Essays on Structural Vector Autoregressions Identified Through Time-Varying Volatility

INAUGURAL-DISSERTATION

zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaft
doctor rerum politicarum
(Dr. rer. pol.)

am Fachbereich Wirtschaftswissenschaft
der Freien Universität Berlin

vorgelegt von
Thore Schlaak, M.Sc.
aus Wildeshausen

Berlin, 2019

Dekan:	Prof. Dr. Dieter Nautz
Erstgutachter:	Prof. Dr. Helmut Lütkepohl
	Professur für Methoden der empirischen Wirtschaftsforschung
	Freie Universität Berlin und DIW Berlin
Zweitgutachter:	Prof. Dr. Lars Winkelmann
	Juniorprofessur für Empirische Makroökonomie
	Freie Universität Berlin

Tag der Disputation: 25.11.2019

Acknowledgments

First and foremost, I am grateful to my supervisor, Helmut Lütkepohl, for his extraordinary and unconditional support: By giving me guidance, motivation, and feedback in the right amounts and moments to keep me on track during the past years. I thank him for dedicating so many hours to my questions and to discussions on research or related topics that shaped my understanding about economics, econometrics, and academic work. It was even earlier during inspiring lectures and seminars that he raised my interest in time series models.

I thank Lars Winkelmann for becoming my second supervisor and for the constructive comments on this dissertation and during seminars at FU Berlin.

I want to express my gratitude to my colleagues in the Forecasting and Economic Policy department of the DIW Berlin. Most of all, I thank Simon Junker who has had an open ear and good advice for many problems and concerns during numerous coffee breaks. I want to thank Ferdinand Fichtner and Claus Michelsen for supporting me way beyond this PhD project. Geraldine Dany-Knedlik, Marius Clemens, and Mathias Klein were available to discuss my progress/regress many times. Working with my co-authors, Maximilian Podstawski and Malte Rieth, was helpful in understanding how to solve the complex riddles of academic work, and I have to thank them for the always open doors.

The DIW Berlin Graduate Center and the persons that fill this institution with life have provided valuable support with many administrative challenges and made it easy to focus on research entirely.

The fellows of the 2014 cohort at the Graduate Center, many of whom became close friends, were an awesome group of people to be around since the beginning of our graduate studies. This holds even more for my office mates in room 4.3.003: Caterina Forti Grazzini, Daniel Bierbaumer, Martin Bruns, Pablo Anaya Longaric, and Sandra Pasch who accompanied me during long hours. Of course, my friend Stefan Etgeton stands out.

Without the support of Annette and Ulrich, this journey – that only had a vague destination at the beginning – would not have started over a decade ago. However, had I not had Johanna, I would never have finished it.

Berlin, September 2019 *Thore Schlaak*

Declaration of Co-Authorship

This dissertation consists of four (working) papers. Three papers were written in collaboration with one or more co-authors. The contribution in conception, implementation and drafting can be summarized as follows:

- Helmut Lütkepohl and Thore Schlaak:

  "Choosing Between Different Time-Varying Volatility Models for Structural Vector Autoregressive Analysis"

  Contribution: 50 percent

- Helmut Lütkepohl and Thore Schlaak:

  "Bootstrapping Impulse Responses of Structural Vector Autoregressive Models Identified through GARCH"

  Contribution: 50 percent

- Thore Schlaak

  "Disentangling the Effects of Uncertainty and Financial Shocks in Structural Vector Autoregressions"

  Contribution: 100 Prozent

- Malte Rieth, Thore Schlaak and Maximilian Podstawski

  "Monetary Policy, External Instruments and Heteroskedasticity"

  Contribution: 40 percent

List of Publications

Publications in Peer-Reviewed Journals

Lütkepohl, H. and Schlaak, T. (2018). Choosing between different time-varying volatility models for structural vector autoregressive analysis, *Oxford Bulletin of Economics and Statistics*, **80**(4): 715–735.
Publication of Chapter 1

Lütkepohl, H. and Schlaak, T. (2019). Bootstrapping impulse responses of structural vector autoregressive models identified through GARCH, *Journal of Economic Dynamics and Control*, **101**: 41–61.
Publication of Chapter 2

Working Paper

Rieth, M., Schlaak, T. and Podstawski, M. (2019). Monetary policy, external instruments and heteroskedasticity, *DIW Discussion Papers 1749*, German Institute for Economic Research, Berlin.
Publication of Chapter 4

Contents

List of Figures

List of Tables

List of Abbreviations

AIC	information criterion by Akaike (1974)
AR	autoregressive
ARCH	autoregressive conditional heteroskedasticity
BIC	Bayesian information criterion by Schwarz (1978)
BM	Bernanke and Mihov's (1998) monetary policy instrument
CB	estimation method conditioning on estimated GARCH parameters
$ComPI$	commodity price index
DGP	data generating process
EBP	Gilchrist and Zakrajšek's (2012) excess bond premium
EM	expectation maximization
Fed	Federal Reserve System
ff	effective nominal federal funds reate
FRED	Federal Reserve Economic Data
FU	Ludvigson et al.'s (2019) financial uncertainty index
GARCH	generalized autoregressive conditional heteroskedasticity
GK	Gertler and Karadi's (2015) monetary policy instrument
GO-GARCH	generalized orthogonal GARCH
H_0	null hypothesis
H_1	alternative hypothesis
H-VAR	heteroskedastic VAR
HQ	information criterion by Hannan and Quinn (1979)
i.e.	id est
i.i.d.	independent and identically distributed
ip	industrial production
LM	Lagrange multiplier
L/S	estimation method based on first step of Lanne and Saikkonen (2007) method
LR	likelihood ratio

MA	moving average
MBB	moving blocks bootstrap
ML	maximum likelihood
MS	Markov switching
MS-VAR	Markov switching in variances VAR
MSE	mean squared error
MU	Jurado et al.'s (2015) macroeconomic uncertainty index
NBER	National Bureau of Economic Research
NID	normally independently distributed
π	inflation rate
p	real price of oil
pp	percentage points
ppi	producer price index
$prod$	global crude oil production
q	index of real economic activity
r	interest rate
RBB	residual-based bootstrap
rr	Romer and Romer's (2004) monetary policy instrumental variable
RU	Ludvigson et al.'s (2019) real uncertainty index
S&P500	stock market index
SEM	simultaneous equation model
SQP	sequential quadratic programming
ST-VAR	smooth transition in variances VAR
SVAR	structural vector autoregressive
$uncert$	uncertainty indicator
US	United States
VAR	vector autoregressive
VIX	option implied volatility of S&P500 stock market index
WB	wild bootstrap
x	output

Summary

In order to employ vector autoregressions (VAR) for the analysis of causal relations between economic quantities – the underlying fundamental structure – researchers must overcome the omnipresent identification challenge. That means, the mutually uncorrelated structural shocks must be uncovered from the estimated reduced form residuals of the model. One possibility to achieve identification and to uncover structural innovations from the data is the use of statistical information extracted from time-varying volatility, which is present in many macroeconomic time series. This approach relies on a minimal set of identifying assumptions and is free of economically motivated restrictions. Chapters 1 and 2 of this dissertation concern model selection and inference in the context of models identified through time-varying volatility. Chapters 3 and 4 use this identification strategy to quantify the economic effects of different structural innovations and evaluate the compatibility of other identification approaches with the data.

The first chapter, which is joint work with Helmut Lütkepohl, assesses the performance of information criteria and tests for residual heteroskedasticity for choosing between different models for time-varying volatility. Although it can be difficult to find the true volatility model with the selection criteria, using them is recommended because they can reduce the mean squared error of impulse response estimates substantially relative to a model that is chosen arbitrarily based on the personal preferences of a researcher. Heteroskedasticity tests are found to be useful tools for deciding whether time-varying volatility is present but do not discriminate well between different types of volatility changes. The selection methods are illustrated by specifying a model for the global market for crude oil.

The second chapter, resulting from joint work with Helmut Lütkepohl, reviews and compares different bootstrap methods and estimation techniques for inference for structural vector autoregressive models identified by generalized autoregressive conditional heteroskedasticity (GARCH) in a Monte Carlo study. Three bootstraps are considered: a wild bootstrap, a moving blocks bootstrap, and a GARCH residual based bootstrap. Estimation is done by Gaussian maximum likelihood, a simplified procedure based on univariate GARCH estimations and a method that does not re-estimate the GARCH parameters in each bootstrap replication. The latter estimation strategy is computationally more efficient than the other methods while still being competitive with the other estimation approaches and often leads to the smallest confidence sets without sacrificing coverage precision. An empirical model for assessing monetary policy in the US is considered as an example. The different inference methods for impulse responses lead to qualitatively very similar results.

The third chapter, a single authored paper, assesses the interrelation of uncertainty and financial conditions and their impact on economic output in the US. Identification via heteroskedasticity offers a convincing alternative to conventional identification strategies in this context to uncover structural innovations because credible identifying assumptions based on economic mechanisms are hard to defend for the subject matter. Additionally, the use of the data-driven identification approach allows for formally testing linear restrictions imposed on structural parameters. This feature is employed to introduce a novel identification scheme using exclusion restrictions for different types of common uncertainty and financial shocks that is in line with the data. The causal dynamic analysis suggests that broad uncertainty shocks from different origins tighten financial conditions and the reverse is usually also true. Moreover, both, uncertainty and financial shocks are important drivers of real economic activity. However, quantitative effects depend on the specific type of uncertainty.

The fourth chapter, based on joint work with Maximilian Podstawski and Malte Rieth, proposes a framework to combine identifying information from time-varying volatility and external instruments for the quantification of US monetary policy shocks. Exploiting both types of information is shown to sharpen structural inference, and allows for testing both the relevance and exogeneity condition of instruments. Moreover, the proposed framework alleviates weak instruments problem from the proxy-VAR approach. Building on this novel framework, surprise monetary contractions are documented to lead to a significant and medium-sized decline in economic activity. Models with external instrument neglecting the identifying information in heteroskedasticity are less efficient and tend to underestimate the effects of monetary policy.

Zusammenfassung

Um mit Hilfe von Vektorautoregressionen kausale Zusammenhänge zwischen ökonomischen Größen – also ihre unterliegende fundamentale Struktur – zu analysieren, müssen ForscherInnen die Herausforderung der Identifizierung bewältigen. Dies bedeutet die wechselseitig unkorrelierten strukturellen Schocks aus den geschätzten Residuen der reduzierten Form zu extrahieren. Eine Möglichkeit zur Identifikation, und somit zur Offenlegung der strukturellen Innovationen, bietet die Nutzung statistischer Informationen aus zeitvariierender Volatilität, welche ein Merkmal vieler makroökonomischer Zeitreihen ist. Dabei werden nur wenige identifizierende Annahmen getroffen und keine ökonomisch motivierten Restriktionen benötigt. Kapitel 1 und 2 drehen sich um Selektion und Inferenz im Kontext von Modellen, die mit zeitvariierender Volatilität identifiziert werden. Kapitel 3 und 4 dieser Dissertation wenden diese Identifikationsstrategie an, um die ökonomischen Effekte verschiedener struktureller Innovationen zu quantifizieren und um die Vereinbarkeit alternativer Identifikationsstrategien mit den Daten zu überprüfen.

Das erste Kapitel, resultierend aus einer Zusammenarbeit mit Helmut Lütkepohl, untersucht die Fähigkeit von Informationskriterien und Tests auf Heteroskedastizität von Residuen zwischen verschiedenen Modellen für zeitvariierende Volatilität zu diskriminieren. Obwohl die Ermittlung des wahren Volatilitätsmodells anhand von Selektionskriterien schwierig sein kann, wird ihre Nutzung empfohlen, da sie den mittleren quadratischen Fehler von geschätzten Impulsantworten im Vergleich zu einer arbiträren Auswahl, etwa aufgrund der subjektiven Präferenz des Forschenden, substanziell reduzieren können. Heteroskedastizitätstests können nützlich sein, um die Existenz von zeitvariierender Volatilität zu verifizieren, sie diskriminieren jedoch nicht gut zwischen verschiedenen Arten von Volatilitätsänderungen. Die Selektionsmethoden werden anhand eines Modells für den globalen Rohölmarkt illustriert.

Das zweite Kapitel, entstanden in Zusammenarbeit mit Helmut Lütkepohl, bietet mittels einer Monte-Carlo-Studie einen Überblick und Vergleich verschiedener Bootstrap-Verfahren und Schätzmethoden für Inferenz bezüglich struktureller autoregressiver Modelle identifiziert durch verallgemeinerte autoregressive bedingte Heteroskedastizität (GARCH). In Betracht gezogen werden hierbei ein Wild-Bootstrap, ein Moving-Block-Bootstrap und ein Bootstrap basierend auf GARCH-Residuen. Die Schätzung wird vorgenommen anhand von Gaußscher Maximum-Likelihood-Methode, anhand einer vereinfachten Proze-

dur basierend auf univariaten GARCH-Schätzungen und mittels einer Methode, bei der die GARCH-Parameter nicht erneut in den einzelnen Bootstrapwiederholungen geschätzt werden. Letztere Methode ist weniger rechenintensiv, erweist sich jedoch als vergleichbar mit den anderen Schätzverfahren und führt oftmals zu den kleinsten Konfidenzbereichen, ohne dabei Überdeckungswahrscheinlichkeit einzubüßen. Als ein empirisches Beispiel wird ein Modell zur Analyse der US-Geldpolitik herangezogen. Dabei liefern die unterschiedlichen Inferenzmethoden für Impulsantworten qualitativ sehr ähnliche Ergebnisse.

Das dritte Kapitel untersucht das Zusammenwirken von Unsicherheits- und Finanzschocks und deren Einfluss auf ökonomische Aktivität in den USA. Identifikation mit Heteroskedastizität stellt in diesem Zusammenhang eine überzeugende Alternative zu konventionellen Identifikationsstrategien basierend auf Annahmen zu ökonomischen Mechanismen dar, um die strukturellen Innovationen zu ermitteln. Die Nutzung einer datengetriebenen Identifikation eröffnet zudem die Möglichkeit formaler Tests von linearen Restriktionen auf strukturellen Parametern. Dies wird genutzt, um ein neues Identifikationsschema basierend auf Nullrestriktionen für verschiedene Typen allgemeiner Unsicherheits- und Finanzschocks einzuführen, welches mit den Daten vereinbar ist. Die dynamische Kausalanalyse legt nahe, dass verschiedene allgemeine Unsicherheitsschocks die finanziellen Konditionen verschlechtern. Gleichermaßen führen negative Finanzschocks im Normalfall zu einem Anstieg von Unsicherheit. Darüber hinaus sind sowohl Unsicherheits- als auch Finanzschocks ein Treiber von realer ökonomischer Aktivität, wobei die genaue Größenordnung von der Art der Unsicherheit abhängt.

Das vierte Kapitel, welches auf einer Zusammenarbeit mit Maximilian Podstawski und Malte Rieth basiert, führt einen Modellrahmen ein, in dem identifizierende Information aus zeitvariierender Volatilität und externen Instrumenten zusammengeführt wird, um geldpolitische Schocks in den USA zu quantifizieren. Es wird gezeigt, dass die gemeinsame Nutzung beider Informationsquellen die strukturelle Inferenz schärft, Tests für Relevanz und Exogenität ermöglicht sowie das Problem schwacher Instrumente im Rahmen des Proxy-VAR-Ansatzes löst. Basierend auf dem vorgeschlagenen Modellrahmen wird dokumentiert, dass eine unerwartete geldpolitische Kontraktion einen signifikanten und mittelstarken Rückgang ökonomischer Aktivität induziert. Modelle, die auf externen Instrumenten basieren, jedoch Heteroskedastizität als Quelle identifizierender Information vernachlässigen, sind weniger effizient und tendieren zu einer Unterschätzung der Effekte von Geldpolitik.

Introduction and Overview

Structural vector autoregressions (SVAR) were introduced to modern empirical macroeconomic research in seminal work by Sims (1980). Since then, SVAR models have become an important tool for the analysis of causal relations between economic quantities. To establish causality, a major challenge is the identification of structural parameters that reflect the underlying fundamental structure of economic processes from the estimated reduced form of the model. More precisely, in the context of SVARs, the mutually uncorrelated and economically interpretable structural innovations must be extracted from the estimated contemporaneously correlated reduced form residuals.

Commonly, a linear relationship between reduced form residuals and structural innovations is assumed. Consequently, identification can be thought of as uniquely solving a system of non-linear equations that depend on the covariance matrix of reduced form residuals and a matrix that captures the contemporaneous endogenous variables' reactions on structural innovations, the instantaneous impact effects matrix. Finding an unique solution to the system of equations that is economically sensible is challenging, because typically the number of structural parameters of interest is higher than the number of parameters that can be inferred from the symmetric reduced form covariance matrix without further identifying restrictions. That means, to assure a solution that is economically meaningful, further *plausible* identifying restrictions are needed to fulfill the order condition that is necessary for identification.

Numerous approaches to overcome this challenge in structural vector autoregressions have been developed in the academcic literature. One popular strategy is to impose restrictions that are motivated by assumptions on the underlying structural relations, e.g., their timing or signs. Such restrictions are most often derived from economic theory but could also be motivated by institutional knowledge or specific ways of construction of the data. Examples building on such strategies are linear restrictions on the instantaneous impact effects matrix or on the immediate relations between the endogenous variables. In this context, commonly used devices are exclusion restrictions (see, e.g., Blanchard and Perotti, 2002), including the widely used recursive pattern (see, e.g., Sims, 1980; Christiano et al., 1999; Bloom, 2009). Another possibility is to impose restrictions on the long-run effects of structural impulse responses (see, e.g., Blanchard and Quah, 1989; King et al., 1991).

In some situations, these types of linear restrictions on the structural parameters are hard to defend and other strategies must be applied for credible identification. For example, Stock and Watson (2012) and Mertens and Ravn (2013) propose the use of instrumental variables. Other approaches identify admissible ranges of the structural parameters by assumptions on the direction of (on-impact) reactions of endogenous variables to structural shocks (see, e.g, Canova and De Nicoló, 2002; Uhlig, 2005) or by narrative arguments (see, e.g., Antolín-Díaz and Rubio-Ramírez, 2018; Ludvigson et al., 2019).

This dissertation – in contrast to the aforementioned identification approaches – builds on alternative models that exploit identifying information exclusively from statistical properties, specifically the second moments, of the data. That means, instead of depending on assumptions derived from theoretical models or other (economic) background knowledge, identifying information is extracted from time-varying volatility that is commonly present in many macroeconomic time series. This feature may be helpful in situations when no or contradictory theoretical motivation for conventional restrictions exists or external instruments are hard to find. Similarly to identification via volatility, other data-driven identification approaches exist, that, for example, exploit non-Gaussian reduced form residuals (see, e.g., Lanne and Lütkepohl, 2010; Gouriéroux and Monfort, 2014; Lanne et al., 2017).

Intuitively explained, heteroskedasticity alleviates identification because the time-varying structure of the reduced form residuals' covariance matrix offers additional parameters for the estimation of underlying mutually uncorrelated innovations. A (locally) unique solution for the structural parameters can be estimated through changes in volatility under fairly parsimonious assumptions. First, as common to all linear SVARs, identification depends on time-invariant impulse responses to unit structural shocks. Second, the mutually uncorrelated innovations must exhibit (distinct) changes of their variances or conditional variances.

As opposed to conventional identification approaches, statistically identified shock series need to be put into economic context by the researcher to have a structural interpretation. Conveniently, however, identification via heteroskedasticity can be combined with many types of conventional identification restrictions that can be additionally imposed on structural parameters, thus facilitating the labeling of the statistically identified shocks. At the same time, the additional restrictions are over-identifying, which opens up the possibility for formal tests of conventional identification schemes.

Sentana and Fiorentini (2001) and Rigobon (2003) establish conditions for identification of structural innovations from the reduced form residuals in different contexts. The former assume the presence of autoregressive conditional heteroskedasticity for the class of factor models, whereas the latter introduce a vector autoregressive framework with exogenous breaks of the variance structure. Since then, other models were introduced to exploit (conditional) heteroskedasticity generated by different processes for identification: Normandin and Phaneuf (2004) develop a SVAR framework with generalized-orthogonal GARCH innovations, Lanne et al. (2010) and Herwartz and Lütkepohl (2014) introduce time-varying volatility that is governed by a discrete Markov process, Lütkepohl and Netšunajev (2017b) link the transition between two volatility regimes to a particular transition variable, whereas Bertsche and Braun (2017) model the variance via a stochastic process.

The first two chapters of the dissertation are based on large-scale Monte Carlo simulation studies to address research questions related to model selection and inference in finite samples that arise in applied work with alternative models identified through heteroskedasticity. Chapters 3 and 4 are empirical papers that assess causal economic relations through identification via heteroskedasticity.

The abundance of different volatility models for identification of structural parameters, as laid out above, motivates the first chapter of this dissertation entitled *Choosing between Different Time-Varying Volatility Models for Structural Vector Autoregressive Analysis*, which is joint work with Helmut Lütkepohl. Usually in applied work no information of the true process that generated the data is available and, hence, it is unclear which model is best suited for a given volatility pattern. We employ a Monte Carlo study to generate data based on different volatility processes and fit a set of four different volatility models to the artificial data. We compare the ability of different procedures, three standard information criteria and tests for (conditional) residual heteroskedasticity, to find the model that actually generated the data. Furthermore, we investigate the impact of model selection on the structural parameters by applying the mean squared error of impulse response function as metric for evaluation. We find that tests for heteroskedasticity are useful for detecting time-varying volatility but not for deciding on a specific model. Concerning the selection criteria, overall, the Akaike information criterion (Akaike, 1974) has a slight advantage over the other criteria in selecting the correct model, even though fairly large sample sizes might be needed for some types of models. Further, models selected by this criterion tend to provide impulse response estimates with relatively small mean squared errors. Generally, using any information criterion is better than choosing the volatility model arbitrarily. A selection strategy is proposed and illustrated by means of an empirical example for the global market for crude oil from Kilian (2009).

The first chapter contributes to the existing literature as it fills the gap of a systematic investigation of model selection in the context of structural VAR models with time-varying volatility that is used to support the identification of structural shocks. Further, the study demonstrates the consequences of model selection for estimates of structural impulse response functions, which are of special interest for the subject matter. Moreover, a two step procedure to conduct model selection is proposed.

The second chapter of the dissertation entitled *Bootstrapping impulse responses of structural vector autoregressive models identified through GARCH*, which is joint work with Helmut Lütkepohl, is similar in its methodological spirit. It is also based on an extensive Monte Carlo simulation. The study focuses on inference for structural impulse response functions in SVAR models where structural parameters are fully identified via the conditional heteroskedasticity of GARCH innovations.

We explore and compare the small sample suitability of alternative bootstrap methods for inference on structural impulse responses. We include a recursive-design wild bootstrap, a recursive-design residual-based moving blocks bootstrap, and a GARCH residual-based bootstrap in the comparison. Further, as full maximum likelihood (ML) estimation is computationally demanding, we also evaluate different estimation methods for the bootstrap algorithms. Additionally to ML, we consider an estimation method that is proposed as a first step in a Gaussian ML procedure by Lanne and Saikkonen (2007), as well as a method that does not re-estimate the GARCH parameters in each bootstrap replication.

We find that the relative coverage frequencies for the impulse responses are quite heterogeneous. Bootstrap and estimation methods designed for more precise estimation of the GARCH structure have no advantages for the coverage precision of the confidence intervals and confidence bands. In fact, the methods that condition on the first round ML estimates of the GARCH parameters in the bootstrap tend to result in smaller intervals and bands with similar coverage properties, even though the latter might still be well below the nominal level. The most accurate coverage is obtained if such a conditional approach is combined with a wild bootstrap.

We use the alternative bootstrap procedures and estimation methods to assess the effects of monetary policy shocks in the United States based on a benchmark study by Caldara and Herbst (2016) and document only minor differences between the alternative setups that do not affect the qualitative results.

The chapter contributes to the literature by exploring inference methods related to impulse responses based on SVARs where identification is obtained via the conditional heteroskedasticity of GARCH innovations. In contrast to previous studies, we investigate inference for the case where identification is obtained via conditional heteroskedasticity, which is why estimating the second moment structure well is of particular importance.

The third chapter of the dissertation, a single authored paper entitled *Disentangling the Effects of Uncertainty and Financial Shocks in Structural Vector Autoregressions*, is an empirical study that exploits the advantages of identification via heteroskedasticity in the context of uncertainty and financial shocks. Both of these structural innovations are important drivers of the business cycle and their consequences should not be assessed independently because of their close interdependencies and comovement. The latter observations illustrate the difficulty that arise when empirically distinguishing the two innovations in SVAR models using conventional identification approaches or external instruments. I employ a Markov switching mechanism to identify both shocks via time-varying volatility. Heteroskedasticity is conceptually well suited for the subject matter as uncertainty is often related to variance changes.

I include three measures of common uncertainty in the study to account for the consequences of different types of uncertainty separately as suggested in many academic papers. Specifically, macroeconomic uncertainty by Jurado et al. (2015), financial uncertainty, and real uncertainty by Ludvigson et al. (2019) are considered. They are constructed using the same methodology, which makes them directly comparable.

In the analysis, changes in the volatility pattern of the data are shown to be a useful tool for identifying uncertainty and financial shocks. I review exclusion restrictions that are imposed on the instantaneous impact effects matrix of the structural shocks in the related literature and extend the restrictions to a novel identification scheme. I do not find evidence against imposing specific exclusion restrictions and use the restricted model in a causal dynamic analysis.

I document that financial uncertainty is a highly relevant exogenous driver of conditions in the financial sector, while real and macroeconomic uncertainty are less important. With respect to real activity, adverse innovations of macroeconomic and financial uncertainty cause more pronounced declines of output compared to innovations in real uncertainty.

The contributions of the paper are twofold. First, departing from a model that exclusively relies on identification through heteroskedasticity, I show that the pass through of common uncertainty shocks to financial conditions and that of financial shocks to common uncertainty is not instantaneous. This finding is exploited in a novel identification scheme that relies on short-run exclusion restrictions that is not rejected by the data. Second, I systematically review and document the differences in the causal effects of three different types of common uncertainty shocks by investigating their interactions with financial shocks as well as their impact on real activity.

The fourth chapter of the dissertation, entitled *Monetary Policy, External Instruments and Heteroskedasticity*, resulting from joint work with Maximilian Podstawski and Malte Rieth, focuses on the combination of identification via heteroskedasticity with identifying information from external instruments. We propose a framework that combines both sources of information in order to improve identification within SVARs and to address some of the limitations that each of the two identification approaches has in isolation.

We show by means of a Monte Carlo simulation study, that combining the two sources of identifying information improves the estimation precision of structural impulse responses. Moreover, in our proposed framework testing the identifying assumption of instruments, the relevance and exogeneity condition, breaks down to simple tests of zero restrictions on the instantaneous impact matrix of the structural innovations. Moreover, the framework dispenses inference problems arising from weak instruments, as present in a traditional proxy-VAR framework. The reason being that, in the presence of sufficient heteroskedasticity in the data, the model is fully identified through the properties of the data's second moments. Hence, the inclusion of a weak instrument does not create problems for inference. At the same time, incorporating valid instruments into a model that otherwise only relies on identification via heteroskedasticity is useful because the instrumented shock is given a structural interpretation by the economic background information of the external instrument.

We apply our framework to conduct a structural analysis of US monetary policy. We find that a model that solely relies on identifying information of the narrative instrument of Romer and Romer (2004) and ignores the identifying information in the data's changes of volatility, underestimates the consequences of contractionary monetary policy shocks on real activity.

We contribute to the literature in three ways. First, the encompassing framework opens up to use a combination of different sources of identifying information, thereby improving the identification of the structural model. Second, our framework allows for testing the validity, that is, exogeneity and relevance, of an instrument. In particular the former problem has been largely unresolved for instrumental variables in SVAR models. Third, our framework also simplifies the economic interpretation of the instrumented shock as it is pinned down by prior economic reasoning stemming from the instrument.

CHAPTER 1

Choosing Between Different Time-Varying Volatility Models for Structural Vector Autoregressive Analysis*

Published as:

Lütkepohl, H. and Schlaak, T. (2018). Choosing between different time-varying volatility models for structural vector autoregressive analysis, *Oxford Bulletin of Economics and Statistics* 80(4): 715–735.

For copyright reasons, this chapter is not included in the online version of the dissertation. An electronic version of the article can be accessed at https://doi.org/10.1111/obes.12238.

*This chapter is joint work with Helmut Lütkepohl.

CHAPTER 2

Bootstrapping Impulse Responses of Structural Vector Autoregressive Models Identified through GARCH*

Published as:

Lütkepohl, H. and Schlaak, T. (2019). Bootstrapping impulse responses of structural vector autoregressive models identified through GARCH, *Journal of Economic Dynamics and Control* 101: 41–61.

For copyright reasons, this chapter is not included in the online version of the dissertation. An electronic version of the article can be accessed at https://doi.org/10.1016/j.jedc.2019.01.008.

*This chapter is joint work with Helmut Lütkepohl.

CHAPTER 3

Disentangling the Effects of Uncertainty and Financial Shocks in Structural Vector Autoregressions

3.1 Introduction

Uncertainty – a catch-all term broadly defined as firms' and households' inability to predict the future – has been shown to be a considerable source of business cycle fluctuations. Since the seminal publication of Bloom (2009), macroeconomic research is devoting increasing attention to the concept.[1] Similarly, as financial markets not only played a decisive role causing the turmoil of the Great Recession, but also triggered and amplified the decline of real economic activity, the role of financial shocks is now a prime focus of macroeconomimc research.[2]

However, uncertainty and financial tightness do not evolve independently of each other; rather, they usually strongly comove. For example, theoretical work suggests that higher uncertainty can be related to tighter credit conditions (Arellano et al., forthcoming; Christiano et al., 2014) and financial markets serve as amplifier for the effects of uncertainty shocks (Alfaro et al., 2018) or even constitute an important link in the transmission mechanism of uncertainty to the real economy (Gilchrist et al., 2014). Against this backdrop, understanding the interrelations of financial and uncertainty shocks is crucial for evaluating their macroeconomic impact in isolation. However, empirically distinguishing between innovations in uncertainty and financial conditions is intricate. For this reason, little empirical work exists to ascertain whether uncertainty arises as a consequence of financial perturbations or if movements in uncertainty cause financial distress (or both) and how innovations in both quantities affect output.

In this paper, I exploit time-varying volatility of monthly US data to disentangle financial shocks and uncertainty shocks to uncover their causal relations in structural vector

[1] Bloom (2014) and Castelnuovo et al. (2017) provide reviews of the literature on uncertainty.

[2] A non-exhaustive list of empirical papers on the economic effects of financial shocks from various sources includes Hristov et al. (2012), Gilchrist and Zakrajšek (2012), Meeks (2012), Peersman (2012), and Gambetti and Musso (2017).

autoregressions (SVAR). I rely on this data-driven identification approach, which depends on statistical properties of the variables' second moments, because it overcomes the drawbacks of many conventional identification schemes for the subject matter: Short-run exclusion restrictions inhibit contemporaneous feedback of variables to shocks that cannot be ruled out *a priori*, the directions of on-impact reactions of many key economic variables to financial and uncertainty shocks are identical and complicate the use of sign-restrictions, and credible exogenous instruments are hard to find.

I estimate a Markov switching (MS) in variances model to capture the data's (conditional) heteroskedasticity. Exploiting second moment properties for identification is based on work of Rigobon (2003) and Rigobon and Sack (2003). The MS framework in this context was introduced by Lanne et al. (2010) and is applied in several empirical macroeconomic studies (Herwartz and Lütkepohl, 2014; Lütkepohl and Velinov, 2016; Podstawski and Velinov, 2018). Chen and Netšunajev (2018) and Netšunajev and Glass (2017) employ it for the analysis of uncertainty shocks.[3] However, these studies do not focus on the interdependencies between financial conditions and uncertainty.

Identification via heteroskedasticity is conceptually well suited for the subject matter, as uncertainty is often related to changes in volatility patterns (Bachmann et al., 2013; Rossi and Sekhposyan, 2015; Ludvigson et al., 2019; Jo and Sekkel, 2019). Hence, it seems natural to incorporate the data's heteroskedasticity in the analysis. From the methodological viewpoint, the setup overcomes many challenges of the joint identification of uncertainty and financial shocks. First, it forgoes any (economic) restrictions on variables' impact reactions to structural shocks and generally allows for contemporaneous feedback between all variables while providing point-estimates of structural impulse responses. Second, as opposed to identification via external instruments or the penalty function approach that have been used in this context (Stock and Watson, 2012; Caldara et al., 2016), it is possible to uncover all underlying structural disturbances simultaneously. Third, exploiting the second moment structure of the data for identification opens up the possibility of formally testing restrictions imposed on structural parameters.

In this context, for example, a recursive identification scheme based on short-run exclusion restrictions that empirical studies rely upon can be assessed. It is applied by Popescu and Smets (2010) for German data of different frequencies and by Gilchrist et al. (2014) for quarterly US data. The former document that perturbations in uncertainty trigger a small effect on financial risk premia, but uncertainty largely remains unaffected by financial disturbances. In contrast, Gilchrist et al. (2014) conclude that uncertainty increases endogenously to changes in credit conditions and that financial frictions reinforce the

[3]Angelini et al. (2017) examine time-varying impact effects of uncertainty shocks using a related model that exploits changes in volatility between pre-specified regimes for identification.

impact of uncertainty shocks. Further empirical work to quantify the impact of shocks to uncertainty and financial conditions along with their interactions is based on alternative identification approaches: Furlanetto et al. (2019) identify credit and uncertainty innovations via sign restrictions on the direction of shocks' instantaneous impact effect ratios in a model with quarterly US data. They find that credit shocks do not trigger much fluctuation in uncertainty, whereas uncertainty drives down financial conditions. A closely related paper to mine is Caldara et al. (2016), which employs the penalty function approach of Faust (1998) and Uhlig (2005). It documents that financial market conditions are an important transmission channel of uncertainty shocks. However, reversing the order of the optimization problems that need to be solved for identification has a first-order effect on the empirical findings; thus, the interrelation of financial and uncertainty shocks remains inconclusive.

To put my empirical results on a broad basis, I pay special attention to exploring the consequences of different types of uncertainty separately as suggested in many academic papers (see, e.g., Ludvigson et al., 2019; Caldara et al., 2016; Angelini et al., 2017). Therefore, I focus this study on three broad uncertainty measures – macroeconomic uncertainty (as by Jurado et al. (2015)), financial uncertainty, and real uncertainty (these last two as by Ludvigson et al. (2019)). All indicators are constructed using the same methodology, thus allowing me to systematically compare their impact on financial conditions and vice versa. The measures are based on a data rich environment and seek to quantify common uncertainty of distinct origin as proposed in a range of theoretical papers (Bloom, 2009; Arellano et al., forthcoming; Gilchrist et al., 2014). The use of broad uncertainty is motivated by the consideration that many uncertainty proxies rely on a limited information set, potentially capturing very specific types of uncertainty that inhibit general conclusions.[4]

Turning to my empirical key findings, first, I revisit exclusion restrictions on the structural impact effects that have been used to disentangle uncertainty and financial shocks in related studies (see Popescu and Smets, 2010; Gilchrist et al., 2014). I add to the existing literature by proposing a novel identification scheme that cannot be rejected by the data for all three uncertainty measures. In contrast to conventional views, I find that broad uncertainty shocks originating from different sources do not trigger an instantaneous reaction of financial conditions and that the reverse also holds true.

Second, structural impulse response analysis based on the novel identification scheme stresses the importance of differentiating between origins of uncertainty: Real and macroeconomic uncertainty can be regarded as cause, but also as consequence, of financial dis-

[4]For example, the widely used VIX, the option-implied volatility of the US S&P500 stock market index, might be well suited to capture stock price uncertainty but not general financial market uncertainty.

tress. In contrast, I find that financial uncertainty is a highly relevant exogenous driver of conditions in the financial sector.

Third, touching upon the debate on the causal directions between uncertainty and recessions (see, e.g., Ludvigson et al., 2019; Angelini et al., 2017; Carriero et al., 2018), I document that adverse innovations of macroeconomic and financial uncertainty cause more pronounced declines in output compared to innovations in real uncertainty. In turn, shocks to production induce a significant decrease in macroeconomic and real uncertainty, whereas financial uncertainty can be considered exogenous to innovations in real activity. The former result is in stark contrast to findings of Ludvigson et al. (2019), who document that higher macroeconomic uncertainty triggers output growth, as suggested by the "growth-options" channel of uncertainty.

The remainder of the paper is organized as follows: Section 3.2 introduces the VAR with volatility changes modeled by Markov-Switching. Section 3.3 discusses the data, model specification, and empirical results. The last section concludes. The appendix contains supplementary results.

3.2 The SVAR Framework

The data generating process is assumed to follow a reduced form vector autoregressive process of order p (VAR(p)) of the form

$$y_t = \Pi X_t' + u_t, \tag{3.1}$$

where $y_t = (y_{1t}, \dots, y_{Kt})'$ is a $(K \times 1)$-vector of observable variables, $\Pi = [\gamma, A_1, \dots, A_p]$, with γ being a vector of constant terms and A_i, for $i = 1, \dots, p$, are $(K \times K)$ autoregressive coefficient matrices. $X_t = [1, y_{t-1}', \dots, y_{t-p}']$ contains a one to capture the intercept term and the lagged endogenous variables. The u_t constitute the K-dimensional reduced form residuals of the model. To construct a mapping from the reduced form VAR to its structural representation, the existence of a linear relationship

$$u_t = B\varepsilon_t \tag{3.2}$$

is assumed, where ε_t are the structural innovations. The matrix B is often termed the (instantaneous) impact effect matrix, as it captures the contemporaneous impact of the structural shocks on the observed variables y_t in the system (3.1).

Furthermore, the VAR-model is assumed to be stable, which implies that the polynomial

$$\det(I_K - A_1 z - \dots - A_p z^p)$$

has no roots inside and on the complex unit circle. This assumption assures the existence of the Wold moving average representation $y_t = \mu + \sum_{i=0}^{\infty} \Phi_i u_{t-i}$, where μ is the unconditional mean of y_t and Φ_i are the Wold moving average coefficient matrices of the VAR.

It is well documented that many macroeconomic time series display time-varying volatility (see, e.g., Stock and Watson, 2002; Gertler and Karadi, 2015) and many concepts of uncertainty are even related to changes in dispersion (Bachmann et al., 2013; Rossi and Sekhposyan, 2015; Carriero et al., 2016; Ludvigson et al., 2019). This observation is reflected in the assumption that the residuals u_t are assumed to be (conditionally) heteroskedastic yet serially uncorrelated. More specifically, the time-varying volatility is assumed to follow a discrete first order Markov switching process s_t that may take values $1, \ldots, M$, hence M is the number of different Markov states. The transition probabilities between MS states are defined as $p_{ij} = \Pr(s_t = j | s_{t-1} = i)$ $(i, j = 1, \ldots, M)$, and $u_t | s_t \sim \text{i.i.d.}(0, \Sigma_m)$ for $m = 1, \ldots M$.

The MS switching in variances framework is appealing because, first, the underlying volatility states are endogenously determined such that they do not have to be assumed, statistically determined, or linked to a certain transition variable as necessary in other models for the variance process.[5] Second, the MS model somewhat shields against a potential misspecification of the volatility process as it was found to be flexible to capture also variance changes from other data generating processes (Lütkepohl and Schlaak, 2018).[6]

One advantage of explicitly modeling heteroskedasticity is that changes in volatility can be exploited as a source of identifying information to overcome the well-known identification problem of homoskedastic structural VARs.

The identification problem is apparent in (3.2): Under the assumption of a constant instantaneous impact matrix B and homoskedastic residuals u_t, i.e., one volatility regime, the reduced form covariance matrix, is symmetric and only provides $K \times (K+1)/2$ distinct parameters, while B contains $K \times K$ unknowns. Therefore, additional restrictions (e.g., short-run exclusion restrictions on the impact effects or long-run restrictions on structural impulse responses) are needed for identification.

In contrast, time-varying volatility can be used for identification, if a decomposition of the regime dependent reduced form covariance matrices into

$$\Sigma_m = B \Lambda_m B' \quad \text{for} \quad m = 1, ..., M, \tag{3.3}$$

[5]For alternative volatility models see Rigobon (2003), Normandin and Phaneuf (2004), Lanne et al. (2010), Lütkepohl and Netšunajev (2017b), and Bertsche and Braun (2017).

[6]To that end, I also conduct an extensive model selection procedure in Subsection 3.3.2.

exists. In the above, $\Lambda_m = \text{diag}(\lambda_{1m}, \ldots, \lambda_{Km})$ for $m = 2, \ldots, M$ is a diagonal matrix whose strictly positive elements can be interpreted as variance change of the k^{th} structural innovation ε_k relative to the first regime Λ_1, which is normalized to I_K without loss of generality.

The state-invariant B matrix is locally – i.e., up to permutations of its columns and multiplication of the columns by -1 – unique if at least one regime exists in which the relative change in volatility with respect to the first regime of variable l is different from the relative change in volatility of variable k for $k, l = 1, \ldots, K$ (see proof in Lanne et al., 2010). The assumption of state-invariance of B implies a constant impact of a unit shock and is imposed in many studies using VAR models to scrutinize the effect of uncertainty shocks (see, e.g., Bloom, 2014; Bachmann et al., 2013; Ludvigson et al., 2019).

Under the assumption of sufficient heteroskedasticity for identification, the parameters of B can be estimated by maximum likelihood. Additional linear restrictions imposed on elements of B, e.g., zero contemporaneous restrictions, become over-identifying and can be tested by standard Likelihood-Ratio (LR) tests (Lanne and Lütkepohl, 2008; Lanne et al., 2010; Lütkepohl and Netšunajev, 2017b). Conveniently, in the MS framework for setups with $M \geq 3$, the assumption of a state-invariant B matrix also becomes empirically testable via LR-tests (see results in Subsection 3.3.3).

Estimation

For the estimation of the MS-VAR, an adaptation of the expectations maximization (EM) algorithm by Krolzig (1997) is used. It was proposed by Herwartz and Lütkepohl (2014) in the context of structural identification via volatility changes. It features state-dependent covariance matrices while all other parameters are regime-invariant. Standard errors for all estimated coefficients are obtained by evaluating the inverse of the negative Hessian matrix of the likelihood function at its optimum.

For inference on the structural impulse response functions, bootstrapped pointwise confidence bands are computed. Given the heteroskedastic pattern of the data, a simple reshuffling of the residuals u_t, as in a classic residual bootstrap, does not preserve the properties of the second moments of the data and potentially invalidates inference. In this context, many empirical studies rely on a wild bootstrap combined with a Rademacher distribution to preserve the data's volatility pattern (see, e.g., Herwartz and Lütkepohl, 2014; Netšunajev, 2013; Lütkepohl and Netšunajev, 2014; Podstawski and Velinov, 2018). In this paper, bootstrap samples are based on a recursive-design wild bootstrap, which performed well for a model with conditional heteroskedasticity that is driven by GARCH processes (see Lütkepohl and Schlaak, 2019). The bootstrap samples are constructed as

$$y_t^* = \hat{\Pi} X_t^{*\prime} + \varphi_t \hat{u}_t, \quad t = 1, 2, \ldots, T, \tag{3.4}$$

where $\hat{\Pi}$ are the estimated (autoregressive) coefficients and $\hat{u}_t$ are the estimated reduced form residuals from the model in (3.1). As suggested by Brüggemann et al. (2016), $\varphi_t \sim N(0,1)$ is an independent random variable with mean zero and unit variance. Thus, using draws of φ_t to transform the estimated reduced form residuals induces variability to the bootstrapped samples but preserves the data's original second moment structure. With $\varphi_t \hat{u}_t$ for $t = 1, 2, \ldots, T$ at hand, y_t^*, the bootstrapped counterpart of y_t, is constructed recursively using (3.4). Specifically, $X_t^* = [1, y_{t-1}^{*\prime}, \ldots, y_{t-p}^{*\prime}]$ contains the lagged bootstrapped values of y_t^*. Each bootstrap sample is based on identical pre-sample values from the original data set as initial values, i.e., $y_{-p+1}^* = y_{-p+1}, \ldots, y_0^* = y_0$.

The bootstrap is conducted conditionally on estimated parameters for the relative variances $\hat{\Lambda}_m$ and transition probabilities $\hat{p}_{ij}$, $i, j = 1, \ldots, M$. Since estimates of B in the bootstrap procedure are only identified up to sign, the shock that exerts the highest initial impact in absolute value on variable 1 (i.e., largest element in absolute terms of the first row of B) is normalized to be positive and all signs within the respective column of (bootstrapped) B are adjusted accordingly. This procedure is repeated for the remaining variables (i.e., the corresponding rows of (bootstrapped) B), where any columns of (bootstrapped) B that have been adjusted in a previous step are excluded from the comparison. The bootstrapped (pointwise) confidence bands are based on 5,000 bootstrap repetitions.

3.3 Empirical Analysis of Uncertainty and Financial Shocks

In this section, the MS-VAR model is taken to the data using trivariate setups for three different uncertainty measures. The first subsection describes the data with a focus on different uncertainty measures. Subsection 3.3.2 verifies the presence and suitability of heteroskedasticity as viable tool for identification. In Subsection 3.3.3 and 3.3.4 an extensive structural analysis, on the basis of hypothesis tests and impulse response functions, is undertaken to shed light on dynamic causal relations between innovations in financial conditions and uncertainty.

3.3.1 Data

The objective of this paper is to examine the interaction of uncertainty and financial conditions. Focusing on uncertainty, a variety of measures for uncertainty are available from the literature reflecting the fact that uncertainty is an opaque concept and quantification

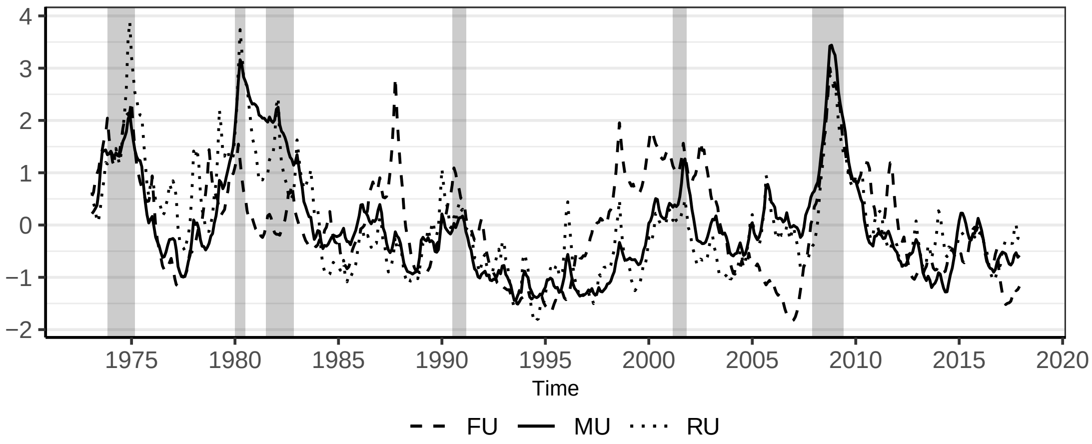


Figure 3.1: Different common uncertainty measures.

Notes: The figure shows the plot of macroeconomic uncertainty (MU – solid line) by Jurado et al. (2015), as well as financial uncertainty (FU – dashed line), and real uncertainty (RU – dotted line) by Ludvigson et al. (2019). The shaded vertical bars mark recession periods defined by the NBER. All measures are standardized after taking the natural logarithm.

is demanding. There are two reasons underlying the decision to conduct the analysis based on macroeconomic uncertainty by Jurado et al. (2015) as well as financial uncertainty and real uncertainty by Ludvigson et al. (2019) as uncertainty indicators. First, all measures can be directly related to a specific source of uncertainty by construction and yet still aim at capturing uncertainty on a broad basis since they are based on high dimensional models. Second, the three measures are based on an identical methodology which makes them directly comparable.

Macroeconomic uncertainty Macroeconomic uncertainty (MU) consists of a mixture of real, price, and financial variables to incorporate various potential sources of uncertainty. Following Jurado et al. (2015), it is supposed to capture uncertainty in the broadest possible sense, that is, in "many economic indicators at the same time, across firms, sectors, markets, and geographic regions."

In order to do so, Jurado et al. (2015) approximate uncertainty as the common variation of the unforecastable component of the development of broad economic conditions. Technically speaking, the authors proxy uncertainty by the aggregated weighted conditional volatility of one-step ahead forecast errors of many macroeconomic indicators. First, Jurado et al. (2015) define the one-step ahead uncertainty, $\mathcal{U}^z_{jt}(1)$ $(j = 1, \ldots, N_{MU})$, of variables z_{jt} of the data set $Z^{MU}_t = (z_{1t}, \ldots, z_{N_{MU}t})'$ as conditional volatility of the unforecastable component of future values of that series. Formally,

$$\mathcal{U}^z_{jt}(1) \equiv \sqrt{\mathbb{E}\Big[(z_{jt+1} - \mathbb{E}\big[z_{jt+1}|I_t\big])^2|I_t\Big]}, \quad j = 1, 2, \dots, N_{MU}, \tag{3.5}$$

where the expectation is taken with respect to the available information set I_t at time t. Second, the measure of macroeconomic uncertainty is then simply constructed as average over individual uncertainty series at each point in time

$$MU_t(1) \equiv \frac{1}{N_{MU}} \sum_{j=1}^{N_{MU}} \mathcal{U}^z_{jt}(1). \tag{3.6}$$

To compute a comprehensive measure for I_t, in practice, Jurado et al. (2015) fit a factor model to a large data set. It is then used to calculate the conditional forecasts of each individual uncertainty series.

In total, macroeconomic uncertainty is based on the one-step ahead forecast errors of $N_{MU} = 134$ time series. For example, the data set Z^{MU}_t contains real output and income, employment and hours, as well as trade, sales, retail, and order statistics. Furthermore, a smaller number of financial series, like commodity and price indices, a small number of bond and stock market indices, and foreign exchange rates are included. The measure is constructed without imposing structure derived from specific theoretical models and, instead of relying on a limited number of variables, it captures aggregate uncertainty in a data rich environment (Jurado et al., 2015). It usually spikes around recessionary phases of the business cycle (see solid line in Figure 3.1).

Financial market uncertainty The concept of financial uncertainty is receiving special attention since the 2007-2009 financial crisis. In the wake of the Great Recession, uncertainty on financial markets has been identified as a potential cause and propagation mechanism for fluctuations of the business cycle (Arellano et al., forthcoming; Christiano et al., 2014; Gilchrist et al., 2014). Ludvigson et al. (2019) propose a measure – financial uncertainty (FU) – of this type of uncertainty that relies on a broad set of variables, as opposed to a single variable like the VIX, for example, to capture financial market developments. It is constructed using an identical methodological framework as Jurado et al. (2015), which is sketched in (3.5) and (3.6). For the construction of financial uncertainty, the conditional volatilities of one-step ahead forecast errors of individual series $\tilde{z}_{jt}$ $(j = 1, \dots, N_{FU})$ from data set $Z^{FU}_t = (\tilde{z}_{1t}, \dots, \tilde{z}_{N_{FU}t})'$ are used. The data Z^{FU}_t exclusively consists of financial variables. It includes time series like valuation ratios, growth rates of dividends, yields of corporate bonds for different rating grades, default spreads, term spreads, as well as a variety of equity returns. Thus, the measure is based on the common conditional volatility of forecast errors from $N_{FU} = 148$ monthly financial time series. It mostly spikes around major financial events like Black Monday in 1987, the LTCM-crisis in 1998, the Enron scandal in 2001, and the 2008 credit crunch (see dashed line in Figure 3.1). However, financial uncertainty is also elevated in recessionary times.

Real uncertainty As Jurado et al.'s (2015) macroeconomic uncertainty measure contains a mixture of real and financial variables, Ludvigson et al. (2019) develop a measure specifically designed to isolate uncertainty from real fundamentals – the real uncertainty indicator (RU). Again, the measure is based on the conditional volatility of the unforecastable component of times series $\hat{z}_{jt}$ ($j = 1, \ldots, N_{RU}$) from the data set $Z_t^{RU} = (\hat{z}_{1t}, \ldots, \hat{z}_{N_{RU}t})'$ that is calculated using (3.5). Z_t^{RU} is a subset of Z_t^{MU} and contains only real quantities that are also used to construct macroeconomic uncertainty. As before, the average of the conditional volatilities of the $N^{RU} = 73$ individual real uncertainty series constitutes the broad real uncertainty measure. As RU is a sub-index of the baseline indicator for macroeconomic uncertainty, it displays a strong comovement with the broader uncertainty series (see the dotted line in Figure 3.1). Thus, it is an interesting exercise to check how much of the fluctuations in Jurado et al.'s (2015) macroeconomic uncertainty can be related to uncertainty in real quantities. This is especially relevant as many theoretical papers suggest that uncertainty arises from economic fundamentals rather than from prices or financial variables (see, e.g., Baker et al., 2016).

Financial conditions and real activity The spread between corporate interest rates and government bonds as risk free assets is used to assess financial conditions. In particular, I rely on Gilchrist and Zakrajšek's (2012) excess bond premium, which is, broadly speaking, the spread between private and public debt net of borrowers' predictable default risk (see plot of the data in Figure 3.6 in Appendix 3.A). More specifically, the variable serves as proxy for the price demanded by investors to bear corporate non-financial credit risk above and beyond the compensation for firm specific fundamentals, in particular expected default costs of the issuer. Thus, the excess bond premium measures the market price of risk freed from the amount of risk inherent in each individual asset. Consequently, an increase of the excess bond premium reflects a decline of risk-bearing capacity of investors resulting in a reduction of lending. The excess bond premium has strong forecasting abilities for real activity and is used in many empirical studies to approximate credit costs (see, e.g. Gilchrist and Zakrajšek, 2012; Gertler and Karadi, 2015; Adrian et al., 2013; Gertler and Kiyotaki, 2015).

To control for real activity in the empirical model, I consider industrial production (see plot of the data in Figure 3.6 in Appendix 3.A). As opposed to GDP, the variable

Table 3.1: Cross-Correlation of Uncertainty Indicators, Excess Bond Premium, and Industrial Production

	ebp	Δip-MA(3)	MU	FU	RU
MU	0.51	-0.53	1	–	–
FU	0.58	-0.34	0.53	1	–
RU	0.39	-0.48	0.87	0.58	1

Notes: The table shows the contemporaneous cross-correlations of the three uncertainty measures and the correlation of the uncertainty measures with the excess bond premium and a 3-month moving average of the growth rate of industrial production.

is available monthly and, even though production captures only about 20% of US GDP, it is highly responsive to economic change; hence it is a good proxy for business cycle fluctuations.

Correlation analysis Table 3.1 displays the contemporaneous cross-correlations of the three uncertainty measures as well as correlations with the excess bond premium and with a 3-month moving average of industrial production growth (Δip-MA(3)) to illustrate the relation between the variables.

The excess bond premium displays the highest correlation with financial uncertainty, expressing the close relatedness of the uncertainty indicator to financial market developments. The excess bond premium and real uncertainty also comove quite closely, but the correlation is smallest among all uncertainty proxies, suggesting that uncertainty about real developments might play only a minor role for the price setting of financial markets in the short run. The differences between the correlations of the uncertainty measures illustrate the importance of considering different sources of uncertainty when assessing their impact on financial market developments.

For industrial production, the contemporaneous correlation increases (in absolute terms) if real uncertainty is augmented with information from financial variables. This follows from the fact that macroeconomic uncertainty displays the strongest comovement (negative correlation of -0.53) of all uncertainty proxies. The finding suggests that fluctuation of uncertainty from financial sources potentially affects the business cycle.

As expected, among the uncertainty indicators, macroeconomic and real uncertainty clearly comove most strongly, whereas macroeconomic and financial uncertainty only exhibit a correlation of 0.58, stressing the differences between the two types of uncertainty.

3.3.2 Specification of Empirical Model

The model consists of monthly US data for three endogenous variables $y_t = [\Delta ip_t, ebp_t, uncert_t]'$, where Δip_t is the first difference of the logarithm of industrial production (multiplied by 100), ebp_t is the excess bond premium and $uncert_t$ refers to the logarithm of one of the three uncertainty indicators described above. The model is a condensed version of a setup considered by Caldara et al. (2016) but focuses on variables of prime interest – especially alternative types of uncertainty – for the analysis due to the computational complexity of the maximization of the likelihood function in the EM algorithm.[7] The estimation sample runs from January 1973 to December 2017, consisting of $T = 539$ observations, where the start of sample is determined by the availability of the excess bond premium. All calculations are based on a lag length of $p = 6$. This choice is motivated by a reasonable trade-off between an over-parametrization of the model and capturing all relevant dynamics of the variables.[8] All variables enter the model standardized to ensure a balanced scaling of the covariance matrices in the EM algorithm.

Table 3.2: Model Selection

Reduced form	log l	AIC	HQ	BIC
		MU		
Linear VAR(6)	-505.30	1136.61	1242.08	1406.15
MS(2)-VAR(6)	-364.82	871.65	990.52	1175.42
MS(3)-VAR(6)	-319.97	795.94	926.53	1129.66
MS(4)-VAR(6)	-303.48	780.97	926.63	1153.20
		FU		
Linear VAR(6)	-674.69	1475.39	1580.87	1744.93
MS(2)-VAR(6)	-467.20	1076.41	1195.28	1380.18
MS(3)-VAR(6)	-422.53	1001.01	1131.66	1334.79
MS(4)-VAR(6)	-399.13	972.27	1117.93	1344.50
		RU		
Linear VAR(6)	-840.14	1806.28	1911.76	2075.83
MS(2)-VAR(6)	-678.06	1498.13	1617.00	1801.90
MS(3)-VAR(6)	-619.01	1394.02	1524.61	1727.74
MS(4)-VAR(6)	-599.79	1373.59	1519.25	1745.82

Notes: log l denotes the log likelihood function evaluated at the optimum, AIC $= -2(\log\, l) + 2f$, HQ $= -2(\log\, l) + 2f \times \log(\log(T))$ and BIC $= -2(\log\, l) + \log(T)f$ where f is the number of free parameters and T the number of observations.

[7] Caldara et al.'s (2016) model consists of ten endogenous variables including employment, personal consumption, a price deflator of personal consumption, short- and long-term interest rates, stock market returns, and a commodity price index.

[8] Standard lag selection criteria even suggest to use shorter lag lengths.

Naturally, a prerequisite of identifying structural shocks using time-varying volatility is the presence of (conditional) heteroskedasticity in the data. In this context, Lütkepohl and Schlaak (2018) propose standard ARCH-tests as a reliable tool to detect conditional heteroskedasticity. The results for the tests on the residuals of linear VAR(6) models with the respective uncertainty indicator clearly indicate the presence of conditional heteroskedasticity (see Table 3.6 in Appendix 3.B). This finding is confirmed by standard model selection criteria that are found to be helpful tools for discriminating between models with time-varying volatility by Lütkepohl and Schlaak (2018). Results for model selection are displayed in Table 3.2 and show that linear, homoskedastic VAR(6) models are not favored by any of the standard selection criteria when compared to different MS-models.[9]

Next, the number of MS states must be specified. Models restricted to $M = 2$ states result in a considerable loss of likelihood and are not preferred by any selection criteria compared to models with more states. For $M = 3$ and $M = 4$, the criteria do not unanimously favor a particular model. While the BIC criterion tends to MS-models with $M = 3$ states, there is also evidence for four volatility states. However, in the following, results are based on $M = 3$ for two reasons: As argued below, MS(3)-VAR models capture all relevant transitions and different regimes of the US economy during the sample period, such that a fourth state does not provide much value added from the economic perspective. Additionally, the number of observations for some states of the MS(4)-VARs is very low, which, in consequence, leads to an unreliable estimation precision.

As noted in Section 3.2, for identification via heteroskedasticity the decomposition in (3.3) relies on a state-invariant impact effects matrix B. This restriction reduces the number of parameters that need to be estimated compared to a reduced form MS-model for $M > 2$. Hence, the data could speak up against the assumption of regime-invariance of B, which can be tested empirically by a conventional LR-test (see, e.g., Herwartz and Lütkepohl, 2014). The p-values of such tests for the three models with different uncertainty indicators are well above conventional significance levels (see Table 3.7 in Appendix 3.B) lending support to the assumption of a constant impact effects matrix.

In order to identify structural shocks by changes in their volatility pattern, their relative variances need to vary sufficiently (for visual inspection of the volatility pattern see the plots of the estimated reduced form residuals in Figure 3.7 in Appendix 3.B). Table 3.3 provides the estimated relative variances and the associated standard errors for the three structural MS(3)-VARs. Since the order of the diagonal elements of Λ_m is arbitrary, they

[9] Other volatility models assuming GARCH-type errors (see, Lütkepohl and Milunovich, 2016) and a smooth transition in variances model with time as transition variable (see, Lütkepohl and Netšunajev, 2017b) were also fitted to the data. These models are favored by all selection criteria over the linear homoskedastic VARs but were inferior to all MS-models.

are ordered from smallest to largest for the second state. Given this convention, the elements of Λ_3 are sorted accordingly.

For macroeconomic and real uncertainty, the heterogeneity of the elements of Λ_3 points toward statistical identification as the confidence intervals constructed using one standard deviation around the respective point estimates do not overlap.[10] Moreover, standard errors are quite small, indicating that estimation precision is reasonable. The differences between λ_{3m} and λ_{im} for $i = 1, 2$ clearly exceed two standard deviations in regimes $m = 2, 3$ for both models. For financial uncertainty, the heterogeneity of the estimated coefficients of Λ_2 clearly lends support to the hypothesis of identification with heteroskedasticity, given the large difference between estimates even if the standard errors are taken into consideration. As the empirical evidence points toward a uniquely identified instantaneous impact effects matrix B, the causal relations between uncertainty and financial conditions can be uncovered by all three MS(3)-VAR(6) models.

Table 3.3: Estimates and Standard Errors of Relative Variances

Param.	Estimate	St. Err.	Param.	Estimate	St. Err.	Param.	Estimate	St. Err.
Macro Uncertainty			Financial Uncertainty			Real Uncertainty		
λ_{12}	0.96	0.13	λ_{12}	0.91	0.12	λ_{12}	0.76	0.10
λ_{22}	1.10	0.15	λ_{22}	3.21	0.41	λ_{22}	0.93	0.13
λ_{32}	11.33	1.50	λ_{32}	10.32	1.30	λ_{32}	12.64	1.65
λ_{13}	8.08	1.63	λ_{13}	9.52	2.49	λ_{12}	5.33	1.17
λ_{23}	4.38	0.89	λ_{23}	31.76	8.38	λ_{22}	9.57	2.09
λ_{33}	28.71	5.75	λ_{33}	30.53	8.31	λ_{32}	32.98	7.15

Notes: The standard errors are obtained from the inverse of the negative Hessian evaluated at the optimum of the likelihood function of MS(3)-VAR(6) with $y_t = [\Delta ip_t, ebp_t, uncert_t]'$ and different uncertainty indicators for $uncert_t$.

To cross-check whether economic developments in the US are plausibly reflected by the estimated MS-states, the state dependent reduced form covariances Σ_m (see Table 3.8 in Appendix 3.B) and estimated MS-state probabilities (see Figure 3.2) are investigated.

The first state, with residual covariance Σ_1 is labeled as low volatility state. This choice is primarily motivated by the residual variances of the excess bond premium that are markedly smaller than in the other states. The residual variances of industrial production, and also that of real uncertainty, are somewhat smaller in the second state. Turning to the estimated state probabilities, the state is predominant during the 1990s (see top panel in Figure 3.2). This period is usually referred to as the Great Moderation because business

[10]For the class of MS-models, currently no formal statistical tests for identification are available. As the model under the null hypothesis is potentially not identified, the derivation of the asymptotic distribution of Wald- or LR-tests is not straightforward. For this reason, in the existing literature, usually the point estimates and standard errors of the respective elements of Λ_m are considered when checking for identification (Lütkepohl and Netšunajev, 2017a; Chen and Netšunajev, 2018).

cycle fluctuations and financial market swings abated compared to the previous decades. During the 2000s, the low volatility state is associated with some periods of economic recovery, for example the boom period of 2005/2006.

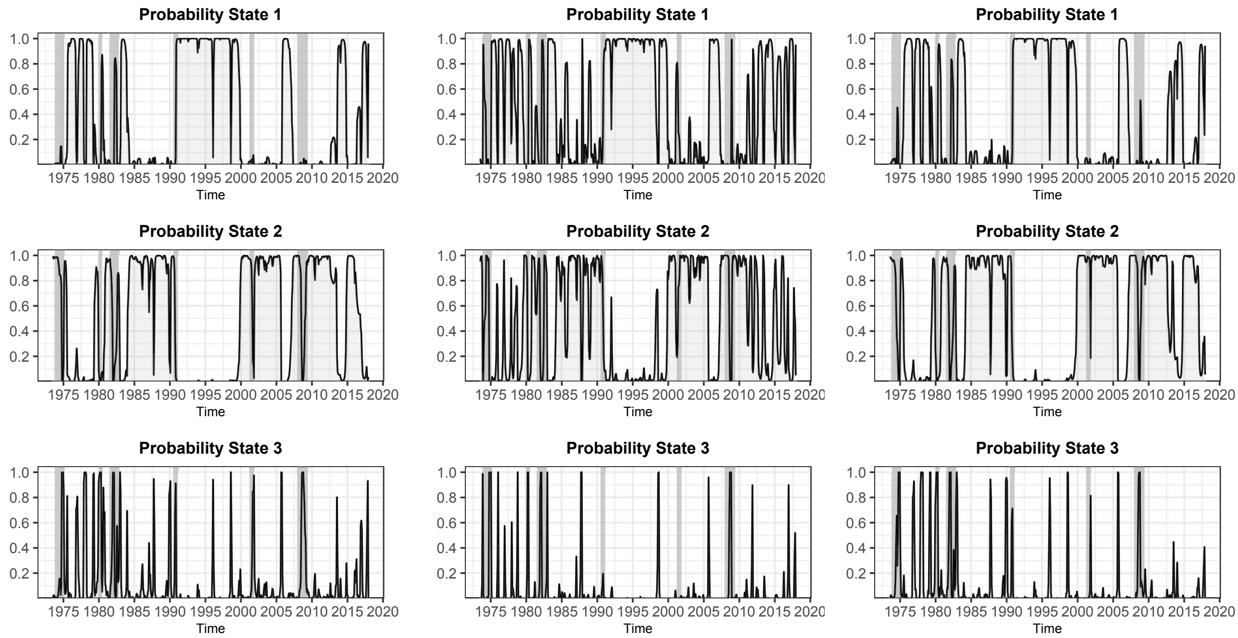
The second regime is linked to the residual covariance Σ_2 and can be interpreted as medium volatility state. Again, this label is justified by the residual variances of the excess bond premium that is clearly below the volatility of the third states but significantly larger than in the first states. Looking at the estimated probabilities supports this interpretation (see middle panel in Figure 3.2): The state is predominant as of the mid-1980s, roughly corresponding with the end of the Great Inflation period that is associated with large swings in output, inflation and, hence, uncertainty. Then again, after the Great Moderation during long periods of the 2000s, the second state prevails except for distinctive times of severe recession or marked boom periods.

The volatility of all residuals is highest in the third state (linked to the largest (co-)variances in absolute terms in Σ_3). This volatility regime is most related to the years of the Great Inflation at the beginning of the sample. Later, it is associated with relatively few observations that usually coincide with times of recession or (economic) crises[11], both of which would naturally be associated with high volatility (see bottom panel in Figure 3.2). All-in-all, this cautious narrative analysis reveals that the endogenously determined states of the covariance coincide with many relevant economic developments in the US. It is reassuring that the states for all three models are quite similar, as only the highly correlated uncertainty indicators are exchanged in the different setups.

3.3.3 Tests of Exclusion Restrictions on Impact Effects

One immense advantage of using data properties for identification of structural vector autoregressions is that additional restrictions imposed on the instantaneous impact effect matrix B can be tested against the data. Under the assumption of full identification of the structural shocks via heteroskedasticity, further restrictions on B reduce the parameter space and can be evaluated against an unrestricted model by means of LR-tests, as mentioned in Section 3.2. I exploit this feature to inspect controversial restrictions – partly used for identification – from the literature to formally investigate structure imposed on the instantaneous effects matrix. The structure of matrix B is crucial for causal dynamic analysis as it (i) can be exploited for identification; and (ii) is key for the shape

[11]For example events like Black Monday in October 1987, the start of the Gulf War in October 1990, the LTCM default in August 1998, and Hurricane Katrina in September 2005 are captured. The terrorist attacks in September 2001 are only captured in the models based on macroeconomic and real uncertainty.



(a) Macro Uncertainty (b) Financial Uncertainty (c) Real Uncertainty

Figure 3.2: Smoothed state probabilities.

Notes: The figure shows the estimated smoothed state probabilities for state 1 in the upper panel and for state 2 in the middle panel and state 3 in the lower panel for MS(3)-VAR(6) with $y_t = [\Delta ip_t, ebp_t, uncert_t]'$ and different uncertainty indicators for $uncert_t$. The shaded vertical bars mark recession periods defined by the NBER.

of structural impulse response functions, i.e., for the transmission of the shocks through the system.

Before turning to the relation of uncertainty and financial conditions, first, I examine the reaction of real activity to innovations in uncertainty and financial conditions. The former is a key question in the related literature and is debated in many papers (see, e.g., Bloom, 2009; Bachmann et al., 2013; Angelini and Fanelli, 2018; Piffer and Podstawski, 2018; Ludvigson et al., 2019). In this context, as outlined above, the instantaneous reaction of real activity to uncertainty shocks plays an important role and only few papers in the literature (e.g., Ludvigson et al., 2019; Chen and Netšunajev, 2018) formally investigate it. Against this backdrop, I test whether the impact effect for various types of uncertainty shocks on real activity is significantly different from zero.

Along the same lines, I scrutinize the instantaneous impact of real activity to a financial shock that – as a type of nominal disturbance – is traditionally restricted to zero in the literature (see, e.g., Bernanke and Blinder, 1992; Christiano et al., 1999; Peersman, 2012).

Both restrictions are jointly tested by placing two zeros in the first row of B, visualized as

$$B_1 = \begin{bmatrix} * & 0 & 0 \\ * & * & * \\ * & * & * \end{bmatrix}. \tag{3.7}$$

Identical restrictions are applied for identification in Gilchrist et al. (2014) for quarterly US data and by Popescu and Smets (2010) for quarterly and monthly German data without formally testing them. In a similar methodological framework to mine, Chen and Netšunajev (2018) for quarterly US data find that a recursive identification that restricts the contemporaneous reaction of real activity to uncertainty perturbations is rejected by the data. However, these findings do not necessarily apply to monthly data as the timing restrictions imposed when using monthly data are less strict by construction. Ludvigson et al. (2019), for a monthly data set including macroeconomic and financial uncertainty as well as industrial production, document that, under their identifying assumptions the impact of uncertainty shocks on real activity is positively bounded away from zero, implying invalidity of a zero-instantaneous effect.

The first panel of Table 3.4 displays the results of testing B_1 against an unrestricted structural MS-VAR. For none of models the null hypothesis of no reaction in real production within a month to exogenous movements in uncertainty or financial conditions can be falsified as all p-values are well above conventional significance levels.

Next, I focus on investigating the reaction of financial conditions to an uncertainty shock. Empirical studies based on structural vector autoregressions that have jointly identified uncertainty and financial shocks via recursive exclusion restrictions must weigh between shutting off the contemporaneous impact of uncertainty shocks on financial conditions or the reverse (Popescu and Smets, 2010; Gilchrist et al., 2014), even though both channels could be at play. I begin with testing the former and check whether financial conditions instantaneously react to an uncertainty shock.

This assumption is reflected placing a zero restriction in second row of the third column B, which is imposed additionally to the restrictions from B_1 such that the restrictions amount to

$$B_2 = \begin{bmatrix} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix}. \tag{3.8}$$

The imposed restrictions represent a Cholesky factorization of B and the model would be identified without heteroskedasticity in a conventional linear VAR.

The results of the tests of the recursive scheme B_2 against B_1 under the alternative are displayed in the second panel of Table 3.4. All p-values are far from conventional significance levels, implying that the hypothesis of no immediate reaction of financial conditions after an uncertainty shock is in line with the data. Thus, there is no evidence of an immediate increase in pricing of credit risk within the same month of an unanticipated shift in uncertainty.

Lastly, I shut off the immediate reaction of common uncertainty to a financial shock. As noted above, restricting the immediate impact of uncertainty to a financial shock is also considered by Popescu and Smets (2010) and Gilchrist et al. (2014). In contrast to these authors, instead of an alternative ordering of the variables in a recursive setup, I propose shutting off the immediate transmission of financial shocks to uncertainty *additionally* to the recursive scheme. This setup is not considered in the literature and amounts to the interesting case where the matrix B is over-identified by the short-run zero restrictions imposed. So far, the previous LR-tests for B_1 and B_2 hinged on the assumption of full identification of B via changes in heteroskedasticity, as thoroughly investigated in Subsection 3.3.2. Even given the solid empirical evidence that the changes in volatility at least carry some identifying information, it cannot be completely ruled out that the non-rejection of the previous tests is due to a lack of identification from heteroskedasticity. As pointed out by Lütkepohl and Netšunajev (2017a), in case of an unidentified model under the null, the asymptotic distribution may not be standard and an incorrect number of degrees of freedom may be applied for the LR-tests. That said, scheme B_3 can, in principle, be rejected by the data when evaluated against B_2 even without any identifying information from heteroskedasticity and, thus, constitutes a valuable double-check to shield against low power of the LR-tests in the previous setups. The restriction of no immediate reaction of uncertainty to a financial shock is visualized as

$$B_3 = \begin{bmatrix} * & 0 & 0 \\ * & * & 0 \\ * & 0 & * \end{bmatrix}, \tag{3.9}$$

where again the restriction is imposed additionally to scheme B_2.

The third panel of Table 3.4 shows results for tests of B_3 against B_2. The minimal p-value for the different uncertainty indicators is 0.36, i.e., the null is not rejected at conventional significance levels. This lends strong support to the hypothesis that uncertainty does not respond instantaneously to a financial shock and reassures that the findings of previous tests do not result from low power due to a lack of identification via heteroskedasticity. The last panel of Table 3.4 shows the results of a joint test of all zero restrictions imposed on B_3 against the alternative of the structural MS-VAR model identified by heteroskedasticity. The p-values for all uncertainty measures back the previous finding that the restrictions imposed on identification B_3 are in line with the data.

Table 3.4: LR-Tests of Different Restriction Schemes

H_0	H_1	uncertainty measure	p-value (LR-statistic)	df
		MU	0.66 (0.66)	
B_1	state-invariant B	FU	0.71 (0.68)	2
		RU	0.57 (1.12)	
		MU	0.87 (0.03)	
B_2	B_1	FU	0.64 (0.22)	1
		RU	0.96 (0.002)	
		MU	0.67 (0.18)	
B_3	B_2	FU	0.36 (0.83)	1
		RU	0.96 (0.003)	
		MU	0.93 (0.87)	
B_3	state-invariant B	FU	0.78 (1.74)	4
		RU	0.89 (1.13)	

Notes: The table shows the p-values with respective LR-statistics in parentheses for different restriction schemes B_1-B_3 of the instantaneous impact effects matrix B of MS(3)-VAR(6) models with $y_t = [\Delta ip_t, ebp_t, uncert_t]'$ and different uncertainty indicators for $uncert_t$.

Additional support for the findings of the LR-tests of this subsection stems from the estimated coefficients of the elements of the unrestricted B matrix (see Table 3.9 in Appendix 3.B): None of the coefficients restricted in scheme B_3 significantly differs from zero if two standard errors of the respective point estimates are considered.

Using the same model specification, I also conduct the various LR-tests using the VIX as uncertainty measure which is one of the most widely used uncertainty indicators (Bloom, 2009; Piffer and Podstawski, 2018; Furlanetto et al., 2019). Table 3.10 in Appendix 3.B reports the test results. The hypotheses of no instantaneous reaction of financial conditions on uncertainty is rejected at the 5% significance level and the reverse is rejected well below the 1% significance level. These findings can be interpreted as further evidence for the special characteristics of the common uncertainty measures.

Taken together, the analysis reveals that the common perception of broad uncertainty measures as fast-moving variables does not hold with respect to changes in financial conditions. Likewise, innovations to common uncertainty do not have immediate consequences for real and financial indicators. Both of these findings might be related to the complexity of the concept of common uncertainty that results in delayed transmissions from and to the respective indicators. These findings can facilitate the identification of structural innovations when working with common uncertainty, as standard short-run exclusion restrictions may be applied for identification.

Since the restrictions imposed by B_3 are compatible with the data, they are implemented in the subsequent structural analysis of Subsection 3.3.4.

3.3.4 Impulse Response Analysis

Structural impulse response analysis is conducted to analyze the dynamic quantitative effects of uncertainty and financial shocks. Imposing identification scheme B_3 comes with the advantage that column permutations of B are ruled out. Therefore, labels can be attached more easily to the identified shocks. Accordingly, the first shock in each model is labeled as a real activity shock. As such, it subsumes all exogenous variation in real activity independent of its fundamental origin and allows for an instantaneous reaction of the excess bond premium and the uncertainty indicator. For example, supply and demand shocks constitute candidate structural shocks that are not disentangled here. The second and the third shocks are labeled as financial and uncertainty shocks in each model. This label is motivated by the assumption that both structural shocks trigger an instantaneous reaction in their respective proxies but not in other endogenous variables.

For comparability, in Figures 3.3-3.5 the same shocks from each model are depicted in columns (MU - left column, FU - middle column, RU - right column) and accordingly the responses of the endogenous variables are displayed in rows. The shaded bands constitute the area between the bootstrapped pointwise 95 % confidence bands. The responses of real activity are always cumulated.

Uncertainty shocks: The impacts of adverse uncertainty shocks from different origins on the models' endogenous variables are displayed in Figure 3.3. The responses are scaled to one standard deviation uncertainty shocks in the first volatility state of each MS-model. The different uncertainty shocks trigger hump-shaped responses for all three uncertainty measures that are very persistent and significant at the 95% significance level.

Financial conditions tighten as investors demand higher compensation for taking risk after adverse uncertainty innovations. How much the excess bond premium increases depends on the type of uncertainty: In response to perturbations of financial uncertainty, firms' financing costs rise persistently for more than 36 months with a peak effect of four

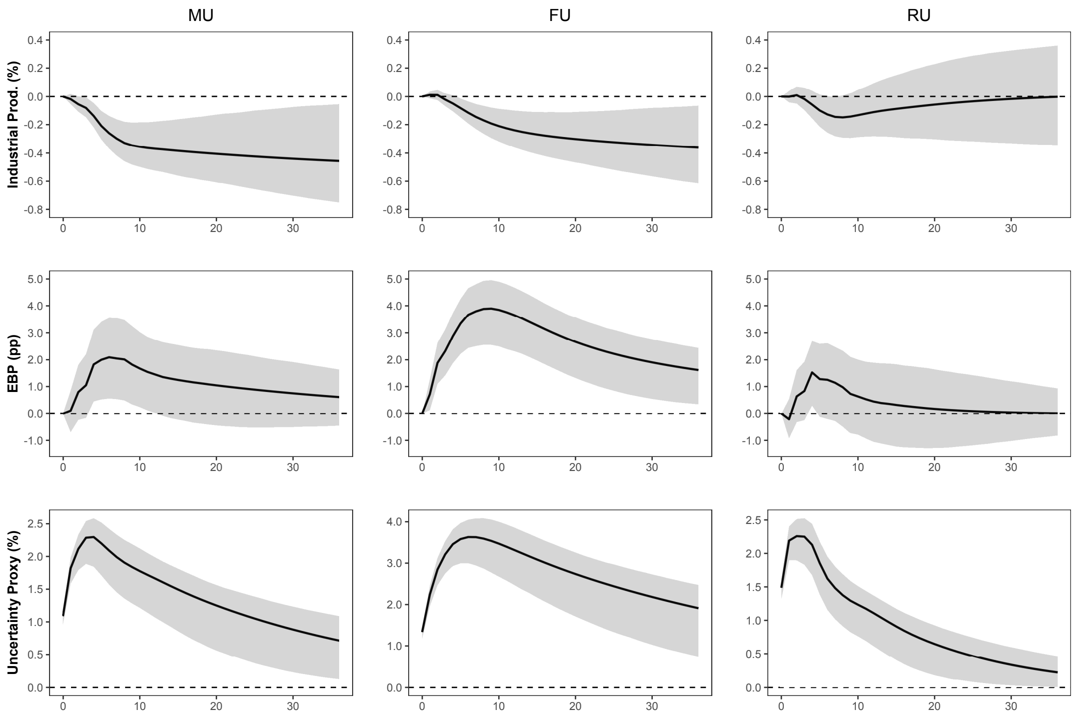


Figure 3.3: Impulse response functions of uncertainty shocks of models with different common uncertainty measures.

Notes: The figure shows impulse response functions of uncertainty shocks in MS(3)-VAR(6) models with $y_t = [\Delta ip_t, ebp_t, uncert_t]'$ and different uncertainty indicators for $uncert_t$. Constraints are imposed on the B-matrix according to the results of Subsection 3.3.3. The impulse responses for Δip_t are cumulated. Scaling of the impulse response functions is according to a one standard deviation uncertainty shock in the first volatility state. Shaded bands are bootstrapped 95% pointwise confidence intervals.

basis points after eight months. This peak effect is roughly more than twice as high compared to the reaction of financial conditions to a real uncertainty shock, which is only marginally significant at the 95% confidence level. Thus, for financial uncertainty, the risk premium channel is an important transmission channel (see, e.g., Arellano et al., forthcoming; Gilchrist et al., 2014), while for real uncertainty this is much less true.

Compared to results of Caldara et al.'s (2016) baseline identification, I find macroeconomic uncertainty causes more sluggish and less strong fluctuations in financing costs, even though the effects are qualitatively comparable. However, my results are at odds with the significant initial decline in financing costs that the authors document for many uncertainty measures under their alternative identification scheme. Compared to the results of a set identified model used by Furlanetto et al. (2019), the implications of uncertainty on financing costs are more pronounced, especially of financial uncertainty, which stresses the importance of uncertainty as one key driver of financial developments. These results also hold if the contemporaneous transmission of the uncertainty shocks to the financial market is not restricted (see Figure 3.8 in Appendix 3.B).

Turning to real effects of uncertainty shocks, as a consequence of exogenous variation in macroeconomic and financial uncertainty, economic activity decelerates quickly. For both shocks, the decrease amounts to about 0.4% of industrial production before stabilizing after roughly one year. Real uncertainty, in contrast, causes a decline of only 0.15% in real activity before reverting to zero. Yet, when confidence bands are taken into consideration, this effect is, if anything, only marginally significant. Thus, the decline in production caused by macroeconomic uncertainty is rather caused by financial factors than by real economic fundamentals.

Generally, the negative reaction of real activity to adverse surprises in uncertainty is consistent with the theoretical argument of a wait-and-see attitude of firms stressed by Bloom (2009) as well as with the empirical findings of Bacchiocchi and Fanelli (2015) and Carriero et al. (2018), who use macroeconomic uncertainty as uncertainty proxy in vector autoregressive frameworks. However, for macroeconomic uncertainty, my findings are in stark contrast to findings of Ludvigson et al. (2019), who report an increase of real activity as consequence of exogenous movements of macroeconomic uncertainty.

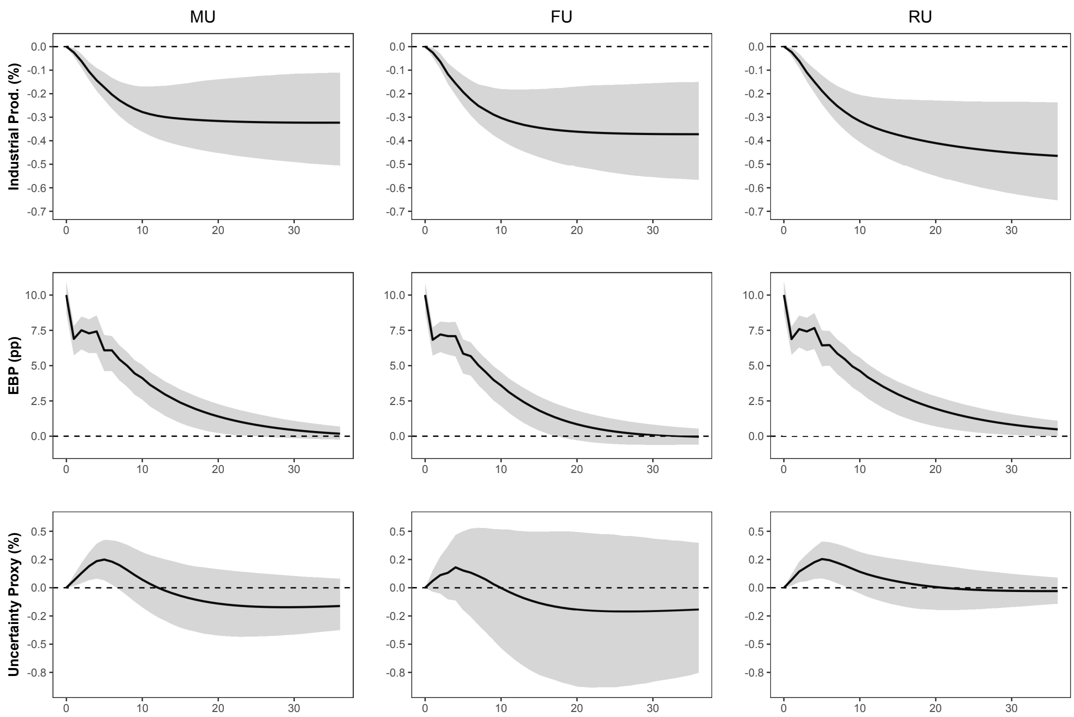


Figure 3.4: Impulse response functions of financial shocks of models with different common uncertainty measures.

Notes: The figure shows impulse response functions of financial shocks in MS(3)-VAR(6) models with $y_t = [\Delta ip_t, ebp_t, uncert_t]'$ and different uncertainty indicators for $uncert_t$. Constraints are imposed on the B-matrix according to the results of Subsection 3.3.3. The impulse responses for Δip_t are cumulated. The instantaneous impact of the financial shock is scaled to an increase of 10 basis points of the excess bond premium. Shaded bands are bootstrapped 95% pointwise confidence intervals.

Financial shocks: Turning to the financial disturbances, the impulse responses of shocks that set off an immediate increase of 10 basis points of the excess bond premium are depicted in Figure 3.4. The impact on the financing conditions fades out rather slowly, staying significant for more than two years in all models.

The reaction of financial uncertainty to the financial shock is indistinguishable from zero at the 95% significance level. This result reveals that the common dispersion of one-step ahead forecast errors of a large set of financial variables is not systematically affected by surprises in investors' pricing of credit risk. This finding is qualitatively in line with recent empirical evidence by Furlanetto et al. (2019), who use sign restrictions in a structural VAR model and find that the VIX as uncertainty measure does not respond significantly to a credit shock. As shown in Figure 3.9 in Appendix 3.B, the findings do not depend on any impact restrictions between the transmission of financial shocks to uncertainty.

In contrast, tighter credit conditions disperse the forecastability of real variables in the economy as real uncertainty increases significantly in the following month after the shock. The higher uncertainty amounts to an increase of roughly 0.2% percent and persists for approximately half a year. One reason for the increase in real uncertainty is the feedback of the decline in industrial production to uncertainty that is suggested by the following analysis of the real activity shocks and supported by the unconditional correlations in Table 3.1. A similar reaction is documented for macroeconomic uncertainty; thus, the positive significant reaction of the broad baseline indicator macroeconomic uncertainty seems to be driven by the real components of that index rather than by financial variables. Generally, the shape of impulse responses of the excess bond premium is virtually identical in all models, implying no distinct feedback of different types of uncertainty to financial conditions.

The adverse financial perturbations trigger a considerable and protracted decline of industrial production in all models that continues for roughly one year before tapering off. The reaction amounts to a sizeable reduction of roughly 0.40% of industrial production. It is roughly similar for all models, which can be seen as an indication that the type of uncertainty plays a minor role of transmitting financial shocks to the real economy.

Given the different identification methods, the impulse responses for the excess bond premium and industrial production are remarkably similar to the results of Caldara et al. (2016) for macroeconomic uncertainty. In this respect, my analysis lends support to their findings under their preferred identification scheme for financial shocks.

Real activity shocks: Figure 3.5 depicts the dynamic effects of exogenous variation in real activity, where the shocks are scaled to an 0.5% instantaneous impact of industrial production. Positive surprises in real activity drive down the excess bond premium

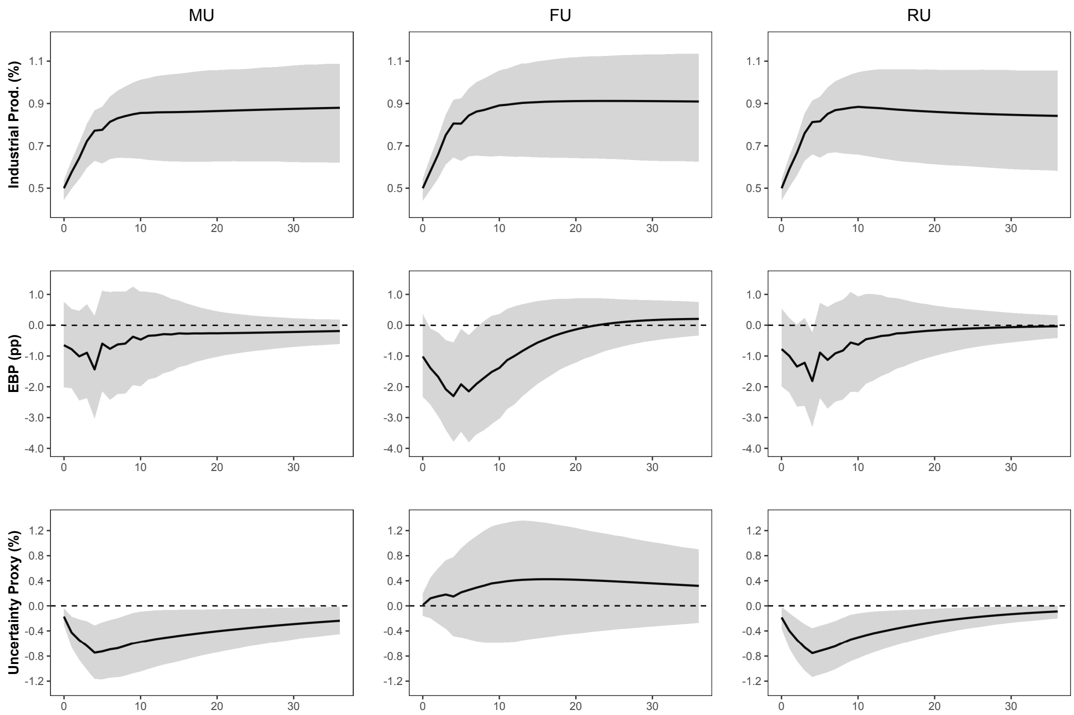


Figure 3.5: Impulse response functions of real activity shocks of models with different common uncertainty measures.

Notes: The figure shows impulse response functions of real activity shocks in MS(3)-VAR(6) models with $y_t = [\Delta ip_t, ebp_t, uncert_t]'$ and different uncertainty indicators for $uncert_t$. Constraints are imposed on the B-matrix according to the results of Subsection 3.3.3. The impulse responses for Δip_t are cumulated. The instantaneous impact of the real activity shock is scaled to an increase of 0.5% of industrial production. Shaded bands are bootstrapped 95% pointwise confidence intervals.

slightly even though the decline is only significant at the 95% significance level for financial uncertainty. The reaction of financial markets to uncertainty is not distinguishable from zero at the 95% significance level conditional on positive innovations in real activity. It can be considered as exogenous to business cycle fluctuations, meaning that it is not a relevant transmission channel of economic distress. This finding is also documented by Ludvigson et al. (2019). Shocks in real activity account for an instantaneous and prolonged decrease of real uncertainty. Thus, real uncertainty serves as an amplifier of movements in real activity, even though quantitative effects seem to remain small given the similar output responses of the model with financial uncertainty. Consequently, the decrease in macroeconomic uncertainty can also be related to the reaction of real variables to the shock in production. Endogenously decreasing macroeconomic uncertainty as a reaction to positive output innovations is also documented by Ludvigson et al. (2019).

Summarizing, the impulse response analysis reveals differences in the interaction of financial conditions and uncertainty from different origins, also stressing the importance of separately examining each distinct type of uncertainty. It underlines that both, uncertainty and financial shocks are important drivers of business cycle fluctuations, as also put forward by Gilchrist et al. (2014) and Caldara et al. (2016).

3.4 Conclusion

Within a structural vector autoregressive framework, exploiting the data's time-varying volatility is found to be a useful tool for disentangling uncertainty shocks from different origins and financial shocks in the US. This statistical identification approach comes with several advantages: First, it provides point estimates of simultaneously identified structural shocks without relying on implausible identifying restrictions. Second, the framework allows for formally testing linear restrictions imposed on the structural parameters. I make use of this feature and propose a new identification scheme based on short-run exclusion restrictions that is not rejected by the data. It opposes the conventional view that innovations in broad uncertainty instantaneously transmit to financial conditions or the reverse. The novel identification scheme is then applied in structural impulse response analysis to examine the causal relations between uncertainty and financial frictions as well as their respective macroeconomic impacts.

I find that adverse shocks to three measures of common uncertainty, macroeconomic uncertainty by Jurado et al. (2015), as well as financial and real uncertainty by Ludvigson et al. (2019), influence financial markets to different degrees. These results extend the analysis of Furlanetto et al. (2019) in the uncertainty dimension. While being qualitatively similar, the quantitative results clearly reveal the importance of systematically differentiating between these distinct types of uncertainty. This finding supports the results of Angelini et al. (2017) and Ludvigson et al. (2019) who argue that differentiating between different types of uncertainty is important. In general, the empirical analysis supports the view that tightened financial conditions and adverse uncertainty shocks play an important role for business cycle fluctuations, as stressed by Gilchrist et al. (2014) and Caldara et al. (2016).

Additionally, the paper adds to the debate on the causal relation between recessions and uncertainty. I find that macroeconomic and real uncertainty respond to business cycle movements endogenously and, hence, are both the cause and effect of recessions. That said, I do not find evidence of the growth-option channel for these types of uncertainty (Ludvigson et al., 2019). In contrast, in line with the empirical findings of Ludvigson et al. (2019) and Angelini et al. (2017), financial uncertainty can be viewed as an exogenous driving force of business cycle fluctuations and does not serve as major propagation mechanism of innovations to real activity.

Appendix

3.A Data

Table 3.5: Data Sources

Data	Sample	Source
IP	1973/01 – 2017/12	Industrial production data is *INDPRO* variable downloaded from the FRED Database of St. Louis Fed
EBP	1973/01 – 2017/12	Monthly updated data for the excess bond premium as modeled by Gilchrist and Zakrajšek (2012) is downloaded from the homepage of the Board of Governors of the Federal Reserve System available at https://www.federalreserve.gov/econresdata/notes/feds-notes/2016/files/ebp_csv.csv
MU	1973/01 – 2017/12	Data for macroeconomic uncertainty as defined by Jurado et al. (2015) is available on Sydney Ludvigson's homepage at https://www.sydneyludvigson.com/s/MacroFinanceUncertainty_2019Feb_update.zip
FU	1973/01 – 2017/12	Data for financial uncertainty as defined by Ludvigson et al. (2019) is available on Sydney Ludvigson's homepage at https://www.sydneyludvigson.com/s/MacroFinanceUncertainty_2019Feb_update.zip
RU	1973/01 – 2017/12	Data for real uncertainty as defined by Ludvigson et al. (2019) is available on Sydney Ludvigson's homepage https://www.sydneyludvigson.com/s/MacroFinanceUncertainty_2019Feb_update.zip
VIX	1973/01 – 2017/12	Data on the (backwards) extended VIX used in (Bloom, 2009) is available on Bloom's webpage https://nbloom.people.stanford.edu/research. It is extended using *VIXCLS* downloaded from the FRED Database of St. Louis Fed

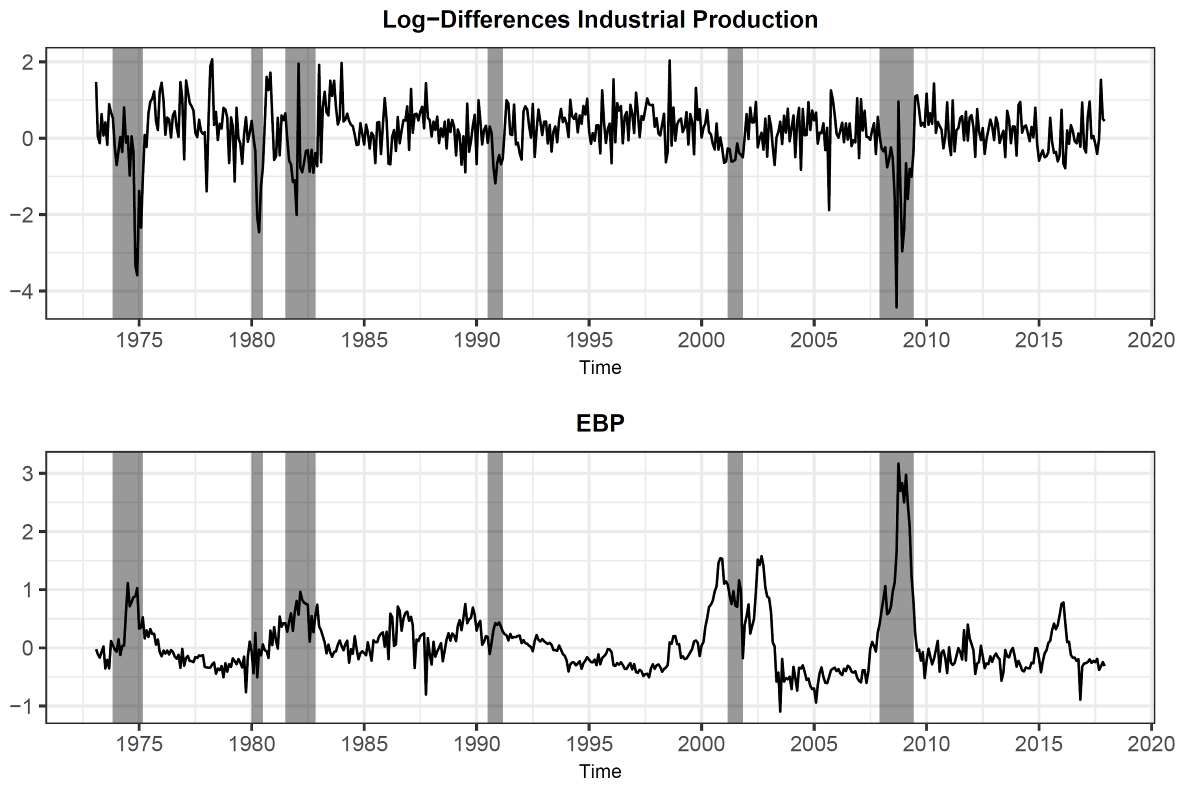


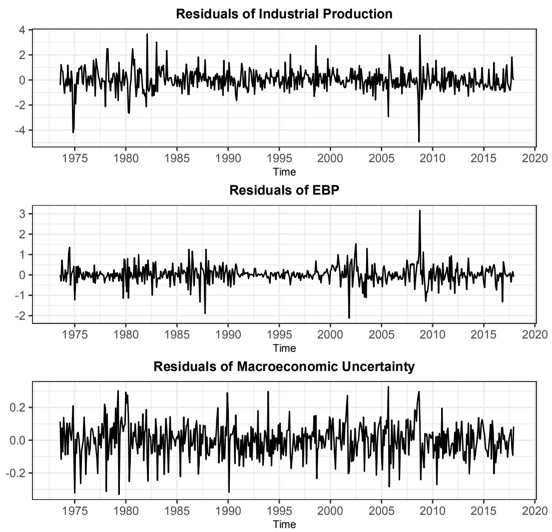
Figure 3.6: Plots of growth rate of industrial production and the excess bond premium.

3.B Additional Results

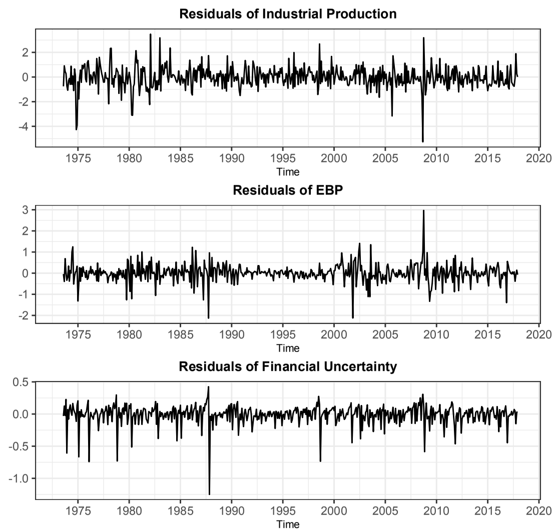
Table 3.6: ARCH-Tests for Linear VAR(6) Models

	p-values			
Test	$H = 1$	$H = 3$	$H = 5$	$H = 10$
	MU			
Portmanteau univariate	0.00	0.00	0.00	0.00
Portmanteau multivariate	0.00	0.00	0.00	0.00
LM	0.00	0.00	0.00	0.00
	FU			
Portmanteau univariate	0.00	0.00	0.00	0.00
Portmanteau multivariate	0.00	0.00	0.00	0.00
LM	0.00	0.00	0.00	0.00
	RU			
Portmanteau univariate	0.00	0.00	0.00	0.00
Portmanteau multivariate	0.00	0.00	0.00	0.00
LM	0.00	0.00	0.00	0.00

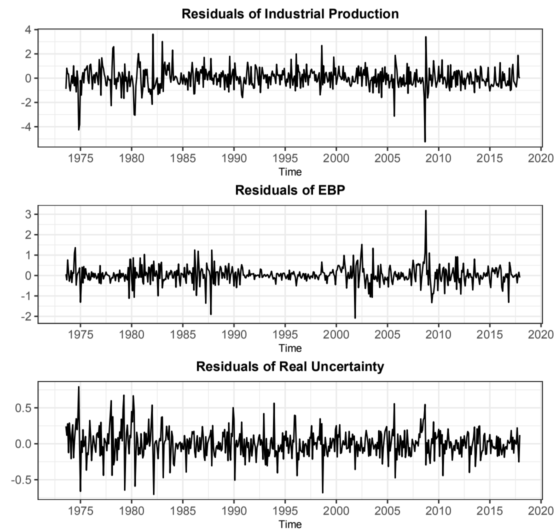
Notes: The table shows p-values of residual ARCH-tests of the estimated standardized residuals $\hat{u}_t^s = \Sigma^{-1/2}\hat{u}_t$ of a linear VAR(6) for $y_t = [\Delta ip_t, ebp_t, uncert_t]'$ and different uncertainty indicators for $uncert_t$. Portmanteau tests are based on the respective univariate and multivariate autocovariances of the sum of squared (standardized) residuals at different horizons H. LM-tests evaluate the joint significance of lagged coefficients in auxiliary regressions where the subdiagonal elements of $\hat{u}_t^s\hat{u}_t^{s\prime}$ are regressed on up to H lagged values of that quantity. See Lütkepohl and Milunovich (2016) for details on the theoretical properties of the tests.



(a) Macroeconomic Uncertainty



(b) Finanical Uncertainty



(c) Real Uncertainty

Figure 3.7: Reduced form residuals.

Notes: The figure shows the reduced form residuals of MS(3)-VAR(6) models with $y_t = [\Delta ip_t, ebp_t, uncert_t]'$ and different uncertainty indicators for $uncert$.

Table 3.7: Tests for State Invariant B-Matrix

Model	log l	# restrictions	LR-statistic	p-value
	MU			
H_0 : MS(3)-VAR(6), state invariant B	-319.97	3	1.93	0.93
H_1 : MS(3)-VAR(6), unrestricted	-319.00			
	FU			
H_0 : MS(3)-VAR(6), state invariant B	-422.38	3	0.32	0.99
H_1 : MS(3)-VAR(6), unrestricted	-422.53			
	RU			
H_0 : MS(3)-VAR(6), state invariant B	-619.01	3	1.50	0.96
H_1 : MS(3)-VAR(6), unrestricted	-618.26			

Notes: Results of LR-tests for parameter constancy of instantaneous impact effects matrix B. Under H_0, MS(3)-VAR(6) with state invariant B refers to model with structure imposed on the M covariance matrices as described in (3.3). This model is tested against the model estimated under H_1, an unrestricted MS(3)-VAR(6) with no structure imposed on the M covariance matrices. Estimates of both models are based on data set $y_t = [\Delta ip_t, ebp_t, uncert_t]'$ with different uncertainty measures for $uncert_t$.

Table 3.8: Estimated State Covariance Matrices ($\times 10^3$)

	MU	FU	RU
Σ_1	$\begin{bmatrix} 394.47 & - & - \\ -7.25 & 19.47 & - \\ -6.98 & 0.73 & 7.06 \end{bmatrix}$	$\begin{bmatrix} 503.95 & - & - \\ -12.16 & 25.34 & - \\ -0.50 & -0.24 & 5.53 \end{bmatrix}$	$\begin{bmatrix} 449.76 & - & - \\ -9.16 & 18.69 & - \\ -5.63 & 0.24 & 24.66 \end{bmatrix}$
Σ_2	$\begin{bmatrix} 380.73 & - & - \\ -7.30 & 218.62 & - \\ -5.48 & 1.08 & 7.70 \end{bmatrix}$	$\begin{bmatrix} 462.59 & - & - \\ -21.42 & 259.11 & - \\ -3.35 & -4.22 & 17.81 \end{bmatrix}$	$\begin{bmatrix} 402.600 & - & - \\ -16.12 & 234.34 & - \\ -12.52 & 0.85 & 19.78 \end{bmatrix}$
Σ_3	$\begin{bmatrix} 3148.17 & - & - \\ -61.20 & 555.19 & - \\ -86.93 & 5.08 & 33.14 \end{bmatrix}$	$\begin{bmatrix} 4809.42 & - & - \\ -137.86 & 768.49 & - \\ 30.60 & -8.10 & 176.52 \end{bmatrix}$	$\begin{bmatrix} 3902.90 & - & - \\ -96.57 & 612.35 & - \\ -230.91 & 6.35 & 158.36 \end{bmatrix}$

Notes: Estimates of state-dependent covariance matrices of MS(3)-VAR(6) reduced form model with $y_t = [\Delta ip_t, ebp_t, uncert_t]'$ with different uncertainty measures for $uncert_t$.

Table 3.9: Unrestricted Estimates of Impact Effect Matrix

MU

$$B = \begin{bmatrix} 0.620\ (0.035) & 0.000\ (0.013) & 0.103\ (0.113) \\ -0.012\ (0.011) & 0.139\ (0.007) & 0.005\ (0.012) \\ -0.025\ (0.015) & 0.000\ (0.002) & 0.080\ (0.006) \end{bmatrix}$$

FU

$$B = \begin{bmatrix} 0.710\ (0.032) & -0.007\ (0.015) & 0.021\ (0.037) \\ -0.016\ (0.011) & 0.158\ (0.007) & 0.002\ (0.016) \\ -0.003\ (0.007) & -0.003\ (0.004) & 0.074\ (0.003) \end{bmatrix}$$

RU

$$B = \begin{bmatrix} 0.596\ (0.082) & -0.005\ (0.012) & 0.308\ (0.186) \\ -0.012\ (0.010) & 0.136\ (0.007) & -0.005\ (0.011) \\ -0.080\ (0.037) & 0.000\ (0.003) & 0.135\ (0.027) \end{bmatrix}$$

Notes: The standard errors in parentheses are obtained from the inverse of the negative Hessian evaluated at the optimum of the likelihood of unrestricted MS(3)-VAR(6) models with $y_t = [\Delta ip_t, ebp_t, uncert_t]'$ and different uncertainty indicators for $uncert_t$.

Table 3.10: LR-Tests of Different Restriction Schemes for VIX Uncertainty Indicator

H_0	H_1	uncertainty measure	p-value (LR-statistic)	df
B_1	state-invariant B	VIX	0.20 (3.18)	2
B_2	B_1		0.05 (3.92)	1
B_3	B_2		0.00 (11.38)	1
B_3	state-invariant B		0.00 (18.48)	4

Notes: The table shows the p-values with respective LR-statistics in parentheses for different restriction schemes B_1-B_3 of the instantaneous impact effects matrix B of MS(3)-VAR(6) models with $y_t = [\Delta ip_t, ebp_t, uncert_t]'$ and the VIX measure as $uncert_t$.

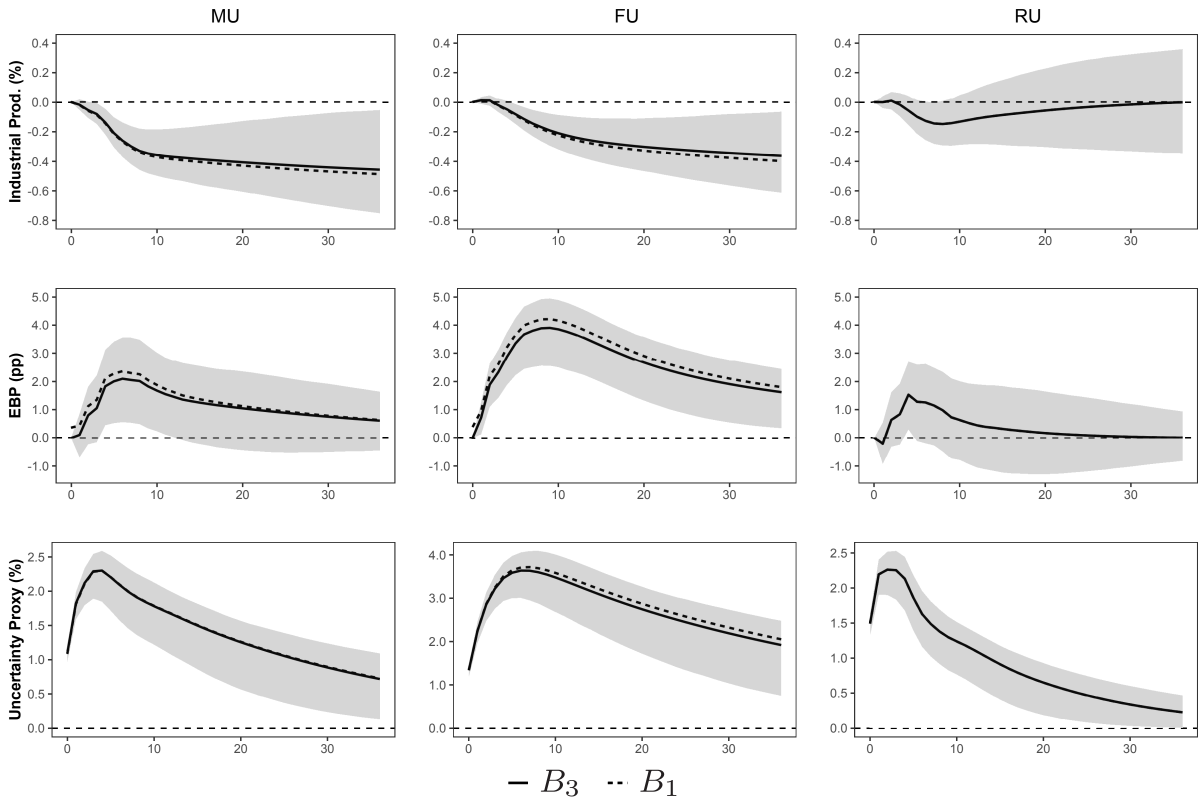


Figure 3.8: Comparison of impulse response functions of uncertainty shocks under identification schemes B_3 and B_1.

Notes: The figure shows comparisons of impulse response functions of uncertainty shocks under identification scheme B_3 (solid line) and B_1 (dashed line) of MS(3)-VAR(6) model with $y_t = [\Delta ip_t, ebp_t, uncert_t]'$ and different uncertainty indicators for $uncert_t$. The impulse responses for Δip_t are cumulated. Scaling of the impulse response functions is according to a one standard deviation uncertainty shock in the first volatility state. Shaded bands are bootstrapped 95% pointwise confidence intervals of the model under identification scheme B_3.

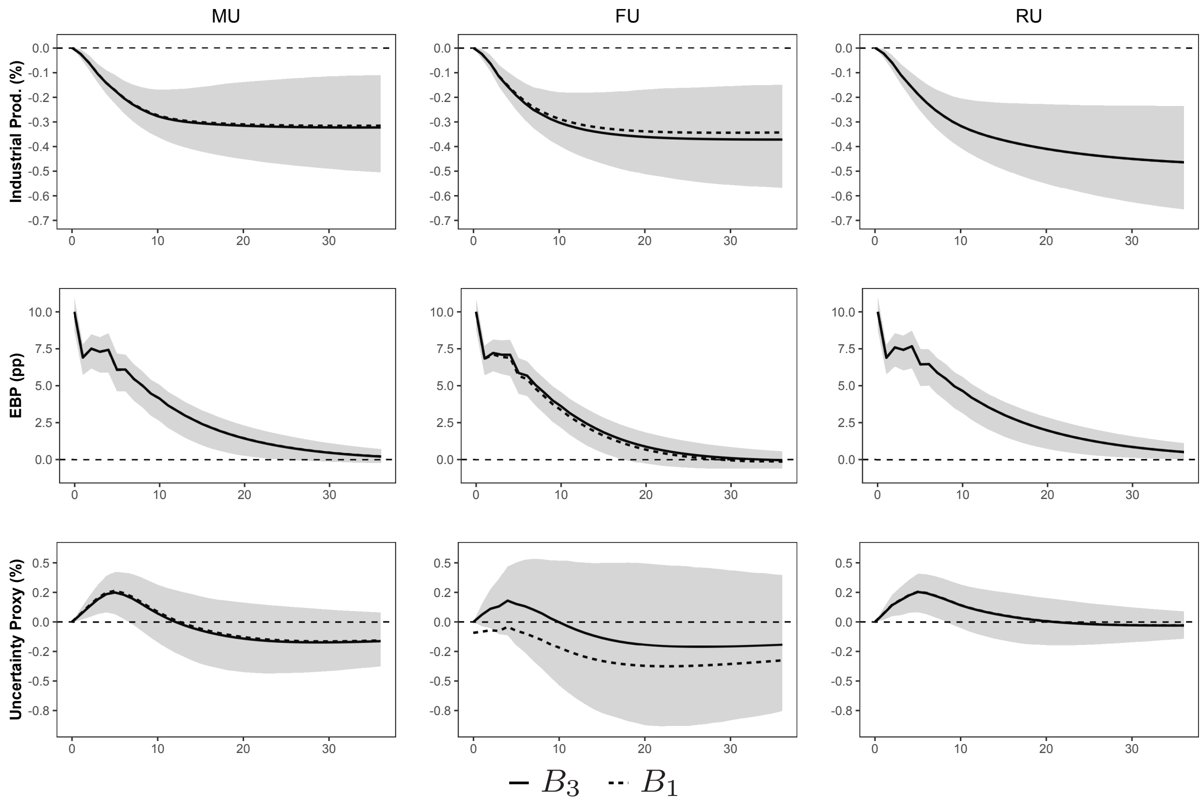


Figure 3.9: Comparison of impulse response functions of financial shocks under identification schemes B_3 and B_1.

Notes: The figure shows comparisons of impulse response functions of financial shocks under identification scheme B_3 (solid line) and B_1 (dashed line) of MS(3)-VAR(6) model with $y_t = [\Delta ip_t, ebp_t, uncert_t]'$ and different uncertainty indicators for $uncert_t$. The impulse responses for Δip_t are cumulated. The instantaneous impact of the financial shock is scaled to an increase of 10 basis points of the excess bond premium. Shaded bands are bootstrapped 95% pointwise confidence intervals of the model under identification scheme B_3.

CHAPTER 4

Monetary Policy, External Instruments and Heteroskedasticity*

4.1 Introduction

Estimating the effects of monetary policy is a central element of macroeconomic analysis. While the economy reacts to policy decisions, monetary policy is also endogenous to the state of the economy, posing the issue of isolating exogenous variation in monetary policy. In the empirical literature, structural vector autoregressions (SVARs) are a main tool for studying the causal effects of monetary interventions. Departing from the classical identification via zero restrictions (Sims, 1980; Christiano et al., 1999), two identification approaches are receiving increasing attention in the literature. On the one hand, authors use extraneous data on monetary surprises to identify latent monetary shocks in SVARs.[1] On the other hand, many papers draw on volatility changes in macroeconomic and financial data to identify monetary shocks.[2] Both identification strategies are popular because they are parsimonious in terms of identifying assumptions and because they incorporate further information into the model.

Identification via external instrument allows for a contemporaneous response of monetary policy to asset prices. Moreover, it adds a potentially large information set to the model through a narrative or financial data-based instrument. Finally, it accounts for measurement error in the instrument, which reduces the attenuation bias in models treating the proxy as the true shock (Mertens and Ravn, 2013; Carriero et al., 2015). However, these advantages rely on the presumption that the instrument is valid, that is, sufficiently strong and exogenous.

*This chapter is joint work with Maximilian Podstawski and Malte Rieth.

[1]See Gertler and Karadi (2015), Cesa-Bianchi et al. (2016), Miranda-Agrippino and Ricco (2017), Stock and Watson (2018), Rogers et al. (2018), Hachula et al. (forthcoming), Caldara and Herbst (2019).

[2]See Rigobon and Sack (2004), Normandin and Phaneuf (2004), Lanne and Lütkepohl (2008), Wright (2012), Herwartz and Lütkepohl (2014), Bacchiocchi and Fanelli (2015), Nakamura and Steinsson (2018).

Identification through heteroskedasticity adds information from time-varying second moments to the model and relies on even weaker identifying assumptions. While an instrument for monetary policy shocks needs to move interest rates without correlating with other structural shocks, a significant relative increase in the variance of monetary shocks is sufficient to trace out the response of the other variables in the system to these shocks. The relative variance shift can be viewed a as 'probabilistic instrument' that increases the likelihood that monetary policy shocks occur (Rigobon, 2003). Again, these minimal assumptions are not costless. The statistically identified shocks are often economically difficult to interpret.

This paper proposes a framework that combines both identification approaches in order to improve inference within SVARs. The framework preserves the attractive features of both approaches but addresses some of the key limitations that each of them has in isolation. It makes use of an external instrument, drawing on instruments for monetary policy shocks proposed in the literature. In addition, it exploits time-variation in the second moments of the data. The combination of both types of identifying information into a 'heteroskedastic proxy-VAR' has three main advantages relative to models using only one type of information.

First, the encompassing framework sharpens the identification of the structural model and, hence, the suitability of the model for policy analysis. We conduct an extensive simulation study. It suggests that the encompassing model yields more accurate estimates of the true model according to the cumulated mean squared errors of impulse response functions than either of the two existing identifcation approaches in isolation. This facet of our model is similar in spirit to Antolín-Díaz and Rubio-Ramírez (2018) who combine narrative information and sign restrictions to enhance inference in SVARs.

We use our framework with tightened grip on the structural model to provide new estimates of the macroeconomic effects of monetary policy shocks in the United States. A common and well documented feature of U.S. real and financial data is time-varying volatility (Stock and Watson, 2002; Justiniano and Primiceri, 2008; Carriero et al., 2016). Standard statistics provide strong evidence that changes in volatility are also present in our sample. We model them within a Markov switching in variances framework and use them for identification. As second central piece of identifying information we include the measure of unanticipated changes in the intended federal funds rate of Romer and Romer (2004) into the model. We find that an unexpected increase in the federal funds rate by 25 basis points leads to a cumulative fall in economic activity of about 0.7 percent. These effects are twice as large as estimates obtained from a standard proxy-VAR that does not exploit the heteroskedasticity. Modeling changes in volatility also allows us to evaluate whether the importance of monetary policy shocks changes across volatility regimes. Our results indicate that monetary shocks were more volatile during the 1970s and 1980s than in the 1990s and 2000s, and that this is associated with a substantially larger role for them in driving real and financial variables in these decades. During the Great Moderation monetary shocks are essentially irrelevant for business cycle fluctuations.

A second contribution of our framework is that it allows testing the validity, that is, exogeneity and relevance, of an instrument. Our framework includes the instrument as an endogenous variable in an augmented SVAR, as in Caldara and Herbst (2019). When using the heteroskedasticity in the residuals of the augmented model, both the exogeneity and the relevance condition become testable. This conveniently reduces to testing zero restrictions on the structural impact matrix of the augmented SVAR. We propose a testing sequence using likelihood ratio (LR) tests for that purpose. Monte Carlo evidence suggests that our exogeneity test has desirable properties in terms of size and power. Testing the exogeneity assumption has so far been unresolved in the literature but is of particular interest as the violation of instrument exogeneity may lead to erroneous conclusions regarding the validity of the instrument and the effects of latent structural shocks.

In the empirical analysis, we first test the narrative measure of Romer and Romer (2004). It is contemporaneously exogenous to technology and financial shocks. Then, we compare alternative instruments for monetary policy shocks proposed in the literature. We find that model-based measures (Bernanke and Mihov, 1998) and high-frequency instruments (Gertler and Karadi, 2015) are also exogenous proxies, and that they produce similar effects as the narrative measure. All three types of instruments imply a significant decline in economic activity during the period of the Great Moderation (see the discussion in Barakchian and Crowe, 2013; Ramey, 2016; Caldara and Herbst, 2019). The exogeneity test complements the invertibility test of the structural moving average representation of SVAR models identified with external instruments of Stock and Watson (2018). Our augmented SVAR provides a natural way for implementing that test as well. Since the instrument enters all equations of the model we can simply test whether it Granger-causes the endogenous variables of the system. We find that it does not, implying invertibility of the structural VAR moving average representation.

The third contribution of the paper is related to the relevance condition for instruments and the literature on identification through heteroskedasticity. Our framework largely dispenses the proxy-VAR approach from weak instrument problems (Lunsford, 2015; Olea et al., 2018). If there is sufficient time-variation for identification in the second moments – a condition that can be checked after estimation – the model is statistically identified. Then, the relevance condition is no longer necessary for statistically valid inference. Whether the instrument is relevant reduces to an economic question about the informational content of the instrument and the interpretation of the structural shock in question. The Monte Carlo evidence suggests that our LR-test reliably discriminates between relevant and irrelevant instruments. The test complements existing versions of F-tests for instrument strength (Stock et al., 2002; Stock and Watson, 2012; Mertens and Ravn, 2013). It has more power than the F-test because it uses all information both under the null and the alternative hypothesis. Thereby, our framework also simplifies the economic interpretation of the shock of interest, addressing a main challenge in the literature on identification through heteroskedasticity (Rigobon and Sack, 2003; Herwartz and Lütkepohl, 2014). In this class of models, structural shocks are identified statistically. They need to be labeled by the researcher after estimation. While the literature has developed several devices for that purpose, this task is often difficult and can leave doubts about the economic meaning of the structural shocks. Through the inclusion of a relevant (and exogenous) proxy into the model, the shock of interest is pinned down by prior economic reasoning.

The remainder of the paper is structured as follows. The next section introduces the heteroskedastic proxy-VAR framework and discusses identification, testing, and estimation of the model. Section 4.3 presents simulation results in support of the framework. In Section 4.4, we use the heteroskedastic proxy-VAR to shed new light on the efficacy of monetary policy and to test a range of instruments discussed in the literature. Finally, Section 4.5 concludes.

4.2 The SVAR Framework

The vector autoregressive (VAR) model is

$$y_t = \gamma + A(L)y_{t-1} + u_t, \tag{4.1}$$

where $y_t = (y_{1t}, \ldots, y_{Kt})'$ is a $(K \times 1)$-vector of observable variables, $A(L)$ is a matrix lag polynomial capturing the autoregressive component of the model, γ collects constant terms, and the u_t are K-dimensional serially uncorrelated residuals. The reduced form residuals u_t are linearly related to white noise structural shocks ε_t, according to

$$u_t = B\varepsilon_t. \tag{4.2}$$

We assume that the VAR is invertible and has a Wold moving average representation $y_t = \gamma + \sum_{i=0}^{\infty} \Phi_i u_{t-i}$.

4.2.1 Identification via External Instrument

We assume that there exists an instrumental variable s_t which is correlated with the structural shock of interest, but uncorrelated with other structural shocks and hence fulfills

$$\mathbb{E}[s_t \varepsilon_{1t}] = \phi \neq 0 \tag{4.3}$$

$$\mathbb{E}[s_t \varepsilon_{jt}] = 0 \quad \forall\, j = 2, \ldots, K, \tag{4.4}$$

where ϕ is an unknown correlation between the instrument s_t and the structural shock of interest ε_{1t}. The latter is ordered first without loss of generality. In the literature, (4.3) is usually called the relevance condition and assumption (4.4) the exogeneity condition. A valid instrument satisfies both (4.3) and (4.4). It allows to recover ε_{1t} and, hence, the corresponding response vector from the reduced form residuals. Rewriting (4.2) with $B = [b_1, B^*]$, where b_1 is the response vector corresponding to ε_{1t} and B^* contains the responses of the remaining shocks, ε_t^*, yields

$$u_t = b_1 \varepsilon_{1t} + B^* \varepsilon_t^*. \tag{4.5}$$

Substituting (4.5) into $\mathbb{E}(s_t u_t)$ and using (4.3) and (4.4) allows uncovering the (relative) impact of the structural shock of interest on every variable in the system, that is, the j^{th} element of b_1 (Stock and Watson, 2012; Mertens and Ravn, 2013; Piffer and Podstawski, 2018). By using the sample moments $\hat{\mathbb{E}}(u_t s_t)$, the instrument s_t implies the following $k-1$ identifying restrictions

$$b_1 = b_{11} \left(1, \frac{\hat{\mathbb{E}}(u_{2t} s_t)}{\hat{\mathbb{E}}(u_{1t} s_t)}, \ldots, \frac{\hat{\mathbb{E}}(u_{Kt} s_t)}{\hat{\mathbb{E}}(u_{1t} s_t)} \right)', \tag{4.6}$$

posing identification of shock ε_{1t} up to the scaling factor b_{11}.

4.2.2 A Heteroskedastic Proxy-VAR

A common feature of macroeconomic and financial data are changes in volatility over time (see, among others, Stock and Watson, 2002; Justiniano and Primiceri, 2008; Carriero et al., 2016). Rigobon and Sack (2004), Normandin and Phaneuf (2004), and Lanne and Lütkepohl (2008) show that this holds in particular for the analysis of monetary policy where changes in volatility of the data feed into heteroskedastic residuals in monetary

SVARs. Against this backdrop, we allow for heteroskedastic residuals in (4.1).[3] We assume that the volatility changes are driven by a first order Markov switching (MS) process S_t that may take values $1, 2, \ldots, M$. The transition probabilities are given by $p_{kl} = P(S_t = l|S_{t-1} = k), k, l = 1, \ldots, M$. Furthermore, the reduced form residuals are normally and independently distributed conditional on a given state $u_t|S_t \sim \text{NID}(0, \Sigma(S_t))$, where all Σ_m, $m = 1, \ldots, M$, are distinct.

Modeling heteroskedasticity in the structural shocks holds direct implications for the external variable instrumenting one of the shocks as the instrument is likely to be heteroskedastic itself. We follow Caldara and Herbst (2019) in assuming that the process generating the potentially heteroskedastic instrument s_t has the following linear form:

$$s_t = \beta\varepsilon_t + \eta\nu_t, \tag{4.7}$$

where ε_t is the $K \times 1$ vector of structural shocks, $\beta = (\beta_1, \beta_2, \ldots, \beta_K)$ is a $1 \times K$-coefficient vector, $\nu_t \sim N(0, \sigma_m^2)$ is a measurement error uncorrelated with the structural shocks ε_t, and η scales the effect of the noise. β_1 and η may be interpreted as weighting parameters of signal to noise, defining the quality of the instrument s_t. The instrument's quality is flawed by the noise ν_t and potentially by the influence of other structural shocks on the instrument through β_j with $j = 2, \ldots, K$, determining the degree of its endogeneity.[4]

We compile the system by appending model (4.1) with the process generating the external instrument in (4.7). The augmented VAR system is

$$z_t = \delta + \Gamma(L)z_{t-1} + e_t, \tag{4.8}$$

where $z_t = [y_t', s_t]'$ is a $((K + 1) \times 1)$-vector of observable variables, $\Gamma(L)$ is a (potentially restricted) lag matrix polynomial capturing the autoregressive component of the model, δ is a $((K + 1) \times 1)$-vector of constant terms, and e_t are $(K + 1)$-dimensional serially

[3] We refrain from introducing additional nonlinearity into the model by allowing state-dependency in the constant or autoregressive parameters as we are interested in the heteroskedasticity features of the data for identification purposes.

[4] Time-variation in the second moments of the data may imply a time-varying correlation ϕ_m in (4.3). However, under the assumption of a time-invariant impact matrix B, which is standard in the literature on external instruments (Stock and Watson, 2012; Mertens and Ravn, 2013; Gertler and Karadi, 2015), this state-dependency does not imply any changes in the use of the instrument for identification, as ϕ (or ϕ_m) does not enter the relative impulse vector (4.6).

uncorrelated residuals. The latter are related to the structural innovations μ_t as

$$\begin{aligned} e_t &= D\mu_t \\ &= \begin{bmatrix} B_{(K\times K)} & 0_{(K\times 1)} \\ \beta_{(1\times K)} & \eta \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ \nu_t \end{bmatrix}. \end{aligned} \tag{4.9}$$

Using (4.9), we rewrite the augmented VAR in (4.8) in structural form as

$$z_t = \Delta + \Gamma(L)z_{t-1} + D\mu_t. \tag{4.10}$$

Since the state dependency in the variances of the reduced form residuals in (4.8), $var(e_t|m)$ $= \tilde{\Sigma}_m$ with $m = 1, \ldots, M$, translates into the structural form, we have $E[\mu_t|m] = 0$ and $E[\mu_t\mu_t'|m] = \Lambda_m$, where Λ_m is a diagonal matrix satisfying the orthogonality condition of the structural innovations.

Beyond identifying information from the external instrument, the heteroskedasticity pattern provides a valuable source of identifying information (Rigobon and Sack, 2004; Normandin and Phaneuf, 2004; Lanne and Lütkepohl, 2008). Under the assumption of a constant instantaneous impact matrix D, for each volatility regime a decomposition

$$\tilde{\Sigma}_m = D\Lambda_m D' \tag{4.11}$$

exists, where $\Lambda_m = \text{diag}(\lambda_{1,m}, \ldots, \lambda_{K+1,m})$. We normalize $\Lambda_1 = I_{K+1}$. For $m \geq 2$, the Λ_m are diagonal matrices with strictly positive elements that can be interpreted as the changes of the variances of the structural innovations in the respective regime relative to the first regime. Lanne et al. (2010) state conditions for local uniqueness of matrix D. Local uniqueness implies that D is identified up to the signs of the parameters in each column as well as to column permutations. The conditions for local uniqueness of D are: (i) the structural impact matrix D is time-invariant; (ii) the structural innovations μ_t are orthogonal; and (iii) there are sufficiently many and distinct changes in the variances of the structural innovations. The first assumption is standard in structural VARs identified with external instruments.[5] The second assumption is common in structural VAR analysis more generally. The third assumption can be checked after estimation by comparing the estimated variances λ_{lm}, with $l = 1, \ldots, K+1$.

[5] Using alternative identification schemes, Owyang and Ramey (2004), Primiceri (2005) and Sims and Zha (2006) examine the role of changes in the monetary policy rule over time. While they find some evidence for regime-switches, they conclude that these changes explain only a small part of U.S. business cycle fluctuations (Ramey, 2016). Other authors find little or no evidence of changes in the policy coefficients (Bernanke and Mihov, 1998; Leeper and Zha, 2003; Hanson, 2006).

To see how the model setup relates to the conventional identification of a proxy-VAR note that the $(K+1, i)$-element of $\tilde{\Sigma}_m$ in (4.11) is $cov(u_{it}s_t)$. This term is equated to the corresponding element of $D\Lambda_m D'$ which is $\beta_1\lambda_{1m}b_{i1}$ for $i = 1, \dots, K$ and $\beta_2 = \cdots = \beta_K = 0$. It follows that $\frac{cov(u_{it}s_t)}{cov(u_{1t}s_t)} = \frac{b_{i1}}{b_{11}}$. This ratio is usually termed the relative impulse vector. It summarizes the restrictions on the structural parameters of the model implied by the instrument.

4.2.3 Testing the Validity of an External Instrument

If the conditions for local uniqueness are met, the heteroskedasticity in the residuals allows for estimating all structural parameters of D of the augmented SVAR model (4.10). Any additional restrictions on D are then over-identifying and, hence, testable. This is particularly interesting in our context as it allows for testing both the relevance and the exogeneity of the instrument and, thus, its validity. Such tests conveniently reduce to testing zero restrictions on β, that is, the last row of the structural impact matrix D. Given full identification of the model via heteroskedasticity, this may be done with likelihood ratio tests (LR-tests), as in Lanne and Lütkepohl (2008) because the elements of β are fixed parameters under the null hypothesis. Testing zero restrictions on these parameters then implies evaluating nested models against each other. This furthermore implies that the distribution of the LR-tests is a standard χ^2-distribution and the degrees of freedom are equal to the number of restrictions.

We propose a two-stage testing sequence. First, we assess the *exogeneity condition* of the instrument by comparing the likelihood of an appropriately restricted version of model (4.10), that is, restricting $\beta = (\beta_1, 0, \dots, 0)$, with an unrestricted model (4.10) where $\beta = (\beta_1, \beta_2, \dots, \beta_K)$. Formally, we test

$$H_0 : \beta_2 = \cdots = \beta_K = 0$$
$$H_1 : \exists j \in \{2, \dots, K\} \text{ s.t. } \beta_j \neq 0.$$

Rejecting the null indicates endogeneity of the instrument. Otherwise we proceed to the second stage.

Here, we assess the *relevance condition* by comparing an appropriately restricted version of model (4.10), that is, setting $\beta_1 = 0$, with model (4.10) where β_1 is unrestricted. Under both the null and the alternative hypothesis $\beta_2 = \cdots = \beta_K = 0$. Formally, we test

$$H_0 : \beta_1 = 0$$
$$H_1 : \beta_1 \neq 0.$$

Rejecting the null indicates the relevance of the instrument and, together with not rejecting the null of instrument exogeneity at the first stage, that it is valid. If the instrument is valid, we set $\beta = (\beta_1, 0, \ldots, 0)$ and refer to model (4.10) as a 'heteroskedastic proxy-VAR'.

Local uniqueness in our setup implies that *a priori* we cannot identify the column of the impact matrix D that belongs to a certain structural shock. Practically and in the simulations this is of little concern because a valid instrument for a monetary policy shock imposes additional restrictions on the covariance matrix. Consequently the shock that is most consistent with these restrictions is ordered to the column with the unrestricted β-coefficient. This pins down the monetary policy shock. Furthermore, assessing the endogeneity of instrument does not require a particular ordering of the remaining structural shocks as the test for exogeneity will reject the null of all but one β-element equal to zero in case of endogeneity. A weak external instrument could potentially prevent the shock of interest to be ordered in the column with the unrestricted β-element. However, this will be detected at the second stage of the testing sequence through not rejecting the null, which indicates an uninformative instrument.

4.2.4 Estimation

The parameters of (4.10) are estimated by means of the expectation maximization (EM) algorithm proposed by Herwartz and Lütkepohl (2014). Crucial for the analysis is to incorporate the regime-switching nature of the covariance matrix described in (4.11), given the restrictions on D and Λ_m. All other parameters are assumed to be regime-independent and do not vary across states. We use the following concentrated out log likelihood function in the maximization step of the EM algorithm and refer to Appendix 4.A for further computational details:

$$\mathcal{L}(D, \Lambda_2, \ldots, \Lambda_M) = \frac{1}{2} \sum_{m=1}^{M} \left[T_m \log(\det(\tilde{\Sigma}_m)) + \mathrm{tr}\left((\tilde{\Sigma}_m)^{-1} \sum_{t=1}^{T} \xi_{mt|T} u_t u_t' \right) \right],$$

where $\xi_{mt|T},\ t = 1, \ldots, T$ are the model smoothed probabilities with $T_m = \sum_{t=1}^{T} \xi_{mt|T}$.

Once the EM algorithm has converged, we obtain standard errors of the point estimates of the parameters through the inverse of the negative Hessian matrix evaluated at the optimum. We use these standard errors as a statistic to determine whether the estimated structural variances change significantly and by differing amounts across states. We check that the one standard error confidence intervals do not overlap. This is a requirement for statistical identification and, hence, for the testing sequence. For the dynamic analysis, we compute bootstrapped impulse responses. Given the heteroskedasticity, classical residual bootstrapping may be problematic in generating reliable confidence intervals. Any resampling scheme needs to preserve the second order characteristics of the data so that the structural parameters are maintained and can be estimated from bootstrapped samples. Therefore, we use a recursive design wild bootstrap with $e_t^\star = \varphi_t \widehat{e}_t$, where φ_t is a random variable independent of z_t following a Rademacher distribution. φ_t is either 1 or −1 with probability 0.5 and the bootstrap procedure is repeated 5,000 times. In our bootstrap procedure, we condition on the original estimates of the relative variance parameters, as well as on the estimates of the transition probabilities. This is a commonly used technique for these types of models (Herwartz and Lütkepohl, 2014; Podstawski and Velinov, 2018). Alternatively, Lütkepohl and Schlaak (2019) show that a related bootstrap method based on a normal distribution performs well for a model with volatility changes driven by GARCH processes. In Figure 4.15 in Appendix 4.C we show that our results are robust to using this bootstrap.

4.3 Simulation Study

To explore the properties of the proposed framework and LR-tests, we conduct an extensive Monte Carlo simulation. First, we evaluate how the tests behave for different degrees of instrument endogeneity and relevance. Second, we assess whether the accuracy of the estimation of the structural model increases systematically due to the explicit modeling of heteroskedasticity. We also discuss how the proposed framework alleviates problems of weak instruments in proxy-VARs.

4.3.1 Setup of Monte Carlo Study

We assume that the data generating process is of the form (4.10). The process implies that y_t and s_t are jointly normally distributed conditional on state $m = 1, \ldots, M$. In the simulation, we generate data for y_t and then for the instrument s_t contingent on the realizations of y_t. We use the following parameters for a first-order autoregressive model with instantaneous effect matrix B, which are taken from the New Keynesian DSGE-model of An and Schorfheide (2007):

$$\begin{bmatrix} r_t \\ x_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} 0.79 & 0.00 & 0.25 \\ 0.19 & 0.95 & -0.46 \\ 0.12 & 0.00 & 0.62 \end{bmatrix} \begin{bmatrix} r_{t-1} \\ x_{t-1} \\ \pi_{t-1} \end{bmatrix} + \begin{bmatrix} 0.69 & 0.61 & 0 \\ -1.10 & 1.49 & 1 \\ -0.75 & 1.49 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_t^r \\ \varepsilon_t^z \\ \varepsilon_t^g \end{bmatrix},$$

where r_t is the interest rate, x_t is output and π_t is the inflation rate. The structural shocks are characterized as a monetary policy shock (ε_t^r), a productivity shock (ε_t^z), and a government spending shock (ε_t^g).

The variances of the structural innovations are driven by a discrete Markov switching process with $M = 2$ states and transition probabilities

$$P = \begin{bmatrix} 0.975 & 0.025 \\ 0.050 & 0.950 \end{bmatrix},$$

which are used to generate the Markov states S_t for $t = 1, \ldots, T$. Following standard conventions, we normalize the variances of the structural innovations in the first state to unity. We set the relative variances in the second state by choosing rather distinct variances in the range of parameters used in comparable studies (Lütkepohl and Schlaak, 2018):

$$\Lambda_2 = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0.5 \end{pmatrix}.$$

Given that the B matrix may be identified up to column signs and permutations only, we assure that the models are uniquely determined by sorting the estimated coefficients of Λ_2 in descending order and by adjusting the columns of the estimated impact matrix correspondingly. With appropriate starting values $y_0 = (0, 0, 0)'$, we generate data recursively by drawing from

$$\varepsilon_t \sim \begin{cases} N(0, I), & \text{for} \quad m = 1 \\ N(0, \Lambda_2), & \text{for} \quad m = 2, \end{cases}$$

using $B\varepsilon_t = u_t$ to calculate the reduced form residuals.

With the structural innovations at hand, we generate the instrument s_t, using (4.7). We set $\eta = 1$ and the variances of the noise parameter ν_t such that

$$\nu_t \sim \begin{cases} N(0, 1), & \text{for} \quad m = 1 \\ N(0, 12), & \text{for} \quad m = 2. \end{cases}$$

This setup implies a time-varying volatility of the instrument which can be observed in many time series of instruments that are used in the literature (see, for example, Romer and Romer, 2004; Gertler and Karadi, 2015).

The relationship between the structural shocks and the artificial instruments is modeled through β. We set $\beta = (\beta_1, \beta_2, 0)$, where β_1 captures the relevance of the instrument for the monetary shock ε_t^r, while β_2 measures the endogeneity to the second structural shock ε_t^z. We equate β_3 to zero to focus the simulation study, concentrating on cases where endogeneity stems only from one source. We construct a set of different instruments for the target monetary policy shock, using the following values: $\beta_1 \in [0, 0.2, 0.4, 0.6]$. Similarly, we consider $\beta_2 \in [0, 0.05, 0.20, 0.30, 0.40]$ and thereby introduce different degrees of endogeneity. These parameter combinations for β_1 and β_2 imply different correlations of the instrument with the monetary shock (ρ_1) and the non-monetary shock (ρ_2), which we list in the simulation results to facilitate a comparison to empirical applications. Finally, we use two sample sizes, $T = 200$ and $T = 500$, which are within the typical range of macroeconomic datasets. The number of replications for each simulation design is $R = 500$.

The setup of our Monte Carlo study closely mimics important features of the data used in our empirical study introduced in Section 4.4. First, in the simulations, the size of the dimensions K and T are chosen to be comparable to the size of the system used in our empirical application. Moreover, the structure imposed on parameter values stems from estimates of a DSGE model and, hence, comprises plausible values for the generation of the artificial data sets. Similarly, the heteroskedasticity induced to the data series is comparable to what has been estimated from data sets in related empirical work (see, e.g., Lütkepohl, 2013; Lütkepohl and Netšunajev, 2017a).

4.3.2 Fitted Models

As a reference model for the test evaluation, we fit a MS(2)-VAR(1) with unrestricted β to the data. Then, we estimate and compare the following three models:

Model A Heteroskedastic proxy-VAR with $\beta = (\beta_1, 0, 0)$, that is, the instrument s_t is assumed to be exogenous.

Model B Heteroskedastic SVAR with $\beta = (0, 0, 0)$, that is, the model is identified by the time-varying volatility only.

Model C Standard proxy-VAR using the identifying information from the external instrument only.

The reference model with unrestricted β and models A and B allow for a time-varying covariance, which we model within a Markov switching in heteroskedasticity framework with two latent states. We place alternative restrictions on β as discussed in Section 4.2.2. Model C fits a standard proxy-VAR with the two stage least squares procedure suggested by Mertens and Ravn (2013) to evaluate a situation where the volatility in the data is ignored. Here, the response of the first variable to the identified structural shock is normalized to a have a positive sign. This model has *a priori* a disadvantage compared to the other models, given that the generated data feature volatility changes. Given that the reference model and models A and B are nested, we can compute χ^2-distributed LR-statistics to test for the exogeneity and the relevance of the generated instrument in each replication.[6] To test the exogeneity condition, we test the heteroskedastic reference VAR with β unrestricted against model A. To test the relevance of the instrument, we test model A against the more restricted model B.

To assess the benefits of combining identification via external instrument and via heteroskedasticity, we calculate the cumulated mean squared errors (MSE) of the estimated structural impulse response functions for models A-C relative to the true parameters of the data generating process. This metric summarizes the accuracy of the structural estimates relative to the true model. The cumulated MSE up to horizon h for variable k induced by shock l is calculated as

$$MSE_h(\theta_{kl,\bullet}) = \sum_{i=0}^{h-1} \left(\frac{1}{R} \sum_{r=1}^{R} (\theta_{kl,i} - \hat{\theta}_{kl,i}(r))^2 \right),$$

where $\hat{\theta}_{kl,i}(r)$ denotes the estimate of the structural impulse response $\theta_{kl,i}$ obtained in the r^{th} replication of our simulation experiment.[7] As the data generating VAR(1) parameters imply substantial persistence, we calculate cumulated MSE for a propagation horizon of up to $h = 25$ such that we capture the impact of differing estimates of both the impact matrix B and the autoregressive part of the model. Finally, we assess the accuracy of the estimates for the monetary policy shock only in order to accommodate the fact that the identification via a single external instrument, as in model C, facilitates the identification of one shock per instrument at most.

4.3.3 Simulation Results

Table 4.1 shows the relative rejection frequencies of the LR-test for exogeneity at a nominal significance level of 10% for the two different sample sizes. The complete set of simulation results may be found in Appendix 4.B. Exogenous instruments with $\rho_2 = 0$ are rejected with frequencies close to their expected nominal levels (see first column). This suggests

[6] If $\beta = 0$, the heteroskedastic proxy-VAR reduces to a standard heteroskedastic SVAR where the distributions of s_t and u_t are independent. In this case, the structural parameters are identified using the heteroskedasticity of the data only.

[7] The structural impulse responses of the models are obtained as the elements of the matrices $\Theta_i = \Phi_i B$, $i = 0, 1, \dots$, where Φ_i is the coefficient matrix of the i^{th} propagation horizon of the Wold moving average representation of the VAR. More precisely, the kl^{th} element of Θ_i, denoted by $\theta_{kl,i}$, is interpreted as the response of variable k to the l^{th} structural shock after a propagation horizon of i periods.

that the test is neither under nor over-sized. When moving to the right across columns, the LR-test has power against the null hypothesis of an exogenous instrument. The rejection frequencies steadily increase with higher instrument endogeneity for both sample sizes. For $T = 200$, endogeneity is detected in roughly 50% of the cases, when the sample correlation of the non-monetary shock with the instrument is elevated. For $T = 500$ and $\rho_2 \geq 0.16$, the rejection frequencies are usually higher than 90%, depending on the relevance of the instrument. For a nominal significance level of 5% the rejection frequencies tend to be lower, but for $T = 500$ and average sample correlations $\rho_2 \geq 0.12$, endogeneity is reliably detected in more than 80% of the cases.

Table 4.1: Relative Rejection Frequencies at Nominal Significance Level of 10% of LR-test for Exogeneity of Instrument

Sample Size	Relevance (β_1,ρ_1)	Endogeneity (β_2,ρ_2)				
		(0.0,0.0)	(0.05,0.03)	(0.20,0.12)	(0.30,0.16)	(0.40,0.22)
200	(0.00,0.00)	0.12	.	.	.	.
	(0.20,0.16)	0.10	0.13	0.39	.	.
	(0.40,0.30)	0.09	0.12	0.31	0.52	0.63
	(0.60,0.43)	0.10	0.10	0.25	0.42	0.54
500	(0.00,0.00)	0.12	.	.	.	.
	(0.20,0.16)	0.10	0.17	0.85	.	.
	(0.40,0.30)	0.09	0.14	0.70	0.90	0.96
	(0.60,0.43)	0.09	0.12	0.50	0.78	0.91

Notes: Each entry in the table is based on 500 replications of each simulation design. Dots (.) denote combinations of values for β_1 and β_2 that produce lower correlations between the instrument s_t and the target structural shock ε_t^r than between the instrument and the endogenous structural shock ε_t^z. These cases are not taken into account in the analysis.

Table 4.2 displays the relative rejection frequencies of the LR-test for instrument relevance. We focus on a 5% nominal significance level as the test displays relative rejection frequencies very close to one against false null hypotheses for a significance level of 10% (see 4.14 in Appendix 4.B). For a white noise instrument without any identifying information ($\rho_1 = \rho_2 = 0$), the test shows an expected nominal rejection rate around 5%. The rejection frequency rapidly increases for higher correlations of the instrument with the monetary shock. The null of an uninformative instrument is rejected reliably in all cases and for both samples if $\rho_1 \geq 0.30$, irrespective of the endogeneity of the instrument.

To obtain an impression of the power of the LR-test and the relevance of the artificial instruments, we compare our test to the well-established F-test for instrument strength (Stock et al., 2002; Stock and Watson, 2012; Mertens and Ravn, 2013). Table 4.3 contains the relative rejection frequencies at a nominal significance level of 5% and the corresponding (heteroskedasticity robust) F-statistics for exogenous instruments of different strength.

Table 4.2: Relative Rejection Frequencies at Nominal Significance Level of 5% of LR-test for Relevance of Instrument

Sample Size	Relevance (β_1,ρ_1)	Endogeneity (β_2,ρ_2)				
		(0.0,0.0)	(0.05,0.03)	(0.20,0.12)	(0.30,0.16)	(0.40,0.22)
200	(0.00,0.00)	0.06	.	.	.	.
	(0.20,0.16)	0.74	0.75	0.82	.	.
	(0.40,0.30)	1.00	1.00	1.00	1.00	1.00
	(0.60,0.43)	1.00	1.00	1.00	1.00	1.00
500	(0.00,0.00)	0.05	.	.	.	.
	(0.20,0.16)	0.98	0.98	0.98	.	.
	(0.40,0.30)	1.00	1.00	1.00	1.00	1.00
	(0.60,0.43)	1.00	1.00	1.00	1.00	1.00

Notes: Each entry in the table is based on 500 replications of the each simulation design. Dots (.) denote combinations of values for β_1 and β_2 that produce lower correlations between the instrument s_t and the target structural shock ε_t^r than between the instrument and the endogenous structural shock ε_t^z. These cases are not taken into account in the analysis.

The first column shows that the size of the F-test is close to its nominal level although the test tends to reject the null slightly too often when it is true and the sample is small. The rejection frequencies increase in instrument relevance, that is, when moving right across columns, for both sample sizes. However, the increase is substantially slower than for the LR-test (see first column of Table 4.2). The latter detects a relevant instrument with 100% probability if $T \geq 200$ and $\rho_1 \geq 0.30$, whereas the F-test does so only if $\rho_1 > 0.43$ for $T = 200$ or if $\rho_1 \geq 0.43$ for $T = 500$. In other words, the LR-test has more power against the null hypothesis of irrelevant instrument when the alternative is true.

The decrease in type-II error is useful for practical purposes. It implies that fewer relevant instrument are discarded. For $T = 200$ and $\rho_1 = 0.30$ the LR-test suggests keeping all candidate relevant instruments, whereas the F-test retains only 75%. For $T = 500$ and $\rho_1 = 0.16$ the LR-test finds 98% of relevant instruments, whereas the F-test detects only 55%. The advantage of having a test with more power becomes even more visible when departing from the 5% significance level for the F-test and using the stricter criterion of an F-statistic of 10, which is commonly used to shield against weak instrument problems. Table 4.3 suggests that between 50% and 60% of relevant instruments are then erroneously labeled as weak instruments in samples of 500 and 200 observations, respectively.

Another feature of the proposed framework is that it largely resolves the problem of weak instruments, provided that there is sufficient heteroskedasticity in the data. Table 4.4 displays the evaluation of models A-C using the MSE of the structural impulse responses as accuracy criterion. We normalize the MSE by those of model A and focus on the

Table 4.3: Relative Rejection Frequencies at Nominal Significance Level of 5% of Robust F-Test for Instrument Relevance

	Sample $T = 200$			Sample $T = 500$		
Relevance (β_1,ρ_1)	Rejection frequency	Frequency F > 10	Robust F-statistic	Rejection frequency	Frequency F > 10	Robust F-statistic
(0.00,0.00)	0.07	0.01	1.19	0.05	0.00	1.02
(0.20,0.16)	0.31	0.06	3.30	0.55	0.17	5.55
(0.30,0.23)	0.66	0.17	5.86	0.86	0.51	11.3
(0.40,0.30)	0.75	0.38	9.33	0.97	0.81	19.19
(0.60,0.43)	0.96	0.78	18.61	1.00	1.00	40.66

Notes: The table shows the relative rejection frequencies of heteroskedasticity robust F-tests for instrument strength at a nominal significance level of 5%, the relative frequencies that F > 10, and the average F-statistics, based on 500 replications for each instrument. Endogeneity is assumed to be absent, that is $\beta_2 = \rho_2 = 0$.

results based on a sample size of $T = 500$.[8] For a white-noise instrument ($\rho_1 = \rho_2 = 0$), models A and B yield essentially the same MSE. This implies that the use of a weak instrument is unproblematic for inference if the data contain changes in volatility and if they are used for identification. For relevant and exogenous instruments, that is moving south across rows, the heteroskedastic proxy-VAR systematically yields the smallest MSE for all variables and parameterizations compared to both competing models. These gains increase with instrument relevance and are substantial. For instruments with 43% correlation (last panel in Table 4.4) with the monetary shock, the improvement relative to model B ranges between 6% and 68% for the respective impulse responses. Model C performs extremely poorly in all cases. Given that the variances of the instrument and of the other endogenous variables are time-varying in our setup, fitting a standard proxy-VAR that does not account for heteroskedasticity leads to serious distortions in the estimates of the structural parameters, irrespective of the relevance of the exogenous instrument. Overall these results suggest that the explicit modeling of volatility changes when they are a feature of the data and using the information of a valid proxy improves structural inference in SVARs.

This conclusion also holds for slightly endogenous instruments ($\rho_2 = 0.03$) if the proxy is sufficiently relevant. When the correlation between the instrument and the monetary shock is larger than $\rho_1 = 0.16$, model A still and consistently yields the smallest MSE. When the endogeneity increases further, the estimation precision of model A deteriorates considerably. Then, model B, which ignores the misspecified instruments for identification, yields more precise estimates. This finding underscores the importance of being able to test for instrument exogeneity. As before, model C performs worst in all cases.

[8] Appendix 4.B shows that the results are robust to changes of the propagation horizon and sample size.

Table 4.4: Comparison of MSE of Impulse Responses to Monetary Policy Shock

Relevance (β_1,ρ_1,F)	Endogeneity (β_2,ρ_2)														
	(0.0,0.0)			(0.05,0.03)			(0.20,0.12)			(0.30,0.16)			(0.40,0.22)		
Model	θ_{11}	θ_{21}	θ_{31}	θ_{11}	θ_{21}	θ_{31}	θ_{11}	θ_{21}	θ_{31}	θ_{11}	θ_{21}	θ_{31}	θ_{11}	θ_{21}	θ_{31}
(0.0,0.0,1.3)															
Model A	1.00	1.00	1.00	.	.	.	.	.	.	.	.	.	.	.	.
Model B	1.00	0.99	1.00	.	.	.	.	.	.	.	.	.	.	.	.
Model C	82.30	50.48	140.88	.	.	.	.	.	.	.	.	.	.	.	.
(0.20,0.16,6.9)															
Model A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.	.	.	.	.	.
Model B	1.02	1.10	1.01	1.01	0.92	0.98	0.86	0.24	0.79	.	.	.	.	.	.
Model C	42.45	21.96	54.48	40.79	18.65	50.87	25.03	9.79	26.27	.	.	.	.	.	.
(0.40,0.30,23.6)															
Model A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Model B	1.05	1.37	1.05	1.04	1.09	1.02	0.85	0.27	0.80	0.50	0.11	0.45	0.36	0.07	0.30
Model C	34.04	15.35	48.45	33.35	10.41	46.90	22.40	3.64	28.69	10.47	2.41	11.88	6.23	1.98	5.86
(0.60,0.43,49.0)															
Model A	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Model B	1.06	1.68	1.08	1.07	1.35	1.08	0.96	0.37	0.95	0.77	0.18	0.75	0.51	0.10	0.47
Model C	32.98	16.26	49.49	32.80	11.16	48.94	26.42	3.15	37.99	18.60	2.14	25.75	10.33	1.68	12.97

Notes: The table shows the cumulated MSE of fitted models (1)-(3) relative to model (1) for a propagation horizon up to $h = 25$ and sample size $T = 500$. Each entry is based on 500 replications of each simulation design. Dots (.) denote combinations of values for β_1 and β_2 that produce lower correlations between the instrument s_t and the target structural shock ε_t^r than between the instrument and the endogenous structural shock ε_t^z. These cases are not taken into account in the analysis.

Summarizing the simulation results, both LR-tests are helpful tools to assess the validity of instruments. Relevant instruments are detected reliably already in small samples at the 5% significance level. Moreover, the LR-test has more power than the widely used F-test. A detection of endogeneity requires somewhat larger samples and higher correlations between the instrument and the non-monetary shocks. The rejection frequencies suggest using the 10% significance level. Regarding the estimation precision, the heteroskedastic proxy-VAR recovers the true model best, even in cases of slightly endogenous instruments. As endogeneity increases, a standard heteroskedastic SVAR not using the instrument performs better and the heteroskedastic proxy-VAR yields seriously distorted estimates. This stresses the importance of having a means for evaluating instrument exogeneity. Finally, both models using time-varying volatility for identification yield sharper inference than a standard proxy-VAR, and the use of heteroskedasticity largely eliminates the problem of weak instruments.

4.4 Monetary Policy Analysis in a Heteroskedastic Proxy-VAR

We use our framework to provide new – and in light of the Monte Carlo evidence potentially sharper – estimates of the impact of monetary policy on the macro-economy. Our baseline model consists of three endogenous variables and an instrument for monetary policy shocks in the vector $z_t = [\Delta ip_t, ff_t, ebp_t, rr_t]'$. We use the first difference of the log of industrial production as a measure of real economic activity Δip_t, the federal funds rate as the monetary policy indicator ff_t, and a measure of corporate bond spreads ebp_t.[9] For the latter, we employ the excess bond premium of Gilchrist and Zakrajšek (2012) which approximates the tightness of credit markets. Caldara and Herbst (2016) show that it is important to control for the endogenous response of monetary policy to financial conditions to identify monetary policy shocks. Following the same authors, we use log differences of industrial production and include six lags of the endogenous variables to account for the persistence in the data. As an instrument for the latent monetary policy shocks, we take the narrative-based measure of unexpected changes in the intended fed funds rate of Romer and Romer (2004), rr_t.[10] This proxy starts in 1969M1 and, thereby, is the longest available potential instrument provided by the literature.

We estimate the VAR on monthly frequency data within the sample 1973M1 to 2007M6. The start is dictated by the availability of data on the excess bond premium, while the end is chosen such as to ensure that our sample is not affected by the zero lower bound or by unconventional monetary policy. We perform an extensive sensitivity analysis of the main results. We change the number of lags as well as states, the sample period, the monetary policy indicator, and the transformation of the variables. We also estimate a model including producer and commodity prices. The robustness results are summarized in Appendix 4.C and show that our main results hold.

4.4.1 Model Specification

An important choice in our framework is the way changes in volatility are modeled. The functional form of the model affects the likelihood and thereby the estimators and tests. Therefore, we perform an extensive model comparison before turning to the model based inference. As candidates we model heteroskedasticity through a smooth transition in variances using either a 12-month trailing moving average of industrial production or

[9]Industrial production is series $INDPRO$ downloaded from the FRED database of the Federal Reserve Bank of St. Louis, the federal funds rate is series $FEDFUNDS$ from the same database. The excess bond premium of non-financial firms is identical to the data used in Gilchrist and Zakrajšek (2012) and was downloaded from Gilchrist's website.

[10]We use the updated version of the original Romer and Romer (2004) constructed by Wieland and Yang (2016) downloaded from Wieland's webpage.

time as the transition variable, an exogenous break point iterating over all potential break points in the sample, a multivariate GARCH process, and a Markov switching framework with two states.[11] Chapter 2 of this dissertation describes all models formally.

Table 4.5: Model Selection

Reduced form models	$\log l$	BIC	AIC	HQ
White noise residuals	-692.91	2046.78	1605.81	1780.32
Smooth transition in variances (IP)	-452.29	1637.66	1148.59	1342.13
Exogenous breakpoint	-400.33	1527.73	1042.67	1234.63
Smooth transition in variances (time)	-293.41	1319.90	830.82	1024.36
GARCH residuals	-286.41	1317.91	820.81	1017.53
Markov switching	-212.60	1158.27	669.20	862.74

Note: $\log l$ denotes the likelihood function evaluated at the optimum, AIC $= -2(\log l) + 2f$, HQ $= -2(\log l) + 2f \times \log(\log(T))$ and BIC $= -2(\log l) + \log(T)f$, where f is the number of free parameters and T the number of observations.

Table 4.5 shows specification statistics for a linear model and the five alternative volatility models. Information criteria are shown to work well for judging the performance of MS models (Psaradakis and Spagnolo, 2006), whereas standard tests are problematic for this purpose as some parameters might not be identified under the null hypothesis of a smaller number of states than under the alternative (Hansen, 1992). Lütkepohl and Schlaak (2018) show that the AIC succesfully chooses between alternative volatility models. The models are ordered ascending according to their log-likelihood values. This ranking coincides with that implied by all three information criteria, leaving little ambiguity about the relative performance of the alternative models. The table conveys two important results. First, the linear model is dominated by all models that allow for changes in volatility. This is strong evidence in support of the assumption of heteroskedasticity. In this case, using any time-varying volatility model seems to be more in line with the data than using a linear model. Second, the Markov switching model is clearly preferred over the other heteroskedastic models according to both the log-likelihood and all three information criteria. Lütkepohl and Schlaak (2018) show that the Markov switching model is usually the best choice even in cases where the volatility specification does not coincide with the data generating process. Modeling changes in volatility through a latent variable gives full voice to the data, reducing the risk of misspecification of the transition variables, functions or break points.

[11] We also estimate smooth transition in variances models using either 6 or 24-month moving averages of industrial production. Both models perform worse than the 12-month version so we do not report them in the table.

Table 4.6: Invertibility Test of Structural MA-Representation of VAR

Equation	Δip	ff	ebp	rr
p-value	0.21	0.56	0.83	0.65

Notes: The table shows p-values for a robust F-statistic testing the null hypothesis that the coefficients on six lags of s_t are jointly equal to zero in each of the VAR equations of the Markov switching proxy-VAR.

Based on Table 4.5 we choose the Markov switching model, but we show that our results are robust to the choice of the volatility model. In Section 4.4.5, we use smooth transition models. They are popular in the VAR literature as they provide economic insights into the driving forces of the states. In Appendix 4.C, we show that our results also hold in a three-state Markov switching model. We favor a two-state model as the baseline since two states are economically more intuitive and easier to interpret. Moreover, a two-state model leads to more stable and precise estimates given that the third state contains only few observations.

Finally, to see whether the matrix polynomial of the structural moving average (MA) representation of the VAR model is invertible, i.e., the shocks we consider in our model are fundamental, we exploit that the augmented model (4.8) naturally lends itself for Granger-causality tests. Invertibility implies that adding past shocks to the system does not improve the forecasting capacity of the VAR model. Consequently, following Stock and Watson (2018), we test the null hypothesis that the six lags of the instrument are jointly equal to zero in each of the VAR equations. Table 4.6 displays that the null is not rejected in any VAR equation, indicating that there is no statistically significant evidence against the hypothesis of invertibility of the structural VAR moving average representation. This result is in line with the findings of Stock and Watson (2018), who do not reject the assumption of invertibility of the MA representation of the model of Gertler and Karadi (2015), which is similar to our specification.

To facilitate the comparison with the literature using the narrative measure of Romer and Romer (2004), we set the autoregressive coefficients for the instrument, that is the respective elements of $\Gamma(L)$ in (4.10), to zero in the subsequent analysis. In Table 4.19 in Appendix 4.C we show that our main results do not change when we leave the autoregressive and constant part of the model fully unrestricted.

4.4.2 Volatility Regimes and Identification

Table 4.7 reports the estimated state-dependent reduced form covariance matrices indicating whether the model detects switches in volatility, which are one important element in our identification and testing strategy. This information also helps us interpret our endogenously and agnostically identified regimes. Clearly, there are increases in volatility in state 2 for all residuals. The error variances in the equations for industrial production, the federal funds rate, the excess bond premium and the instrument increase by factors of approximately 3, 54, 10 and 20, respectively. In particular, the volatility of the error in the interest rate equation changes strongly across regimes. We read this as further evidence that the sample is characterized by changes in volatility. Moreover, the model seems able to detect and separate them.

The table also shows that the covariances increase (in absolute value) in state 2 as well, and by larger factors than the variances. These changes in the covariances illustrate the idea behind identification through heteroskedasticity. In a period where interest rates are highly volatile, we learn more about the relation between the federal funds rate and economic activity as the covariance between both temporarily increases. Monetary policy shocks are then more likely to occur and can be used as a 'probabilistic instrument' (see Rigobon, 2003) to trace out the response of production.

Table 4.7: Estimated State Covariance Matrices ($\times 10^3$) of Reduced Form Model

State 1: $\tilde{\Sigma}_1$	State 2: $\tilde{\Sigma}_2$
$\begin{bmatrix} 281.54 & & & \\ 10.60 & 32.96 & & \\ -1.03 & 0.40 & 22.00 & \\ 0.18 & 7.73 & -0.92 & 21.08 \end{bmatrix}$	$\begin{bmatrix} 864.71 & & & \\ 497.08 & 1835.58 & & \\ -46.06 & 27.88 & 194.85 & \\ 206.86 & 415.70 & 57.08 & 413.46 \end{bmatrix}$

To achieve identification from a statistical point of view and to be able to test the validity of the external instrument, we need significant and differential changes in the volatility of the structural innovations μ_t. Table 4.8 shows the estimated variances of the structural model (4.10) with unrestricted β in state 2, which are contained in Λ_2. Given the restrictions on D and that the instrument is ordered last in z_t, λ_{42} captures the change in the variance of the noise in the measurement of the instrument. As the ordering of the remaining $\lambda_{\cdot 2}$s is arbitrary, we simply order them from largest to smallest. All of these three estimates are significantly larger than one, implying that the volatility of all structural shocks increases when switching from state 1 to state 2. Thus, together with the evidence in Table 4.7, we label state 2 the high volatility state. Identification requires that the variance shifts are all distinct from each other. According to their estimated standard errors and the respective one standard deviation confidence bands around the point estimates, there is evidence for the assumption of distinct variance shifts. In consequence, the decomposition in (4.11) is locally unique and can be used to test the validity of the instrument.

Table 4.8: Estimates and Standard Errors of Relative Variances

Parameter	Estimate	Standard error
λ_{12}	55.57	10.16
λ_{22}	8.90	3.03
λ_{32}	2.62	0.53
λ_{42}	16.79	4.80

Notes: The standard errors are obtained from the inverse of the negative Hessian evaluated at the optimum of the structural model (4.10) with $z_t = [\Delta ip_t, ff_t, ebp_t, rr_t]'$.

To develop an economic notion about the statistically identified regimes we plot the smoothed state probabilities in Figure 4.1. The upper part corresponds to state 1 and the lower part to state 2. State 1 dominates the sample, in particular its second half. The model detects a long spell of low volatility during a period that is often referred to as the 'Great Moderation' in the 1990s and 2000s with stable growth and inflation under the chairmanship of Alan Greenspan. The high volatility regime appears more often during the first part of the sample. Many of the spikes in the probability of being in state 2 are associated with specific events in the economic history of the U.S. For instance, there are peaks around the energy crisis and the subsequent recession in the middle of the 1970s and at the beginning of the 1980s. There is also a longer-lasting switch to state 2 which coincides with the chairmanship of Paul Volcker at the end of the 1970s and the first half of the 1980s. In the second part of the sample, there are peaks around the burst of the dot-com bubble in 2001, the 9/11 attacks, and the subsequent recession. Overall, this short narrative, while only suggestive, indicates that the endogenously determined volatility regimes capture relevant developments in the U.S. real economy and in the conduct of monetary policy in our sample.

4.4.3 Instrument Validity

We now use these significant and distinct changes in the variances of the structural innovations to test the validity of the instrument. Given validity, we combine the information in the instrument with that in the second moments of the data and estimate the dynamic effects of monetary policy shocks.

First, we test for exogeneity. We leave β_1 unrestricted, thereby ordering the monetary shock first. The null hypothesis is that the instrument is exogenous to the non-monetary policy shocks, $H_0 : \beta_2 = \beta_3 = 0$, against the alternative that the instrument is endogenous. Table 4.9 shows that the data do not reject the assumption of exogeneity. The LR-statistic is quite small and the p-value not close to conventional significance levels. Thus, the instrument passes the first stage.

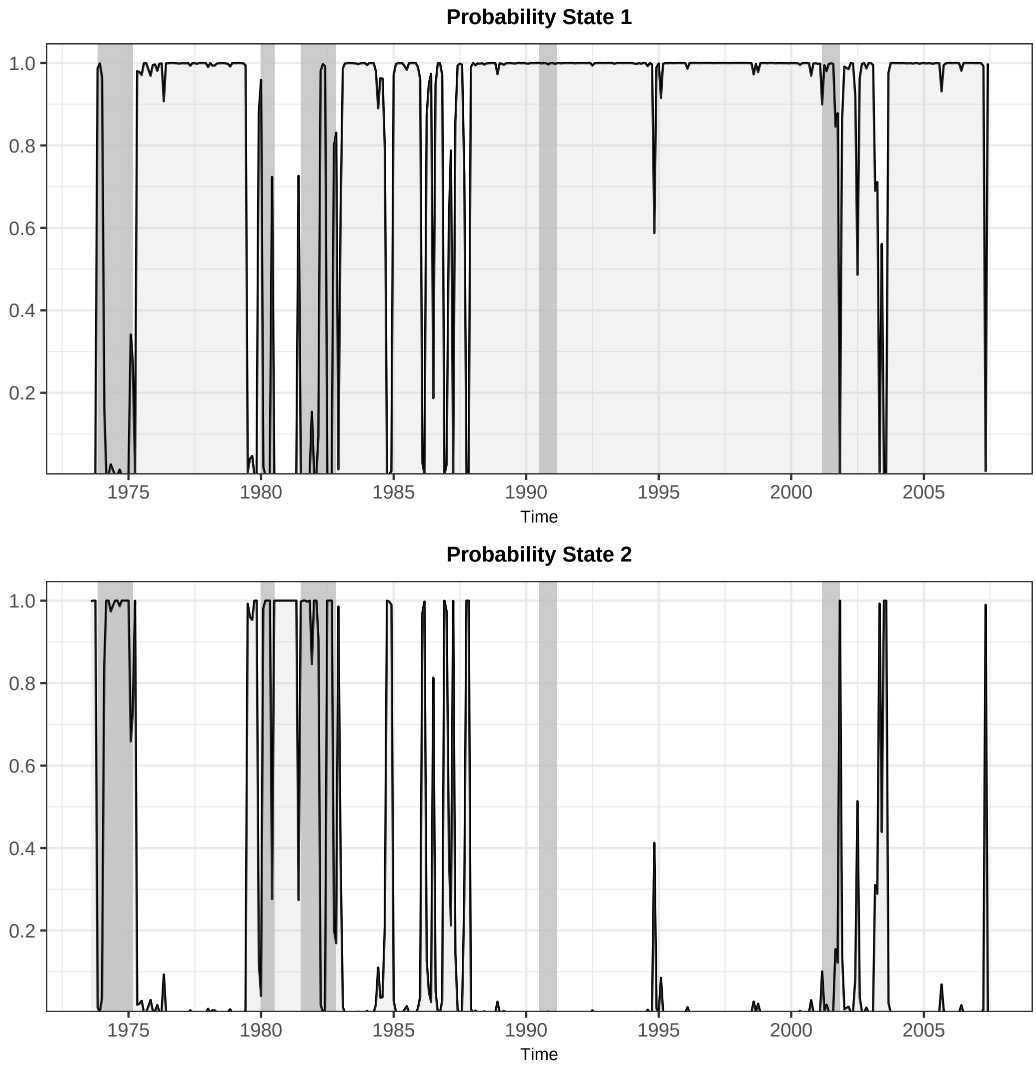


Figure 4.1: Smoothed state probabilities.

Notes: The figure shows the smoothed state probabilities for $m = 1$ in the upper panel and for $m = 2$ in the lower panel of model (4.10) with $z_t = [\Delta ip_t, ff_t, ebp_t, rr_t]'$. The shaded vertical bars mark recession periods defined by the NBER.

Table 4.9: Test for Instrument Validity

	Exogeneity	Relevance
LR statistic	0.04	43.32
p-value	0.97	0.00
Restrictions	2	1

Notes: The table shows the LR statistic, the p-value and the number of restrictions for the tests of instrument exogeneity ($H_0 : \beta_2 = \cdots = \beta_K = 0, \quad H_1 : \beta$ unrestricted) and instrument relevance ($H_0 : \beta_1 = 0, \quad H_1 : \beta_1 \neq 0$). The instrument is the narrative-based measure of monetary surprises of Romer and Romer (2004).

To assess its relevance, we test the null hypothesis that the instrument is unrelated to all structural shocks, $H_0 : \beta = 0$, against the alternative that it is significantly related to at least one structural shock. If the null is rejected, this will be the monetary policy shock given the first stage result and that the instrument is constructed to have a high correlation with the monetary shock and a low (zero) correlation with the other shocks. The table shows that the instrument is highly relevant. The null is rejected at the 1% significance level. We conclude that the narrative-based measure of Romer and Romer (2004) is relevant and, thus, a valid instrument for monetary policy shocks.

4.4.4 Dynamic Effects and Importance of Monetary Shocks

We now estimate the dynamic effects of monetary shocks and quantify their economic importance for output and credit spread fluctuations. Based on the testing sequence, we leave β_1 unrestricted and set $\beta_2 = \beta_3 = 0$. This implies that the estimation combines the information contained in the instrument and in the second moments of the data for identification of a heteroskedastic proxy-VAR.

Figure 4.2 shows the impulse responses to all three shocks in columns on the endogenous variables in rows. The inclusion of the proxy into the model is a key advantage over traditional identification through heteroskedasticity where a main challenge is the economic labeling of the statistically identified shocks (Herwartz and Lütkepohl, 2014). Due to our restrictions on β, we can clearly label the monetary shock, which is pinned down in the first column of D. A one standard deviation monetary surprise corresponds to an increase in the federal funds rate by 18 basis points. According to the 95 percent confidence bands that are computed using the bootstrap method outlined in Section 4.2.4, credit spreads increase significantly a few months after the shock and remain elevated for several years, before gradually returning back to trend. Economic activity declines significantly a year and a half after the occurrence of the shock. The response bottoms toward the end of the propagation horizon with a cumulative effect of –0.5 percent.

Quantitatively, the dynamics of real activity are similar to those implied by the hybrid VAR models of Romer and Romer (2004) and Coibion (2012), who include the cumulated monetary policy surprise measure directly into a VAR together with industrial production and producer prices. They document a 2.9 percent trough effect for a 100 basis points shock. Qualitatively, our results differ from their estimates which suggest a short-lived increase in real activity but then a quick and significant decline after about six months. Romer and Romer (2004) trace the initial hump back to a single observation (a large negative surprise in the intended fed funds rate on April 1980 coupled with a strong decline in production) and the sampling error due to this extreme event. While we also

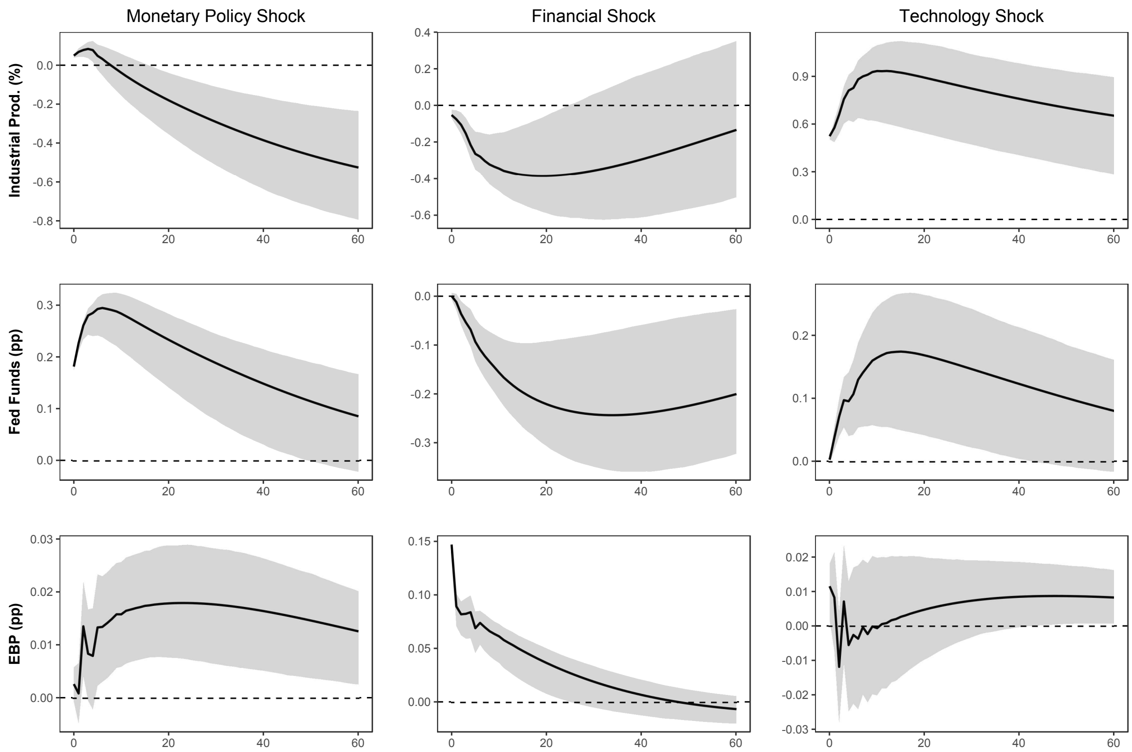


Figure 4.2: Impulse responses for heteroskedastic proxy-VAR.

Notes: The figure shows the impulse responses to one standard deviation shocks in state $m = 1$ of the heteroskedastic proxy-VAR(6) model with $M = 2$ states for $z_t = [\Delta ip_t, ff_t, ebp_t, rr_t]'$. The sample is 1973M1-2007M6 and the instrument for monetary policy shocks is the narrative-based measure of Romer and Romer (2004). The shaded bands denote 95 percent pointwise confidence intervals based on 5,000 bootstrap replications.

obtain the hump, the following decline is more sluggish and more persistent. Output falls below trend only after a year and a half. These dynamics are similar to Ramey (2016).

We next assess the economic importance of monetary policy shocks for output and credit spread fluctuations. The regime-specific forecast error variance decompositions in Figure 4.3 show that the contribution of monetary shocks to the variability of the endogenous variables is highly state-dependent.[12] In the high volatility regime, monetary shocks account for up to 40 percent of the variance of production and spreads at longer horizons. In the low volatility regime, they each explain less than 10 percent. They also account for a much larger share of the variance in the federal funds rate in the high volatility regime than in the low volatility regime.

We briefly discuss the effects of the other two structural shocks, which are mainly identified using the changes in volatility and yet need to be labeled (if they are of interest to the researcher). Our framework also simplifies this task compared to traditional identifi-

[12] We evaluate the forecast error variance decompositions conditional on the respective MS-states, i.e., we assume that no transition between the states occurs.

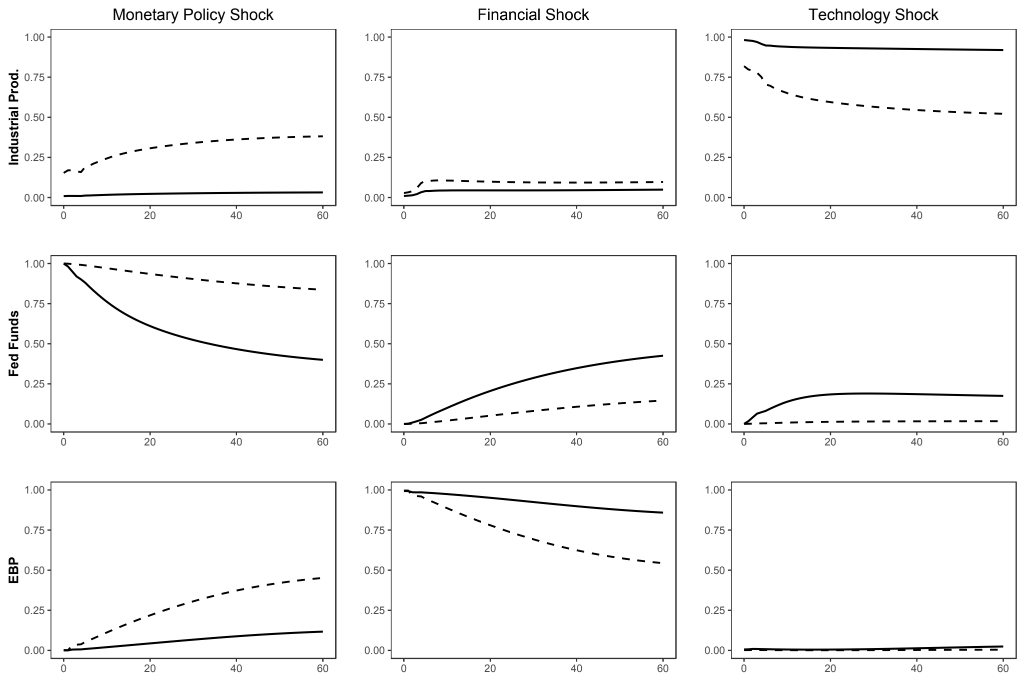


Figure 4.3: Variance decompositions for heteroskedastic proxy-VAR.

Notes: The figure shows the regime-specific forecast error variance decompositions (solid line - state 1; dashed line - state 2) for the structural shocks in columns on the endogenous variables in rows for the heteroskedastic proxy-VAR(6) model with $M = 2$ states and $z_t = [\Delta ip_t, ff_t, ebp_t, rr_t]'$. The sample is 1973M1-2007M6 and the instrument for monetary policy shocks is the narrative-based measure of Romer and Romer (2004).

cation through heteroskedasticity. The inclusion of a valid proxy for monetary policy into the model gives an economic interpretation to the main shock of interest and separates it from the remaining shocks. Therefore, the latter are easier to label. This is reflected in relatively clear sign patterns of the impulse responses for the other two shocks and in the forecast error variance decomposition, which both suggest two simple labels.

The second shock accounts for virtually all of the variability in the excess bond premium upon impact and for more than 60 percent in both states in the long-run. Thus, we label it a financial shock. An exogenous 15 basis points increase in credit spreads leads to an immediate contraction in real activity (see Figure 4.2). This is followed by a hump-shaped negative response, with a trough of –0.4 percent, and a gradual return back to trend after about three years. The monetary authority responds by lowering the policy rate by up to 20 basis points after two years to offset the adverse effect of tighter financial conditions on production.

The remaining shock accounts for a minimum of 80 percent of the variance of industrial production on impact. At longer horizons, it explains more than half of the variance in the high volatility state and more than 90 percent in the low volatility state. Thus, we label it a technology shock. In response, industrial production jumps up immediately and reaches a peak after one year. Monetary policy aims to counter the expansion by raising the interest rate by 15 basis points after one year, which thereafter reverts to trend together with real activity. The excess bond premium hardly responds to the shock. Toward the end of the horizon there is a mildly positive response, consistent with the monetary tightening.

In light of the impulse response analysis and the results from the testing sequence, it is not surprising that we find similar effects of monetary shocks as Romer and Romer (2004). The tests show that the narrative-based measure is a valid instrument for latent monetary policy shocks. The impulse response analysis shows that the instrument is neither endogenous to economic activity shocks nor to financial shocks. Thus, a main contribution of our approach is to increase the confidence in estimates based on this instrument. The results suggest that it can be used in models like ours or in the Bayesian model of Caldara and Herbst (2016) to reliably estimate the dynamic effects of monetary policy shocks.[13]

The simulation study suggests that another advantage of modeling heteroskedasticity – which is present in our sample as shown in Section 4.4.2 – is that the information contained in the second moments helps identifying the structural model and, hence, the effects of monetary policy more accurately. This is corroborated by Figure 4.4, which compares the impulse responses to a monetary policy shock from the heteroskedastic (left column) to those from a standard proxy-VAR (right column). The shock is scaled to 25 basis points for comparison. Qualitatively, both models yield the same conclusions. Corporate bond spreads increase and production declines.

However, quantitatively and in terms of economic significance, there are notable differences. In the heteroskedastic proxy-VAR, the monetary shock is more hump-shaped, reaching a peak only after ten months, and estimated to be more persistent. The federal funds rate remains significantly above trend for about 48 months, whereas in the standard proxy-VAR it is indistinguishable from zero after roughly 24 months. This stronger and longer-lasting monetary contraction leads to a quicker, larger and more persistent drop in industrial production, which falls significantly below trend after a year and a half, declining cumulatively 0.7 percent. In contrast, in the standard model, the decline in economic activity is more sluggish, the effect is only borderline significant after three years, and the trough is only –0.4 percent. Similarly, the effect of the monetary shock on credit

[13] The advantages of using the measure as an instrument and accounting for measurement error instead of using it directly as a variable in a regression or VAR are discussed in Stock and Watson (2012) and Mertens and Ravn (2013).

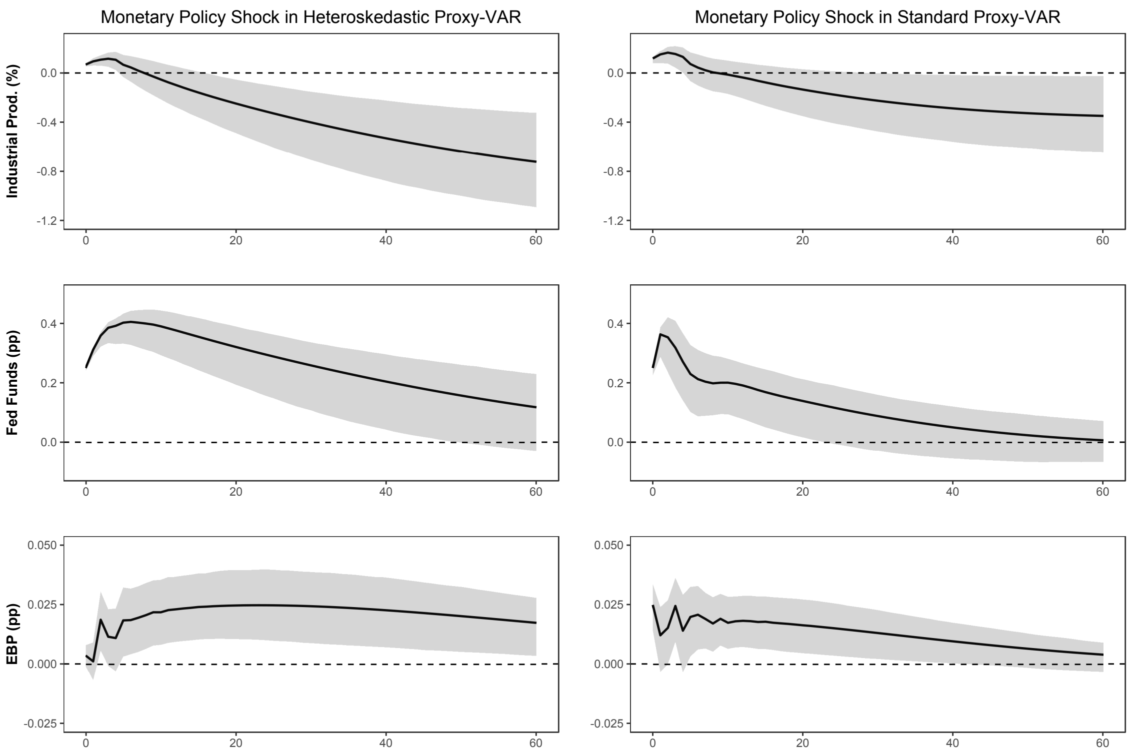


Figure 4.4: Comparison of heteroskedastic and standard proxy-VAR.

Notes: The figure shows the impulse responses to a monetary policy shock of 25 basis points in a heteroskedastic (left column) and standard proxy-VAR (right column) on the endogenous variables in rows. The model contains $z_t = [\Delta ip_t, ff_t, ebp_t, rr_t]'$, the sample is 1973M1-2007M6 and the instrument for latent monetary shocks is the narrative-based measure of Romer and Romer (2004). The shaded bands denote 95 percent pointwise confidence intervals based on 5,000 bootstrap replications.

spreads is stronger, longer-lasting, and more statistically significant in the model using the time-varying volatility.

These differences are due to alternative uses of the existing information in the data. While the heteroskedastic proxy-VAR draws on both the instrument and changes in volatility for identification, the standard proxy-VAR discards the second piece of information. To see whether the assumption of heteroskedastic structural innovations is consistent with the data we now employ information criteria for the comparison of the *structural* models. Since the standard proxy-VAR point-identifies only one column, it is under-identified and its likelihood and information scores are the same as for the linear reduced form model. They are shown in Table 4.5. In contrast, the heteroskedastic proxy-VAR is over-identified, such that its likelihood deteriorates slightly (to –216.29) relative to the reduced form Markov switching model. Nevertheless, the information criteria favor the more parsimonious structural over the reduced form model. All three criteria improve; to 1135.62, 666.69, and 852.20 for the SC, AIC, and HQ, respectively. This reflects the earlier result from the LR-tests for instrument exogeneity of not rejecting the over-identifying restrictions. More importantly, all information criteria clearly prefer the heteroskedastic over the standard proxy-VAR. We conclude that the former provides sharper inference because it exploits time-variation in second moments and that the latter underestimates the effects of monetary policy.

4.4.5 Smooth Transition in Variances

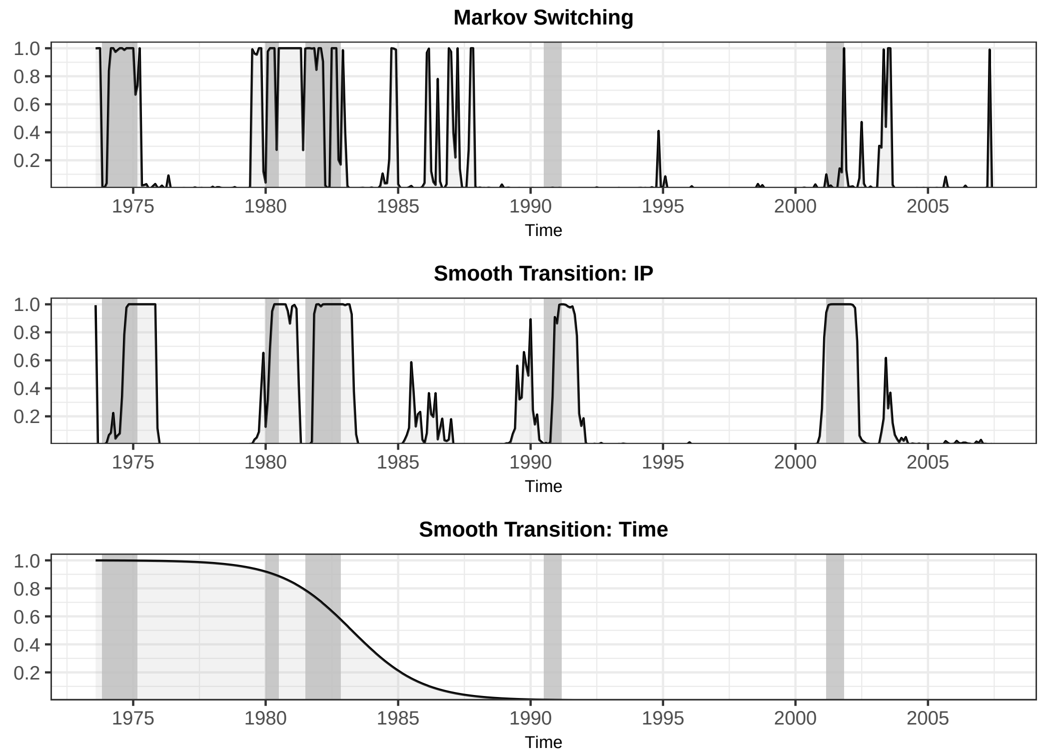


Figure 4.5: Volatility states 2 of Markov switching and smooth transition models.

Notes: The figure shows the probability volatility state 2 of heteroskedastic proxy-VARs, using Markov switching (upper panel) and smooth transition in variances based on a 12-month moving average of industrial production (middle panel) or on time (lower panel) as transition variable. The model is $z_t = [\Delta ip_t, ff_t, ebp_t, rr_t]'$. The shaded vertical bars mark recession periods defined by the NBER.

In this subsection, we assess whether our results are sensitive to using an alternative volatility model as both the test results and the impulse responses depend on the functional form of the heteroskedasticity. We focus on the smooth transition in variances specifications as they add economic insight to the drivers of the states. We consider the version with a 12-month trailing moving average of industrial production and with time as transition variable to see whether there is a systematic relation between high volatility and recessions and a transition in shock variances from the Volcker to the Greenspan area, respectively.

Figure 4.5 compares the volatility states across models. The smooth transition based on industrial production estimates roughly similar states as the Markov switching model.

The correlation between the two states is 0.36. Moreover, it captures most of the NBER recessions, although with some delay which reflects the use of a trailing moving average as transition variable. The model based on time also does an acceptable job. It matches the transition in the Fed Chairmanship occurring in 1987. The correlation with the Markov switching state is 0.43.

Table 4.10 shows the estimated structural variances and their standard errors to assess whether the smooth transition in variances adds identifying information to the model. The variance increases in state 2 have the same ranking in both models and as in the Markov switching model. The monetary policy shock and the measurement error shift most in variances, followed by the financial and technology shock, although the variances tend to be less precisely separated between regimes than in the Markov switching model. Importantly, the structural shocks are all statistically identified given the usual metric of non-overlapping one standard error confidence intervals that are constructed around the respective point estimates.

Table 4.10: Estimates and Standard Errors of Relative Variances for Smooth Transition Models

Transition variable	Industrial production		Time	
Parameter	Estimate	Standard Error	Estimate	Standard Error
λ_{12}	22.41	1.99	69.61	13.32
λ_{22}	3.01	0.34	2.93	0.51
λ_{32}	2.25	0.26	0.94	0.19
λ_{42}	6.18	0.59	11.14	2.64

Therefore, we proceed to test whether the instrument is valid. Table 4.11 shows that this is the case. Both models come to the same conclusion as the Markov switching model. The narrative-measure of Romer and Romer (2004) is exogenous and relevant. Finally, we study the impulse responses implied by the smooth transition models. Figure 4.6 shows that they are not statistically distinguishable from those of the Markov switching model, although both imply smaller effects more similar to the proxy-VAR that does not exploit the heteroskedasticity. The latter finding reflects that both models are less successful in separating the volatility regimes, using less of the identifying information in the data and yielding blunter inference. Overall, however, the subsection shows that our framework and results do not depend on a specific volatility model.

Table 4.11: Test for Instrument Validity Based on Smooth Transition in Variances

Transition variable	Exogeneity p-value	Relevance p-value
Industrial production	0.313	0.000
Time	0.715	0.000

Notes: The table shows the p-values of LR-tests for the exogeneity and relevance of the measure of Romer and Romer (2004) as instrument s_t for monetary policy shocks in smooth transition in variances models using different transition variables. The model is $z_t = [\Delta ip_t, ff_t, ebp_t, s_t]'$ and sample period is 1973M1-2007M6.

4.4.6 Testing Alternative Proxies for Monetary Shocks

In this section, we test and compare alternative measures of monetary surprises proposed in the literature to study the effects of monetary policy. In addition to the narrative measure of Romer and Romer (2004), we consider the identified monetary shocks from the SVAR of Bernanke and Mihov (1998) and monetary surprises identified using high(er) frequency data. For the latter, we employ measures derived from changes in federal funds futures data around policy announcements using a daily window (see Barakchian and Crowe, 2013), a 30-minutes window (see Gertler and Karadi, 2015), and a 30-minutes window including further cleaning of the surprises by regression on a range of control variables (see Miranda-Agrippino and Ricco, 2017). We consider the potential instruments one at a time.

To establish a level playing field and to facilitate a clean comparison, we use a common sample period for the evaluation of the proxies although they are available for slightly different periods. As most high(er) frequency proxies start only in the 1990s, we use the same sample (and model) as Caldara and Herbst (2016), which is 1994M1-2007M6. The first row in Table 4.12 shows the test results when re-estimating the model using the narrative-based measure on the shorter sample period. The conclusion from above based on the longer sample hold. The instrument is valid. The next row shows that the same assertion applies to the model-based measure.

The picture is more mixed for the instruments based on high-frequency data. While none appear to suffer from endogeneity, only the plain changes in the fourth federal funds futures in a 30-minute window around policy announcements are a relevant proxy.[14] Both the instrument using daily data and the cleaned instrument do not meet the relevance condition according to the LR-test and conventional significance levels. One possible

[14] Bertsche and Braun (2017) aim at assessing the estimated moment conditions implied by the instruments of Romer and Romer (2004) and Gertler and Karadi (2015) in non-nested models using a stochastic volatility framework. However, the distribution and properties of their test statistic are unclear as the parameters under the null hypothesis are estimated and thus random variables.

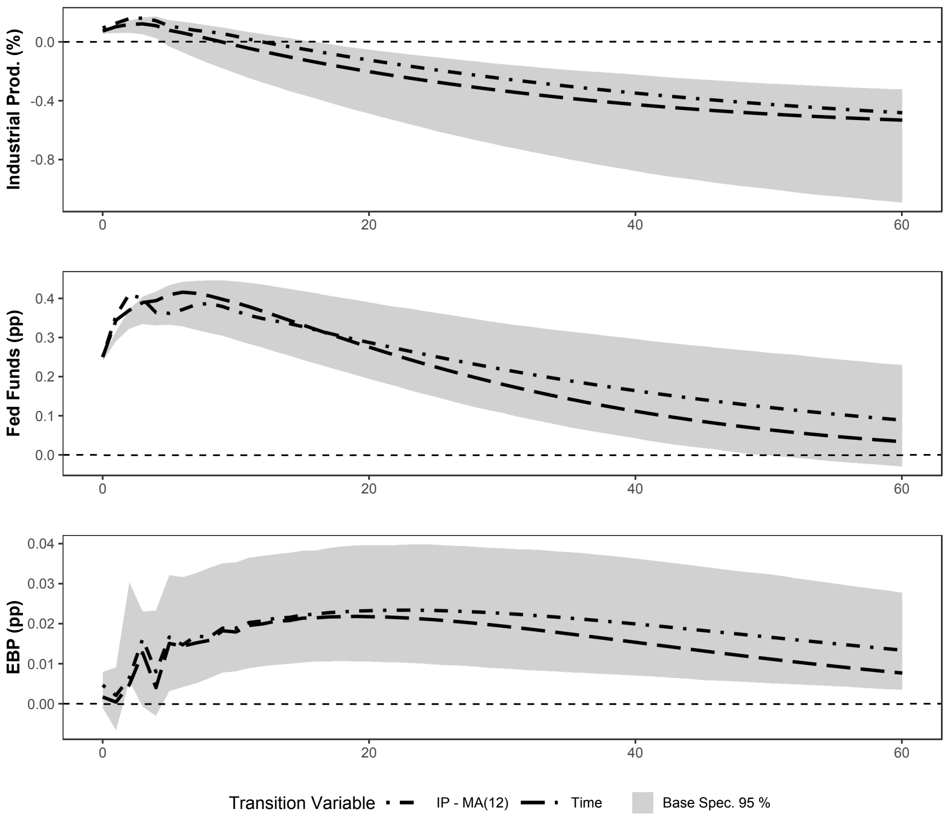


Figure 4.6: Comparison of smooth transition with baseline Markov switching model.

Notes: The figure shows impulse responses to a 25 basis points monetary policy shock of the smooth transition in variances SVAR using industrial production as transition variable (dashed line) or time (dash-dotted line). The shaded area denotes 95 percent pointwise confidence intervals based on 5,000 bootstrap replications of the baseline Markov switching proxy-VAR(6).

Table 4.12: Testing Validity of Alternative Instruments

Instrument	Exogeneity p-value	Relevance p-value
Narrative-based	0.427	0.003
Model-based	0.830	0.000
High(er) frequency data		
Daily data	0.485	0.302
30-minute window	0.599	0.002
30-minute window and cleaned	0.897	0.238

Notes: The table shows the p-values of LR-tests for the exogeneity and relevance of different instruments s_t in the model with $z_t = [\Delta ip_t, ff_t, ebp_t, s_t]'$, testing them one at a time. The sample period is 1994M1-2007M6. The narrative-based measure is of Romer and Romer (2004), the model-based measure is of Bernanke and Mihov (1998), and measures based on high(er) frequency data are taken from Barakchian and Crowe (2013), Gertler and Karadi (2015) and Miranda-Agrippino and Ricco (2017), respectively.

explanation for this finding is that the former instrument may be too noisy and the latter stripped of too much relevant information through the regression on further control variables. These results, of course, are conditional on the model specification and sample used.

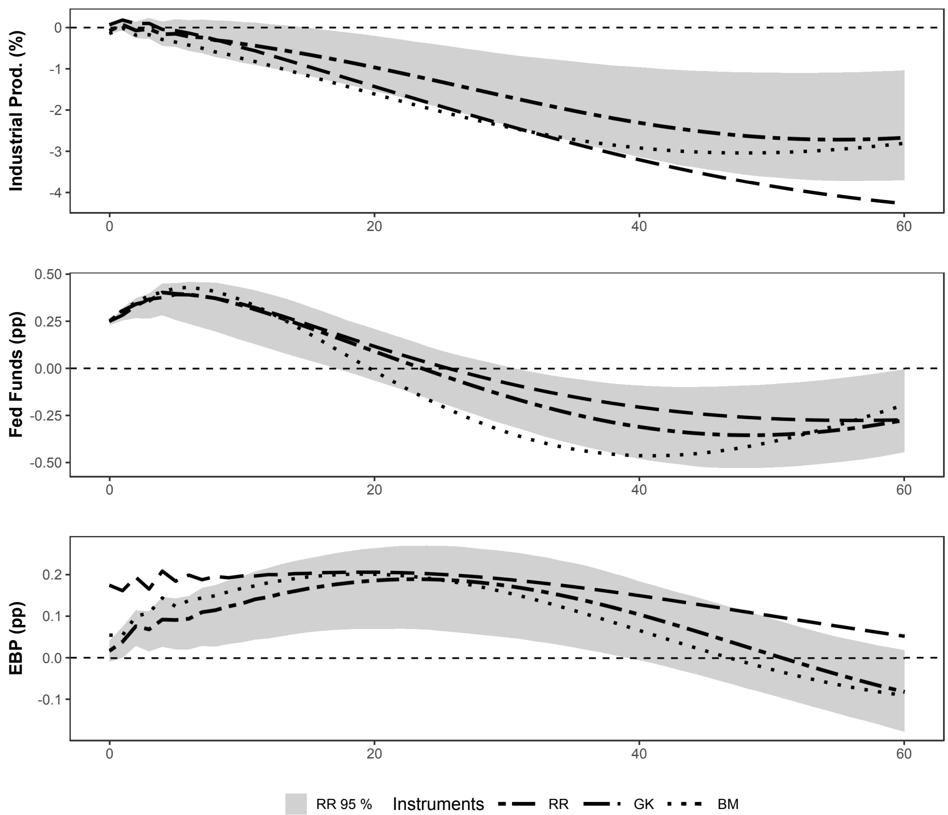


Figure 4.7: Impulse Responses for heteroskedastic proxy-VAR using different instruments.

Notes: The figure shows the impulse responses to a 25 basis points monetary policy shock. The sample is 1994M1-2007M6 and the different instruments, using one at a time, are the narrative measure of Romer and Romer (2004) (dashed-dotted line with shaded 95 percent pointwise confidence intervals based on 5,000 bootstrap replications), the high-frequency proxy of Gertler and Karadi (2015) (dashed line), and the model-based measure of Bernanke and Mihov (1998) (dotted line).

Focusing on the valid instruments for the given model and sample, Figure 4.7 compares the implied effects of a monetary policy shock on the endogenous variables. The solid lines show the estimates using the narrative-based measure together with 95 percent confidence bands as a baseline for the comparison. Relative to the estimates using the full sample, the initial short increase in industrial production vanishes in the shorter sample starting in January 1994, consistent with the argument of Romer and Romer (2004) that this hump is related to the April 1980 observation. The other two proxies produce similar effects. Over most of the propagation horizon, the responses are not distinguishable from those implied by the narrative-based measure. The monetary shock identified using the high-frequency proxy has a larger effect on the excess bond premium upon impact and on industrial production toward the end of the propagation horizon. Overall, however, the three valid proxies lead to similar conclusions. This is reassuring and suggests that the main findings of the previous sections based on a longer sample provide a reasonable description of the effects of monetary policy. A natural next step is to include several valid proxies for monetary shocks simultaneously into the model. However, this is beyond the scope of the paper and left for future research.

4.5 Conclusions

We propose an econometric framework in the form of a structural vector autoregression that combines the information contained in an external instrument and in time-varying second moments of the data for identification of latent monetary policy shocks in the U.S. We show that the framework improves the identification of the structural model and leads to sharper inference. Moreover, it allows testing the validity of the chosen instrument, thereby increasing the credibility and reliability of the estimation results for policy analysis. Given sufficient heteroskedasticity in the data the framework also largely dispenses the proxy-VAR approach from problems arising from weak instruments. Finally, it facilitates an economic interpretation of the structural shock of interest, which is not only identified statistically through heteroskedasticity but also through prior economic reasoning contained in the instrument.

We apply the framework to test the validity of using the narrative measure of monetary surprises of Romer and Romer (2004) as an instrument for monetary policy shocks. We find that it is a valid instrument in our model and sample. We use it and combine it with the heteroskedasticity in the data to provide new and potentially sharper estimates of the dynamic effects of monetary policy on the macro-economy. We find that a surprise monetary contraction of 25 basis points in the federal funds rate leads to a significant increase in corporate bond spreads and to a significant decline in real economic activity of cumulatively 0.7 percent. In contrast, a standard proxy-VAR that does not use the time-variation in second moments implies substantially smaller effects. The results further suggest significant changes in the volatility of monetary shocks over time and that the shocks explain a large share of real and financial fluctuations in the 1970s and 1980s, but only a small share during the Great Moderation under the chairmanship of Alan Greenspan.

Finally, we evaluate different proxies for monetary policy proposed in the literature and find that instruments based on intra-daily data that are not further cleaned (Gertler and Karadi, 2015) and instruments from time-series models (Bernanke and Mihov, 1998) are also valid in our model and sample. They lead to qualitatively and quantitatively similar results as the narrative-based proxy.

Appendix

4.A Notes on Computation

All estimations of this paper use the statistical software R3.4.1. For maximization of the log-likelihood function the R-package 'nloptr' provides the optimization routine 'slsqp', a sequential (least-squares) quadratic programming algorithm for nonlinearly constrained, gradient-based optimization. This algorithm supports equality constraints and inequality constraints. The former are needed to implement zero restrictions in our model setup on the structural impact matrix. The latter are used to impose a lower bound of 0.001 on the diagonal elements of Λ_m for $m = 1, \ldots, M$ to avoid singularity of the covariance matrix (see Herwartz and Lütkepohl, 2014).

The zero restrictions in rows $(K+1)$ and columns $(K+1)$ of the autoregressive coefficient matrices Γ_p for $p = 1, \ldots, P$ and row $(K + 1)$ of the intercept term δ of model (4.8) are implemented using restricted ordinary least squares. The restrictions are updated at the end of each maximization step of the EM algorithm.

To generate starting values for the structural parameters D and Λ_m for $m = 2, \ldots, M$ for the estimation algorithm we follow Herwartz and Lütkepohl (2014) with two exceptions. First, we choose starting values of $D = \hat{\Sigma}^{1/2}\Omega$ with $\hat{\Sigma}$ being the estimated reduced form covariance matrix of the respective model and Ω being a random orthogonal matrix.[15] Choosing an orthogonal matrix Ω instead of adding a matrix of small random numbers as suggested by Herwartz and Lütkepohl (2014) covers a wider range of the parameter space of possible starting values for D. Second, starting values of Λ_m for $m = 2, \ldots, M$ are chosen as $\Lambda_m = \text{diag}(0.5k, 2.0k, 3.5k, 5.0k)^{m-1}$ with $k = 1, \ldots, 10$. For each k we draw 50 random orthogonal matrices Ω as starting values for D. Thus, the total number of distinct initial parameters for each model amounts to 500. We check convergence of the estimation algorithm using the relative changes of the log-likelihood function for each estimated model and choose the model that maximizes the likelihood among all converged models.

In the Monte Carlo study we rely on one draw of starting values for D to limit the computational burden. To make up for choosing only one initial parameter for D we set starting values for $\Lambda = (0.5, 2.0, 3.5, 5.0)$ to start the estimation algorithm in the proximity of the true parameter values. We did not encounter convergence problems of the estimation algorithm in the simulation study. The simulations are conducted with 50

[15] The matrix square root of $\hat{\Sigma}$ is computed by taking the square root of its eigenvalues in the well-known decomposition of a positive definite symmetric matrix L into its eigenvectors Z and eigenvalues E, $L = Z\text{diag}(E)Z'$.

cores (Intel Xeon Westmere X5650 processors) on the high performance computing server at Freie Universität Berlin.

4.B Supplementary Results of Monte Carlo Study

This section contains supplementary material to the Monte Carlo study carried out in Section 3 of the paper. All relevant information is in the captions and notes of the respective tables.

Table 4.13: Relative Rejection Frequencies at Nominal Significance Level of 5% of LR-Tests on Exogeneity of Instrument

Sample Size	Relevance (β_1,ρ_1)	Endogeneity (β_2,ρ_2)				
		(0.0,0.0)	(0.05,0.03)	(0.20,0.12)	(0.30,0.16)	(0.40,0.22)
$T = 200$	(0.0,0.0)	0.05	.	.	.	.
	(0.20,0.16)	0.05	0.06	0.27	.	.
	(0.40,0.30)	0.04	0.05	0.22	0.40	0.50
	(0.60,0.43)	0.04	0.05	0.14	0.29	0.43
$T = 500$	(0.0,0.0)	0.06	.	.	.	.
	(0.20,0.16)	0.05	0.09	0.76	.	.
	(0.40,0.30)	0.05	0.07	0.56	0.86	0.93
	(0.60,0.43)	0.05	0.06	0.37	0.68	0.85

Notes: Based on 500 replications of simulation experiment. Dots (.) denote combinations of values for β_1 and β_2 that produce lower correlations between the instrument s_t and the target structural shock of interest (ε_t^r) than between the instrument s_t and the endogenous structural shock (ε_t^z). These cases are not taken into account in the analysis.

Table 4.14: Relative Rejection Frequencies at Nominal Significance Level of 10% of LR-tests for Relevance of Instrument

Sample Size	Relevance (β_1,ρ_1)	Endogeneity (β_2,ρ_2)				
		(0.0,0.0)	(0.05,0.03)	(0.20,0.12)	(0.30,0.16)	(0.40,0.22)
$T = 200$	(0.0,0.0)	0.13	.	.	.	.
	(0.20,0.16)	0.83	0.83	0.86	.	.
	(0.40,0.30)	1.00	1.00	1.00	1.00	1.00
	(0.60,0.43)	1.00	1.00	1.00	1.00	1.00
$T = 500$	(0.0,0.0)	0.09	.	.	.	.
	(0.20,0.16)	1.00	1.00	1.00	.	.
	(0.40,0.30)	1.00	1.00	1.00	1.00	1.00
	(0.60,0.43)	1.00	1.00	1.00	1.00	1.00

Notes: Based on 500 replications of each simulation design. Dots (.) denote combinations of values for β_1 and β_2 that produce lower correlations between the instrument s_t and the target structural shock of interest (ε_t^r) than between the instrument s_t and the endogenous structural shock (ε_t^z). These cases are not taken into account in the analysis.

Table 4.15: Comparison of MSE of Impulse Responses to Monetary Policy Shock for $T = 200$ and Propagation Horizon up to $h = 25$

Relevance (β_1,ρ_1)	Endogeneity (β_2,ρ_2)														
	(0.0,0.0)			(0.05,0.03)			(0.20,0.12)			(0.30,0.16)			(0.40,0.22)		
Model	θ_{11}	θ_{21}	θ_{31}	θ_{11}	θ_{21}	θ_{31}	θ_{11}	θ_{21}	θ_{31}	θ_{11}	θ_{21}	θ_{31}	θ_{11}	θ_{21}	θ_{31}
(0.0,0.0)															
Model (A)	1.00	1.00	1.00	.	.	.	.	.	.	.	.	.	.	.	.
Model (B)	0.99	0.98	1.00	.	.	.	.	.	.	.	.	.	.	.	.
Model (C)	28.89	21.16	48.05	.	.	.	.	.	.	.	.	.	.	.	.
(0.20,0.16)															
Model (A)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.	.	.	.	.	.
Model (B)	0.96	0.99	0.97	0.97	0.92	0.98	0.81	0.42	0.86	.	.	.	.	.	.
Model (C)	19.47	13.10	23.71	20.05	12.21	24.80	14.89	7.73	18.49	.	.	.	.	.	.
(0.40,0.30)															
Model (A)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Model (B)	1.05	1.35	1.09	1.05	1.17	1.08	0.80	0.45	0.80	0.60	0.25	0.63	0.53	0.19	0.53
Model (C)	15.73	10.35	20.09	15.63	8.19	19.93	10.52	3.65	12.87	7.16	2.75	8.49	5.65	2.57	6.11
(0.60,0.43)															
Model (A)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Model (B)	1.12	1.60	1.16	1.12	1.45	1.16	0.94	0.66	1.00	0.75	0.37	0.78	0.55	0.22	0.56
Model (C)	15.24	9.89	20.00	15.08	8.07	20.08	11.99	3.39	15.98	8.53	2.26	11.08	5.61	1.76	7.08

Notes: The table shows the cumulated MSE of fitted models (1)-(3) relative to model (1) for a propagation horizon up to $h = 25$ and sample size $T = 200$. Each entry is based on 500 replications of each simulation design. Dots (.) denote combinations of values for β_1 and β_2 that produce lower correlations between the instrument s_t and the target structural shock of interest (ε_t^r) than between the instrument s_t and the endogenous structural shock (ε_t^z). These cases are not taken into account in the analysis.

Table 4.16: Comparison of MSE of Impulse Responses to Monetary Policy Shock for $T = 200$ and Propagation Horizon up to $h = 5$

Relevance (β_1,ρ_1)	Endogeneity (β_2,ρ_2)														
	(0.0,0.0)			(0.05,0.03)			(0.20,0.12)			(0.30,0.16)			(0.40,0.22)		
Model	θ_{11}	θ_{21}	θ_{31}	θ_{11}	θ_{21}	θ_{31}	θ_{11}	θ_{21}	θ_{31}	θ_{11}	θ_{21}	θ_{31}	θ_{11}	θ_{21}	θ_{31}
(0.0,0.0)															
Model (A)	1.00	1.00	1.00	.	.	.	.	.	.	.	.	.	.	.	.
Model (B)	0.99	0.97	1.00	.	.	.	.	.	.	.	.	.	.	.	.
Model (C)	31.74	24.14	54.25	.	.	.	.	.	.	.	.	.	.	.	.
(0.20,0.16)															
Model (A)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.	.	.	.	.	.
Model (B)	0.98	1.01	0.99	1.01	0.94	1.00	0.79	0.41	0.87	.	.	.	.	.	.
Model (C)	25.19	14.94	27.04	26.06	13.84	28.21	15.90	7.77	19.82	.	.	.	.	.	.
(0.40,0.30)															
Model (A)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Model (B)	1.10	1.47	1.10	1.10	1.23	1.10	0.86	0.46	0.85	0.64	0.27	0.68	0.51	0.20	0.55
Model (C)	24.98	14.04	23.58	24.83	10.18	23.41	16.13	3.34	15.30	9.94	2.44	10.05	6.26	2.40	6.54
(0.60,0.43)															
Model (A)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Model (B)	1.19	1.78	1.17	1.19	1.57	1.18	1.09	0.68	1.06	0.85	0.38	0.85	0.60	0.24	0.63
Model (C)	26.36	15.01	23.91	26.08	11.41	24.03	21.50	3.44	19.62	14.61	1.88	13.79	8.68	1.39	8.76

Notes: The table shows the cumulated MSE of fitted models (1)-(3) relative to model (1) for a propagation horizon up to $h = 5$ and sample size $T = 200$. Each entry is based on 500 replications of each simulation design. Dots (.) denote combinations of values for β_1 and β_2 that produce lower correlations between the instrument s_t and the target structural shock of interest (ε_t^r) than between the instrument s_t and the endogenous structural shock (ε_t^z). These cases are not taken into account in the analysis.

Table 4.17: Comparison of MSE of Impulse Responses to Monetary Policy Shock for $T = 500$ and Propagation Horizon up to $h = 5$

Relevance (β_1,ρ_1)	Endogeneity (β_2,ρ_2)														
	(0.0,0.0)			(0.05,0.03)			(0.20,0.12)			(0.30,0.16)			(0.40,0.22)		
Model	θ_{11}	θ_{21}	θ_{31}	θ_{11}	θ_{21}	θ_{31}	θ_{11}	θ_{21}	θ_{31}	θ_{11}	θ_{21}	θ_{31}	θ_{11}	θ_{21}	θ_{31}
(0.0,0.0)															
Model (1)	1.00	1.00	1.00	.	.	.	.	.	.	.	.	.	.	.	.
Model (2)	1.00	0.98	1.00	.	.	.	.	.	.	.	.	.	.	.	.
Model (3)	93.11	61.52	162.31	.	.	.	.	.	.	.	.	.	.	.	.
(0.20,0.16)															
Model (1)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.	.	.	.	.	.
Model (2)	1.01	1.09	1.00	1.02	0.87	0.98	0.83	0.24	0.79	.	.	.	.	.	.
Model (3)	65.51	27.72	63.94	63.01	20.30	59.47	31.01	9.44	28.81	.	.	.	.	.	.
(0.40,0.30)															
Model (1)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Model (2)	1.06	1.35	1.04	1.06	1.05	1.02	0.94	0.24	0.85	0.48	0.10	0.47	0.29	0.06	0.28
Model (3)	61.76	24.04	58.79	60.77	14.15	56.94	42.17	2.66	36.16	16.14	1.80	14.38	6.68	1.73	6.04
(0.60,0.43)															
Model (1)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Model (2)	1.12	1.69	1.08	1.12	1.34	1.08	1.08	0.35	0.99	0.91	0.17	0.82	0.56	0.09	0.51
Model (3)	63.62	28.23	60.84	63.53	17.88	60.35	53.81	2.79	48.32	39.29	1.43	34.05	19.52	1.10	16.64

Notes: The table shows the cumulated MSE of fitted models (A)-(C) relative to model (C) for a propagation horizon up to $h = 5$ and sample size $T = 500$. Each entry is based on 500 replications of each simulation design. Dots (.) denote combinations of values for β_1 and β_2 that produce lower correlations between the instrument s_t and the target structural shock of interest (ε_t^r) than between the instrument s_t and the endogenous structural shock (ε_t^z). These cases are not taken into account in the analysis.

4.C Sensitivity Analysis of Baseline Model

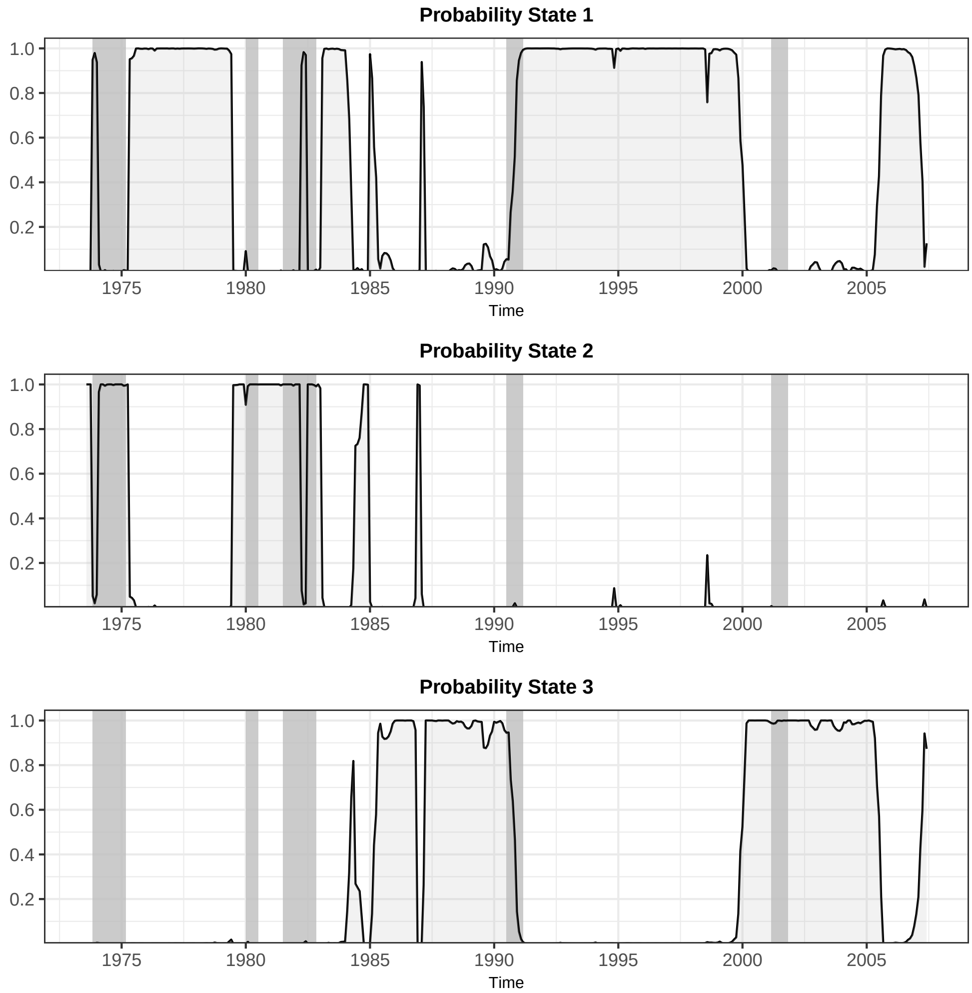


Figure 4.8: Smoothed state probabilities of Markov switching proxy-VAR(6) with $M = 3$ states.

Notes: The figure shows the estimated smoothed state probabilities for $m = 1$ in the upper panel, for $m = 2$ in the middle panel, and for $m = 3$ in the lower panel, where $t = 1, \ldots, T$, of the Markov switching proxy-VAR(6) model with $m = 3$ states. The dataset is $z_t = [\Delta ip_t, ff_t, ebp_t, rr_t]'$. The shaded vertical bars mark recession periods as defined by the NBER.

Table 4.18: Instrument Validity for MS(3)-VAR(6)

	Exogeneity	Relevance
LR statistic	0.02	44.84
p-value	0.99	0.00
Restrictions	2	1

Notes: The table shows the LR statistic, p-value and number of restrictions of the test for instrument exogeneity ($H_0: \ \beta_2 = \cdots = \beta_K = 0, \ \ H_1 : \beta$ unrestricted) and for instrument relevance ($H_0 : \beta_1 = 0, \ \ H_1 : \beta_1 \neq 0$) for a Markov switching proxy-VAR(6) model with $M = 3$ states. The instrument is the narrative-based measure of monetary surprises of Romer and Romer (2004).

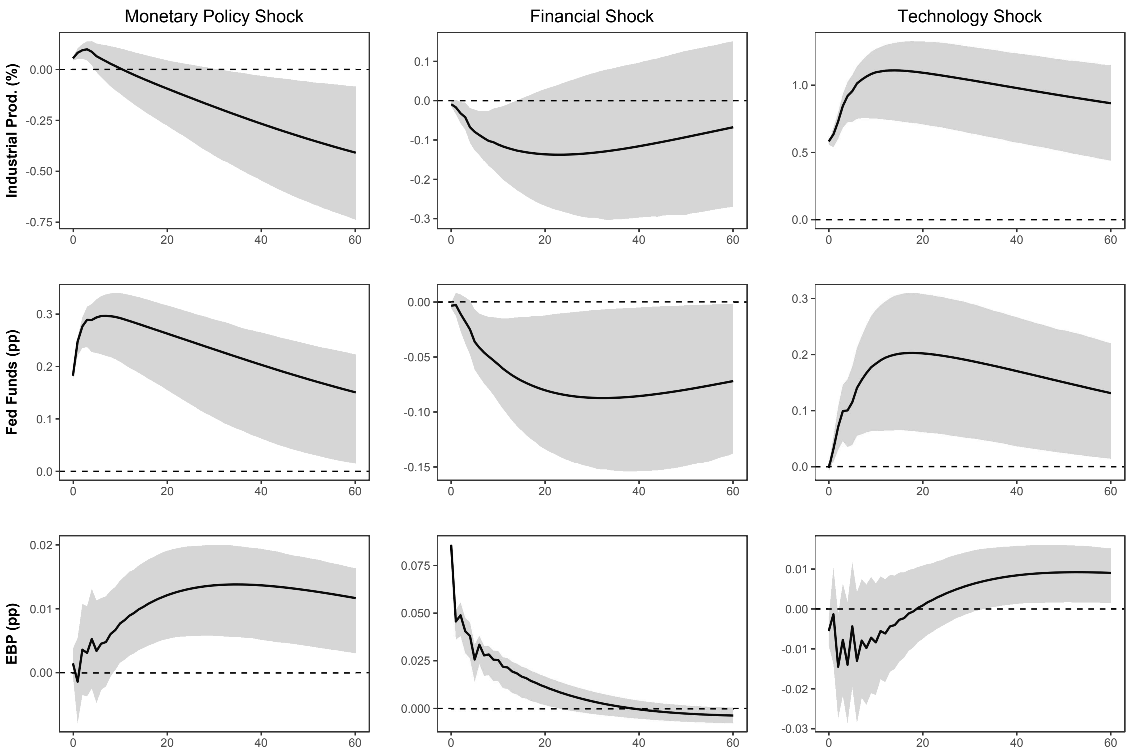


Figure 4.9: Impulse responses for Markov switching proxy-VAR(6) with $M = 3$ states.

Notes: The figure shows the impulse responses to one standard deviation shocks in state $m = 1$ of the Markov switching proxy-VAR(6) model with $M = 3$ states. The dataset is $z_t = [\Delta ip_t, ff_t, ebp_t, rr_t]'$. The sample is 1973M1-2007M6 and the instrument for monetary policy shocks is the narrative-based measure of Romer and Romer (2004). The shaded bands denote 95 percent pointwise confidence intervals based on 5,000 bootstrap replications.

Table 4.19: Validity of Cleaned Instrument in Heteroskedastic Proxy-VAR

Instrument	Exogeneity p-value	Relevance p-value
Narrative-based	0.791	0.000

Notes: The table shows the p-values of LR-tests for the exogeneity and relevance of Romer and Romer (2004) instrument s_t based on an uncrestricted (that is, imposing no restrictions on the constant and autoregressive parts of the model) heteroskedastic proxy-VAR(6) model with $M = 2$ states and $z_t = [\Delta ip_t, ff_t, ebp_t, s_t]'$. The sample period is 1973M1-2007M6.

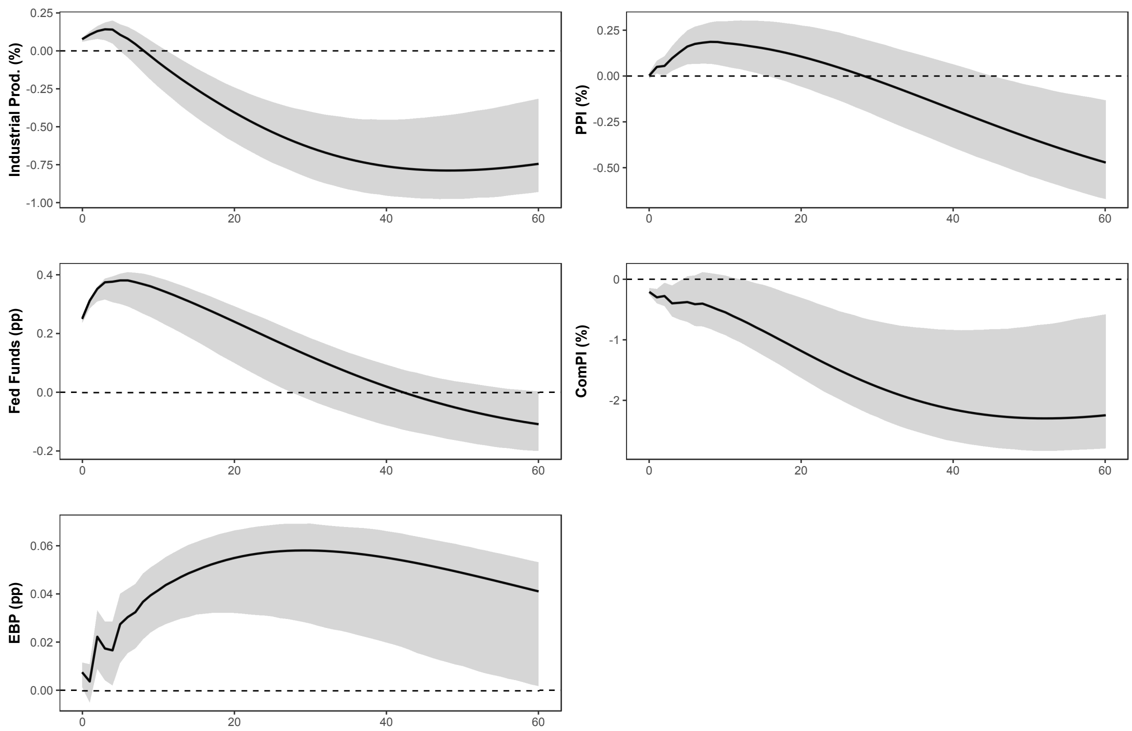


Figure 4.10: Sensitivity analysis of main results to adding price variables.

Notes: The figure shows the impulse responses to a monetary policy shock of 25 basis points in state $m = 1$ of a heteroskedastic proxy-VAR(6) with $M = 2$ states. The dataset is $\hat{z}_t = [x_t, ff_t, ebp_t, ppi_t, ComPI_t, rr_t]'$, where ppi_t refers to the log of producer prices (St. Louis FRED series $PPIACO$) and $ComPI_t$ the log of a commodity price index by the United Nations Conference on Trade and Development (UNCTAD Stat). The sample is 1973M1-2007M6 and the instrument for monetary policy shocks is the narrative-based measure of Romer and Romer (2004). The shaded bands denote 95 percent pointwise confidence intervals based on 5,000 bootstrap replications.

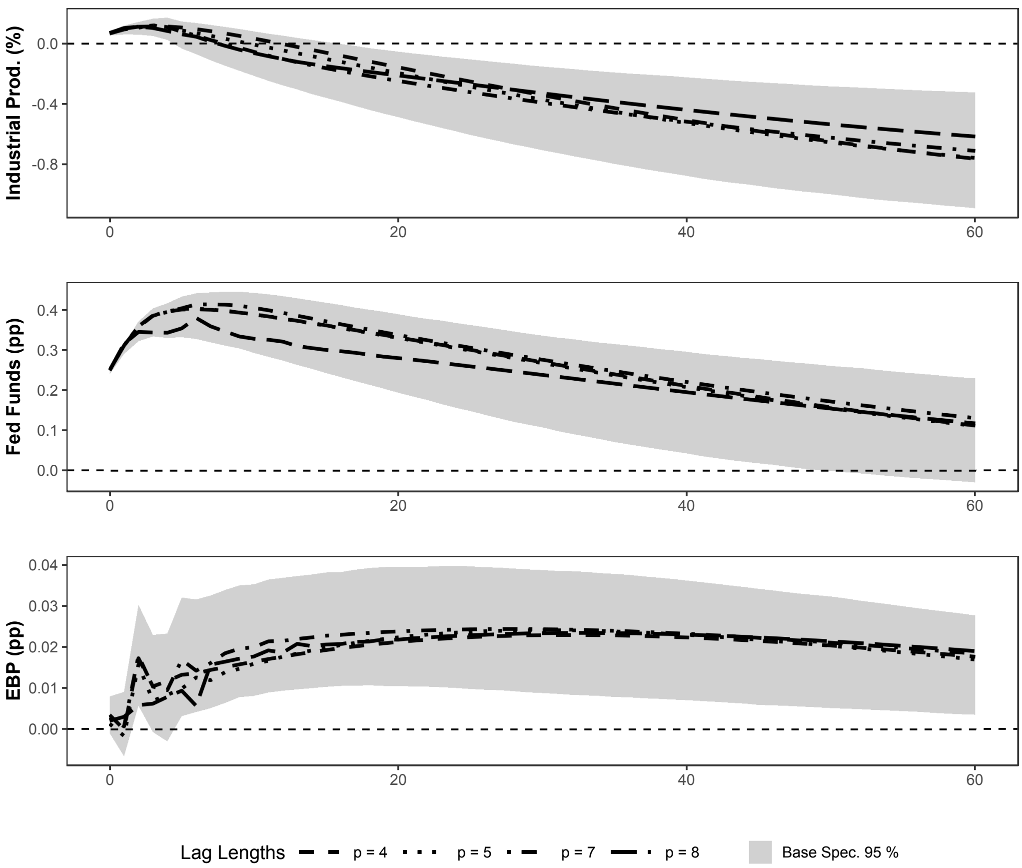


Figure 4.11: Sensitivity analysis of baseline model using different lag lengths.

Notes: The figure shows the impulse responses to a monetary policy shock of 25 basis points in state $m = 1$ of the heteroskedastic proxy-VAR(p) with $M = 2$ states with $p = 4, 5, 7, 8$. The dataset is $z_t = [\Delta ip_t, ff_t, ebp_t, rr_t]'$. The sample is 1973M1-2007M6 and the instrument for monetary policy shocks is the narrative-based measure of Romer and Romer (2004). The shaded bands denote 95 percent pointwise confidence intervals based on 5,000 bootstrap replications of the baseline heteroskedastic proxy-VAR(6) model with $M = 2$ states.

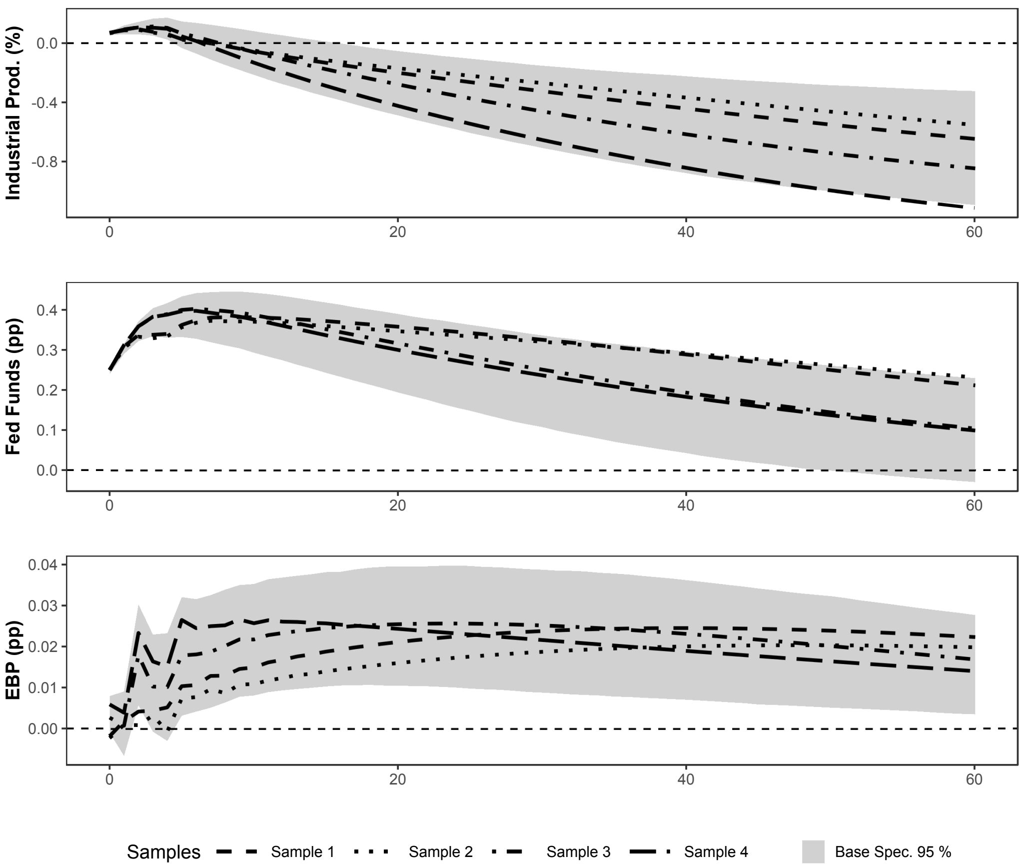


Figure 4.12: Sensitivity analysis of baseline model using different sample periods.

Notes: The figure shows the impulse responses to a monetary policy shock of 25 basis points in state $m = 1$ of the heteroskedastic proxy-VAR(6) model with $M = 2$ states. The dataset is $z_t = [\Delta ip_t, ff_t, ebp_t, rr_t]'$ for different sample specifications. Sample 1 refers to 1975M1-2007M6, Sample 2 refers to 1978M1-2007M6, Sample 3 refers to 1973M1-2005M6, Sample 4 refers to 1973M1-2002M6. The instrument for monetary policy shocks is the narrative-based measure of Romer and Romer (2004). The shaded bands denote 95 percent pointwise confidence intervals based on 5,000 bootstrap replications of the baseline heteroskedastic proxy-VAR(6) model with $M = 2$ states for the sample 1973M1-2007M6.

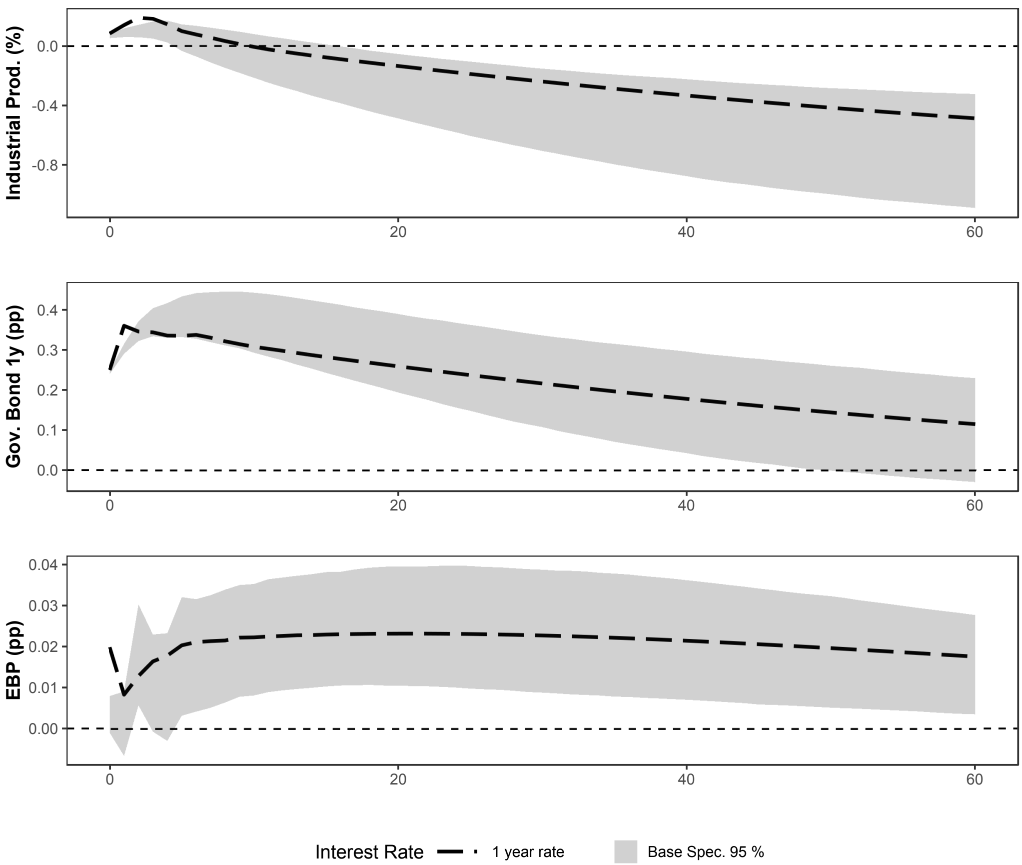


Figure 4.13: Sensitivity analysis of baseline model using alternative indicator for monetary policy.

Notes: The figure shows the impulse responses to a monetary policy shock of 25 basis points in state $m = 1$ of the heteroskedastic proxy-VAR(6) model with $M = 2$ states. The dataset is $\hat{z}_t = [\Delta ip_t, GovBond1_t, ebp_t, rr_t]'$, where $GovBond1_t$ refers to the US government bond yield with one year maturity. The sample is 1973M1-2007M6 and the instrument for monetary policy shocks is the narrative-based measure of Romer and Romer (2004). The shaded bands denote 95 percent pointwise confidence intervals based on 5,000 bootstrap replications of the baseline heteroskedastic proxy-VAR(6) model with $M = 2$ states and $z_t = [\Delta ip_t, ff_t, ebp_t, rr_t]'$.

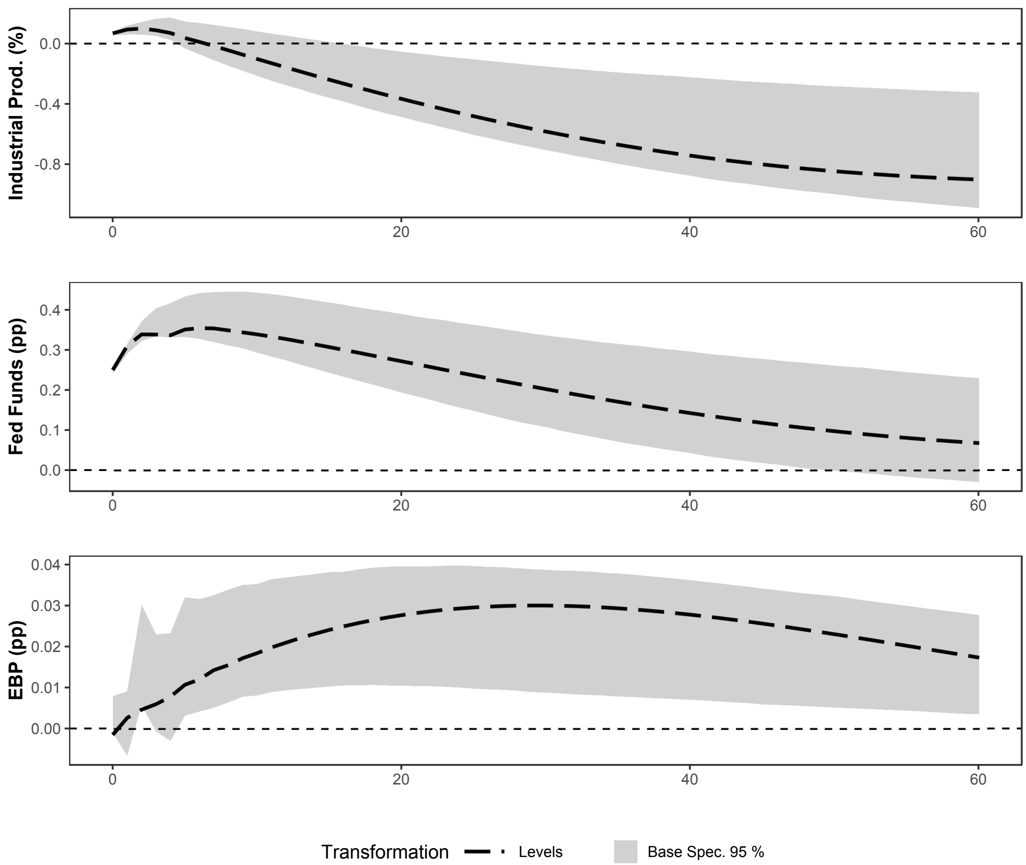


Figure 4.14: Sensitivity analysis of baseline model using level of industrial production.

Notes: The figure shows the impulse responses to a monetary policy shock of 25 basis points in state $m = 1$ of the heteroskedastic proxy-VAR(6) model with $M = 2$ states . The dataset is $\tilde{z}_t = [x_t, ff_t, ebp_t, rr_t]'$, that is, the industrial production x_t enters the model in (logged) levels. The sample is 1973M1-2007M6 and the instrument for monetary policy shocks is the narrative-based measure of Romer and Romer (2004). The shaded bands denote 95 percent pointwise confidence intervals based on 5,000 bootstrap replications of the baseline heteroskedastic proxy-VAR(6) model with $M = 2$ states and $z_t = [\Delta ip_t, ff_t, ebp_t, rr_t]'$.

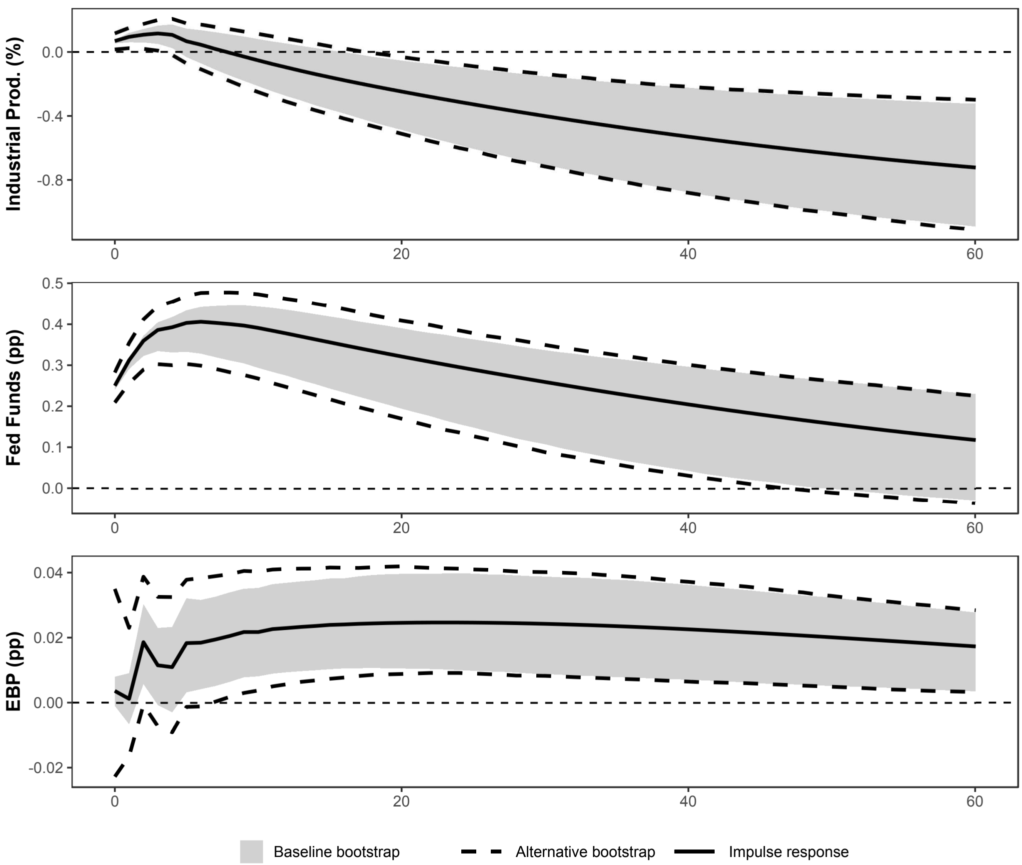


Figure 4.15: Sensitivity analysis of baseline model using an alternative bootstrap method.

Notes: The figure shows the impulse responses to a monetary policy shock of 25 basis points in state $m = 1$ of the baseline heteroskedastic proxy-VAR(6) model with $M = 2$ states. The dataset is $z_t = [\Delta ip_t, ff_t, ebp_t, rr_t]'$. The shaded bands denote 95 percent pointwise confidence intervals based on 5,000 bootstrap replications of the baseline heteroskedastic proxy-VAR(6) model with $M = 2$ states. Under the baseline bootstrap, the bootstrapped residuals $e_t^\star$ are computed from the estimated reduced form residuals $\widehat{e}_t$ as $e_t^\star = \varphi_t \widehat{e}_t$, where φ_t are independent random draws from a Rademacher distribution. The dashed lines denote the 95 percent pointwise confidence intervals based on 5,000 bootstrap replications of an alternative recursive design wild bootstrap. The bootstrapped residuals are $\tilde{e}_t^\star = \tilde{\varphi}_t \widehat{e}_t$, where $\widehat{e}_t$ are the model's estimated reduced form residuals and $\tilde{\varphi} \sim N(0, 1)$ is independently standard normal distributed.

Bibliography

Adrian, T., Colla, P. and Song Shin, H. (2013). Which financial frictions? Parsing the evidence from the financial crisis of 2007 to 2009, *NBER Macroeconomics Annual* **27**(1): 159–214.

Akaike, H. (1974). A new look at the statistical model identification, *IEEE Transactions on Automatic Control* **19**(6): 716–723.

Alfaro, I., Bloom, N. and Lin, X. (2018). The finance uncertainty multiplier, *Working Paper 24571*, National Bureau of Economic Research.

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

Angelini, G., Bacchiocchi, E., Caggiano, G. and Fanelli, L. (2017). Uncertainty across volatility regimes, *Research Discussion Papers 35/2017*, Bank of Finland.

Angelini, G. and Fanelli, L. (2018). Exogenous uncertainty and the identification of structural vector autoregressions with external instruments, *MPRA Paper 93864*, University Library of Munich, Germany.

Antolín-Díaz, J. and Rubio-Ramírez, J. F. (2018). Narrative sign restrictions for SVARs, *American Economic Review* **108**(10): 2802–2829.

Arellano, C., Bai, Y. and Kehoe, P. J. (forthcoming). Financial frictions and fluctuations in volatility, *Journal of the Political Economy* **forthcoming**.

Bacchiocchi, E. and Fanelli, L. (2015). Identification in structural vector autoregressive models with structural changes, with an application to US monetary policy, *Oxford Bulletin of Economics and Statistics* **77**(6): 761–779.

Bachmann, R., Elstner, S. and Sims, E. R. (2013). Uncertainty and economic activity: Evidence from business survey data, *American Economic Journal: Macroeconomics* **5**(2): 217–49.

Baker, S. R., Bloom, N. and Davis, S. J. (2016). Measuring economic policy uncertainty, *The Quarterly Journal of Economics* **131**(4): 1593–1636.

Barakchian, S. M. and Crowe, C. (2013). Monetary policy matters: Evidence from new shocks data, *Journal of Monetary Economics* **60**(8): 950–966.

Bates, J. M. and Granger, C. W. J. (1969). The combination of forecasts, *Operations Research Quarterly* **20**: 451–468.

Bauwens, L., Laurent, S. and Rombouts, J. V. K. (2006). Multivariate GARCH models: a survey, *Journal of Applied Econometrics* **21**(1): 79–109.

Becker, R. and Clements, A. E. (2008). Are combination forecasts of S&P 500 volatility statistically superior?, *International Journal of Forecasting* **24**: 122–133.

Becker, R., Clements, A. E., Doolan, M. B. and Hurn, A. S. (2015). Selecting volatility forecasting models for portfolio allocation purposes, *International Journal of Forecasting* **31**: 849–861.

Belongia, M. T. and Ireland, P. N. (2015). Interest rates and money in the measurement of monetary policy, *Journal of Business & Economic Statistics* **33**(2): 255–269.

Berkowitz, J., Birgean, I. and Kilian, L. (2000). On the finite-sample accuracy of nonparametric resampling algorithms for economic time series, *Advances in Econometrics* **14**: 77–105.

Bernanke, B. S. and Blinder, A. S. (1992). The federal funds rate and the channels of monetary transmission, *American Economic Review* **82**(4): 901–921.

Bernanke, B. S. and Mihov, I. (1998). Measuring monetary policy, *The Quarterly Journal of Economics* **113**(3): 869–902.

Bertsche, D. and Braun, R. (2017). Identification of structural vector autoregressions by stochastic volatility, Discussion Paper, University of Konstanz.

Blanchard, O. J. and Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances, *The American Economic Review* **79**(4): 655–673.

Blanchard, O. and Perotti, R. (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output, *The Quarterly Journal of Economics* **117**(4): 1329–1368.

Bloom, N. (2009). The impact of uncertainty shocks, *Econometrica* **77**(3): 623–685.

Bloom, N. (2014). Fluctuations in uncertainty, *Journal of Economic Perspectives* **28**(2): 153–176.

Boswijk, H. P. and van der Weide, R. (2011). Method of moments estimation of GO-GARCH models, *Journal of Econometrics* **163**: 118–126.

Bouakez, H. and Normandin, M. (2010). Fluctuations in the foreign exchange market: How important are monetary policy shocks?, *Journal of International Economics* **81**(1): 139–153.

Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco.

Bruder, S. (2018). Inference for structural impulse responses in SVAR-GARCH models, Working Paper 281, Department of Economics, University of Zurich.

Brüggemann, R., Jentsch, C. and Trenkler, C. (2016). Inference in VARs with conditional heteroskedasticity of unknown form, *Journal of Econometrics* **191**: 69–85.

Caldara, D., Fuentes-Albero, C., Gilchrist, S. and Zakrajšek, E. (2016). The macroeconomic impact of financial and uncertainty shocks, *European Economic Review* **88**(C): 185–207.

Caldara, D. and Herbst, E. (2016). Monetary policy, real activity, and credit spreads: Evidence from Bayesian proxy SVARs, *Finance and Economics Discussion Series 2016-049*, Board of Governors of the Federal Reserve System (US).

Caldara, D. and Herbst, E. (2019). Monetary policy, real activity, and credit spreads: Evidence from Bayesian proxy SVARs, *American Economic Journal: Macroeconomics* **11**(1): 157–192.

Canova, F. and De Nicoló, G. (2002). Monetary disturbances matter for business fluctuations in the G-7, *Journal of Monetary Economics* **49**(6): 1131–1159.

Caporin, M. and McAleer, M. (2011). Banking multivariate GARCH models by problem dimension: An empirical evaluation, *Working Paper No. 23/2011*, University of Canterbury, Christchurch.

Caporin, M. and McAleer, M. (2012). Model selection and testing of conditional and stochastic volatility models, *in* L. Bauwens, C. Hafner and S. Laurent (eds), *Handbook of Volatility Models and Their Applications*, Wiley, New York, chapter 8, pp. 199–224.

Carriero, A., Clark, T. E. and Marcellino, M. (2016). Common drifting volatility in large Bayesian VARs, *Journal of Business & Economic Statistics* **34**(3): 375–390.

Carriero, A., Clark, T. E. and Marcellino, M. (2018). Endogenous uncertainty, *Working Paper no. 18-05*, Federal Reserve Bank of Cleveland.

Carriero, A., Mumtaz, H., Theodoridis, K. and Theophilopoulou, A. (2015). The impact of uncertainty shocks under measurement error: A proxy SVAR approach, *Journal of Money, Credit and Banking* **47**(6): 1223–1238.

Castelnuovo, E., Lim, G. and Pellegrino, G. (2017). A short review of the recent literature on uncertainty, *Australian Economic Review* **50**(1): 68–78.

Cavaliere, G., Pedersen, R. S. and Rahbek, A. (2018). Fixed volatility bootstrap for a class of ARCH(q) models, *Journal of Time Series Analysis* **39**: 920–941.

Cesa-Bianchi, A., Thwaites, G. and Vicondoa, A. (2016). Monetary policy transmission in an open economy: New data and evidence from the United Kingdom, Mimemo, London School of Economics and Political Science, LSE Library.

Chen, W. and Netšunajev, A. (2018). Structural vector autoregression with time varying transition probabilities: Identifying uncertainty shocks via changes in volatility, *Bank of Estonia Working Papers wp2018-02*, Bank of Estonia.

Christiano, L. J., Eichenbaum, M. and Evans, C. L. (1999). Monetary policy shocks: What have we learned and to what end?, *Handbook of macroeconomics* **1**: 65–148.

Christiano, L. J., Motto, R. and Rostagno, M. (2014). Risk shocks, *American Economic Review* **104**(1): 27–65.

Coibion, O. (2012). Are the effects of monetary policy shocks big or small?, *American Economic Journal: Macroeconomics* **4**(2): 1–32.

Cooper, R. L. (1972). The predictive performance of quarterly econometric models of the United States, *in* B. G. Hickman (ed.), *Econometric Models of Cyclical Behavior*, NBER, pp. 813–947.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* **50**: 987–1007.

Faust, J. (1998). The robustness of identified VAR conclusions about money, *Carnegie-Rochester Conference Series on Public Policy* **49**: 207–244.

Francq, C. and Zakoïan, J.-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes, *Bernoulli* **10**: 605–637.

Fuller, W. A. (1976). *Introduction to Statistical Time Series*, John Wiley & Sons, New York.

Furlanetto, F., Ravazzolo, F. and Sarferaz, S. (2019). Identification of financial factors in economic fluctuations, *The Economic Journal* **129**: 311–337.

Gambetti, L. and Musso, A. (2017). Loan supply shocks and the business cycle, *Journal of Applied Econometrics* **32**(4): 764–782.

Gertler, M. and Karadi, P. (2015). Monetary policy surprises, credit costs, and economic activity, *American Economic Journal: Macroeconomics* **7**(1): 44–76.

Gertler, M. and Kiyotaki, N. (2015). Banking, liquidity, and bank runs in an infinite horizon economy, *American Economic Review* **105**(7): 2011–43.

Gilchrist, S., Sim, J. W. and Zakrajšek, E. (2014). Uncertainty, financial frictions, and investment dynamics, *Working Paper 20038*, National Bureau of Economic Research.

Gilchrist, S. and Zakrajšek, E. (2012). Credit spreads and business cycle fluctuations, *American Economic Review* **102**(4): 1692–1720.

Gonçalves, S. and Kilian, L. (2004). Bootstrapping autoregressions with conditional heteroskedasticity of unknown form, *Journal of Econometrics* **123**: 89–120.

Gonçalves, S. and Kilian, L. (2007). Asymptotic and bootstrap inference for AR(∞) processes with conditional heteroskedasticity, *Econometric Reviews* **26**: 609–641.

Gouriéroux, C. and Monfort, A. (2014). Revisiting Identification and estimation in Structural VARMA Models, *Working Papers 2014-30*, Center for Research in Economics and Statistics.

Granger, C. W. J. and Newbold, P. (1974). Spurious regressions in eonometrics, *Journal of Econometrics* **2**: 111–120.

Granger, C. W. J. and Newbold, P. (1975). Economic forecasting: The atheist's viewpoint, *in* G. A. Renton (ed.), *Modelling the Economy*, Heinemann Educational Books, London, pp. 131–148.

Granger, C. W. J. and Newbold, P. (1977). *Forecasting Economic Time Series*, Academic Press, New York.

Hachula, M., Piffer, M. and Rieth, M. (forthcoming). Unconventional monetary policy, fiscal side effects and euro area (im)balances, *Journal of the European Economic Association* **forthcoming**.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*, Springer, New York.

Hannan, E. J. and Quinn, B. G. (1979). The determination of the order of an autoregression, *Journal of the Royal Statistical Society. Series B* **41**(2): 190–195.

Hansen, B. E. (1992). The likelihood ratio test under nonstandard conditions: Testing the Markov switching model of GNP, *Journal of Applied Econometrics* **7**(S1): 61–82.

Hansen, P. R. and Lunde, A. (2005). A forecast comparison of volatility models: Does anything beat a GARCH(1,1)?, *Journal of Applied Econometrics* **20**: 873–889.

Hanson, M. S. (2006). Varying monetary policy regimes: A vector autoregressive investigation, *Journal of Economics and Business* **58**(5-6): 407–427.

He, C. and Teräsvirta, T. (1999). Fourth moment structure of the GARCH(p, q) process, *Econometric Theory* **15**: 824–846.

Herwartz, H. and Lütkepohl, H. (2014). Structural vector autoregressions with Markov switching: Combining conventional with statistical identification of shocks, *Journal of Econometrics* **183**(1): 104–116.

Hidalgo, J. and Zaffaroni, P. (2007). A goodness-of-fit test for ARCH(∞) models, *Journal of Econometrics* **141**: 835–875.

Hristov, N., Hülsewig, O. and Wollmershäuser, T. (2012). Loan supply shocks during the financial crisis: Evidence for the euro area, *Journal of International Money and Finance* **31**(3): 569–592.

Jeong, M. (2017). Residual-based GARCH bootstrap and second order asymptotic refinements, *Econometric Theory* **33**: 779–790.

Jo, S. and Sekkel, R. (2019). Macroeconomic uncertainty through the lens of professional forecasters, *Journal of Business & Economic Statistics* **37**(3): 436–446.

Jurado, K., Ludvigson, S. C. and Ng, S. (2015). Measuring uncertainty, *American Economic Review* **105**(3): 1177–1216.

Justiniano, A. and Primiceri, G. E. (2008). The time-varying volatility of macroeconomic fluctuations, *American Economic Review* **98**(3): 604–41.

Kilian, L. (1998). Small-sample confidence intervals for impulse response functions, *Review of Economics and Statistics* **80**: 218–230.

Kilian, L. (1999). Finite-sample properties of percentile and percentile-t bootstrap confidence intervals for impulse responses, *Review of Economics and Statistics* **81**: 652–660.

Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market, *American Economic Review* **99**: 1053–1069.

Kilian, L. and Lütkepohl, H. (2017). *Structural Vector Autoregressive Analysis*, Cambridge University Press, Cambridge.

Kilian, L. and Murphy, D. (2012). Why agnostic sign restrictions are not enough: Understanding the dynamics of oil market VAR models, *Journal of the European Economic Association* **10**: 1166–1188.

Kim, D. (2014). Maximum likelihood estimation for vector autoregressions with multivariate stochastic volatility, *Economics Letters* **123**: 282–286.

King, R., Plosser, C., Stock, J. and Watson, M. (1991). Stochastic trends and economic fluctuations, *American Economic Review* **81**(4): 819–40.

Kreiss, J.-P. (1997). Asymptotical properties of residual bootstrap for autoregressions, *Technical report*, TU Braunschweig.

Krolzig, H.-M. (1997). *Markov-Switching Vector Autoregressions. Modelling, Statistical Inference, and Application to Business Cycle Analysis*, Vol. 454 of *Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag, Berlin and Heidelberg.

Lanne, M. and Lütkepohl, H. (2010). Structural vector autoregressions with nonnormal residuals, *Journal of Business & Economic Statistics* **28**(1): 159–168.

Lanne, M. and Lütkepohl, H. (2008). Identifying monetary policy shocks via changes in volatility, *Journal of Money, Credit and Banking* **40**(6): 1131–1149.

Lanne, M., Lütkepohl, H. and Maciejowska, K. (2010). Structural vector autoregressions with Markov switching, *Journal of Economic Dynamics and Control* **34**(2): 121–131.

Lanne, M., Meitz, M. and Saikkonen, P. (2017). Identification and estimation of non-gaussian structural vector autoregressions, *Journal of Econometrics* **196**(2): 288–304.

Lanne, M. and Saikkonen, P. (2007). A multivariate generalized orthogonal factor GARCH model, *Journal of Business & Economic Statistics* **25**: 61–75.

Laurent, S., Rombouts, J. V. K. and Violante, F. (2012). On the forecasting accuracy of multivariate GARCH models, *Journal of Applied Econometrics* **27**: 934–955.

Leeper, E. M. and Zha, T. (2003). Modest policy interventions, *Journal of Monetary Economics* **50**(8): 1673–1700.

Ludvigson, S. C., Ma, S. and Ng, S. (2019). Uncertainty and business cycles: Exogenous impulse or endogenous response?, *NBER Working Papers 21803*, National Bureau of Economic Research, Inc.

Lunsford, K. (2015). Identifying structural vars with a proxy variable and a test for a weak proxy, *Technical report*, Federal Reserve Bank of Cleveland.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.

Lütkepohl, H. (2013). Identifying structural vector autoregressions via changes in volatility, *Advances in Econometrics* **32**: 169–203.

Lütkepohl, H. and Milunovich, G. (2016). Testing for identification in SVAR-GARCH models, *Journal of Economic Dynamics and Control* **73**: 241–258.

Lütkepohl, H. and Netšunajev, A. (2014). Disentangling demand and supply shocks in the crude oil market: How to check sign restrictions in structural VARs, *Journal of Applied Econometrics* **29**(3): 479–496.

Lütkepohl, H. and Netšunajev, A. (2017a). Structural vector autoregressions with heteroskedasticity: A review of different volatility models, *Econometrics and Statistics* **1**: 2–18.

Lütkepohl, H. and Netšunajev, A. (2017b). Structural vector autoregressions with smooth transition in variances, *Journal of Economic Dynamics and Control* **84**(C): 43–57.

Lütkepohl, H. and Schlaak, T. (2018). Choosing between different time-varying volatility models for structural vector autoregressive analysis, *Oxford Bulletin of Economics and Statistics* **80**(4): 715–735.

Lütkepohl, H. and Schlaak, T. (2019). Bootstrapping impulse responses of structural vector autoregressive models identified through GARCH, *Journal of Economic Dynamics and Control* **101**: 41–61.

Lütkepohl, H., Staszewska-Bystrova, A. and Winker, P. (2015a). Comparison of methods for constructing joint confidence bands for impulse response functions, *International Journal for Forecasting* **31**: 782–798.

Lütkepohl, H., Staszewska-Bystrova, A. and Winker, P. (2015b). Confidence bands for impulse responses: Bonferroni versus Wald, *Oxford Bulletin of Economics and Statistics* **77**: 800–821.

Lütkepohl, H. and Velinov, A. (2016). Structural vector autoregressions: Checking identifying long-run restrictions via heteroskedasticity, *Journal of Economic Surveys* **30**: 377–392.

Meeks, R. (2012). Do credit market shocks drive output fluctuations? Evidence from corporate spreads and defaults, *Journal of Economic Dynamics and Control* **36**(4): 568–584.

Mertens, K. and Ravn, M. O. (2013). The dynamic effects of personal and corporate income tax changes in the united states, *American Economic Review* **103**(4): 1212–1247.

Milunovich, G. and Yang, M. (2013). On identifying structural VAR models via ARCH effects, *Journal of Time Series Econometrics* **5**: 117–131.

Miranda-Agrippino, S. and Ricco, G. (2017). The transmission of monetary policy shocks, Mimemo, London School of Economics and Political Science, LSE Library.

Nakamura, E. and Steinsson, J. (2018). High-frequency identification of monetary non-neutrality: The information effect, *The Quarterly Journal of Economics* **133**(3): 1283–1330.

Nelson, C. R. (1972). The prediction performance of the FRB-MIT-PENN model of the U.S. economy, *American Economic Review* **62**: 902–917.

Netšunajev, A. and Glass, K. (2017). Uncertainty and employment dynamics in the euro area and the US, *Journal of Macroeconomics* **51**: 48–62.

Netšunajev, A. (2013). Reaction to technology shocks in Markov-switching structural VARs: Identification via heteroskedasticity, *Journal of Macroeconomics* **36**: 51–62.

Newbold, P. and Granger, C. W. J. (1974). Experience with forecasting univariate time series and combination of forecasts, *Journal of the Royal Statistical Society* **A137**: 131–146.

Normandin, M. and Phaneuf, L. (2004). Monetary policy shocks: Testing identification conditions under time-varying conditional volatility, *Journal of Monetary Economics* **51**(6): 1217–1243.

Olea, M., Stock, J. and Watson, M. (2018). Inference in structural vector autoregressions identified with an external instrument, *Technical report*, mimemo, Columbia University.

Owyang, M. T. and Ramey, G. (2004). Regime switching and monetary policy measurement, *Journal of Monetary Economics* **51**(8): 1577–1597.

Peersman, G. (2012). Bank lending shocks and the euro area business cycle, *Working papers of faculty of economics and business administration, ghent university, belgium*, Ghent University, Faculty of Economics and Business Administration.

Piffer, M. and Podstawski, M. (2018). Identifying uncertainty shocks using the price of gold, *The Economic Journal* **128**(616): 3266–3284.

Podstawski, M. and Velinov, A. (2018). The state dependent impact of bank exposure on sovereign risk, *Journal of Banking & Finance* **88**(C): 63–75.

Politis, D. N. and Romano, J. P. (1994). The stationary bootstrap, *Journal of the American Statistical Association* **89**: 1303–1313.

Popescu, A. and Smets, F. R. (2010). Uncertainty, risk-taking, and the business cycle in Germany, *CESifo Economic Studies* **56**(4): 596–626.

Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy, *The Review of Economic Studies* **72**(3): 821–852.

Psaradakis, Z. and Spagnolo, N. (2006). Joint determination of the state dimension and autoregressive order for models with Markov regime switching, *Journal of Time Series Analysis* **27**(5): 753–766.

Ramey, V. A. (2016). Macroeconomic shocks and their propagation, *Handbook of Macroeconomics*, Vol. 2, Elsevier, pp. 71–162.

Rigobon, R. (2003). Identification through heteroskedasticity, *Review of Economics and Statistics* **85**(4): 777–792.

Rigobon, R. and Sack, B. (2003). Measuring the reaction of monetary policy to the stock market, *The Quarterly Journal of Economics* **118**(2): 639–669.

Rigobon, R. and Sack, B. (2004). The impact of monetary policy on asset prices, *Journal of Monetary Economics* **51**(8): 1553–1575.

Rogers, J. H., Scotti, C. and Wright, J. H. (2018). Unconventional monetary policy and international risk premia, *Journal of Money, Credit and Banking* **50**(8): 1827–1850.

Romer, C. D. and Romer, D. H. (2004). A new measure of monetary shocks: Derivation and implications, *American Economic Review* **94**(4): 1055–1084.

Rossi, B. and Sekhposyan, T. (2015). Macroeconomic uncertainty indices based on nowcast and forecast error distributions, *American Economic Review* **105**(5): 650–655.

Schwarz, G. (1978). Estimating the dimension of a model, *Annals of Statistics* **6**(2): 461–464.

Sentana, E. and Fiorentini, G. (2001). Identification, estimation and testing of conditionally heteroskedastic factor models, *Journal of Econometrics* **102**(2): 143–164.

Silvennoinen, A. and Teräsvirta, T. (2009). Multivariate GARCH models, *in* T. G. Andersen, R. A. Davis, K.-P. Kreiß and T. Mikosch (eds), *Handbook of Financial Time Series*, Springer-Verlag, Berlin, pp. 201–229.

Sims, C. A. (1980). Macroeconomics and reality, *Econometrica: Journal of the Econometric Society* **48**(1): 1–48.

Sims, C. A. and Zha, T. (2006). Were there regime switches in US monetary policy?, *American Economic Review* **96**(1): 54–81.

Stock, J. H. and Watson, M. W. (2002). Has the business cycle changed and why?, *NBER macroeconomics annual* **17**: 159–218.

Stock, J. H. and Watson, M. W. (2012). Disentangling the channels of the 2007-09 recession, *Brookings Papers on Economic Activity* pp. 120–157.

Stock, J. H. and Watson, M. W. (2018). Identification and estimation of dynamic causal effects in macroeconomics using external instruments, *The Economic Journal* **128**(610): 917–948.

Stock, J. H., Wright, J. H. and Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments, *Journal of Business & Economic Statistics* **20**(4): 518–529.

Uhlig, H. (1997). Bayesian vector autoregressions with stochastic volatility, *Econometrica* **65**: 59–73.

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.

van der Weide, R. (2002). GO-GARCH: a multivariate generalized orthogonal GARCH model, *Journal of Applied Econometrics* **17**(5): 549–564.

van Dijk, D., Teräsvirta, T. and Franses, P. H. (2002). Smooth transition autoregressive modesl - a survey of recent developments, *Econometric Reviews* **21**(1): 1–47.

Velinov, A. and Chen, W. (2015). Do stock prices reflect their fundamentals? New evidence in the aftermath of the financial crisis, *Journal of Economics and Business* **80**: 1–20.

Wieland, J. F. and Yang, M.-J. (2016). Financial dampening, *NBER Working Papers 22141*, National Bureau of Economic Research, Inc.

Wright, J. H. (2012). What does monetary policy do to long-term interest rates at the zero lower bound?, *The Economic Journal* **122**(564): 447–466.

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorgelegte Dissertation auf Grundlage der angegebenen Quellen und Hilfsmittel selbstständig verfasst habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Schriften entnommen sind, sind als solche kenntlich gemacht. Die vorgelegte Dissertation hat weder in der gleichen noch einer anderen Fassung bzw. Überarbeitung einer anderen Fakultät, einem Prüfungsausschuss oder einem Fachvertreter an einer anderen Hochschule zum Promotionsverfahren vorgelegen.

Thore Schlaak
Berlin, den 11. September 2019

Liste verwendeter Hilfsmittel

- RStudio 1.2.1335 basierend auf R 3.3.3 - R 3.5.1
- Microsoft Excel
- LaTeX
- Siehe auch Literatur- und Quellenangaben