on technology. Note that there is no distinction between imposing these restrictions appropriately on $M(0)$ and $M(1)$ or on $\widehat{M}(0)$ and $\widehat{M}(1)$. Because the preference shock does not affect the nonfundamentalness property of the system, I follow the standard approach with regard to these two additional restrictions. Hence, the only source of error in the standard approach comes from the incorrect restriction on the news shock. Furthermore, the theory-based root-flipping procedure does not have an effect upon the estimated impulse responses to the preference shock. Thus, as before, I only concentrate on the impulse responses to the two technology shocks in the analysis.

Figures 7.7 and 7.8 illustrate the estimated impulse responses for the VECM and ACCA.[10] As in the bivariate exercises, the visual comparison between standard and root flipping technique mainly permits a qualitative rather than a quantitative appraisal. The impulse responses of the VECM with root flipping show jagged behavior, which is slightly seen in the ACCA-related functions. The VECM exhibits wider confidence bounds while the ACCA is subject to a lower degree of uncertainty with respect to the surprise shock responses. Regarding the news shock responses, a significant difference is not visible for both VECM and ACCA.

To shed more light on the quantitative effects, Table 7.6 presents the computed MAWE statistics. Because the root flipping procedure does not affect the estimates with respect to the preference shock, I only present the statistics associated with the two technology shocks, i.e., each cell contains 3×2 matrices, where the order of the three rows corresponds to the order of the vector of observables (technology, stock prices and consumption). As in most of the previous exercises, I experience precision gains due to root flipping, whereby the success occurs primarily on the side of the surprise shock. Root flipping provides an improvement of approximately 25% in the ACCA and approximately 20% in the VECM according to the computed MAWE ratio. The dominance of the subspace algorithm over the VAR method (i.e., VECM here) is reduced in the current example in comparison to the exercises above. Now, it needs root flipping to maintain its superior position over the VECM. Nevertheless, ACCA can outperform VECM in terms of MAWE levels. This is not the case for responses to the news shock for which the VECM offers higher accuracy, albeit root flipping helps only slightly. This result is also in line with my previous findings.

---

[10]For various lag lengths of the VECM, I find similar outcomes to the bivariate cases above. Therefore, I specify the VECM according to the lag length suggested by the AIC.

Figure 7.7: EXERCISE 7.18 - IMPULSE RESPONSES (VECM RESULTS)

STANDARD VECM
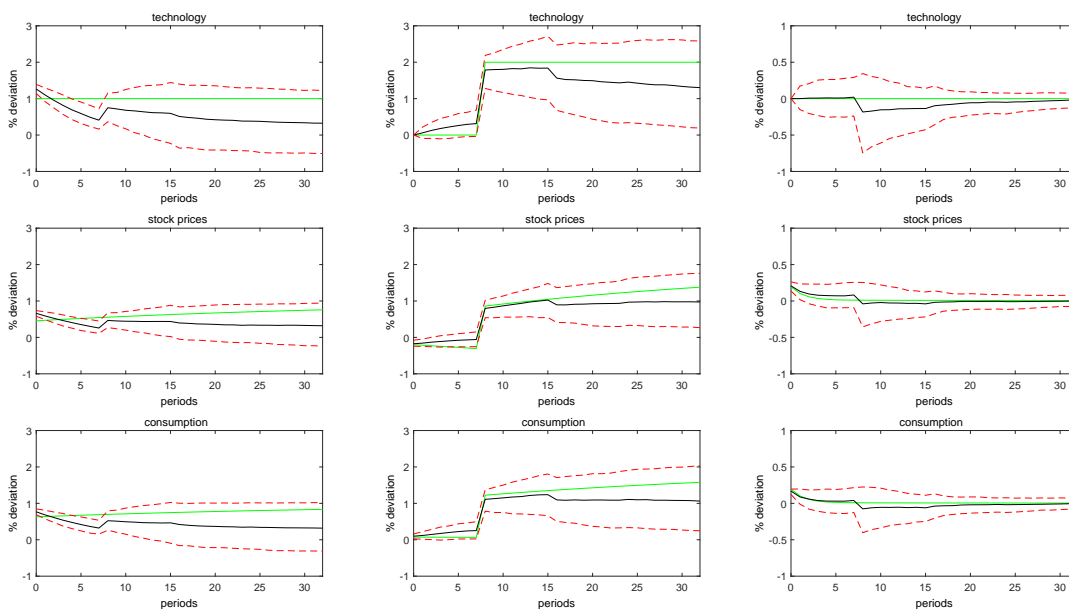
VECM WITH FLIP
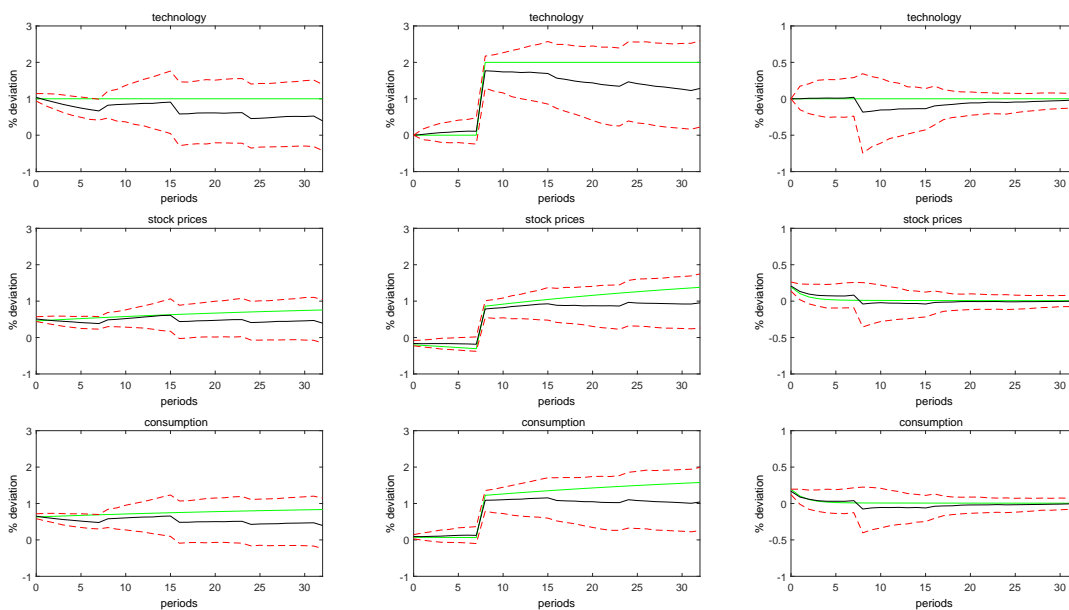
Note: The figure depicts FIA model impulse responses (in green) and estimated impulse responses (in black). Dashed red lines mark two-sided 90% confidence bounds. Impulse responses to $w_{0,t}^{\theta}$ are shown in the left column, impulse responses to $w_{8,t}^{\theta}$ in the middle column and impulse responses to w_t^{ψ} in the right column of both panels.

Figure 7.8: EXERCISE 7.18 - IMPULSE RESPONSES (ACCA RESULTS)

STANDARD ACCA

ACCA WITH FLIP

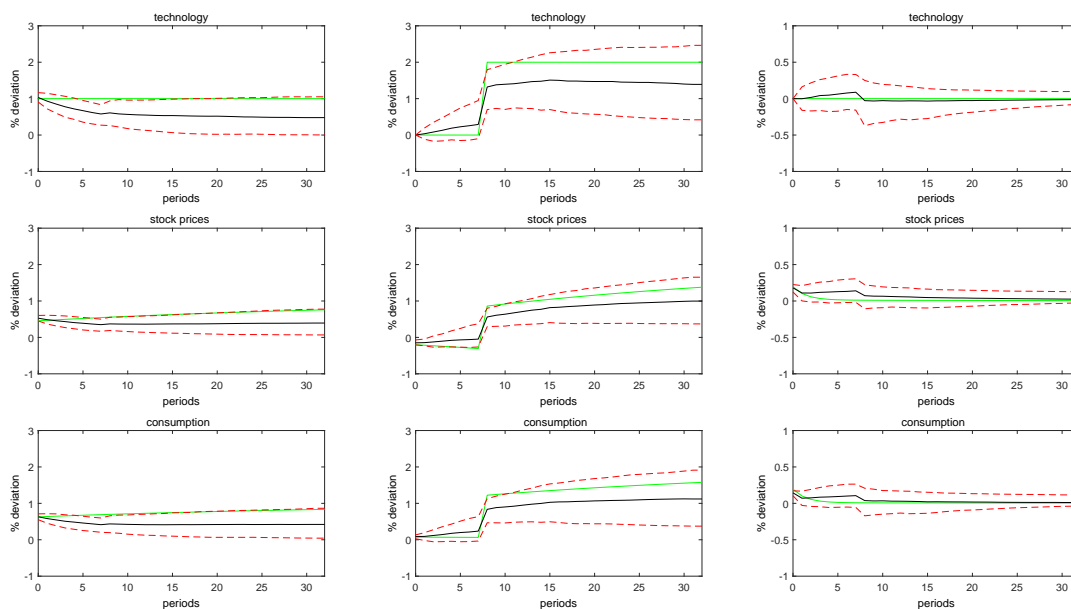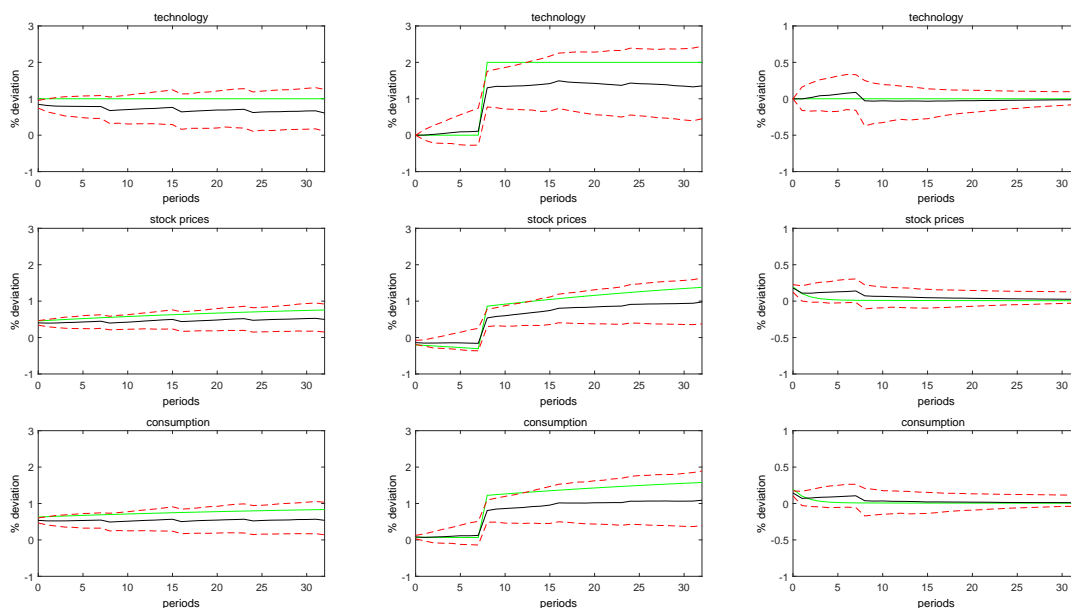**Note**: The figure depicts FIA model impulse responses (in green) and estimated impulse responses (in black). Dashed red lines mark two-sided 90% confidence bounds. Impulse responses to $w_{0,t}^{\theta}$ are shown in the left column, impulse responses to $w_{8,t}^{\theta}$ in the middle column and impulse responses to w_t^{ψ} in the right column of both panels.

Table 7.6: EXERCISE 7.18 - MAWE RESULTS (NONSTATIONARY SYSTEM)

	MAWE (VECM)		MAWE (ACCA)	
Standard	1.5811	1.0307	1.3125	1.1100
	1.3919	0.9248	1.2021	1.0257
	1.4861	0.9706	1.2633	1.0504
With flip	1.2624	1.0090	0.9921	1.0910
	1.1030	0.9071	0.9036	1.0173
	1.1836	0.9632	0.9515	1.0471
Ratio (flip/standard)	0.7984	0.9789	0.7559	0.9829
	0.7925	0.9809	0.7517	0.9919
	0.7965	0.9924	0.7532	0.9968

Note: Analogously to Table 7.1, the table reports the computed MAWEs (and ratios) resulting from the simulation exercise in which the set of observables comprises three variables (technology, stock prices and consumption), which are driven by one common stochastic trend. MAWE values are calculated as in equation (5.8) multiplied by 100.

It should be noted that my foregoing conclusions are not independent of the chosen values for γ_θ, ρ_ψ and σ_ψ. The value of γ_θ is selected such that it roughly matches the average quarterly growth rate of the TFP series used in the empirical analysis in Chapter 8. While γ_θ affects the anticipation rate, the two parameters related to the preference shock do not.[11] Nevertheless, persistence and volatility of the preference shock influence the performance of the estimation methods in small sample experiments, but in contrast to γ_θ, they are not relevant for the specification of the theory-based root flipping estimator when it is applied practically. I therefore do not deepen the analysis here and move on to the empirical examination after summarizing the main aspects of this chapter.

7.5 Summary and discussion

Nonfundamentalness implies that the econometrician's information set is smaller than the information set of the agents in the model economy. As a consequence, the econometrician is not able to uncover the true model dynamics. One needs to find estimation methods that are not vulnerable to that problem and/or use additional information that helps eliminate misalignment between the different information sets. In this chapter, I have explored an alternative way of addressing nonfundamentalness in a situation in which only a small number of observables is available and, thus, the econometrician's access to information is strongly limited.

As a possible solution, I presented an estimator that is primarily targeted to environments in

---

[11]Including the preference shock in the workhorse model adds expressions such as the $\widetilde{s}_t$ and $E\left[\widetilde{s}_{t+1}\right]$ terms to the second order difference equation (5.9). However, as it only represents a surprise shock, solving the unstable root forward offsets any future expected preference shocks so that the anticipation rate remains untouched.

which economic agents anticipate shocks to future fundamentals. Originally introduced by Mertens and Ravn (2010), I applied their procedure to a new setting – when agents receive news about future technology growth. This context has attracted much attention in recent years in both theoretical and empirical business cycle research. The estimator combines theory with empirical methods to deal with the difficulties caused by nonfundamentalness. Endowing the econometrician with auxiliary information based on theory can reduce the discrepancy in relation to the agents' information set. Although this approach provokes limitations in terms of additional restrictions on the regression model, the proposed estimator maintains the basic idea behind traditional SVAR models to impose a minimum structure on the data because it only requires limited information about the anticipation rate and horizon as well as the relative weight of news compared to other structural shocks. In contrast, alternative methods that can tackle nonfundamentalness, such as dynamic factor models or Bayesian estimation of fully specified models, either necessitate a sizable set of observable variables or complete knowledge of the entire economic model.

A contribution of this chapter is in embedding of the theory-based estimator into the framework of subspace algorithm analysis. In the simulation experiments, I demonstrated that this technique can provide improvements over its SVAR counterpart and constitutes a step toward coping with the problem of nonfundamentalness in small scale systems without using external information.

It must be noted that my workhorse model might be viewed as being not realistic enough in terms of typical DSGE model features, e.g., nominal or more real frictions, to study the performance of the presented estimators in Monte Carlo experiments. This has been done by Sims (2012), Seymen (2013) and Beaudry et al. (2013), for instance, but not with respect to the subspace algorithm analysis or root flipping procedure. A possible modification of the workhorse model could also be the specification of the technology process such that it exhibits a delayed but slow diffusion news as proposed by Portier (2015) rather than a distinct jump after the anticipation period.

Another outcome of this chapter has already been indicated in the previous chapters and in the Monte Carlo studies by Sims (2012), Seymen (2013) and Beaudry et al. (2013): nonfundamentalness seems not to be a matter of serious concern for the econometrician in the context of technology-related news because the bias in the estimated impulse responses is relatively small compared to other sources of bias, such as small sample uncertainty.[12] Does this mean that testing the theory-based root flipping procedure has been worthless? No, not if one thinks of the Mertens and Ravn (2010) puzzle mentioned in Section 7.1 or the exemplification of Leeper et al. (2013) that the econometrician can infer mistaken conclusions from her impulse responses, which contradict the true responses in the environment of the

---

[12]See also Beaudry and Portier (2014b) and Beaudry et al. (2015) for a discussion of the importance of nonfundamentalness in a setting with technological news shocks. Beaudry et al. (2015) suggest a diagnostic test to check whether the quantitative implications of nonfundamentalness are relevant for the empirical analysis.

neoclassical growth model augmented with anticipated tax shocks. I resume the example of Leeper et al. (2013) in the appendix, wherein I disclose this finding in another version of my workhorse model extended with tax news. By applying the tools developed in Chapter 5, I show that the responses to a tax news shock, as detected by the econometrician, differ not only in shape but also in sign from the FIA model impulse responses. Therefore, this chapter's results concerning the possible gain from reducing the bias in the impulse responses due to root flipping can indeed be seen as promising, and it is worth studying the concept of theory-based root flipping further. However, as my focus is the idea of news shocks in the spirit of Beaudry and Portier (2006), I leave the last example as motivation for future related research and turn to the empirical investigation of news shocks in the following and final chapter.

Chapter 8

Exploring the empirical relevance of news shocks

In this chapter, I conduct an empirical analysis of the US data focusing on the role of news shocks in explaining business cycles. The objective is to bring together the estimation methods I have presented in this thesis in a common framework and to conclude my thesis with a practical application of these tools.

8.1   Motivation and related literature

News shocks as triggers of macroeconomic fluctuation have become an important topic in empirical business cycle analysis in recent years. Their factual relevance is a controversial subject and has attracted considerable debate in the profession. Since the influential work by Beaudry and Portier (2005, 2006), the subject of technology-related news has become a central point of the discussion.[1]

Structural innovations in stock prices identified in a structural VECM that comprises a proxy for technology, stock prices and macroeconomic aggregates can be a decisive source of business cycle fluctuation. In examining quarterly postwar data for the US, Beaudry and Portier (2006) find that anticipated neutral technology shocks account for more than one-half of the forecast error variance of output, consumption and hours at business cycle frequencies. Beaudry and Lucke (2010) extend the set of possible shocks as drivers of the business cycle to candidates such as IST, monetary policy and preference shocks, and they confirm an outstanding role for anticipated technology shocks in a five-dimensional system including TFP, stock prices, relative price of investment, interest rate and an activity measure. Ac-

---

[1]My focus is the empirical relevance of rather than a theoretical perspective on news shocks. One of the challenges of the latter context is building a model in which news shocks generate comovement between macroeconomic quantities. For related contributions, see the articles by Beaudry and Portier (2007), Jaimovich and Rebelo (2009), Christiano et al. (2010), Pavlov and Weder (2013) and Dupor and Mehkari (2014).

cording to their findings, news about TFP growth explain approximately 70% of variation in hours worked at forecast horizons that exceed 1–2 years, whereas surprise changes in neutral and investment-specific technology or monetary shocks play a minor role at all horizons. The only exception is what Beaudry and Lucke (2010) interpret as a preference shock, which dominantly contributes to very short-run movements. Similar results are obtained with respect to other activity variables, such as output, consumption and investment, where the technological news shock explains at least a 50% share in the variance decomposition at horizons longer than one year.

The results of Beaudry and Portier (2006) and Beaudry and Lucke (2010) are challenged by several studies that relate to three key areas of concern: the identification procedure, the problems caused by nonfundamentalness and the estimation methodology (and its results) in general. Barsky and Sims (2011) undertake an SVAR-based analysis that differs in the identification methodology (and the choice of the data series) from the approach in the aforementioned articles. While Beaudry and Portier (2006) and Beaudry and Lucke (2010) deploy traditional zero short-run and long-run restrictions to identify all structural shocks in their systems, Barsky and Sims (2011) develop an alternative partial identification strategy to uncover surprise and anticipated changes in technology. They identify the news shock as being orthogonal to the unanticipated shock in technology and maximizing the forecast error variance share of TFP over a fixed horizon. Moreover, they use the utilization-corrected TFP series from Basu et al. (2006) included in a seven-dimensional VAR with output, consumption and hours as well as inflation, stock prices and consumer confidence. Their findings do not debunk the pivotal role of anticipated neutral technology shocks in fluctuations over the business cycle horizon. For example, news shocks dominate hours variation only over the very short-run with a fraction of 60% of the FEVD, whereas this share declines to below 10% at lower frequencies. In the case of output and consumption, TFP news explain approximately 40% to 50% of the FEVD but not until a medium-term horizon of four years.

As explicated in the preceding chapters, VAR-based identification of news shocks faces the criticism that the underlying theoretical model might feature a nonfundamental representation, so it cannot uncover the true shocks and their dynamics. This point is raised by, e.g., Schmitt-Grohé (2010) regarding the VECM of Beaudry and Lucke (2010), whereas Forni et al. (2014) address the nonfundamental representation underlying the VECM of Beaudry and Portier (2006). Based on an empirical test of detecting nonfundamentalness in the data, Forni et al. (2014) infer that the inclusion of consumer sentiment in the set of observables, as in the case of Barsky and Sims (2011), prevents the consequences of nonfundamentalness.[2] They further propose to estimate a FAVAR model and apply the Barsky and Sims (2011) identification procedure. Forni et al. (2014) find that TFP news explain from 5% to 43% of the volatility in hours (5%), investment (17%), output (21%) and consumption (43%)

---

[2]See also Forni and Gambetti (2014), who introduce the empirical test of detecting nonfundamentalness.

at business cycle frequencies.

A different approach to quantifying the contribution of news shocks to aggregate fluctuations can be found in, e.g., Fujiwara et al. (2011), Schmitt-Grohé and Uribe (2012) and Khan and Tsoukalas (2012), who rely on the application of a full-information method by estimating a fully specified DSGE model via Bayesian techniques. These studies all have a common outcome that the relative importance of anticipated technology shocks is limited to a small portion of the variance decomposition of the main macroeconomic aggregates. Fujiwara et al. (2011) find that the corresponding fraction of technological news shocks in macroeconomic fluctuations lies approximately between 3% and 13% while Schmitt-Grohé and Uribe (2012) and Khan and Tsoukalas (2012) estimate variance shares of technological news shocks between 1% and 3%. The latter two analyses expose the potential relevance of non-technological news shocks, which account for a sizable fraction ranging from 60% to 70% in the variation in hours and, to a lesser extent, in consumption and output. The main part of the variation in the latter variables is captured by wage markup shocks.

It is worth mentioning that there are also various contributions in the literature, which, in turn, give rise to reasonable doubt about the critique of the news shock hypothesis of Beaudry and Portier (2006). Beaudry et al. (2011) argue that the Barsky and Sims (2011) identification method is sensitive to the choice of the truncation horizon that is used in this procedure. Moreover, Beaudry et al. (2013) demonstrate in a Monte Carlo simulation study that the discrepancy between the findings of Beaudry and Portier (2006) and Barsky and Sims (2011) is related more to the informational content of the variables used in the SVARs or the small sample uncertainty rather than to the different identification schemes.[3]

The second issue of concern, i.e., the consequences of nonfundamentalness, is somehow refuted by means of Monte Carlo evidence by Beaudry and Lucke (2010) and Sims (2012). Beaudry and Lucke (2010) show in their web appendix that their estimation procedure works successfully in a simulation study. Sims (2012) emphasizes that nonfundamentalness is not an "either/or proposition" because VAR models still perform quite convincingly in his Monte Carlo experiments (see also the previous chapter of this thesis).

Recently, Beaudry and Portier (2014a) revisit their own study and address inter alia the third point of concern, i.e., the DSGE model-based findings that wage markup shocks play a dominant role in explaining business cycles while the role of technological news is negligible. Beaudry and Portier (2014a) propose that the importance of wage markup shocks is disputed in the literature and may hint at model misspecification. The authors use an updated data set and replicate the Beaudry and Portier (2006) results with various types of SVAR models that differ, e.g., in the number and choice of variables. Furthermore, they question the FAVAR model outcome of Forni et al. (2014) by including the (first two) es-

---

[3]See also Sims (2016), who demonstrates that using the most recent vintage of the TFP series constructed by Fernald (2014) in the Barsky and Sims (2011) model can produce results in favor of the Beaudry and Portier (2006) findings.

timated factors of Forni et al. (2014) in their own SVAR model and discovering that stock prices do not immediately respond to the news shock. As stock prices are chosen in order to segregate the news shock from other structural shocks, it is therefore unconvincing that stock prices are not affected by this kind of shock on impact.

The aim of this chapter is to contribute to the previously presented literature and to reinvestigate the empirical relevance of technological news. Thereby, I use the core model in Beaudry and Portier (2014a) as a benchmark and test the estimators of the preceding chapter in an empirical environment. The chapter brings together the findings and tools considered in the foregoing chapters of this thesis.

8.2 Empirical application

8.2.1 Course of action

My course of action follows Beaudry and Portier (2014a). I consider the trivariate system presented in Chapter 7, i.e., TFP, stock prices and consumption, which is the baseline system of Beaudry and Portier (2014a). The data set is described in the following subsection.

The goal is to identify three structural shocks in the system. One shock is supposed to be a surprise technology shock while the second is presumed to be an anticipated technology shock. The interpretation of the third shock is left open at this stage and will be discussed in a separate subsection. Prior to that, like Beaudry and Portier (2014a), I focus on the implications of the technological news shock in terms of the corresponding impulse responses and the FEVD. In a first step, I reproduce the results of Beaudry and Portier (2014a) in estimating a SVECM comprising the three variables with three lags in differences and two cointegrating relations. The identification scheme corresponds to the scheme in Chapter 7. That is, I impose two zero short-run restrictions: the anticipated technology shock and the third structural shock do not affect TFP on impact. I also impose a zero long-run restriction: the third shock has no effect on TFP in the longer term.

The results are supplemented by the findings for alternative estimation techniques. Similar to Chapter 7, I use the subspace algorithm analysis and apply the theory-based root flipping procedures associated with the SVECM and the subspace algorithm introduced therein. Some of the details regarding the procedures are addressed below when presenting the results. In the final step, I shift the focus and elaborate the interpretation of the third structural shock. This step completes the circle by referring to the first part of my thesis and provides closing statements.

8.2.2   Data set

I use the quarterly data set of Beaudry and Portier (2014a), which is available on the companion website for the article. The sample covers the 1947Q1–2012Q3 period. The TFP series is taken from Fernald (2014). Stock prices are represented by the Standard & Poor's 500 index series drawn from Federal Reserve Economic Data (FRED). Real personal consumption expenditures on goods and services from the NIPA serve as the consumption series. For deflating stock prices, the NIPA output deflator is used. To form per capita values, stock prices and consumption are divided by the series of total population from the FRED database.

8.2.3   Results

At first, I begin with the outcomes of the SVECM replicating the findings of Beaudry and Portier (2014a). The top panel of Figure 8.1 displays the estimated impulse response of TFP to the news shock (black lines) with the corresponding confidence intervals (dashed red lines), which shows the typical feature that TFP increases with a delay of approximately ten quarters.[4] There is some negative reaction in the first quarters, which Portier (2015) interprets as a "consequence of an excessive correction for utilization." The FEVD of TFP, as presented in (the upper panel of the second column of) Table 8.1, supports the technological news interpretation of the identified shock. The news shock does not explain a substantial part of the TFP forecast error variation for the first years, but its share steadily increases to more than 80% over the long-run.

In the middle panel of Figure 8.1, I present the estimated impulse response of stock prices to the news shock and the associated confidence bounds. Stock prices jump by 7.5% on impact and remain at approximately this level thereafter. The FEVD of stock prices (in the middle panel of the second column of Table 8.1) indicates that stock prices capture the news shock adequately because the news shock accounts for nearly 90% or more of the variation at all horizons.

The estimated impulse response of consumption to the news shock is illustrated in the lower panel of Figure 8.1. Consumption increases slightly on impact and increases further in the first year after the shock occurs, remaining at a level of approximately 1% higher thereafter. Within the first year, the news shock explains only a small part of the forecast error variance of consumption, which is reversed at longer horizons (see the bottom panel of the second column of Table 8.1). After four quarters, the news shock is the dominant source of movements in consumption.

Next, I turn to the alternative approaches, i.e., the subspace algorithm analysis and theory-based root flipping procedures applied to the VECM and the subspace algorithm. The find-

---

[4]The first positive point estimate of the impulse response is in the ninth quarter while the first significant positive response of TFP occurs after eleven quarters.

Figure 8.1: ESTIMATED IMPULSE RESPONSES TO NEWS SHOCK IN US DATA
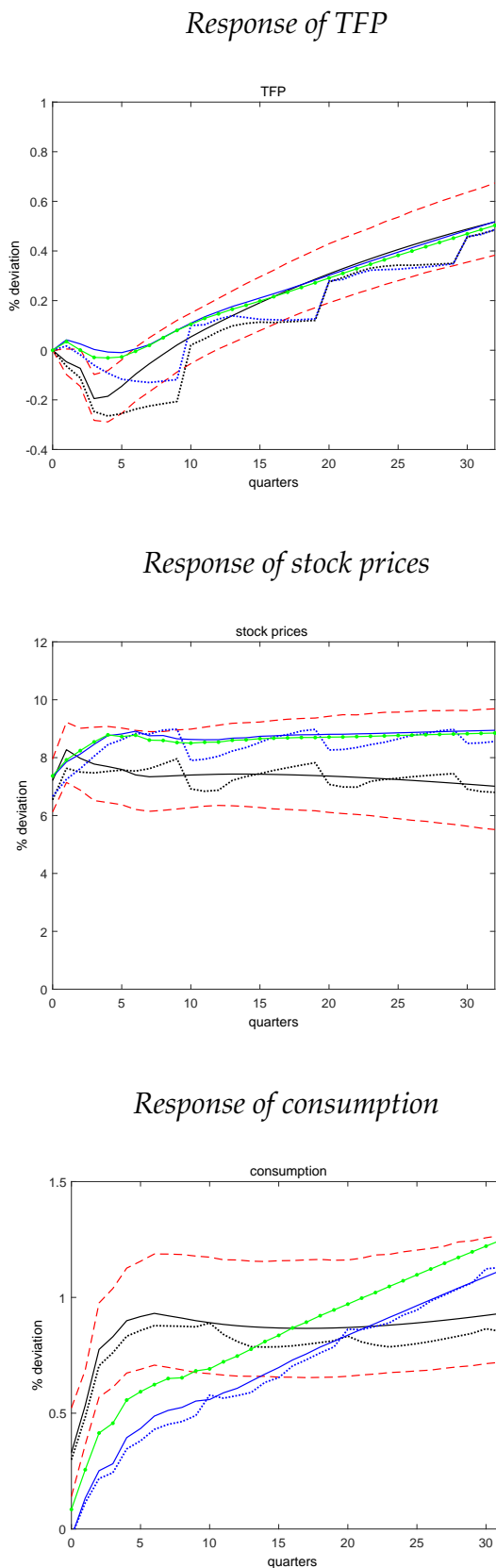
Response of TFP

Response of stock prices

Response of consumption

Note: The figure depicts the estimated impulse responses resulting from different estimation approaches. These approaches are standard SVECM (solid black), SVECM with root flipping (dotted black), standard structural CCA based on the [1,3]-restriction (solid blue), structural CCA based on the [1,3]-restriction with root flipping (dotted blue) and standard structural CCA based on the [3,3]-restriction (solid green). Dashed red lines mark the 16% and 84% quantiles of the distribution of the impulse responses w.r.t. the standard SVECM, obtained by the approach as discussed in Doan (1992).

Table 8.1: SHARE OF FORECAST ERROR VARIANCE EXPLAINED BY NEWS SHOCK

		SVECM		Structural CCA		
	Horizon	Standard	Root-flip	Standard [1,3]-restr.	Root-flip [1,3]-restr.	Standard [3,3]-restr.
TFP	1	0.00	0.00	0.00	0.00	0.00
	4	0.02	0.05	0.00	0.00	0.00
	8	0.03	0.09	0.00	0.02	0.00
	16	0.03	0.07	0.03	0.03	0.03
	24	0.10	0.10	0.11	0.07	0.10
	32	0.22	0.18	0.23	0.16	0.21
	50	0.48	0.40	0.53	0.44	0.50
	120	0.83	0.76	0.91	0.91	0.88
Stock prices	1	0.87	0.82	0.98	0.93	0.99
	4	0.91	0.87	0.98	0.92	1.00
	8	0.91	0.89	0.98	0.93	0.98
	16	0.92	0.90	0.98	0.93	0.97
	24	0.93	0.92	0.99	0.93	0.97
	32	0.94	0.93	0.99	0.94	0.97
	50	0.95	0.94	0.99	0.94	0.97
	120	0.96	0.96	0.99	0.95	0.97
Consumption	1	0.21	0.17	0.00	0.00	0.01
	4	0.53	0.45	0.05	0.04	0.15
	8	0.61	0.54	0.15	0.11	0.28
	16	0.64	0.57	0.26	0.23	0.42
	24	0.65	0.57	0.37	0.34	0.53
	32	0.67	0.58	0.45	0.44	0.62
	50	0.72	0.61	0.58	0.60	0.74
	120	0.88	0.78	0.75	0.81	0.84

Note: The table reports the FEVD shares of TFP (top panel), stock prices (middle panel) and consumption (bottom panel) due to the technological news shock identified via the various estimation techniques as labeled in each column.

ings are added to the aforementioned figure and table to allow a direct comparison with the previous results. With the subspace algorithm analysis, it is difficult to identify a cointegration relationship among the chosen variables. Note that performing a standard Johansen cointegration test does not provide a fully convincing picture either, as it implies a cointegration rank of one only at the 10% significance level. The difficulty concerning the subspace algorithm is reflected in the result that three eigenvalues of the $\mathcal{A}$-matrix in the state equation are estimated to be very close to one, i.e., the system might be driven by three common trends instead of one as Beaudry and Portier (2014a) presume.

A possible step could be to impose a cointegration rank of one by estimating the system matrices of the state equation via a reduced rank procedure (see Bauer and Wagner (2009)) so that (only) one eigenvalue of the $\mathcal{A}$-matrix is restricted to an exact value of one while the other two eigenvalues are not fixed at this value but instead estimated to be smaller than one. However, in this case, the corresponding estimated values are nearly one. Then, proceeding with the above-mentioned identification scheme, where the zero long-run restriction is placed on the total impact matrix derived from the MA representation of the LIE model (as described in Chapter 3), leads to results that are not convincing and seem highly sensitive.

A natural strategy when faced with nonstationary time series without cointegrating features would be to estimate the system in first differences, but this would ignore the long-run properties of the data. I therefore choose a different way inspired by the common procedure to estimate a VAR in levels as an alternative to a VECM.[5] I adopt this approach to the subspace algorithm analysis, i.e., I deploy the CCA instead of the ACCA but set up the corresponding observation equation for the vector of observables in terms of levels and not in terms of first differences. This technique is an ad hoc and unconventional way of dealing with the data, but it produces more plausible results than in the case of the ACCA. Moreover, it allows using the same identification scheme as Beaudry and Portier (2014a) because three restrictions are necessary for the identification of the three structural shocks in a trivariate system in levels. The difference to the approaches based on the cointegration assumption is that the long-run effect of the third shock is only constrained with respect to one variable instead of all three variables in the system. I impose the zero long-run restriction of the third shock on TFP that is the variable ordered first in the system, i.e., I restrict the [1,3]-element of the total impact matrix to zero. In the discussion below (Subsection (8.2.4)), I also consider the case wherein the restriction is applied to the [3,3]-element of the total impact matrix (i.e., to consumption).

---

[5]Asymptotically, a VAR in levels and the corresponding VECM representation are equivalent. See also the derivation of equation (3.10). On the one hand, estimating a VAR in levels ignores restrictions that are imposed on the coefficients when using a VECM. On the other hand, Sims et al. (1990) propose to apply a VAR in levels rather than a VECM when the degree of integration in the system is uncertain. I follow their suggestion and focus on the estimation of the system in levels. I thereby proceed without providing a formal treatment of that issue, as this would exceed the scope of this thesis.

Before turning to the corresponding results, I point to a particular issue that has come up in Chapter 7 and has been suppressed thus far: the presumption that the roots related to the MA representation, which are associated with the nonfundamentalness property of the underlying theoretical model, lie on a circle that has a radius smaller than one. Mertens and Ravn (2010) address this subject in their empirical application and essentially find an empirical circle of roots in their VAR system. As a by-product of the subspace algorithm analysis, I can also uncover the (estimated) MA roots in the data. Recall that the econometrician estimates the reciprocals of the true MA roots in the case of nonfundamentalness. Because, in turn, the estimated MA roots are equivalent to the reciprocals of the estimated (nonzero) eigenvalues of the matrix $(\mathcal{A} - K\mathcal{C})$, I simply use the estimated eigenvalues of $(\mathcal{A} - K\mathcal{C})$ as direct counterparts of the true MA roots.

The first attempt of producing results that are consistent with the theoretical framework above is not successful. When letting the data speak and using the criterion by Bauer and Wagner (2002) to estimate the number of states, I obtain four states yielding eigenvalues of $(\mathcal{A} - K\mathcal{C})$ equal to absolute values of 0.08, 0.10 and 0.60, where the latter value appears in a pair. My second attempt is justified by theory and guided by the VECM result from above. The theoretical model implies $q + 2$ states. As I use $q = 10$ as a benchmark based on the observation in the case of the VECM that TFP does not respond positively to the news shock within a horizon of approximately ten quarters, I fix the number of states at $n = 12$ and reestimate the state space model by applying the subspace algorithm.

Figure 8.2 plots the estimated eigenvalues of $(\mathcal{A} - K\mathcal{C})$ in the complex plane. Ten eigenvalues lie close to a circle with a radius of approximately 0.75. Although this is relatively low compared to the value of ω in the theoretical models in the previous chapters, it verifies the findings of Mertens and Ravn (2010), who identify an empirical circle with radius 0.77 in the US postwar data. As regards the anticipation horizon, it should be stressed that only nine eigenvalues are relevant. The reason is that in the case of the polynomial expression $\Theta(L)$ from equation (7.2), the corresponding roots appear in complex conjugate pairs except for even q, where one of these roots is real and equal to $-\omega$. Thus, the positive real eigenvalue lying close to the red circle cannot be taken into account. It follows that the remaining nine eigenvalues indicate an anticipation horizon of ten quarters.

The estimated impulse responses are added to Figure 8.1 (see the solid blue lines). They confirm the findings of Beaudry and Portier (2014a) with a few exceptions. In the case of the TFP response, the adjustment within the first periods, as has also been found for the VECM, is only weakly indicated. The gradual increase (with positive point estimates) starts after approximately two years, approaching very close to the response from the VECM. Likewise, the response of stock prices from the subspace algorithm analysis is qualitatively similar to the VECM finding, but the point estimates of the former surpass the ones of the VECM by around two percentage points at the end of the displayed horizon. There is no

Figure 8.2: ESTIMATED EIGENVALUES OF $(\mathcal{A} - K\mathcal{C})$ IN US DATA
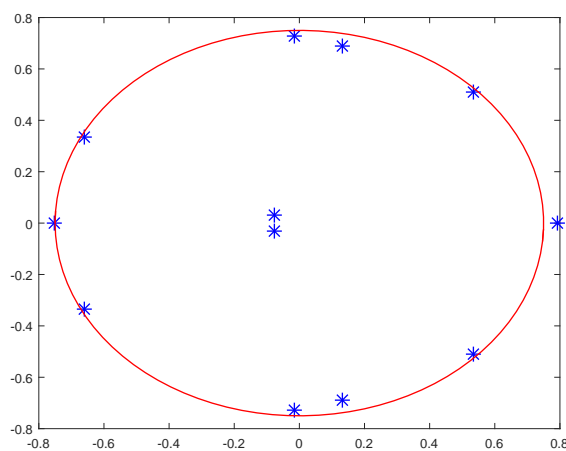
Note: The figure depicts the eigenvalues of the estimated matrix $(\mathcal{A} - K\mathcal{C})$ in the case of the subspace algorithm analysis (in blue). The red line marks a circle with a radius of 0.75.

significant difference between the two responses because the solid blue line is embedded in the confidence interval of the VECM.

The most obvious deviation from the VECM result can be seen in the response of consumption to the news shock. In the subspace algorithm analysis, consumption has virtually no positive response on impact, followed by a steep increase that continues two years later as in case of the VECM. Nevertheless, there is no significant difference between the subspace algorithm analysis and the VECM at the end of the plotted horizon. Note that the near-zero impact effect on consumption does not refute the news shock interpretation completely. Beaudry and Portier (2014a) show that their baseline dynamic general equilibrium model generates a zero impact response of consumption to a technological news shock unless the model is extended by certain properties, e.g., decreasing returns to capital, which allow richer dynamics of the relevant variables.

In terms of the FEVD, the VECM findings are supported for the TFP and stock price variation but only to a lesser extent for consumption (see the fourth column of Table 8.1). For consumption, the dominant role of the news shock in explaining the forecast error variance seems to be a more medium- and long-term phenomenon. The main reason for this is the strong sustainable effect of the third structural shock in the system to which I will turn below.

For completeness, Figure 8.1 (see the dotted black and blue lines) and Table 8.1 (see the columns three and five) are supplemented by the outcomes for the root flipping techniques. The relevant parameters are set as in the Monte Carlo experiment of Chapter 7, i.e., the anticipation rate is fixed at the value implied by my theoretical model, $\omega = 0.9627$, and the relative weight of the news shock, λ_θ, equals two.[6] In general, the results of the root

---

[6] I follow Mertens and Ravn (2010) in using the theoretical value for ω instead of the estimated value for the

flipping procedures do not deviate much from their standard counterparts except for the jagged pattern of the impulse responses. The issue has already been discussed in the Monte Carlo simulation study regarding the VAR models, where the number of lags in the VAR is smaller than the anticipation horizon of the news shock. This is the case here because the VECM is estimated with three lags in differences. The jagged movements also apply to the impulse responses in the subspace algorithm analysis because the VECM, and its number of lags in particular, serve as benchmark for the subspace algorithm to pin down the number of lags and leads used in the procedure. The Monte Carlo simulation exercises in Chapter 7 hint at a direct coherence between the length of the anticipation horizon in the underlying theoretical model and the number of VAR lags suggested by the AIC.[7] Unfortunately, this coherence cannot be maintained in this empirical analysis.

The jagged pattern of the impulse responses might also depend on the fact that the theoretical values of ω and λ_θ used in the root flipping procedure differ from their empirical counterparts. While the subspace algorithm analysis offers a way to estimate a value for ω (as well as for q), the identification of λ_θ is unclear at this stage. A solution to the latter problem would help move from theory-based to (an appropriate method of) empirically based root flipping. However, a further investigation of the aforementioned aspects is left for future work. Instead, I continue with a final discussion related to the third structural shock identified in the system above.

8.2.4   Discussion

The implications of the third structural shock identified above are discussed because, as revealed in the previously presented FEVD, the TFP news shock does not play the superior role in the variation in economic activity during the first quarters. Such a candidate has been introduced in the first part of this thesis: the market rush shock. Therefore, the natural question arises whether the third structural shock identified in the trivariate system comprising TFP, stock prices and consumption reflects a market rush phenomenon.

A first glimpse at the FEVD findings in Table 8.2 related to the third structural shock might lead to a positive answer to this question (at least for the VECM results in the second column of the table).[8] According to the VECM, this shock is replaced by the anticipated TFP shock as the major determinant of the forecast error variance of consumption not before the fourth quarter. According to the subspace algorithm analysis, the dominance of the third shock is maintained for at least the first eight years (see the third column of Table 8.2), thus contradicting the property of the market rush shock as being (only) a very short-run driver

---

root flipping procedure.

[7]Note that the AIC suggests only one lag more than Beaudry and Portier (2014a) have selected for their VECM, i.e., four lags in differences instead of three.

[8]Recall from Chapter 7 that the identification of the third structural shock is not affected by whether the standard procedure or the root flipping method is applied.

Table 8.2: SHARE OF FORECAST ERROR VARIANCE DUE TO 3RD STRUCTURAL SHOCK

		SVECM	Structural CCA	
	Horizon	Standard	Standard [1,3]-restr.	Standard [3,3]-restr.
Consumption	1	0.69	0.86	0.85
	4	0.43	0.90	0.80
	8	0.35	0.80	0.67
	16	0.29	0.70	0.55
	24	0.25	0.61	0.45
	32	0.21	0.53	0.37
	50	0.15	0.38	0.23
	120	0.05	0.15	0.06

Note: The table reports the FEVD share of consumption (bottom panel) due to the third structural shock identified via the various estimation techniques as labeled in each column.

of economic activity. One reason for the difference between the two results could be that the VECM forces the third structural shock to have no long-run effect on any of the variables by construction, whereas in the case of the subspace algorithm analysis, it has only been restricted to have no long-run impact on TFP. Hence, I also conduct the subspace algorithm analysis in combination with a different long-run restriction on the third shock that it has no long-run effect on consumption instead of TFP, i.e., restricting the [3,3]-element of the total impact matrix to zero. As the fourth column of Table 8.2 shows, there is no decisive change in the results (between the [1,3]- and the [3,3]-restriction) with respect to the first years. Consequently, I concentrate on the outcomes of the VECM in order to deepen the discussion about the market rush interpretation.[9]

Figure 8.3 plots the estimated response of consumption to the third structural shock (in black).[10] Two aspects have to be addressed in this context. First, one might argue that the identified shock cannot reflect a market rush, as characterized in Chapters 3 and 4, because consumption exhibits a significant reaction to the shock, which has not been the case in the Beaudry et al. (2011)-type examination of the response of consumption to the market rush shock therein. An explanation for this discrepancy could be that, in the gold rush fever analysis, a bivariate VECM is used, where consumption captures the technology shock and output is driven, over the very short-run, by the market rush shock, whereas TFP encompasses the (surprise) technology shock and consumption represents the activity measure that might be influenced by a kind of a market rush shock in the trivariate system of this chapter. The second aspect is independent of the aforementioned issue and can be seen as

---

[9]Note that there is a mentionable change in the FEVD share of consumption w.r.t. the technological news shock (see the sixth column of Table 8.1), indicating a higher relevance of this type of shock than found in the case of the [1,3]-restriction.

[10]The estimated responses stemming from the subspace algorithm analysis are included for completeness.

Figure 8.3: ESTIMATED RESPONSE OF CONSUMPTION TO 3RD STRUCTURAL SHOCK
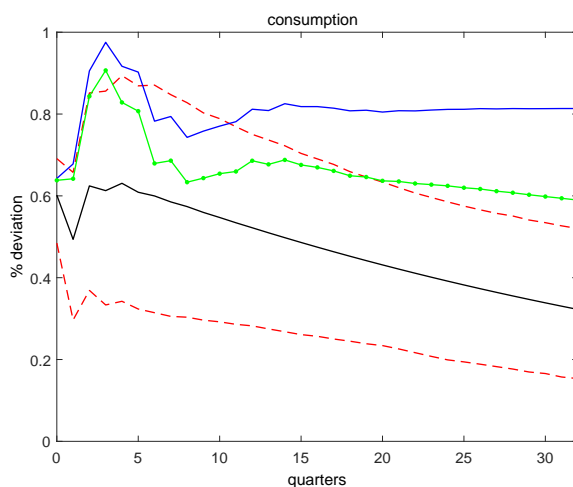
Note: The figure depicts the estimated impulse responses resulting from different estimation approaches. These approaches are standard SVECM (black), standard structural CCA based on the [1,3]-restriction (blue) and standard structural CCA based on the [3,3]-restriction (green). Dashed red lines mark the 16% and 84% quantiles of the distribution of the impulse responses w.r.t. the standard SVECM, obtained by the approach as discussed in Doan (1992).

a stronger argument against the market rush hypothesis with respect to the third structural shock: the estimated response of economic activity (in terms of consumption) displays a very persistent reaction as opposed to the very short-run effect of the market rush shock in Chapter 4. Beaudry et al. (2011) emphasize the nonproductive characteristic of the market rush shock as a very short-run driver, which has to be distinguished from a productive version of this type of shock for which they do not find a relevant role in explaining US business cycles.

Thus, a final point could be raised in the sense that the third structural shock mirrors not the (nonproductive) market rush shock in the spirit of Beaudry et al. (2011) but a productive one. In this case, finding a proper label for that shock can be difficult. Beaudry and Portier (2014a) define nonproductive market rush shocks as those that "create cycles driven by competition of monopoly rents, which are socially inefficient as investment only redistributes rents without having any productive impact." Productive market rushes can arise in models that allow for increasing returns to variety in the goods market, for instance. Beaudry and Portier (2014a) give a corresponding example and state that changes in the number of firms "will play the exact same role as changes in productivity" in this case. As a consequence, news about the opening of new markets and news about future productivity growth can have the same implications, which make them difficult to differentiate in empirical investigations.

Note that there are not only empirical but also theoretical challenges when combining firm dynamics with the concept of news-driven business cycles. Fan and Xu (2014) show that a typical DSGE model, such as the one used by Jaimovich and Rebelo (2009) in the news shock context, cannot generate comovement between stock prices and the number of firms, as can be found in the data, unless the model is extended by an endogenous survival rate

governing the entrance of new firms into the market. However, fully endogenizing the process of firm entry in a business cycle model somehow disposes of the idea of market rushes as introduced by Beaudry et al. (2011), who model market rushes as firm dynamics resulting from an exogenous expansion of the product variety.

8.3   Concluding remarks

This chapter concludes my thesis, which provided a suitable framework to compare the standard VAR/VECM estimator with an alternative tool, subspace algorithm (cointegration) analysis. The framework was based on state space representations that helped reconcile the commonalities and differences between both approaches. Moreover, the usage of state space representations facilitated the illustration of the implications of nonfundamentalness when the information sets of the agents and the econometrician do not match. My analysis relied on a prototypical, theoretical business cycle workhorse model, which is flexible enough to provide intuitive and analytical examples on the one hand and to allow appropriate extensions of the model to offer a comprehensive setting on the other hand. The workhorse model (with its several corresponding specifications) served as the DGP in the Monte Carlo simulation studies conducted in the thesis.

The Monte Carlo simulations showed that subspace algorithm cointegration analysis can be an adequate alternative to the standard VECM for structural estimation. However, I disclosed that it also suffers from lag truncation bias in small samples. While lag truncation bias is a well-known issue, I consider another source of systematic bias in the second part of the thesis wherein I placed special attention to the situation in which nonfundamentalness occurs. In this situation, estimates can be biased due to misalignment between the agent and econometrician's information sets. The Monte Carlo experiments revealed that a convenient adjustment of the estimation procedures, i.e., the theory-based root flipping technique, improves the performance of the subspace algorithm analysis and the VAR/VECM.

Nevertheless, two warnings have to be stressed in that context. First, the systematic errors and the corresponding success of the modified estimator depend on the underlying DGP. Second, one could suspect that (at least in the case of a setting in which news about future technology growth are apparent) nonfundamentalness does not appear to be the most influential determinant of the estimation results. Other factors, such as small sample uncertainty or the general accuracy of the estimation technique, are more decisive. Hence, there might not be an "either/or proposition" regarding the issue of considering nonfundamentalness, but this proposition might be applicable to the basic estimation method, VAR or subspace algorithm analysis. My simulation results highlighted the relatively good performance of the latter but also demonstrated that this finding, in turn, is driven by the underlying DGP.

Finally, the empirical investigations using US macro data led to the conclusion that news

shocks can be seen as important triggers of macroeconomic fluctuation over the entire business cycle horizon. Over the very short-term, market rushes, as introduced by Beaudry et al. (2011) and interpreted as one kind of news shock, explain a sizable fraction of US economic activity. At horizons beyond the first year, technological news shocks represent a dominant driver of US business cycles when a VECM is used for estimation, whereas such shocks are more a medium- and long-run determinant of fluctuations according to the subspace algorithm analysis. The examination in this chapter has also indicated that there is an open question in terms of how to address both types of shocks, market rushes and technological news, in a common environment. This task appears challenging because the general concept of market rushes can imply outcomes similar to those of technological news shocks. Thus, one has to find a proper theoretical and empirical framework to obtain a clear distinction between them, which guarantees unambiguous identification in the data. In this context, the final discussion in the last subsection laid a foundation for future research.

Chapter A

Appendix

Appendix to Chapter 2

Basic remarks

In order to show how to compute the roots of the characteristic (and/or lag) polynomials, I begin with some definitions that can be found in Lütkepohl (1996), page 12:

Definition 3. *The equation $det\left[\lambda I_n - A\right] = 0$ is the **characteristic equation** of the $(n \times n)$ matrix A.*

Definition 4. *The polynomial in λ given by $det\left[\lambda I_n - A\right]$ is the **characteristic polynomial** of the $(n \times n)$ matrix A.*

Definition 5. *A number λ for which $p\left(\lambda\right) = p_0 + p_1\lambda + ... + p_n\lambda^n = 0$ is called a **root of the polynomial** $p\left(\cdot\right)$.*

Based on these definitions, I use the terms characteristic equation and characteristic polynomial to refer to the expressions that include $z \in \mathbb{C}$. Note that the root(s) of the characteristic polynomial above equal/s the eigenvalue(s) of A. Moreover, I call the expression that includes the lag operator L the (VAR or MA) lag polynomial, e.g., $det\left[I_n - AL\right]$. Hence, $\lambda = \frac{1}{L}$ is called a root of the (VAR or MA) lag polynomial. This root is the reciprocal of the corresponding root of the characteristic polynomial $det\left[\lambda I_m - A\right]$. Thus, the root of the (VAR or MA) lag polynomial equals the inverse of the corresponding eigenvalue of A.

According to Lütkepohl (2005), I use the terms VAR or MA operator to refer to equations such as $A\left(L\right)y_t = M\left(L\right)a_t$, where $A\left(L\right)$ denotes the VAR operator, and $M\left(L\right)$ represents the MA operator. $det\left[A\left(L\right)\right]$ is called the VAR lag polynomial and $det\left[M\left(L\right)\right]$ the MA lag polynomial. Furthermore, I apply a rule related to the algebra of partitioned matrices (see Lütkepohl (1996) or Hansen and Sargent (2014)) as follows.

Assume a and d as nonsingular matrices and appropriate dimensions of all four matrices to write

$$det\,[a]\,det\left[d - ca^{-1}b\right] \;=\; det\,[d]\,det\left[a - bd^{-1}c\right]\,.$$

Another rule is given by the following identity, where the assumptions with respect to the four matrices still hold:

$$\left(a - bd^{-1}c\right)^{-1} \;=\; a^{-1} + a^{-1}b\left(d - ca^{-1}b\right)^{-1}ca^{-1}\,.$$

Computation of roots

Consider the vector MA representation in equation (2.5) in the main text

$$y_t \;=\; \left[\mathcal{D} + \mathcal{C}\left(L^{-1} - \mathcal{A}\right)^{-1}\mathcal{B}\right]w_t\,. \tag{A.1}$$

I apply the first rule from above to calculate the roots of the characteristic polynomial $det[\mathcal{D} + \mathcal{C}(zI_n - \mathcal{A})^{-1}\mathcal{B}]$. Setting

$$
\begin{aligned}
a &= zI_n - \mathcal{A}\,,\\
b &= \mathcal{B}\,,\\
c &= -\mathcal{C}\,,\\
d &= \mathcal{D}\,,
\end{aligned}
$$

it follows that

$$
\begin{aligned}
det\left[\mathcal{D} + \mathcal{C}\left(zI_n - \mathcal{A}\right)^{-1}\mathcal{B}\right] &= \frac{det\left[zI_n - \mathcal{A} + \mathcal{B}\mathcal{D}^{-1}\mathcal{C}\right]}{det\left[zI_n - \mathcal{A}\right]}\\
&= \frac{det\left[zI_n - (\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})\right]}{det\left[zI_n - \mathcal{A}\right]}\,.
\end{aligned}
$$

Because the roots of the characteristic polynomial $det\left[zI_n - (\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})\right]$ are the eigenvalues of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$, these eigenvalues are the roots of the characteristic polynomial $det\left[\mathcal{D} + \mathcal{C}\left(zI_n - \mathcal{A}\right)^{-1}\mathcal{B}\right]$.

Next, using the lag operator L by setting $z = L^{-1}$, I obtain the MA operator $[\mathcal{D} + \mathcal{C}(L^{-1} - \mathcal{A})^{-1}\mathcal{B}]$ as given in equation (A.1) or more conveniently as

$$M\,(L) \;=\; \mathcal{D} + \mathcal{C}\left(I_n - \mathcal{A}L\right)^{-1}\mathcal{B}L\,.$$

Hence, the roots of the MA lag polynomial $det\,[M\,(L)]$ equal the reciprocals of the corresponding eigenvalues of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$.

Derivation of Riccati difference equation

Starting from equation (2.7), I derive the Riccati difference equation (2.9). In substituting y_t by the observation equation of the FIA model, equation (2.7) can be written as

$$\widehat{x}_t \;=\; (\mathcal{A} - K_t\mathcal{C})\,\widehat{x}_{t-1} + K_t\,(\mathcal{C}x_{t-1} + \mathcal{D}w_t)\,. \tag{A.2}$$

Next, I subtract this equation from the state equation of the FIA model to obtain

$$x_t - \widehat{x}_t \;=\; (\mathcal{A} - K_t\mathcal{C})\,(x_{t-1} - \widehat{x}_{t-1}) + (\mathcal{B} - K_t\mathcal{D})\,w_t\,. \tag{A.3}$$

Multiplying both sides by their transposes and using the expectation operator, I formulate

$$\begin{aligned}
E\left\{(x_t - \widehat{x}_t)\,(x_t - \widehat{x}_t)'\right\} \;=\;\; & E\left\{(\mathcal{A} - K_t\mathcal{C})\,(x_{t-1} - \widehat{x}_{t-1})\,(x_{t-1} - \widehat{x}_{t-1})'\,(\mathcal{A} - K_t\mathcal{C})' \right. \\
& \left. + (\mathcal{B} - K_t\mathcal{D})\,w_t w_t'\,(\mathcal{B} - K_t\mathcal{D})'\right\}\,.
\end{aligned} \tag{A.4}$$

Because $E\left(w_t w_t'\right) = I_k$ and $\Sigma_t \equiv E\left\{(x_t - \widehat{x}_t)\,(x_t - \widehat{x}_t)'\right\}$, I write this equation as

$$\Sigma_t \;=\; (\mathcal{A} - K_t\mathcal{C})\,\Sigma_{t-1}\,(\mathcal{A} - K_t\mathcal{C})' + (\mathcal{B} - K_t\mathcal{D})\,(\mathcal{B} - K_t\mathcal{D})'\,. \tag{A.5}$$

Next, I factor out K_t on the right-hand side of the equation:

$$\begin{aligned}
\Sigma_t \;=\;\; & \mathcal{A}\Sigma_{t-1}\mathcal{A}' + \mathcal{B}\mathcal{B}' + K_t\,(\mathcal{C}\Sigma_{t-1}\mathcal{C}' + \mathcal{D}\mathcal{D}')\,K_t' \\
& - K_t\,(\mathcal{C}\Sigma_{t-1}\mathcal{A}' + \mathcal{D}\mathcal{B}') - (\mathcal{A}\Sigma_{t-1}\mathcal{C}' + \mathcal{B}\mathcal{D}')\,K_t'\,.
\end{aligned} \tag{A.6}$$

Substitution of K_t by equation (2.8) and some rearrangements yield equation (2.9) in the main text:

$$\begin{aligned}
\Sigma_t \;=\;\; & \mathcal{A}\Sigma_{t-1}\mathcal{A}' + \mathcal{B}\mathcal{B}' \\
& + (\mathcal{A}\Sigma_{t-1}\mathcal{C}' + \mathcal{B}\mathcal{D}')\,(\mathcal{C}\Sigma_{t-1}\mathcal{C}' + \mathcal{D}\mathcal{D}')^{-1}\,(\mathcal{A}\Sigma_{t-1}\mathcal{C}' + \mathcal{B}\mathcal{D}')' \\
& - (\mathcal{A}\Sigma_{t-1}\mathcal{C}' + \mathcal{B}\mathcal{D}')\,(\mathcal{C}\Sigma_{t-1}\mathcal{C}' + \mathcal{D}\mathcal{D}')^{-1}\,(\mathcal{C}\Sigma_{t-1}\mathcal{A}' + \mathcal{D}\mathcal{B}') \\
& - (\mathcal{A}\Sigma_{t-1}\mathcal{C}' + \mathcal{B}\mathcal{D}')\,(\mathcal{C}\Sigma_{t-1}\mathcal{C}' + \mathcal{D}\mathcal{D}')^{-1}\,(\mathcal{A}\Sigma_{t-1}\mathcal{C}' + \mathcal{B}\mathcal{D}')'
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \Sigma_t \;=\;\; & \mathcal{A}\Sigma_{t-1}\mathcal{A}' + \mathcal{B}\mathcal{B}' \\
& - (\mathcal{A}\Sigma_{t-1}\mathcal{C}' + \mathcal{B}\mathcal{D}')\,(\mathcal{C}\Sigma_{t-1}\mathcal{C}' + \mathcal{D}\mathcal{D}')^{-1}\,(\mathcal{A}\Sigma_{t-1}\mathcal{C}' + \mathcal{B}\mathcal{D}')'\,.
\end{aligned} \tag{A.7}$$

Inverting the MA operator of the LIE model representation

The vector MA representation is given in equation (2.18), where L denotes the lag operator.

Applying the second algebraic rule from above by setting

$$
\begin{aligned}
a &= I_k \,, \\
b &= -\mathcal{C} \,, \\
c &= KL \,, \\
d &= I_n - \mathcal{A}L \,,
\end{aligned}
$$

it follows that the inverse of the MA operator can be written as

$$
\left[I_k + \mathcal{C} \left(I_n - \mathcal{A} \right)^{-1} KL \right]^{-1} \;=\; I_k - \mathcal{C} \left(I_n - \left(\mathcal{A} - K\mathcal{C} \right) L \right)^{-1} KL \,.
$$

Using this identity, the VAR representation given in equation (2.19) can be derived.

Appendix to Chapter 3

Jordan decomposition

In the case of a Jordan decomposition of matrix T (that is supposed to have q distinct eigenvalues $\lambda_1, ..., \lambda_q$) such that $T = P\Lambda P^{-1}$, Λ has block-diagonal structure (i.e. Jordan form),

$$\Lambda \;\; = \;\; \begin{bmatrix} \Lambda_1 & & 0 \\ & \ddots & \\ 0 & & \Lambda_s \end{bmatrix},$$

with (Jordan) blocks defined as

$$\Lambda_i \;\; = \;\; \begin{bmatrix} \lambda_{q_i} & 1 & 0 & \cdots & 0 \\ 0 & \lambda_{q_i} & 1 & \ddots & 0 \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & \cdots & \cdots & \lambda_{q_i} \end{bmatrix}$$

for $i = 1, ...s \geq q$ and $\{q_1, ...q_s\} = \{1, ..., q\}$ (see, e.g., Lütkepohl (1996)). It can be the case that more than one Jordan block is associated with a particular eigenvalue. Then, the corresponding Jordan blocks form a Jordan segment (with the additional zeros off the block diagonal).

Illustrative examples

I provide some illustrative examples to show the connection between the integration order of a (state) process and the multiplicity of the corresponding unit root. Algebraic multiplicity is the multiplicity of an eigenvalue that solves the characteristic polynomial. Geometric multiplicity is the number of Jordan blocks associated with an eigenvalue. I draw on an example by Bauer and Wagner (2003) and present some related variations.

The first system includes only random walks, the second system contains an $I(2)$ process, and the third system involves a process integrated of order three. With respect to the description of a Jordan decomposition above, the geometric and algebraic multiplicity can easily be detected.

System 1:

$$
\begin{pmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} w_t \tag{A.8}
$$

$$
\begin{aligned}
\Longrightarrow \triangle x_{1,t} &= w_t, \\
\triangle x_{2,t} &= w_t, \\
\triangle x_{3,t} &= w_t.
\end{aligned}
$$

System 2:

$$
\begin{pmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} w_t \tag{A.9}
$$

$$
\begin{aligned}
\Longrightarrow \triangle x_{1,t} &= x_{2,t-1}, \\
\triangle x_{2,t} &= w_t, \\
\triangle x_{3,t} &= w_t.
\end{aligned}
$$

System 3:

$$
\begin{pmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} w_t \tag{A.10}
$$

$$
\begin{aligned}
\Longrightarrow \triangle x_{1,t} &= x_{2,t-1}, \\
\triangle x_{2,t} &= x_{3,t-1}, \\
\triangle x_{3,t} &= w_t.
\end{aligned}
$$

In all three systems, the algebraic multiplicity is equal to three. In System 1, the geometric multiplicity is also equal to three. All three processes are $I(1)$. In System 2, the geometric multiplicity is equal to two, $x_{1,t}$ is $I(2)$, and the other two processes are $I(1)$. In System 3, the geometric multiplicity equals one, $x_{1,t}$ is $I(3)$, $x_{2,t}$ is $I(2)$, and $x_{3,t}$ is $I(1)$. Note that only

System 1 is minimal.

Remarks on MA roots in the nonstationary case

Note that I refer to nonstationary cases in which I consider cointegrated, observable variables driven by common stochastic trends. In my framework, perpetual growth arises because of external $I(1)$ impulse processes, i.e., there is no intrinsic propagation mechanism that generates (co)integration properties of the endogenous variables. I model these external $I(1)$ impulse processes as conventional random walks.

This section is decomposed into three steps. In the first step, I present an alternative way to derive the vector MA representation in first differences associated with the FIA model when facing cointegration. Thereby, I use the rule for a square $(m \times m)$ matrix A, as given in Lütkepohl (1996) (on page 27),

$$A \cdot adj\,[A] \;=\; adj\,[A] \cdot A = det\,[A] \cdot I_m \,, \tag{A.11}$$

where $det\,[\bullet]$ and $adj\,[\bullet]$ denote the determinant and adjoint, respectively. The first step allows me to detect the origin of the roots of the corresponding MA lag polynomial. In the second step, I exemplify that the number of endogenous states determines the number of the aforementioned roots. The final step demonstrates the linkage between these roots and the eigenvalues of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$ by taking into account the findings of the previous steps.

My explanations refer to the theoretical model solution, i.e., I proceed in terms of the FIA model. Because the *invertibility condition* is supposed to hold, the conclusions of this section can be carried over to the LIE model. In the main text, I show how to derive a vector MA representation in first differences for the LIE model in the case of cointegration (see equation (3.29)). I can proceed similarly for the FIA model to obtain the following vector MA representation

$$\triangle y_t = \underbrace{\left[\mathcal{D} + \left(\tilde{\mathcal{A}}(L)^{-1} \mathcal{C} \breve{\mathcal{A}}(L) \mathcal{B} - \mathcal{D} \right) L \right]}_{= \widetilde{M}(L)} w_t \,, \tag{A.12}$$

where the roots of the MA lag polynomial $det\left[\widetilde{M}(L)\right]$ are the reciprocals of the nonzero eigenvalues of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$ and a unit root.

An equivalent expression can be derived from an alternative representation of the FIA model, where the state vector is partitioned into endogenous and exogenous states:

$$y_t \;=\; \mathcal{R}s_{t-1} + \mathcal{S}z_t \,, \tag{A.13}$$

$$s_t \;=\; \mathcal{P}s_{t-1} + \mathcal{Q}z_t \,, \tag{A.14}$$

$$z_t = \mathcal{N} z_{t-1} + w_t. \tag{A.15}$$

I omit possible constants to simplify the notation. s_t is the $(n_s \times 1)$ vector of endogenous states. z_t is the $(n_z \times 1)$ vector of exogenous states. y_t and w_t are the same as before. The coefficient (and further) matrices have the corresponding dimensions. Suppose that all exogenous states evolve according to standard AR(1) processes. Because I consider some of the exogenous states as random walks, some of the eigenvalues of $\mathcal{N}$ are equal to one, whereas all the remaining eigenvalues are smaller than one (in modulus). Furthermore, the number of exogenous states equals the number of shocks and thus the number of observables, i.e., $n_z = k$. For the following explanations, I assume that $\mathcal{P}$ and $\mathcal{S}$ are nonsingular. As it will turn out, $\mathcal{S}$ is equivalent to the system matrix $\mathcal{D}$, so *Assumption 1* covers the invertibility property of $\mathcal{S}$.

Solve equation (A.14) for s_{t-1} and use backward substitution:

$$s_{t-1} = \left[\mathcal{P}^{-1} \left(I_{n_s} - \mathcal{P} L \right)^{-1} \mathcal{Q} - \mathcal{P}^{-1} \mathcal{Q} \right] z_t. \tag{A.16}$$

Insert equation (A.16) into equation (A.13) and rearrange:

$$
\begin{aligned}
y_t &= \left[\mathcal{R} \mathcal{P}^{-1} \left(I_{n_s} - \mathcal{P} L \right)^{-1} \mathcal{Q} - \mathcal{R} \mathcal{P}^{-1} \mathcal{Q} + \mathcal{S} \right] z_t \\
&= \left[\mathcal{S} - \mathcal{R} \mathcal{P}^{-1} \left(I_{n_s} - \left(I_{n_s} - \mathcal{P} L \right)^{-1} \right) \mathcal{Q} \right] z_t \\
&= \left[I_k - \mathcal{R} \mathcal{P}^{-1} \left(I_{n_s} - \left(I_{n_s} - \mathcal{P} L \right)^{-1} \right) \mathcal{Q} \mathcal{S}^{-1} \right] \mathcal{S} z_t \\
&= \left[I_k - \mathcal{R} \mathcal{P}^{-1} \left(-\mathcal{P} L \right) \left(I_{n_s} - \mathcal{P} L \right)^{-1} \mathcal{Q} \mathcal{S}^{-1} \right] \mathcal{S} z_t \\
&= \left[I_k + \mathcal{R} \left(I_{n_s} - \mathcal{P} L \right)^{-1} \mathcal{Q} \mathcal{S}^{-1} L \right] \mathcal{S} z_t. \tag{A.17}
\end{aligned}
$$

Applying the rule as given in equation (A.11), I can write equation (A.15) as

$$det \left[I_{n_z} - \mathcal{N} L \right] z_t = adj \left[I_{n_z} - \mathcal{N} L \right] w_t. \tag{A.18}$$

Multiply both sides of equation (A.17) by (the scalar polynomial) $det \left[I_{n_z} - \mathcal{N} L \right]$ and substitute equation (A.18):

$$det \left[I_{n_z} - \mathcal{N} L \right] y_t = \left[I_k + \mathcal{R} \left(I_{n_s} - \mathcal{P} L \right)^{-1} \mathcal{Q} \mathcal{S}^{-1} L \right] \mathcal{S} \, adj \left[I_{n_z} - \mathcal{N} L \right] w_t. \tag{A.19}$$

If $\mathcal{N}$ has c eigenvalues equal to one, then $det \left[I_{n_z} - \mathcal{N} L \right]$ and $adj \left[I_{n_z} - \mathcal{N} L \right]$ have $c - 1$ common factors $(1 - L)$. Thus, I can formulate

$$
\begin{aligned}
det \left[I_{n_z} - \mathcal{N} L \right] &= (1 - L)^c \, \widetilde{N} (L), \\
adj \left[I_{n_z} - \mathcal{N} L \right] &= (1 - L)^{c-1} \, \breve{N} (L),
\end{aligned}
$$

where $\widetilde{N} (L)$ denotes the scalar polynomial determined by the eigenvalues of $\mathcal{N}$ less than

one, and $\check{\mathcal{N}}(L)$ represents a matrix polynomial that remains after eliminating the common factors $(1 - L)$ in $adj \left[I_{n_z} - \mathcal{N}L \right]$. Note that there is still one unit root left in $det \left[\check{\mathcal{N}}(L) \right]$.

Using the expressions from above, I can write equation (A.19) as

$$\widetilde{N}(L) \Delta y_t \;\; = \;\; \left[I_k + \mathcal{R}(I_{n_s} - \mathcal{P}L)^{-1} \mathcal{Q}\mathcal{S}^{-1}L \right] \mathcal{S}\check{\mathcal{N}}(L) w_t. \tag{A.20}$$

Multiplication of both sides of equation (A.20) by the inverse of $\widetilde{N}(L)$ yields the vector MA representation in first differences

$$\triangle y_t = \left[I_k + \mathcal{R}(I_{n_s} - \mathcal{P}L)^{-1} \mathcal{Q}\mathcal{S}^{-1}L \right] \mathcal{S}\widetilde{N}(L)^{-1} \check{\mathcal{N}}(L) w_t, \tag{A.21}$$

which is equivalent to equation (A.12), i.e., $\widetilde{M}(L) = [I_k + \mathcal{R}(I_{n_s} - \mathcal{P}L)^{-1}\mathcal{Q}\mathcal{S}^{-1}L]\mathcal{S}\widetilde{N}(L)^{-1}\check{\mathcal{N}}(L)$. The roots of $det \left[\widetilde{M}(L) \right]$ can be computed as follows. First, notice that $det \left[\widetilde{N}(L) \right]$ and $det \left[\check{\mathcal{N}}(L) \right]$ have the same roots except the unit root in $det \left[\check{\mathcal{N}}(L) \right]$. Hence, $det[\widetilde{N}(L)^{-1}\check{\mathcal{N}}(L)]$ and thus $det \left[\widetilde{M}(L) \right]$ have one unit root that indicates the cointegration relationship among the variables in y_t. Second, the remaining roots of $det \left[\widetilde{M}(L) \right]$ are given as the roots of $det \left[I_k + \mathcal{R}(I_{n_s} - \mathcal{P}L)^{-1} \mathcal{Q}\mathcal{S}^{-1}L \right]$. These roots can be calculated analogously to the steps presented in the appendix to Chapter 2. Hence, the roots of the characteristic polynomial $det \left[I_k + \mathcal{R}(zI_{n_s} - \mathcal{P})^{-1} \mathcal{Q}\mathcal{S}^{-1} \right]$ are the eigenvalues of $(\mathcal{P} - \mathcal{Q}\mathcal{S}^{-1}\mathcal{R})$, and the reciprocals of the (nonzero) eigenvalues of $(\mathcal{P} - \mathcal{Q}\mathcal{S}^{-1}\mathcal{R})$ are the additional roots of the MA lag polynomial $det \left[\widetilde{M}(L) \right]$. Furthermore, this means that the number of endogenous states determines the number of non-unit roots of $det \left[\widetilde{M}(L) \right]$.

In the following, I show two examples supporting the claim that the number of non-unit roots of $det \left[\widetilde{M}(L) \right]$ is equal to the number of endogenous states that are not in the set of observables. The examples should be considered only as tentative rather than a general proof. These examples illustrate that my claim holds for the particular models used in the chapter.

Example 1:

Consider the simple RBC model as specified in Subsection (3.5.2). I can formulate the FIA model in the state space form of equations (A.13)-(A.15) as

$$\begin{pmatrix} \widetilde{C}_t \\ \widetilde{Y}_t \end{pmatrix} \;\; = \;\; \underbrace{\begin{pmatrix} \phi_{CK} \\ \phi_{YK} \end{pmatrix}}_{= \, \mathcal{R}} \widetilde{K}_{t-1} + \underbrace{\begin{pmatrix} \phi_{CA} & \phi_{C\psi} \\ \phi_{YA} & \phi_{Y\psi} \end{pmatrix}}_{= \, \mathcal{S}} \begin{pmatrix} log\,(A_t) \\ \widetilde{\psi}_t \end{pmatrix}, \tag{A.22}$$

$$\widetilde{K}_t \;=\; \underbrace{\phi_{KK}}_{=\,\mathcal{P}} \widetilde{K}_{t-1} + \underbrace{\left(\begin{array}{cc} \phi_{KA} & \phi_{K\psi} \end{array}\right)}_{=\,\mathcal{Q}} \left(\begin{array}{c} log\left(A_t\right) \\ \widetilde{\psi}_t \end{array}\right), \qquad (\text{A.23})$$

and

$$\left(\begin{array}{c} log\left(A_t\right) \\ \widetilde{\psi}_t \end{array}\right) \;=\; \underbrace{\left(\begin{array}{cc} 1 & 0 \\ 0 & \rho_\psi \end{array}\right)}_{=\,\mathcal{N}} \left(\begin{array}{c} log\left(A_{t-1}\right) \\ \widetilde{\psi}_{t-1} \end{array}\right) + \left(\begin{array}{c} w_t^A \\ w_t^\psi \end{array}\right). \qquad (\text{A.24})$$

Using the definitions of the system matrices, I obtain

$$\begin{aligned} \mathcal{P} - \mathcal{Q}\mathcal{S}^{-1}\mathcal{R} \;=\; & [\phi_{CK}\left(\phi_{K\psi}\phi_{YA} - \phi_{KA}\phi_{Y\psi}\right) \\ & + \phi_{CA}\left(\phi_{KK}\phi_{Y\psi} - \phi_{K\psi}\phi_{YK}\right) \\ & + \phi_{C\psi}\left(\phi_{KA}\phi_{YK} - \phi_{KK}\phi_{YA}\right)] \\ & \times \left(\phi_{CA}\phi_{Y\psi} - \phi_{C\psi}\phi_{YA}\right)^{-1}. \end{aligned}$$

If capital belongs to the vector of observables instead of output, it is easy to see that, in this case, this expression would be zero. Following Ravenna (2007), it is possible to show that the model solution can be expressed as a VAR(2) model in this case. For this purpose, define the vector $\overline{y}_t \equiv \left(\begin{array}{cc} \widetilde{C}_t & \widetilde{K}_t \end{array}\right)'$ and combine the system equations to obtain

$$\overline{y}_t \;=\; A\overline{y}_{t-1} + Bz_t, \qquad (\text{A.25})$$

where $A = \left(\begin{array}{cc} 0_{2\times 1} & \mathcal{R} \end{array}\right)$ and $B = \mathcal{S}$. Given that B has full rank, it is possible to solve equation (A.25) for z_t and substitute it into equation (A.24). Finally, the resulting expression is plugged into equation (A.25), so I end up with a VAR(2) representation

$$\overline{y}_t \;=\; \left(A + B\mathcal{N}B^{-1}\right)\overline{y}_{t-1} - B\mathcal{N}B^{-1}A\overline{y}_{t-2} + Bw_t. \qquad (\text{A.26})$$

Example 2:

Suppose a model with two observables and two endogenous states so that all system matrices have a 2×2 dimension. Moreover, assume that one of the endogenous states (which is ordered at first position in the state vector, for instance) is included in the vector of observ-

ables. Then, the system matrices can be partitioned as

$$
\mathcal{P} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \quad , \quad \mathcal{Q} = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} ,
$$

$$
\mathcal{R} = \begin{pmatrix} r_{11} & r_{12} \\ p_{11} & p_{12} \end{pmatrix} \quad , \quad \mathcal{S} = \begin{pmatrix} s_{11} & s_{12} \\ q_{11} & q_{12} \end{pmatrix} ,
$$

$$
\Longrightarrow \mathcal{P} - \mathcal{Q}\mathcal{S}^{-1}\mathcal{R} = \begin{pmatrix} 0 & 0 \\ \begin{matrix}[p_{11}(q_{22}s_{11} - q_{21}s_{12}) \\ +p_{21}(q_{11}s_{12} - q_{12}s_{11}) \\ +r_{11}(q_{12}q_{21} - q_{11}q_{22})] \\ \times (q_{11}s_{12} - q_{12}s_{11})^{-1} \end{matrix} & \begin{matrix}[p_{12}(q_{22}s_{11} - q_{21}s_{12}) \\ +p_{22}(q_{11}s_{12} - q_{12}s_{11}) \\ +r_{12}(q_{12}q_{21} - q_{11}q_{22})] \\ \times (q_{11}s_{12} - q_{12}s_{11})^{-1} \end{matrix} \end{pmatrix} .
$$

Simple calculation of $(\mathcal{P} - \mathcal{Q}\mathcal{S}^{-1}\mathcal{R})$ reveals that its first row has only zero entries, i.e., one eigenvalue of $(\mathcal{P} - \mathcal{Q}\mathcal{S}^{-1}\mathcal{R})$ is equal to zero.[1] If both endogenous states form the set of observables, then $\mathcal{P}$ is equal to $\mathcal{R}$, and $\mathcal{Q}$ is equal to $\mathcal{S}$. Example 2 could easily be extended to a higher-dimensional case. However, because this would complicate the illustrative goal, I stop at this stage.

Finally, I point to the linkage between the eigenvalues of $(\mathcal{P} - \mathcal{Q}\mathcal{S}^{-1}\mathcal{R})$ and $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$. It is convenient to start with the state space representation in equations (A.13)-(A.15) again and derive the FIA model representation given in the main text. This can be done by stacking the endogenous and exogenous states into one vector:

$$
\begin{pmatrix} s_t \\ z_t \end{pmatrix} = \underbrace{\begin{pmatrix} \mathcal{P} & \mathcal{Q}\mathcal{N} \\ 0_{n_z \times n_s} & \mathcal{N} \end{pmatrix}}_{=\,\mathcal{A}} \begin{pmatrix} s_{t-1} \\ z_{t-1} \end{pmatrix} + \underbrace{\begin{pmatrix} \mathcal{Q} \\ I_{n_z} \end{pmatrix}}_{=\,\mathcal{B}} w_t , \tag{A.27}
$$

$$
y_t = \underbrace{\begin{pmatrix} \mathcal{R} & \mathcal{S}\mathcal{N} \end{pmatrix}}_{=\,\mathcal{C}} \begin{pmatrix} s_{t-1} \\ z_{t-1} \end{pmatrix} + \underbrace{\mathcal{S}}_{=\,\mathcal{D}} w_t . \tag{A.28}
$$

---

[1] Note that this result does not require the number of endogenous states to be equal to two.

Using the definitions, I can compute

$$\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C} \;=\; \begin{pmatrix} \mathcal{P} - \mathcal{Q}\mathcal{S}^{-1}\mathcal{R} & 0_{n_s \times n_z} \\ -\mathcal{S}^{-1}\mathcal{R} & 0_{n_z \times n_z} \end{pmatrix} .$$

It is straightforward to show that the eigenvalues of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$ are given as the eigenvalues of $(\mathcal{P} - \mathcal{Q}\mathcal{S}^{-1}\mathcal{R})$ (plus n_z zero eigenvalues).

Let $M = \mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C}$, $M_{11} = \mathcal{P} - \mathcal{Q}\mathcal{S}^{-1}\mathcal{R}$ and $M_{21} = -\mathcal{S}^{-1}\mathcal{R}$. If λ is an eigenvalue, it solves the characteristic equation

$$det\left[\lambda I_{n_s} - M_{11}\right] \;=\; 0.$$

If $\overline{\lambda}$ is an eigenvalue of M, it solves

$$det\left[\overline{\lambda}I_{n_s+n_z} - M\right] = 0.$$

I can write this equation, using the blockwise notation of M, as

$$det\left[\begin{pmatrix} \overline{\lambda}I_{n_s+n_z} - M_{11} & 0_{n_s \times n_z} \\ M_{21} & \overline{\lambda}I_{n_z} \end{pmatrix}\right] = 0$$

Applying the rule for determinants of triangular partitioned matrices (see Lütkepohl (1996)), I formulate

$$det\left[\overline{\lambda}I_{n_s} - M_{11}\right] det\left[\overline{\lambda}I_{n_z}\right] = 0.$$

Hence, if λ (or $\overline{\lambda}$) is an eigenvalue of the submatrix M_{11}, it is also an eigenvalue of M, i.e., an eigenvalue of $(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C})$.

Appendix to Chapter 5

Equivalent representations of the matrix Riccati equation

I show that equation (5.7) is equivalent to equation (2.12). Note that Σ is symmetric. It thus holds that

$$\left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1} = \left(\mathcal{D}\mathcal{D}'\right)^{-1}\left(I_k - \mathcal{C}\Sigma\mathcal{C}'\left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}\right)$$

and

$$\left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1} = \left(I_k - \left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\Sigma\mathcal{C}'\right)\left(\mathcal{D}\mathcal{D}'\right)^{-1}.$$

I use these identities in the following calculations.

Starting from equation (5.7) and using the definitions of $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{Q}}$ in the main text yields

$$\begin{aligned}
\Sigma = &\left(\mathcal{A} - \mathcal{B}\mathcal{D}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\right)\Sigma\left(\mathcal{A} - \mathcal{B}\mathcal{D}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\right)' \\
&+ \mathcal{B}\mathcal{B}' - \mathcal{B}\mathcal{D}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{D}\mathcal{B}' \\
&- \left(\mathcal{A} - \mathcal{B}\mathcal{D}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\right)\Sigma\mathcal{C}'\left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\Sigma\left(\mathcal{A} - \mathcal{B}\mathcal{D}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\right)'. \quad \text{(A.29)}
\end{aligned}$$

Next, I document some computational steps to illustrate the equivalence to equation (2.12):

$$\begin{aligned}
\Sigma = \ &\mathcal{A}\Sigma\mathcal{A}' + \mathcal{B}\mathcal{B}' - \mathcal{A}\Sigma\mathcal{C}'\left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\Sigma\mathcal{A}' \\
&- \mathcal{A}\Sigma\mathcal{C}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{D}\mathcal{B}' + \mathcal{A}\Sigma\mathcal{C}'\left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\Sigma\mathcal{C}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{D}\mathcal{B}' \\
&- \mathcal{B}\mathcal{D}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{D}\mathcal{B}' + \mathcal{B}\mathcal{D}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\Sigma\mathcal{C}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{D}\mathcal{B}' \\
&- \mathcal{B}\mathcal{D}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\Sigma\mathcal{C}'\left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\Sigma\mathcal{C}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{D}\mathcal{B}' \\
&- \mathcal{B}\mathcal{D}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\Sigma\mathcal{A}' + \mathcal{B}\mathcal{D}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\Sigma\mathcal{C}'\left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\Sigma\mathcal{A}'
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \Sigma = \ &\mathcal{A}\Sigma\mathcal{A}' + \mathcal{B}\mathcal{B}' - \mathcal{A}\Sigma\mathcal{C}'\left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\Sigma\mathcal{A}' \\
&- \mathcal{A}\Sigma\mathcal{C}'\left(I_k - \left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\Sigma\mathcal{C}'\right)\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{D}\mathcal{B}' \\
&- \mathcal{B}\mathcal{D}'\left(I_k - \left(\mathcal{D}\mathcal{D}'\right)^{-1}\left(I_k - \mathcal{C}\Sigma\mathcal{C}'\left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}\right)\mathcal{C}\Sigma\mathcal{C}'\right)\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{D}\mathcal{B}' \\
&- \mathcal{B}\mathcal{D}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\left(I_k - \mathcal{C}\Sigma\mathcal{C}'\left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}\right)\mathcal{C}\Sigma\mathcal{A}'
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \Sigma = \ &\mathcal{A}\Sigma\mathcal{A}' + \mathcal{B}\mathcal{B}' - \mathcal{A}\Sigma\mathcal{C}'\left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C}\Sigma\mathcal{A}' \\
&- \mathcal{A}\Sigma\mathcal{C}'\left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{D}\mathcal{B}' \\
&- \mathcal{B}\mathcal{D}'\left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{D}\mathcal{B}'
\end{aligned}$$

$$-\mathcal{BD}' \left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{DD}'\right)^{-1} \mathcal{C}\Sigma\mathcal{A}'$$

$$\Leftrightarrow \Sigma = \mathcal{A}\Sigma\mathcal{A}' + \mathcal{BB}' - \left(\mathcal{A}\Sigma\mathcal{C}' + \mathcal{BD}'\right)\left(\mathcal{C}\Sigma\mathcal{C}' + \mathcal{DD}'\right)^{-1}\left(\mathcal{A}\Sigma\mathcal{C}' + \mathcal{BD}'\right)'. \qquad \text{(A.30)}$$

Note that formulating the matrix Riccati equation as in equation (5.7) implies that the associated Kalman gain is given by

$$\widetilde{K} = \widetilde{\mathcal{A}}\overline{\Sigma}\mathcal{C}' \left(\mathcal{C}\overline{\Sigma}\mathcal{C}' + \mathcal{DD}'\right)^{-1},$$

where $\overline{\Sigma}$ is the solution to the matrix Riccati equation. Hence, it is straightforward to show that the matrix $\left(\widetilde{\mathcal{A}} - \widetilde{K}\mathcal{C}\right)$ is equal to the matrix $(\mathcal{A} - K\mathcal{C})$ by using the expressions of $\widetilde{\mathcal{A}}$ and $\widetilde{K}$:

$$\begin{aligned}
\widetilde{\mathcal{A}} - \widetilde{K}\mathcal{C} &= \left(\mathcal{A} - \mathcal{BD}'\left(\mathcal{DD}'\right)^{-1}\mathcal{C}\right)\left(I_n - \overline{\Sigma}\mathcal{C}'\left(\mathcal{C}\overline{\Sigma}\mathcal{C}' + \mathcal{DD}'\right)^{-1}\mathcal{C}\right) \\
&= \mathcal{A} - \mathcal{A}\overline{\Sigma}\mathcal{C}'\left(\mathcal{C}\overline{\Sigma}\mathcal{C}' + \mathcal{DD}'\right)^{-1}\mathcal{C} \\
&\quad - \mathcal{BD}'\left(\mathcal{DD}'\right)^{-1}\left(I_k - \mathcal{C}\overline{\Sigma}\mathcal{C}'\left(\mathcal{C}\overline{\Sigma}\mathcal{C}' + \mathcal{DD}'\right)^{-1}\right)\mathcal{C} \\
&= \mathcal{A} - \mathcal{A}\overline{\Sigma}\mathcal{C}'\left(\mathcal{C}\overline{\Sigma}\mathcal{C}' + \mathcal{DD}'\right)^{-1}\mathcal{C} \\
&\quad - \mathcal{BD}'\left(\mathcal{DD}'\right)^{-1}\left(\mathcal{C}\overline{\Sigma}\mathcal{C}' + \mathcal{DD}'\right)^{-1}\mathcal{C} \\
&= \mathcal{A} - \left(\mathcal{A}\overline{\Sigma}\mathcal{C}' + \mathcal{BD}'\right)\left(\mathcal{C}\overline{\Sigma}\mathcal{C}' + \mathcal{DD}'\right)^{-1}\mathcal{C}.
\end{aligned}$$

Recalling that $K = \left(\mathcal{A}\overline{\Sigma}\mathcal{C}' + \mathcal{BD}'\right)\left(\mathcal{C}\overline{\Sigma}\mathcal{C}' + \mathcal{DD}'\right)^{-1}$ (see equation (2.13)) completes the computations.

Derivation of the LIE model representation in the multivariate example of Subsection (5.3.3)

The observability matrix is

$$\mathcal{O} = \left(\begin{array}{ccc} \mathcal{C}' & \mathcal{A}'\mathcal{C}' & \left(\mathcal{A}'\right)^2\mathcal{C}' \end{array} \right)'$$

$$= \begin{pmatrix} 0 & 0 & \lambda_s \\ \phi_{kk} & \phi_{ks,1}\lambda_s & \phi_{ks}\lambda_s \\ 0 & \lambda_s & 0 \\ \phi_{kk}^2 & \left(\phi_{kk}\phi_{ks,1} + \phi_{ks}\right)\lambda_s & \phi_{kk}\phi_{ks}\lambda_s \\ 0 & 0 & 0 \\ \phi_{kk}^3 & \phi_{kk}\lambda_s\left(\phi_{kk}\phi_{ks,1} + \phi_{ks}\right) & \phi_{kk}^2\phi_{ks}\lambda_s \end{pmatrix},$$

and it has full column rank of $n = 3$. The reachability matrix is

$$\mathcal{R} = \begin{pmatrix} \mathcal{B} & \mathcal{A}\mathcal{B} & \mathcal{A}^2\mathcal{B} \end{pmatrix}$$

$$= \begin{pmatrix} \phi_{ks} & \omega\phi_{ks,1}\lambda_s & \phi_{kk}\phi_{ks} & \phi_{ks,1}\lambda_s\left(1 + \omega\phi_{kk}\right) & \phi_{kk}^2\phi_{ks} & \left(\phi_{kk}\phi_{ks,1}\left(1 + \omega\phi_{kk}\right) + \phi_{ks}\right)\lambda_s \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix},$$

and it has full row rank of $n = 3$.

My calculations are based on Pappas et al. (1980). I present only the necessary steps to compute the solution to the algebraic Riccati equation. For further details, the reader is directed to the article by Pappas et al. (1980). The basic idea is to compute the generalized eigenvectors (and generalized principal vectors) of the generalized eigenvalue problem

$$Mz = \Lambda Nz, \tag{A.31}$$

with

$$M = \begin{pmatrix} \widetilde{\mathcal{A}} & 0_{3\times3} \\ -\widetilde{\mathcal{Q}} & I_3 \end{pmatrix} \text{ and } N = \begin{pmatrix} I_3 & \mathcal{C}'\left(\mathcal{D}\mathcal{D}'\right)^{-1}\mathcal{C} \\ 0_{3\times3} & \widetilde{\mathcal{A}} \end{pmatrix},$$

where the system matrices $\widetilde{\mathcal{A}}$, $\mathcal{C}$, $\mathcal{D}$ and $\widetilde{\mathcal{Q}}$ are defined in the main text. The solution of the Riccati equation can be derived from the stable generalized eigenspace of the problem.

Note that a generalized principal vector is characterized as follows. Suppose Λ is a generalized eigenvalue with multiplicity $r > 1$. A chain of generalized principal vectors is a set of vectors that satisfy

$$Mz_1 = \Lambda Nz_1,$$
$$(M - \Lambda N)z_i = \Lambda z_{i-1} \text{ for } i = 2, 3, ..., j \text{ with } j \leq r.$$

The vector z_i is called a generalized principal vector of grade i.

In a first step, I solve the generalized characteristic equation to find the generalized eigenvalues. Because

$$det\left[M - \Lambda N\right] = -\frac{\Lambda^2\left(1 + \omega\Lambda\right)\left(\Lambda + \omega\right)}{\omega^2},$$

the generalized eigenvalues are $\Lambda_1 = \Lambda_2 = 0$, $\Lambda_3 = -\omega$ and $\Lambda_4 = -\omega^{-1}$. The generalized eigenvector corresponding to the zero eigenvalue is easy to find as $u_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}'$. Due to the multiplicity of two, I calculate one generalized principal vector u_2, satisfying

$$(M - \Lambda_2 N)u_2 = Nu_1,$$

as $u_2 = \begin{pmatrix} 0 & -\frac{\omega\phi_{ks,1}\lambda_s}{\phi_{kk}} & -\frac{\phi_{ks,1}\lambda_s}{\phi_{kk}} & 0 & 0 & 0 \end{pmatrix}'$. The last generalized eigenvector conforms to

$$(M - \Lambda_3 N) u_3 = 0_{6\times 1},$$

and it is computed as $u_3 = \begin{pmatrix} -\frac{\phi_{kk}(1+\lambda_s^2)}{\lambda_s^2\omega\phi_{ks,1}(1-\omega^2)} & \frac{1+\lambda_s^2\omega^2}{\omega^3-\omega} & -\lambda_s^2 & 0 & -\frac{1}{\omega} & 1 \end{pmatrix}'$.

Finally, define the stable generalized eigenspace and partition it into two 3×3 blocks, U_1 and U_2, as

$$\begin{aligned} U &= \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} \\ &\equiv \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}. \end{aligned}$$

Hence, I compute the solution to the algebraic Riccati equation as

$$\begin{aligned} \overline{\Sigma} &= U_2 U_1^{-1} \\ &= \left(\frac{1-\omega^2}{1+\lambda_s^2\omega^4}\right) \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -\omega \\ 0 & -\omega & \omega^2 \end{pmatrix}. \end{aligned}$$

Computation of the wedge of nonfundamentalness for $q \geq 2$

For this purpose, I have to resort to the explanations of Chapters 6 and 7. Starting with the true model solution as in the main text

$$y_t = M(L) w_t, \tag{A.32}$$

which is supposed to exhibit a nonfundamental representation, I show how to derive the Wold representation that would be estimated by the econometrician. This procedure uses Blaschke matrices (see Chapter 6 for a detailed definition and explanation). Choosing an appropriate Blaschke matrix $B(L)$, the Wold representation can be calculated as

$$y_t = T(L) a_t, \tag{A.33}$$

where $T(L) = M(L) B(L)^{-1} B(0) M(0)^{-1}$ and $a_t = M(0) B(0)^{-1} \varepsilon_t$ with $\varepsilon_t = B(L) w_t$. $B(L)$ is defined such that the lag polynomial associated with $M(L) B(L)^{-1}$ has all roots outside of the unit circle. Note that, due to the characteristics of $B(L)$ and that w_t has unit standard deviation, $E(\varepsilon_t \varepsilon_t') = I_k$. Hence, the variance-covariance matrix of the forecast errors in y_t is $E(a_t a_t') = M(0) B(0)^{-1} (M(0) B(0)^{-1})'$. In the main text, I define their direct counterpart in the FIA model v_t, i.e., related to the vector MA representation above, I ob-

tain $v_t = M(0) w_t$. Thus, I can apply the rules of linear algebra to compute the wedge of nonfundamentalness as follows:

$$
\begin{aligned}
\varphi &= log\left(det\left[M(0)\right] det\left[B(0)^{-1}\right] det\left[\left(B(0)^{-1}\right)'\right] det\left[M(0)'\right]\right) \\
&\quad -log\left(det\left[M(0)\right] det\left[M(0)'\right]\right) \\
&= log\left(det\left[M(0)\right] det\left[B(0)^{-1}\right] det\left[B(0)^{-1}\right] det\left[M(0)\right]\right) \\
&\quad -log\left(det\left[M(0)\right] det\left[M(0)\right]\right) \\
&= 2log\left(det\left[M(0)\right]\right) + 2log\left(det\left[B(0)^{-1}\right]\right) - 2log\left(det\left[M(0)\right]\right) \\
&= 2log\left(det\left[B(0)^{-1}\right]\right).
\end{aligned}
$$

For my workhorse model, the definition of $B(L)^{-1}$ is given in Chapter 7 (see equation (7.3)) by

$$
B(L)^{-1} = G \cdot R_1(L)^{-1} R_2(L)^{-1} \dots R_{q-1}(L)^{-1}
$$

with

$$
R_i(L)^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1-\overline{\omega}_i L}{L-\omega_i} \end{pmatrix},
$$

$$
G = \frac{1}{\sqrt{1+(\lambda_\theta \omega^q)^2}} \begin{pmatrix} 1 & -\lambda_\theta \omega^q \\ \lambda_\theta \omega^q & 1 \end{pmatrix},
$$

where $\overline{\omega}_i$ denotes the complex conjugate of ω_i. I can use this Blaschke matrix for the model in Chapter 5 (but replace λ_θ by λ_s).

If q is even, $q-2$ roots appear as complex conjugate pairs, and the $(q-1)th$ root is simply given by $-\omega$. Then

$$
\begin{aligned}
det\left[B(0)^{-1}\right] &= det\left[G \cdot R_1(0)^{-1} R_2(0)^{-1} \dots R_{q-1}(0)^{-1}\right] \\
&= \underbrace{det\left[G\right]}_{=1} \underbrace{det\left[R_1(0)^{-1} R_2(0)^{-1} \dots R_{q-2}(0)^{-1}\right]}_{=\frac{1}{\omega^{q-2}}} \underbrace{det\left[R_{q-1}(0)^{-1}\right]}_{=\frac{1}{\omega}} \\
&= \frac{1}{\omega^{q-1}}.
\end{aligned}
$$

If q is uneven, there are $q-1$ roots that appear as complex conjugate pairs and

$$
det\left[B(0)^{-1}\right] = det\left[G \cdot R_1(0)^{-1} R_2(0)^{-1} \dots R_{q-1}(0)^{-1}\right]
$$

$$= \underbrace{det\,[G]}_{=\,1} \; \underbrace{det\left[R_1\,(0)^{-1}\,R_2\,(0)^{-1}\,...R_{q-1}\,(0)^{-1}\right]}_{=\,\frac{1}{\omega^{q-1}}}$$

$$= \frac{1}{\omega^{q-1}}\,.$$

In both cases, $\varphi = -2\,(q-1)\,log\,(\omega)$. Notice that this result is independent of the relative weight of the news shock.

Appendix to Chapter 6

Additional remarks to Subsection 6.5.2

Lippi and Reichlin (1994) propose an alternative way to calculate an appropriate rotation matrix that is part of the Blaschke matrix in order to flip the root of a two-dimensional MA lag polynomial $det\,[T\,(L)]$. Let $t_{ij}\,(z_0)$ denotes the $[i,j]$-element of $T\,(z)$ evaluated at z_0. The rotation matrix can be computed as

$$
G \;=\; h^{-1}
\begin{pmatrix}
t_{12}\,(z_0) & \overline{t_{11}\,(z_0)} \\
-t_{11}\,(z_0) & \overline{t_{12}\,(z_0)}
\end{pmatrix},
$$

where $h = \sqrt{\left|t_{11}\,(z_0)\right|^2 + \left|t_{12}\,(z_0)\right|^2}$, and upper bars denote the complex conjugate elements. However, G is written such that the first column of $T\,(z)\,G$ is determined. If I want to determine the second column, I have to change the columns of G.

Because the second column is associated with the news shock in my Lucas asset tree example, I apply the formula above (whereby changing the columns) to determine the rotation matrix when transforming the nonfundamental representation into the fundamental representation. The corresponding Blaschke matrix is then given by (the inverse of)

$$
B_1\,(L) =
\begin{pmatrix}
1 & 0 \\
0 & \frac{L+\beta}{1+\beta L}
\end{pmatrix}
G_1^{-1}
$$

with $G_1 =
\begin{pmatrix}
\frac{1}{\sqrt{1+\beta^4}} & \frac{\beta^2}{\sqrt{1+\beta^4}} \\
\frac{\beta^2}{\sqrt{1+\beta^4}} & -\frac{1}{\sqrt{1+\beta^4}}
\end{pmatrix}.$

The Blaschke matrix that I use for flipping the root back to $-\beta$ is

$$
B_2\,(L) \;=\; G_2 \cdot
\begin{pmatrix}
1 & 0 \\
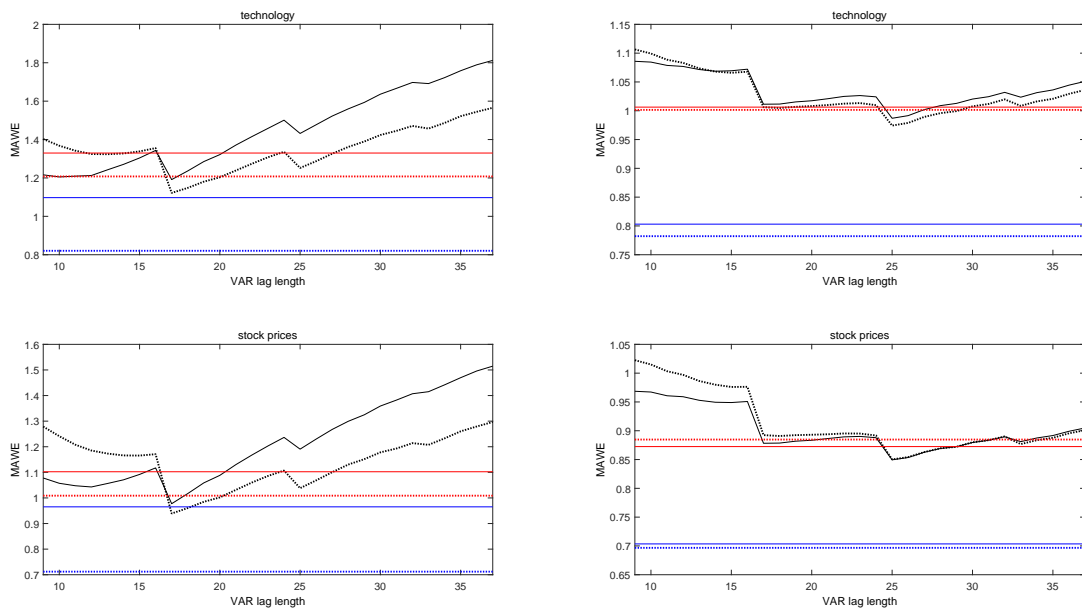0 & \frac{L+\beta}{1+\beta L}
\end{pmatrix}
$$

with $G_2 = \frac{\sqrt{2}}{2}
\begin{pmatrix}
1 & 1 \\
1 & -1
\end{pmatrix}.$

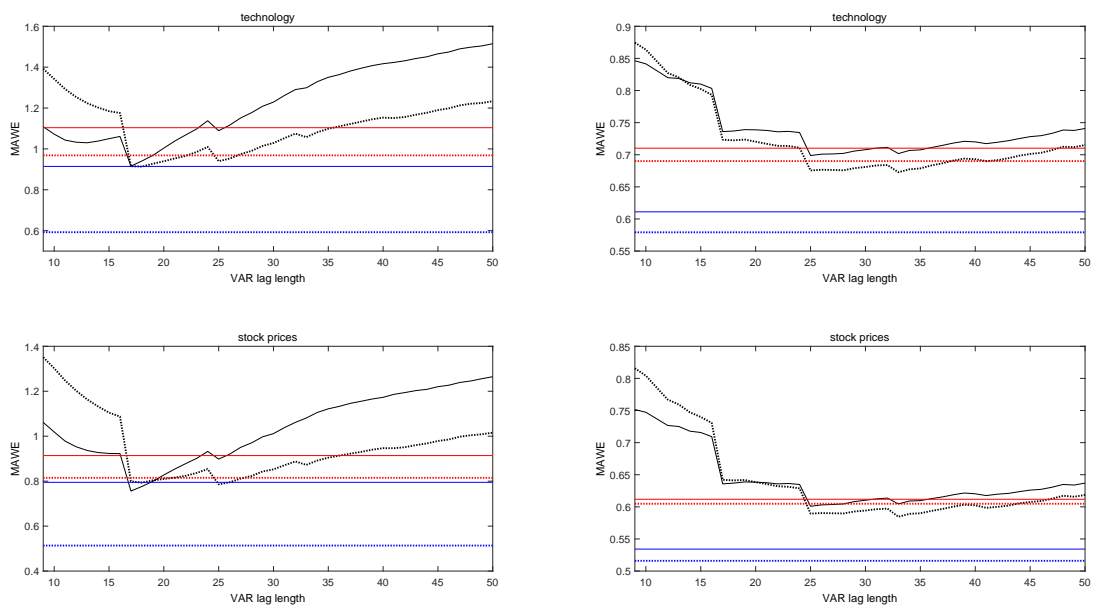# Appendix to Chapter 7

Additional figures w.r.t. Section 7.4

Figure A.1: EXERCISES 7.3 & 7.4 - MAWE RESULTS (SELECTED VAR LAG LENGTHS, STOCK PRICES)
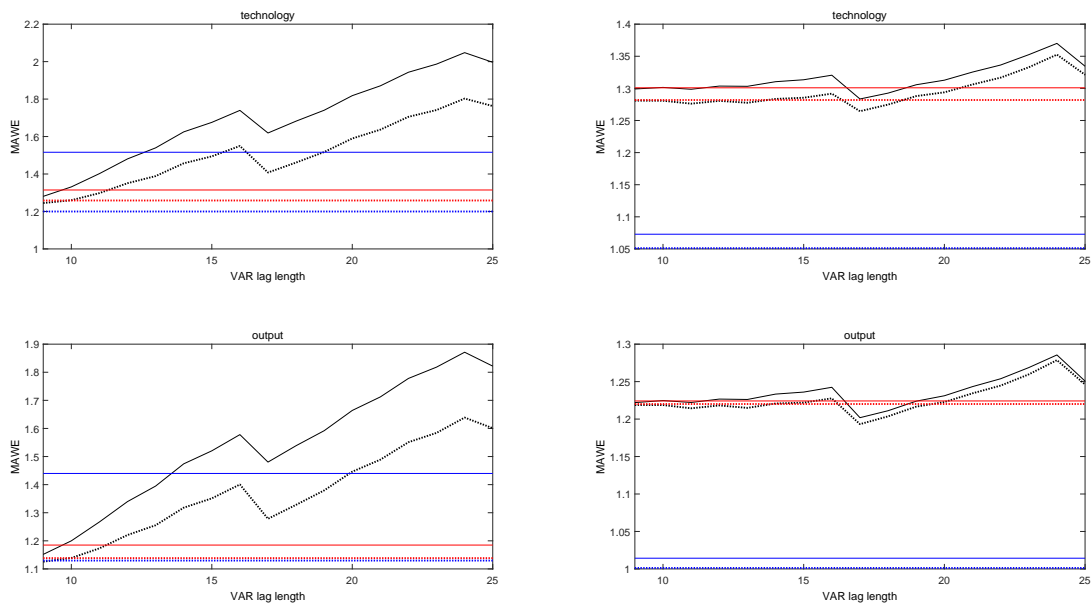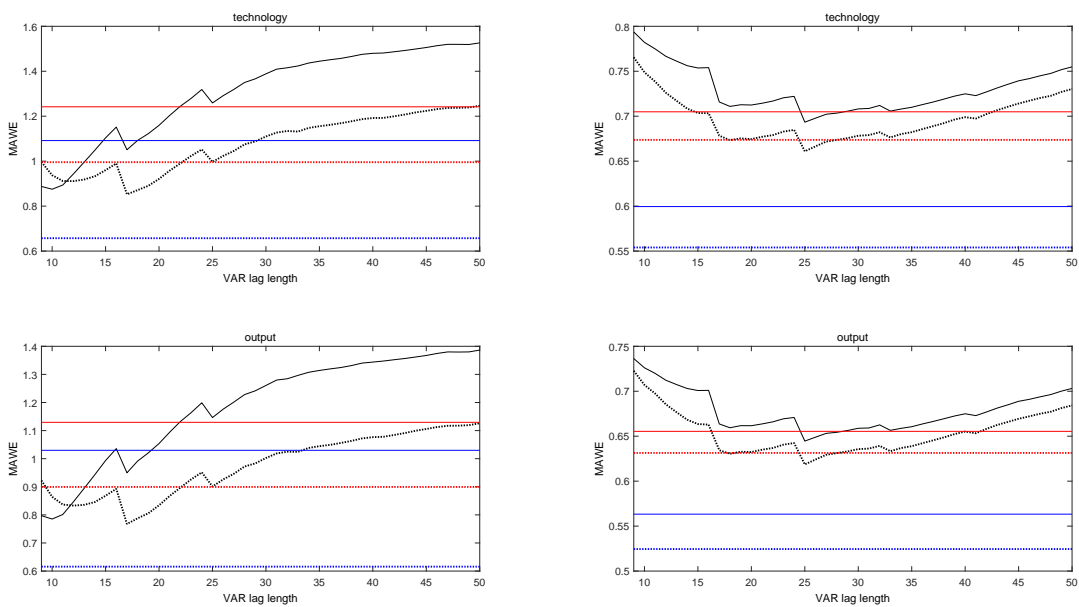
Exercise 7.3 ($t = 300$)

Exercise 7.4 ($t = 500$)

Note: The figure depicts the computed MAWEs associated with the estimated impulse responses resulting from different estimation approaches. These approaches are based on the SVAR model depending on selected lag lengths (in black), SVAR using the lag length suggested by the AIC (in red) and CCA (in blue). Solid lines refer to the standard application of these techniques and dotted lines to the corresponding root flipping procedure. MAWE values are calculated as in equation (5.8) multiplied by 100. MAWEs associated with the impulse responses to $w_{0,t}^{\theta}$ are shown in the left column and MAWEs w.r.t. impulse responses to $w_{8,t}^{\theta}$ in the right column.

Figure A.2: Exercises 7.5 & 7.6 - MAWE results (selected VAR lag lengths, output)
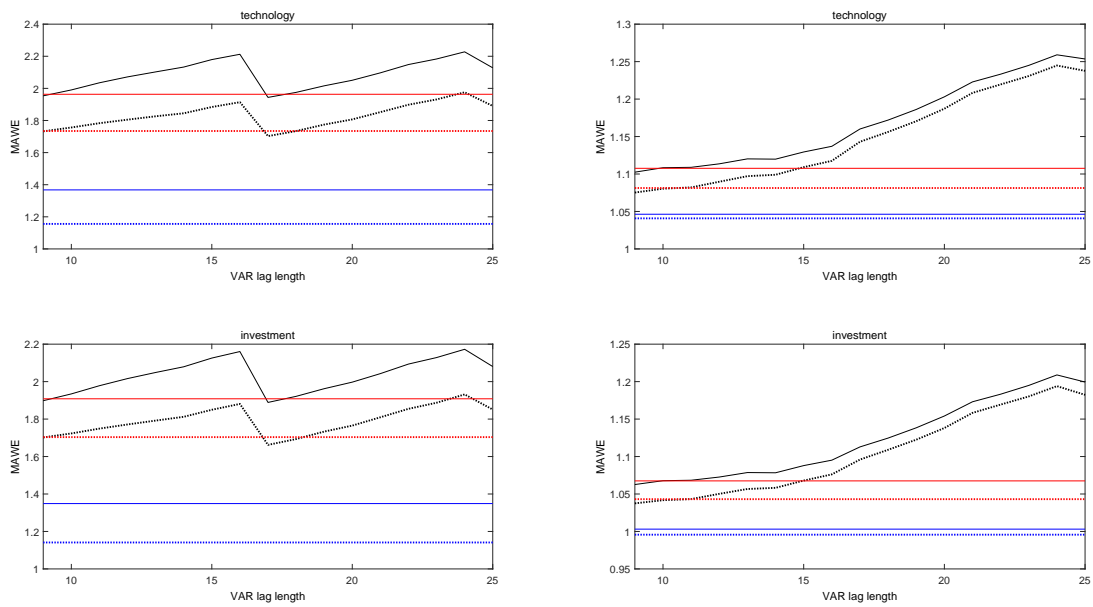
Exercise 7.5 ($t = 200$)
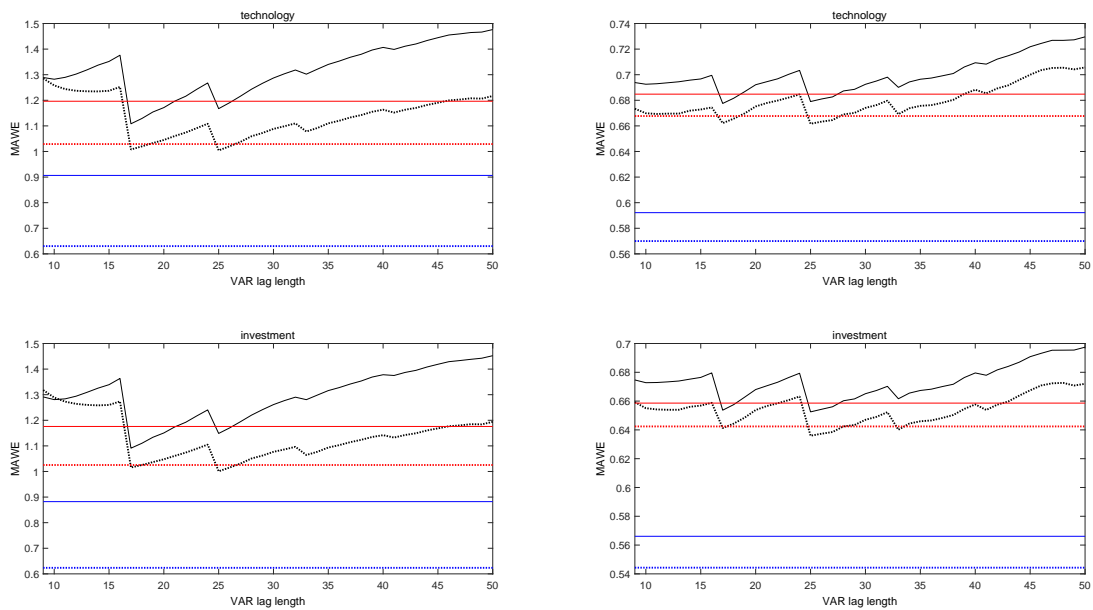

Exercise 7.6 ($t = 500$)

Note: The figure depicts the computed MAWEs associated with the estimated impulse responses resulting from different estimation approaches. These approaches are based on the SVAR model depending on selected lag lengths (in black), SVAR using the lag length suggested by the AIC (in red) and CCA (in blue). Solid lines refer to the standard application of these techniques and dotted lines to the corresponding root flipping procedure. MAWE values are calculated as in equation (5.8) multiplied by 100. MAWEs associated with the impulse responses to $w_{0,t}^{\theta}$ are shown in the left column and MAWEs w.r.t. impulse responses to $w_{8,t}^{\theta}$ in the right column.

Figure A.3: EXERCISES 7.7 & 7.8 - MAWE RESULTS (SELECTED VAR LAG LENGTHS, IN-VESTMENT)

Exercise 7.7 ($t = 200$)

Exercise 7.8 ($t = 500$)

Note: The figure depicts the computed MAWEs associated with the estimated impulse responses resulting from different estimation approaches. These approaches are based on the SVAR model depending on selected lag lengths (in black), SVAR using the lag length suggested by the AIC (in red) and CCA (in blue). Solid lines refer to the standard application of these techniques and dotted lines to the corresponding root flipping procedure. MAWE values are calculated as in equation (5.8) multiplied by 100. MAWEs associated with the impulse responses to $w_{0,t}^{\theta}$ are shown in the left column and MAWEs w.r.t. impulse responses to $w_{8,t}^{\theta}$ in the right column.

The workhorse model with tax news shocks

To illustrate the implications of tax news in the workhorse model, I use the standard RBC model as specified by Chari et al. (2008), which is an extended version of the workhorse model in Chapter 2. To incorporate tax shocks into the model, the workhorse model is augmented by a government sector. Public spending in terms of lump-sum transfers to the households is financed by a tax on labor. The exogenous labor tax rate follows a conventional AR(1) process with a white noise shock. As opposed to Chari et al. (2008), I model this shock not as a surprise shock but as a news shock that is anticipated by the agents q periods before it materializes. I refer the reader to the technical appendix for Chari et al. (2008) for details of the model and its solution.[2]

I derive the corresponding state space representation of the model by solving the second order difference equation of the capital stock analogously to my explanations in Chapter 5. The resulting FIA model representation is very similar to the one given in equations (5.29) and (5.30):

$$
\begin{pmatrix}
\widetilde{k}_{t+1} \\
\widetilde{\tau}_t \\
w^\tau_{q,t} \\
w^\tau_{q,t-1} \\
\vdots \\
w^\tau_{q,t-q+2} \\
w^\tau_{q,t-q+1}
\end{pmatrix}
=
\underbrace{
\begin{pmatrix}
\phi_{kk} & \rho_\tau \phi_{k\tau} & \omega^{q-2}\phi_{k\tau,1} & \cdots & \omega\phi_{k\tau,1} & \phi_{k\tau,1} & \phi_{k\tau} \\
0 & \rho_\tau & 0 & \cdots & 0 & 0 & 1 \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & 0 & \frac{1}{\sigma_\tau} & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \frac{1}{\sigma_\tau} & 0 & 0 \\
0 & 0 & 0 & \cdots & 0 & \frac{1}{\sigma_\tau} & 0
\end{pmatrix}
}_{=\,\mathcal{A}}
\Sigma_1
\begin{pmatrix}
\widetilde{k}_t \\
\widetilde{\tau}_{t-1} \\
w^\tau_{q,t-1} \\
w^\tau_{q,t-2} \\
\vdots \\
w^\tau_{q,t-q+1} \\
w^\tau_{q,t-q}
\end{pmatrix}
$$

$$
+
\underbrace{
\begin{pmatrix}
\phi_{k\theta} & \omega^{q-1}\phi_{k\tau,1} \\
0 & 0 \\
0 & \frac{1}{\sigma_\tau} \\
0 & 0 \\
\vdots & \vdots \\
0 & 0 \\
0 & 0
\end{pmatrix}
}_{=\,\mathcal{B}}
\Sigma_w
\begin{pmatrix}
w^\theta_{0,t} \\
w^\tau_{q,t}
\end{pmatrix}
\tag{A.34}
$$

---

[2]The technical appendix can be found on Ellen McGrattan's website: http://www.econ.umn.edu/~erm/research.php.

$$
\left(\begin{array}{c} (\widetilde{y}_t - \widetilde{n}_t) \\ \widetilde{n}_t \end{array} \right) = \underbrace{\left(\begin{array}{ccccccc} \phi_{ynk} & \rho_\tau \phi_{yn\tau} & \omega^{q-2}\phi_{yn\tau,1} & \cdots & \omega\phi_{yn\tau,1} & \phi_{yn\tau,1} & \phi_{yn\tau} \\ \phi_{nk} & \rho_\tau \phi_{n\tau} & \omega^{q-2}\phi_{n\tau,1} & \cdots & \omega\phi_{n\tau,1} & \phi_{n\tau,1} & \phi_{n\tau} \end{array} \right)}_{=\,\mathcal{C}} \Sigma_1 \left(\begin{array}{c} \widetilde{k}_t \\ \widetilde{\tau}_{t-1} \\ w^\tau_{q,t-1} \\ w^\tau_{q,t-2} \\ \vdots \\ w^\tau_{q,t-q+1} \\ w^\tau_{q,t-q} \end{array} \right)
$$

$$
+ \underbrace{\left(\begin{array}{cc} \phi_{yn\theta} & \omega^{q-1}\phi_{yn\tau,1} \\ \phi_{n\theta} & \omega^{q-1}\phi_{n\tau,1} \end{array} \right)}_{=\,\mathcal{D}} \Sigma_w \left(\begin{array}{c} w^\theta_{0,t} \\ w^\tau_{q,t} \end{array} \right), \tag{A.35}
$$

where

$$
\Sigma_1 = \left(\begin{array}{cc} I_2 & 0_{2\times q} \\ 0_{q\times 2} & \sigma_\tau \cdot I_q \end{array} \right) \text{ and } \Sigma_w = \left(\begin{array}{cc} \sigma_\theta & 0 \\ 0 & \sigma_\tau \end{array} \right).
$$

The notation is equivalent to the notation used in Chapter 5. τ denotes the tax rate and its associated parameters. The vector of observables comprises (the log deviations of detrended) labor productivity, i.e., output Y_t divided by labor N_t, and labor itself. The parameters labeled as ϕ and the anticipation rate ω are functions of the deep model parameters. I calibrate the deep model parameters as in the original model.[3]

The resulting impulse responses for the FIA and the LIE model are plotted in Figure A.4, where q is set equal to four. The right column of A.4 reveals that there is a remarkable difference between the responses to the tax news shock in the FIA model and the responses to the tax news shock in the LIE model over the anticipation horizon.

---

[3]See the Matlab file "pu.m" in the additional files for Chari et al. (2008) on Ellen McGrattan's website (http://www.econ.umn.edu/~erm/research.php) for the original calibration.

Figure A.4: IMPULSE RESPONSES IN RBC MODEL WITH TAX NEWS SHOCK $(q = 4)$
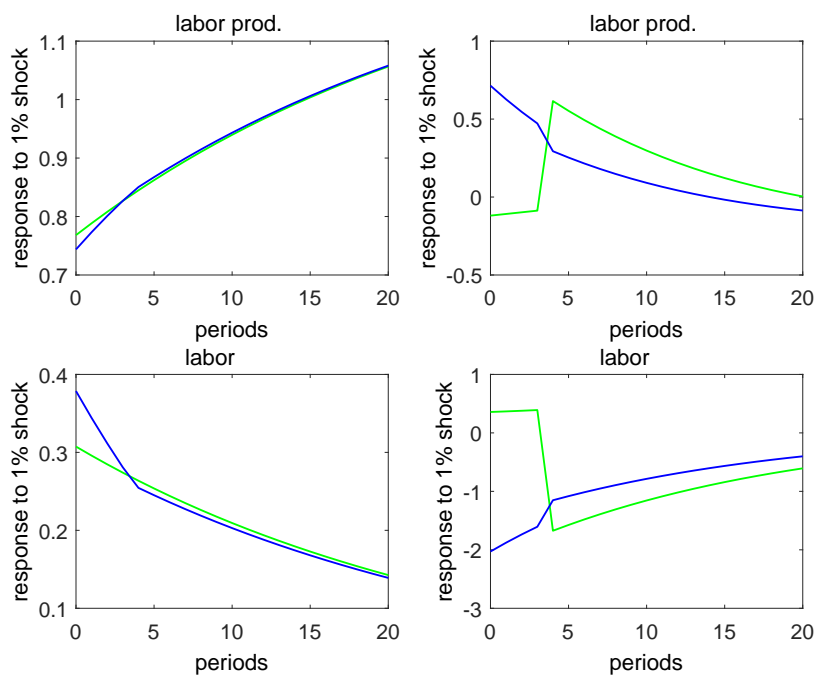
Note: The figure depicts the FIA model impulse responses (in green) and the LIE model impulse responses (in blue). Impulse responses to $w^{\theta}_{0,t}$ are shown in the left column and impulse responses to $w^{\tau}_{4,t}$ in the right column.

Bibliography

Adolfson, M., S. Laséen, J. Lindé, and M. Villani (2008). Evaluating an estimated new Keynesain small open economy model. *Journal of Economic Dynamics and Control 32*(8), 2690–2721.

Alessi, L., M. Barigozzi, and M. Capasso (2011). Non-Fundamentalness in Structural Econometric Models: A Review. *International Statistical Review 79*(1), 16–47.

Altug, S. (1989). Time-to-Build and Aggregate Fluctuations: Some New Evidence. *International Economic Review 30*(4), 889–920.

Amisano, G. and C. Giannini (1997). *Topics in Structural VAR Econometrics* (2nd ed.). Springer.

Anderson, B. D. O. and J. B. Moore (1979). *Optimal Filtering*. Prentice-Hall.

Anderson, E. W., L. P. Hansen, E. R. McGrattan, and T. J. Sargent (1996). Mechanics of forming and estimating dynamic linear economies. In H. M. Amman, D. A. Kendrick, and J. Rust (Eds.), *Handbook of Computational Economics*, Volume 1, Chapter 4, pp. 171–252. Elsevier.

Aoki, M. (1987). *State Space Modeling of Time Series*. Springer.

Aoki, M. and A. Havenner (1991). State space modeling of multiple time series. *Econometric Reviews 10*(1), 1–59.

Archontakis, F. (1998). An Alternative Proof Of Granger's Representation Theorem For I(1) Systems Through Jordan Matrices. *Journal of the Italian Statistical Society 7*(2), 111–127.

Aruoba, S. B., J. Fernández-Villaverde, and J. F. Rubio-Ramírez (2006). Comparing solution methods for dynamic equilibrium economies. *Journal of Economic Dynamics and Control 30*(12), 2477–2508.

Barakchian, S. M. and C. Crowe (2013). Monetary policy matters: Evidence from new shocks data. *Journal of Monetary Economics 60*(8), 950–966.

Barro, R. J. and R. G. King (1984). Time-Separable Preferences and Intertemporal-Substitution Models of Business Cycles. *Quarterly Journal of Economics 99*(4), 817–839.

Barsky, R. B. and E. R. Sims (2011). News shocks and business cycles. *Journal of Monetary Economis 58*(3), 273–289.

Barsky, R. B. and E. R. Sims (2012). Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence. *American Economic Review 102*(4), 1343–1377.

Basu, S., J. G. Fernald, and M. S. Kimball (2006). Are Technology Improvements Contractionary? *American Economic Review 96*(5), 1418–1448.

Bauer, D. (2005a). Comparing the CCA Subspace Method to Pseudo Maximum Likelihood Methods in the case of No Exogenous Inputs. *Journal of Time Series Analysis 26*(5), 631–668.

Bauer, D. (2005b). Estimating Linear Dynamical Systems Using Subspace Methods. *Econometric Theory 21*(1), 181–211.

Bauer, D. and M. Wagner (2002). Estimating cointegrated systems using subspace algorithms. *Journal of Econometrics 111*, 47–84.

Bauer, D. and M. Wagner (2003). A Canonical Form for Unit Root Processes in the State Space Framework. Discussion paper 03-12, University of Bern.

Bauer, D. and M. Wagner (2009). Using subspace algorithm cointegration analysis: Simulation performance and application to the term structure. *Computational Statistics and Data Analysis 53*(6), 1954–1973.

Bauer, D. and M. Wagner (2012). A State Space Canonical Form For Unit Root Processes. *Econometric Theory 28*(6), 1313–1349.

Beaudry, P., F. Collard, and F. Portier (2006). Gold Rush Fever in Business Cycles. Working paper 12710, National Bureau of Economic Research.

Beaudry, P., F. Collard, and F. Portier (2011). Gold rush fever in business cycles. *Journal of Monetary Economics 58*(2), 84–97.

Beaudry, P., P. Fève, A. Guay, and F. Portier (2015). When is Nonfundamentalness in VARs A Real Problem? An Application to News Shocks. Working paper 21466, National Bureau of Economic Research.

Beaudry, P. and B. Lucke (2010). Letting Different Views about Business Cycles Compete. In D. Acemoglu, K. Rogoff, and M. Woodford (Eds.), *NBER Macroeconomics Annual 2009, Volume 24*, pp. 413–455. National Bureau of Economic Research.

Beaudry, P., D. Nam, and J. Wang (2011). Do Mood Swings Drive Business Cycles and is it Rational? Working paper 17651, National Bureau of Economic Research.

Beaudry, P. and F. Portier (2005). The 'news view' of economic fluctuations: Evidence from aggregate Japanese data and sectoral US data. *Journal of the Japanese and International Economies 19*, 635–652.

Beaudry, P. and F. Portier (2006). Stock Prices, News, and Economic Fluctuations. *American Economic Review 96*(4), 1293–1307.

Beaudry, P. and F. Portier (2007). When can changes in expectations cause business cycle fluctuations in neo-classical settings? *Journal of Economic Theory 135*, 458–477.

Beaudry, P. and F. Portier (2014a). News-Driven Business Cycles: Insights and Challenges. *Journal of Economic Literature 52*(4), 993–1074.

Beaudry, P. and F. Portier (2014b). On The Quantitative Importance of Non-fundamentalness for News Shocks. mimeo.

Beaudry, P., F. Portier, and A. Seymen (2013). Comparing Two Methods for the Identification of News Shocks. Discussion paper 13-110, Centre for European Economic Research.

Benes, J. and D. Vavra (2005). Eigenvalue Filtering In VAR Models With Application To The Czech Business Cycle. Working paper no. 549, European Central Bank.

Bernanke, B. S., J. Boivin, and P. Eliasz (2005). Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach. *Quarterly Journal of Economics 120*(1), 387–422.

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

Blanchard, O. J. and D. Quah (1989). The Dynamic Effects of Aggregate Demand and Supply Disturbances. *American Economic Review 79*(4), 655–673.

Braun, P. A. and S. Mittnik (1993). Misspecifications in vector autoregressions and their effects on impulse responses and variance decompositions. *Journal of Econometrics 59*(3), 319–341.

Breitung, J., R. Brüggemann, and H. Lütkepohl (2004). Structural Vector Autoregressive Modeling and Impulse Responses. In M. Krätzig and H. Lütkepohl (Eds.), *Applied Time Series Econometrics*, pp. 159–196. Cambridge University Press.

Burns, A. F. and W. C. Mitchell (1946). *Measuring Business Cycles*. National Bureau of Economic Research.

Campbell, J. Y. (1987). Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis. *Econometrica 55*(6), 1249–1273.

Canova, F. and G. D. Nicoló (2002). Monetary disturbances matter for business fluctuations in the G-7. *Journal of Monetary Economics 49*(6), 1131–1159.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2005). A Critique of Structural VARs Using Business Cycle Theory. Working paper 631, Federal Reserve Bank of Minneapolis.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2008). Are structural VARs with long-run restrictions useful in developing business cycle theory? *Journal of Monetary Economics 55*(8), 1337–1352.

Christiano, L., C. Illut, R. Motto, and M. Rostagno (2010). Monetary Policy and Stock Market Booms. Working paper 16402, National Bureau of Economic Research.

Christiano, L. J., M. Eichenbaum, and R. Vigfusson (2007). Assessing Structural VARs. In D. Acemoglu, K. Rogoff, and M. Woodford (Eds.), *NBER Macroeconomics Annual 2006, Volume 21*, pp. 1–106. National Bureau of Economic Research.

Cochrane, J. H. (1994a). Permanent and Transitory Components of GNP and Stock Prices. *Quarterly Journal of Economics 109*(1), 241–265.

Cochrane, J. H. (1994b). Shocks. *Carnegie-Rochester Conference Series on Public Policy 41*, 295–364.

Cochrane, J. H. (1998). What do the VARs mean? Measuring the output effects of monetary policy. *Journal of Monetary Economics 41*(2), 277–300.

Cooley, T. F. (1995). *Frontiers of Business Cycle Research*. Princeton University Press.

Cooley, T. F. and M. Dwyer (1998). Business cycle analysis without much theory: A look at structural VARs. *Journal of Econometrics 83*(1-2), 57–88.

d'Andréa Novel, B. and M. D. Lara (2013). *Control Theory for Engineers - A Primer*. Springer.

Danthine, J.-P. and J. B. Donaldson (2002). Decentralizing the Stochastic Growth Model. Discussion paper 3348, Centre for Economic Policy Research.

Danthine, J.-P., J. B. Donaldson, and T. Johnsen (1998). Productivity growth, consumer confidence and the business cycle. *European Economic Review 42*(6), 1113–1140.

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

Díaz-Giménez, J. (1999). Linear Quadratic Approximations - An introduction. In R. Marimon and A. Scott (Eds.), *Computational Methods for the Study of Dynamic Economies*, Chapter 2, pp. 13–29. Oxford University Press.

Doan, T. A. (1992). *RATS Manual*. Estima.

Dupor, B. and M. S. Mehkari (2014). The analytics of technology news shocks. *Journal of Economic Theory 153*, 392–427.

Durdu, C. B., R. Nunes, and H. Sapriza (2013). News and sovereign default risk in small open economies. *Journal of International Economics 91*(1), 1–17.

Erceg, C. J., L. Guerrieri, and C. Gust (2005). Can Long-Run Restrictions Identify Technology Shocks? *Journal of the European Economic Association 3*(6), 1237–1278.

Fama, E. F. (1990). Stock Returns, Expected Returns, and Real Activity. *Journal of Finance 45*(4), 1089–1108.

Fan, H. and Z. Xu (2014). Firm dynamics in news-driven business cycles: the role of endogenous survival rate. *Applied Economics 46*(15), 1767–1777.

Faust, J. (1998). The robustness of identified VAR conclusions about money. *Carnegie-Rochester Conference Series on Public Policy 49*, 207–244.

Faust, J. and E. M. Leeper (1997). When Do Long-Run Identifying Restrictions Give Reliable Results? *Journal of Business and Economic Statistics 15*(3), 345–353.

Fernald, J. (2014). A Quarterly, Utilization-Adjusted Series on Total Factor Productivity. Working paper 2012-19, Federal Reserve Bank of San Francisco.

Fernández-Villaverde, J., J. F. Rubio-Ramírez, T. J. Sargent, and M. W. Watson (2005). A, B, C's (and D)'s for Understanding VARs. Technical working paper 308, National Bureau of Economic Research.

Fernández-Villaverde, J., J. F. Rubio-Ramírez, T. J. Sargent, and M. W. Watson (2007). ABCs (and Ds) of Understanding VARs. *American Economic Review 97*(3), 1021–1026.

Fève, P. and A. Jidoud (2012). Identifying News Shocks from SVARs. *Journal of Macroeconomics 34*(4), 919–932.

Fève, P. and A. Jidoud (2014). News Shocks, Information Flows and SVARs. *Annals of Economics and Statistics 113-114*, 293–308.

Fischer, S. (1977). Long-Term Contracts, Rational Expectations, and the Optimal Money Supply Rule. *Journal of Political Economy 85*(1), 191–205.

Fisher, J. D. M. (2003). Technology Shocks Matter. Working paper 2002-14, Federal Reserve Bank of Chicago.

Fisher, J. D. M. (2006). The Dynamic Effects of Neutral and Investment-Specific Technology Shocks. *Journal of Political Economy 114*(3), 413–451.

Flavin, M. A. (1981). The Adjustment of Consumption to Changing Expectations About Future Income. *Journal of Political Economy 89*(5), 974–1009.

Forni, M. and L. Gambetti (2010). Macroeconomic Shocks and the Business Cycle: Evidence from a Structural Factor Model. Discussion paper 7692, Centre for Economic Policy Research.

Forni, M. and L. Gambetti (2014). Sufficient information in structural VARs. *Journal of Monetary Economics 66*, 124–136.

Forni, M., L. Gambetti, and L. Sala (2014). No News in Business Cycles. *The Economic Journal 124*(581), 1168–1191.

Forni, M., D. Giannone, M. Lippi, and L. Reichlin (2009). Opening the Black Box: Structural Factor Models with Large Cross Section. *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.

Francis, N., M. T. Owyang, J. E. Roush, and R. DiCecio (2014). A Flexible Finite-Horizon Alternative to Long-run Restrictions with an Application to Technology Shocks. *The Review of Economics and Statistics 96*(4), 638–647.

Fry, R. and A. Pagan (2011). Sign Restrictions in Structural Vector Autoregressions: A Critical Review. *Journal of Economic Literature 49*(4), 938–960.

Fujiwara, I., Y. Hirose, and M. Shintani (2011). Can News Be a Major Source of Aggregate Fluctuations? A Bayesian DSGE Approach. *Journal of Money, Credit and Banking 43*(1), 1–29.

Galí, J. (1999). Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations? *American Economic Review 89*(1), 249–271.

Galí, J. (2008). *Monetary Policy, Inflation, and the Business Cycle: An introduction to the New Keynesian Framework*. Princeton University Press.

Gershun, N. and S. G. Harrison (2008). Asset Pricing in Dynamic Stochastic General Equilibrium Models with Indeterminacy. *Macroeconomic Dynamics 12*(1), 50–71.

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.

Girardi, A. (2014). Expectations and macroeconomic fluctuations in the euro area. *Economics Letters 125*(2), 315–318.

Gonzalo, J. and S. Ng (2001). A systematic framework for analyzing the dynamic effects of permanent and transitory shocks. *Journal of Economic Dynamics and Control 25*, 1527–1546.

Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988). Investment, Capacity Utilization, and the Real Business Cycles. *American Economic Review 78*(3), 402–417.

Greenwood, J., Z. Hercowitz, and P. Krusell (1997). Long-Run Implications of Investment-Specific Technological Change. *American Economic Review 87*(3), 342–362.

Greenwood, J., Z. Hercowitz, and P. Krusell (2000). The role of investment-specific technological change in the business cycle. *European Economic Review 44*(1), 91–115.

Gu, G. (2012). *Discrete-Time Linear Systems: Theory and Design with Applications*. Springer.

Gunn, C. M. and A. Johri (2013). An expectations-driven interpretation of the 'Great Recession'. *Journal of Monetary Economics 60*(4), 391–407.

Haertel, T. and B. Lucke (2008). Do News Shocks Drive Business Cycles? Evidence from German Data. *Economics: The Open-Access, Open-Assessment E-Journal 2*(10), 1–21.

Hall, G. J. (1996). Overtime, effort, and the propagation of business cycle shocks. *Journal of Monetary Economics 38*(1), 139–160.

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

Hannan, E. J. and M. Deistler (1988). *The Statistical Theory of Linear Systems*. John Wiley & Sons.

Hansen, G. D. (1985). Indivisible Labor and the Business Cycle. *Journal of Monetary Economics 16*(3), 309–327.

Hansen, L. P. and T. J. Sargent (2014). *Recursive Models of Dynamic Linear Economies*. Princeton University Press.

Heer, B. and A. Maußner (2008). Computation of Business Cycle Models: A Comparison of Numerical Methods. *Macroeconomic Dynamics 12*(5), 641–663.

Ingram, B. F., N. R. Kocherlakota, and N. E. Savin (1994). Explaining business cycles: A multiple-shock approach. *Journal of Monetary Economics 34*(3), 415–428.

Ireland, P. N. (2004). Technology Shocks in the New Keynesian Model. *Review of Economics and Statistics 86*(4), 923–936.

Jaimovich, N. and S. T. Rebelo (2009). Can News about the Future Drive the Business Cycle? *American Economic Review 99*(4), 1097–1118.

Jermann, U. J. (1998). Asset pricing in production economies. *Journal of Monetary Economics 41*(2), 257–275.

Johansen, S. (1995). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

Justiniano, A., C. L. Evans, J. R. Campbell, and J. D. Fisher (2012). Macroeconomic Effects of Federal Reserve Forward Guidance. *Brookings Papers on Economic Activity*, 1–54.

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME - Journal of Basic Engineering 82*(Series D), 35–45.

Kascha, C. J. and K. Mertens (2009). Business cycle analysis and VARMA models. *Journal of Economic Dynamics and Control 33*(2), 267–282.

Keynes, J. M. (1936). *The General Theory Of Employment, Interest And Money*. Macmillan and Co.

Khan, H. and J. Tsoukalas (2012). The Quantitative Importance of News Shocks in Estimated DSGE Models. *Journal of Money, Credit and Banking 44*(8), 1535–1561.

Kilian, L. (2001). Impulse Response Analysis in Vector Autoregressions with Unknown Lag Order. *Journal of Forecasting 20*(3), 161–179.

Kilian, L. (2011). Structural Vector Autoregressions. Discussion paper 8515, Centre for Economic Policy Research.

King, R. G., C. I. Plosser, and S. T. Rebelo (1988). Production, Growth and Business Cycles: I. The Basic Neoclassical Model. *Journal of Monetary Economics 21*(2), 195–232.

King, R. G., C. I. Plosser, J. H. Stock, and M. W. Watson (1991). Stochastic Trends and Economic Fluctuations. *American Economic Review 81*(4), 819–840.

Klein, P. (2000). Using the generalized Schur form to solve a multivariate linear rational expectations model. *Journal of Economic Dynamics and Control 24*(10), 1405–1423.

Kriwoluzky, A. (2012). Pre-announcement and timing: The effects of a government expenditure shock. *European Economic Review 56*(3), 373–388.

Kwakernaak, H. and R. Sivan (1972). *Linear Optimal Control Systems*. John Wiley & Sons.

Kydland, F. E. and E. C. Prescott (1982). Time To Build And Aggregate Fluctuations. *Econometrica 50*(6), 1345–1370.

Larimore, W. E. (1983). System identification, reduced-order filters and modeling via canonical variate analysis. In P. Dorato and H. S. Rao (Eds.), *Proceedings of the 1983 American Control Conference*, pp. 445–451. Piscataway.

Larsen, V. H. and L. A. Thorsrud (2015). The Value of News. Working paper 6/2015, Centre for Applied Macro- and Petroleum Economics.

Leeper, E. M. and C. A. Sims (1994). Toward a Modern Macroeconomic Model Usable for Policy Analysis. In S. Fischer and J. J. Rotemberg (Eds.), *NBER Macroeconomics Annual 1994, Volume 9*, pp. 81–140. MIT Press.

Leeper, E. M. and T. B. Walker (2011). Information flows and news driven business cycles. *Review of Economic Dynamics 14*(1), 55–71.

Leeper, E. M., T. B. Walker, and S.-C. S. Yang (2008). Fiscal Foresight: Analytics and Econometrics. Working paper 14028, National Bureau of Economic Research.

Leeper, E. M., T. B. Walker, and S.-C. S. Yang (2013). Fiscal Foresight And Information Flows. *Econometrica 81*(3), 1115–1145.

Lewis, V. and A. Stevens (2012). The Competition Effect in Business Cycles. mimeo.

Lippi, M. and L. Reichlin (1993). The Dynamic Effects of Aggregate Demand and Supply Disturbances: Comment. *American Economic Review 83*(3), 644–652.

Lippi, M. and L. Reichlin (1994). VAR analysis, nonfundamental representations, Blaschke matrices. *Journal of Econometrics 63*(1), 307–325.

Ljungqvist, L. and T. J. Sargent (2004). *Recursive Macroeconomic Theory* (2nd ed.). MIT Press.

Long, J. B. and C. I. Plosser (1983). Real Business Cycles. *Journal of Political Economy 91*(1), 39–69.

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

Lucke, B. (2013). Testing the technology interpretation of news shocks. *Applied Economics 45*(1), 1–13.

Lütkepohl, H. (1996). *Handbook of Matrices*. John Wiley & Sons.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer.

Matheson, T. and E. Stavrev (2014). News and monetary shocks at a high frequency: A simple approach. *Economics Letters 125*(2), 282–286.

McGrattan, E. R. (1994). The macroeconomic effects of distortionary taxation. *Journal of Monetary Economics 33*(3), 573–601.

McGrattan, E. R., R. Rogerson, and R. Wright (1997). An Equilibrium Model of the Business Cycle with Household Production and Fiscal Policy. *International Economic Review 38*(2), 267–290.

Mertens, K. and M. O. Ravn (2010). Measuring the Impact of Fiscal Policy in the Face of Anticipation: A Structural VAR Approach. *The Economic Journal 120*(540), 393–413.

Milani, F. (2012). Sentiment and the U.S. Business Cycle. mimeo.

Milani, F. and J. Treadwell (2012). The Effects of Monetary Policy 'News' and 'Surprises'. *Journal of Money, Credit and Banking 44*(8), 1667–1692.

Nelson, C. R. and C. I. Plosser (1982). Trends and random walks in macroeconmic time series: Some evidence and implications. *Journal of Monetary Economics 10*(2), 139–162.

Overschee, P. V. and B. D. Moor (1994). N4sid: Subspace Algorithms for the Identification of Combined Deterministic-Stochastic Systems. *Automatica 30*(1), 75–93.

Pappas, T., A. J. Laub, and N. R. Sandell (1980). On the Numerical Solution of the Discrete-Time Algebraic Riccati Equation. *IEEE Transactions On Automatic Control 25*(4), 631–641.

Pavlov, O. and M. Weder (2013). Countercyclical markups and news-driven business cycles. *Review of Economic Dynamics 16*(2), 371–382.

Pigou, A. C. (1927). *Industrial Fluctuations*. Macmillan and Co.

Portier, F. (2015). Comment on 'Whither News Shocks?'. In J. A. Parker and M. Woodford (Eds.), *NBER Macroeconomics Annual 2014, Volume 29*, pp. 265–278. National Bureau of Economic Research.

Poskitt, D. S. and W. Yao (2016). Vector Autoregressions and Macroeconomic Modeling: An Error Taxonomy. *Journal of Business and Economic Statistics*, forthcoming.

Ravenna, F. (2007). Vector autoregressions and reduced form representations of DSGE models. *Journal of Monetary Economics 54*(7), 2048–2064.

Rebelo, S. T. (2005). Real Business Cycle Models: Past, Present and Future. *The Scandinavian Journal of Economics 107*(2), 217–238.

Rozanov, Y. A. (1967). *Stationary Random Processes*. Holden-Day.

Ruge-Murcia, F. J. (2007). Methods to estimate dynamic stochastic general equilibrium models. *Journal of Economic Dynamics and Control 31*(8), 2599–2636.

Sadeghi, A. (2008). The births and deaths of business establishments in the United States. *Monthly Labor Review 131*(12), 3–18.

Sargent, T. J. (1987). *Macroeconomic Theory* (2nd ed.). Academic Press, Inc.

Sargent, T. J. (1989). Two Models of Measurements and the Investment Accelerator. *Journal of Political Economy 97*(2), 251–287.

Schmitt-Grohé, S. (2010). Comment on 'Letting Different Views about Business Cycles Compete'. In D. Acemoglu, K. Rogoff, and M. Woodford (Eds.), *NBER Macroeconomics Annual 2009, Volume 24*, pp. 475–489. National Bureau of Economic Research.

Schmitt-Grohé, S. and M. Uribe (2012). What's News in Business Cylces. *Econometrica 80*(6), 2733–2764.

Schwert, G. W. (1990). Stock Returns and Real Activity: A Century of Evidence. *Journal of Finance 45*(4), 1237–1257.

Seymen, A. (2013). Sequential Identification of Technological News Shocks. Discussion paper 13-111, Centre for European Economic Research.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica 48*(1), 1–48.

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

Sims, C. A., J. H. Stock, and M. W. Watson (1990). Inference in Linear Time Series Models With Some Unit Roots. *Econometrica 58*(1), 113–144.

Sims, E. R. (2012). News, Non-Invertibility, and Structural VARs. In N. Balke, F. Canova, F. Milani, and M. A. Wynne (Eds.), *DSGE Models in Macroeconomics: Estimation, Evaluation, and New Developments (Advances in Econometrics, Volume 28)*, pp. 81–135. Emerald Group Publishing Limited.

Sims, E. R. (2016). Differences in Quarterly Utilization-Adjusted TFP by Vintage, with an Application to News Shocks. Working paper 22154, National Bureau of Economic Research.

Sirbu, A.-I. (2013). News About Taxes and Expectations-Driven Business Cycles. mimeo.

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

Smets, F. and R. Wouters (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review 97*(3), 586–606.

Stock, J. H. and M. W. Watson (2005). Implications of Dynamic Factor Models for VAR Analysis. Working paper 11467, National Bureau of Economic Research.

Stock, J. H. and M. W. Watson (2010). Dynamic Factor Models. In M. P. Clements and D. F. Hendry (Eds.), *The Oxford Handbook of Economic Forecasting*, pp. 35–60. Oxford University Press.

Thorsrud, L. A. (2016). Words are the new numbers: A newsy coincident index of business cycles. Working paper 4/2016, Centre for Applied Macro- and Petroleum Economics.

Tovar, C. E. (2009). DSGE Models and Central Banks. *Economics: The Open-Access, Open-Assessment E-Journal 3*(16), 1–31.

Uhlig, H. (1999). A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily. In R. Marimon and A. Scott (Eds.), *Computational Methods for the Study of Dynamic Economies*, Chapter 3, pp. 30–61. Oxford University Press.

Uhlig, H. (2004). Do Technology Shocks Lead to a Fall in Total Hours Worked? *Journal of the European Economic Association 2*(2-3), 361–371.

Uhlig, H. (2005). What are the effects of monetary policy on output? Results from an agnostic identification procedure. *Journal of Monetary Economics 52*(2), 381–419.

Verhaegen, M. (1994). Identification of the Deterministic Part of MIMO State Space Models given in Innovations Form from Input-Output Data. *Automatica 30*(1), 61–74.

Viberg, M. (1995). Subspace-based Methods for the Identification of Linear Time-invariant Systems. *Automatica 31*(12), 1835–1851.

von Haberler, G. (1937). *Prosperity and Depression - A Theoretical Analysis of Cyclical Movements*. League of Nations.

Wagner, M. (2004). A Comparison of Johansen's, Bierens' and the Subspace Algorithm Method for Cointegration Analysis. *Oxford Bulletin of Economics and Statistics 66*(3), 399–424.

Wagner, M. (2010). Cointegration analysis with state space models. *Advances in Statistical Analysis 94*, 273–305.

Zeev, N. B. and H. Khan (2015). Investment-Specific News Shocks and U.S. Business Cycles. *Journal of Money, Credit and Banking 47*(7), 1443–1464.

Zellner, A. and F. Palm (1974). Time Series Analysis and Simultaneous Equation Econometric Models. *Journal of Econometrics 2*, 17–54.

Declarations

Eidesstattliche Versicherung

Hiermit erkläre ich, Thomas Haertel, an Eides statt, dass ich die Dissertation mit dem Titel "New Issues in Empirical Business Cycle Analysis: News Shocks and Structural Inference" selbständig und ohne fremde Hilfe verfasst habe. Andere als die von mir angegebenen Quellen und Hilfsmittel habe ich nicht benutzt. Die den herangezogenen Werken wörtlich oder sinngemäß entnommenen Stellen sind als solche gekennzeichnet. Es wurde keine kommerzielle Promotionsberatung in Anspruch genommen.

Ich habe mich anderweitig noch keiner Doktorprüfung unterzogen oder um Zulassung zu einer solchen beworben. Diese Dissertation hat noch keiner Fachvertreterin, keinem Fachvertreter und keinem Prüfungsausschuss einer anderen Hochschule vorgelegen; sie wurde nicht schon einmal in einem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

Ort/Datum: Frankfurt, 21. Mai 2016

Unterschrift: Thomas Haertel

* Gemäß § 6 Abs. 4 der Promotionsordnung der Fakultät Wirtschafts- und Sozialwissenschaften der Universität Hamburg vom 24. August 2010.