Structural Vector Autoregressions and Information in Moments Beyond the Variance

Sascha Alexander Keweloh

Veröffentlichung als Dissertation in der Fakultät Wirtschaftswissenschaften der Technischen Universität Dortmund

Dortmund, August 2022

ACKNOWLEDGMENTS

This thesis was written as a doctoral student at the Chair of Applied Economics at the TU Dortmund University. First, I would like to thank my supervisor Prof. Dr. Ludger Linnemann for his support, guidance, and for the freedom I was granted in my research. I am grateful for the discussions, comments, and critique which helped me to improve my work. I would also like to thank Prof. Dr. Carsten Jentsch for co-refereeing this dissertation and Prof. Dr. Christoph Hanck for agreeing to be the third member of the doctoral committee. Furthermore, I would like to thank my co-authors, Andre Seepe and Stephan Hetzenecker, for the productive collaboration. I also thank the associate editor of the Journal of Business & Economic Statistics and the anonymous reviewers for their comments and recommendations. Moreover, I am thankful for the comments and recommendations at several conferences and seminars, including the 12th and 14th RGS Doctoral Conference in Economics, the 2019 IAAE Annual Conference, the 2019 European Winter Meeting of the Econometric Society, the UAR Econometrics Seminar, and the Helsinki Graduate School of Economics Econometrics Seminar.

I gratefully acknowledge the computing time provided on the Linux HPC cluster at Technical University Dortmund (LiDO3), partially funded in the course of the Large-Scale Equipment Initiative by the German Research Foundation (DFG) as project 271512359. Moreover, financial support of the RGS Econ and the German Science Foundation, DFG - SFB 823, is gratefully acknowledged.

Last but not least, I want to express my gratitude to my wife Evelyn for her encouragement and support throughout the last years.

Contents

Introduction **1**

1 A Generalized Method of Moments Estimator for Structural Vector Autoregressions Based on Higher Moments[1] **6**

1.1 Introduction 6

1.2 SVAR model 9

1.3 Illustration: Identification and higher moments 10

1.4 Notation 12

1.5 SVAR-GMM estimator 14

1.6 Fast SVAR-GMM estimator 19

1.7 Finite sample properties 24

1.8 Economic activity, oil, and stock prices 28

1.9 Conclusion 29

1.10 Appendix 30

1.10.1 Appendix - Proofs 30

1.10.2 Appendix - Notes on identification 37

1.10.3 Appendix - Notes on PML 39

1.10.4 Appendix - Finite sample performance 40

1.10.5 Appendix - Application: Data and robustness checks 41

2 A Feasible Approach to Incorporate Information in Higher Moments in Structural Vector Autoregressions[2] **44**

[1]This is an Accepted Manuscript of an article published by Taylor & Francis in the Journal of Business & Economic Statistics on March 2020, available online: http://wwww.tandfonline.com/10.1080/07350015.2020.1730858

[2]A slightly different version of the chapter appeared as Keweloh, S. A. (2021), *A feasible approach to incorporate information in higher moments in structural vectorautoregressions*, SFB 823 Discussion Paper series No. 22/2021, https://doi.org/10.17877/DE290R-22416

2.1 Introduction 44

2.2 SVAR models 47

2.3 SVAR GMM with higher-order moment conditions 49

2.4 Estimating S and G 52

2.5 Monte Carlo Simulation 58

2.6 Conclusion 66

2.7 Appendix 66

2.7.1 Appendix - Simulation with mixture of Gaussian distributions 66

2.7.2 Appendix - Simulation with t-distributed shocks 72

3 Monetary Policy and the Stock Market - A Partly Recursive SVAR Estimator[3] 76

3.1 Introduction 76

3.2 Monetary policy and the stock market 78

3.2.1 The unavailability of common identifying restrictions 78

3.2.2 Monetary policy and the stock market SVAR models 82

3.3 A partly recursive, partly non-Gaussian SVAR estimator 85

3.3.1 Derivation of the estimator 86

3.3.2 Finite sample performance 90

3.4 The interdependence of U.S. monetary policy and the stock market 94

3.5 Conclusion 100

3.6 Appendix 100

3.6.1 Appendix - White SVAR estimators with partly recursive constraints 100

3.6.2 Appendix - Finite sample performance 104

3.6.3 Appendix - Application 108

[3]joint work with Andre Seepe, a slightly different version of the chapter appeared as Keweloh, S.A. and A. Seepe (2020), *Monetary policy and the stock market - A partly recursive SVAR estimator*, SFB 823 Discussion Paper series No. 32/2020, https://doi.org/10.17877/DE290R-21722

4 **Block-Recursive Non-Gaussian Structural Vector Autoregressions: Identification, Efficiency, and Moment Selection**[4] **122**

4.1 Introduction 122

4.2 Overview SVAR 126

4.3 Imposing structure in a SVAR 128

4.3.1 Imposing structure on the interaction of shocks 128

4.3.2 Imposing structure on the stochastic properties of shocks 130

4.4 Estimation of a block-recursive SVAR 132

4.4.1 Identification 132

4.4.2 Overidentification and efficiency gains 136

4.4.3 Data-driven moment selection 140

4.5 Finite sample performance 143

4.5.1 Block-Recursive Structure 145

4.5.2 Recursive Structure 147

4.6 Disentangling speculative demand and supply shocks in the oil market 152

4.7 Conclusion 158

4.8 Appendix - Supplementary Notation and Proofs 160

4.8.1 Appendix - Asymptotic variance of the block-recursive SVAR GMM estimator 160

4.8.2 Appendix - Asymptotic variance of the (overidentified) block-recursive SVAR GMM estimator 160

4.8.3 Appendix - Identification in the block-recursive SVAR 161

4.8.4 Appendix - Redundant and relevant moment conditions in the recursive SVAR 165

[4]joint work with Stephan Hetzenecker, a slightly different version of the chapter appeared as Keweloh, S.A. and S. Hetzenecker (2021), *Efficiency gains in structural vector autoregressions by selecting informative higher-order moment conditions*, SFB 823 Discussion Paper series No. 26/2021, http://dx.doi.org/10.17877/DE290R-22447

4.8.5 Appendix - Asymptotic variance of the block-recursive SVAR pGMM estimator . . . 166
4.8.6 Appendix - Choice of maximum λ in the cross-validation . . . 171
4.9 Appendix - Details on Proof of Proposition 4.3 . . . 173
4.9.1 Appendix - Notation and preparations part 1 . . . 173
4.9.2 Appendix - Preparations part 2 . . . 188
4.9.3 Appendix - Final Lemma . . . 204
4.10 Appendix - Finite sample performance . . . 206
4.11 Appendix - Empirical illustration . . . 209
References **220**

List of Figures

1.1 Economic activity, oil, and stock prices - Impulse responses. 30

1.2 Economic activity, oil, and stock prices - Impulse responses (robustness checks). . 43

2.1 Finite sample performance - GMM loss at B_0 for different weighting schemes. . . 60

2.2 Finite sample performance - Bias of the estimated elements. 62

2.3 Finite sample performance - Interquartile range of the estimated elements. 64

2.4 Finite sample performance - Rejection rates. 65

2.5 Finite sample performance - Bias of the estimated elements (one-step estimators). 70

2.6 Finite sample performance - Interquartile range of the estimated elements (one-step estimators). 71

2.7 Finite sample performance - Rejection rates (one-step estimators). 72

2.8 Finite sample performance - Bias of the estimated elements (t-distribution). . . . 73

2.9 Finite sample performance - Interquartile range of the estimated elements (t-distribution). 74

2.10 Finite sample performance - Rejection rates (t-distribution). 75

3.1 Simulated response of real stock prices to a one-time exogenous interest rate increase of about one percentage point as implied by the exogenous and endogenous growth model. 81

3.2 Monetary policy and the stock market - Impulse responses. 97

3.3 Monetary policy and the stock market - Impulse responses (unrestricted SVAR). . 110

3.4 Monetary policy and the stock market - Impulse responses (PML estimator). . . . 113

3.5 Monetary policy and the stock market - Impulse responses (whitened fast SVAR GMM estimator). 114

3.6 Monetary policy and the stock market - Impulse responses (no trend). 115

3.7 Monetary policy and the stock market - Impulse responses (no trend and 1983Q1 to 2007Q4). 116

3.8 Monetary policy and the stock market - Impulse responses (without investments). 117

3.9 Monetary policy and the stock market - Impulse responses (controlled for commodity price inflation). 118

3.10 Monetary policy and the stock market - Impulse responses (two lags). 119

3.11 Monetary policy and the stock market - Impulse responses (zero lags for real stock returns). 121

4.1 Examples of different block-recursive SVAR models. 129

4.2 Finite sample performance - Share of selected moments by the pGMM estimator. . 151

4.3 Finite sample performance - Illustration of influence of λ on β for the pGMM estimator. 152

4.4 Oil market SVAR - Impulse responses. 155

4.5 Oil market SVAR - Real oil price evolution explained by the estimated structural shocks. 157

4.6 Finite sample performance - Relationship of chosen λ_{CV} and the number of selected moments across Monte Carlo runs in the recursive SVAR. 207

4.7 Finite sample performance - Average weight of moments across Monte Carlo runs. 208

4.8 Real oil price. 209

4.9 Oil market SVAR - Estimated structural shocks, averaged to annual frequency. . . 210

4.10 Oil market SVAR - Impulse responses (recursive SVAR). 211

4.11 Oil market SVAR - Impulse responses (unrestricted SVAR with GMM). 212

4.12 Oil market SVAR - Impulse responses (unrestricted SVAR with oGMM). 213

4.13 Oil market SVAR - Impulse responses (block-recursive SVAR with GMM). 214

4.14 Oil market SVAR - Impulse responses (block-recursive SVAR with oGMM). 215

4.15 Oil market SVAR - Impulse responses (using 24 instead of 12 lags). 216

4.16 Oil market SVAR - Impulse responses (using the percentage deviation of industrial production from a linear trend). 217

4.17 Oil market SVAR - Impulse responses (using the percentage deviation of industrial production from a linear trend and the log of the real oil price). 218

4.18 Oil market SVAR - Impulse responses (using the log of the real oil price). 219

List of Tables

1.1 Finite sample performance - Comparison of unrestricted non-Gaussian SVAR estimators. 25

1.2 Relative non-Gaussianity. 26

1.3 Finite sample performance - The impact of non-Gaussianity. 27

1.4 Economic activity, oil, and stock prices - Reduced form residuals. 29

1.5 Finite sample performance - Comparison of unrestricted non-Gaussian SVAR estimators (global optimization). 41

1.6 Economic activity, oil, and stock prices - Descriptive statistics. 42

2.1 Illustration of moment conditions. 50

2.2 Number of moments. 56

2.3 Finite sample performance - All estimated elements for $n = 2$. 67

2.4 Finite sample performance - All estimated elements for $n = 4$ (Part 1). 68

2.5 Finite sample performance - All estimated elements for $n = 4$ (Part 2). 69

3.1 Finite sample performance - Partly recursive SVAR. 92

3.2 Finite sample performance - Hypothesis tests in the partly recursive SVAR. 93

3.3 Monetary policy and the stock market - Non-Gaussianity of adjusted reduced form shocks. 96

3.4 Finite sample performance - Partly recursive SVAR (GMM white fast). 105

3.5 Finite sample performance - Partly recursive SVAR (PML). 106

3.6 Finite sample performance - Partly recursive SVAR (t-distribution). 107

3.7 Finite sample performance - Hypothesis tests in the partly recursive SVAR (t-distribution). 108

3.8 Monetary policy and the stock market - Descriptive statistics. 109

3.9 Monetary policy and the stock market - Correlation of reduced form and estimated structural shocks (unrestricted SVAR). 109

3.10 Monetary policy and the stock market - Skewness, Kurtosis and p-value of the Jarque-Bera test of estimated structural shocks (unrestricted SVAR). 110

3.11 Monetary policy and the stock market - Reduced form estimation of real stock returns. 120

4.1 Finite sample performance - Block-recursive SVAR. 146

4.2 Finite sample performance - Recursive SVAR and the pGMM estimator. 149

4.3 Oil market SVAR - Non-Gaussianity of the estimated structural shocks. 155

4.4 Finite sample performance - Recursive SVAR and the pGMM estimator including Post-LASSO. 206

4.5 Oil market SVAR - Descriptive statistics. 209

4.6 Oil market SVAR - Correlation of reduced form and estimated structural shocks. . 209

Introduction

The structural vector autoregressive (SVAR) model is an important tool in applied macroeconomics to analyze the effects of structural economic shocks on macroeconomic variables. Macroeconomic variables are typically driven by multiple structural shocks. The main difficulty in estimating the SVAR is to identify the structural shocks and their simultaneous impact on the variables of interest. Traditional identification approaches rely on restrictions on the responses to the structural shocks, see, e.g., Sims (1980) for short-run restrictions, Blanchard (1989) for long-run restrictions, and Uhlig (2005) for sign restrictions. More recently, several data-driven identification approaches have been proposed in the literature. These approaches do not require any restrictions on the responses to the structural shocks, instead, identification is based on stochastic properties of the shocks, see, e.g., Rigobon (2003), Lanne et al. (2010), Lütkepohl and Netšunajev (2017), and Lewis (2021) for time-varying volatility, and Gouriéroux et al. (2017), Lanne et al. (2017), Lanne and Luoto (2021), and Guay (2021) for non-Gaussian and independent shocks.

This thesis is concerned with the estimation of the simultaneous interaction in non-Gaussian SVAR models using generalized method of moments (GMM) estimators with higher-order moment conditions. The thesis contributes to the literature by providing global identification results using coskewness and cokurtosis moment conditions, by proposing modifications to the GMM estimation procedure to improve the small sample performance of the SVAR GMM estimator in the presence of higher-order moment conditions, and by developing a framework to combine traditional short-run restrictions with data-driven identification and estimation approaches.

In Chapter 1, I propose a SVAR GMM estimator which minimizes the dependencies of the shocks measured by covariance, coskewness, and cokurtosis conditions. The moment conditions are derived from the assumption of independent structural shocks. The identification result neither requires restrictions on the interaction nor does it require to specify the distribution of the structural shocks a priori. Instead, identification requires that at most one shocks has zero skewness and zero excess kurtosis. Moreover, identification requires valid moment conditions, which holds if the structural shocks are independent. To reduce the computational burden of the estimator, I propose a specific weighting scheme which allows to minimize the dependencies of the shocks by maximizing the non-Gaussianity of the shocks measured by the skewness and excess kurtosis of the shocks. A Monte Carlo experiment analyzes the finite sample performance of the proposed SVAR GMM estimator and compares it to alternative data-driven estimators in the SVAR literature. In the empirical illustration, I study the interaction of economic activity, oil prices, and stock prices. I present evidence that oil and stock prices interact simultaneously and cannot be ordered recursively.

In Chapter 2, I study the finite sample performance of the SVAR GMM estimator from the previous chapter and propose modifications to the GMM estimation procedure to improve the small sample performance. A Monte Carlo experiment shows that the bias, variance, and distortion of the test statistics of the SVAR GMM estimator (using standard two-step or continuous updating weighting schemes) increases with the number of variables included in the SVAR. I demonstrate that a large part of these performance issues is related to imprecise estimates of the asymptotically optimal weighting matrix and the asymptotic variance of the SVAR GMM estimator. Both values require to estimate the variance of the coskewness and cokurtosis conditions, which are co-moments up to order eight and difficult to estimate in small samples. Moreover, the number of these higher-order co-moments increases quickly with the number of variables included in the SVAR. I propose to use the assumption of independent shocks not only to derive moment conditions but also to derive alternative estimators for the asymptotically optimal weighting matrix and the asymptotic variance of the SVAR GMM estimator. With the assumption of independent shocks, the higher-order co-moments contained in the weighting matrix and asymptotic variance can be decomposed into a product of lower-order moments. I demonstrate in a Monte Carlo experiment that this modification to the estimation of the optimal weighting matrix and variance of the estimator greatly improve the performance of the SVAR GMM estimators in terms of bias, variance, and distortion of the test statistics, especially in large SVAR models.

While the modifications to the GMM estimation procedure lead to a better finite sample performance of the SVAR GMM estimators, the bias, variance, and distortion of the test statistics still increase with the number of variables included in the SVAR. Therefore, estimates based on unrestricted purely data-driven SVAR estimators tend to become less reliable when the number of variables included in the SVAR increases. The following two chapters aim to solve this problem by combining traditional restriction based approaches with data-driven estimation approaches based on information in moments beyond the variance.

In Chapter 3, which is joint work with Andre Seepe, we propose a combination of restriction based and data-driven estimation based on information in higher-order moments to analyze the interaction of U.S. monetary policy and the stock market. In particular, we propose an estimator which allows to order and identify some shocks recursively, while other shocks can remain unrestricted and are identified based on independence and non-Gaussianity. A Monte Carlo experiment illustrates how the performance of unrestricted purely data-driven SVAR estimators based on independence and non-Gaussianity deteriorates with an increasing model size. However, we show that ordering some shocks recursively can stop the performance decline of data-driven SVAR estimators in large SVAR models. In the application, we assume that that output, investment, and inflation behave sluggishly such that they cannot respond to stock market and monetary policy shocks within the same quarter. However, interest rates and stock returns remain unrestricted and can simultaneously respond to all shocks. Therefore, we impose a partly recursive order such that output, investment, and inflation are ordered recursively, while interest rates and stock returns remain unrestricted. We find that a positive stock market shock contemporaneously increases the nominal interest rate, while a contractionary monetary policy shock leads to lower real stock returns on impact. Furthermore, we present evidence that monetary policy is non-neutral with respect to real stock prices in the long-run.

In Chapter 4, which is joint work with Stephan Hetzenecker, we propose a rigorous framework to combine restrictions with higher-order moment conditions to identify and estimate SVAR models. The framework nests several SVAR estimators as special cases: i) the unrestricted SVAR GMM estimator proposed in Chapter 1, ii) the unrestricted SVAR GMM estimator proposed proposed by Lanne and Luoto (2021), iii) the partly recursive estimator proposed in Chapter 3, and iv) the frequently used recursive SVAR estimator based on the Cholesky decomposition.

Our framework allows the researcher to specify an arbitrary block-recursive order, such that that shocks in a given block can only influence variables in the same block or blocks order below. For a given block-recursive order we derive a set of identifying moment conditions based on the assumption of uncorrelated shocks across blocks and mean independent shocks within the blocks. We then derive overidentifying higher-order moment conditions from the assumption of independent shocks and show that these conditions can decrease the asymptotic variance of the estimator.

In particular, we derive conditions under which the frequently applied estimator based on the Cholesky decomposition is asymptotically inefficient. We discuss the trade-off between using only the identifying moment conditions which yields robust estimates since it allows for some dependencies of the structural shocks, and additionally using the overidentifying moment conditions which yields efficient estimates, but can lead to biased estimates if the overidentifying moment conditions contain invalid conditions. We propose to use a LASSO-type SVAR GMM estimator to select the relevant and valid and to unselect the invalid and redundant overidentifying moment conditions in a data-driven way. Our proposed LASSO-type SVAR GMM estimator is as robust as the SVAR GMM estimator using only the identifying moment conditions and as efficient as the SVAR GMM estimator which additionally uses all valid and relevant overidentifying moment conditions. A Monte Carlo experiment illustrates how block-recursive restrictions can mitigate the performance decline of non-Gaussian SVAR GMM estimators in the number of variables included in the SVAR. The simulation also shows that the LASSO-type SVAR GMM estimator successfully selects relevant moment conditions and increases the small sample performance compared to other block-recursive SVAR GMM estimators.

We use the block-recursive framework to analyze the impact of flow and speculative shocks in the oil market. In his seminal work, Kilian (2009) highlights that it is necessary to distinguish between oil supply and oil demand shocks rather than including solely an oil price shock in the oil market SVAR. We contribute to the oil market SVAR literature by explicitly distinguishing between speculative demand and speculative supply shocks. We use our proposed block-recursive framework and assume that oil production and economic activity is simultaneously only influenced by flow supply and flow demand shocks, while the real oil price and real stock returns can additionally contemporaneously be influenced by speculative supply and speculative demand shocks. We find that flow supply shocks lead to an increase of oil production and to a decrease of the real oil price. Flow demand shocks lead to an increase of economic activity and to an increase of the real oil price. Moreover, a speculative supply shock leads to an immediate decrease of the real oil price and to an increase of economic activity and oil production in the medium term. Speculative demand shocks lead to an immediate increase of the real oil price and to an increase of economic activity and oil production in the medium term. We demonstrate that an estimation based on a fully recursive order cannot not distinguish between speculative supply and demand shocks but rather contains a single speculative oil price shock and leads to counterintuitive results. Additionally, we show that an unrestricted estimation based only on non-Gaussianity leads to estimates with large confidence bands and no significant response of the real oil price to flow supply and demand shocks. Therefore, the application illustrates how the combination of data-driven and restrictions based identification allows to gain deeper insights into the transmission of demand and supply shocks in the oil market.

1 A Generalized Method of Moments Estimator for Structural Vector Autoregressions Based on Higher Moments[5]

1.1 Introduction

One of the most important tools to estimate the effects of economic shocks on a set of variables is the structural vector autoregressive (SVAR) model. Due to the simultaneous nature of the interaction among economic variables, the identification of the underlying structural shocks generally requires the researcher to impose restrictions on the system. A variety of identifying restrictions have been proposed in the SVAR literature. The one factor all identification approaches have in common is the assumption of uncorrelated structural shocks. Unfortunately, uncorrelatedness is not sufficient to identify the simultaneous interaction.

A large part of the SVAR literature eliminates this lack of identification with short- or long-run restrictions. For example, the often used recursive SVAR employs short-run zero restrictions between the included variables. However, these restrictions are often difficult to identify, or hardly justifiable based on economic theory. Therefore, a number of proposals have been made to avoid these restrictions. The general idea is to exploit the independence of the shocks and not merely their uncorrelatedness. With independent and non-Gaussian shocks, results from the independent component analysis (ICA) literature can be applied to identify the SVAR.

This study presents a generalized method of moments (GMM) estimator for non-Gaussian SVAR models with independent shocks. The identification is derived as a straightforward extension of traditional approaches that rely on the assumption of uncorrelated shocks to independent shocks. The approach is purely data-driven and does not require any assumptions or restrictions apart from independent and non-Gaussian shocks (more precisely: at most one shock is allowed to have zero skewness or zero excess kurtosis). In macroeconomic applications, where restrictions are scarce and traditional identification approaches fail, the proposed estimator allows identification and estimation of a given SVAR by exploiting information contained in moments beyond the variance.

Independence has rarely been used to identify SVAR models. A few authors use independent

[5] This is an Accepted Manuscript of an article published by Taylor & Francis in the Journal of Business & Economic Statistics on March 2020, available online: http://wwww.tandfonline.com/10.1080/07350015.2020.1730858

shocks to evaluate the fit of different causal orders (Hyvärinen et al. (2010) or Moneta et al. (2013)). More recently, independence has been used to identify SVAR models without restrictions on the interaction of the included variables. Lanne et al. (2017) and Gouriéroux et al. (2017) propose a maximum likelihood (ML) and a pseudo maximum likelihood (PML) estimator for non-Gaussian SVAR models. Lanne and Luoto (2021) use cokurtosis conditions to derive a GMM estimator for non-Gaussian SVAR models. The authors relax the assumption of independent structural shocks and instead assume uncorrelated shocks with a few shocks additionally satisfying cokurtosis restrictions. Herwartz (2018) proposes a method to find the least dependent shocks, measured by the difference between the empirical copula and the copula under independence. Herwartz and Plödt (2016) apply the method and analyze the interaction of real economic activity, oil production, and the real price of oil.

The SVAR-GMM estimator proposed in this study requires no distributional assumptions, apart from independent and non-Gaussian shocks. In contrast, the estimators proposed by Lanne et al. (2017) and Gouriéroux et al. (2017) require the distribution of the structural shocks to be specified a priori. In macroeconomic applications, distributional restrictions are probably even harder to derive from economic theory than traditional short- or long-run restrictions. The PML estimator proposed by Gouriéroux et al. (2017) is robust to distributional misspecification to some extent. Based on a Monte Carlo study, I demonstrate that misspecifying the distribution can lead to a serious deterioration of the finite sample performance of the PML estimator. I find that the SVAR-GMM estimator performs more robustly across different error term specifications and it performs better than the misspecified PML estimator.

The moment conditions derived in this study ensure global identification up to sign and permutation. However, the number of moment conditions increases quickly with the dimension of the SVAR, which makes the estimator computationally expensive in large models. Lanne and Luoto (2021) propose a GMM estimator that estimates the simultaneous relationships based on a subset of the moment conditions derived herein. Relying only on a subset of the moment conditions yields a computationally cheap estimator, but it also destroys the global identification result. Therefore, the estimator proposed by Lanne and Luoto (2021) is only locally identified. I propose an alternative way to decrease the computational burden of the estimator. By employing a specific weighting scheme, one can gain a consistent, asymptotically normally distributed and computationally cheap estimator, which is denoted as the fast SVAR-GMM estimator. In particular, the fast SVAR-GMM estimator can be derived as the limit of the SVAR-GMM estimator when the weights of the variance and covariance conditions tend to infinite weight. While this weighting scheme obviously deviates from the asymptotically efficient weighting, the Monte Carlo study shows that it leads to a good small sample performance of the fast SVAR-GMM estimator. I find that neither the GMM estimator proposed by Lanne and Luoto (2021) nor the PML estimator proposed by Gouriéroux et al. (2017) (with a pseudo distribution equal to a t-distribution) are able to exploit information contained in the skewness of the structural shocks. Instead, both estimators primarily rely on the excess kurtosis of the shocks. The estimator proposed in this study is more general and can use information contained in the skewness and excess kurtosis. I provide empirical evidence that macroeconomic variables such as economic activity, oil, or stock prices are driven by skewed shocks. The Monte Carlo study reveals that estimators based on the skewness have desirable small sample properties. In particular, I find that the small sample bias and standard deviation of an estimator based on the skewness are driven by the relative skewness of the shocks and I find no deterioration with a decreasing sample size.

In an empirical application, the estimator is applied to analyze the interaction of economic activity, oil, and stock prices. SVAR models with oil and stock prices have often been identified by imposing a recursive order on both variables, see Sadorsky (1999) or Kilian and Park (2009). I challenge this practice and provide evidence that no zero restrictions on the simultaneous relationship between oil and stock prices are feasible.

The remainder of the paper is organized as follows. Section 1.2 presents the SVAR model and derives the identification problem. Section 1.3 illustrates how the standard SVAR identification approach relying on uncorrelated shocks can be extended to independent shocks. Section 1.4 introduces the notation. Section 1.5 derives the identification of the SVAR model based on higher moments and introduces the SVAR-GMM estimator. Section 1.6 derives the fast SVAR-GMM estimator and Section 1.7 analyzes the finite sample properties of the estimators in a Monte Carlo study. The estimator is applied in Section 1.8 to examine the interaction of economic activity, oil, and stock prices. Concluding remarks are provided in Section 1.9.

Throughout this study, real numbers are denoted by $\mathbb{R}$, natural numbers are denoted by $\mathbb{N}$, and the identity matrix is denoted by I. Moreover, the function $vec(.)$ denotes the vectorization of a matrix, $det(.)$ denotes the determinant of a matrix, and the factorial of an integer n is denoted by $n!$.

1.2 SVAR model

This section briefly summarizes the identification problem in SVAR models. A detailed explanation can be found in Kilian and Lütkepohl (2017). The SVAR is given by

$$A_0 y_t = A_1 y_{t-1} + ... + A_p y_{t-p} + \epsilon_t, \tag{1.1}$$

with constant parameter matrices $A_0, ..., A_p \in \mathbb{R}^{n \times n}$, the n-dimensional vector of time series $y_t = [y_{1,t}, ..., y_{n,t}]'$ and the vector of structural shocks $\epsilon_t = [\epsilon_{1,t}, ..., \epsilon_{n,t}]'$. The structural shocks are supposed to satisfy the following assumptions.

Assumption 1.1. *(i) ϵ_t is a vector of i.i.d. random variables.*
(ii) ϵ_t has mutually independent components, meaning that $\epsilon_{i,t}$ is independent of $\epsilon_{j,t}$ for $i \neq j$.
(iii) Each component of ϵ_t has zero mean, unit variance, and finite third and fourth moments.
(iv) At most one component of ϵ has zero skewness and/or at most one component of ϵ has zero excess kurtosis.

The parameter matrix governing the simultaneous interaction is assumed to be invertible.

Assumption 1.2. $A_0 \in \mathcal{A} := \{A \in \mathbb{R}^{n \times n} | det(A) \neq 0\}$.

Equation (1.1) cannot be estimated consistently by OLS since a non-diagonal matrix A_0 leads to endogenous regressors. The reduced form vector autoregression (VAR) is given by

$$y_t = C_1 y_{t-1} + ... + C_p y_{t-p} + u_t. \tag{1.2}$$

The reduced form shocks u_t are i.i.d. and the VAR can be estimated by OLS. However, the estimated reduced form parameters and the reduced form shocks are of limited interest for the structural analysis, which focuses on the structural parameters and the structural shocks. The reduced form shocks can be written as a linear combination of the structural shocks

$$u_t = A_0^{-1} \epsilon_t. \tag{1.3}$$

However, neither the parameters of the matrix governing the simultaneous interaction, nor the structural shocks are known. That is, the structural shocks cannot be directly recovered from the estimated reduced form VAR, leaving us with an identification problem.

Define the unmixed innovations as the vector of random variables obtained by unmixing the reduced form shocks by a matrix $A \in \mathcal{A}$ as

$$e_t(A) := Au_t. \tag{1.4}$$

If the unmixing matrix A is equal to A_0, the unmixed innovations are equal to the structural shocks. I show how to derive a system of moment conditions that globally identifies the matrix governing the simultaneous interaction and the structural shocks up to sign and permutation. The identification requires independent and non-Gaussian structural shocks. Intuitively, if the unmixed innovations and the structural shocks have the same covariance, coskewness, and cokurtosis, the unmixed innovations and structural shocks are equal up to sign and permutation.

Note that equations (1.3) and (1.4) contain no lag structure and the shocks are i.i.d. over time. Therefore, the time index is suppressed whenever possible.

1.3 Illustration: Identification and higher moments

This section uses a bivariate SVAR to illustrate the intuition behind the identification approach presented in Section 1.5. The approach is a straightforward extension of the standard identification scheme relying on uncorrelated shocks to higher moments and independent shocks. Basically, the structural shocks are assumed to be mutually independent and thus uncorrelated, which allows us to postulate moment conditions. However, these second-order moment conditions are not sufficient for identification. Exploiting the implications of independent shocks concerning higher-order moments allows us to postulate additional moment conditions and to identify the SVAR.

Consider a bivariate SVAR such that the unmixed innovations are given by

$$\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \tag{1.5}$$

One can now use Assumption 1.1 to derive the stochastic properties of the unknown structural shocks and choose an unmixing matrix such that the unmixed innovations fulfill the same stochastic properties. Basically, every SVAR identification approach exploits the second-order properties of the structural shocks. In particular, the structural shocks have unit variance and the unmixed innovations should therefore satisfy the following variance conditions

$$1 \stackrel{!}{=} E\left[e_1^2\right] = E\left[(a_{11}u_1 + a_{12}u_2)^2\right] \tag{1.6}$$

$$1 \stackrel{!}{=} E\left[e_2^2\right] = E\left[(a_{21}u_1 + a_{22}u_2)^2\right]. \tag{1.7}$$

Moreover, the components ϵ_1 and ϵ_2 are assumed to be (second-order) uncorrelated and therefore satisfy the covariance condition

$$0 \stackrel{!}{=} E\left[e_1 e_2\right] = E\left[(a_{11}u_1 + a_{12}u_2)(a_{21}u_1 + a_{22}u_2)\right]. \tag{1.8}$$

Exploiting all second-order properties yields three equations in the four unknown coefficients of A. Therefore, infinitely many unmixing matrices A generate unmixed innovations satisfying the second-order properties and second-order statistics are therefore not sufficient to identify A_0.

If one coefficient of A_0 is known a priori, the corresponding coefficient of the unmixing matrix A can be restricted and the system is just identified, for example, see Rubio-Ramírez et al. (2010). Using short-run restrictions of this kind is probably the most common way to solve the identification problem. However, the approach requires the researcher to know half of the simultaneous structure a priori and by relying solely on second-order moments, these restrictions cannot be tested. Technically, short-run restrictions reduce the number of unknowns. Alternatively, one could attempt to increase the number of equations while keeping the number of unknowns constant. Increasing the number of equations until no short-run restrictions are required seems appealing since one does not need to restrict the simultaneous structure a priori. Additional equations can only be generated by imposing more structure on the stochastic properties of structural shocks. The following argument shows how independence accomplishes that.

So far, only second-order properties of the structural shocks have been used. Independent and non-Gaussian structural shocks allow information contained in moments beyond the variance and covariance to be exploited. For independent structural shocks, it is fairly straightforward to generate as many equations as desired. In particular, independence implies third-order uncorrelatedness, $E\left[\epsilon_1^2\epsilon_2\right] = E\left[\epsilon_1^2\right] E\left[\epsilon_2\right] = 0$ and analogously $E\left[\epsilon_1\epsilon_2^2\right] = 0$, and therefore the unmixed innovations should satisfy the coskewness conditions

$$0 \overset{!}{=} E\left[e_1^2 e_2\right] = E\left[(a_{11}u_1 + a_{12}u_2)^2(a_{21}u_1 + a_{22}u_2)\right] \tag{1.9}$$

$$0 \overset{!}{=} E\left[e_1 e_2^2\right] = E\left[(a_{11}u_1 + a_{12}u_2)(a_{21}u_1 + a_{22}u_2)^2\right]. \tag{1.10}$$

Thus, independence allows us to generate further moment conditions analogously to the usual approach based on second moments. If the shocks are non-Gaussian, these moment conditions contain further information that allow the SVAR to be identified. The system of equations (1.6) - (1.10) now contains five equations in the four unknowns. However, the system contains nonlinear equations and thus, it is not obvious whether the system globally identifies the SVAR. Section 1.5 shows that the system indeed identifies the SVAR up to sign and permutation, given that the structural shocks are independent and non-Gaussian.

1.4 Notation

The identification approach requires (co-)moments of order two, three, and four to be calculated. The following notation yields a short expression to collect all (co-)moments of a given order r. For an n-dimensional random variable x, define a moment generating index $W = [w_1, ..., w_r] \in \{1, ..., n\}^r$ with $r \in \mathbb{N}$ and let

$$x_W := [x_{w_1}, ..., x_{w_r}]. \tag{1.11}$$

Let the expected value be denoted by $E\left[x_W\right] := E\left[x_{w_1}...x_{w_r}\right]$ and let $E_T\left[x_W\right]$ denote the respective sample counterpart. Furthermore, define a moment generating set $\mathcal{W} = \{W_1, ..., W_l\}$ with the moment generating indices $W_i \in \{1, ..., n\}^r$ for $i = 1, ..., l$ and define

$$x_{\mathcal{W}} := \begin{bmatrix} x_{W_1} \\ ... \\ x_{W_l} \end{bmatrix}, \; E\left[x_{\mathcal{W}}\right] := \begin{bmatrix} E\left[x_{W_1}\right] \\ ... \\ E\left[x_{W_l}\right] \end{bmatrix} \text{ and } E_T\left[x_{\mathcal{W}}\right] := \begin{bmatrix} E_T\left[x_{W_1}\right] \\ ... \\ E_T\left[x_{W_l}\right] \end{bmatrix}. \tag{1.12}$$

This notation can be used to generate a vector containing the r-th (co-)moments of an n-dimensional random variable.

For $r \in \mathbb{N}$, define the r-th moments generating set as

$$\mathcal{M}(r) = \{[m_1, ..., m_r] \in \{1, ..., n\}^r | m_1 = ... = m_r\}. \tag{1.13}$$

The set contains n elements and the vector $E\left[\epsilon_{\mathcal{M}(r)}\right]$ contains all r-th moments of the structural shocks and the vector $E\left[e_{\mathcal{M}(r)}(A)\right]$ contains all r-th moments of the unmixed innovations. In the bivariate example, the second moments generating set is equal to $\mathcal{M}(2) = \{[1,1], [2,2]\}$ and the variance of the structural shocks is given by

$$E\left[\epsilon_{\mathcal{M}(2)}\right] = E\begin{bmatrix} \epsilon_1\epsilon_1 \\ \epsilon_2\epsilon_2 \end{bmatrix}. \tag{1.14}$$

For $r \in \mathbb{N}$, define the r-th co-moments generating set as

$$\mathcal{C}(r) = \{[c_1, ..., c_r] \in \{1, ..., n\}^r | [c_1, ..., c_r] \notin \mathcal{M}(r) \text{ and } c_i \leq c_j \text{ for } i < j\}. \tag{1.15}$$

The set contains $\frac{(n+r-1)!}{(n-1)!r!} - n$ elements and can be used to generate the corresponding co-moments of an n-dimensional random variable. In particular, the vector $E\left[\epsilon_{\mathcal{C}(r)}\right]$ contains all r-th co-moments of the structural shocks and the vector $E\left[e_{\mathcal{C}(r)}(A)\right]$ contains all r-th co-moments of the unmixed innovations. In the bivariate example, the second co-moments generating set is equal to $\mathcal{C}(2) = \{[1,2]\}$ and the covariance of the structural shocks is given by

$$E\left[\epsilon_{\mathcal{C}(2)}\right] = E\left[\epsilon_1\epsilon_2\right]. \tag{1.16}$$

Analogously, the third co-moment generating set is equal to $\mathcal{C}(3) = \{[1,1,2], [1,2,2]\}$ and the coskewness of the structural shocks is given by

$$E\left[\epsilon_{\mathcal{C}(3)}\right] = E\begin{bmatrix} \epsilon_1\epsilon_1\epsilon_2 \\ \epsilon_1\epsilon_2\epsilon_2 \end{bmatrix}. \tag{1.17}$$

Additionally, for $X = [x_1, ..., x_r] \in \{1, ..., n\}^r$ define the index counting function

$$\#X := \left[\sum_{x \in X} 1_{x=1}, ..., \sum_{x \in X} 1_{x=n} \right], \tag{1.18}$$

where $1_{x=i} = \begin{cases} 1 & \text{, if } x = i \\ 0 & \text{, else} \end{cases}$ such that the i-th element of $\#X$ counts how often the index i appears in X.

1.5 SVAR-GMM estimator

This section generalizes the identification technique outlined in Section 1.3 to an n-dimensional non-Gaussian SVAR. I first derive a system of variance, covariance, coskewness, and cokurtosis conditions, which globally identify the non-Gaussian SVAR up to sign and permutations. The SVAR is then estimated by matching the moments via a GMM estimator.

First, the (co-)moments of the unknown structural shocks need to be derived. The (co-)moments follow from Assumption 1.1.

Proposition 1.1. *Let ϵ satisfy Assumption 1.1. It holds that*

1. *For $[m_1, m_2] \in \mathcal{M}(2)$: $E\left[\epsilon_{[m_1,m_2]}\right] = 1$*
2. *For $[c_1, c_2] \in \mathcal{C}(2)$: $E\left[\epsilon_{[c_1,c_2]}\right] = 0$*
3. *For $[c_1, c_2, c_3] \in \mathcal{C}(3)$: $E\left[\epsilon_{[c_1,c_2,c_3]}\right] = 0$*
4. *For $[c_1, c_2, c_3, c_4] \in \mathcal{C}(4)$: $E\left[\epsilon_{[c_1,c_2,c_3,c_4]}\right] = \begin{cases} 1, \text{ if } c_1 = c_2 \text{ and } c_3 = c_4 \\ 0, \text{ else} \end{cases}$*

Proof. Independence embedded in Assumption 1.1 implies that for $C = [c_1, ..., c_r]$

$$E\left[\epsilon_C\right] = \prod_{i=1}^{n} E\left[\epsilon_i^{\#C_i}\right], \tag{1.19}$$

where $\#C_i$ is the i-th element of the index counting function, which counts how often the index i

appears in C. All statements follow by plugging in the value of each factor implied by Assumption 1.1 (iii). $\square$

One can now match the (co-)moments of the structural shocks with the (co-)moments of the unmixed innovations, which yields n variance conditions

$$E\left[e_{\mathcal{M}(2)}(A)\right] = E\left[\epsilon_{\mathcal{M}(2)}\right] \iff E\left[e_{\mathcal{M}(2)}(A)\right] - 1 = 0, \tag{1.20}$$

$\frac{n(n+1)}{2} - n$ covariance conditions

$$E\left[e_{\mathcal{C}(2)}(A)\right] = E\left[\epsilon_{\mathcal{C}(2)}\right] \iff E\left[e_{\mathcal{C}(2)}(A)\right] = 0, \tag{1.21}$$

$\frac{n(n+1)(n+2)}{6} - n$ coskewness conditions

$$E\left[e_{\mathcal{C}(3)}(A)\right] = E\left[\epsilon_{\mathcal{C}(3)}\right] \iff E\left[e_{\mathcal{C}(3)}(A)\right] = 0 \tag{1.22}$$

and $\frac{n(n+1)(n+2)(n+3)}{24} - n$ cokurtosis conditions

$$E\left[e_{\mathcal{C}(4)}(A)\right] = E\left[\epsilon_{\mathcal{C}(4)}\right], \tag{1.23}$$

with $E\left[\epsilon_{\mathcal{C}(4)}\right]$ as defined in Proposition 1.1.

The SVAR will only be identified up to sign and permutations. Let $\mathcal{P}$ be the set containing all $n \times n$ signed permutation matrices. For any signed permutation matrix $P \in \mathcal{P}$, the shocks $\tilde{\epsilon} := P\epsilon$ and the mixing matrix $\tilde{A}_0 := PA_0$ generate the same reduced form shocks as the shocks ϵ and the mixing matrix A_0. This can easily be verified since $u = A_0^{-1}\epsilon = A_0^{-1}P^{-1}P\epsilon = \tilde{A}_0^{-1}\tilde{\epsilon}$. Moreover, since $\tilde{\epsilon}$ is only a signed permutation of ϵ, both vectors of shocks share the same dependence structure and hence, the identification approach cannot identify the correct sign and permutation. Identification up to permutation is equivalent to the problem of labeling the structural shocks, see Lanne et al. (2017) or Gouriéroux et al. (2017). Labeling and thus attaching a meaning to the shocks cannot be done statistically but remains for the researcher. Since the identification approach cannot identify the correct sign and permutation, I redefine the problem such that the indeterminacy of sign and permutation no longer appears in the new identification problem. Obviously, the indeterminacy is only removed from the statistical side of the problem and the researcher still needs to label the shocks. Note that the indeterminacy of scaling the shocks is already excluded from the identification problem by assuming shocks with unit variance. Define the set of sign-permutation representatives analogously to the set guaranteeing global identification in Lanne et al. (2017) as

$$\mathcal{A}^* := \{A \in \mathcal{A} | \forall i, \, a_{ii} > 0 \text{ and } \forall i < j, \, |a_{ii}| \geq |a_{ji}|\}. \tag{1.24}$$

An element $A \in \mathcal{A}^*$ is called a unique sign-permutation representative if, for any signed permutation matrix $P \in \mathcal{P}$ with $P \neq I$, it holds that $PA \notin \mathcal{A}^*$. The set $\mathcal{A}^*$ fulfills the following properties:

Proposition 1.2. *Almost all elements $A \in \mathcal{A}^*$ are unique sign-permutation representatives. For any matrix $A \in \mathcal{A}$ there exists at least one signed permutation matrix P with $PA \in \mathcal{A}^*$.*

Proof. An inner point $A \in \mathcal{A}^*$ satisfies that $\forall i < j$, $|a_{ii}| > |a_{ji}|$. Let $A \in \mathcal{A}^*$ be an inner point. For any $P \in \mathcal{P}$ with $P \neq I$, it holds that for $\tilde{A} := PA$ indices $i < j$ exist with $|\tilde{a}_{ii}| < |\tilde{a}_{ji}|$ and thus, $\tilde{A} \notin \mathcal{A}^*$. Therefore, an inner point of $\mathcal{A}^*$ is a unique sign-permutation representative. Only the boundary of $\mathcal{A}^*$ contains elements which are not unique sign-permutation representatives. However, the boundary of the n^2 dimensional manifold $\mathcal{A}^*$ has dimension $n^2 - 1$ and is thus a null set in $\mathcal{A}^*$. The second statement is trivial. □

Since A_0 almost surely has a unique representative in $\mathcal{A}^*$, I replace Assumption 1.1 .

Assumption 1.3. *$A_0 \in \mathcal{A}^*$ is a unique sign-permutation representative.*

The following proposition is based on Comon (1994) and shows that the variance, covariance, coskewness, and cokurtosis conditions globally identify the SVAR.

Proposition 1.3. *Let $\epsilon = A_0 u$ satisfy Assumption 1.1 and 1.3. For $A \in \mathcal{A}^*$, it holds that*

$$E \begin{bmatrix} e_{\mathcal{M}(2)}(A) - 1 \\ e_{\mathcal{C}(2)}(A) \\ e_{\mathcal{C}(3)}(A) \\ e_{\mathcal{C}(4)}(A) - E\left[\epsilon_{\mathcal{C}(4)}\right] \end{bmatrix} = 0 \iff A = A_0, \tag{1.25}$$

with $E\left[\epsilon_{\mathcal{C}(4)}\right]$ as defined in Proposition 1.1.

Proof. Let $\tilde{A} \in \mathcal{A}^*$ solve the moment conditions. Let $\tilde{Q} := \tilde{A}A_0^{-1}$ and thus $\tilde{e} = \tilde{A}u = \tilde{Q}\epsilon$. Assumption 1.1 implies that $I = E[\epsilon\epsilon']$ and since $\tilde{e}$ solves the variance and covariance condition, it follows that $\tilde{Q}$ is orthogonal. The coskewness and cokurtosis conditions then imply that all third- and fourth-order cross-cumulants of $\tilde{e}$ are zero. Assumption 1.1 ensures that the shocks are non-Gaussian and have finite moments up to order four. One can therefore apply Comon (1994) Theorem 16 and Comon (1994) equation (3.10), which yields that Q is in $\mathcal{P}$ and thus $\tilde{A}$ is a signed permutation of A_0. With Assumption 1.3, it follows that $\tilde{A} = A_0$. The other direction is trivial. $\square$

If only the first (second) part of Assumption 1.1 (iv) is fulfilled, the cokurtosis (coskewness) conditions can be dropped and the SVAR is still globally identified. However, even if the cokurtosis (or alternatively the coskewness) conditions are dropped, there are still more moment conditions than unknown parameters. Importantly, dropping additional moment conditions immediately destroys the global identification result, see the Appendix 1.10.2. Lanne and Luoto (2021) basically identify the simultaneous interaction with a subsystem of the moment conditions used in Proposition 1.3. In particular, their system contains the variance, covariance, and a subset of the cokurtosis conditions. Therefore, their system is only locally identified. Without the global identification result, the GMM objective function can have minima converging to solutions unequal to signed permutations of A_0. Additionally, it is difficult to derive the consistency and asymptotic normality of the estimator. In fact, the authors argue that the asymptotic properties can be derived under standard assumptions. However, one of these standard assumptions is a globally identified system.

With Proposition 1.3, the matrix A_0 can be estimated by the SVAR-GMM estimator

$$\hat{A}_T(W) := \arg \min_{A \in \mathcal{A}^*} J_T(A, W), \tag{1.26}$$

with a positive semidefinite weighting matrix W and the objective function

$$J_T(A,W) := \begin{pmatrix} E_T\left[e_{\mathcal{M}(2)}(A)\right] - 1 \\ E_T\left[e_{\mathcal{C}(2)}(A)\right] \\ E_T\left[e_{\mathcal{C}(3)}(A)\right] \\ E_T\left[e_{\mathcal{C}(4)}(A)\right] - E\left[\epsilon_{\mathcal{C}(4)}\right] \end{pmatrix}' W \begin{pmatrix} E_T\left[e_{\mathcal{M}(2)}(A)\right] - 1 \\ E_T\left[e_{\mathcal{C}(2)}(A)\right] \\ E_T\left[e_{\mathcal{C}(3)}(A)\right] \\ E_T\left[e_{\mathcal{C}(4)}(A)\right] - E\left[\epsilon_{\mathcal{C}(4)}\right] \end{pmatrix}. \tag{1.27}$$

If the estimator only contains the variance, covariance, and coskewness conditions, it is denoted as the SVAR-GMM estimator based on the coskewness. The SVAR-GMM estimator based on the cokurtosis is defined analogously. The asymptotic properties of the estimator follow from standard arguments. Given standard assumptions, which in particular include global identification ensured by Proposition 1.3, the estimator $\hat{A}_T$ is a consistent estimator for A_0 and it is asymptotically normally distributed with $\sqrt{T}\left(vec(\hat{A}_T) - vec(A_0)\right) \xrightarrow{d} N(0, MSM')$, where M and S are defined as usual, see Hall (2005). Additionally, the weighting matrix $W_{opt} = S^{-1}$ yields the estimator with the minimum asymptotic variance, see Hall (2005). The two-step SVAR-GMM estimator denotes the SVAR-GMM estimator under the standard two-step GMM procedure. Moreover, parameter hypothesis tests can be performed as usual, see Hall (2005). Of course, the interpretation of a test is always subject to the labeling of the shocks, see Lanne et al. (2017). Bonhomme and Robin (2009) note that the asymptotic standard errors depend on the variances of third- and fourth-order moments. In small samples, these moments are difficult to estimate, and the authors suggest using bootstrap based confidence intervals instead of the estimated asymptotic standard errors.

To summarize, if the structural shocks are independent and non-Gaussian, the identification problem can be solved by extending the identification approach from covariance to coskewness and cokurtosis conditions. The SVAR can then be estimated by the SVAR-GMM estimator. However, the number of moment conditions used in Proposition 1.3 increases quickly with the dimension of the SVAR. Therefore, the computational burden of the SVAR-GMM estimator can be high in large models. The next section shows how a specific weighting scheme applied to the SVAR-GMM estimator leads to a consistent, asymptotically normally distributed and computationally cheap estimator.

1.6 Fast SVAR-GMM estimator

This section derives the fast SVAR-GMM estimator. The fast SVAR-GMM estimator can be derived by applying a specific weighting scheme to the SVAR-GMM estimator. First, the estimator is whitened, meaning that the weights of the variance and covariance conditions of the SVAR-GMM estimator tend to infinity. Second, a specific weighting matrix is applied to the higher-order co-moment conditions. Put together, both steps allow a computationally cheap expression of the estimator to be derived. Obviously, when the weighting is used to derive a computationally cheap expression of the estimator, it can no longer be used to achieve asymptotic efficiency of the estimator. Therefore, in large SVAR models, there is a trade-off between asymptotic efficiency and computational expense. However, the Monte Carlo simulations in the next section reveal that the asymptotic efficiency loss does not translate to finite samples. Instead, putting infinite weights on the less volatile second-order moment conditions leads to a good small sample performance of the fast SVAR-GMM estimator.

Consider the SVAR-GMM estimator from equation (1.26) with the weighting matrix

$$W_m := \begin{bmatrix} mI & 0 \\ 0 & W^{(3\&4)} \end{bmatrix}, \tag{1.28}$$

where I is an $\frac{n(n+1)}{2} \times \frac{n(n+1)}{2}$ dimensional identity matrix, such that the variance and covariance conditions receive the weight m, and $W^{(3\&4)}$ is the weighting matrix corresponding to the coskewness and cokurtosis conditions. The weighting matrix W_m splits the objective function of the SVAR-GMM estimator into a bi-objective function, where the first objective is to minimizes the variance and covariance conditions and the second objective is to minimize the higher-order co-moment conditions. The SVAR-GMM estimator can then be written as

$$\hat{A}_T(W_m) = \arg \min_{A \in \mathcal{A}^*} m J_T^{(2)}(A, I) + J_T^{(3,4)}(A, W^{(3,4)}), \tag{1.29}$$

with

$$J_T^{(2)}(A, I) = \begin{pmatrix} E_T\left[e_{\mathcal{M}(2)}(A)\right] - 1 \\ E_T\left[e_{\mathcal{C}(2)}(A)\right] \end{pmatrix}' I \begin{pmatrix} E_T\left[e_{\mathcal{M}(2)}(A)\right] - 1 \\ E_T\left[e_{\mathcal{C}(2)}(A)\right] \end{pmatrix} \tag{1.30}$$

and

$$J_T^{(3,4)}(A, W^{(3,4)}) = \begin{pmatrix} E_T\left[e_{\mathcal{C}(3)}(A)\right] \\ E_T\left[e_{\mathcal{C}(4)}(A)\right] - E\left[\epsilon_{\mathcal{C}(4)}\right] \end{pmatrix}' W^{(3,4)} \begin{pmatrix} E_T\left[e_{\mathcal{C}(3)}(A)\right] \\ E_T\left[e_{\mathcal{C}(4)}(A)\right] - E\left[\epsilon_{\mathcal{C}(4)}\right] \end{pmatrix}. \tag{1.31}$$

When the weight m of the variance and covariance conditions tends to infinity, these moment conditions become constraints to the optimization problem and yield the whitened SVAR-GMM estimator defined as

$$\begin{aligned} \hat{A}_T^{white}(W^{(3,4)}) &:= \arg\min_{A \in \mathcal{A}^*} J_T^{(3,4)}(A, W^{(3,4)}) \\ \text{s.t. } & J_T^{(2)}(A, I) = 0. \end{aligned} \tag{1.32}$$

Proposition 1.4. *Let $\epsilon = A_0 u$ satisfy Assumption 1.1 and 1.3. Then any limit point of the series $\hat{A}_T(W_m)$ for m to infinity solves the optimization problem in equation (1.32).*

Proof. Note that the term $mJ_T^{(2)}(A, I)$ in equation (1.29) is a quadratic penalty term which penalizes violations of the variance and covariance conditions. The statement then follows from the convergence of penalty function methods, see Luenberger et al. (2016). □

The whitened estimator puts infinite weight on the variance and covariance conditions. While this weighting clearly deviates from the asymptotically efficient weighting matrix of the SVAR-GMM estimator, the next section shows that putting higher weights on the less volatile second-order moment conditions results in a good small sample performance of the estimator.

The whitened SVAR-GMM estimator is a GMM estimator subject to restrictions. Let $V_T V_T' = E_T[uu']$ be the Cholesky decomposition of the sample variance-covariance matrix of the reduced form shocks. The restriction in equation (1.32) is then equivalent to the restriction

$$(AV_T)(AV_T)' = I, \tag{1.33}$$

meaning that AV_T is restricted to be an orthogonal matrix. In practice, it is not necessary to keep track of this restriction. Instead, the whitened SVAR-GMM estimator can be calculated as $\hat{A}_T^{white}(W^{(3,4)}) = \hat{O}_{V_T,T}(W^{(3,4)})V_T^{-1}$ with

$$\hat{O}_{V_T,T}(W^{(3,4)}) := \arg\min_{O\in\mathcal{O}} J_T^{(3,4)}(OV_T^{-1}, W^{(3,4)}), \tag{1.34}$$

where $\mathcal{O}$ denotes the set of orthogonal matrices. Therefore, the constraint is an orthogonal constraint and the optimization problem can be transformed into an unconstrained problem over an Euclidean space, for example, see Lezcano-Casado and Martınez-Rubio (2019). For simplicity, the indeterminacy of sign and permutation has been ignored, but could easily be fixed by defining $\mathcal{O}^* := \{O \in \mathcal{O} | OV_T^{-1} \in \mathcal{A}^*\}$.

Due to the random variable V_T, the restriction of the whitened SVAR-GMM estimator is not static. Therefore, consistency and asymptotic normality of the whitened SVAR-GMM estimator do not follow from standard arguments but can be derived analogously.

Proposition 1.5. *Let $\epsilon = A_0 u$ satisfy Assumption 1.1 and 1.3 and let the standard assumptions used in Lemma 5.2 and Lemma 5.4 in Hall (2005) hold. Furthermore, assume uniform convergence in probability, $\sup_{O\in\mathcal{O}} |J_T^{(3,4)}(OV_T^{-1}) - J_T^{(3,4)}(OV^{-1})| \xrightarrow{p} 0$ for $m \to \infty$, where V is the Cholesky decomposition of the variance-covariance matrix, $VV' = E[uu']$, and V_T is the sample analog. The whitened SVAR-GMM estimator then satisfies the following properties:*

- *Consistency: $\hat{A}_T^{white}(W^{(3,4)}) \xrightarrow{P} A_0$*
- *Asymptotic normality: $\sqrt{T}(vec(\hat{A}_T^{white}(W^{(3,4)})) - vec(A_0)) \xrightarrow{d} N(0, Z)$, where the variance Z is given in Appendix 1.10.1.*

Proof. Replacing the restriction $(AV_T)(AV_T)' = I$ with $(AV)(AV)' = I$ yields the estimator $\hat{A}_{V,T}(W^{(3,4)})$. The restriction of $\hat{A}_{V,T}(W^{(3,4)})$ no longer contains a random variable and the consistency of $\hat{A}_{V,T}(W^{(3,4)})$ follows from Lemma 5.2 in Hall (2005). With V_T being a consistent estimator for V and the uniform convergence assumption, the consistency of $\hat{A}_T^{white}(W^{(3,4)})$ can be established. Asymptotic normality can be shown analogous to Lemma 5.4 in Hall (2005). A detailed proof can be found in Appendix 1.10.1. $\square$

Having established the asymptotic properties of the whitened SVAR-GMM, the next step in deriving the fast SVAR-GMM estimator is to apply a specific weighting matrix $W^{(3,4)}_{fast}$ to the third- and fourth-order co-moment conditions. The fast weighting matrix allows a computationally cheap expression for the whitened SVAR-GMM estimator to be established. The key idea is related to a fundamental idea of ICA; minimizing the dependency means to maximize non-Gaussianity. Define the fast weighting matrix $W^{(3,4)}_{fast}$ as the diagonal matrix where the diagonal element corresponding to the co-moment $C \in \mathcal{C}(r)$ with $r \in \{3, 4\}$ is defined as

$$w^{(r)}_{fast}(C) := \binom{r}{\#C} = \frac{r!}{\prod_{i=1}^{n} \#C_i!}, \tag{1.35}$$

where $\#C_i$ denotes the i-th element of the counting function $\#C$. Thus, for example, the weight corresponding to the coskewness condition $E\left[e_1^2 e_2^1\right]$ in equation (1.9) is equal to $\frac{3!}{2!1!} = 3$. The fast weighting matrix can now be used to derive a computationally cheap expression of the whitened estimator, which is henceforth denoted as the fast SVAR-GMM estimator.

Proposition 1.6. *Let $\epsilon = A_0 u$ satisfy Assumption 1.1 and 1.3. The estimator $\hat{A}^{white}_T(W^{(3,4)}_{fast})$ is equal to*

$$\begin{aligned} \hat{A}^{white}_T(W^{(3,4)}_{fast}) =& arg \max_{A \in \mathcal{A}^*} H^{(3,4)}_T(A) \\ & s.t.\ J^{(2)}_T(A, I) = 0. \end{aligned} \tag{1.36}$$

with

$$H^{(3,4)}_T(A) := \sum_{M \in \mathcal{M}(3)} E_T\left[e_M(A)\right]^2 + \sum_{M \in \mathcal{M}(4)} \left(E_T\left[e_M(A)\right] - 3\right)^2. \tag{1.37}$$

Proof. The weights are constructed such that for a given sample of size T there exists a constant $\omega_T \in \mathbb{R}$ with $H^{(3,4)}_T(A) + J^{(3,4)}_T(A, W^{(3,4)}_{fast}) = \omega_T$ for all $A \in \mathcal{A}^*$ with $J^{(2)}_T(A, I) = 0$, which is the sample counterpart of equation (3.10) in Comon (1994). Rearranging yields $J^{(3,4)}_T(A, W^{(3,4)}_{fast}) = \omega_T - H^{(3,4)}_T(A)$, which proves the Proposition. The proof is written down in more detail in Appendix 1.10.1. $\square$

The whitened SVAR-GMM estimator with the fast weighting matrix can therefore be calculated by equation (1.32) or equation (1.36). The former calculates the estimator by minimizing the dependencies of the shocks captured by the objective function $J_T^{(3,4)}(A, W_{fast}^{(3,4)})$ and the latter by maximizing the non-Gaussianity of the shocks captured by the objective function $H_T^{(3,4)}(A)$. When the fast weighting matrix is used, Proposition 1.6 yields that both optimization problems have the same solution.

In a bivariate SVAR, the objective function $J_T^{(3,4)}(A, W_{fast}^{(3,4)})$ is equal to

$$\begin{aligned} J_T^{(3,4)}(A, W_{fast}^{(3,4)}) =& \frac{3!}{2!1!} E_T\left[e_1(A)^2 e_2(A)\right]^2 + \frac{3!}{1!2!} E_T\left[e_1(A) e_2(A)^2\right]^2 \\ &+ \frac{4!}{3!1!} E_T\left[e_1(A)^3 e_2(A)\right]^2 + \frac{4!}{2!2!} E_T\left[e_1(A)^2 e_2(A)^2 - 1\right]^2 + \frac{4!}{1!3!} E_T\left[e_1(A) e_2(A)^3\right]^2 . \end{aligned} \tag{1.38}$$

and the objective function $H_T^{(3,4)}(A)$ is equal to

$$H_T(A) = E_T\left[e_1(A)^3\right]^2 + E_T\left[e_2(A)^3\right]^2 + E_T\left[e_1(A)^4 - 3\right]^2 + E_T\left[e_2(A)^4 - 3\right]^2 . \tag{1.39}$$

In the bivariate example, the computational advantage of maximizing the non-Gaussianity of the shocks compared to minimizing the dependencies of the shocks is small, since the former requires five co-moments to be calculated and the latter requires four moments. However, the computational advantage pays out in larger SVAR models. For example, an SVAR with five variables leads to an objective function $J_T^{(3,4)}(A, W_{fast}^{(3,4)})$ that contains 172 terms measuring the unmixed innovations dependencies, compared to an objective function $H_T^{(3,4)}(A)$ containing 10 terms measuring the unmixed innovations' non-Gaussianity.

The fast SVAR-GMM estimator was derived by applying a weighting scheme to the SVAR-GMM estimator. The fast SVAR-GMM estimator remains computationally cheap in large SVAR models. However, as a trade-off, one can no longer use the weighting matrix to achieve asymptotic efficiency of the estimator. The next section shows that even though the fast SVAR-GMM estimator may be asymptotically inefficient, it performs well in small samples and often outperforms the asymptotically efficient two-step SVAR-GMM estimator.

1.7 Finite sample properties

This section analyzes the finite sample performance of the SVAR-GMM estimators and compares it to the GMM estimator proposed by Lanne and Luoto (2021) and to the PML estimator proposed by Gouriéroux et al. (2017). I find that the SVAR-GMM estimators perform more robustly throughout different specifications than the two alternatives. Moreover, the simulations show that putting infinite weight on second-order moment conditions leads to a good performance of the fast SVAR-GMM estimator in small samples. Additionally, the Monte Carlo simulation sheds light on the impact of the degree of non-Gaussianity on the finite sample performance of SVAR-GMM estimators. I find that estimators based on the skewness have desirable small sample properties as the bias and standard deviation of the estimators is found to be almost solely determined by the relative skewness and show no deterioration with a decreasing sample size.

The setup of the Monte Carlo study is similar to the setup in Gouriéroux et al. (2017) with $u = A_0^{-1}\epsilon$ and

$$A_0^{-1} = \begin{bmatrix} cos(\phi) & -sin(\phi) \\ sin(\phi) & cos(\phi) \end{bmatrix}, \tag{1.40}$$

where $\phi = -\pi/5$. The shocks are drawn from a distribution of the Pearson distribution family with mean zero, unit variance, and different skewness/kurtosis parameters. A high excess kurtosis (skewness) is defined as the excess kurtosis (skewness) a standard normally distributed shock does not exceed in a sample of 200 observations with a probability of 99.99% and is equal to 2.33 (0.53). The first Monte Carlo study contains three specifications. In the first specification, both structural shocks have zero skewness and a high excess kurtosis. In the second specification, both shocks have a high skewness and zero excess kurtosis. In the third specification, the first shock is Gaussian, and the second shock has a high skewness and high excess kurtosis. Table 1.1 shows the mean bias and scaled standard deviation of the two-step SVAR-GMM estimator, the fast SVAR-GMM estimator, the GMM estimator proposed by Lanne and Luoto (2021) with the cokurtosis condition $E\left[\epsilon_1\epsilon_2^3\right] = 0$, and the PML estimator proposed by Gouriéroux et al. (2017) using a t-distribution with twelve degrees of freedom. The results are largely unaffected by the chosen degrees of freedom.

Table 1.1: Finite sample performance - Comparison of unrestricted non-Gaussian SVAR estimators.

	$T=200$			$T=500$			$T=5000$		
Specification	1	2	3	1	2	3	1	2	3
$\hat{A}(W^{\text{2-step}})$	-0.03	-0.01	-0.04	-0.02	-0.01	-0.02	0	0	0
	(6.43)	(0.99)	(5.46)	(4.96)	(0.69)	(5.45)	(2.97)	(0.62)	(3.83)
$\hat{A}(W^{fast})$	-0.02	-0.02	-0.03	-0.01	0	-0.01	0	0	0
	(4.02)	(3.05)	(5.02)	(3.84)	(2.51)	(4.8)	(3.48)	(1.22)	(4.53)
GMM_{LL}	-0.03	-0.04	-0.03	-0.01	-0.03	-0.01	0	-0.02	0
	(4.18)	(7.48)	(4.29)	(5.61)	(18.08)	(5.07)	(5.41)	(177.26)	(4.8)
PML	-0.01	-0.1	-0.04	0	-0.1	-0.02	0	-0.15	0
	(3.1)	(16.14)	(6.05)	(2.57)	(43.25)	(7.89)	(2.15)	(591.88)	(4.97)

Monte Carlo simulation with sample sizes 200, 500, *and* 5000 *each with* 10000 *iterations. For an estimator* $\hat{A}$ *of* A_0*, define the estimator* $\hat{B} := \hat{A}^{-1}$ *of* $B := A_0^{-1}$*. Each entry shows the mean bias,* $E\left[\hat{b}_{1,1} - b_{1,1}\right]$*, and the standard deviation of* $\sqrt{T}\left(\hat{b}_{1,1} - b_{1,1}\right)$ *is shown in parentheses. Moreover, the element* $b_{1,1}$ *of* B *is equal to* $cos(-\pi/5)$*. In specification one, both structural shocks have zero skewness and a high excess kurtosis. In the specification two, both structural shocks have a high skewness and zero excess kurtosis. In specification three, the first shock is Gaussian and the second shock has a high skewness and high excess kurtosis. The SVAR-GMM estimator denoted by* $\hat{A}(W^{2\text{-}step})$ *is the two-step SVAR-GMM estimator and the fast SVAR-GMM estimator is denoted by* $\hat{A}(W^{fast})$*. The GMM estimator proposed by Lanne and Luoto (2021) is denoted by* GMM_{LL} *and uses the cokurtosis condition* $E\left[\epsilon_1\epsilon_2^3\right] = 0$*. The PML estimator proposed by Gouriéroux et al. (2017) is denoted by PML and assumes a t-distribution with twelve degrees of freedom.*

The PML estimator performs best in the first specification. This is not too surprising since the PML estimator is essentially correctly specified (the shocks are drawn from a Pearson Type VII distribution, which contains the t-distribution). However, the advantage of the PML estimator compared to the SVAR-GMM estimators decreases with an increasing sample size. Moreover, the performance of the PML estimator deteriorates with the degree of misspecification and it performs worse than the SVAR-GMM estimators in specification two and three. In the second specification, the misspecification of the PML estimator has particularly severe consequences as the estimator cannot exploit the skewness to identify and consistently estimate the parameter (see Appendix 1.10.3). The GMM estimator proposed by Lanne and Luoto (2021) performs best in the third specification and small samples. However, its advantage vanishes with an increasing sample size and in the largest sample, it performs worse than the SVAR-GMM estimators. Just like the PML estimator, the GMM estimator proposed by Lanne and Luoto (2021) cannot identify and consistently estimate the parameter in the second specification. Moreover, the estimator is only locally identified and its performance depends on an optimization algorithm with a starting value close to the true solution; see the Appendix 1.10.4 for an extension to a grid of starting values.

The two-step SVAR-GMM estimator uses an estimate of the asymptotically efficient weighting matrix and in large samples it performs better than the fast SVAR-GMM estimator. However, the Monte Carlo simulation shows that this result does not translate to small samples. In fact, when the sample size is small the fast SVAR-GMM estimator performs better than the two-step SVAR-GMM estimator in two out of three specifications. The two-step SVAR-GMM estimator estimates the variances of the moment conditions and adjusts the weights, such that less volatile moments receive higher weights. This weighting is asymptotically efficient, however, in small samples the variances of the moment conditions are difficult to estimate (compare this to Bonhomme and Robin (2009)). Therefore, the weighting used by the two-step SVAR-GMM estimator may rely on a few outliers and can be far away from being efficient. In contrast, the fast SVAR-GMM estimator does not rely on estimated variances, but by construction, it puts infinite weight on the less volatile second-order moment conditions, which explains the good finite sample performance of the fast SVAR-GMM estimator.

In a second Monte Carlo simulation, I analyze the impact of the degree of non-Gaussianity and the sample size on the SVAR-GMM estimators. I find that the results do not depend on the weighting matrix and I therefore only report the results for the fast SVAR-GMM estimator. The degree of non-Gaussianity is now chosen in relative terms. A low relative skewness (excess kurtosis) is defined as the sample skewness (excess kurtosis) a standard normally distributed shock will not exceed in a sample of size T with a probability of 90%. A medium and high relative skewness (excess kurtosis) is defined analogously for a probability of 99% and 99.99%. The values are calculated by bootstrap and are shown in Table 1.2. Defining the non-Gaussianity in relative terms allows us to compare the impact of the skewness and excess kurtosis. Moreover, it allows us to disentangle the effects of the non-Gaussianity and the sample size. The first part of

Table 1.2: Relative non-Gaussianity.

	rel. skewness			rel. excess kurtosis		
	low	medium	high	low	medium	high
$T = 200$	0.22	0.4	0.68	0.4	0.97	2.33
$T = 500$	0.14	0.26	0.41	0.27	0.6	1.25
$T = 5000$	0.04	0.08	0.13	0.09	0.17	0.3

The table shows the quantiles (low is 0.9*, medium is* 0.99*, high is* 0.9999*) of the sample skewness and sample excess kurtosis of standard normally distributed shocks in a sample of size T in a Monte Carlo simulation with* 5000000 *simulated samples for each sample size.*

Table 1.3 shows the impact of the relative skewness on the fast SVAR-GMM estimators based on coskewness conditions and the second part of the table shows the impact of the relative excess kurtosis on the fast SVAR-GMM estimators based on cokurtosis conditions. In the first part of the table, the structural shocks have zero excess kurtosis and a low, medium, or high relative skewness. In the second part of the table, the shocks have zero skewness and a low, medium, or high relative excess kurtosis.

Keeping the sample size unchanged and increasing the degree of non-Gaussianity has a positive impact on the finite sample properties of the estimators. Intuitively, this finding is comparable to the strength or weakness of an instrument in an instrumental variables estimation. Decreasing the sample size and keeping the degree of non-Gaussianity unchanged reveals an important difference between both estimators. The bias and standard deviation of the estimator based on the coskewness appear to be entirely determined by the relative skewness and do not vary across sample sizes. In contrast, the bias and standard deviation of the estimator based on the cokurtosis increase with a decreasing sample size. This finding can be explained by the sample variance of the moment conditions, which increases with an increase of the excess kurtosis. The effect is more pronounced in small samples and partly offsets the positive impact of a higher excess kurtosis.

Table 1.3: Finite sample performance - The impact of non-Gaussianity.

	$T = 200$			$T = 500$			$T = 5000$		
Skewness	low	med	high	low	med	high	low	med	high
Exc. kurtosis	zero	zero	zero	zero	zero	zero	zero	zero	zero
$\hat{A}_{r=3}(W^{fast})$	-0.05 (0.23)	-0.03 (0.16)	-0.01 (0.08)	-0.06 (0.23)	-0.03 (0.15)	-0.01 (0.09)	-0.06 (0.24)	-0.03 (0.16)	-0.01 (0.09)
Skewness	zero	zero	zero	zero	zero	zero	zero	zero	zero
Exc. kurtosis	low	med	high	low	med	high	low	med	high
$\hat{A}_{r=4}(W^{fast})$	-0.05 (0.22)	-0.03 (0.18)	-0.02 (0.14)	-0.05 (0.21)	-0.03 (0.16)	-0.01 (0.11)	-0.05 (0.2)	-0.02 (0.13)	-0.01 (0.07)

Monte Carlo simulation with sample sizes 200, 500, and 5000 each with 10000 iterations. For an estimator $\hat{A}$ of A_0, define the estimator $\hat{B} := \hat{A}^{-1}$ of $B := A_0^{-1}$. Each entry shows the mean bias, $E\left[\hat{b}_{1,1} - b_{1,1}\right]$, and the standard deviation of $\left(\hat{b}_{1,1} - b_{1,1}\right)$ is shown in parentheses. Moreover, the element $b_{1,1}$ of B is equal to $\cos(-\pi/5)$. In the first part of the table, the structural shocks have zero excess kurtosis and a low, medium, or high relative skewness. In the second part of the table, the shocks have zero skewness and a low, medium, or high relative excess kurtosis. The relative skewness and relative excess kurtosis values are shown in Table 1.2. The fast SVAR-GMM estimator based on the skewness is denoted by $\hat{A}_{r=3}(W^{fast})$ and the fast SVAR-GMM estimator based on the kurtosis is denoted by $\hat{A}_{r=4}(W^{fast})$.

In summary, I find that the SVAR-GMM estimators perform more robustly than the PML estimator proposed by Gouriéroux et al. (2017) and the GMM estimator proposed by Lanne and Luoto (2021). Moreover, putting infinite weight on the less volatile lower-order moment conditions leads to a good performance of the fast SVAR-GMM estimator in small samples. Finally, the simulations show how the degree of non-Gaussianity influences the finite sample performance of the SVAR-GMM estimators and reveals desirable small sample properties of estimators based on the skewness.

1.8 Economic activity, oil, and stock prices

This section applies the SVAR-GMM estimator to analyze the simultaneous relationship between economic activity, oil, and stock prices. It may be convincing to argue that the real economy behaves sluggishly and does not respond to stock and oil price shocks contemporaneously. However, convincing arguments on how to contemporaneously restrict the stock and oil market are lacking. Nevertheless, in applications (e.g., Sadorsky (1999), Kilian and Park (2009), or Apergis and Miller (2009)), stock market shocks have been restricted to have no simultaneous impact on oil prices. I provide evidence that the real economy may indeed behave sluggishly, but oil and stock prices do not.

The SVAR is estimated with monthly US data from 1990 to 2018 and contains three variables: a measure of real economic activity (EA), monthly real S&P 500 returns (SP), and the monthly growth rates of real oil prices (OP). Real economic activity is measured as 100 times the log difference of the monthly US Industrial Production Index. Real S&P 500 returns are calculated as 100 times the log difference of the S&P 500 closing price deflated by the US CPI. The monthly growth rates of real oil price are calculated as 100 times the log difference of the crude oil composite acquisition cost by refiners deflated by the US CPI. See Appendix 1.10.5 for more information on the data sources.

The SVAR is given by

$$A_0 \begin{bmatrix} EA_t \\ SP_t \\ OP_t \end{bmatrix} = \alpha + \sum_{i=1}^{p} A_i \begin{bmatrix} EA_{t-i} \\ SP_{t-i} \\ OP_{t-i} \end{bmatrix} + \begin{bmatrix} \epsilon_t^{EA} \\ \epsilon_t^{SP} \\ \epsilon_t^{OP} \end{bmatrix}. \tag{1.41}$$

Based on the AIC criterion, I estimate the reduced form with a lag length of four months. The moments of the residuals and the p-value of the Jarque-Bera test are shown in Table 1.4. The Jarque-Bera test indicates non-Gaussian residuals with non-zero skewness and a positive excess kurtosis. Based on the Monte Carlo simulation, I use the two-step SVAR-GMM estimator

Table 1.4: Economic activity, oil, and stock prices - Reduced form residuals.

	Variance	Skewness	Kurtosis	JB-Test
u^{EA}	0.28	-0.84	10.71	0.00
u^{OP}	44.49	0.14	3.73	0.02
u^{SP}	15.53	-0.37	3.55	0.01

The JB-Test shows the p-value of the Jarque-Bera test.

to estimate the simultaneous relationship. The results presented below are robust to different specifications and estimators, see Appendix 1.10.5.

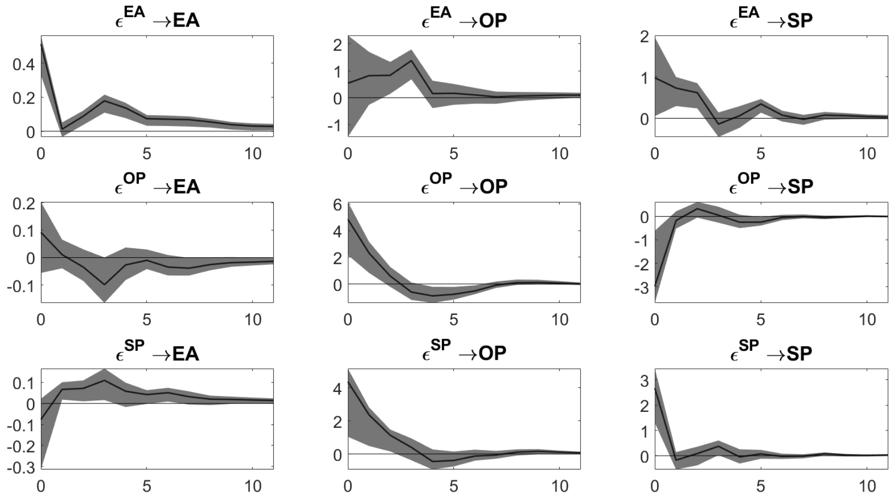
The impulse response function (IRF) is shown in Figure 1.1. The shocks are labeled such that economic activity shocks have a positive impact on oil and stock prices, oil price shocks have a long-run negative impact on economic activity, and the remaining shock is labeled as the stock market shock. According to the IRF, oil price shocks lead to a lagged decrease of economic activity and an immediate decrease of stock returns. Stock market shocks lead to a long-run increase of economic activity and an immediate increase of oil prices. Therefore, real economic activity behaves sluggishly with no simultaneous response to stock or oil market shocks. However, the stock and oil markets are found to interact simultaneously and no recursive order of both variables appears viable.

1.9 Conclusion

This study proposes an identification approach based on higher moments, which is derived as a straightforward extension of the usual SVAR identification approach relying on uncorrelated shocks to independent shocks. Exploiting the skewness and excess kurtosis of the shocks allows a non-Gaussian SVAR with independent structural shocks to be identified. The identification result is used to derive the SVAR-GMM estimator and, as a computationally cheap alternative, the fast SVAR-GMM estimator.

In the Monte Carlo simulation, I find that the SVAR-GMM estimators perform more robustly

Figure 1.1: Economic activity, oil, and stock prices - Impulse responses.



Confidence intervals are calculated by bootstrap with 10000 *replications. The interval shows the upper* 0.9 *and lower* 0.1 *percentiles. The reduced form VAR is estimated with four lags and the simultaneous interaction is estimated by the two-step SVAR-GMM estimator.*

than two alternatives proposed in the literature. Moreover, due to its asymptotic efficiency, the two-step SVAR-GMM estimator performs superior to the fast SVAR-GMM estimator in large samples. However, in small samples the fast SVAR-GMM estimator often performs better than the two-step SVAR-GMM estimator.

Finally, the empirical applications analyze the interaction between real economic activity, stock and oil markets. I find that stock and oil prices interact simultaneously, while the real economy appears to behave sluggishly with no contemporaneous reaction to oil and stock market shocks. The application therefore illustrates how an SVAR can be estimated based on higher moments without relying on incredible short-run restrictions.

1.10 Appendix

1.10.1 Appendix - Proofs

Proof of Proposition 1.5. Let $V_T V_T' = E_T[uu']$ be the Cholesky decomposition of the sample variance-covariance matrix of the reduced form shocks and let $VV' = E[uu']$ be the Cholesky decomposition of the variance-covariance matrix of the reduced form shocks. Define

$$\hat{A}^{white}_{V,T}(W^{(3,4)}) := \arg \min_{A \in \mathcal{A}^*} J^{(3,4)}_T(A, W^{(3,4)}) \tag{1.42}$$
$$\text{s.t. } J^{(2)}(A, I) = 0,$$

where

$$J^{(2)}(A, I) = \begin{pmatrix} E\left[e_{\mathcal{M}(2)}(A)\right] - 1 \\ E\left[e_{\mathcal{C}(2)}(A)\right] \end{pmatrix}' I \begin{pmatrix} E\left[e_{\mathcal{M}(2)}(A)\right] - 1 \\ E\left[e_{\mathcal{C}(2)}(A)\right] \end{pmatrix} \tag{1.43}$$

and note that $J^{(2)}(A, I) = 0$ is equivalent to $(AV)(AV)' = I$, meaning that (AV) is a orthogonal. Consistency of $\hat{A}^{white}_{V,T}(W^{(3,4)})$ follows from Lemma 5.2 of Hall (2005) and uses global identification analogous to Proposition 1.3.
Let $\mathcal{O}^* \subset \mathcal{O}$ be a set of unique sign-permutation representatives of orthogonal matrices and let $O_0 := A_0 V \in \mathcal{O}^*$. Given a weighting matrix $W^{(3,4)}$ define

$$\hat{O}_{V,T} := \arg \min_{O \in \mathcal{O}^*} J^{(3,4)}_T(OV^{-1}, W^{(3,4)}) \tag{1.44}$$

and note that $\hat{A}^{white}_{V,T}(W^{(3,4)}) = \hat{O}_{V,T} V^{-1}$ and therefore $\hat{O}_{V,T} \xrightarrow{p} O_0$. Define

$$\hat{O}_{V_T,T} := \arg \min_{O \in \mathcal{O}^*} J^{(3,4)}_T(OV_T^{-1}, W^{(3,4)}) \tag{1.45}$$

and note that $\hat{A}^{white}_T(W^{(3,4)}) = \hat{O}_{V_T,T} V_T^{-1}$. Consistency of $\hat{O}_{V_T,T}$ requires to show that

$$\underset{T\to\infty}{lim} P\left[0 \leq J^{(3,4)}_T(\hat{O}_{V_T,T} V^{-1}) < \epsilon\right] = 1, \tag{1.46}$$

for $\epsilon > 0$, meaning that $\hat{O}_{V_T,T}$ minimizes $J_T(OV^{-1})$ with probability one as $T \to \infty$. The uniform convergence assumption implies that

$$\underset{T\to\infty}{lim} P\left[J^{(3,4)}_T(\hat{O}_{V_T,T} V^{-1}) < J^{(3,4)}_T(\hat{O}_{V_T,T} V_T^{-1}) + \frac{4}{\epsilon}\right] = 1. \tag{1.47}$$

Moreover, $\hat{O}_{V_T,T}$ minimizes $J_T(OV_T^{-1})$ and thus

$$\underset{T\to\infty}{lim} P\left[J_T^{(3,4)}(\hat{O}_{V_T,T}V_T^{-1}) < J_T^{(3,4)}(\hat{O}_{V,T}V_T^{-1}) + \frac{4}{\epsilon}\right] = 1. \tag{1.48}$$

Additionally, the uniform convergence assumption implies that

$$\underset{T\to\infty}{lim} P\left[J_T^{(3,4)}(\hat{O}_{V,T}V_T^{-1}) < J_T^{(3,4)}(\hat{O}_{V,T}V^{-1}) + \frac{4}{\epsilon}\right] = 1. \tag{1.49}$$

Finally, consistency of $\hat{O}_{V,T}$ implies

$$\underset{T\to\infty}{lim} P\left[J_T^{(3,4)}(\hat{O}_{V,T}V^{-1}) < \frac{4}{\epsilon}\right] = 1. \tag{1.50}$$

Equation (1.46) then follows from $J_T(\hat{O}_{V_T,T}V^{-1}) \geq 0$ and equations (1.47), (1.48), (1.49), (1.50). Consistency of $\hat{O}_{V_T,T}$ can then be established analogous to part (ii) in the proof of Theorem 3.1. in Hall (2005). With $V_T \overset{p}{\to} V$ and $\hat{O}_{V_T,T} \overset{p}{\to} O_0$ it follows that $\hat{A}_T^{white}(W^{(3,4)}) = \hat{O}_{V_T,T}V_T^{-1} \overset{P}{\to} O_0V^{-1} = A_0$.

The proof of the asymptotic normality is analogous to Lemma 5.4 in Hall (2005). Denote the whitened SVAR-GMM estimator as $\hat{A} := \hat{A}_T^{white}(W^{(3,4)})$ and for simplicity assume that $\hat{A}$ and A_0 are already vectorized and suppress the $vec(.)$ notation. Note that the whitened SVAR-GMM estimator is equal to

$$\begin{aligned} \hat{A} =& \arg \min_{A\in\mathcal{A}^*} J_T^{(3,4)}(A, W^{(3,4)}) \\ & \text{s.t. } g_T^{(2)}(A) = 0, \end{aligned} \tag{1.51}$$

with

$$g_T^{(2)}(A) := \begin{pmatrix} E_T\left[e_{\mathcal{M}(2)}(A)\right] - 1 \\ E_T\left[e_{\mathcal{C}(2)}(A)\right] \end{pmatrix}. \tag{1.52}$$

Furthermore, let

$$g_T^{(3,4)}(A) := \begin{pmatrix} E_T\left[e_{\mathcal{C}(3)}(A)\right] \\ E_T\left[e_{\mathcal{C}(4)}(A)\right] - E\left[\epsilon_{\mathcal{C}(4)}\right] \end{pmatrix}. \tag{1.53}$$

and let $G_T^{(2)}(A) := E_T\left[\partial \begin{bmatrix} e_{\mathcal{M}(2)}(A) \\ e_{\mathcal{C}(2)}(A) \end{bmatrix} /\partial A'\right]$, $G_0^{(2)}(A) := E\left[\partial \begin{bmatrix} e_{\mathcal{M}(2)}(A) \\ e_{\mathcal{C}(2)}(A) \end{bmatrix} /\partial A'\right]$ and $G_T^{(3,4)}(A) := E_T\left[\partial \begin{bmatrix} e_{\mathcal{C}(3)}(A) \\ e_{\mathcal{C}(4)}(A) \end{bmatrix} /\partial A'\right]$ as well as $G_0^{(3,4)}(A) := E\left[\partial \begin{bmatrix} e_{\mathcal{C}(3)}(A) \\ e_{\mathcal{C}(4)}(A) \end{bmatrix} /\partial A'\right]$ respectively. The Lagrangian function corresponding to equation (1.51) is equal to

$$\mathcal{L}(A,\rho) = J_T^{(3,4)}(A, W^{(3,4)}) - 2g_T^{(2)}(A)'\rho \tag{1.54}$$

and the first order conditions are given by

$$G_T^{(3,4)}(\hat{A})'W^{(3,4)}g_T^{(3,4)}(\hat{A}) - G_T^{(2)}(\hat{A})'\hat{\rho}_T = 0 \tag{1.55}$$

$$g_T^{(2)}(\hat{A}) = 0. \tag{1.56}$$

Using $\hat{A} \xrightarrow{p} A_0$ and $\hat{\rho} \xrightarrow{p} 0$ (analogous to Newey and McFadden (1994) [p.2218]) as well as $G_T^{(2)}(\hat{A}) \xrightarrow{p} G_0^{(2)}(A_0)$ and $G_T^{(3,4)}(\hat{A}) \xrightarrow{p} G_0^{(3,4)}(A_0)$, equation (1.55) can be rewritten as

$$G_0^{(3,4)}(A_0)'W^{(3,4)}T^{1/2}g_T^{(3,4)}(\hat{A}) - G_0^{(2)}(A_0)'T^{1/2}\hat{\rho}_T \tag{1.57}$$

and equation (1.56) can be written as

$$-T^{1/2}g_T^{(2)}(\hat{A}) = 0. \tag{1.58}$$

Using the Mean Value Theorem to linearize $T^{1/2}g_T^{(2)}(\hat{A})$ and $T^{1/2}g_T^{(3,4)}(\hat{A})$ around A_0 yields

$$\begin{aligned}\begin{bmatrix}0\\0\end{bmatrix} &= \begin{bmatrix}T^{1/2}g_T^{(2)}(A_0)\\ G_0^{(3,4)}(A_0)'W^{(3,4)}T^{1/2}g_T^{(3,4)}(A_0)\end{bmatrix}\\ &+ \begin{bmatrix}G_0^{(2)}(A_0)' & 0\\ G_0^{(3,4)}(A_0)'W^{(3,4)}G_0^{(3,4)}(A_0) & -G_0^{(2)}(A_0)'\end{bmatrix}\begin{bmatrix}T^{1/2}(\hat{A}-A_0)\\ T^{1/2}(\hat{\rho})\end{bmatrix} + o_p(1)\end{aligned} \tag{1.59}$$

Define

$$H := \begin{bmatrix}G_0^{(2)}(A_0)' & 0\\ G_0^{(3,4)}(A_0)'W^{(3,4)}G_0^{(3,4)}(A_0) & -G_0^{(2)}(A_0)'\end{bmatrix}^{-1} \tag{1.60}$$

and let

$$M_{(2)} := H_{1:n^2,1:n^2+n(n-1)/2} \tag{1.61}$$

$$M_{(3,4)} := H_{1:n^2,n^2+n(n-1)/2+1:end}G_0^{(3,4)}(A_0)'W^{(3,4)}, \tag{1.62}$$

where $H_{i:j,k:l}$ denotes the matrix containing rows i to j and columns k to l of H. With equation (1.59) and the asymptotic normality of

$$\begin{bmatrix}T^{1/2}g_T^{(2)}(A_0)\\ T^{1/2}g_T^{(3,4)}(A_0)\end{bmatrix} \sim N(0, \begin{bmatrix}S_{(2),(2)} & S_{(2),(3,4)}\\ S_{(3,4),(2)} & S_{(3,4),(3,4)}\end{bmatrix}), \tag{1.63}$$

where $S_{(2),(2)} := \underset{T\to\infty}{lim} Var(T^{1/2}g_T^{(2)}(A_0))$, $S_{(2),(3,4)} := \underset{T\to\infty}{lim} Cov(T^{1/2}g_T^{(2)}(A_0), T^{1/2}g_T^{(3,4)}(A_0))$, $S_{(3,4),(2)} := \underset{T\to\infty}{lim} Cov(T^{1/2}g_T^{(3,4)}(A_0), T^{1/2}g_T^{(2)}(A_0))$ and $S_{(3,4),(3,4)} := \underset{T\to\infty}{lim} Var(T^{1/2}g_T^{(3,4)}(A_0))$ it follows that

$$T^{1/2}(\hat{A}-A_0) \sim N(0, Z), \tag{1.64}$$

with $Z = M_{(2)}S_{(2),(2)}M_{(2)}' + M_{(2)}S_{(2),(3,4)}M_{(3,4)}' + M_{(3,4)}S_{(3,4),(2)}M_{(2)}' + M_{(3,4)}S_{(3,4),(3,4)}M_{(3,4)}'$. $\square$

Let x be a an n-dimensional random variable with $E[x] = 0$ and $E[xx'] = I$. For $C = [c_1, c_2, c_3] \in$

$\{1, ..., n\}^3$ the third order (cross-)cumulant of x is equal to

$$Cum(x_C) = Cum(x_{c_1}, x_{c_2}, x_{c_3}) = E\left[x_{c_1}x_{c_2}x_{c_3}\right]. \tag{1.65}$$

For $c = [c_1, c_2, c_3, c_4] \in \{1, ..., n\}^4$ the fourth order (cross-)cumulant of x is equal to

$$\begin{aligned}
Cum(x_C) = Cum(x_{c_1}, x_{c_2}, x_{c_3}, x_{c_4}) =& E\left[x_{c_1}x_{c_2}x_{c_3}x_{c_4}\right] && (1.66)\\
&- E\left[x_{c_1}x_{c_2}\right]E\left[x_{c_3}x_{c_4}\right] && (1.67)\\
&- E\left[x_{c_1}x_{c_3}\right]E\left[x_{c_2}x_{c_4}\right] && (1.68)\\
&- E\left[x_{c_1}x_{c_4}\right]E\left[x_{c_2}x_{c_3}\right]. && (1.69)
\end{aligned}$$

Consider a sample of the random variable x for which $E_T[x] = 0$ and $E_T[xx'] = I$, then the same equalities hold for the sample counterparts $Cum_T(.)$ and $E_T(.)$.

Lemma 1.1. *Let x be an n-dimensional random variable with $E\left[x_{\mathcal{M}(1)}\right] = 0$, $E\left[x_{\mathcal{M}(2)}\right] = 1$, $E\left[x_{\mathcal{C}(2)}\right] = 0$ and $E\left[x_{\mathcal{M}(3)}\right] < \infty$. Let ϵ satisfy Assumption 1.*

1) For $c \in \mathcal{C}(3)$ it holds that $E[x_c] = Cum(x_c)$.

2) For $c \in \mathcal{C}(4)$ it holds that $E[x_c] - E[\epsilon_c] = Cum(x_c)$.

3) For $m \in \mathcal{M}(3)$ it holds that $E[x_m] = Cum(x_m)$.

4) For $m \in \mathcal{M}(4)$ it holds that $E[x_m] - 3 = Cum(x_m)$.

Consider a sample of the random variable x with $E_T\left[x_{\mathcal{M}(1)}\right] = 0$, $E_T\left[x_{\mathcal{M}(2)}\right] = 1$, $E_T\left[x_{\mathcal{C}(2)}\right] = 0$ and $E_T\left[x_{\mathcal{M}(3)}\right] < \infty$, then the same statements hold for the sample counterparts $E_T[.]$ and $Cum_T(.)$.

Proof. Statements 1) and 3) are trivial. Statement 2) holds since for $C = [c_1, c_2, c_3, c_4] \in \mathcal{C}(4)$

$$E[\epsilon_c] = \begin{cases} 1 & \text{, if } c_1 = c_2 \text{ and } c_3 = c_4 \\ 0 & \text{, else} \end{cases} \tag{1.70}$$

and

$$Cum(x_c) = \begin{cases} E\left[x_c\right] - 1 & \text{, if } c_1 = c_2 \text{ and } c_3 = c_4 \\ E\left[x_c\right] & \text{, else} \end{cases}. \tag{1.71}$$

Statement 4) holds since for $C = [c_1, c_2, c_3, c_4] \in \mathcal{M}(4)$

$$Cum(x_c) = E\left[x_c\right] - 3 \tag{1.72}$$

The sample counterpart can be shown analogously. $\square$

Proof of Proposition 1.6. For $A \in \mathcal{A}^*$ with $J_T^{(2)}(A, I) = 0$, there exists $O \in \mathcal{O}$ such that $A = OV_T^{-1}$. Therefore, Proposition 6 requires to show that there exists a constant ω_T invariant to O, such that

$$J_T^{(3,4)}(OV_T^{-1}, W_{fast}^{(3,4)}) = \omega_T - H_T^{(3,4)}(OV_T^{-1}). \tag{1.73}$$

Define

$$\tilde{e}(O) := e(OV_T^{-1}) = OV_T^{-1}u = OV_T^{-1}A_0^{-1}\epsilon = O\tilde{\epsilon}, \tag{1.74}$$

with $V_T^{-1}A_0^{-1}\epsilon = \tilde{\epsilon}$ and by construction $I = E_T\left[\tilde{\epsilon}\tilde{\epsilon}'\right]$. Then the objective function $J_T^{(3,4)}(OV_T^{-1}, W_{fast}^{(3,4)})$ can be written as

$$J_T^{(3,4)}(OV_T^{-1}, W_{fast}^{(3,4)}) = \sum_{C \in \mathcal{C}(3)} \binom{3}{\#C} E_T\left[\tilde{e}_C(O)\right]^2 + \sum_{C \in \mathcal{C}(4)} \binom{4}{\#C} \left(E_T\left[\tilde{e}_C(O)\right] - E\left[\epsilon_C\right]\right)^2 \tag{1.75}$$

and the objective function $H_T^{(3,4)}(OV_T^{-1})$ can be written as

$$H_T^{(3,4)}(OV_T^{-1}) = \sum_{M \in \mathcal{M}(3)} \binom{3}{\#M} E_T\left[\tilde{e}_M(O)\right]^2 + \sum_{M \in \mathcal{M}(4)} \binom{4}{\#M} \left(E_T\left[\tilde{e}_M(O) - 3\right]\right)^2, \tag{1.76}$$

since $\binom{r}{\#M} = 1$ for $M \in \mathcal{M}(r)$ and $r \in \{3, 4\}$. With Lemma 1.1 it follows that

$$J_T^{(3,4)}(OV_T^{-1}, W_{fast}^{(3,4)}) = \sum_{C \in \mathcal{C}(3)} \binom{3}{\#C} Cum_T(\tilde{e}_C(O))^2 + \sum_{C \in \mathcal{C}(4)} \binom{4}{\#C} Cum_T(\tilde{e}_C(O))^2 \tag{1.77}$$

and

$$H_T^{(3,4)}(OV_T^{-1}) = \sum_{M \in \mathcal{M}(3)} \binom{3}{\#M} Cum_T(\tilde{e}_M(O))^2 + \sum_{M \in \mathcal{M}(4)} \binom{4}{\#M} Cum_T(\tilde{e}_M(O))^2. \tag{1.78}$$

The weights are constructed such that

$$\begin{aligned} J_T^{(3,4)}(OV_T^{-1}, W_{fast}^{(3,4)}) + H_T^{(3,4)}(OV_T^{-1}) = &\sum_{s_1,...,s_3=1}^{n} Cum_T(\tilde{e}_{s_1}(O), ..., \tilde{e}_{s_3}(O))^2 \\ &+ \sum_{s_1,...,s_4=1}^{n} Cum_T(\tilde{e}_{s_1}(O), ..., \tilde{e}_{s_4}(O))^2. \end{aligned} \tag{1.79}$$

Equation (3.10) in Comon (1994) states that on the population level, the right and side of the equation (1.79) is invariant with respect to $O \in \mathcal{O}$. However, the same statement also holds for the sample counterpart and thus the right hand side of the equation (1.79) is invariant with respect to $O \in \mathcal{O}$. Let $\omega_T := \sum_{s_1,...,s_3=1}^{n} Cum_T(\tilde{e}_{s_1}(O), ..., \tilde{e}_{s_3}(O))^2 + \sum_{s_1,...,s_4=1}^{n} Cum_T(\tilde{e}_{s_1}(O), ..., \tilde{e}_{s_4}(O))^2$ and thus

$$J_T^{(3,4)}(OV_T^{-1}, W_{fast}^{(3,4)}) = \omega_T - H_T^{(3,4)}(OV_T^{-1}). \tag{1.80}$$

Therefore equation (1.73) holds, which proves the proposition. □

1.10.2 Appendix - Notes on identification

This section uses a bivariate example to illustrate why global identification up to sign and permutation requires to include all coskewness conditions. Analogously, one can construct an example based on cokurtosis conditions and show that including only $n(n-1)/2$ cokurtosis conditions (as proposed by Lanne and Luoto (2021)) does not globally identify the SVAR.

Let $E[\epsilon_i] = 0$, $E\left[\epsilon_i^2\right] = 1$ and $E\left[\epsilon_i^3\right] = 1$ for $i \in \{1, 2\}$ and let

$$\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} = A_0 \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \tag{1.81}$$

with $A_0 = I$. Define the unmixed innovations as

$$\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} 1 & a_{1,2} \\ a_{2,1} & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \tag{1.82}$$

To simplify calculations, the moment condition $E\left[e^2\right] = 1$ has been replaced by the assumption $a_{1,1} = a_{2,2} = 1$. The covariance and coskewness conditions are given by

$$0 = E[e_1 e_2] \tag{1.83}$$

$$0 = E\left[e_1^2 e_2\right] \tag{1.84}$$

$$0 = E\left[e_1 e_2^2\right]. \tag{1.85}$$

The system thus contains three equations in two unknowns (e_1 and e_2 are functions of $a_{1,2}$ and $a_{2,1}$). Omitting one of the coskewness conditions leads to a system not identifying A_0 up to sign and permutations. Consider the system of two equations and two unknowns

$$0 = E[e_1 e_2] \tag{1.86}$$

$$0 = E\left[e_1^2 e_2\right]. \tag{1.87}$$

The first equation yields

$$0 = E[e_1 e_2] = E[(\epsilon_1 + a_{1,2}\epsilon_2)(a_{2,1}\epsilon_1 + \epsilon_2)] \tag{1.88}$$

$$\implies 0 = a_{2,1} + a_{1,2}. \tag{1.89}$$

The second equation yields

$$0 = E\left[e_1^2 e_2\right] = E\left[(\epsilon_1 + a_{1,2}\epsilon_2)^2(a_{2,1}\epsilon_1 + \epsilon_2)\right] \tag{1.90}$$

$$\implies 0 = a_{2,1} + a_{1,2}^2. \tag{1.91}$$

Therefore, $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ solves the system and thus $A_0 = I$ is not identified up to sign and permutations.

1.10.3 Appendix - Notes on PML

Also the PML estimator proposed by Gouriéroux et al. (2017) is closely related to the maximization of certain moments. To see this, consider the PML estimator with the pseudo distribution of the i-th shock being equal to $g_i \sim t(v)$, where $t(v)$ denotes a t-distribution with v degrees of freedom. The PML estimator is given by

$$\hat{O}^{PML} = \arg\max_{O \in \mathcal{O}} \sum_{t=1}^{T} \sum_{i=1}^{n} log\ g_i(e_{t,i}) \tag{1.92}$$

$$= \arg\max_{O \in \mathcal{O}} \sum_{t=1}^{T} \sum_{i=1}^{n} -\frac{1-v}{2} log\left(1 + \frac{e_{t,i}^2}{v-2}\right) \tag{1.93}$$

Ignoring the weighting implied by the degrees of freedom and using $log(1+x) \approx x - \frac{x^2}{2}$ yields

$$\hat{O}^{PML} \approx \arg\max_{O \in \mathcal{O}} \sum_{t=1}^{T} \sum_{i=1}^{n} -e_{t,i}^2 + \frac{e_{t,i}^4}{2} \tag{1.94}$$

$$= \arg\max_{O \in \mathcal{O}} T \sum_{i=1}^{n} \left(-T^{-1} \sum_{t=1}^{T} e_{t,i}^2\right) + T \sum_{i=1}^{n} \left(\frac{1}{2} T^{-1} \sum_{t=1}^{T} e_{t,i}^4\right) \tag{1.95}$$

The shocks are normalized and thus $T^{-1}\sum_{t=1}^{T} e_{t,i}^2 = 1$ for $i = 1, ..., n$. It follows that the PML estimator can be written as

$$\hat{O}^{PML} \approx \arg\max_{O\in\mathcal{O}} \sum_{i=1}^{n}\left(T^{-1}\sum_{t=1}^{T} e_{t,i}^4\right) \tag{1.96}$$

$$= \arg\max_{O\in\mathcal{O}} \sum_{M\in\mathcal{M}(4)} E_T\left[e_M(O)\right] \tag{1.97}$$

Therefore, maximizing the pseudo log likelihood function is approximately equal to maximizing the kurtosis of the unmixed innovations. Moreover, Equation (1.97) shows that the PML estimator based on a t-distribution cannot utilize any information contained in third moments.

1.10.4 Appendix - Finite sample performance

This section contains the results of the first Monte Carlo study presented in Section 7, when the estimators are calculated with an optimization routine using a grid of starting values. The simulations show, that using a grid of starting values has almost no impact on the globally identified SVAR-GMM estimators proposed in this paper, but obviously has a severe impact on the locally identified GMM estimator proposed by Lanne and Luoto (2021).
The Monte Carlo study again uses $u = A_0^{-1}\epsilon$ with

$$A_0^{-1} = \begin{bmatrix} cos(\phi) & -sin(\phi) \\ sin(\phi) & cos(\phi) \end{bmatrix}, \tag{1.98}$$

and $\phi = -\pi/5$. The shocks are drawn from a distribution of the Pearson distribution family with mean zero, unit variance, and different skewness/kurtosis parameters. In the first specification, both structural shocks have zero skewness and a high excess kurtosis. In the second specification, both shocks have a high skewness and zero excess kurtosis. In the third specification, the first shock is Gaussian, and the second shock has a high skewness and high excess kurtosis. In Section 7 all estimators where calculated with local optimization routines using starting values close to the true parameter Matrix A_0. In this section, the optimization routine calculates several minima based on different starting values. The estimator is then equal to the solution leading to the smallest minima. The starting values are equal to

$$A_{start}(\phi_{start}) = \begin{bmatrix} cos(\phi_{start}) & -sin(\phi_{start}) \\ sin(\phi_{start}) & cos(\phi_{start}) \end{bmatrix}, \tag{1.99}$$

with $\phi_{start} \in \{2, 1.8, 1.6, 1.4, 1.2, 1, 0.8, 0.6, 0.4, 0.2, 0\}$. The simulation results are shown in Table 1.5.

Table 1.5: Finite sample performance - Comparison of unrestricted non-Gaussian SVAR estimators (global optimization).

	$T = 200$			$T = 500$			$T = 5000$		
Specification	1	2	3	1	2	3	1	2	3
$\hat{A}(W^{2\text{-step}})$	-0.02 (6.51)	-0.01 (1.51)	-0.03 (6.08)	-0.01 (5.11)	0 (0.96)	-0.01 (5.58)	0 (3.7)	0 (0.83)	0 (3.91)
$\hat{A}(W^{fast})$	-0.02 (4.49)	-0.02 (3.32)	-0.03 (5.53)	-0.01 (3.91)	0 (2.61)	-0.01 (4.99)	0 (3.48)	0 (1.22)	0 (4.52)
GMM_{LL}	-0.08 (15.38)	-0.06 (14.05)	-0.03 (10.3)	-0.09 (41.09)	-0.06 (35.39)	-0.01 (18.15)	-0.19 (582.78)	-0.06 (384.64)	0.01 (70.28)
PML	-0.01 (3.11)	-0.1 (16.12)	-0.04 (6.06)	0 (2.57)	-0.1 (43.27)	-0.02 (7.87)	0 (2.15)	-0.15 (592.72)	0 (4.94)

Monte Carlo simulation with sample sizes 200, 500, and 5000 each with 10000 iterations. For an estimator $\hat{A}$ of A_0, define the estimator $\hat{B} := \hat{A}^{-1}$ of $B := A_0^{-1}$. Each entry shows the mean bias, $E\left[\hat{b}_{1,1} - b_{1,1}\right]$, and the standard deviation of $\sqrt{T}\left(\hat{b}_{1,1} - b_{1,1}\right)$ is shown in parentheses. Moreover, the element $b_{1,1}$ of B is equal to $cos(-\pi/5)$. In specification one, both structural shocks have zero skewness and a high excess kurtosis. In the specification two, both structural shocks have a high skewness and zero excess kurtosis. In specification three, the first shock is Gaussian and the second shock has a high skewness and high excess kurtosis. The SVAR-GMM estimator denoted by $\hat{A}(W^{2\text{-}step})$ is the two-step SVAR-GMM estimator and the fast SVAR-GMM estimator is denoted by $\hat{A}(W^{fast})$. The GMM estimator proposed by Lanne and Luoto (2021) is denoted by GMM_{LL} and uses the cokurtosis condition $E\left[\epsilon_1\epsilon_2^3\right] = 0$. The PML estimator proposed by Gouriéroux et al. (2017) is denoted by PML and assumes a t-distribution with twelve degrees of freedom. Each estimator is calculated as the minima of all local minima found by a local optimization routine based on the grid of starting values $\phi_{start} \in \{2, 1.8, 1.6, 1.4, 1.2, 1, 0.8, 0.6, 0.4, 0.2, 0\}$.

1.10.5 Appendix - Application: Data and robustness checks

U.S. Industrial Production Index

Source: Board of Governors of the Federal Reserve System (US)

Retrieved from: FRED, Federal Reserve Bank of St. Louis

Link: `https://fred.stlouisfed.org/series/INDPRO`, August 25, 2018

Crude oil composite acquisition cost by refiner

Source: U.S. Energy Information Administration

Link: https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=R0000____3&f=M, August 12, 2018

S&P 500

Source: Yahoo! Finance

Link: https://finance.yahoo.com/quote/%5EGSPC?p=%5EGSPC, August 12, 2018

U.S. CPI

Source: U.S. Bureau of Labor Statistics

Retrieved from: FRED, Federal Reserve Bank of St. Louis

Link: https://fred.stlouisfed.org/series/CPIAUCSL, August 12, 2018

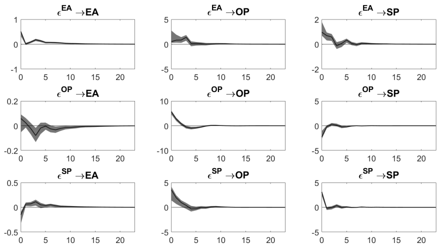
Table 1.6: Economic activity, oil, and stock prices - Descriptive statistics.

	Mean	Median	Mode	Std. deviation	Variance	Skewness	Kurtosis	Range
EA	0.13	0.15	0.41	0.58	0.34	-1.97	15.11	6.15
OP	0.03	0.14	-10.98	1.89	3.56	-1.35	9.42	16.48
SP	0.44	0.82	-17.7	4.12	17	-0.75	4.66	27.99

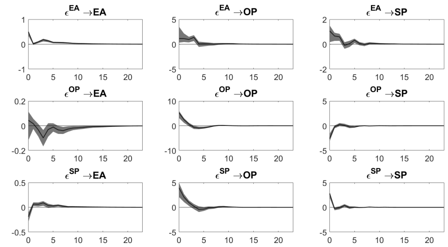
Real economic activity is measured as 100 times the log difference of the monthly US Industrial Production Index. Real S&P 500 returns are calculated as 100 times the log difference of the the S&P 500 closing price deflated by the US CPI. The monthly growth rates of real oil prices are calculated as 100 times the log difference of the crude oil composite acquisition cost by refiners deflated by the US CPI.

Figure 1.2: Economic activity, oil, and stock prices - Impulse responses (robustness checks).

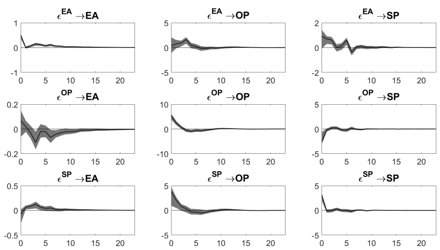
1) Fast SVAR-GMM



2) PML



3) SVAR-GMM (2-step) with six lags



The first and second set of IRFs are based on the reduced form described in Section 8. The first set of IRFs is based on the fast SVAR-GMM estimator and the second set of IRFs is based on the PML estimator proposed by Gouriéroux et al. (2017) and uses a t-distribution with twelve degrees of freedom. The third set of IRFs estimates the reduced form with six lags and the simultaneous relationship is estimated by the two-step SVAR-GMM estimator. Confidence intervals are calculated by bootstrap with 1000 *replications and the interval shows the upper* 0.9 *and lower* 0.1 *percentiles.*

2 A Feasible Approach to Incorporate Information in Higher Moments in Structural Vector Autoregressions[6]

2.1 Introduction

In a non-Gaussian structural vector autoregression (SVAR) independent structural shocks imply higher-order moment conditions which identify the simultaneous relationship without any restrictions on the simultaneous interaction. These higher-order moment conditions can be used to estimate the SVAR with a generalized of moments (GMM) or continuous updating estimator (CUE), see, e.g., Lanne and Luoto (2021), Keweloh (2021b), or Guay (2021). However, with higher-order moment conditions the long-run covariance matrix of the sample average of the moment conditions is difficult to estimate in small samples. Nevertheless, an accurate estimation of the covariance matrix is crucial for the estimation of the asymptotically optimal weighting matrix, the estimation of the asymptotic variance, and inference.

This study analyzes the small sample behavior of CUE and GMM estimators with higher-order moment conditions in SVAR models. I find that standard approaches to estimate the long-run covariance matrix lead to volatile and biased CUE and GMM estimators with distorted J and Wald test statistics. Moreover, the performance of the estimators decreases with the model size, to the point of limiting the usefulness of the approach for specifications usually considered in macroeconometrics. I propose to use the assumption of mutually independent structural shocks not only to derive moment conditions but also to estimate the asymptotically efficient weighting matrix and the asymptotic variance. I demonstrate that this simple modification substantially increases the small sample performance of the estimators.

The small sample behavior of CUE and GMM estimators in general has been studied extensively. GMM estimators are known to exhibit a small sample bias and the CUE is associated with a smaller bias, see, e.g., Hansen et al. (1996), Donald and Newey (2000), Han and Phillips (2006), or Newey and Windmeijer (2009). Moreover, the inability to precisely estimate the asymptotic variance leads to oversized Wald test statistics, see Burnside and Eichenbaum (1996). Therefore,

[6] A slightly different version of the chapter appeared as Keweloh, S. A. (2021), *A feasible approach to incorporate information in higher moments in structural vectorautoregressions*, SFB 823 Discussion Paper series No. 22/2021, https://doi.org/10.17877/DE290R-22416

Burnside and Eichenbaum (1996) propose to use restrictions implied by the underlying model to calculate test statistics. In the context of SVAR models, Bonhomme and Robin (2009), Keweloh (2021b), and Guay (2021) recognize that due to higher-order moment conditions, the long-run covariance matrix of the sample average of the moment conditions is particularly difficult to estimate. For example, the covariance of cokurtosis moment conditions is of order eight and therefore, difficult to estimate in samples with a few hundred observations. I show that exploiting the assumption of mutually independent structural shocks simplifies the problem of estimating the covariance of higher-order moment conditions. In particular, with independent structural shocks, higher-order moments of the covariance matrix can be calculated as products of lower-order moments. Therefore, I propose the SVAR CUE-MI and SVAR GMM-MI estimators, which are SVAR CUE and SVAR GMM estimators exploiting the assumption of mutually independent structural shocks to estimate the asymptotically optimal weighting matrix and the asymptotic variance. A Monte Carlo simulation demonstrates that the SVAR CUE-MI and SVAR GMM-MI outperform SVAR CUE and SVAR GMM estimators, which are not exploiting the assumption of mutually independent shocks to estimate the optimal weighting and asymptotic variance.

It is well known that the number coskewness and cokurtosis conditions implied by independent structural shocks increases quickly with the dimension of the SVAR. For example, with $n = 2$ variables, independent structural shocks imply two variance, one covariance, two coskewness and three cokurtosis conditions, and with $n = 4$ variables independent structural shocks imply four variance, six covariance, 16 coskewness and 31 cokurtosis conditions, see Keweloh (2021b).[7] While the possible number of moment conditions increases quickly with the dimension of the SVAR, the number of moments contained in the covariance matrix of the moment conditions increases even more rapidly. In particular, for $n = 2$ variables the covariance matrix of all second- to fourth-order moment conditions implied by mutually independent shocks is a 8×8 matrix with five co-moments of order four, six co-moments of order five, seven co-moments of order six, eight co-moments of order seven, and nine co-moments of order eight. However, for $n = 4$ variables the covariance matrix of all second- to fourth-order moment conditions implied

[7] Lanne and Luoto (2021) and Keweloh (2021b) propose GMM estimators which minimize the second- and higher-order dependencies of the unmixed innovations. In contrast to that, the GMM estimator proposed by Guay (2021) minimizes the distance of the second- and higher-order co-moments of the reduced form shocks to the second- and higher-order co-moments implied by a mixture of independent structural shocks. The GMM estimator proposed by Guay (2021) has even more higher-order moment conditions.

by mutually independent shocks is a 57×57 matrix with 35 co-moments of order four, 56 co-moments of order five, 84 co-moments of order six, 120 co-moments of order seven, and 165 co-moments of order eight. Exploiting the assumption of mutually independent shocks allows to estimate the covariance matrix of all second- to fourth-order moment conditions implied by mutually independent shocks as a product of n moments of order one, n moments of order two, n moments of order three, n moments of order four, n moments of order five, and n moments of order six. Therefore, with mutually independent shocks the researcher only needs to estimate moments up to order six instead of order eight and the number of these moments increases linearly in the dimension of the SVAR.

Mutually independent structural shocks simplify the estimation of the asymptotically optimal weighting matrix and the asymptotic variance. In many cases, mutually independent structural shocks are no additional assumption but assumed anyway to derive the identifying higher-order moment conditions, see, e.g., Keweloh (2021b) and Guay (2021).[8] However, some authors argue that the assumption of mutually independent shocks is too strong, see, e.g., Kilian and Lütkepohl (2017, Chapter 14), Lewis (2021), or Lanne and Luoto (2021). In particular, independence implies that the volatility processes of the shocks are independent. Lanne and Luoto (2021) show that a suitable subset of $n(n-1)/2$ asymmetric cokurtosis conditions is sufficient to ensure local identification in a non-Gaussian SVAR.[9] These asymmetric cokurtosis conditions can be motivated by mutually mean independent shocks, which allows a dependence of the volatility processes. Therefore, I derive analogous results for the estimation of the weighting matrix and variance depending on mutually mean independent shocks. More generally, the approach proposed in this study does not rely on a specific set of moment conditions. Instead, I argue that the same statistical properties used to derive the moment conditions should be used to estimate the asymptotically optimal weighting matrix and the asymptotic variance of the estimator.

The remainder of this article is organized as follows. Section 2.2 summarizes the SVAR model and the main assumptions. Section 2.3 defines the GMM estimator for SVAR models based on

[8]Note that also the (pseudo) maximum-likelihood estimators proposed by Lanne et al. (2017) and Gouriéroux et al. (2017) or the Bayesian approaches proposed by Lanne and Luoto (2020) and Anttonen et al. (2021) assume independent shocks to ensure identification.

[9]Additionally, Lanne et al. (2021) show that the second-order moment conditions together with $n(n-1)/2$ symmetric cokurtosis conditions can be sufficient to ensure global identification. However, the result only holds if the structural shocks satisfy all cokurtosis moment conditions implied by independent shocks.

higher-order moment conditions. Section 2.4 proposes novel estimators for the long-run covariance matrix and the asymptotic variance by exploiting mutually independent shocks. Section 2.5 demonstrates the advantages of the proposed estimators over traditional estimators in a Monte Carlo simulation. Section 2.6 concludes.

2.2 SVAR models

This section briefly explains the identification problem and common identification approaches of SVAR models. A detailed overview can be found in Kilian and Lütkepohl (2017). Consider the SVAR $y_t = \sum_{p=1}^{P} A_p y_{t-p} + u_t$ with an n-dimensional vector of observable variables $y_t = [y_{1,t}, ..., y_{n,t}]'$, the reduced form shocks $u_t = [u_{1,t}, ..., u_{n,t}]'$, and

$$u_t = B_0 \epsilon_t \tag{2.1}$$

describing the impact of an n-dimensional vector of unknown structural shocks $\epsilon_t = [\epsilon_{1,t}, ..., \epsilon_{n,t}]'$. The matrix $B_0 \in \mathbb{R}^{n \times n}$ governs the simultaneous interaction and is assumed to be invertible.

Assumption 2.1. *$B_0 \in \mathbb{B} := \{B \in \mathbb{R}^{n \times n} | det(B) \neq 0\}$.*

The reduced form shocks can be estimated consistently, and for the sake of simplicity, I focus on the simultaneous interaction in Equation (2.1) and treat the reduced form shocks as observable random variables. The identically distributed structural shocks satisfy the following assumptions.

Assumption 2.2. *ϵ_t is serially independent (ϵ_t is independent of $\epsilon_{\tilde{t}}$ for $t \neq \tilde{t}$)*

Assumption 2.3. *ϵ_t has mutually uncorrelated components ($\epsilon_{i,t}$ is uncorrelated with $\epsilon_{j,t}$ for $i \neq j$).*

Assumption 2.4. *Each component of ϵ_t has zero mean, unit variance, and finite third- and fourth-order moments.*

Based on the assumptions used so far, neither the matrix B_0 nor the structural shocks ϵ_t are identified. Several identifying assumptions have been proposed in the literature ranging from short- or long-run restrictions on the interaction of the variables (see, e.g., Sims (1980), or Blanchard (1989)), over Proxy-SVAR models (see, e.g., Mertens and Ravn (2013)) up to sign restrictions (see, e.g., Uhlig (2005) or Peersman (2005)). A novel branch of the SVAR identification literature uses non-Gaussian and independent shocks to identify the SVAR, see, e.g., Lanne et al. (2017), Gouriéroux et al. (2017), Herwartz (2018), Lanne and Luoto (2021), Keweloh (2021b), or Guay (2021)). These data-driven identification schemes do not require to impose any short- or long-run restrictions on the interaction of the variables, instead, identification is based on statistical properties of the shocks. The most commonly used statistical property is the assumption of mutually independent structural shocks.

Assumption 2.5. *ϵ_t has mutually independent components ($\epsilon_{i,t}$ is independent of $\epsilon_{j,t}$ for $i \neq j$).*

The independence assumption can be used to derive moment conditions, see, e.g., Keweloh (2021b) or Guay (2021). However, the assumption of mutually independent shocks has been criticized by several authors for being too restrictive, see, e.g., Kilian and Lütkepohl (2017, Chapter 14), Lewis (2021), or Lanne and Luoto (2021). In particular, it appears plausible that multiple macroeconomic shocks are driven by the same volatility process, which is not possible with mutually independent shocks. The independence assumption can be relaxed to the weaker assumption of mutually mean independent shocks.

Assumption 2.6. *ϵ_t has mutually mean independent components ($E\left[\epsilon_{i,t}|\epsilon_{-i,t}\right] = 0$ for $i = 1, ..., n$).*

If the shocks are mutually mean independent, no shock contains any information on the mean of another shock, however, shocks can contain information on the variance of other shocks and thus, the assumption may be more plausible in some applications. Note that mutually mean independent structural shocks are sufficient to derive the identifying coskewness conditions proposed in Keweloh (2021b) and the identifying asymmetric cokurtosis conditions used in Lanne and Luoto (2021). For the sake of simplicity, this study focuses on mutually independent shocks. Results under the weaker assumption of mutually mean independent shocks can be obtained analogously and are briefly sketched.

The assumption of mutually (mean) independent shocks can be used to generate higher-order moment condition, however, these conditions are only informative if the structural shocks are non-Gaussian. Therefore, identification requires non-Gaussian shocks embedded in the following assumption.

Assumption 2.7. *At most one component of ϵ_t is Gaussian.*

Note that depending on the particular identification approach a slight modification of Assumption 2.7 is required. For example, the GMM estimators proposed by Lanne and Luoto (2021), Keweloh (2021b) or Guay (2021) require that the third- and/or fourth-order moments of at most one structural shock is equal to the corresponding moment of a Gaussian shock.

2.3 SVAR GMM with higher-order moment conditions

This section briefly summarizes the SVAR estimators based on higher-order moment conditions derived from mutually independent structural shocks. A detailed description can be found in Lanne and Luoto (2021), Keweloh (2021b), or Guay (2021).

The reduced form shocks are equal to an unknown mixture of the unknown structural shocks, $u_t = B\epsilon_t$. Reversing this relationship yields the unmixed innovations $e(B)_t$, defined as the innovations obtained by unmixing the reduced form shocks with some invertible matrix B

$$e(B)_t := B^{-1}u_t. \tag{2.2}$$

If B is equal to the true mixing matrix B_0, the unmixed innovations are equal to the structural shocks. Assumption 2.4 and 2.5 can be used to derive moment conditions. In particular, the structural shocks are uncorrelated with unit variance. Therefore, the unmixing matrix B should yield uncorrelated unmixed innovations with unit variance, see Table 2.1. Moreover, independent structural shocks yield coskewness or third-order moment conditions and cokurtosis or fourth-order moment conditions, see Table 2.1.

In general, all variance, covariance, coskewness, and cokurtosis moment conditions derived from independent structural shocks embedded in Assumption 2.5 can be written as

$$E[f_m(B, u_t)] = 0 \quad \text{with} \quad f_m(B, u_t) := \prod_{i=1}^{n} e(B)_{i,t}^{m_i} - m_0, \tag{2.3}$$

where $f_m(B, u_t)$ contains all variance and covariance conditions for $m \in \mathbf{2}$, all coskewness condi-

Table 2.1: Illustration of moment conditions.

covariance / second-order conditions		
$E[\epsilon_{1,t}^2] = 1$	$\Longrightarrow$	$E[e(B)_{1,t}^2] \overset{!}{=} 1$
	$\vdots$	
$E[\epsilon_{n,t}^2] = 1$	$\Longrightarrow$	$E[e(B)_{n,t}^2] \overset{!}{=} 1$
$E[\epsilon_{1,t}\epsilon_{2,t}] = 0$	$\Longrightarrow$	$E[e(B)_{1,t}e(B)_{2,t}] \overset{!}{=} 0$
	$\vdots$	

coskewness / third-order conditions		
$E[\epsilon_{1,t}^2\epsilon_{2,t}] = 0$	$\Longrightarrow$	$E[e(B)_{1,t}^2 e(B)_{2,t}] \overset{!}{=} 0$
$E[\epsilon_{1,t}\epsilon_{2,t}^2] = 0$	$\Longrightarrow$	$E[e(B)_{1,t} e(B)_{2,t}^2] \overset{!}{=} 0$
$E[\epsilon_{1,t}\epsilon_{2,t}\epsilon_{3,t}] = 0$	$\Longrightarrow$	$E[e(B)_{1,t}e(B)_{2,t}e(B)_{3,t}] \overset{!}{=} 0$
	$\vdots$	

cokurtosis / fourth-order conditions		
$E[\epsilon_{1,t}^3\epsilon_{2,t}] = 0$	$\Longrightarrow$	$E[e(B)_{1,t}^3 e(B)_{2,t}] \overset{!}{=} 0$
$E[\epsilon_{1,t}^2\epsilon_{2,t}^2] = 1$	$\Longrightarrow$	$E[e(B)_{1,t}^2 e(B)_{2,t}^2] \overset{!}{=} 1$
$E[\epsilon_{1,t}\epsilon_{2,t}^3] = 0$	$\Longrightarrow$	$E[e(B)_{1,t} e(B)_{2,t}^3] \overset{!}{=} 0$
$E[\epsilon_{1,t}^2\epsilon_{2,t}\epsilon_{3,t}] = 0$	$\Longrightarrow$	$E[e(B)_{1,t}^2 e(B)_{2,t}e(B)_{3,t}] \overset{!}{=} 0$
$E[\epsilon_{1,t}\epsilon_{2,t}\epsilon_{3,t}\epsilon_{4,t}] = 0$	$\Longrightarrow$	$E[e(B)_{1,t}^2 e(B)_{2,t}e(B)_{3,t}e(B)_{4,t}] \overset{!}{=} 0$
	$\vdots$	

tions for $m \in \mathbf{3}$, and all cokurtosis conditions for $m \in \mathbf{4}$ with

$$\mathbf{2} := \{[m_0, m_1, ...m_n] \in \{\{0,1\},\{0,1,2\}^n\} | \sum_{i=1}^{n} m_i = 2 \text{ and} \tag{2.4}$$

$$m_0 = \begin{cases} 0, & \text{if } \exists m_i = 1 \text{ for } m_i \in m_1, ...m_n \\ 1, & \text{otherwise} \end{cases} \},$$

$$\mathbf{3} := \{[m_0, m_1, ...m_n] \in \{\{0,1\},\{0,1,2\}^n\} | \sum_{i=1}^{n} m_i = 3 \text{ and } m_0 = 0\}, \tag{2.5}$$

$$\mathbf{4} := \{[m_0, m_1, ...m_n] \in \{\{0,1\},\{0,1,2,3\}^n\} | \sum_{i=1}^{n} m_i = 4 \text{ and} \tag{2.6}$$

$$m_0 = \begin{cases} 0, & \text{if } \exists m_i = 1 \text{ for } m_i \in m_1, ...m_n \\ 1, & \text{otherwise} \end{cases} \}.$$

Note that all moment conditions except the symmetric cokurtosis conditions $E[\epsilon_{i,t}^2\epsilon_{j,t}^2] = 1$ for $i \neq j$ can be derived from Assumption 2.6 and therefore, only require mutually mean independent shocks.

Based on all or a subset of the moment conditions presented above, Lanne and Luoto (2021) and Keweloh (2021b) provide different local and global identification results for the following SVAR GMM estimator

$$\hat{B}_T := \underset{B \in \mathbb{B}}{\arg\min}\, g_T(B)' W g_T(B), \tag{2.7}$$

where $g_T(B) = \frac{1}{T}\sum_{t=1}^{T} f(B, u_t)$, and W is a positive semi-definite weighting matrix. Suppose that $f(B, u_t)$ contains all or a subset of the moment conditions $f_m(B, u_t)$ with $m \in \mathbf{2} \cup \mathbf{3} \cup \mathbf{4}$ such that the SVAR GMM estimator (2.7) is identified. Consistency and asymptotic normality of the estimator follow from standard assumptions

$$\begin{aligned} \hat{B}_T &\xrightarrow{p} B_0 \\ \sqrt{T}(\hat{B}_T - B_0) &\xrightarrow{d} \mathcal{N}(0, MSM') \end{aligned} \qquad \text{with} \qquad \begin{aligned} M &:= \left(G'S^{-1}G\right)^{-1} G'W \\ G &:= E\left[\frac{\partial f(B_0, u_t)}{\partial vec(B)'}\right] \\ S &:= \underset{T\to\infty}{lim} E\left[T g_T(B_0) g_T(B_0)'\right], \end{aligned} \tag{2.8}$$

see Hall (2005). In particular, asymptotic normality requires that the matrix S exists and is finite. For the SVAR GMM estimator based on second- to fourth-order moment conditions this holds if ϵ_t has finite moments up to order eight. The weighting matrix $W^* := S^{-1}$ leads to the estimator $\hat{B}_T^*$ with the asymptotic variance $\sqrt{T}(\hat{B}_T^* - B_0) \xrightarrow{d} \mathcal{N}(0, (G'S^{-1}G)^{-1})$, which is the lowest possible asymptotic variance, see Hall (2005). Han and Phillips (2006) proposed the continuous updating estimator estimator (CUE)

$$\hat{B}_T := \underset{B \in \mathbb{B}}{\arg\min}\, g_T(B)' \hat{W}(B) g_T(B), \tag{2.9}$$

where $\hat{W}(B)$ is a consistent estimator for the asymptotically optimal weighting matrix W^*.

Han and Phillips (2006) and Newey and Windmeijer (2009) show that for i.i.d. observations and a nonrandom weighting matrix W the expected value of a GMM objective function is equal to

$$\begin{aligned} E\left[g_T(B)' W g_T(B)\right] =& E\left[\sum_{t \neq \tilde{t}} f(B, u_t)' W f(B, u_{\tilde{t}})\right] + E\left[\sum_{t} f(B, u_t)' W f(B, u_t)\right] && (2.10) \\ =& (1 - T^{-1}) E[f(B, u_t)]' W E[f(B, u_t)] + trace(WS(B))/T, && (2.11) \end{aligned}$$

where the second equality uses the nonrandom weighting matrix W and $S(B) := E[f(B, u_t)f(B, u_t)']$. The first term in Equation (2.11) is called signal term and is minimized at B_0 since $E[f(B_0, u_t)] = 0$. The second term in Equation (2.11) is called noise term and is not minimized at B_0. The impact of the noise term vanishes with $T \to \infty$. Nevertheless, in a finite sample the noise term can dominate the signal term and lead to a bias, especially in large SVAR models with many moment conditions. If the weighting matrix $W(B)$ is equal to $S(B)^{-1}$, the noise term in Equation (2.11) collapses to m/T, where m is equal to the number of moment conditions. Therefore, the noise term no longer depends on B and hence leads to no bias. The CUE is a feasible version of this approach and replaces $W(B) = S(B)^{-1}$ with some estimator $\hat{W}(B) = \hat{S}(B)^{-1}$.

2.4 Estimating S and G

In practice, S and G are unknown and need to be estimated for inference and asymptotically optimal weighting. These matrices can be difficult to estimate in small samples. In a GMM setup not related to SVAR models Burnside and Eichenbaum (1996) propose to impose restrictions of the underlying economic model on the estimator for S and G. They show that exploiting additional information on S and G can improve the rejection rates of Wald tests in small samples. This section shows how the structure of the SVAR can be used to improve the estimation of S and G. In particular, I propose to exploit the assumption of serially and mutually independent structural shocks. In the SVAR with second- to fourth-order moment conditions, estimation of the long-run covariance matrix S is particularly difficult, since it requires to estimate moments up to order eight. Bonhomme and Robin (2009), Keweloh (2021b), and Guay (2021) recognize that the presence of these higher-order moments makes it difficult to estimate the asymptotically optimal weighting matrix and the asymptotic variance of the estimator. I show that exploiting the assumption of serially and mutually independent shocks largely simplifies the estimation of the long-run covariance matrix S and yields more precise estimates of the asymptotically optimal weighting matrix and the asymptotic variance. Additionally, for the first weighting step of the SVAR GMM estimator, I propose an approximation of the asymptotically optimal weighting matrix based on the assumption of mutually independent shocks not requiring any prior estimates of B_0.

In the SVAR, the long-run covariance matrix of two arbitrary moment conditions $f_m(B, u_t) = \prod_{i=1}^{n} e(B)_{i,t}^{m_i} - m_0$ and $f_{\tilde{m}}(B, u_t) = \prod_{i=1}^{n} e(B)_{i,t}^{\tilde{m}_i} - \tilde{m}_0$ with $m, \tilde{m} \in \mathbf{2} \cup \mathbf{3} \cup \mathbf{4}$ at $B = B_0$ is equal to

$$
\begin{aligned}
S_{m,\tilde{m}} :=& \lim_{T\to\infty} E\left[T\left(\frac{1}{T}\sum_{t=1}^{T} f_m(B_0, u_t)\right)\left(\frac{1}{T}\sum_{t=1}^{T} f_{\tilde{m}}(B_0, u_t)\right)\right] && (2.12)\\
=& \lim_{T\to\infty} E\left[T\left(\frac{1}{T}\sum_{t=1}^{T}\prod_{i=1}^{n} e(B_0)_{i,t}^{m_i} - m_0\right)\left(\frac{1}{T}\sum_{t=1}^{T}\prod_{i=1}^{n} e(B_0)_{i,t}^{\tilde{m}_i} - \tilde{m}_0\right)\right] && (2.13)\\
=& E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{m_i+\tilde{m}_i}\right] - m_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{\tilde{m}_i}\right] - \tilde{m}_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{m_i}\right] + m_0\tilde{m}_0 && (2.14)\\
&+\sum_{j=1}^{\infty} E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{m_i}\epsilon_{i,t-j}^{\tilde{m}_i}\right] - m_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t-j}^{\tilde{m}_i}\right] - \tilde{m}_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{m_i}\right] + m_0\tilde{m}_0\\
&+\sum_{j=1}^{\infty} E\left[\prod_{i=1}^{n}\epsilon_{i,t-j}^{m_i}\epsilon_{i,t}^{\tilde{m}_i}\right] - m_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t}^{\tilde{m}_i}\right] - \tilde{m}_0 E\left[\prod_{i=1}^{n}\epsilon_{i,t-j}^{m_i}\right] + m_0\tilde{m}_0,
\end{aligned}
$$

where the last equality follows from identically distributed shocks and $e(B_0)_t = \epsilon_t$. Therefore, with fourth order moments $m, \tilde{m} \in \mathbf{4}$ such that $\sum_{i=1}^{n} m_i = \sum_{i=1}^{n} \tilde{m}_i = 4$, the long-run covariance matrix $S_{m,\tilde{m}}$ contains co-moments of the structural shocks up to order eight. In practice, $S_{m,\tilde{m}}$ in Equation (2.14) can be estimated by replacing ϵ_t with $e(B)_t$ and some initial estimate or guess B of B_0 and a heteroscedasticity and autocorrelation consistent covariance (HAC) estimator, see Newey and West (1994).

However, with serially independent structural shocks implied by Assumption 2.2 the expression of $S_{m,\tilde{m}}$ simplifies to[10]

$$S^{SI}_{m,\tilde{m}} = E\left[\prod_{i=1}^{n} \epsilon_{i,t}^{m_i+\tilde{m}_i}\right] - m_0 E\left[\prod_{i=1}^{n} \epsilon_{i,t}^{\tilde{m}_i}\right] - \tilde{m}_0 E\left[\prod_{i=1}^{n} \epsilon_{i,t}^{m_i}\right] + m_0\tilde{m}_0, \tag{2.18}$$

where the superscript SI indicates that the equality $S_{m,\tilde{m}} = S^{SI}_{m,\tilde{m}}$ only holds for serially independent shocks. Let S^{SI} denote S under the assumption of serially independent shocks. Based on Equation (2.18), the long-run covariance under serially independent shocks can be estimated by

$$\hat{S}^{SI}_{m,\tilde{m}}(B) := \frac{1}{T}\sum_{t=1}^{T}\left[\prod_{i=1}^{n} e(B)_{i,t}^{m_i+\tilde{m}_i}\right] - m_0\frac{1}{T}\sum_{t=1}^{T}\left[\prod_{i=1}^{n} e(B)_{i,t}^{\tilde{m}_i}\right] - \tilde{m}_0\frac{1}{T}\sum_{t=1}^{T}\left[\prod_{i=1}^{n} e(B)_{i,t}^{m_i}\right] + m_0\tilde{m}_0, \tag{2.19}$$

where B is an initial guess or a consistent estimator for B_0. Let $\hat{S}^{SI}(B)$ denote the estimator where each element of the long-run covariance matrix is estimated by Equation (2.18). Note that serially independent shocks imply that $S^{SI} = E\left[f(B_0,u_t)f(B_0,u_t)'\right]$ and $\hat{S}^{SI}(B) = \frac{1}{T}\sum_{t=1}^{T} f(B,u_t)f(B,u_t)'$, which corresponds to the estimator for S under the frequently used assumption of serially uncorrelated moment conditions.

With serially independent structural shocks the expression of the long-run covariance matrix S simplifies to the covariance matrix S^{SI}. Nevertheless, the covariance matrix $S^{SI}_{m,\tilde{m}}$ of two fourth-order moments $m,\tilde{m} \in \mathbf{4}$ is still of order eight and remains difficult to estimate in small samples. Analogously, one can now exploit that the shocks are mutually independent to further simplify the estimation of S. Note that many non-Gaussian identification approaches rely on the assumption of mutual independent shocks to ensure identification, see, e.g., Lanne et al. (2017),

[10]To see this note that for $j > 1$

$$E\left[\prod_{i=1}^{n} \epsilon_{i,t}^{m_i}\epsilon_{i,t-j}^{\tilde{m}_i}\right] - m_0 E\left[\prod_{i=1}^{n} \epsilon_{i,t-j}^{\tilde{m}_i}\right] - \tilde{m}_0 E\left[\prod_{i=1}^{n} \epsilon_{i,t-j}^{m_i}\right] + m_0\tilde{m}_0 \tag{2.15}$$

$$= E\left[\prod_{i=1}^{n} \epsilon_{i,t}^{m_i}\right] E\left[\prod_{i=1}^{n} \epsilon_{i,t-j}^{\tilde{m}_i}\right] - m_0 E\left[\prod_{i=1}^{n} \epsilon_{i,t-j}^{\tilde{m}_i}\right] - \tilde{m}_0 E\left[\prod_{i=1}^{n} \epsilon_{i,t-j}^{m_i}\right] + m_0\tilde{m}_0 \tag{2.16}$$

$$= E\left[\prod_{i=1}^{n} \epsilon_{i,t}^{m_i} - m_0\right] E\left[\prod_{i=1}^{n} \epsilon_{i,t-j}^{\tilde{m}_i} - \tilde{m}_0\right] = 0, \tag{2.17}$$

where the first equality follows from serially independent shocks.

Gouriéroux et al. (2017), or Keweloh (2021b). In this case, the researcher already relies on the assumption of mutual independent shocks and may thus as well use it to simplify the estimation of S.

With serially and mutually independent shocks implied by Assumption 2.2 and 2.5 the expression of $S_{m,\tilde{m}}$ simplifies to

$$S_{m,\tilde{m}}^{SMI} = \prod_{i=1}^{n} E\left[\epsilon_{i,t}^{m_i+\tilde{m}_i}\right] - m_0 \prod_{i=1}^{n} E\left[\epsilon_{i,t}^{\tilde{m}_i}\right] - \tilde{m}_0 \prod_{i=1}^{n} E\left[\epsilon_{i,t}^{m_i}\right] + m_0\tilde{m}_0, \tag{2.20}$$

where the superscript SMI indicates that the equality $S_{m,\tilde{m}} = S_{m,\tilde{m}}^{SMI}$ only holds for serially and mutually independent shocks. Let S^{SMI} denote S under the assumption of serially and mutually independent shocks. Based on Equation (2.20), the long-run covariance under serially and mutually independent shocks can be estimated by

$$\hat{S}_{m,\tilde{m}}^{SMI}(B) := \prod_{i=1}^{n} \frac{1}{T}\sum_{t=1}^{T} \left[e(B)_{i,t}^{m_i+\tilde{m}_i}\right] - m_0 \prod_{i=1}^{n} \frac{1}{T}\sum_{t=1}^{T} \left[e(B)_{i,t}^{\tilde{m}_i}\right] - \tilde{m}_0 \prod_{i=1}^{n} \frac{1}{T}\sum_{t=1}^{T} \left[e(B)_{i,t}^{m_i}\right] + m_0\tilde{m}_0, \tag{2.21}$$

where B is an initial guess or a consistent estimator for B_0.[11] Let $\hat{S}^{SMI}(B)$ denote the estimator where each element of the long-run covariance matrix is estimated by Equation (2.21).

Exploiting mutually independent shocks allows to transform higher-order co-moments into a product of lower-order moments. For example, consider the two moment conditions $E[\epsilon_{1,t}^3\epsilon_{2,t}] = 0$ and

[11] Consistency of $\hat{S}_{m,\tilde{m}}^{SMI}(\hat{B}_T) \xrightarrow{p} S_{m,\tilde{m}}^{SMI}$ for $\hat{B}_T \xrightarrow{p} B_0$ follow as usual from continuity of $e(B)_{i,t}$, consistency of $\hat{B}_T$ which implies

$$P\left(|E\left[e(\hat{B}_T)_{i,t}^s\right] - E\left[\epsilon_{i,t}^s\right]| > \gamma/2\right) \to 0 \tag{2.22}$$

and uniform convergence of $\frac{1}{T}\sum_{t=1}^{T} e(B)_{i,t}^s$, such that $\underset{B\in}{SIp}|\frac{1}{T}\sum_{t=1}^{T} e(B)_{i,t}^s - E\left[e(B)_{i,t}^s\right]| \xrightarrow{p} 0$ which implies

$$P\left(|\frac{1}{T}\sum_{t=1}^{T} e(\hat{B}_T)_{i,t}^s - E\left[e(\hat{B}_T)_{i,t}^s\right]| > \gamma/2\right) \to 0 \tag{2.23}$$

and therefore

$$P\left(|\frac{1}{T}\sum_{t=1}^{T} e(\hat{B}_T)_{i,t}^s - E\left[\epsilon_{i,t}^s\right]| > \gamma\right) \to 0. \tag{2.24}$$

Therefore, if the structural shocks are serially and mutually independent, it holds that $S_{m,\tilde{m}}^{SMI} = S_{m,\tilde{m}}$ and hence $\hat{S}_{m,\tilde{m}}^{SMI}(\hat{B}_T)$ is also a consistent estimator for $S_{m,\tilde{m}}$.

$E[\epsilon_{3,t}^3\epsilon_{4,t}] = 0$, such that the covariance of both moment conditions is equal to $E[\epsilon_{1,t}^3\epsilon_{2,t}\epsilon_{3,t}^3\epsilon_{4,t}]$, a co-moment of order eight. However, with mutually independent shocks the covariance is equal to $E[\epsilon_{1,t}^3]E[\epsilon_{2,t}]E[\epsilon_{3,t}^3]E[\epsilon_{4,t}]$, a product of moments of order one and three. In general, the covariance matrix under serially independent shocks S^{SI} requires to calculate co-moments of ϵ_t of order four to eight and the covariance matrix under serially and mutually independent shocks S^{SMI} requires to estimate moments of ϵ_t of order one to six. Table 2.2 shows the number of co-moments of ϵ_t contained in S^{SI} and S^{SMI} of a SVAR GMM estimator using all second- to fourth- order moment conditions. The number of higher-order co-moments increases quickly with the dimension of ϵ_t. For example, in an SVAR with $n = 2$ variables S^{SI} requires to estimate nine co-moments of order eight, however, in an SVAR with $n = 4$ this number grows to 156, and with $n = 6$ variables S^{SI} require to estimate 1287 co-moments of order eight. In contrast to that, the number of higher-order moments in S^{SMI} grows linearly in n. Therefore, using mutually independent shocks to estimate S appears particularly beneficial in larger SVARs.

Table 2.2: Number of moments.

		$n=2$	$n=3$	$n=4$	$n=5$	$n=6$
Number of GMM moment conditions:	second-order	3	6	10	10	15
	third-order	2	7	16	30	50
	fourth-order	3	12	31	65	120
	S dimension	8×8	25×25	57×57	105×105	185×185
Number of co-moments in S^{SI}:	fourth-order	5	15	35	70	126
	fifth-order	6	21	56	126	252
	sixth-order	7	28	84	210	462
	seventh-order	8	36	120	330	792
	eighth-order	9	45	156	495	1287
Number of moments in S^{SMI}:	first-order	2	3	4	5	6
	second-order	2	3	4	5	6
	third-order	2	3	4	5	6
	fourth-order	2	3	4	5	6
	fifths-order	2	3	4	5	6
	sixth-order	2	3	4	5	6

The table shows the number of GMM moment conditions implied by mutually independent shocks and the number of co-moments of ϵ_t contained in S^{SI} and S^{SMI} in a SVAR with two to six variables.

In this study, I simultaneously use the assumption of serially and mutually independent shocks to estimate S. However, one could also directly exploit mutually independent shocks to simplify $S_{m,\tilde{m}}$ in Equation (2.14) without assuming serially independent shocks. Moreover, the weaker assumption of mutually mean independent shocks can also be used to simplify the estimation of

S. In particular, with serially independent and mutually mean independent shocks implied by Assumption 2.2 and 2.6 the expression of $S_{m,\tilde{m}}$ simplifies to

$$S_{m,\tilde{m}}^{SMMI} = E\left[\prod_{\substack{i=1 \\ m_i+\tilde{m}_i \neq 1}}^{n} \epsilon_{i,t}^{m_i+\tilde{m}_i}\right] \prod_{\substack{i=1 \\ m_i+\tilde{m}_i=1}}^{n} E\left[\epsilon_{i,t}^{m_i+\tilde{m}_i}\right] - m_0 E\left[\prod_{\substack{i=1 \\ \tilde{m}_i \neq 1}}^{n} \epsilon_{i,t}^{\tilde{m}_i}\right] \prod_{\substack{i=1 \\ \tilde{m}_i=1}}^{n} E\left[\epsilon_{i,t}^{\tilde{m}_i}\right] \tag{2.25}$$

$$-\tilde{m}_0 E\left[\prod_{\substack{i=1 \\ m_i \neq 1}}^{n} \epsilon_{i,t}^{m_i}\right] \prod_{\substack{i=1 \\ m_i=1}}^{n} E\left[\epsilon_{i,t}^{m_i}\right] + m_0\tilde{m}_0,$$

where the superscript $SMMI$ indicates that the equality $S_{m,\tilde{m}} = S_{m,\tilde{m}}^{SMMI}$ only holds for serially independent and mutually mean independent shocks. For the sake of simplicity, the remainder of the paper focuses on the assumption of mutually independent shocks.

The assumption of mutually independent structural shocks can also be used to estimate G required to estimate the asymptotic variance of the CUE or GMM estimator. For an arbitrary moment condition $f_m(B, u_t) = \prod_{i=1}^{n} e(B)_{i,t}^{m_i} - m_0$ with $m \in \mathbf{2} \cup \mathbf{3} \cup \mathbf{4}$ the derivative with respect to b_{pq} the element at row p and column q of B evaluated at $B = B_0$ corresponds to an element of G and is equal to

$$G_{m,b_{ql}} := E\left[\frac{\partial f_m(B_0, u_t)}{\partial b_{pq}}\right] \tag{2.26}$$

$$= \sum_{j=1, j\neq q}^{n} -m_j a_{jp} E\left[\epsilon_{j,t}^{m_j-1} \epsilon_{q,t}^{m_q+1} \prod_{i=1, i\neq j,q}^{n} \epsilon_{i,t}^{m_i}\right] - m_q a_{qp} E\left[\prod_{i=1}^{n} \epsilon_{i,t}^{m_i}\right], \tag{2.27}$$

with $A = B_0^{-1}$ and a_{jp} are the elements of A. The equality follows from $e(B_0)_t = \epsilon_t$, the product rule, and $\frac{\partial e(B_0)_{i,t}}{\partial b_{pq}} = -a_{ip}\epsilon_{q,t}$. Again, for mutually independent structural shocks implied by Assumption 2.5 if follows

$$G_{m,b_{ql}}^{MI} = \sum_{j=1, j\neq q}^{n} -m_j a_{jp} E\left[\epsilon_{j,t}^{m_j-1}\right] E\left[\epsilon_{q,t}^{m_q+1}\right] \prod_{i=1, i\neq j,q}^{n} E\left[\epsilon_{i,t}^{m_i}\right] - m_q a_{qp} \prod_{i=1}^{n} E\left[\epsilon_{i,t}^{m_i}\right], \tag{2.28}$$

where the superscript MI indicates that the equality $G_{m,b_{ql}} = G_{m,b_{ql}}^{MI}$ only holds for mutually independent shocks. Let G^{MI} denote G under the assumption of mutually independent shocks and let $\hat{G}$ and $\hat{G}^{MI}$ denote the corresponding estimators. Again, mutual independence allows to calculate higher-order co-moments as a product of lower-order moments. For example, in a SVAR with $n = 4$ variables and a GMM estimator including all 4 variance, 6 covariance, 16 coskewness, and 31 cokurtosis conditions implied by independent shocks with unit variance, the covariance matrix G is a 57×16 dimensional matrix containing 10 co-moments of order two, 16 co-moments of order three, and 31 co-moments of order four. In contrast to that, G^{MI} contains four moments of order one, four moments of order two, four moments of order three, four moments of order four.

The assumption of serially and mutually independent shocks can also be used to derive a guess for the optimal weighting matrix W^* without requiring an initial guess or estimate of the unknown simultaneous interaction B_0. Instead, the researcher can guess the distribution of each structural shock $\epsilon_{i,t}$ for $i = 1, ..., n$ and if the guess is correct Equation (2.20) directly yields the true covariance matrix S, which can be used to calculate the optimal weighting matrix. In practice, I recommend starting with the assumption of t- or normally distributed shocks to approximate S and hence W^*. I find that even if the initially assumed distributions are incorrect, the corresponding one-step GMM estimator performs similarly in terms of bias and interquartile range to the one-step GMM estimator using the true asymptotically optimal weighting matrix. This might be related to the fact that due to the normalization to mean zero and unit variance shocks, the guess of higher-order moments is irrelevant for many moments. For example the two moment conditions $E[\epsilon_{1,t}^3 \epsilon_{2,t}] = 0$ and $E[\epsilon_{1,t}^3 \epsilon_{3,t}] = 0$ require to estimate $E[\epsilon_{1,t}^6 \epsilon_{2,t} \epsilon_{3,t}]$. However, this co-moment of order eight is equal to zero for all independent shocks with mean zero and finite second to six moments.

2.5 Monte Carlo Simulation

This section compares the impact of the estimates $\hat{S}^{SI}$, $\hat{S}^{SMI}$, $\hat{G}$, and $\hat{G}^{MI}$ on the finite sample performance of CUE and GMM estimators. I simulate a SVAR $u_t = B_0 \epsilon_t$ with $n = 2$ and $n = 4$

variables with

$$B_0 = \begin{bmatrix} 1 & 0 \\ 0.5 & 1 \end{bmatrix} \text{ and } B_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.5 & 1 & 0 & 0 \\ 0.5 & 0.5 & 1 & 0 \\ 0.5 & 0.5 & 0.5 & 1 \end{bmatrix}. \tag{2.29}$$

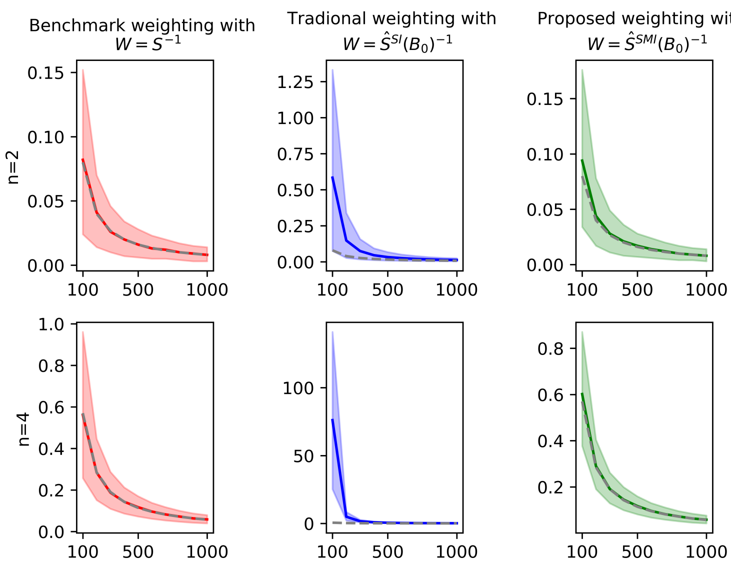
The structural shocks are drawn from a mixture of Gaussian distributions with mean zero, unit variance, skewness equal to 0.89 and an excess kurtosis of 2.35. In particular, the shocks satisfy

$$\varepsilon_i = z\phi_1 + (1-z)\phi_2 \text{ with } \phi_1 \sim \mathcal{N}(-0.2, 0.7),\ \phi_2 \sim \mathcal{N}(0.75, 1.5),\ z \sim \mathcal{B}(0.79), \tag{2.30}$$

where $\mathcal{B}(p)$ indicates a Bernoulli distribution and $\mathcal{N}(\mu, \sigma^2)$ indicates a normal distribution. Simulations based on t-distributed shocks are shown in the appendix.

Before turning to different CUE and GMM estimators, I analyze the impact of the estimated asymptotically efficient weighting matrix on the GMM loss. Figure 2.1 compares the average and quantiles of the GMM objective function $g_T(B)'Wg_T(B)$ at $B = B_0$ with all second- to fourth-order moment conditions implied by mutually independent shocks for different weighting matrices. The red loss serves as a benchmark and uses the true but in practice unknown asymptotically efficient weighting matrix, $W^* = S^{-1}$. The blue loss uses the traditional estimator for the asymptotically efficient weighting matrix relying on serially uncorrelated moment conditions equivalent to serially independent shocks, $\hat{W}^{SI} = \hat{S}^{SI}(B_0)^{-1}$. The green loss corresponds to the proposed estimator for the asymptotically efficient weighting matrix using serially and mutually independent shocks, $\hat{W}^{SMI} = \hat{S}^{SMI}(B_0)^{-1}$. The simulation shows that the standard estimator for the asymptotically efficient weighting matrix $\hat{W}^{SI}$ is ill suited to approximate the asymptotically efficient weighting matrix in small samples. In the small SVAR with $n = 2$ and the smallest sample size $T = 100$, the average GMM loss based on $\hat{W}^{SI}$ is approximately seven times larger than the average GMM loss based on the asymptotically efficient weighting matrix W^* and in the large SVAR with $n = 4$ the average loss is approximately 132 times larger. In contrast to that, the weighting scheme proposed in this study, $\hat{W}^{SMI}$, which exploits the mutual independence of the structural shocks closely approximates the infeasible asymptotically efficient weighting scheme

Figure 2.1: Finite sample performance - GMM loss at B_0 for different weighting schemes.



Average, 10%, and 90% quantiles of the GMM loss $g_T(B)'Wg_T(B)$ with all second-, third-, and fourth-order moment conditions implied by mutually independent shocks at $B = B_0$ for $W = S^{-1}$ in red, $W = \hat{S}_T^{SI}(B_0)^{-1}$ in blue, and $W = \hat{S}_T^{SMI}(B_0)^{-1}$ in green with 5000 simulations and sample sizes $T = 100, 200, ..., 1000$. The dotted gray line shows the expected value of the GMM objective function at B_0 and $W = S^{-1}$ which is equal to m/T where m denotes the number of moment conditions, compare Han and Phillips (2006) and Newey and Windmeijer (2009).

with $W^* = S^{-1}$.

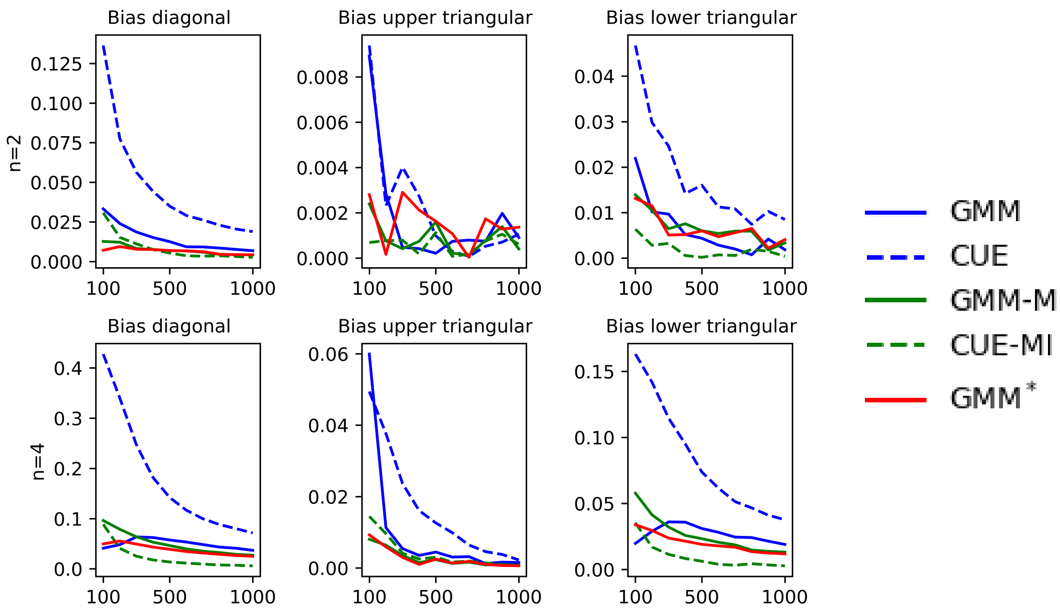
In the following, I analyze the impact of the weighting scheme on the following asymptotically efficient estimators:

- GMM*: A one-step GMM estimator with the asymptotically efficient weighting matrix $W = S^{-1}$.
- GMM: A two-step GMM estimator with $W = I$ in the first step and $W = \hat{S}^{SI}(\hat{B})^{-1}$ in the second step.
- CUE: A continuous updating estimator with $W(B) = \hat{S}^{SI}(B)^{-1}$.
- GMM-MI: A two-step GMM estimator with $W = S_{Norm}^{SMI^{-1}}$ in the first step, $W = \hat{S}^{SI}(\hat{B})^{-1}$ in the second step and S_{Norm}^{SMI} denotes the long-run covariance matrix under serially and mutually independent Gaussian shocks.
- CUE-MI: A continuous updating estimator with $W(B) = \hat{S}^{SMI}(B)^{-1}$.

The estimator GMM* is infeasible since it uses the unknown asymptotically efficient weighting matrix $W = S^{-1}$ and it serves as a benchmark. The CUE and GMM estimators only use the assumption of serially independent shocks or serially uncorrelated moment conditions to estimate S and therefore, represent the traditional estimation approaches. The CUE-MI and GMM-MI estimators are the novel estimators proposed in this study and rely on the assumption of serially and mutually independent shocks to estimate S. All estimators use all second-, third-, and fourth-order moment conditions implied by mutually independent shocks. In particular, for $n = 2$ the estimators use three second-, two third-, and three fourth-order moment conditions and for $n = 4$ the estimators use 10 second-, 16 third-, and 31 fourth-order moment conditions. Note that all analyzed estimators have the same asymptotic variance and are asymptotically efficient. The appendix contains further results for the one-step GMM estimator using the identity weighting matrix $W = I$, the approximation of the asymptotically efficient weighting matrix based on serially and mutually independent Gaussian shocks $W = S_{Norm}^{SMI^{-1}}$, and the white fast weighting matrix proposed in Keweloh (2021b).

Firstly, I analyze the impact of the weighting scheme on the finite sample bias and the interquartile range. The elements of B are equal on the diagonal, upper- and lower- triangular. Therefore,

Figure 2.2: Finite sample performance - Bias of the estimated elements.



Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, ..., 1000$. The figure shows the average of the median absolute bias of the elements on the diagonal, upper-, lower- triangular of $\hat{B}$. Let $\hat{B}^m$ be the estimator in one simulation m and $1 \leq m \leq M$. Let $\bar{B}^m := abs(\hat{B}^m - B_0)$ be the absolute bias in simulation m. The median absolute bias over all simulations M is then denoted by $bias := med(\bar{B}^m)$, which is a $n \times n$ matrix containing the median absolute bias over all simulations for each element $\hat{B}_{ij}$. The average of the median absolute bias of the elements on the diagonal, upper-, lower- triangular of $\hat{B}$ is then the average of all elements on the diagonal, upper-, lower- triangular of $bias$.

I summarizes the results for all elements on the diagonal, upper- and lower- triangular of B. Figure 2.2 shows the average of the median absolute bias and Figure 2.3 shows the average of the interquartile range (IQR) of the estimated elements on the diagonal, upper- and lower- triangular matrix $\hat{B}$. The results for all individual elements of B for $T = 100$ and $T = 1000$ are shown in the appendix. The computation of the average of the median absolute bias and the average of the interquartile range of the diagonal, upper-, lower-triangular elements is shown in the description of the corresponding figure.

In small samples, all estimators are biased and the bias increases with the dimension of the SVAR. The elements on the diagonal show the largest bias and the elements in the upper triangular have the smallest bias. This pattern can be explained by a bias due to scaling, meaning that for small T the GMM loss $E\left[g_T(B)'Wg_T(B)\right]$ is not minimized at $B = B_0$ but at $B = DB_0$

where $D = diag(d_1, ..., d_n)$ is a scaling matrix.[12] Moreover, Figure 2.2 shows that the CUE has the largest bias. In contrast to that, the CUE-MI performs notably better. In fact, in the large SVAR it has the lowest bias on the diagonal and lower triangular of all estimators for sample sizes with more than 100 observations. This behavior can be explained by Figure 2.1, which suggests that in small samples, the estimator $\hat{S}^{SI}(B_0)$ poorly approximates S while $\hat{S}^{SMI}(B_0)$ yields a much better approximation. Additionally, the derivation of the signal term in Equation (2.11) requires a nonrandom weighting matrix W, which is not satisfied by the CUE. The signal term of the CUE can only be written as $E\left[\sum_{t\neq\tilde{t}} f(B, u_t)'\hat{S}(B)^{-1} f(B, u_{\tilde{t}})\right]$, which is not necessarily minimized at B_0. Therefore, the CUE can be biased since it searches for solutions which minimize the GMM loss by manipulating the weights $W(B) = \hat{S}(B)^{-1}$.

Figure 2.3 shows that the average IQR of the estimated elements increases with the dimension of the SVAR. Again, the CUE-MI and GMM-MI estimators perform better than the CUE and GMM estimator. In particular, for $n = 4$ and $T = 100$ the average IQR of the GMM estimator is two to three times larger than the average IQR of the CUE-MI.

Secondly, I analyze the impact of the weighting scheme and the estimation of the asymptotic variance on the rejection frequencies for different tests. Figure 2.4 shows the rejection frequencies at the 10% nominal level for a J-Test, the Wald test with $H_0 : B_{i,j} = 0$ for $j > i$ testing the null hypothesis of a recursive SVAR, and a Wald test with the null hypothesis $H_0 : B_{1,n} = 0$ for $n = 2$ and $n = 4$. The Wald tests require an estimate of the asymptotic variance. The estimators GMM*, GMM, and CUE use the standard estimators $\hat{S}^{SI}_T(\hat{B})$ and $\hat{G}(\hat{B})$ and the estimators GMM*-MI, GMM-MI, and CUE-MI use the proposed estimators $\hat{S}^{SMI}_T(\hat{B})$ and $\hat{G}^{MI}(\hat{B})$

[12] For example with $W = I$ the noise term in Equation 2.11 is equal to

$$\frac{1}{T} E\left[f(B_0, u_t)' W f(B_0, u_t)\right] = \frac{1}{T} E\left[f(B_0, u_t)' f(B_0, u_t)\right], \tag{2.31}$$

which is the sum of the variances of the variance, covariance, and coskewness conditions. The variance of any moment condition m of the type $\prod_{i=1}^{n} e(B)^{m_i}_{i,t}$ at B_0 is equal to

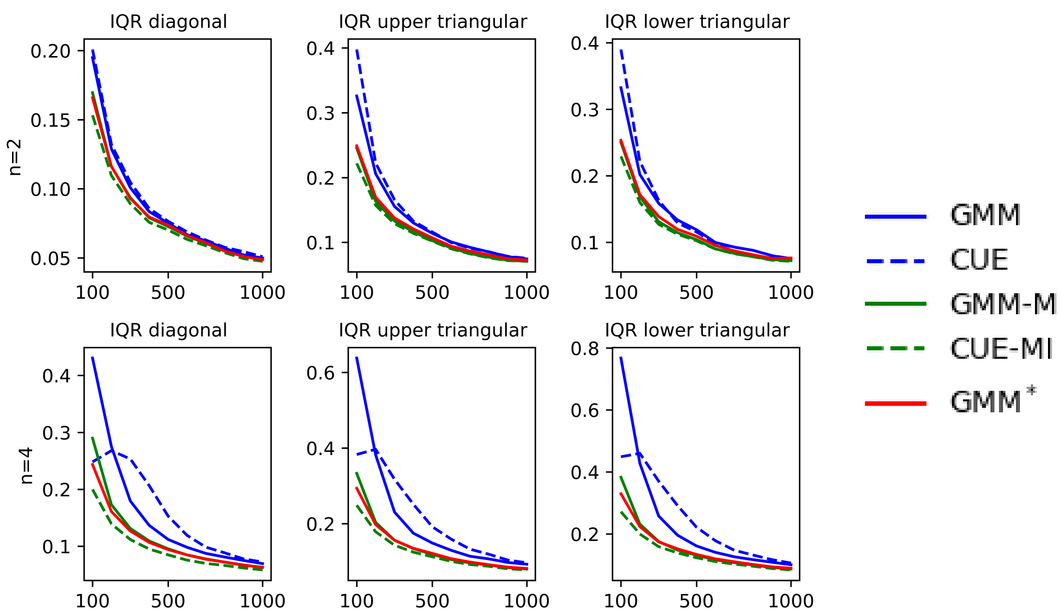
$$E\left[\prod_{i=1}^{n} e(B_0)^{2m_i}_{i,t})\right] = E\left[\prod_{i=1}^{n} \epsilon^{2m_i}_{i,t}\right]. \tag{2.32}$$

Let $S = diag(d_1, ..., d_n)$ be a scaling matrix and note that $De(B_0) = DB_0 u = e(DB_0)$, such that

$$E\left[\prod_{i=1}^{n} e^{2m_i}_{i,t}(SB_0)\right] = E\left[\prod_{i=1}^{n} \frac{\epsilon^{2m_i}_{i,t}}{d_i^{2m_i}}\right] = \frac{1}{d_1^{2m_1} ... d_n^{2m_n}} E\left[\prod_{i=1}^{n} \epsilon^{2m_i}_{i,t}\right]. \tag{2.33}$$

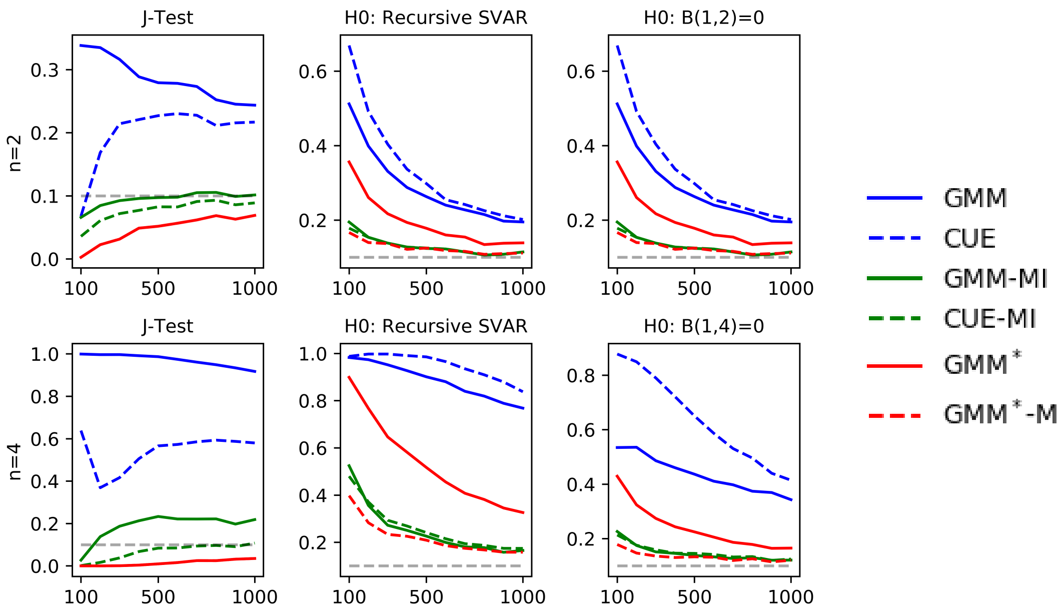
It is easy to see that at B_0 an increase in the scaling parameter d_i, which corresponds to a decrease of the sample variance of the i-th estimated structural shock, decreases the noise term and therefore, leads to a bias.

Figure 2.3: Finite sample performance - Interquartile range of the estimated elements.



Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, ..., 1000$. The figure shows the average of the interquartile range (IQR) of the elements on the diagonal, upper-, lower- triangular of $\hat{B}$. Let $\hat{B}^m$ be the estimator in one simulation m and $1 \leq m \leq M$. The interquartile range over all simulations M is then denoted by $iqr := quartile(\hat{B}^m, 0.75) - quartile(\hat{B}^m, 0.25)$, which is a $n \times n$ matrix containing the interquartile range over all simulations for each element $\hat{B}_{ij}$. The average of the interquartile range of the elements on the diagonal, upper-, lower- triangular of $\hat{B}$ is then the average of all elements on the diagonal, upper-, lower- triangular of iqr.

Figure 2.4: Finite sample performance - Rejection rates.



Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, ..., 1000$. The figure shows the rejection rates at $\alpha = 10\%$ for J-Test, recursive SVAR Wald test, and Wald test with $H_0 : B_{1,n} = 0$ for $n = 2$ and $n = 4$.

to estimate the asymptotic variance. The GMM*-MI estimator is equal to the GMM* estimator using the true asymptotically optimal weighting matrix $W = S^{-1}$. The tests for GMM* and GMM*-MI only differ in the way the asymptotic variance is estimated. Figure 2.4 shows that the distortion of the tests increases with an increase of the dimension of the SVAR, a decrease of the sample size, and an increase of the number of hypothesis being jointly tested. The CUE and GMM estimators have the largest distortions and they decrease only slowly with an increase of the sample size. For example, even in the largest sample with $T = 1000$ observations the CUE and GMM estimator reject the null hypothesis of a recursive SVAR in roughly 80% of all cases. The rejection rates of CUE-MI and GMM-MI estimators are also distorted in small samples, however, the distortion is much smaller and decreases more quickly with an increase of the sample size. GMM* and GMM*-MI both use the same estimator, the tests only differ due to the estimated asymptotic variance. The smaller distortions of the GMM*-MI tests compared to the GMM* tests can solely be attributed to the impact of exploiting the assumption of mutually independent shocks to estimate the asymptotic variance.

2.6 Conclusion

This paper argues that the assumption of mutually independent shocks should be used to estimate the asymptotically efficient weighting matrix and the asymptotic variance of SVAR CUE and SVAR GMM estimators based on higher-order moment conditions. Without exploiting the assumption of mutually independent shocks, estimating the covariance of fourth-order moment conditions requires to estimate co-moments of order eight. This leads to biased and volatile estimates and oversize test statistics in finite samples. With mutually independent shocks, the covariance of the higher-order moment conditions can be estimated as the product of moments up to order six. A Monte Carlo simulation demonstrates that the propose approach improves the finite sample performance of the SVAR CUE and SVAR GMM estimators, especially in larger SVAR models.

2.7 Appendix

2.7.1 Appendix - Simulation with mixture of Gaussian distributions

This section supplements the Monte Carlo simulation shown in Section 2.5. Table 2.3, 2.4, and 2.5 show the median (med), 0.25 quantile ($q25$), 0.75 quantile ($q75$), the 0.25 confidence level ($c25$), and the 0.75 confidence level ($c75$).

Table 2.3: Finite sample performance - All estimated elements for $n = 2$.

	$T = 100$	$T = 1000$
GMM^*	$\begin{bmatrix} \underset{\substack{0.93\ 1.09\\0.93\ 1.07}}{1.0} & \underset{\substack{-0.13\ 0.12\\-0.11\ 0.11}}{-0.01} \\ \underset{\substack{0.39\ 0.64\\0.39\ 0.61}}{0.52} & \underset{\substack{0.92\ 1.1\\0.92\ 1.08}}{1.02} \end{bmatrix}$	$\begin{bmatrix} \underset{\substack{0.99\ 1.03\\0.98\ 1.02}}{1.01} & \underset{\substack{-0.04\ 0.04\\-0.03\ 0.03}}{0.0} \\ \underset{\substack{0.47\ 0.54\\0.47\ 0.53}}{0.51} & \underset{\substack{0.98\ 1.04\\0.97\ 1.03}}{1.01} \end{bmatrix}$
GMM	$\begin{bmatrix} \underset{\substack{0.88\ 1.06\\0.93\ 1.07}}{0.97} & \underset{\substack{-0.17\ 0.17\\-0.11\ 0.11}}{0.0} \\ \underset{\substack{0.31\ 0.65\\0.39\ 0.61}}{0.48} & \underset{\substack{0.86\ 1.09\\0.92\ 1.08}}{0.98} \end{bmatrix}$	$\begin{bmatrix} \underset{\substack{0.97\ 1.02\\0.98\ 1.02}}{1.0} & \underset{\substack{-0.04\ 0.04\\-0.03\ 0.03}}{-0.0} \\ \underset{\substack{0.46\ 0.54\\0.47\ 0.53}}{0.5} & \underset{\substack{0.97\ 1.03\\0.97\ 1.03}}{1.0} \end{bmatrix}$
CUE	$\begin{bmatrix} \underset{\substack{0.75\ 0.94\\0.93\ 1.07}}{0.84} & \underset{\substack{-0.2\ 0.2\\-0.11\ 0.11}}{-0.01} \\ \underset{\substack{0.26\ 0.65\\0.39\ 0.61}}{0.46} & \underset{\substack{0.76\ 0.97\\0.92\ 1.08}}{0.88} \end{bmatrix}$	$\begin{bmatrix} \underset{\substack{0.96\ 1.01\\0.98\ 1.02}}{0.99} & \underset{\substack{-0.04\ 0.04\\-0.03\ 0.03}}{0.0} \\ \underset{\substack{0.46\ 0.53\\0.47\ 0.53}}{0.49} & \underset{\substack{0.96\ 1.01\\0.97\ 1.03}}{0.99} \end{bmatrix}$
$GMM - MI$	$\begin{bmatrix} \underset{\substack{0.93\ 1.1\\0.93\ 1.07}}{1.01} & \underset{\substack{-0.13\ 0.12\\-0.11\ 0.11}}{-0.0} \\ \underset{\substack{0.39\ 0.64\\0.39\ 0.61}}{0.52} & \underset{\substack{0.93\ 1.11\\0.92\ 1.08}}{1.02} \end{bmatrix}$	$\begin{bmatrix} \underset{\substack{0.99\ 1.03\\0.98\ 1.02}}{1.01} & \underset{\substack{-0.03\ 0.04\\-0.03\ 0.03}}{0.0} \\ \underset{\substack{0.47\ 0.54\\0.47\ 0.53}}{0.51} & \underset{\substack{0.98\ 1.04\\0.97\ 1.03}}{1.01} \end{bmatrix}$
$CUE - MI$	$\begin{bmatrix} \underset{\substack{0.89\ 1.04\\0.93\ 1.07}}{0.97} & \underset{\substack{-0.12\ 0.11\\-0.11\ 0.11}}{-0.0} \\ \underset{\substack{0.38\ 0.61\\0.39\ 0.61}}{0.5} & \underset{\substack{0.89\ 1.06\\0.92\ 1.08}}{0.98} \end{bmatrix}$	$\begin{bmatrix} \underset{\substack{0.98\ 1.03\\0.98\ 1.02}}{1.0} & \underset{\substack{-0.03\ 0.04\\-0.03\ 0.03}}{0.0} \\ \underset{\substack{0.47\ 0.54\\0.47\ 0.53}}{0.5} & \underset{\substack{0.98\ 1.03\\0.97\ 1.03}}{1.0} \end{bmatrix}$

Monte Carlo simulation with 5000 iterations and sample sizes. Median (med), 0.25 quantile ($q25$), 0.75 quantile ($q75$), the 0.25 confidence level ($c25$), and the 0.75 confidence level ($c75$), where the confidence levels are calculated according to $B_{i,j} \pm z^* \frac{\sigma_{i,j}}{\sqrt{T}}$ with $z^* = 0.67$ and $\sigma_{i,j}$ is the square root of the variance of the element i, j according to $\sqrt{T}(\hat{B}_T - B_0) \xrightarrow{d} \mathcal{N}(0, MSM')$. For each element the data is shown as $\underset{\substack{q25\ q75\\c25\ c75}}{med}$.

Table 2.4: Finite sample performance - All estimated elements for $n = 4$ (Part 1).

	$T = 100$				$T = 1000$			
GMM^*	1.04 0.95 1.14 0.93 1.07	0.0 −0.12 0.14 −0.11 0.11	−0.0 −0.13 0.13 −0.11 0.11	−0.02 −0.14 0.13 −0.11 0.11	1.02 1.0 1.05 0.98 1.02	−0.0 −0.04 0.04 −0.03 0.03	−0.0 −0.04 0.04 −0.03 0.03	0.0 −0.04 0.04 −0.03 0.03
	0.53 0.39 0.66 0.39 0.61	1.05 0.93 1.16 0.92 1.08	−0.01 −0.16 0.15 −0.12 0.12	−0.02 −0.17 0.15 −0.12 0.12	0.51 0.48 0.55 0.47 0.53	1.03 1.0 1.05 0.97 1.03	0.0 −0.04 0.04 −0.04 0.04	−0.0 −0.04 0.04 −0.04 0.04
	0.54 0.37 0.68 0.38 0.62	0.53 0.37 0.69 0.38 0.62	1.05 0.92 1.18 0.9 1.1	−0.02 −0.19 0.16 −0.14 0.14	0.51 0.47 0.56 0.46 0.54	0.51 0.47 0.55 0.46 0.54	1.02 0.99 1.06 0.97 1.03	−0.0 −0.05 0.05 −0.04 0.04
	0.54 0.35 0.71 0.36 0.64	0.54 0.35 0.72 0.36 0.64	0.54 0.35 0.7 0.36 0.64	1.05 0.89 1.19 0.89 1.11	0.51 0.46 0.56 0.46 0.54	0.51 0.46 0.56 0.46 0.54	0.51 0.46 0.56 0.46 0.54	1.03 0.99 1.07 0.96 1.04
GMM	1.0 0.85 1.19 0.93 1.07	0.13 −0.16 0.44 −0.11 0.11	0.05 −0.24 0.37 −0.11 0.11	0.01 −0.3 0.34 −0.11 0.11	0.96 0.94 0.99 0.98 1.02	−0.0 −0.04 0.04 −0.03 0.03	−0.0 −0.04 0.04 −0.03 0.03	0.0 −0.04 0.04 −0.03 0.03
	0.46 0.12 0.81 0.39 0.61	1.0 0.83 1.21 0.92 1.08	0.1 −0.2 0.44 −0.12 0.12	0.02 −0.31 0.37 −0.12 0.12	0.48 0.44 0.52 0.47 0.53	0.96 0.93 1.0 0.97 1.03	−0.0 −0.05 0.05 −0.04 0.04	−0.0 −0.05 0.05 −0.04 0.04
	0.47 0.1 0.84 0.38 0.62	0.51 0.13 0.89 0.38 0.62	0.98 0.77 1.21 0.9 1.1	0.05 −0.26 0.39 −0.14 0.14	0.48 0.44 0.53 0.46 0.54	0.48 0.43 0.53 0.46 0.54	0.96 0.92 1.0 0.97 1.03	−0.0 −0.06 0.05 −0.04 0.04
	0.46 0.06 0.85 0.36 0.64	0.5 0.11 0.91 0.36 0.64	0.5 0.08 0.92 0.36 0.64	0.87 0.58 1.13 0.89 1.11	0.48 0.43 0.53 0.46 0.54	0.48 0.43 0.54 0.46 0.54	0.48 0.42 0.53 0.46 0.54	0.96 0.92 1.01 0.96 1.04
CUE	0.57 0.5 0.66 0.93 1.07	0.1 −0.08 0.26 −0.11 0.11	0.05 −0.13 0.22 −0.11 0.11	0.01 −0.19 0.21 −0.11 0.11	0.93 0.9 0.95 0.98 1.02	0.0 −0.04 0.04 −0.03 0.03	0.0 −0.04 0.04 −0.03 0.03	−0.0 −0.04 0.04 −0.03 0.03
	0.32 0.13 0.5 0.39 0.61	0.6 0.5 0.7 0.92 1.08	0.09 −0.1 0.28 −0.12 0.12	0.02 −0.19 0.22 −0.12 0.12	0.46 0.42 0.51 0.47 0.53	0.93 0.89 0.96 0.97 1.03	−0.0 −0.05 0.05 −0.04 0.04	−0.0 −0.05 0.05 −0.04 0.04
	0.33 0.11 0.53 0.38 0.62	0.35 0.13 0.57 0.38 0.62	0.59 0.45 0.72 0.9 1.1	0.04 −0.17 0.24 −0.14 0.14	0.46 0.41 0.51 0.46 0.54	0.46 0.41 0.51 0.46 0.54	0.93 0.89 0.97 0.97 1.03	−0.01 −0.06 0.05 −0.04 0.04
	0.33 0.09 0.56 0.36 0.64	0.35 0.1 0.6 0.36 0.64	0.33 0.08 0.58 0.36 0.64	0.53 0.33 0.7 0.89 1.11	0.46 0.4 0.52 0.46 0.54	0.46 0.41 0.52 0.46 0.54	0.46 0.41 0.52 0.46 0.54	0.93 0.88 0.97 0.96 1.04

Monte Carlo simulation with 5000 iterations and sample sizes. Median (med), 0.25 quantile ($q25$), 0.75 quantile ($q75$), the 0.25 confidence level ($c25$), and the 0.75 confidence level ($c75$), where the confidence levels are calculated according to $B_{i,j} \pm z^* \frac{\sigma_{i,j}}{\sqrt{T}}$ with $z^* = 0.67$ and $\sigma_{i,j}$ is the square root of the variance of the element i, j according to $\sqrt{T}(\hat{B}_T - B_0) \xrightarrow{d} \mathcal{N}(0, MSM')$. For each element the data is shown as $\underset{\substack{q25\ q75 \\ c25\ c75}}{med}$.

Table 2.5: Finite sample performance - All estimated elements for $n = 4$ (Part 2).

	$T = 100$				$T = 1000$			
$GMM - MI$	1.1 0.98 1.22 0.93 1.07	0.0 −0.14 0.16 −0.11 0.11	−0.0 −0.15 0.15 −0.11 0.11	−0.01 −0.16 0.14 −0.11 0.11	1.03 1.0 1.05 0.98 1.02	−0.0 −0.04 0.04 −0.03 0.03	−0.0 −0.04 0.04 −0.03 0.03	0.0 −0.04 0.04 −0.03 0.03
	0.56 0.39 0.71 0.39 0.61	1.1 0.96 1.24 0.92 1.08	−0.0 −0.18 0.18 −0.12 0.12	−0.01 −0.19 0.17 −0.12 0.12	0.52 0.48 0.55 0.47 0.53	1.03 1.0 1.06 0.97 1.03	0.0 −0.04 0.04 −0.04 0.04	−0.0 −0.04 0.04 −0.04 0.04
	0.56 0.38 0.73 0.38 0.62	0.56 0.38 0.74 0.38 0.62	1.1 0.94 1.25 0.9 1.1	−0.02 −0.2 0.19 −0.14 0.14	0.51 0.47 0.56 0.46 0.54	0.51 0.47 0.55 0.46 0.54	1.03 0.99 1.06 0.97 1.03	−0.0 −0.05 0.05 −0.04 0.04
	0.56 0.34 0.77 0.36 0.64	0.56 0.35 0.77 0.36 0.64	0.55 0.34 0.75 0.36 0.64	1.09 0.91 1.25 0.89 1.11	0.51 0.46 0.56 0.46 0.54	0.51 0.47 0.56 0.46 0.54	0.51 0.47 0.56 0.46 0.54	1.03 0.99 1.07 0.96 1.04
$CUE - MI$	0.91 0.83 0.99 0.93 1.07	−0.0 −0.11 0.11 −0.11 0.11	−0.01 −0.12 0.11 −0.11 0.11	−0.01 −0.13 0.1 −0.11 0.11	0.99 0.97 1.01 0.98 1.02	−0.0 −0.04 0.04 −0.03 0.03	−0.0 −0.04 0.03 −0.03 0.03	0.0 −0.03 0.04 −0.03 0.03
	0.46 0.34 0.57 0.39 0.61	0.91 0.82 1.0 0.92 1.08	−0.01 −0.14 0.13 −0.12 0.12	−0.02 −0.15 0.11 −0.12 0.12	0.5 0.46 0.53 0.47 0.53	0.99 0.97 1.02 0.97 1.03	0.0 −0.04 0.04 −0.04 0.04	−0.0 −0.04 0.04 −0.04 0.04
	0.47 0.33 0.59 0.38 0.62	0.46 0.32 0.59 0.38 0.62	0.92 0.8 1.02 0.9 1.1	−0.03 −0.17 0.12 −0.14 0.14	0.5 0.46 0.54 0.46 0.54	0.5 0.45 0.54 0.46 0.54	0.99 0.96 1.03 0.97 1.03	−0.0 −0.05 0.04 −0.04 0.04
	0.48 0.33 0.62 0.36 0.64	0.47 0.31 0.6 0.36 0.64	0.45 0.31 0.6 0.36 0.64	0.91 0.78 1.02 0.89 1.11	0.5 0.45 0.54 0.46 0.54	0.5 0.45 0.54 0.46 0.54	0.5 0.45 0.54 0.46 0.54	0.99 0.96 1.03 0.96 1.04

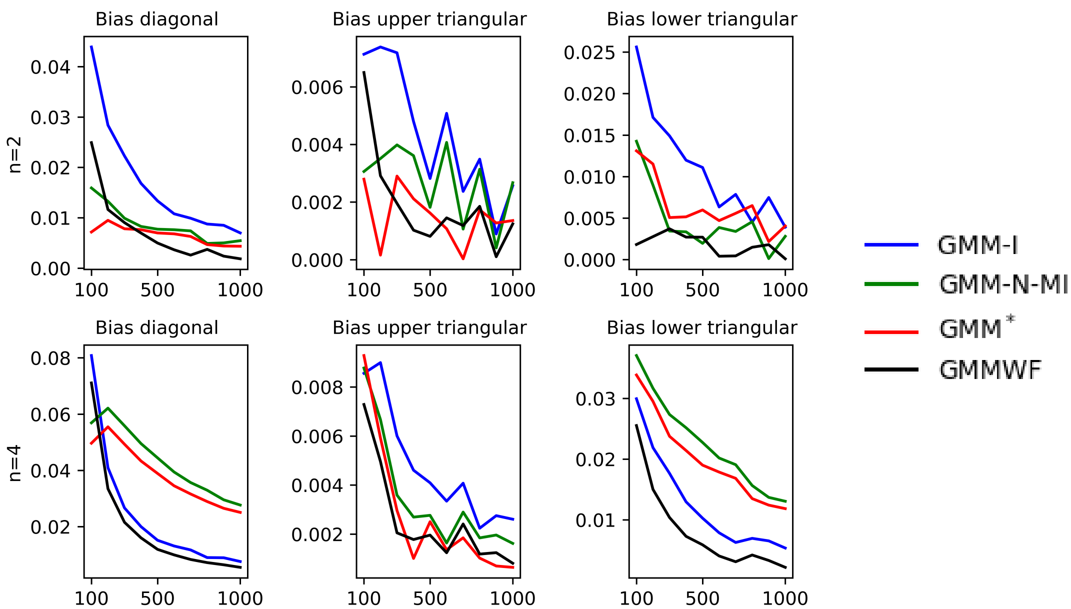
Monte Carlo simulation with 5000 iterations and sample sizes. Median (med), 0.25 quantile ($q25$), 0.75 quantile ($q75$), the 0.25 confidence level ($c25$), and the 0.75 confidence level ($c75$), where the confidence levels are calculated according to $B_{i,j} \pm z^* \frac{\sigma_{i,j}}{\sqrt{T}}$ with $z^* = 0.67$ and $\sigma_{i,j}$ is the square root of the variance of the element i, j according to $\sqrt{T}(\hat{B}_T - B_0) \xrightarrow{d} \mathcal{N}(0, MSM')$. For each element the data is shown as $\underset{\substack{q25\ q75 \\ c25\ c75}}{med}$.

Figure 2.5, 2.6, and 2.7 show the average of the median absolute bias, the average of the interquartile range (IQR) of the estimated elements, and the rejection frequencies for the hypothesis tests. The Figures contain results for the following estimators:

- GMM*: A one-step GMM estimator with the asymptotically efficient weighting matrix $W = S^{-1}$.
- GMM-I: A one-step GMM estimator with $W = I$.
- GMM-N-MI: A one-step GMM estimator with $W = S_{Norm}^{SMI^{-1}}$, where S_{Norm}^{SMI} denotes the long-run covariance matrix under serially and mutually independent Gaussian shocks.
- GMMWF: A one-step GMM estimator with the fast weighting matrix proposed in Keweloh (2021b).

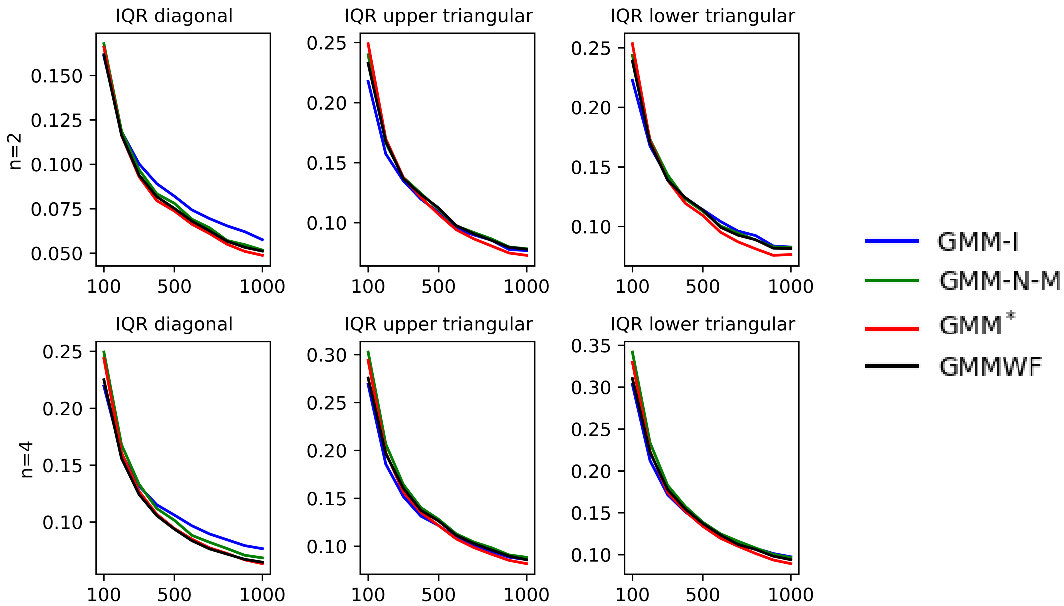
All estimators use all second- to fourth-order moment conditions implied by independent shocks. The Wald tests require an estimate of the asymptotic variance. The estimators GMM* and GMM-I use the standard estimators $\hat{S}_T^{SI}(\hat{B})$ and $\hat{G}(\hat{B})$ and the estimators GMM*-MI, GMM-N-MI, and GMMWF use the proposed estimators $\hat{S}_T^{SMI}(\hat{B})$ and $\hat{G}^{MI}(\hat{B})$ to estimate the asymptotic variance. The GMM*-MI estimator is equal to the GMM* estimator using the true asymptotically optimal weighting matrix $W = S^{-1}$. The tests for GMM* and GMM*-MI only differ in the way the asymptotic variance is estimated.

Figure 2.5: Finite sample performance - Bias of the estimated elements (one-step estimators).



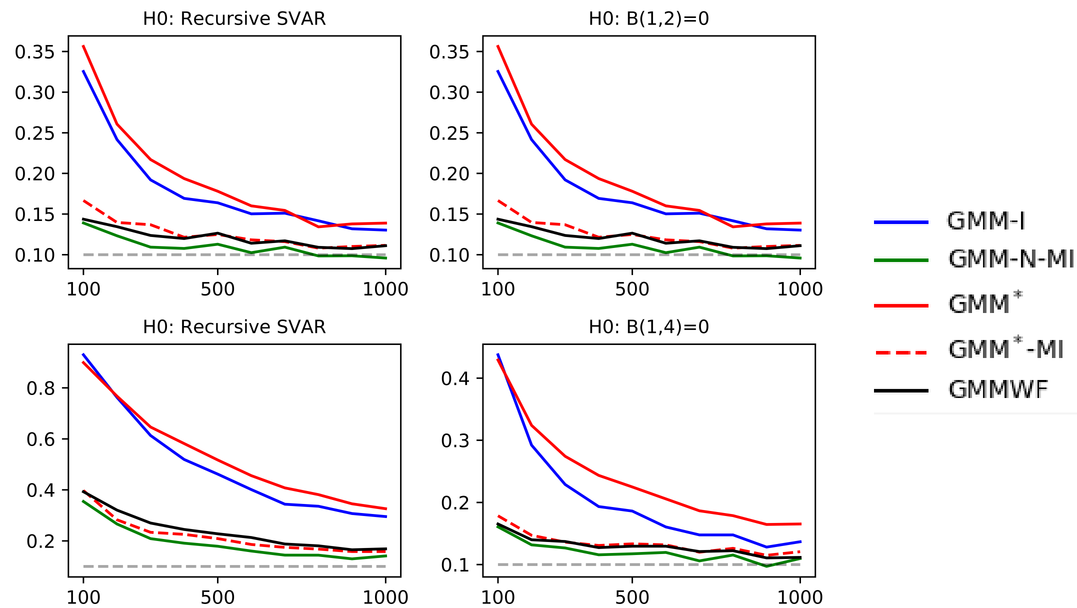
Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, ..., 1000$. The figure shows the average of the median absolute bias of the elements on the diagonal/upper-/lower- triangular of $\hat{B}$. Let $\hat{B}^m$ be the estimator in one simulation m and $1 \leq m \leq M$. Let $\bar{B}^m := abs(\hat{B}^m - B_0)$ be the absolute bias in simulation m. The median absolute bias over all simulations M is then denoted by $bias := med(\bar{B}^m)$, which is a $n \times n$ matrix containing the median absolute bias over all simulations for each element $\hat{B}_{ij}$. The average of the median absolute bias of the elements on the diagonal/upper-/lower- triangular of $\hat{B}$ is then the average of all elements on the diagonal/upper-/lower- triangular of $bias$.

Figure 2.6: Finite sample performance - Interquartile range of the estimated elements (one-step estimators).



Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, ..., 1000$. The figure shows the average of the interquartile range (IQR) of the elements on the diagonal/upper-/lower- triangular of $\hat{B}$. Let $\hat{B}^m$ be the estimator in one simulation m and $1 \leq m \leq M$. The interquartile range over all simulations M is then denoted by $iqr := quartile(\hat{B}^m, 0.75) - quartile(\hat{B}^m, 0.25)$, which is a $n \times n$ matrix containing the interquartile range over all simulations for each element $\hat{B}_{ij}$. The average of the interquartile range of the elements on the diagonal/upper-/lower- triangular of $\hat{B}$ is then the average of all elements on the diagonal/upper-/lower- triangular of iqr.

Figure 2.7: Finite sample performance - Rejection rates (one-step estimators).

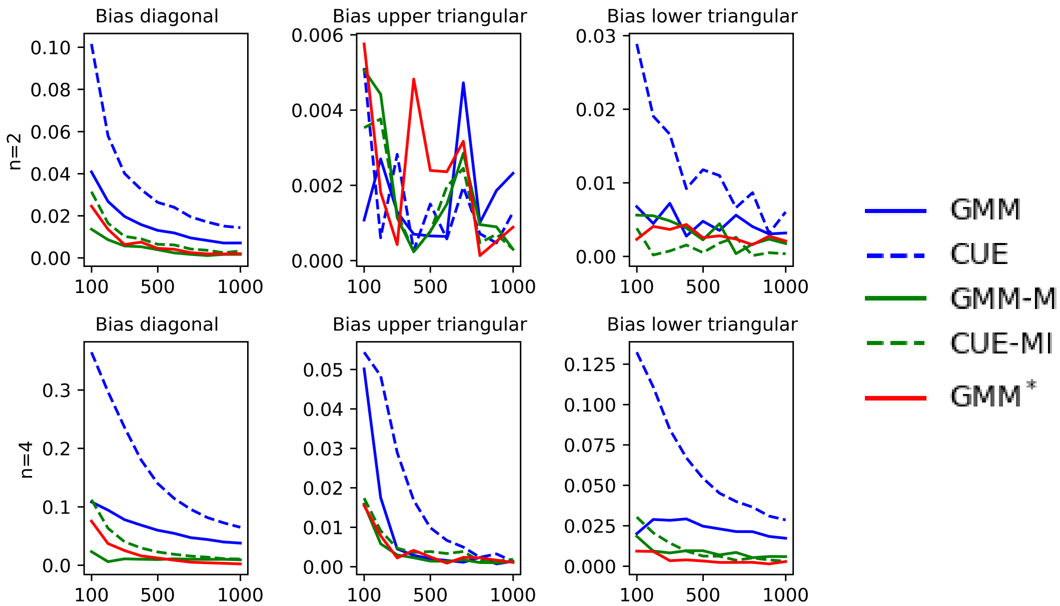


Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, ..., 1000$. The figure shows the rejection rates at $\alpha = 10\%$ for J-Test, recursive SVAR Wald test, and Wald test with $H_0 : B_{1,n} = 0$ for $n = 2$ and $n = 4$.

2.7.2 Appendix - Simulation with t-distributed shocks

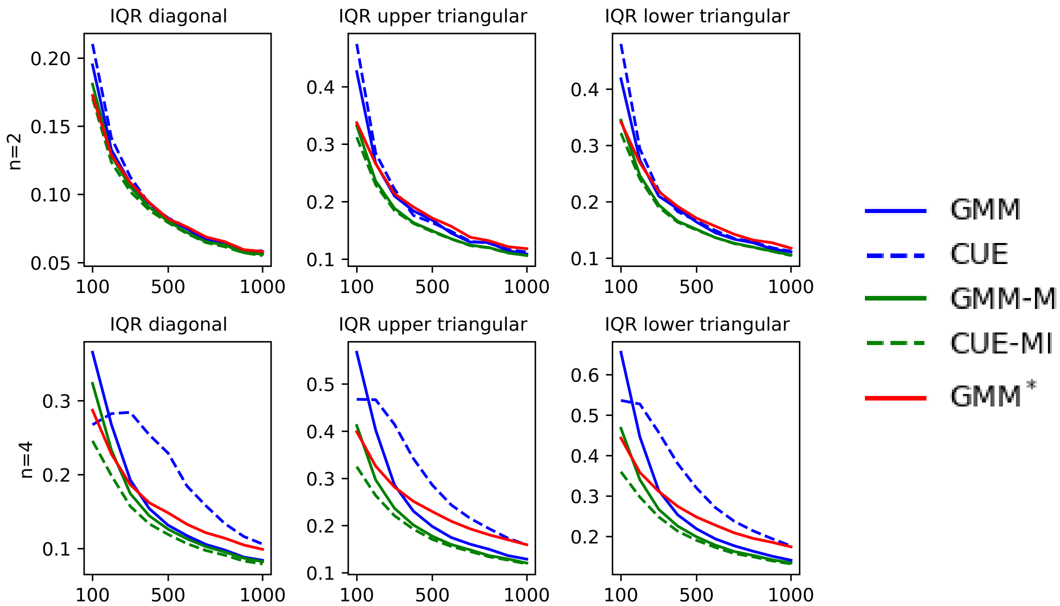
This section shows analogous results to the Monte Carlo simulation in Section 2.5. However, the structural shocks are drawn from a t-distribution with seven degrees of freedom. Additionally, the shocks have been normalized to unit variance by multiplying each shock with $1/\sqrt{(v/(v-2))}$ and $v = 7$.

Figure 2.8: Finite sample performance - Bias of the estimated elements (t-distribution).



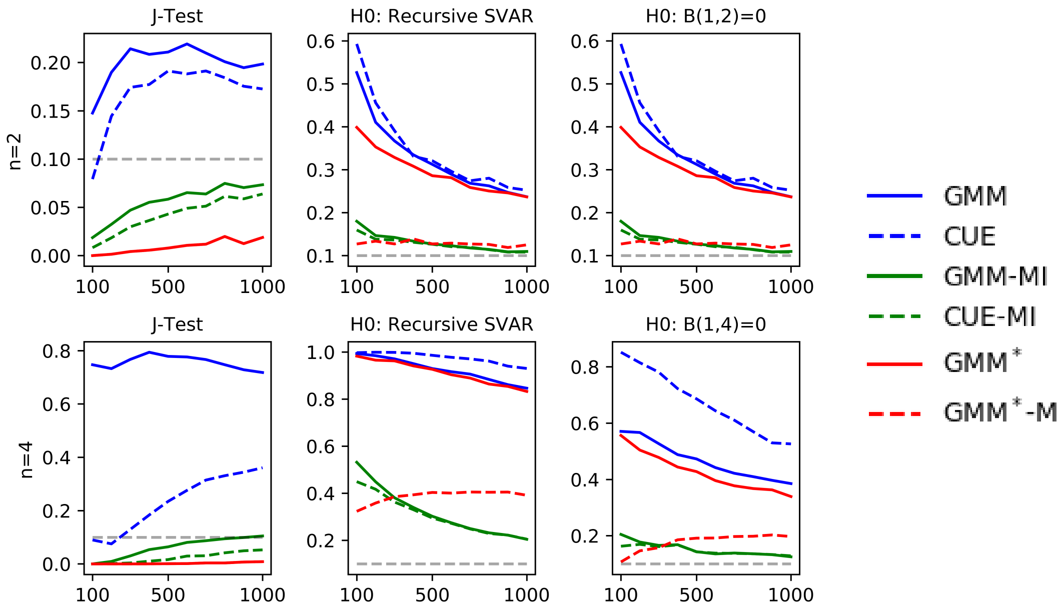
Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, ..., 1000$. The figure shows the average of the median absolute bias of the elements on the diagonal/upper-/lower- triangular of $\hat{B}$. Let $\hat{B}^m$ be the estimator in one simulation m and $1 \leq m \leq M$. Let $\bar{B}^m := abs(\hat{B}^m - B_0)$ be the absolute bias in simulation m. The median absolute bias over all simulations M is then denoted by $bias := med(\bar{B}^m)$, which is a $n \times n$ matrix containing the median absolute bias over all simulations for each element $\hat{B}_{ij}$. The average of the median absolute bias of the elements on the diagonal/upper-/lower- triangular of $\hat{B}$ is then the average of all elements on the diagonal/upper-/lower- triangular of *bias*.

Figure 2.9: Finite sample performance - Interquartile range of the estimated elements (t-distribution).



Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, ..., 1000$. The figure shows the average of the interquartile range (IQR) of the elements on the diagonal/upper-/lower- triangular of $\hat{B}$. Let $\hat{B}^m$ be the estimator in one simulation m and $1 \leq m \leq M$. The interquartile range over all simulations M is then denoted by $iqr := quartile(\hat{B}^m, 0.75) - quartile(\hat{B}^m, 0.25)$, which is a $n \times n$ matrix containing the interquartile range over all simulations for each element $\hat{B}_{ij}$. The average of the interquartile range of the elements on the diagonal/upper-/lower- triangular of $\hat{B}$ is then the average of all elements on the diagonal/upper-/lower- triangular of iqr.

Figure 2.10: Finite sample performance - Rejection rates (t-distribution).



Monte Carlo simulation with $M = 5000$ iterations and sample sizes $T = 100, 200, ..., 1000$. The figure shows the rejection rates at $\alpha = 10\%$ for J-Test, recursive SVAR Wald test, and Wald test with $H_0 : B_{1,n} = 0$ for $n = 2$ and $n = 4$.

3 Monetary Policy and the Stock Market - A Partly Recursive SVAR Estimator[13]

3.1 Introduction

Simultaneously identifying monetary policy and stock market shocks in a structural vector autoregression (SVAR) is an ongoing challenge for econometricians. Identifying both shocks requires to impose an a priori structure. Most of the literature covers one of two extreme cases: I) identifying all shocks based on restrictions concerning the short- or long-run interaction, or II) data-driven approaches without restrictions, but based on heteroskedastic or non-Gaussian shocks. We argue that neither of the two extreme cases is suited for the application at hand. In particular, we show that commonly used short- and long-run restrictions on the interaction of monetary policy and the stock market are not available. Additionally, purely data-driven estimators depend on latent, volatile, or hardly observable features of the data, which leads to biased and volatile estimates and distorted test statistics in small samples and reasonably large SVAR models.

The estimator proposed in this study combines the traditional identification approach based on restrictions with the more recent data-driven approach based on non-Gaussianity. Our estimator allows the researcher to rely on recursiveness restrictions if possible and to be agnostic on the interaction of the variables and rely on data-driven estimates when necessary. The estimator is applied to analyze the interaction of monetary policy and the stock market. We find evidence against commonly used short- and long-run restrictions and demonstrate that a purely data-driven estimator leads to imprecise estimates.

In the literature, the interaction of monetary policy and the stock market has been estimated based on short-run restrictions (see e.g. Laopodis (2013)) and long-run restrictions (see Bjørnland and Leitemo (2009) or Kontonikas and Zekaite (2018)). We argue that neither the short- nor the long-run restrictions are plausible. In particular, stock market shocks can contain news about future business cycle fluctuations (see e.g. Beaudry and Portier (2006)) and assuming that the central bank does not react simultaneously to these shocks is, at least with low frequency data,

[13]joint work with Andre Seepe, a slightly different version of the chapter appeared as Keweloh, S.A. and A. Seepe (2020), *Monetary policy and the stock market - A partly recursive SVAR estimator*, SFB 823 Discussion Paper series No. 32/2020, https://doi.org/10.17877/DE290R-21722

debatable. Moreover, recent studies (see for instance Moran and Queralto (2018), Bianchi et al. (2019) and Jordà et al. (2022)) find evidence against the long-run neutrality of monetary policy, which casts doubt on long-run restrictions used to identify monetary policy shocks.

Due to the unavailability of short- and long-run restrictions, several authors use data-driven approaches to estimate the interaction of monetary policy and the stock market. These approaches do not require any restrictions on the interaction of the variables, but instead exploit a structure imposed on the statistical properties of the shocks. Lütkepohl and Netšunajev (2017) estimate the interaction of monetary policy and the stock market based on time-varying volatility and find a negative impact of a tightening of monetary policy on stock prices. However, the authors are unable to clearly label a stock market shock. Moreover, a tightening of monetary policy appears to have an unexpected initial positive impact on output and inflation and therefore even the labeling of the monetary policy shock is debatable. Lanne et al. (2017) estimate a SVAR based on non-Gaussianity and find that a tightening of monetary policy has an immediate negative impact on financial conditions. However, they are also unable to label any other shock and in particularly cannot label a stock market shock.

Neither the traditional restriction based approaches nor the more recent purely data-driven approaches yield conclusive insights into the interaction of monetary policy and the stock market. The restriction based approaches fail due to the unavailability of sufficiently many short- or long-run restrictions and the data-driven approaches fail, since they impose such little structure that finite sample estimates become volatile, up to the point that it becomes difficult to even label the shocks.

We argue that the key to gain insight into the interaction of monetary policy and the stock market is a combination of the traditional restriction based and the more recent data-driven approach. In particular, the estimator proposed in this study allows to divide the variables of the SVAR into a first block of recursively ordered variables and a second block of non-recursive variables. Only the non-recursive block relies on data-driven estimates based on non-Gaussian and independent shocks. The more recursiveness restrictions the researcher applies, the less the estimator depends on moments beyond the variance. In a Monte Carlo simulation we show how the performance of a purely data-driven estimator based on non-Gaussianity deteriorates with a decreasing sample size and an increasing model size. However, the simulation also shows that exploiting the partly recursive order can stop the performance decline. Therefore, the estimator proposed in this study allows the researcher to rely on all available recursiveness restrictions, which reduces the dependence of the estimator on moments beyond the variance and thereby increases the finite sample performance of the estimator.

In our application the variables output, investment, and inflation are assumed to be rigid and are restricted such that they cannot respond to stock market and monetary policy shocks within the same quarter. However, interest rates and stock returns remain unrestricted and can simultaneously respond to all shocks. We find that stock prices immediately decrease in response to a tightening of monetary policy. Moreover, output, investment and stock prices show a persistent negative reaction to monetary policy shocks. Given standard confidence bands we reject the null hypothesis of a long-run zero effect of a monetary policy shock on stock prices. We thus find evidence against the long-run restrictions used in Bjørnland and Leitemo (2009). In contrast to the fully data-driven approaches in the literature we are able to label a stock market shock. We find that positive stock market shocks behave like news shocks and indicate a future business cycle expansion and the central bank reacts immediately with a tightening of monetary policy.

The remainder of this article is structured as follows. Section 3.2 shows that commonly used identification schemes in the related literature come with caveats that render them not applicable to analyze the interaction of monetary policy and the stock market. Section 3.3 derives our estimator for partly recursive, partly non-Gaussian SVAR models and contains a Monte-Carlo study illustrating how exploiting the partly recursive order increases the finite sample performance of the estimator. In Section 3.4 we use the proposed partly recursive, partly non-Gaussian SVAR estimator to analyze the interaction of the stock market and monetary policy. Section 3.5 concludes.

3.2 Monetary policy and the stock market

3.2.1 The unavailability of common identifying restrictions

In this section, we use a simple asset pricing model to illustrate that there is no indisputable answer about the short- and long-run interaction of stock markets and monetary policy. We keep the model intentionally simple to show that only a small deviation in basic assumptions can cast common short- or long-run restrictions inappropriate.

Consider that households can save by buying firm stocks of firm i at price $v_{i,t}$, yielding dividend $d_{i,t+1}$ in the next period or by a non-contingent bond b_t^f yielding a guaranteed real interest at rate r_t. The no-arbitrage condition then is

$$1 + r_t = E_t \frac{v_{i,t+1} + d_{i,t+1}}{v_{i,t}}. \tag{3.1}$$

From this, one can acquire the central asset pricing equation of the form

$$v_{i,t} = E_t \sum_{s=1}^{\infty} \frac{d_{i,t+s}}{\prod_{j=1}^{s}(1 + r_{t+j-1})}, \tag{3.2}$$

so the current stock price is the expected discounted sum of future dividends. On the firm side, assume a continuum of infinitely small firms with mass 1 and dividends of firm i are given by

$$d_{i,t+s} = y_{i,t+s} - j_{i,t+s} + b^f_{i,t+1+s} - (1 + r_{t+s-1})b^f_{i,t+s} - \bar{w}\bar{n}, \tag{3.3}$$

where $y_{i,t}$ is output, $j_{i,t}$ investment in the physical capital stock, $b^f_{i,t}$ are debt sales (where $\int_0^1 b^f_{i,t} di = b^f_t$), $\bar{w}$ the constant real wage and $\bar{n}$ labor input, also assumed constant for simplicity. We assume further an accumulation of physical capital $k_{i,t}$ of the form

$$k_{i,t+1} = (1 - \delta)k_{i,t} + j_{i,t}, \quad \delta \in (0, 1). \tag{3.4}$$

The production function reads

$$y_{i,t} = A k^{\alpha}_{i,t} (Z_t \bar{n})^{1-\alpha}, \quad \alpha \in (0, 1), \tag{3.5}$$

with A an exogenous scaling factor and Z_t an aggregate productivity factor exogenous to the individual firm. Consequently, the firm maximization problem reads

$$\max_{\left\{k_{i,t+1+s}, b^f_{i,t+s}\right\}} \sum_{s=0}^{\infty} E_t \Lambda_{t+s} d_{i,t+s}, \tag{3.6}$$

with Λ_t the firm's discount factor and subject to (3.4)-(3.5). The optimality conditions yield the common interest rate parity condition of the form

$$E_t A \alpha k_{i,t+1}^{\alpha-1} (Z_{t+1}\bar{n})^{1-\alpha} + (1-\delta) = 1 + r_t, \tag{3.7}$$

which says that in equilibrium the interest rate on foreign capital and the return on capital investment will coincide. Now inserting (3.3)-(3.5) into (3.2) yields

$$v_{i,t} = E_t \sum_{s=1}^{\infty} \frac{A k_{i,t+s}^{\alpha} (Z_{t+s}\bar{n})^{1-\alpha} - k_{i,t+s+1} + (1-\delta) k_{i,t+s} + b_{i,t+1+s}^{f} - (1 + r_{t+s-1}) b_{i,t+s}^{f} - \bar{w}\bar{n}}{\prod_{j=1}^{s} (1 + r_{t+j-1})}. \tag{3.8}$$

Imposing the limiting condition $\lim_{T \to \infty} b_T = 0$ then leads to future debt sales dropping out from the asset pricing equation, as dividends cannot be debt-financed indefinitely. As becomes evident, the dynamics of the numerator are then entirely driven by the evolution of capital. Using Equation (3.7) then allows to find the evolution of capital as

$$k_{i,t+1} = E_t \left[\left(\frac{\alpha A}{r_t + \delta} \right)^{\frac{1}{1-\alpha}} \bar{n} Z_{t+1} \right]. \tag{3.9}$$

Now consider that the real interest rate increases once such that $r_t' > r_t^*$ and for the rest of the time $r_{t+s}' = r_{t+s}^*, \forall s > 0$ (primes denote variables after the shock, asterisks variables without the shock). The resulting response of dividends and stock prices now crucially depends on what we assume about the productivity factor Z_t:

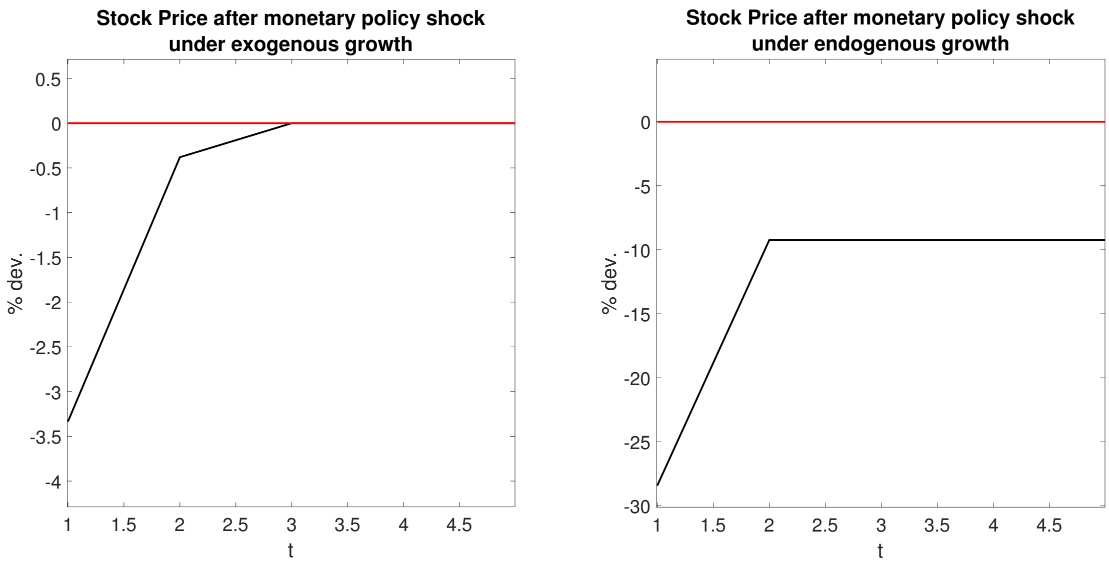
1.) Exogenous growth: Assume a neoclassical growth model with decreasing marginal returns to capital, so Z_t is some exogenously growing variable.

2.) Endogenous growth: Assume an endogenous growth model, for instance a standard learning-by-doing technology with $Z_t = \int_0^1 k_{i,t-1} di = K_{t-1}$.

Figure 3.1 shows the effect of an interest rate shock on stock prices for an exogenous and endogenous growth model[14]. Assuming sticky prices, thus nominal and real variables move in the same direction in the short-run, we can interpret the exogenous real interest rate increase as equivalent to a monetary policy shock. Both models imply an immediate reaction of stock prices

[14]For simplicity we assume $\bar{n} = 1$, the initial debt $b_t^f = 0$, $\bar{w} = 0$ and use a standard calibration of $\alpha = \frac{1}{3}$, $\delta = 0.1$ and setting $A = 0.46$ to ensure a long-run output growth rate of about 3%.

Figure 3.1: Simulated response of real stock prices to a one-time exogenous interest rate increase of about one percentage point as implied by the exogenous and endogenous growth model.



to the monetary policy shock. However, in the exogenous growth model with decreasing returns to capital, stock prices revert back to their long-run level, while under endogenous growth with the learning-by-doing technology, the decrease in stock prices is permanent. This is because in the first case the lower capital stock implies a higher marginal return of capital in the future, which drives back capital to its old steady state, while in the second case it does not, because the lower aggregate capital stock implies lower capital investment return for the individual firm.

Furthermore, interpret a stock price shock as news about higher future productivity that is not realized today like in Beaudry and Portier (2006). For instance assume A is no longer a constant, but time dependent. Assume now that in the next period $A'_{t+1} > A^*_{t+1}$. From Equation (3.8) it becomes evident that an increase in future dividends leads to an increase in stock prices now. Because of $A'_{t+1} > A^*_{t+1}$, we also know that $y'_{t+1} > y^*_{t+1}$. A central bank aiming at flattening business cycle fluctuations would immediately adjust its policy rate. Consequently, stock prices will contemporaneously react to monetary policy shocks, as will monetary policy to stock market shocks.

Now the task of the econometrician would be to let the data decide, which of the two theoretical approaches is correct. Of course, we need to make some assumptions to identify the structural shocks. However, we know that a monetary policy shock will immediately influence stock prices and vice versa, so we cannot impose a short-run restriction. Imposing a long-run restriction means to ex ante decide that the model with decreasing returns is the right one and not the endogenous growth model and strips us of the ability to let the data decide. In particular, our application later on shows that indeed the endogenous growth model is favored by the data and not the exogenous growth one. Thus we are in need of a data-driven identification approach, which is the objective of the present paper.

3.2.2 Monetary policy and the stock market SVAR models

As a second step we review the approaches of the related literature to estimate the interaction of monetary policy and the stock market in a SVAR. We show that there is a lack of a compelling estimation approach that is both feasible, but not too restrictive for the problem at hand.

In a SVAR a vector of time series is explained by its past values and a linear combination of structural shocks

$$y_t = A_1 y_{t-1} + ... + A_p y_{t-p} + u_t, \tag{3.10}$$

$$u_t = B\varepsilon_t, \tag{3.11}$$

with an n-dimensional vector of macroeconomic variables y_t, parameter matrices $A_1, ..., A_p$, a non-singular matrix B, the n-dimensional vector of structural shocks ε_t and the n-dimensional vector of reduced form shocks u_t. Here, the vector of structural shocks will contain a monetary policy and a stock market shock. The goal is to identify both shocks and estimate their impact on the macroeconomic variables. The SVAR imposes only little a priori structure, however, without further assumptions the structural shocks are not identified.

In general, the probably most frequently used identifying assumption is a recursive ordering, meaning zero restrictions on the impact of some shocks, such that each variable is simultaneously only influenced by shocks ordered in rows below the variable. However, in the case of monetary policy and the stock market zero restrictions on the interaction of both variables are hardly credible. In particular, stock prices can contain news about future productivity, see Beaudry and Portier (2006). Therefore, a positive stock price shock might indicate a future boom accompanied by inflationary pressure and a stabilizing central bank would respond immediately. Consequently, a zero restriction on the response of monetary policy to stock market shocks is difficult to defend. Nevertheless, zero restrictions on the interaction of monetary policy and the stock market have been used to estimate the interaction of both variables, see e.g. Laopodis (2013). However, these estimates only reflect the interaction of monetary policy and the stock market if the identifying assumptions are correct, which is at best questionable.

Due to the unavailability of credible short-run restrictions on the interaction of monetary policy and the stock market, several authors identify the shocks based on restrictions on the long-run interaction of both variables (see e.g. Bjørnland and Leitemo (2009) or Kontonikas and Zekaite (2018)). In particular, the authors assume long-run neutrality of monetary policy, meaning the monetary policy shock by construction has no long-run impact on real stock prices. Bjørnland and Leitemo (2009) find that monetary policy and the stock market interact simultaneously. In particular, a tightening of monetary policy leads to an immediate decrease of stock prices and a positive stock market shock leads to an immediate tightening of monetary policy. Again, these results only reflect the true interaction of both variables if the identifying long-run restriction is correct. In contrast to the short-run restriction used in Laopodis (2013), the long-run restriction used in Bjørnland and Leitemo (2009) is at least based on an underlying theory yielding long-run neutrality of monetary policy. However, as shown in Section 3.2.1, a slight modification of the theory from exogenous to endogenous growth already breaks the long-run neutrality result. In fact, recent studies (see e.g. Moran and Queralto (2018), Bianchi et al. (2019) and Jordà et al. (2022)) consistently find that monetary policy affects real variables much longer than usually assumed[15]. These results cast doubt on the long-run restriction and the corresponding estimated interaction of monetary policy and the stock market.

[15] For instance, Moran and Queralto (2018) and Bianchi et al. (2019) find that the impulse response of TFP is significantly positive even 15 years after a negative monetary policy shock has hit the economy. Again as in the previous section, higher productivity goes hand in hand with higher expected dividends. Therefore, stock prices should not only decrease immediately, but permanently in response to an unexpected tightening of monetary policy.

Rigobon and Sack (2004) propose an estimator, which does not require any restrictions on the short- or long-run interaction of the stock market and monetary policy. Instead, it is based on heteroskedastic shocks and requires to a priori specify periods of different variances of the monetary policy shock. The identification is thus based on a stochastic property of the structural shocks and not on a restriction on the impact of the shocks. Specifying volatility regimes of monetary policy may be straight-forward on a daily basis (by choosing all days with FOMC announcements), however, with lower frequency data it becomes increasingly difficult. Therefore, the estimator becomes infeasible in a typical macroeconomic application with monthly, quarterly or even lower frequency data.

In general, identification based on time-varying volatility does not require to a priori specify volatility periods (see e.g. Rigobon and Sack (2003), Lanne et al. (2010), Lütkepohl and Netšunajev (2017) or Lewis (2021)). In fact, a latent volatility process can be used for identification without imposing much structure on the latent process. However, Lütkepohl and Netšunajev (2017) argue that reliable estimators based on GARCH or Markov switching processes are only available in small models and few volatility states. The intuition is simple: The more (correct) structure we impose on the latent process, the more precise the corresponding estimate. Therefore, Lütkepohl and Netšunajev (2017) propose an estimator which imposes a parametric smooth transition function between two states of the variance-covariance matrix of the reduced form shocks. The estimator is applied to analyze the interaction of monetary policy and the stock market. The authors find a small simultaneous negative response of the stock market to a tightening of monetary policy. However, a tightening of monetary policy is also found to lead to an initial increase of inflation and output. Due to the counterintuitive response of output and inflation to the shock, the authors admit that labeling the shock as a monetary policy shock in a "conventional" sense may be misleading. Additionally, the authors cannot label a stock market shock and hence it remains unclear how monetary policy reacts to a stock market shock.

Another branch of the SVAR literature uses non-Gaussian and independent shocks for identification (see e.g. Lanne et al. (2017), Gouriéroux et al. (2017), Lanne and Luoto (2021), Guay (2021) and Keweloh (2021b)). Theses approaches are also data-driven and do not require to impose any short- or long-run restrictions. Instead, these approaches require that the structural shocks are mutually independent and at most one shock is allowed to be Gaussian. Intuitively, non-Gaussian shocks do contain information in moments beyond the variance, which allows to identify the simultaneous interaction. In a short application Lanne et al. (2017) use a data-driven identification approach imposing non-Gaussian and independent shocks to estimate the interdependence of monetary policy and the stock market. The authors find that an unexpected tightening of monetary policy has an immediate negative impact on financial conditions. However, they are unable to label a stock market shock. Therefore, it again remains unclear how stock market shocks influence monetary policy.

To sum up, the commonly used short- and long-run restrictions regarding the interaction of monetary policy and the stock market have implications on the underlying data generating process. Until now there is no consensus about which theoretical model is correct and the estimation should not depend on an a priori restriction to one or another model, but rather be able to decide which fits the data best. On the other side, there are identification approaches that do not rely on short- or long-run restrictions, but they are either not able to be generalized to a broader macroeconomic setup or become less feasible the more variables are included into the VAR. Ideally, the SVAR estimator should allow to factor in a priori restrictions that we are certain about, but also allow a data-driven identification, when we are not certain about the underlying theory. In the following section we propose an estimator that fulfills these criteria.

3.3 A partly recursive, partly non-Gaussian SVAR estimator

A non-Gaussian SVAR with independent shocks can be estimated based on restrictions governing the interaction of the variables or based on information contained in moments beyond the variance and without any assumptions on the interaction of the variables. At first glance, in a non-Gaussian SVAR and from an asymptotic point of view, the traditional identification approach based on restrictions appears to be unnecessarily restrictive. However, we show that in small samples the performance of data-driven estimators based on non-Gaussianity quickly deteriorates with an increasing model size, while the performance of a restriction based estimator is less affected by the model and sample size. In macroeconomic applications, we can oftentimes derive some credible restrictions based on economic theory. However, in many cases we cannot derive sufficiently many restrictions to identify the SVAR based on second moments and the researcher is forced to rely on additional, less credible restrictions or to use an unreliable purely data-driven estimator.

The estimator proposed in this section combines the traditional restriction based approach with the more recent data-driven approach based on non-Gaussianity. Our estimator allows the researcher to rely on recursiveness restrictions if possible and to be agnostic on the interaction of the variables and relying on data-driven estimates when necessary. In particular, the proposed estimator allows to order some, but not all, shocks recursively. While the impact of the recursive shocks is estimated based on second moments, the impact of the non-recursive shocks is estimated based on non-Gaussianity. We show that in comparison to an unrestricted estimator solely based on non-Gaussian and independent shocks, exploiting the partly recursive structure i) improves the finite sample performance of the estimator, ii) reduces the burden of labeling the shocks, and iii) relaxes the non-Gaussianity and independence assumptions.

3.3.1 Derivation of the estimator

Consider a partly recursive SVAR, meaning there exists $m \in \mathbb{N}$ with $0 \leq m \leq n$ and

$$B = \begin{bmatrix} b_{11} & 0 & & \dots & & 0 \\ \vdots & \ddots & \ddots & & & \vdots \\ b_{m1} & \dots & b_{mm} & 0 & \dots & 0 \\ b_{m+1,1} & \ddots & b_{m+1,m} & b_{m+1,m+1} & & b_{m+1,n} \\ \vdots & & & \vdots & & \vdots \\ b_{n1} & \dots & b_{nm} & b_{n,m+1} & & b_{nn} \end{bmatrix}. \tag{3.12}$$

Therefore, the first m variables are ordered recursively, meaning they cannot contemporaneously be influenced by structural shocks in rows ordered below. However, the last $n - m$ variables are not ordered recursively and can contemporaneously be influenced by all structural shocks. Since the matrix B is only partly recursive, it cannot be identified solely by second moments. However, the partly recursive structure can be combined with estimators based on independent and non-Gaussian shocks.

The partly recursive SVAR can be estimated in three steps. For simplicity, consider a SVAR with four variables and the following partly recursive structure

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} b_{11} & 0 & 0 & 0 \\ b_{21} & b_{22} & 0 & 0 \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix}. \tag{3.13}$$

The recursive part can be written as

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}, \tag{3.14}$$

which is a simple recursive SVAR and can be identified and estimated based on second moments (e.g. by applying the Cholesky decomposition to the variance-covariance matrix of the reduced form shocks, see Kilian and Lütkepohl (2017)). The non-recursive part can be written as

$$\begin{bmatrix} u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} b_{31} & b_{32} \\ b_{41} & b_{42} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} + \begin{bmatrix} \nu_3 \\ \nu_4 \end{bmatrix}, \tag{3.15}$$

with the adjusted reduced form shocks

$$\begin{bmatrix} \nu_3 \\ \nu_4 \end{bmatrix} = \begin{bmatrix} b_{33} & b_{34} \\ b_{43} & b_{44} \end{bmatrix} \begin{bmatrix} \varepsilon_3 \\ \varepsilon_4 \end{bmatrix}. \tag{3.16}$$

Using the estimated structural shocks $\hat{\varepsilon}_1$ and $\hat{\varepsilon}_2$ from the first step allows to estimate the lower-left block of B in Equation (3.15) by OLS. The adjusted reduced form shocks ν in Equation (3.15) represent the variation in u_3 and u_4, which is unexplained by the structural shocks in the recursive block. Therefore, the adjusted reduced form shocks only contain information from shocks in the non-recursive block. However, Equation (3.16) is just an unrestricted SVAR in the two structural shocks from the non-recursive block. The identification results and estimators based on non-Gaussianity proposed by Lanne et al. (2017), Gouriéroux et al. (2017), Lanne and Luoto (2021), Keweloh (2021b), or Guay (2021) can then be used to identify and estimate the non-recursive block if the structural shocks in the non-recursive block are mutually independent and at most one shock is Gaussian. Note that the different estimators require slightly different assumptions on the non-Gaussianity and independence of the structural shocks in the non-recursive block.

The partly recursive, partly non-Gaussian estimator can also be calculated in a single step. For example, a partly recursive version of the GMM estimators proposed in Keweloh (2021b) and Keweloh (2021a) can be obtained by including the second-order moment conditions of all shocks and the higher-order moment conditions associated with the shocks in the non-recursive block. Some estimators based on non-Gaussianity rely on an initial whitening step, see e.g. the PML estimator proposed in Gouriéroux et al. (2017) or the whitened GMM estimators proposed in Keweloh (2021b). In the preliminary whitening step the reduced form shocks are transformed into uncorrelated shocks with unit variance and in the second step the optimization is performed over orthogonal matrices, which correspond to rotations of the transformed reduced form shocks[16]. Whitening is equivalent to an optimization subject to the constraint that the estimated structural shocks are uncorrelated with unit variance in the given sample, compare Keweloh (2021b). However, in the partly recursive SVAR defined in Equation (3.12), the first m columns of B are uniquely determined by the whitening constraint, imposing that the estimated structural shocks have to be uncorrelated with unit variance. Therefore, a whitened estimator with partly recursive constraints by definition only relies on second moments to identify and estimate the impact of the shocks in the recursive block, see Appendix 3.6.1 for more details.

The partly recursive SVAR estimation approach presented above does not require independence or non-Gaussianity of the structural shocks in the recursive block. Additionally, if we are only interested in the impact of the shocks in the non-recursive block, the partly recursive SVAR estimation approach is robust to misspecifications of the recursive order. In particular, if the

[16]Optimizing over orthogonal matrices is computationally simple, since it can be pulled back to an optimization problem over the euclidean space, see Lezcano-Casado and Martınez-Rubio (2019). In Appendix 3.6.1 we propose a similar transformation for the optimization problem over orthogonal matrices with partly recursive constraints.

SVAR is only block recursive, such that there exists $m \in \mathbb{N}$ with $0 \leq m \leq n$ and

$$B = \begin{bmatrix} b_{11} & ... & b_{m1} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots & & \vdots \\ b_{m1} & \dots & b_{mm} & 0 & \dots & 0 \\ b_{m+1,1} & \ddots & b_{m+1,m} & b_{m+1,m+1} & & b_{m+1,n} \\ \vdots & & & \vdots & & \vdots \\ b_{n1} & \dots & b_{nm} & b_{n,m+1} & & b_{nn} \end{bmatrix}, \tag{3.17}$$

the partly recursive SVAR estimation approach proposed above yields inconsistent estimates for the upper-left and lower-left block of B, but remains consistent for the lower-right block which represents the impact of the shocks in the non-recursive block. To see this, note that falsely imposing a recursive order in Equation (3.14) yields inconsistent estimates of the upper-left block of B. Additionally, using the shocks of the first step, here $\hat{\varepsilon}_1$ and $\hat{\varepsilon}_2$, to estimate Equation (3.15) will also yield inconsistent estimates of the lower-left block of B. However, if the shocks in the non-recursive block, here ε_3 and ε_4, have no simultaneous impact on the variables in the recursive block, the shocks $\hat{\varepsilon}_1$ and $\hat{\varepsilon}_2$ obtained from the first step are equal to a linear combination of the true shocks ε_1 and ε_2. Therefore, the adjusted reduced form shocks ν in Equation (3.15) still represent the variation in u_3 and u_4 which is unexplained by the structural shocks in the recursive block and hence, the non-recursive SVAR in Equation (3.16) remains valid. The proposed estimator thus allows to identify and consistently estimate the impact of a non-recursive block of variables if Equation (3.17) holds, meaning that all the shocks in the non-recursive block have no simultaneous impact on the variables in the recursive block of variables, and the shocks in the non-recursive block satisfy the independence and non-Gaussianity assumptions.

Exploiting the partly recursive structure yields several advantages compared to an unrestricted estimator solely identified by independence and non-Gaussianity. First, the Monte Carlo study in Section 3.3.2 shows that exploiting the partly recursive order and thus decreasing the dependence of the estimator on higher moments leads to an increase of the small sample performance of the estimator. Second, every identification approach requires to impose an a priori structure. In particular, if no restrictions on the interaction of the variables are imposed, the researcher has to impose that all shocks are independent and at most one shock is allowed to be Gaussian. Both assumptions have bee criticized in the literature, see, e.g., Kilian and Lütkepohl (2017, Chapter 14) or Olea et al. (2022). The partly recursive estimation approach decrease the dependence on these non-Gaussianity and independence assumptions. Third, a data-driven identification scheme based only on non-Gaussian and independent shocks only identifies the shocks up to labeling. Therefore, the researcher has to decide which impulse response belongs to which shock. The task of labeling the shocks becomes increasingly difficult the more shocks are identified by this procedure, especially if the impulse responses of the variables are quite similar with respect to two or more shocks. Imposing a partly recursive structure alleviates this burden on the econometrician, since the shocks in the recursive block are already labeled by the identifying assumptions of the partly recursive order.

3.3.2 Finite sample performance

In the following Monte Carlo study, we show that data-driven estimators based on non-Gaussianity suffer from a curse of dimensionality, i.e. the bias and mean squared error (MSE) increases quickly with an increasing model size and a decreasing sample size. However, we show that exploiting the partly recursive structure can stop the curse of dimensionality.

We first simulate a small SVAR with $n = 2$ variables and

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} 10 & 5 \\ 5 & 10 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}. \tag{3.18}$$

We then add two additional shocks ε_{3t} and ε_{4t} and obtain a SVAR with $n = 4$ variables with

$$\begin{bmatrix} u_{3t} \\ u_{4t} \\ u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} 10 & 0 & 0 & 0 \\ 5 & 10 & 0 & 0 \\ 5 & 5 & 10 & 5 \\ 5 & 5 & 5 & 10 \end{bmatrix} \begin{bmatrix} \varepsilon_{3t} \\ \varepsilon_{4t} \\ \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}. \tag{3.19}$$

With this setup, the impact of ε_{1t} and ε_{2t} on u_{1t} and u_{1t} is identical in both SVARs and given by the B matrix for $n = 2$ or the lower-right block of the B matrix for $n = 4$. Moreover, the structural shocks ε_{it}, $i = 1, \ldots, n$, $t = 1, \ldots, T$, are drawn independently and identically from the two-component mixture

$$\varepsilon_{it} \sim 0.79\, \mathcal{N}(-0.2, 0.7^2) + 0.21\, \mathcal{N}(0.75, 1.5^2),$$

where $\mathcal{N}(\mu, \sigma^2)$ indicates a normal distribution with mean μ and standard deviation σ. The shocks have skewness 0.9 and excess kurtosis 2.4.

In the small SVAR with $n = 2$ the reported estimator uses no restrictions and thus contains only a single non-recursive block. In the large SVAR with $n = 4$ one estimator uses no restrictions meaning all shocks are contained in a single non-recursive block. Additionally, for $n = 4$ a second estimator uses the restrictions from the partly recursive order such that the non-recursive block only contains ε_{1t} and ε_{2t}. The non-recursive block is estimated with the SVAR GMM estimator proposed by Keweloh (2021b) using all second- to fourth-order moment conditions corresponding to the shocks in the non-recursive block. Additionally, we update the weighting matrix continuously and use the assumption of serially and mutually independent shocks to estimate the asymptotically optimal weighting matrix and the asymptotic variance as proposed in Keweloh (2021a). In Appendix 3.6.2 we report analogous results for the PML estimator proposed by Gouriéroux et al. (2017) and the whitened fast GMM estimator proposed by Keweloh (2021b). Additionally, we show results for a similar simulation using t-distributed structural shocks. None of the conclusions drawn in this section is sensitive to the alternative simulations.

Table 3.1 shows the average and MSE of each estimated element. The simulations show how the performance in terms of bias and MSE of estimates based entirely on non-Gaussianity decreases with an increasing model size. In particular, the elements in the lower-right block of the unrestricted SVAR GMM estimator for $n = 4$ are more biased and have a large MSE compared corresponding elements of unrestricted SVAR GMM estimator for $n = 2$. This curse of dimensionality is more pronounced in smaller samples. However, the simulation also shows that exploiting the partly recursive structure stops the deterioration of the performance induced by a larger model. In particular, the elements in the lower-right block of the partly recursive SVAR GMM estimator for $n = 4$ exhibit a similar bias and MSE compared to corresponding elements of the unrestricted SVAR GMM estimator for $n = 2$.

Table 3.1: Finite sample performance - Partly recursive SVAR.

T	$n=2$ GMM	$n=4$ GMM	$n=4$ partly recursive GMM
100	$\begin{bmatrix} \underset{(2.04)}{9.73} & \underset{(4.52)}{4.67} \\ \underset{(4.21)}{4.97} & \underset{(2.5)}{9.55} \end{bmatrix}$	$\begin{bmatrix} \underset{(2.48)}{9.13} & \underset{(3.64)}{-0.02} & \underset{(3.71)}{0.06} & \underset{(4.08)}{-0.04} \\ \underset{(4.3)}{4.58} & \underset{(3.07)}{9.14} & \underset{(4.54)}{-0.0} & \underset{(4.85)}{-0.1} \\ \underset{(6.23)}{4.52} & \underset{(5.94)}{4.6} & \underset{(4.91)}{9.23} & \underset{(7.63)}{4.22} \\ \underset{(6.64)}{4.57} & \underset{(6.01)}{4.61} & \underset{(6.6)}{4.81} & \underset{(6.03)}{8.89} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.14)}{9.94} & . & . & . \\ \underset{(1.34)}{4.98} & \underset{(1.02)}{9.88} & . & . \\ \underset{(1.86)}{4.98} & \underset{(1.54)}{4.95} & \underset{(2.3)}{9.6} & \underset{(5.06)}{4.57} \\ \underset{(1.86)}{4.99} & \underset{(1.59)}{4.96} & \underset{(4.59)}{4.95} & \underset{(2.75)}{9.42} \end{bmatrix}$
250	$\begin{bmatrix} \underset{(0.86)}{9.86} & \underset{(1.58)}{4.94} \\ \underset{(1.64)}{4.93} & \underset{(0.75)}{9.86} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.61)}{9.68} & \underset{(1.51)}{0.0} & \underset{(1.6)}{-0.02} & \underset{(1.45)}{-0.05} \\ \underset{(1.67)}{4.83} & \underset{(0.94)}{9.68} & \underset{(2.0)}{0.02} & \underset{(1.95)}{-0.08} \\ \underset{(2.54)}{4.87} & \underset{(2.49)}{4.81} & \underset{(1.8)}{9.73} & \underset{(2.83)}{4.7} \\ \underset{(2.45)}{4.89} & \underset{(2.43)}{4.85} & \underset{(2.7)}{4.93} & \underset{(1.96)}{9.58} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.4)}{9.98} & . & . & . \\ \underset{(0.52)}{4.99} & \underset{(0.44)}{9.97} & . & . \\ \underset{(0.75)}{4.99} & \underset{(0.62)}{4.98} & \underset{(0.76)}{9.88} & \underset{(1.69)}{4.87} \\ \underset{(0.74)}{5.0} & \underset{(0.63)}{4.97} & \underset{(1.57)}{5.0} & \underset{(0.88)}{9.8} \end{bmatrix}$
500	$\begin{bmatrix} \underset{(0.32)}{9.94} & \underset{(0.63)}{4.95} \\ \underset{(0.6)}{4.98} & \underset{(0.34)}{9.93} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.24)}{9.87} & \underset{(0.58)}{0.01} & \underset{(0.61)}{0.0} & \underset{(0.62)}{-0.02} \\ \underset{(0.63)}{4.93} & \underset{(0.34)}{9.85} & \underset{(0.83)}{0.02} & \underset{(0.85)}{0.0} \\ \underset{(1.01)}{4.94} & \underset{(1.03)}{4.91} & \underset{(0.64)}{9.87} & \underset{(1.09)}{4.92} \\ \underset{(1.01)}{4.95} & \underset{(1.07)}{4.92} & \underset{(1.04)}{4.94} & \underset{(0.72)}{9.85} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.22)}{9.99} & . & . & . \\ \underset{(0.25)}{5.0} & \underset{(0.22)}{9.96} & . & . \\ \underset{(0.36)}{5.0} & \underset{(0.3)}{4.99} & \underset{(0.34)}{9.92} & \underset{(0.67)}{4.96} \\ \underset{(0.36)}{5.0} & \underset{(0.31)}{4.99} & \underset{(0.64)}{4.96} & \underset{(0.36)}{9.92} \end{bmatrix}$
1000	$\begin{bmatrix} \underset{(0.16)}{9.99} & \underset{(0.28)}{4.99} \\ \underset{(0.28)}{5.0} & \underset{(0.16)}{9.97} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.11)}{9.93} & \underset{(0.29)}{-0.0} & \underset{(0.3)}{-0.01} & \underset{(0.29)}{-0.01} \\ \underset{(0.31)}{4.98} & \underset{(0.17)}{9.93} & \underset{(0.39)}{0.01} & \underset{(0.39)}{0.0} \\ \underset{(0.49)}{4.99} & \underset{(0.5)}{4.95} & \underset{(0.32)}{9.95} & \underset{(0.51)}{4.96} \\ \underset{(0.47)}{4.99} & \underset{(0.5)}{4.96} & \underset{(0.5)}{4.99} & \underset{(0.32)}{9.94} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.11)}{9.99} & . & . & . \\ \underset{(0.12)}{5.0} & \underset{(0.11)}{9.99} & . & . \\ \underset{(0.18)}{5.01} & \underset{(0.16)}{5.0} & \underset{(0.16)}{9.97} & \underset{(0.3)}{4.97} \\ \underset{(0.18)}{5.0} & \underset{(0.16)}{5.0} & \underset{(0.3)}{5.01} & \underset{(0.16)}{9.97} \end{bmatrix}$
5000	$\begin{bmatrix} \underset{(0.03)}{10.0} & \underset{(0.06)}{4.99} \\ \underset{(0.06)}{5.0} & \underset{(0.03)}{9.99} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.02)}{9.98} & \underset{(0.05)}{0.0} & \underset{(0.05)}{-0.0} & \underset{(0.05)}{-0.0} \\ \underset{(0.06)}{4.99} & \underset{(0.03)}{9.99} & \underset{(0.07)}{-0.01} & \underset{(0.07)}{0.0} \\ \underset{(0.09)}{5.0} & \underset{(0.09)}{5.0} & \underset{(0.06)}{9.99} & \underset{(0.09)}{4.99} \\ \underset{(0.09)}{5.0} & \underset{(0.09)}{5.0} & \underset{(0.09)}{4.99} & \underset{(0.06)}{9.99} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.02)}{10.0} & . & . & . \\ \underset{(0.02)}{5.0} & \underset{(0.02)}{10.0} & . & . \\ \underset{(0.04)}{5.0} & \underset{(0.03)}{5.0} & \underset{(0.03)}{10.0} & \underset{(0.06)}{4.99} \\ \underset{(0.04)}{5.0} & \underset{(0.03)}{5.0} & \underset{(0.06)}{5.0} & \underset{(0.03)}{9.99} \end{bmatrix}$

Monte Carlo simulation with $M = 2000$ replications. The table shows the average, $1/M \sum_{m=1}^{M} \hat{b}_{ij}^{m}$, and the estimated mean squared error, $1/M \sum_{m=1}^{M} \left(\hat{b}_{ij}^{m} - b_{ij}\right)^2$, of each estimated element $\hat{b}_{ij}^{m}$ of b_{ij} denoting the element of B in row i and column j. The table reports results for the SVAR GMM estimator without restrictions for $n = 2$ and $n = 4$, and the partly recursive SVAR GMM estimator for $n = 4$, which uses zero restrictions highlighted by the dots.

Table 3.2: Finite sample performance - Hypothesis tests in the partly recursive SVAR.

		$n=2$ GMM	$n=4$ GMM	$n=4$ partly recursive GMM
$J-Test$	$T=100$	5.6	4.1	6.1
	$T=250$	7.5	4.8	6.9
	$T=500$	8.5	7.9	8.8
	$T=1000$	9.2	11.4	9.2
	$T=5000$	10.5	10.8	8.9
Wald $H_0: b=0$	$T=100$	85.8	70.4	83.6
	$T=250$	99.0	93.5	98.2
	$T=500$	99.9	99.6	99.8
	$T=1000$	100	100	100
	$T=5000$	100	100	100
Wald $H_0: b=5$	$T=100$	17.6	21.8	19
	$T=250$	15.0	18.1	14.6
	$T=500$	11.7	13.9	12.7
	$T=1000$	9.4	13.5	11.4
	$T=5000$	10.3	9.9	9.9

Monte Carlo simulation with $M=2000$ replications. The table shows the average the rejection rates at $\alpha=10\%$ for the J-Test, a Wald test with $H_0: b=0$, and Wald test with $H_0: b=5$, where b is equal to the element b_{12} for $n=2$ and b_{34} for $n=4$. The table reports results for the SVAR GMM estimator without restrictions for $n=2$ and $n=4$, and the partly recursive SVAR GMM estimator for $n=4$. We use the assumption of serially and mutually independent shocks to estimate the asymptotic variance required for the Wald tests as proposed in Keweloh (2021a).

Table 3.2 reports the rejection rates at $\alpha=10\%$ for the J-Test, a Wald test with $H_0: b=0$, and Wald test with $H_0: b=5$, where b is equal to the element b_{12} for $n=2$ and b_{34} for $n=4$, representing the impact of ε_{it} on u_{1t}. In small samples, the rejection rate of the J-Test is distorted for all estimators. We find a similar distortion of the unrestricted SVAR GMM estimator for $n=2$ and the partly recursive SVAR GMM estimator for $n=4$, while the distortion of the unrestricted SVAR GMM estimator for $n=4$ is larger. The Wald test with the null hypothesis $H_0: b=0$ tests the null hypothesis of a recursive SVAR order of ε_{1t} and ε_{2t}. The null hypothesis is incorrect and should be rejected. However, in small samples the unrestricted SVAR GMM estimator for $n=4$ rejects the null hypothesis less frequently than the unrestricted SVAR GMM estimator for $n=2$ and the partly recursive SVAR GMM estimator for $n=4$. The Wald test with the null hypothesis $H_0: b=5$ should only be rejected in 10% of all simulations. However, Table 3.2 shows that in small samples, all estimators reject the null hypothesis too often. The distortion of the two unrestricted estimators increases with the model size and imposing the partly recursive order decreases the distortion.

In macroeconomic applications, we oftentimes face relatively large models but only small samples with at best a few hundred observations. In this case, purely data-driven estimates based on non-Gaussianity become volatile and in a given application it can become difficult to draw any conclusions on the interaction of the variables or even label the shocks. However, econometricians have put much work into deriving and defending restrictions on the interaction of macroeconomic variables. The simulations show how including traditional zero restrictions increases the finite sample performance of a data-driven estimator based on non-Gaussianity. Therefore, we argue that in a given application the researcher should include restrictions when possible and only rely on a data-driven estimation when necessary.

3.4 The interdependence of U.S. monetary policy and the stock market

In this section, we apply the partly recursive estimation approach to analyze the interaction of monetary policy and the stock market. We consider a SVAR in five variables with quarterly U.S. data from 1983Q1 to 2019Q1 of the form

$$\begin{bmatrix} y_t \\ I_t \\ \pi_t \\ s_t \\ i_t \end{bmatrix} = \alpha + \gamma t + \sum_{i=1}^{4} A_i \begin{bmatrix} y_{t-i} \\ I_{t-i} \\ \pi_{t-i} \\ s_{t-i} \\ i_{t-i} \end{bmatrix} + \begin{bmatrix} u_t^y \\ u_t^I \\ u_t^\pi \\ u_t^s \\ u_t^i \end{bmatrix}, \tag{3.20}$$

where y denotes real output growth, I real investment growth, π the inflation rate, i the federal funds rate and s real stock returns[17]. The consideration of the variables follows from Section 3.2, where we constructed a model that either features long-run monetary neutrality or non-neutrality based on the role of physical capital concerning productivity. The inclusion of output, investment and stock prices allows to discriminate between the two models, as we would predict them to

[17]The inflation rate is defined as the quarter to quarter growth rate in the quarterly chain-type GDP price index retrieved from the FRED. The GDP growth rate is given by the quarterly log-difference of real GDP retrieved from the FRED. Real investment growth is given by the quarterly growth rate of real gross private domestic investment obtained from the FRED. The nominal interest rate is defined as the Federal Funds Rate (FFR), where the effective FFR (retrieved from FRED) is replaced by the shadow FFR provided by Wu and Xia (2016) for the Zero Lower Bound observations during the Great Recession. Stock returns are defined as the quarterly log-difference in real stock prices, where real stock prices are given by the S&P 500 index (retrieved from macrotrends.net) divided by the chain-type GDP price index.

either be influenced temporarily or permanently by a monetary policy shock. Furthermore, we include output, investment and stock prices in growth rates in order to check on the validity of potential long-run restrictions. The linear time trend is added to account for an eventual long-term decline in the interest rate as discussed by for instance Carvalho et al. (2016).

We assume that real investment growth, real output growth and inflation behave sluggishly, meaning they cannot react to monetary policy or stock market shocks within the same quarter. These restrictions can be justified by price rigidities and capital adjustment costs as oftentimes used in standard DSGE models, see for example Smets and Wouters (2007). However, interest rates and stock returns are unrestricted and can contemporaneously respond to all shocks. Therefore, we impose the following partly recursive structure

$$\begin{bmatrix} u_t^y \\ u_t^I \\ u_t^\pi \\ u_t^s \\ u_t^i \end{bmatrix} = \begin{bmatrix} b_{11} & 0 & 0 & 0 & 0 \\ b_{21} & b_{22} & 0 & 0 & 0 \\ b_{31} & b_{32} & b_{33} & 0 & 0 \\ b_{41} & b_{42} & b_{43} & b_{44} & b_{45} \\ b_{51} & b_{52} & b_{53} & b_{54} & b_{55} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_t^{sm} \\ \varepsilon_t^{mp} \end{bmatrix}, \tag{3.21}$$

where the non-recursive block contains the stock market shock ε_t^{sm} and the monetary policy shock ε_t^{mp}. Identification of the shocks in the non-recursive block requires that neither the stock market, nor the monetary policy shock simultaneously affect real output growth, real investment growth, or inflation. Importantly, identification of the non-recursive block is robust to misspecifications in the recursive block, meaning even if the data generating process is given by

$$\begin{bmatrix} u_t^y \\ u_t^I \\ u_t^\pi \\ u_t^s \\ u_t^i \end{bmatrix} = \begin{bmatrix} \tilde{b}_{11} & \tilde{b}_{12} & \tilde{b}_{13} & 0 & 0 \\ \tilde{b}_{21} & \tilde{b}_{22} & \tilde{b}_{23} & 0 & 0 \\ \tilde{b}_{31} & \tilde{b}_{32} & \tilde{b}_{33} & 0 & 0 \\ \tilde{b}_{41} & \tilde{b}_{42} & \tilde{b}_{43} & \tilde{b}_{44} & \tilde{b}_{45} \\ \tilde{b}_{51} & \tilde{b}_{52} & \tilde{b}_{53} & \tilde{b}_{54} & \tilde{b}_{55} \end{bmatrix} \begin{bmatrix} \tilde{\varepsilon}_{1t} \\ \tilde{\varepsilon}_{2t} \\ \tilde{\varepsilon}_{3t} \\ \varepsilon_t^{sm} \\ \varepsilon_t^{mp} \end{bmatrix}, \tag{3.22}$$

the partly recursive estimation approach based on the restrictions in Equation (3.21) is still able to identify the stock market shock and the monetary policy shock.

The SVAR is estimated by the partly recursive estimation approach with the restrictions imposed in Equation (3.21). The non-recursive block containing the monetary policy and stock market shock is estimated by the SVAR GMM estimator proposed in Keweloh (2021a) using two coskewness and three cokurtosis conditions. We update the weighting matrix continuously and use the assumption of serially and mutually independent shocks to estimate the asymptotically optimal weighting matrix and the asymptotic variance as proposed in Keweloh (2021a).

Table 3.3 shows the skewness, kurtosis and the p-value of the Jarque-Bera test of the adjusted reduced form shocks equal to

$$\begin{bmatrix} \nu_t^s \\ \nu_t^i \end{bmatrix} = \begin{bmatrix} u_t^s \\ u_t^i \end{bmatrix} - \begin{bmatrix} \hat{b}_{41} & \hat{b}_{42} & \hat{b}_{43} \\ \hat{b}_{51} & \hat{b}_{52} & \hat{b}_{53} \end{bmatrix} \begin{bmatrix} e(\hat{B})_{1t} \\ e(\hat{B})_{2t} \\ e(\hat{B})_{3t} \end{bmatrix}, \tag{3.23}$$

where $e(\hat{B})_{1t}$, $e(\hat{B})_{2t}$, and $e(\hat{B})_{3t}$ are the unmixed innovations of the recursive block. The adjusted reduced form shocks are equal to the variation in u_t^s and u_t^i unexplained by the unmixed innovations in the recursive block. The unmixed innovations in the first block are linear combinations of the structural shocks in the recursive block. Therefore, the two adjusted reduced form shocks are equal to a linear combination of the structural shocks in the non-recursive block, which contains the stock market and monetary policy shock. We find strong evidence that one

Table 3.3: Monetary policy and the stock market - Non-Gaussianity of adjusted reduced form shocks.

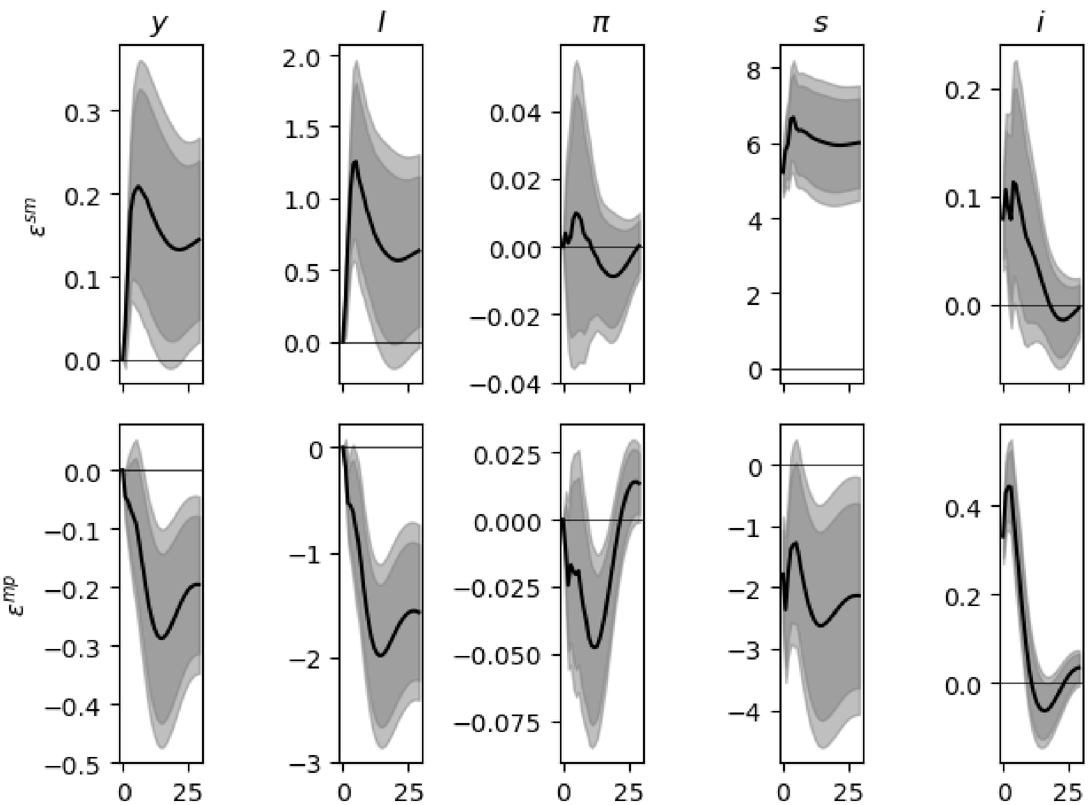
	ν_t^s	ν_t^i
Skewness	-0.313	-0.407
Kurtosis	3.413	10.663
JB-Test	0.192	0

Skewness, kurtosis and the p-value of the Jarque-Bera test of the adjusted reduced form shocks.

of the adjusted reduced form shocks is non-Gaussian. Non-Gaussianity of the adjusted reduced form shocks implies that at least one of the two structural shocks in the non-recursive block is non-Gaussian, which is sufficient to identify the stock market and monetary policy shock.

Figure 3.2 shows the corresponding impulse response functions (IRFs) to the estimated stock market and monetary policy shocks. The responses of stock returns, real investment growth and real GDP growth are integrated to show the associated level effects, which makes it possible to visually check the validity of long-run restrictions. Exploiting the partly recursive order makes

Figure 3.2: Monetary policy and the stock market - Impulse responses.



The figure shows the responses to one standard deviation shocks in the partly recursive SVAR. The columns y, I, and s show the cumulative responses of investment growth, output growth, and stock returns. Confidence bands are symmetrical 68% and 80% bootstrap bands with 2000 replications in the bootstrap algorithm.

labeling straightforward: There is only one shock which leads to an increase of the interest rate together with a decrease of output and a medium-run decrease of inflation in the non-recursive block, which is what we would expect from a monetary policy shock. The remaining shock is then labeled as the stock market shock.

The estimated lower-right block of the B matrix reads (asymptotic variance, Wald test statistic with $H_0 : B_{ij} = 0$ and p-value for the elements in parentheses)[18]

$$\hat{B}_{lower-right} = \begin{bmatrix} 5.22 & -1.78 \\ (27.21/141.4/0.0) & (67.54/6.63/0.01) \\ 0.08 & 0.33 \\ (0.17/5.08/0.02) & (0.31/50.06/0.0) \end{bmatrix}. \tag{3.24}$$

A one standard deviation stock market shock which increases stock prices by about 5.22% leads to an immediate interest rate increase of approximately 0.08 percentage points. A one standard deviation monetary policy shock leads to a decrease in stock prices of about 1.78%, while the nominal interest rate increases about 0.33 percentage points. The estimated simultaneous interaction is qualitatively comparable to the results in Bjørnland and Leitemo (2009). The Wald tests suggest that the simultaneous response of stock returns and interest rates to monetary policy and stock market shocks is significant at the 5% level. Additionally, also the confidence bands in Figure 3.2 suggest that both variables interact simultaneously and cannot be ordered recursively.

Consistent with the news literature around Beaudry and Portier (2006), we find that a positive stock market shock is followed by a business cycle expansion with an increase in real output growth and real investment growth. Therefore, even if the central bank is not interested in stock prices in the first place, a stock market shock implies a business cycle boom and the central bank, trying to stabilize output, will increase the FFR. Additionally, we find that a contractionary monetary policy shock induces a recession with a decrease in output, investment and prices. The future recession and an efficient stock market, which immediately incorporates all available information, then explains the initial negative response of stock prices to the monetary policy shock.

In Section 3.2 we presented two models: An exogenous growth model where monetary policy shocks lead to transitory deviations in real output, real investment and real stock returns, and an endogenous growth model where monetary policy shocks lead to permanent deviations in real output, real investment and real stock returns. The estimator based on long-run restrictions used in Bjørnland and Leitemo (2009) ex ante chooses the exogenous growth model and thus

[18] Based on Keweloh (2021a) we use the assumption of serially and mutually independent shocks to estimate the asymptotic variance.

does not allow to distinguish between both models. The partly recursive estimation approach presented above does not require any restrictions on the short- or long-run impact of monetary policy and stock market shocks. Therefore allows to discriminate between both models. Based on the resulting impulse responses we find evidence that monetary policy is not neutral w.r.t stock prices in the long-run. In particular, a contractionary monetary policy shock leads to a permanent and significant decrease of output, investment and stock prices.

Moreover, the ability to identify structural shocks based on long-run restrictions depends on the explanatory power of lagged variables. In particular, identifying the impact of monetary policy shocks on real stock returns based on a long-run restriction requires that lagged variables in the reduced form Equation (3.20) explain variation in real stock returns. However, we find that real stock returns are only insignificantly affected by nearly all lagged variables and we cannot reject the null hypothesis of the joint F-test that all considered lags have no impact on real stock returns, see Appendix 3.6.3. If the null hypothesis is true such that real stock returns only react simultaneously to structural shocks, imposing a long-run zero restriction on the response of real stock prices to a monetary policy shock is equal to imposing a short-run zero restriction on the response of real stock returns. Importantly, the explanatory power of lagged variables for real stock returns has no impact on the partly recursive estimator proposed in this study. We re-estimate the SVAR with the additional restriction that stock returns do not respond to any lagged variables. The results are shown in Appendix 3.6.3 and we find an immediate and by construction permanent effect of a monetary policy shock on stock prices, which is in line with our results from before.

In Appendix 3.6.3 we conduct robustness checks regarding the observation period, the exclusion of the linear time trend, the lag structure, the estimation method within the blocks, different specifications, the inclusion of further control variables, and regarding the viability of our recursiveness restrictions by estimating the SVAR without imposing any restrictions. Our main results are found to be robust and remain qualitatively unchanged: Real stock returns and the nominal interest rate both react immediately to monetary policy and stock market shocks and the long-run impact of monetary policy shocks on real stock prices is persistently negative.

3.5 Conclusion

The present paper proposes a partly recursive, partly non-Gaussian SVAR estimation approach, which generalizes between the traditional restriction based and the more recent data-driven identification approach based on non-Gaussianity. In particular, we demonstrate how the SVAR can be identified by combining traditional short-run zero restrictions and information retrieved from moments beyond the variance. We show that purely data-driven estimators based on non-Gaussianity suffer from a curse of dimensionality in small samples and large models. Exploiting the partly recursive order can break the curse of dimensionality and increase the finite sample performance, rendering our estimator to be more flexible and feasible for larger applications with only limited data availability.

We apply the proposed partly recursive SVAR estimator to analyze the interaction of monetary policy and the stock market. We argue that on the one hand there are not enough credible short or long-run restrictions available to identify the SVAR based on restrictions on the interaction of the variables. On the other hand, purely data-driven approaches typically become less and less precise the more variables are considered in the SVAR. Employing our new estimator, we find that real stock returns and interest rates both react simultaneously to stock market and monetary policy shocks. In particular, a tightening of monetary policy leads to a recession and a decrease of real stock returns and a positive stock market shock indicates a future business cycle expansion and an immediate increase in interest rates. We find evidence against the long-run neutrality of monetary policy: The estimates show that a tightening of monetary policy leads to permanently lower output, investment, and real stock prices.

3.6 Appendix

3.6.1 Appendix - White SVAR estimators with partly recursive constraints

Let $Q(B, u)$ be the objective function of a non-Gaussian SVAR estimator. Moreover, define the unmixed innovation $e_t(B) = B^{-1}u_t$. A whitened SVAR estimator then requires that $\frac{1}{T}\sum_{t=1}^{T} e_t(B)e_t'(B) = I$, such that in a given sample the unmixed innovations are mutually uncorrelated with unit variance.

Estimating an n dimensional SVAR with m partly recursive constraints and T observations yields the following optimization problem

$$\hat{B} := \underset{B \in \mathbb{R}^{n \times n}}{\arg\min}\, Q(B, u) \tag{3.25}$$
$$s.t. \quad b_{i,j} = 0, \quad \text{for } i < j \text{ and } i \leq m.$$

A whitened SVAR estimator has an additional constraint

$$\hat{B} := \underset{B \in \mathbb{R}^{n \times n}}{\arg\min}\, Q(B, u) \tag{3.26}$$
$$s.t. \quad b_{i,j} = 0, \quad \text{for } i < j \text{ and } i \leq m \tag{3.27}$$
$$\frac{1}{T}\sum_{t=1}^{T} e_t(B) e_t'(B) = I. \tag{3.28}$$

However, due to the whitening constraint (3.28) the optimization problem (3.26) is difficult to solve numerically.

First, we ignore the partly recursive constraint (3.27) and consider a white SVAR estimator with the corresponding optimization problem

$$\hat{B} := \underset{B \in \mathbb{R}^{n \times n}}{\arg\min}\, Q(B, u) \tag{3.29}$$
$$s.t. \quad \frac{1}{T}\sum_{t=1}^{T} e_t(B) e_t'(B) = I$$

The constrained optimization problem (3.29) can be transformed into an unconstrained optimization problem over orthogonal matrices. Let $VV' = \frac{1}{T}\sum_{t=1}^{T} u_t u_t'$ be the sky decomposition of the sample variance-covariance matrix of the reduced form shocks. For simplicity, we ignore the indeterminacy of sign and permutation. It holds that $\hat{B} = V\hat{O}$ with

$$\hat{O} := \underset{O \in \mathbb{O}^{n \times n}}{\arg\min}\, Q(VO, u), \tag{3.30}$$

where $\mathbb{O}^{n \times n}$ denotes the set of $n \times n$ dimensional orthogonal matrices. The optimization problem over orthogonal matrices in Equation (3.30) has no constraints and can be pulled back to an optimization problem over the euclidean space, see Lezcano-Casado and Martınez-Rubio (2019). Therefore, let $exp(\cdot)$ denote the matrix exponential function, let $s(\cdot)$ be the function which maps a vector into a lower skew-symmetric matrix. It then holds that

$$\hat{O} := \underset{\theta \in \mathbb{R}^{\frac{n(n-1)}{2}}}{\arg\min}\ Q(V\mathcal{O}(\theta), u), \tag{3.31}$$

where $\mathcal{O}(\theta) = exp(s(\theta))$ maps the $\frac{n(n-1)}{2}$ dimensional vector θ into an orthogonal matrix.

Similar to the case without the partly recursive constraints, the optimization problem (3.26) with the partly recursive constraints (3.27) can be transformed into an optimization problem over orthogonal matrices such that $\hat{B} = V\hat{O}$ with

$$\hat{O} := \underset{O \in \mathbb{O}^{n \times n}}{\arg\min}\, Q(VO, u), \tag{3.32}$$

$$s.t. \quad (VO)_{i,j} = 0, \quad \text{, for } i < j \text{ and } i \leq m \tag{3.33}$$

Let $d = \frac{(n-m)(n-m-1)}{2}$ and define the mapping between a d dimensional vector into an orthogonal matrix which preserves the partly recursive constraint (3.33)

$$\mathcal{O}_m : \mathbb{R}^d \to \mathbb{O}^{n \times n},\, \theta \longmapsto \begin{bmatrix} I_m & 0 \\ 0 & exp(s(\theta)) \end{bmatrix}, \tag{3.34}$$

where I_m denotes and m dimensional identity matrix. The optimization problem (3.32) can now be pulled back to an unconstrained optimization problem over the euclidean space

$$\hat{O} := \underset{\theta \in \mathbb{R}^d}{\arg\min}\, Q(V\mathcal{O}_m(\theta), u), \tag{3.35}$$

which simplifies the numerical optimization problem.

We now show that in a SVAR with a whitening constraint, the first m columns of the B matrix and therefore the first m recursively ordered shocks are determined by second moments due to the whitening constraint. Put differently, no information in moments beyond the variance can affect the estimated impact of the first m recursively ordered shocks since it is entirely determined by the whitening constraint. For simplicity, consider the four dimensional example with $m = 2$

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} b_{11} & 0 & 0 & 0 \\ b_{21} & b_{22} & 0 & 0 \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix} \tag{3.36}$$

which can be written as

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \tag{3.37}$$

$$\begin{bmatrix} u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} b_{31} & b_{32} \\ b_{41} & b_{42} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} + \begin{bmatrix} \nu_3 \\ \nu_4 \end{bmatrix} \tag{3.38}$$

$$\begin{bmatrix} \nu_3 \\ \nu_4 \end{bmatrix} = \begin{bmatrix} b_{33} & b_{34} \\ b_{43} & b_{44} \end{bmatrix} \begin{bmatrix} \varepsilon_3 \\ \varepsilon_4 \end{bmatrix}. \tag{3.39}$$

In a whitened SVAR, the unmixed innovations have to satisfy the condition $\frac{1}{T}\sum_{t=1}^{T} e_t(B)e_t'(B) = I$. In particular, the matrix B has to satisfy

$$\frac{1}{T}\sum_{t=1}^{T} e_{1,t}(B)e_{1,t}(B) = 1 \tag{3.40}$$

$$\frac{1}{T}\sum_{t=1}^{T} e_{2,t}(B)e_{2,t}(B) = 1 \tag{3.41}$$

$$\frac{1}{T}\sum_{t=1}^{T} e_{1,t}(B)e_{2,t}(B) = 0. \tag{3.42}$$

However, Equation (3.37) is a recursive SVAR which is uniquely determined by the variance and covariance conditions (3.40)-(3.42). Therefore, in a whitened SVAR the parameters b_{11}, b_{21}, and b_{22} and hence the first m estimated structural shocks, here $\hat{e}_1$ and $\hat{e}_2$, are uniquely determined by second moments. Note that this solution is equal to the solution obtained by applying the Cholesky decomposition to the variance covariance matrix of the reduced form shocks. Moreover, the whitening constraint implies

$$\frac{1}{T}\sum_{t=1}^{T}\hat{e}_{1,t}v_{3,t}(B) = 0 \tag{3.43}$$

$$\frac{1}{T}\sum_{t=1}^{T}\hat{e}_{2,t}v_{3,t}(B) = 0 \tag{3.44}$$

$$\frac{1}{T}\sum_{t=1}^{T}\hat{e}_{1,t}v_{4,t}(B) = 0 \tag{3.45}$$

$$\frac{1}{T}\sum_{t=1}^{T}\hat{e}_{2,t}v_{4,t}(B) = 0. \tag{3.46}$$

Replacing ε_1 and ε_2 with $\hat{e}_1$ and $\hat{e}_2$ in Equation (3.38) and exploiting the four conditions (3.43)-(3.46) implies that the parameters b_{31}, b_{32}, b_{41}, and b_{42} are again uniquely determined by second moments. Therefore, the estimated impact of the first m recursively ordered shocks is uniquely determined by second-order moment conditions derived from the whitening constraint.

3.6.2 Appendix - Finite sample performance

Table 3.4 and Table 3.5 report the results of the simulation in Section 3.3.2 using the white fast SVAR GMM estimator proposed by Keweloh (2021b) and the PML estimator proposed by Gouriéroux et al. (2017) to estimate the non-recursive block. Table 3.6 and 3.7 report the results of a simulation analogous to the simulation in Section 3.3.2 using t-distributed structural shocks with seven degrees of freedom.

Table 3.4: Finite sample performance - Partly recursive SVAR (GMM white fast).

T	$n=2$ GMMWF	$n=4$ GMMWF	$n=4$ partly recursive GMMWF
100	$\begin{bmatrix} \underset{(2.38)}{9.77} & \underset{(4.76)}{4.83} \\ \underset{(4.71)}{4.91} & \underset{(2.52)}{9.69} \end{bmatrix}$	$\begin{bmatrix} \underset{(2.67)}{9.15} & \underset{(4.44)}{-0.05} & \underset{(4.82)}{0.13} & \underset{(5.07)}{0.02} \\ \underset{(5.1)}{4.64} & \underset{(3.66)}{9.13} & \underset{(5.94)}{0.04} & \underset{(6.13)}{0.01} \\ \underset{(7.73)}{4.49} & \underset{(7.76)}{4.6} & \underset{(5.88)}{9.21} & \underset{(8.05)}{4.51} \\ \underset{(8.04)}{4.56} & \underset{(7.73)}{4.59} & \underset{(7.71)}{4.67} & \underset{(6.34)}{9.11} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.14)}{9.94} & . & . & . \\ \underset{(1.34)}{4.98} & \underset{(1.02)}{9.88} & . & . \\ \underset{(1.79)}{4.98} & \underset{(1.46)}{4.95} & \underset{(2.62)}{9.59} & \underset{(5.24)}{4.75} \\ \underset{(1.79)}{4.99} & \underset{(1.49)}{4.96} & \underset{(5.1)}{4.84} & \underset{(2.77)}{9.56} \end{bmatrix}$
250	$\begin{bmatrix} \underset{(1.13)}{9.86} & \underset{(2.28)}{4.96} \\ \underset{(2.32)}{4.92} & \underset{(0.98)}{9.88} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.85)}{9.62} & \underset{(2.31)}{-0.01} & \underset{(2.33)}{-0.02} & \underset{(2.18)}{-0.03} \\ \underset{(2.57)}{4.8} & \underset{(1.43)}{9.63} & \underset{(2.82)}{0.04} & \underset{(2.83)}{-0.04} \\ \underset{(3.61)}{4.86} & \underset{(3.53)}{4.76} & \underset{(2.64)}{9.68} & \underset{(3.62)}{4.75} \\ \underset{(3.52)}{4.87} & \underset{(3.39)}{4.8} & \underset{(3.73)}{4.89} & \underset{(2.56)}{9.59} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.4)}{9.98} & . & . & . \\ \underset{(0.52)}{4.99} & \underset{(0.43)}{9.97} & . & . \\ \underset{(0.73)}{4.99} & \underset{(0.61)}{4.97} & \underset{(1.02)}{9.88} & \underset{(2.07)}{4.91} \\ \underset{(0.71)}{5.0} & \underset{(0.62)}{4.97} & \underset{(2.13)}{4.97} & \underset{(0.98)}{9.84} \end{bmatrix}$
500	$\begin{bmatrix} \underset{(0.44)}{9.94} & \underset{(0.94)}{4.96} \\ \underset{(0.91)}{4.97} & \underset{(0.44)}{9.93} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.3)}{9.85} & \underset{(0.92)}{0.02} & \underset{(0.93)}{-0.01} & \underset{(0.93)}{-0.02} \\ \underset{(1.01)}{4.92} & \underset{(0.49)}{9.83} & \underset{(1.19)}{0.03} & \underset{(1.18)}{-0.0} \\ \underset{(1.45)}{4.95} & \underset{(1.54)}{4.89} & \underset{(0.91)}{9.86} & \underset{(1.58)}{4.89} \\ \underset{(1.45)}{4.95} & \underset{(1.53)}{4.91} & \underset{(1.5)}{4.94} & \underset{(1.05)}{9.82} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.22)}{9.99} & . & . & . \\ \underset{(0.25)}{5.0} & \underset{(0.22)}{9.96} & . & . \\ \underset{(0.35)}{5.0} & \underset{(0.3)}{4.99} & \underset{(0.43)}{9.93} & \underset{(0.94)}{4.96} \\ \underset{(0.36)}{5.0} & \underset{(0.3)}{4.99} & \underset{(0.91)}{4.96} & \underset{(0.46)}{9.92} \end{bmatrix}$
1000	$\begin{bmatrix} \underset{(0.19)}{9.99} & \underset{(0.4)}{5.0} \\ \underset{(0.39)}{4.99} & \underset{(0.2)}{9.98} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.13)}{9.93} & \underset{(0.42)}{0.02} & \underset{(0.41)}{-0.02} & \underset{(0.42)}{-0.0} \\ \underset{(0.46)}{4.95} & \underset{(0.21)}{9.94} & \underset{(0.5)}{0.01} & \underset{(0.52)}{-0.01} \\ \underset{(0.64)}{4.99} & \underset{(0.65)}{4.97} & \underset{(0.41)}{9.94} & \underset{(0.69)}{4.95} \\ \underset{(0.64)}{4.97} & \underset{(0.65)}{4.99} & \underset{(0.67)}{4.99} & \underset{(0.45)}{9.94} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.11)}{9.99} & . & . & . \\ \underset{(0.12)}{5.0} & \underset{(0.11)}{9.99} & . & . \\ \underset{(0.18)}{5.01} & \underset{(0.16)}{5.0} & \underset{(0.2)}{9.97} & \underset{(0.41)}{4.98} \\ \underset{(0.18)}{5.0} & \underset{(0.16)}{5.0} & \underset{(0.4)}{5.0} & \underset{(0.21)}{9.98} \end{bmatrix}$
5000	$\begin{bmatrix} \underset{(0.04)}{10.0} & \underset{(0.07)}{4.99} \\ \underset{(0.07)}{5.0} & \underset{(0.04)}{9.99} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.02)}{9.99} & \underset{(0.06)}{0.0} & \underset{(0.06)}{-0.01} & \underset{(0.06)}{-0.0} \\ \underset{(0.07)}{4.99} & \underset{(0.04)}{9.99} & \underset{(0.08)}{-0.01} & \underset{(0.08)}{0.0} \\ \underset{(0.1)}{5.0} & \underset{(0.1)}{5.01} & \underset{(0.07)}{9.98} & \underset{(0.11)}{4.99} \\ \underset{(0.1)}{5.0} & \underset{(0.11)}{5.0} & \underset{(0.1)}{4.99} & \underset{(0.07)}{9.99} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.02)}{10.0} & . & . & . \\ \underset{(0.02)}{5.0} & \underset{(0.02)}{10.0} & . & . \\ \underset{(0.04)}{5.0} & \underset{(0.03)}{5.0} & \underset{(0.04)}{10.0} & \underset{(0.07)}{5.0} \\ \underset{(0.04)}{5.0} & \underset{(0.03)}{5.0} & \underset{(0.07)}{5.0} & \underset{(0.04)}{10.0} \end{bmatrix}$

Monte Carlo simulation with $M = 2000$ replications. The table shows the average, $1/M \sum_{m=1}^{M} \hat{b}_{ij}^m$, and the estimated mean squared error, $1/M \sum_{m=1}^{M} \left(\hat{b}_{ij}^m - b_{ij}\right)^2$, of each estimated element $\hat{b}_{ij}^m$ of b_{ij} denoting the element of B in row i and column j. The table reports results for the white fast SVAR GMM estimator proposed by Keweloh (2021b). The partly recursive white fast SVAR GMM estimator for $n = 4$ uses the zero restrictions highlighted by the dots.

Table 3.5: Finite sample performance - Partly recursive SVAR (PML).

T	$n=2$ PML	$n=4$ PML	$n=4$ partly recursive PML
100	$\begin{bmatrix} \underset{(4.37)}{9.76} & \underset{(9.13)}{4.22} \\ \underset{(6.1)}{5.26} & \underset{(8.13)}{9.16} \end{bmatrix}$	$\begin{bmatrix} \underset{(2.59)}{9.32} & \underset{(3.61)}{-0.04} & \underset{(3.66)}{0.08} & \underset{(3.83)}{0.02} \\ \underset{(4.13)}{4.72} & \underset{(3.62)}{9.25} & \underset{(4.86)}{0.08} & \underset{(4.96)}{-0.06} \\ \underset{(6.18)}{4.61} & \underset{(6.44)}{4.64} & \underset{(4.63)}{9.48} & \underset{(12.55)}{3.79} \\ \underset{(6.5)}{4.63} & \underset{(6.54)}{4.68} & \underset{(6.97)}{5.27} & \underset{(11.51)}{8.65} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.14)}{9.94} & \cdot & \cdot & \cdot \\ \underset{(1.34)}{4.98} & \underset{(1.02)}{9.88} & \cdot & \cdot \\ \underset{(1.79)}{4.98} & \underset{(1.46)}{4.95} & \underset{(3.36)}{9.67} & \underset{(9.84)}{4.06} \\ \underset{(1.79)}{4.99} & \underset{(1.49)}{4.96} & \underset{(5.78)}{5.3} & \underset{(8.21)}{9.02} \end{bmatrix}$
250	$\begin{bmatrix} \underset{(3.29)}{9.78} & \underset{(3.23)}{4.81} \\ \underset{(2.98)}{4.94} & \underset{(3.6)}{9.7} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.52)}{9.79} & \underset{(1.18)}{-0.0} & \underset{(1.33)}{0.0} & \underset{(1.18)}{-0.01} \\ \underset{(1.32)}{4.9} & \underset{(0.84)}{9.8} & \underset{(1.52)}{0.01} & \underset{(1.65)}{-0.05} \\ \underset{(2.07)}{4.91} & \underset{(1.89)}{4.9} & \underset{(1.47)}{9.84} & \underset{(2.74)}{4.77} \\ \underset{(1.99)}{4.92} & \underset{(1.96)}{4.93} & \underset{(2.16)}{4.99} & \underset{(2.12)}{9.69} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.4)}{9.98} & \cdot & \cdot & \cdot \\ \underset{(0.52)}{4.99} & \underset{(0.43)}{9.97} & \cdot & \cdot \\ \underset{(0.73)}{4.99} & \underset{(0.61)}{4.97} & \underset{(2.1)}{9.85} & \underset{(2.14)}{4.86} \\ \underset{(0.71)}{5.0} & \underset{(0.62)}{4.97} & \underset{(2.42)}{4.97} & \underset{(1.88)}{9.77} \end{bmatrix}$
500	$\begin{bmatrix} \underset{(1.91)}{9.87} & \underset{(0.88)}{4.94} \\ \underset{(1.62)}{4.92} & \underset{(1.22)}{9.88} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.23)}{9.92} & \underset{(0.47)}{-0.01} & \underset{(0.5)}{-0.0} & \underset{(0.49)}{0.0} \\ \underset{(0.54)}{4.97} & \underset{(0.34)}{9.89} & \underset{(0.66)}{-0.0} & \underset{(0.67)}{0.01} \\ \underset{(0.78)}{4.98} & \underset{(0.83)}{4.94} & \underset{(0.67)}{9.9} & \underset{(0.82)}{4.96} \\ \underset{(0.77)}{4.97} & \underset{(0.84)}{4.94} & \underset{(0.86)}{4.94} & \underset{(0.6)}{9.91} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.22)}{9.99} & \cdot & \cdot & \cdot \\ \underset{(0.25)}{5.0} & \underset{(0.22)}{9.96} & \cdot & \cdot \\ \underset{(0.35)}{5.0} & \underset{(0.3)}{4.99} & \underset{(2.01)}{9.84} & \underset{(0.67)}{4.99} \\ \underset{(0.36)}{5.0} & \underset{(0.3)}{4.99} & \underset{(2.09)}{4.87} & \underset{(0.55)}{9.9} \end{bmatrix}$
1000	$\begin{bmatrix} \underset{(0.42)}{9.98} & \underset{(0.26)}{5.0} \\ \underset{(0.49)}{4.99} & \underset{(0.19)}{9.98} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.12)}{9.95} & \underset{(0.24)}{0.01} & \underset{(0.24)}{-0.01} & \underset{(0.24)}{-0.0} \\ \underset{(0.27)}{4.98} & \underset{(0.17)}{9.96} & \underset{(0.31)}{0.0} & \underset{(0.3)}{-0.01} \\ \underset{(0.39)}{5.0} & \underset{(0.38)}{4.98} & \underset{(0.4)}{9.95} & \underset{(0.4)}{4.98} \\ \underset{(0.39)}{4.99} & \underset{(0.37)}{4.99} & \underset{(0.5)}{4.98} & \underset{(0.3)}{9.97} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.11)}{9.99} & \cdot & \cdot & \cdot \\ \underset{(0.12)}{5.0} & \underset{(0.11)}{9.99} & \cdot & \cdot \\ \underset{(0.18)}{5.01} & \underset{(0.16)}{5.0} & \underset{(0.62)}{9.95} & \underset{(0.34)}{4.99} \\ \underset{(0.18)}{5.0} & \underset{(0.16)}{5.0} & \underset{(0.55)}{4.98} & \underset{(0.41)}{9.97} \end{bmatrix}$
5000	$\begin{bmatrix} \underset{(0.03)}{10.0} & \underset{(0.05)}{4.99} \\ \underset{(0.05)}{5.0} & \underset{(0.03)}{9.99} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.02)}{9.99} & \underset{(0.04)}{0.0} & \underset{(0.04)}{-0.0} & \underset{(0.04)}{-0.0} \\ \underset{(0.05)}{5.0} & \underset{(0.03)}{9.99} & \underset{(0.06)}{-0.01} & \underset{(0.06)}{0.0} \\ \underset{(0.07)}{5.0} & \underset{(0.07)}{5.0} & \underset{(0.05)}{9.99} & \underset{(0.07)}{4.99} \\ \underset{(0.07)}{5.0} & \underset{(0.08)}{5.0} & \underset{(0.07)}{5.0} & \underset{(0.05)}{9.99} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.02)}{10.0} & \cdot & \cdot & \cdot \\ \underset{(0.02)}{5.0} & \underset{(0.02)}{10.0} & \cdot & \cdot \\ \underset{(0.04)}{5.0} & \underset{(0.03)}{5.0} & \underset{(0.03)}{10.0} & \underset{(0.05)}{4.99} \\ \underset{(0.04)}{5.0} & \underset{(0.03)}{5.0} & \underset{(0.05)}{5.01} & \underset{(0.03)}{9.99} \end{bmatrix}$

Monte Carlo simulation with $M = 2000$ replications. The table shows the average, $1/M \sum_{m=1}^{M} \hat{b}_{ij}^{m}$, and the estimated mean squared error, $1/M \sum_{m=1}^{M} \left(\hat{b}_{ij}^{m} - b_{ij}\right)^2$, of each estimated element $\hat{b}_{ij}^{m}$ of b_{ij} denoting the element of B in row i and column j. The table reports results for the PML estimator proposed by Gouriéroux et al. (2017) using a t-distribution with seven degrees of freedom. The partly recursive PML estimator for $n = 4$ uses the zero restrictions highlighted by the dots.

Table 3.6: Finite sample performance - Partly recursive SVAR (t-distribution).

T	$n=2$ GMM	$n=4$ GMM	$n=4$ partly recursive GMM
100	$\begin{bmatrix} \underset{(3.65)}{9.6} & \underset{(9.64)}{4.46} \\ \underset{(8.74)}{5.01} & \underset{(4.37)}{9.35} \end{bmatrix}$	$\begin{bmatrix} \underset{(3.43)}{8.88} & \underset{(6.3)}{-0.04} & \underset{(6.58)}{-0.09} & \underset{(6.47)}{-0.12} \\ \underset{(7.05)}{4.5} & \underset{(5.24)}{8.82} & \underset{(8.15)}{-0.02} & \underset{(7.77)}{-0.03} \\ \underset{(10.12)}{4.56} & \underset{(10.61)}{4.34} & \underset{(8.34)}{8.88} & \underset{(12.5)}{4.1} \\ \underset{(10.18)}{4.6} & \underset{(10.65)}{4.39} & \underset{(10.83)}{4.64} & \underset{(9.67)}{8.63} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.98)}{9.97} & . & . & . \\ \underset{(1.33)}{4.98} & \underset{(0.95)}{9.89} & . & . \\ \underset{(1.8)}{4.89} & \underset{(1.48)}{4.93} & \underset{(3.24)}{9.64} & \underset{(9.99)}{4.19} \\ \underset{(1.78)}{4.9} & \underset{(1.45)}{4.97} & \underset{(8.33)}{5.22} & \underset{(4.96)}{9.15} \end{bmatrix}$
250	$\begin{bmatrix} \underset{(1.62)}{9.77} & \underset{(4.27)}{4.85} \\ \underset{(4.03)}{4.9} & \underset{(1.78)}{9.74} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.38)}{9.37} & \underset{(3.51)}{0.03} & \underset{(3.66)}{-0.0} & \underset{(3.88)}{-0.02} \\ \underset{(3.96)}{4.69} & \underset{(2.12)}{9.44} & \underset{(4.06)}{0.01} & \underset{(4.6)}{-0.04} \\ \underset{(5.95)}{4.72} & \underset{(5.29)}{4.69} & \underset{(3.67)}{9.45} & \underset{(6.66)}{4.5} \\ \underset{(5.89)}{4.74} & \underset{(5.68)}{4.7} & \underset{(5.76)}{4.82} & \underset{(4.86)}{9.23} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.4)}{9.97} & . & . & . \\ \underset{(0.5)}{4.98} & \underset{(0.38)}{9.99} & . & . \\ \underset{(0.7)}{4.97} & \underset{(0.59)}{4.96} & \underset{(1.44)}{9.8} & \underset{(4.41)}{4.72} \\ \underset{(0.7)}{5.0} & \underset{(0.59)}{4.96} & \underset{(3.93)}{4.98} & \underset{(1.91)}{9.65} \end{bmatrix}$
500	$\begin{bmatrix} \underset{(0.63)}{9.93} & \underset{(2.09)}{4.84} \\ \underset{(1.82)}{5.04} & \underset{(0.9)}{9.84} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.55)}{9.7} & \underset{(1.88)}{-0.01} & \underset{(1.84)}{0.02} & \underset{(2.0)}{0.02} \\ \underset{(1.96)}{4.87} & \underset{(1.05)}{9.67} & \underset{(2.38)}{0.04} & \underset{(2.38)}{-0.02} \\ \underset{(2.75)}{4.85} & \underset{(2.99)}{4.84} & \underset{(1.68)}{9.75} & \underset{(3.36)}{4.76} \\ \underset{(2.93)}{4.85} & \underset{(2.99)}{4.86} & \underset{(2.72)}{4.95} & \underset{(2.4)}{9.63} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.2)}{10.01} & . & . & . \\ \underset{(0.24)}{5.01} & \underset{(0.2)}{9.97} & . & . \\ \underset{(0.37)}{5.01} & \underset{(0.31)}{4.98} & \underset{(0.62)}{9.91} & \underset{(1.78)}{4.9} \\ \underset{(0.36)}{5.01} & \underset{(0.31)}{4.99} & \underset{(1.67)}{4.99} & \underset{(0.75)}{9.86} \end{bmatrix}$
1000	$\begin{bmatrix} \underset{(0.29)}{9.97} & \underset{(0.78)}{4.98} \\ \underset{(0.77)}{4.99} & \underset{(0.32)}{9.96} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.14)}{9.88} & \underset{(0.78)}{0.02} & \underset{(0.85)}{-0.0} & \underset{(0.76)}{-0.01} \\ \underset{(0.83)}{4.93} & \underset{(0.35)}{9.87} & \underset{(1.02)}{0.04} & \underset{(1.01)}{-0.03} \\ \underset{(1.23)}{4.95} & \underset{(1.32)}{4.91} & \underset{(0.75)}{9.88} & \underset{(1.4)}{4.91} \\ \underset{(1.17)}{4.95} & \underset{(1.25)}{4.93} & \underset{(1.28)}{4.96} & \underset{(0.91)}{9.84} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.1)}{10.0} & . & . & . \\ \underset{(0.13)}{5.01} & \underset{(0.1)}{9.98} & . & . \\ \underset{(0.18)}{5.01} & \underset{(0.15)}{5.0} & \underset{(0.28)}{9.94} & \underset{(0.88)}{4.96} \\ \underset{(0.18)}{5.01} & \underset{(0.15)}{4.99} & \underset{(0.79)}{4.97} & \underset{(0.37)}{9.94} \end{bmatrix}$
5000	$\begin{bmatrix} \underset{(0.06)}{10.0} & \underset{(0.15)}{4.99} \\ \underset{(0.15)}{5.0} & \underset{(0.06)}{9.99} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.02)}{9.98} & \underset{(0.14)}{0.0} & \underset{(0.15)}{-0.02} & \underset{(0.16)}{0.0} \\ \underset{(0.15)}{4.98} & \underset{(0.06)}{9.98} & \underset{(0.2)}{-0.01} & \underset{(0.18)}{0.0} \\ \underset{(0.23)}{5.0} & \underset{(0.24)}{5.0} & \underset{(0.13)}{9.96} & \underset{(0.22)}{5.0} \\ \underset{(0.24)}{4.99} & \underset{(0.23)}{4.99} & \underset{(0.21)}{4.97} & \underset{(0.13)}{9.98} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.02)}{10.01} & . & . & . \\ \underset{(0.03)}{5.0} & \underset{(0.02)}{10.0} & . & . \\ \underset{(0.03)}{5.0} & \underset{(0.03)}{5.0} & \underset{(0.05)}{9.99} & \underset{(0.14)}{5.0} \\ \underset{(0.03)}{5.0} & \underset{(0.03)}{5.0} & \underset{(0.13)}{4.99} & \underset{(0.05)}{9.99} \end{bmatrix}$

Monte Carlo simulation with $M = 2000$ replications. The structural shocks are drawn from a t-distribution with seven degrees of freedom and normalized to unit variance by multiplying each shock with $1/\sqrt{7/(7-2)}$. The table shows the average, $1/M \sum_{m=1}^{M} \hat{b}_{ij}^m$, and the estimated mean squared error, $1/M \sum_{m=1}^{M} \left(\hat{b}_{ij}^m - b_{ij}\right)^2$, of each estimated element $\hat{b}_{ij}^m$ of b_{ij} denoting the element of B in row i and column j. The table reports results for the SVAR GMM estimator without restrictions for $n = 2$ and $n = 4$, and the partly recursive SVAR GMM estimator for $n = 4$, which uses zero restrictions highlighted by the dots.

Table 3.7: Finite sample performance - Hypothesis tests in the partly recursive SVAR (t-distribution).

		$n=2$ CUE	$n=4$ CUE	$n=4$ partly recursive CUE
$J-Test$	$T=100$	3.4	1.0	2.15
	$T=250$	2.75	0.3	2.4
	$T=500$	3.7	1.9	4.1
	$T=1000$	6.35	5.8	6.15
	$T=5000$	8.1	8.5	7.6
Wald $H_0 : b = 0$	$T=100$	59.7	49.2	57.0
	$T=250$	85.3	72.7	84.0
	$T=500$	97.0	92.0	97.5
	$T=1000$	99.6	99.3	99.8
	$T=5000$	100	100	100
Wald $H_0 : b = 5$	$T=100$	18.8	17.1	19.4
	$T=250$	14.7	17.45	14.9
	$T=500$	12.3	15.1	12.2
	$T=1000$	11.9	14.1	11.4
	$T=5000$	10.8	9.7	9.2

Monte Carlo simulation with $M = 2000$ replications. The structural shocks are drawn from a t-distribution with seven degrees of freedom. The shocks are normalized to unit variance by multiplying each shock with $1/\sqrt{7/(7-2)}$. The table shows the average the rejection rates at $\alpha = 10\%$ for the J-Test, a Wald test with $H_0 : b = 0$, and Wald test with $H_0 : b = 5$, where b is equal to the element b_{12} for $n = 2$ and b_{34} for $n = 4$. The table reports results for the SVAR GMM estimator without restrictions for $n = 2$ and $n = 4$, and the partly recursive SVAR GMM estimator for $n = 4$. We use the assumption of serially and mutually independent shocks to estimate the asymptotic variance required for the Wald tests as proposed in Keweloh (2021a).

3.6.3 Appendix - Application

This section contains supplementary material and robustness checks for the application presented in Section 3.4. The estimated interaction of the stock market and monetary policy is found to be robust to all applied robustness checks.

Table 3.8 shows descriptive statistics of the variables used in the SVAR.

Table 3.8: Monetary policy and the stock market - Descriptive statistics.

	Mean	Median	Std. deviation	Variance	Skewness	Kurtosis
y	0.71	0.74	0.61	0.37	-0.83	6.46
I	1.1	0.95	3.15	9.90	-0.28	5.30
π	2.25	2.07	0.85	0.72	0.34	2.71
s	1.78	2.14	6.39	40.8	-0.73	4.99
i	3.69	4.02	3.43	11.76	-0.03	2.07

To check on our previous results, we estimate the SVAR without restrictions. In particular, we use the unrestricted SVAR GMM estimator proposed in Keweloh (2021a) using all coskewness and cokurtosis conditions. We update the weighting matrix continuously and use the assumption of serially and mutually independent shocks to estimate the asymptotically optimal weighting matrix and the asymptotic variance as proposed in Keweloh (2021a). Table 3.9 shows the correlation between the estimated structural shocks from the non-recursive SVAR and the reduced form shocks.

Table 3.9: Monetary policy and the stock market - Correlation of reduced form and estimated structural shocks (unrestricted SVAR).

	u^y	u^I	u^π	u^s	u^i
ε^y	0.58	0.33	0.62	-0.29	0.08
ε^I	0.29	0.66	-0.26	-0.41	0.13
ε^π	-0.5	0.06	0.73	-0.06	0.13
ε^{sm}	0.57	0.66	0.04	0.79	0.44
ε^{mp}	-0.09	-0.19	-0.06	-0.4	0.85

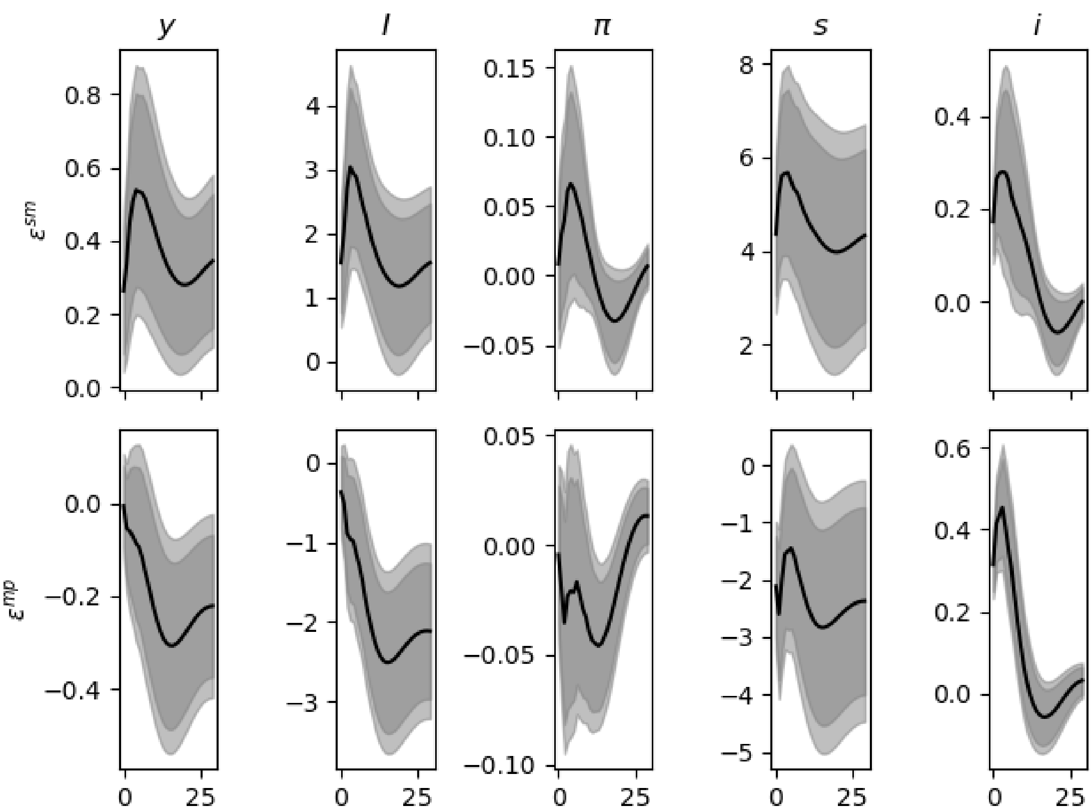
Based on the correlation with the reduced form shocks we label the fourth shock as the stock market shock and the fifth shock as the monetary policy shock. Table 3.10 shows the skewness, kurtosis and p-value of the Jarque-Bera test of all estimated structural shocks in the unrestricted SVAR.

Table 3.10: Monetary policy and the stock market - Skewness, Kurtosis and p-value of the Jarque-Bera test of estimated structural shocks (unrestricted SVAR).

	ε^{y}	ε^{I}	ε^{π}	ε^{sm}	ε^{mp}
Skewness	-0.064	0.632	-0.186	-1.185	-0.14
Kurtosis	2.412	4.848	2.968	6.629	12.325
JB-Test	0.345	0	0.663	0	0

Figure 3.3 shows the impulse responses for stock market and monetary policy shocks estimated by the unrestricted SVAR.

Figure 3.3: Monetary policy and the stock market - Impulse responses (unrestricted SVAR).



The figure shows the responses to one standard deviation shocks in the unrestricted SVAR. The columns y, I, and s show the cumulative responses of investment growth, output growth, and stock returns. Confidence bands are symmetrical 68% and 80% bootstrap bands with 2000 replications in the bootstrap algorithm.

The results are similar to the ones obtained using the partly recursive structure. Stock market shocks have an immediate positive effect on the FFR, while a monetary policy shock has an immediate and permanent negative effect on real stock prices. However, the confidence bands are larger compared to the partly recursive estimation in the main text. The results illustrate how combining a data-driven approach with zero restrictions on the short-run interaction allows to decrease the variance of the estimator and to gain deeper insights into the interrelationship between monetary policy and the stock market.

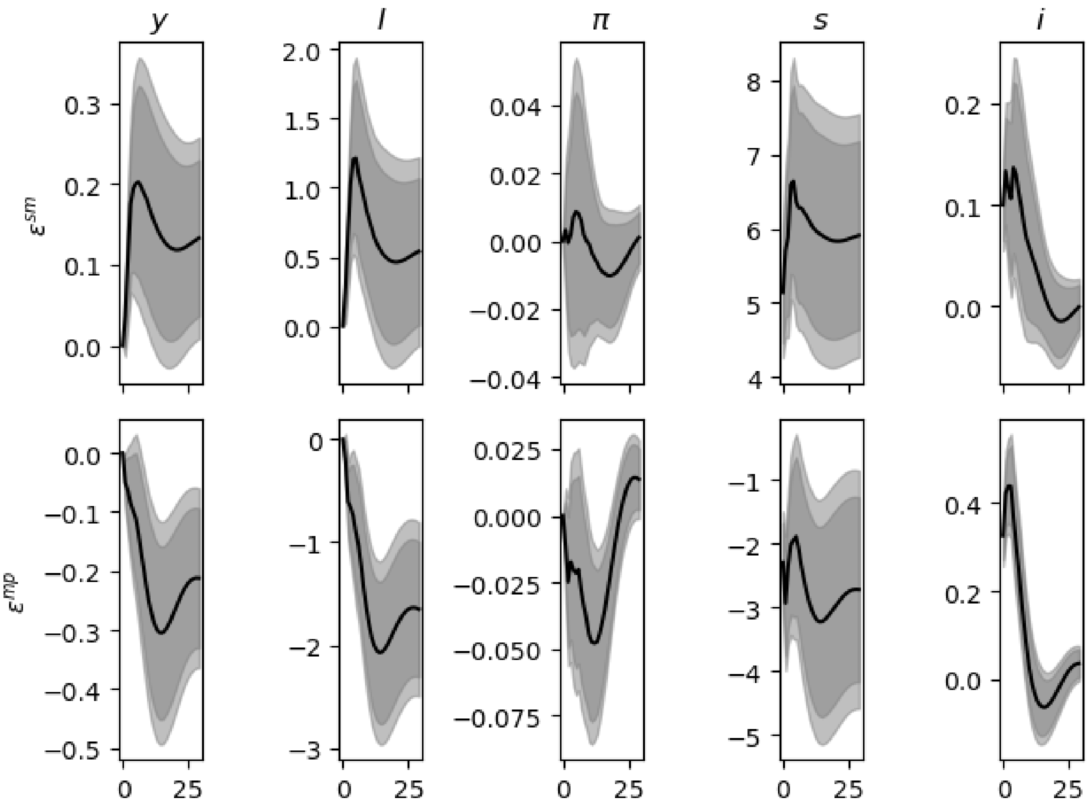
The last two columns of the estimated B matrix corresponding to the stock market and monetary policy shock read (asymptotic variance, Wald test statistic with $H_0 : B_{ij} = 0$ and p-value for the elements in parentheses)

$$\hat{B}_{columns\,4/5} = \begin{bmatrix} 0.26 & -0.01 \\ (0.56/17.43/0.0) & (0.42/0.01/0.92) \\ 1.55 & -0.37 \\ (12.09/27.84/0.0) & (10.21/1.87/0.17) \\ 0.01 & 0.0 \\ (0.14/0.07/0.79) & (0.1/0.03/0.87) \\ 4.36 & -2.11 \\ (48.94/54.71/0.0) & (59.38/10.61/0.0) \\ 0.17 & 0.31 \\ (0.17/24.57/0.0) & (0.33/41.67/0.0) \end{bmatrix}. \tag{3.47}$$

The unrestricted estimation confirms our finding on the interaction of monetary policy and the stock market: A tightening of monetary policy induces a recession with a decrease in output, investment, inflation and stock prices, and a positive stock market shock is accompanied by an immediate increase in interest rates. Additionally, we again find evidence that no recursive ordering of both variables is viable. Turning to the validity of the partly recursiveness assumption used in Section 3.4, we find mixed results. In particular, we perform a joint Wald test with the null hypothesis that the stock market and monetary policy shock have no simultaneous impact on output growth, investment growth and inflation. The Wald statistic of this test is 27.063 with a p-value < 0.01 and, therefore, we would reject the hypothesis of the partly recursive order used in Section 3.4 at any conventional significance level. However, Keweloh (2021a) shows that in larger SVARs the Wald tests reject the null hypothesis too often, especially when multiple hypotheses are tested jointly. Based on the element wise Wald tests shown above, we reject the null hypothesis that stock market shocks have no simultaneous impact on investment and output at the 1% level. However, the Wald test statistics are distorted and tend to reject the null hypothesis too often, see Section 3.3.2. We conclude that due to biased and volatile estimates and distorted test statistics in larger unrestricted SVAR models, it is difficult to judge whether the partly recursive structure is correct.

We now check if our results are dependent on our estimation technique for the non-recursive block. Therefore, we employ the PML estimator proposed by Gouriéroux et al. (2017) using the assumption of t-distributed shocks with seven degrees of freedom to estimate the non-recursive block. Figure 3.4 shows the results.

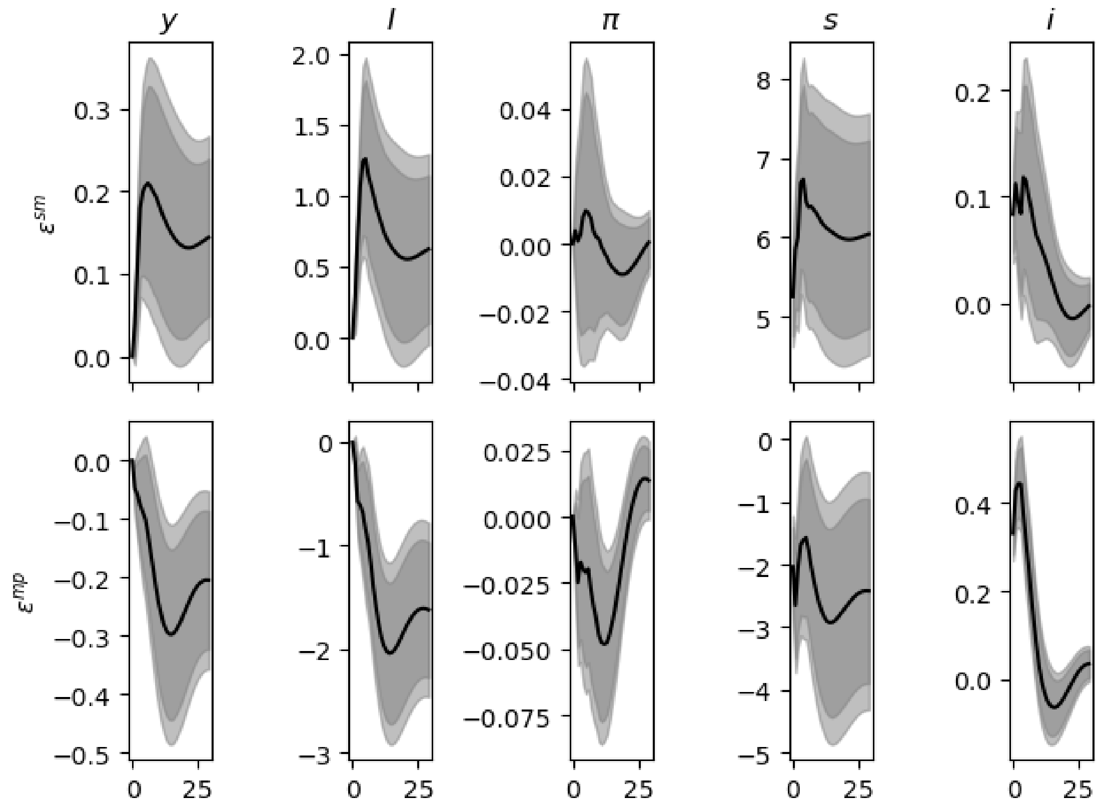
Figure 3.4: Monetary policy and the stock market - Impulse responses (PML estimator).



The figure shows the responses to one standard deviation shocks in the partly recursive SVAR using the PML estimator proposed by Gouriéroux et al. (2017) assuming t-distributed shocks with seven degrees of freedom to estimate the non-recursive block. The columns y, I, and s show the cumulative responses of investment growth, output growth, and stock returns. Confidence bands are symmetrical 68% and 80% bootstrap bands with 2000 replications in the bootstrap algorithm.

Additionally, we use the whitened fast SVAR GMM estimator proposed in Keweloh (2021a) to estimate the non-recursive block. Figure 3.5 shows the results.

Figure 3.5: Monetary policy and the stock market - Impulse responses (whitened fast SVAR GMM estimator).

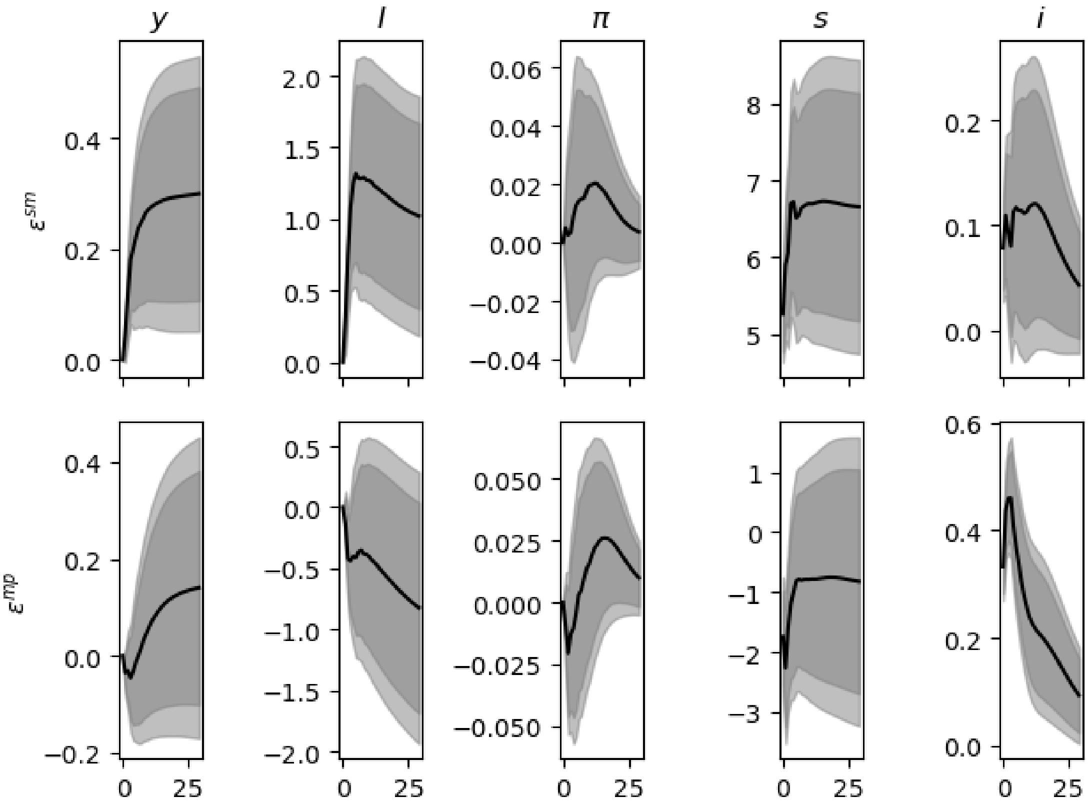


The figure shows the responses to one standard deviation shocks in the partly recursive SVAR using the whitened fast SVAR GMM estimator proposed by Keweloh (2021b) to estimate the non-recursive block. The columns y, I, and s show the cumulative responses of investment growth, output growth, and stock returns. Confidence bands are symmetrical 68% and 80% bootstrap bands with 2000 replications in the bootstrap algorithm.

The change of the estimation technique does not change our results from Section 3.4: The interest rate increases in response to a stock market shock and stock prices immediately decrease after a monetary policy shock and stay permanently below the level without the shock.

Regarding our specification, we first check on the relevance of the time trend included in our specification. We included a linear time trend to account for a potential drift in the nominal interest rate as noted by Carvalho et al. (2016). Here we exclude the linear time trend from the estimation procedure and assume that the nominal interest rate is stationary like in standard economic theory. Figure 3.6 shows the impulse responses of the stock price and FFR to a stock market and monetary policy shock under the specification without a linear time trend.

Figure 3.6: Monetary policy and the stock market - Impulse responses (no trend).

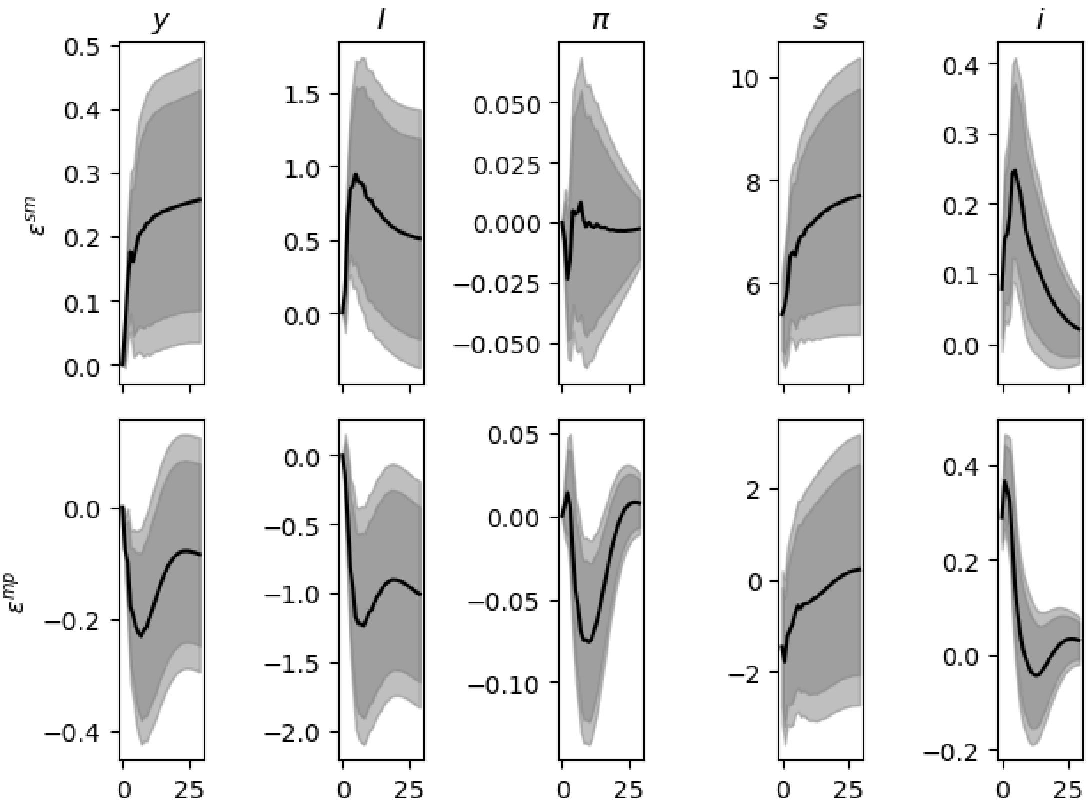


The figure shows the responses to one standard deviation shocks in the partly recursive SVAR without a trend. The columns y, I, and s show the cumulative responses of investment growth, output growth, and stock returns. Confidence bands are symmetrical 68% and 80% bootstrap bands with 500 replications in the bootstrap algorithm.

The estimated interaction of monetary policy and the stock market is similar to the results presented in Section 3.4. However, the confidence band of the stock price response to a monetary policy shock is broader and the response becomes insignificant in the long run. Furthermore, the response of output growth, investment growth, and the inflation rate to monetary policy shocks is not significant. Therefore, excluding the time trend has an impact on the estimated response of the macroeconomic variables to the monetary policy shock. We now additionally exclude observation from 2007Q4 onward and obtain a similar observation period as Bjørnland

and Leitemo (2009) and Kontonikas and Zekaite (2018). Figure 3.7 shows the resulting IRFs.

Figure 3.7: Monetary policy and the stock market - Impulse responses (no trend and 1983Q1 to 2007Q4).

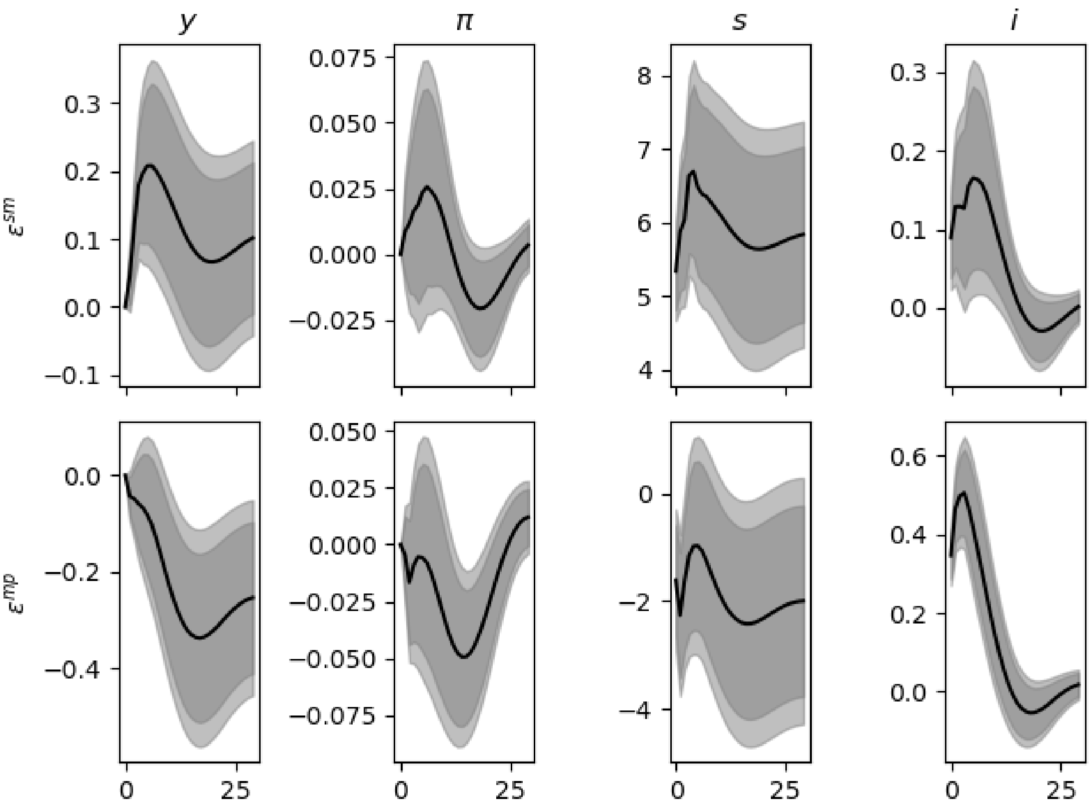


The figure shows the responses to one standard deviation shocks in the partly recursive SVAR without a trend and with data from 1983Q1 to 2007Q4. The columns y, I, and s show the cumulative responses of investment growth, output growth, and stock returns. Confidence bands are symmetrical 68% and 80% bootstrap bands with 500 replications in the bootstrap algorithm.

Again, the simultaneous interaction of monetary policy and the stock market is not affected and remains similar to the estimated interaction in Section 3.4. We also find the expected negative response of output, investment, and inflation to a tightening of monetary policy. Additionally, the assumption of long-run neutrality of monetary policy with respect to real stock prices cannot be rejected in the shorter sample. However, we again find large confidence bands. Long-run neutrality is part of the confidence band, but only one of several outcomes. Considering that including the time trend and the data from 2007Q4 onward results in more precise and conclusive results we opt to keep them in the specification of the main body of the paper.

We also consider the exclusion of investment growth. We introduced investment due to the model in Section 3.2, but of course investment does not have to be the driver behind endogenous growth. Figure 3.8 shows the resulting IRFs.

Figure 3.8: Monetary policy and the stock market - Impulse responses (without investments).



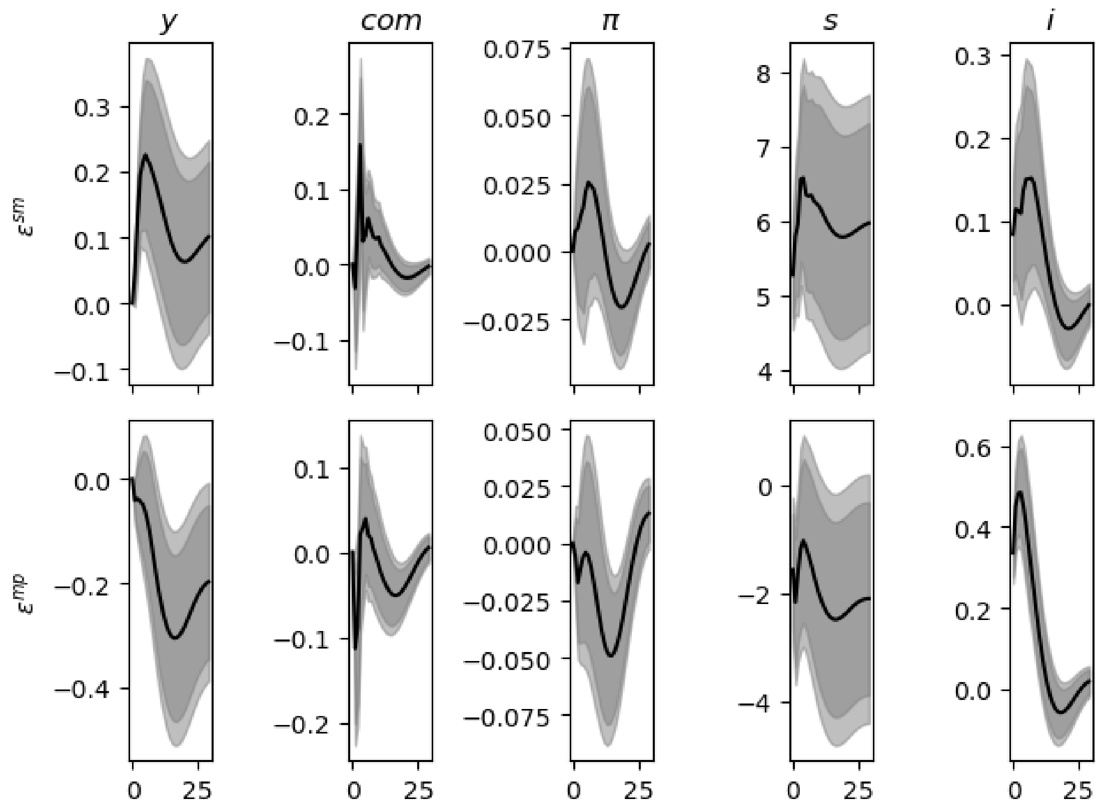
The figure shows the responses to one standard deviation shocks in the partly recursive SVAR without investments. The columns y and s show the cumulative responses of investment growth, output growth, and stock returns. Confidence bands are symmetrical 68% and 80% bootstrap bands with 500 replications in the bootstrap algorithm.

As it turns out, our results from the main body of the paper remain unchanged.

Additionally, we include commodity price inflation (named *com*), defined as the logarithmic difference in the producer price index (also taken from the FRED). For instance, Bjørnland and

Leitemo (2009) argue that the inclusion of commodity price inflation helps to reduce the price puzzle and thus should be included into the SVAR specification. We assume that commodity price inflation can not react to stock market and monetary policy shocks. Figure 3.9 shows the resulting IRFs.

Figure 3.9: Monetary policy and the stock market - Impulse responses (controlled for commodity price inflation).

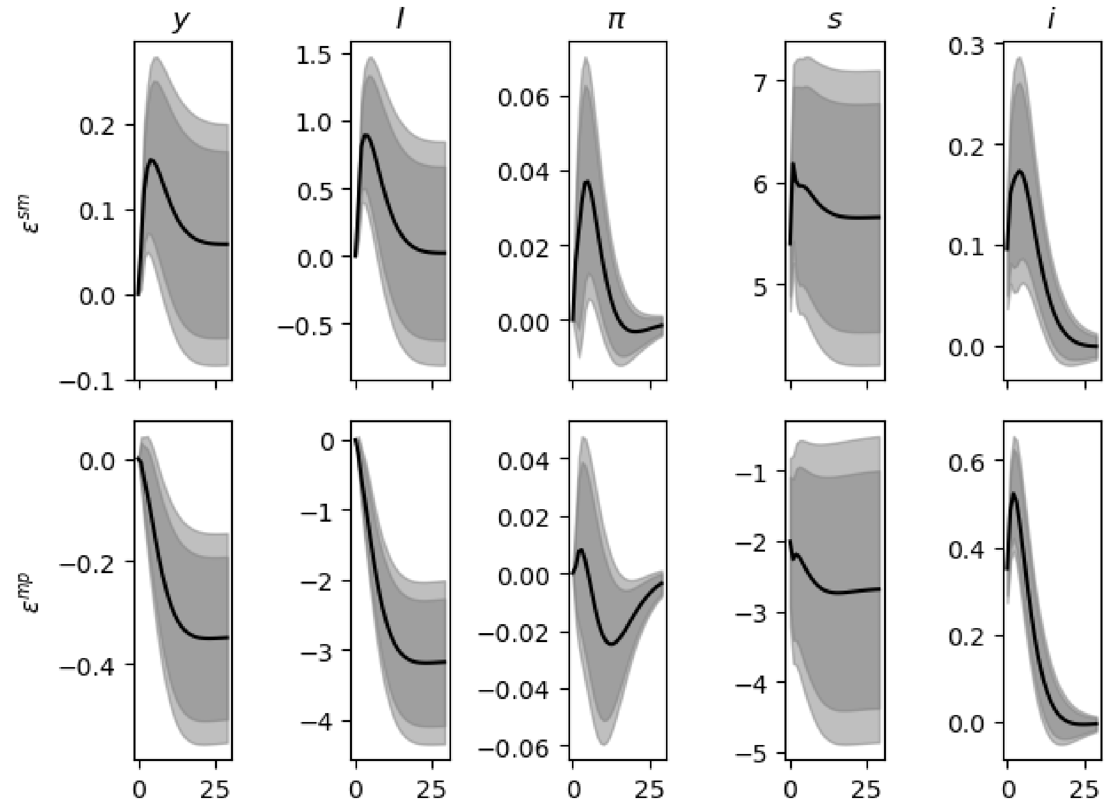


The figure shows the responses to one standard deviation shocks in the partly recursive SVAR controlled for commodity price inflation. The columns y and s show the cumulative responses of investment growth, output growth, and stock returns. Confidence bands are symmetrical 68% and 80% bootstrap bands with 500 replications in the bootstrap algorithm.

The inclusion of commodity price inflation has no impact on the estimated interaction of monetary policy and stock markets compared to Section 3.4.

Moreover, we use a lag length of two lags as suggested by the AIC. In Section 3.4 we use a higher lag order to allow for more dynamics. Figure 3.10 shows the resulting IRFs with two lags.

Figure 3.10: Monetary policy and the stock market - Impulse responses (two lags).



The figure shows the responses to one standard deviation shocks in the partly recursive SVAR with two lags. The columns y, I, and s show the cumulative responses of investment growth, output growth, and stock returns. Confidence bands are symmetrical 68% and 80% bootstrap bands with 500 replications in the bootstrap algorithm.

The results are robust to the change in the lag order.

Finally, we assume that real stock returns immediately incorporate all information and are not affected by lagged variables. Table 3.11 shows the results of the OLS regression of stock prices on four lags of all variables considered in the reduced form VAR.

Table 3.11: Monetary policy and the stock market - Reduced form estimation of real stock returns.

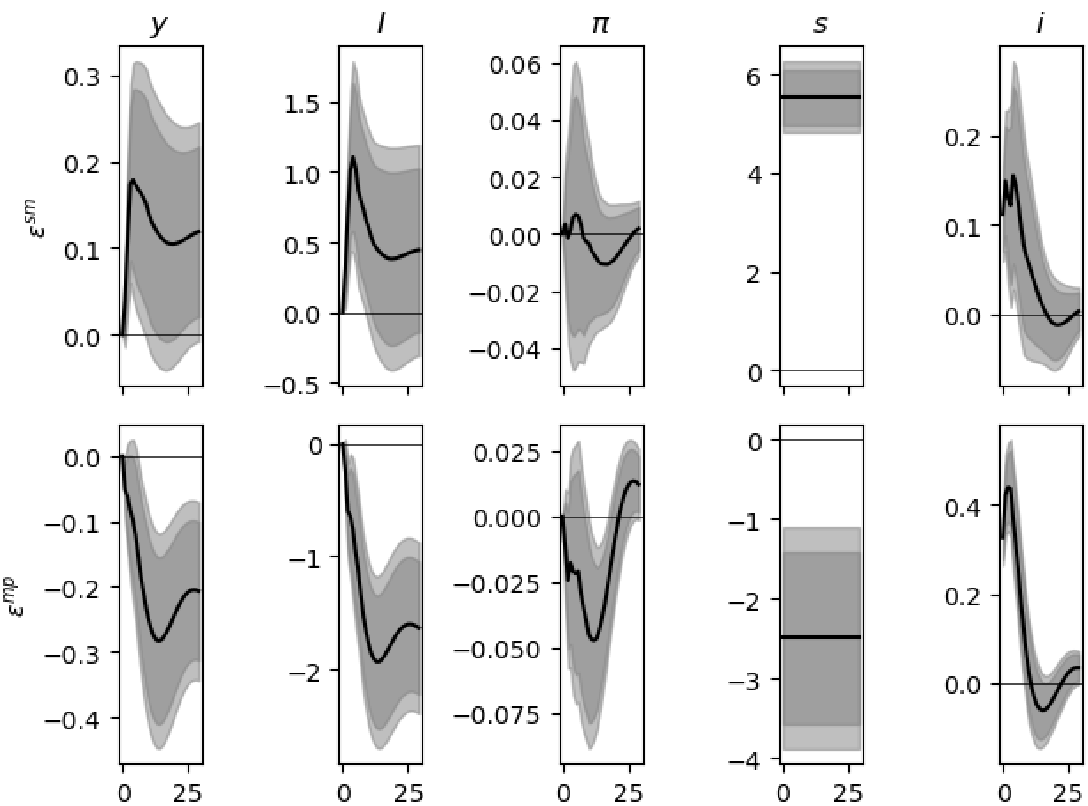
Dep. Variable: s_t	No. Observations: 141
F-statistic: 1.210	Prob (F-statistic): 0.255
R-squared: 0.176	Adj. R-squared: 0.031

	coef.	std.err.	t-value	p-value
y_{t-1}	4.0194	1.601	2.510	0.013
I_{t-1}	-0.3623	0.311	-1.164	0.247
π_{t-1}	-2.6205	2.583	-1.014	0.312
s_{t-1}	0.1299	0.094	1.377	0.171
i_{t-1}	-1.0525	1.508	-0.698	0.487
y_{t-2}	1.0216	1.695	0.603	0.548
I_{t-2}	-0.1637	0.301	-0.543	0.588
π_{t-2}	2.5525	3.992	0.639	0.524
s_{t-2}	-0.0264	0.094	-0.282	0.778
i_{t-2}	3.4615	2.484	1.394	0.166
y_{t-3}	-2.1150	1.689	-1.252	0.213
I_{t-3}	-0.1494	0.299	-0.500	0.618
π_{t-3}	0.4863	4.014	0.121	0.904
s_{t-3}	0.0800	0.096	0.837	0.404
i_{t-3}	-2.0406	2.453	-0.832	0.407
y_{t-4}	-0.5387	1.814	-0.297	0.767
I_{t-4}	0.1995	0.300	0.665	0.507
π_{t-4}	-1.2912	2.538	-0.509	0.612
s_{t-4}	-0.0105	0.093	-0.112	0.911
i_{t-4}	-0.6746	1.460	-0.462	0.645

As it can be seen, all but one of the lag coefficients are insignificant at standard significance levels. The joint F-test of the lagged observations of all considered variables being not able to explain anything about current real stock prices has an F-value of 1.21, which corresponds to a p-value of

0.255. Therefore, we cannot reject the null hypothesis that real stock returns are not affected by lagged variables. If this is true, stock prices will incorporate the effects of the shocks immediately and will not adjust dynamically over time. We proceed by excluding all lagged variables from the reduced form estimation of real stock prices and re-estimate the SVAR. Figure 3.11 shows the resulting impulse responses for a stock market and monetary policy shock.

Figure 3.11: Monetary policy and the stock market - Impulse responses (zero lags for real stock returns).



The figure shows the responses to one standard deviation shocks in the partly recursive SVAR with zero lags for real stock returns. The columns y, I, and s show the cumulative responses of investment growth, output growth, and stock returns. Confidence bands are symmetrical 68% and 80% bootstrap bands with 500 replications in the bootstrap algorithm.

Again, our main results presented in Section 3.4 are robust to the changes in the lag order.

4 Block-Recursive Non-Gaussian Structural Vector Autoregressions: Identification, Efficiency, and Moment Selection[19]

4.1 Introduction

Identification of a structural vector autoregression (SVAR) requires to assume an a priori structure of the model. Traditionally, identification is based on imposing structure on the interaction of the variables, ideally derived from macroeconomic theory (e.g., short-run restrictions Sims (1980) or long-run restrictions Blanchard and Quah (1993)). However, uncontroversial theoretical restrictions are rare. More recently, data-driven approaches allow to identify the SVAR without imposing any restrictions on the interaction. Instead, identification is achieved by imposing structure on the stochastic properties of the shocks (e.g., time-varying volatility as discussed in Rigobon (2003), Lanne et al. (2010), Lütkepohl and Netšunajev (2017), and Lewis (2021) or non-Gaussian and independent shocks as discussed in Gouriéroux et al. (2017), Lanne et al. (2017), Lanne and Luoto (2021), Keweloh (2021b), and Guay (2021)).

Traditional identification approaches may appear unnecessarily restrictive compared to novel data-driven approaches. However, Olea et al. (2022) stress that these data-driven approaches rely on information in higher moments, while traditional approaches only rely on second moments. The data-driven approaches are sensitive to the imposed statistical properties on the higher moments, while the traditional approaches are not and hence, are robust to these statistical properties. Additionally, they argue that using economic theory for identification is a feature and not a handicap and conclude that traditional identification approaches remain relevant.

We agree with their reasoning and recognize the advantages of identification approaches based on economic theory. However, in many applications we can derive some, but not sufficiently many convincing restrictions from economic theory to ensure identification. Therefore, with a traditional purely restriction based approach, even the most plausible restrictions are worthless if

[19]joint work with Stephan Hetzenecker, a slightly different version of the chapter appeared as Keweloh, S.A. and S. Hetzenecker (2021), *Efficiency gains in structural vector autoregressions by selecting informative higher-order moment conditions*, SFB 823 Discussion Paper series No. 26/2021, http://dx.doi.org/10.17877/DE290R-22447

there are not sufficiently many. We propose a Generalized Method of Moments (GMM) estimator that combines the traditional identification approach based on restrictions with the more recent data-driven approach based on non-Gaussianity. Our approach allows to impose a block-recursive structure, meaning that shocks in a given block only influence variables in the same block or blocks ordered below. The block-recursive structure seems plausible in many macroeconomic applications. Examples include applications analyzing (i) the interaction of macroeconomic and financial variables, where the former respond sluggishly while the latter respond quickly, or (ii) the interaction of small and large open economies, where large economies may have an immediate impact on small economics but not vice versa. Additionally, the block-recursive structure nests two important special cases: a recursive and an unrestricted SVAR.

Identification based on higher moments and non-Gaussian shocks oftentimes relies on the assumption of independent shocks which is criticized as too restrictive (see, e.g., Kilian and Lütkepohl (2017, Chapter 14)). Importantly, our identification result does not rely on independent shocks but is robust in the sense that it allows for various kinds of dependencies of the shocks. In particular, for a given block-recursive structure identification of the shocks within a given block is based on a small (subset) of cokurtosis conditions derived from mean independence of the shocks in the corresponding block.[20] Therefore, identification within a block follows from Lanne and Luoto (2021). Moreover, the impact of the shocks in one block on variables in another block is identified based only on covariance conditions and not on higher-order moment conditions and requires only uncorrelated shocks. Therefore, imposing a finer block-recursive structure reduces the dependency of identification on higher-order moment conditions.

However, if the shocks are independent, using only the set of identifying conditions, which are derived from mean independent shocks within blocks and uncorrelated shocks across blocks, can be inefficient. To demonstrate this, we prove that in a recursive SVAR with independent shocks the set of overidentifying higher-order moment conditions can contain additional information

[20]A common critique to the assumption of independent shocks is that it does not allow for multiple shocks to be driven by the same volatility process. Thereby, it rules out a case which may be encountered for some macroeconomic shocks. However, mean independent shocks and, in particular, the set of cokurtosis conditions used for identification allow for these kinds of dependencies.

and allows to decrease the asymptotic variance of the GMM estimator.[21] Efficient estimation requires to detect and select the valid and relevant overidentifying conditions. To this end, Lanne and Luoto (2021) suggest to calculate the information and moment selection criteria proposed by Andrews (1999) and Hall et al. (2007) for all possible combinations of moment conditions. However, they note that this approach becomes infeasible in higher-dimensional SVARs.

In a general GMM setup, Cheng and Liao (2015) propose a LASSO-type GMM estimator, hereafter referred to as the penalized GMM estimator (pGMM), which consistently selects only relevant and valid overidentifying conditions in a data-driven way. We apply the pGMM estimator to the block-recursive SVAR to exploit potential efficiency gains from overidentifying moment conditions. Our block-recursive SVAR pGMM estimator is consistent, asymptotically normal and as efficient as the asymptotically efficient block-recursive SVAR GMM estimator, including all valid and relevant overidentifying moment conditions. Importantly, these properties also hold if there are invalid overidentifying moment condition which could arise due to dependent structural shocks. Additionally, the SVAR pGMM estimator refrains from selecting valid but redundant overidentifying conditions which would neither increase nor decrease the asymptotic variance of the estimator but lead to imprecise estimates in small samples due to a many moments problem.

Guay (2021) also proposes to combine restrictions with non-Gaussian identification. In particular, he tests which shocks of the SVAR are identified based on non-Gaussianity and subsequently, his approach only uses restrictions to identify the remaining part of the SVAR. In this approach, if all shocks are non-Gaussian, no restrictions have to be used and the SVAR can be estimated solely by higher-order moment conditions. Consequently, the identification approach relies as heavily on non-Gaussianity as possible and as little on restrictions as necessary. In contrast to that, our identification approach relies as much as possible on economically justified restrictions and on non-Gaussianity only when needed.[22] To be precise, the more block-recursiveness restrictions the

[21] Note that this is not trivial. For example, in a linear regression model $y_t = \beta_1 x_t + \epsilon_t$ the GMM estimator with the moment condition $E[x_t \epsilon_t] = 0$ is identified and efficient under (conditional) homoscedastic errors. Therefore, including additional higher-order moment conditions like $E[x_t^2 \epsilon_t] = 0$ does not decrease the asymptotic variance of the GMM estimator even if the shocks or variables are non-Gaussian.

[22] Keweloh and Seepe (2020) propose a similar approach to combine restrictions and data-driven estimation. Their approach allows to identify and estimate a first block of shocks based on covariance conditions and recursiveness assumptions and a second block of unrestricted shocks based on higher moments and independent non-Gaussian shocks. They use a Monte Carlo simulation to show that imposing recursiveness restrictions leads to a better performance of the estimator in terms of bias and variance. Note that their proposed partly-recursive ordering is a special block recursive ordering and, thus, contained as a special case of the block recursive framework proposed in this study.

researcher imposes, the less identification depends on higher order-moment conditions.

We conduct two Monte Carlo experiments. In the first one, we demonstrate that the performance of a purely data-driven estimator based on non-Gaussianity deteriorates substantially with both a decreasing sample size and an increasing model size. However, exploiting the block-recursive order can mitigate this performance decline. In the second Monte Carlo experiment, we illustrate that the SVAR pGMM estimator successfully selects relevant moment conditions and increases the finite sample performance compared to other block-recursive SVAR estimators for a given block-recursive structure.

We use the block-recursive SVAR pGMM estimator to analyze the impact of oil supply and oil demand shocks, including speculative oil supply and demand shocks, on the oil price. In his seminal work, Kilian (2009) highlights that it is necessary to distinguish between oil supply and demand shocks rather than including solely an oil price shock in the SVAR for the oil market. However, oil prices are not only affected by supply and demand shocks, but also by speculative shocks causing shifts in the expectations of forward-looking traders (see, e.g., Baumeister and Kilian (2016)). In particular, new oil production technologies, anticipated wars, or news about oil discoveries or about the (future) state of the economy can shift expectations of future oil supply and future oil demand. The studies of Kilian and Murphy (2014), Juvenal and Petrella (2015), Byrne et al. (2019), and Moussa and Thomas (2021) extend the original oil market SVAR from Kilian (2009) to include speculative shocks. We contribute to this literature by explicitly distinguishing between speculative supply and speculative demand shocks.

The remainder of the paper is organized as follows: Section 4.2 reviews the SVAR and different identification schemes. Section 4.3 introduces the block-recursive SVAR. Section 4.4 derives identifying and overidentifying moment conditions in a block-recursive SVAR, analyzes which of the overidentifying conditions are redundant or relevant in a recursive SVAR, and describes the block-recursive SVAR GMM and pGMM estimators. In Section 4.5, we present the Monte Carlo experiments. In Section 4.6, we use the proposed block-recursive estimator to analyze the impact of flow and speculative supply and demand shocks in the oil market. Section 4.7 concludes.

4.2 Overview SVAR

This section briefly recalls the identification problem and common identification approaches for SVAR models. A detailed overview can be found in Kilian and Lütkepohl (2017). Consider the SVAR

$$y_t = A_1 y_{t-1} + ... + A_p y_{t-p} + B_0 \varepsilon_t, \tag{4.1}$$

with parameter matrices $A_1, ..., A_p \in \mathbb{R}^{n \times n}$, an invertible matrix B_0, an n-dimensional vector of time series $y_t = [y_{1,t}, ..., y_{n,t}]'$ and an n-dimensional vector of i.i.d. structural shocks $\varepsilon_t = [\varepsilon_{1,t}, ..., \varepsilon_{n,t}]'$ with mean zero and unit variance.

W.l.o.g. we focus on the simultaneous interaction of the SVAR given by

$$u_t = B_0 \varepsilon_t, \tag{4.2}$$

with the reduced form shocks $u_t = y_t - A_1 y_{t-1} - ... - A_p y_{t-p}$, which can be estimated consistently by OLS. The reduced form shocks are an unknown mixture B_0 of the unknown structural shocks ε_t. So far, neither the mixing matrix B_0 nor the structural shocks ε_t are identified. To see this, define the unmixed innovations $e(B)$ as the innovations obtained by unmixing the reduced form shocks with some matrix B

$$e_t(B) := B^{-1} u_t. \tag{4.3}$$

Note that for $B = B_0$, the unmixed innovations are equal to the structural shocks, i.e., $e_t(B_0) = \varepsilon_t$. Additionally, given an estimate $\hat{B}$ of B_0 we refer to $e_t(\hat{B})$ as the estimated structural shocks. The true structural shocks ε_t and the true mixing matrix B_0 are unknown and without imposing further structure, we cannot verify whether the mixing matrix B and the unmixed innovations $e_t(B)$ are equal to the true mixing matrix B_0 and the true structural shocks ε_t.

To identify B_0 and the shocks ε_t, the researcher has to impose structure on the SVAR. The structure can be specified in two ways: We may

(i) impose more structure on the interaction of the shocks (see Sims (1980) for short-run restrictions, Blanchard (1989) for long-run restrictions, and Uhlig (2005) for sign restrictions),

(ii) impose more structure on the stochastic properties of the structural shock (see Lanne et al. (2010) for time-varying volatility or Gouriéroux et al. (2017), Lanne et al. (2017), Lanne and Luoto (2021) Keweloh (2021b), and Guay (2021) for non-Gaussian shocks).

Imposing structure on the stochastic properties of the shocks can be used to derive conditions for the unmixed innovations, while imposing structure on the interaction narrows the space of possible mixing matrices used to unmix the reduced form shocks.

In applied work, the probably most frequently imposed structure are uncorrelated structural shocks (meaning $\varepsilon_{i,t}$ is restricted to be uncorrelated with $\varepsilon_{j,t}$ for $i \neq j$) and a recursive interaction (meaning restricting B_0 such that $b_{ij} = 0$ for $i < j$ where b_{ij} denotes the element at row i and column j of B_0). Uncorrelated shocks with unit variance can be used to derive $(n+1)n/2$ (co-)variance conditions from $I = E\left[\varepsilon_t \varepsilon_t'\right] \stackrel{!}{=} E\left[e_t(B) e_t(B)'\right]$. A recursive interaction implies that $n(n-1)/2$ parameters of B_0 are known a priori, leaving only $(n+1)n/2$ unknown parameters in the mixing matrix B. It is then straightforward to show that, if the remaining $(n+1)n/2$ parameters of the restricted B matrix generate unmixed innovations $e_t(B)$ which satisfy the $(n+1)n/2$ (co-)variance conditions, the matrix B has to be equal to B_0 and, hence, the unmixed innovations are equal to the structural shocks, meaning the SVAR is identified.[23]

However, economic theory rarely allows to derive the required $n(n-1)/2$ parameter restrictions to ensure identification. More recently, identification methods based on non-Gaussian and independent shocks have been put forward in the literature (see Gouriéroux et al. (2017), Lanne et al. (2017), Lanne and Luoto (2021), Keweloh (2021b), or Guay (2021)). These identification schemes do not require to impose any restrictions on the impact of the shocks, in particular on the matrix B_0. Instead, the researcher has to impose structure on the stochastic properties of the shocks. If the structural shocks are not only mutually uncorrelated but mutually independent, we can derive additional moment conditions. For example, independent and mean zero shocks imply that all entries of coskewness matrices $E\left[\varepsilon_t \varepsilon_t' \varepsilon_{i,t}\right]$ for $i = 1, \ldots, n$ are zero except for the ith diagonal element, which contains the (unknown) skewness of the shock $\varepsilon_{i,t}$. Hence, we can exploit that the mixing matrix B has to generate unmixed innovations, which satisfy the coskewness

[23]Note that this GMM approach is equivalent to the the frequently used estimator obtained by applying the Cholesky decomposition to the variance-covariance matrix of the reduced form shocks.

moment conditions derived from $E\left[\varepsilon_t\varepsilon_t'\varepsilon_{i,t}\right] \overset{!}{=} E\left[e_t(B)e_t(B)'e_{i,t}(B)\right]$. Similarly, we can use that the mixing matrix B has to generate unmixed innovations which satisfy the cokurtosis moment conditions derived from $E\left[\varepsilon_t\varepsilon_t'\varepsilon_{i,t}\varepsilon_{j,t}\right] \overset{!}{=} E\left[e_t(B)e_t(B)'e_{i,t}(B)e_{j,t}(B)\right]$.

4.3 Imposing structure in a SVAR

This section introduces the framework of the block-recursive SVAR. First, we discuss various structures of the interaction of the shocks allowed in this framework and then, assumptions on the stochastic properties of the shocks.

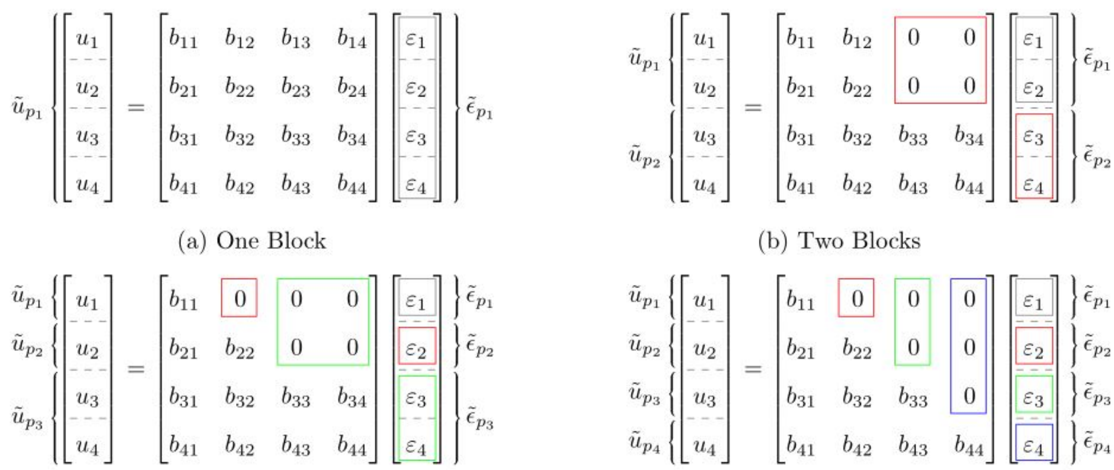
4.3.1 Imposing structure on the interaction of shocks

Traditionally, identification of a SVAR is based on the structure imposed on the interaction of the shocks (see Section 4.2). These restriction based approaches require restrictions on the interaction of the shocks to ensure identification, e.g., a recursive structure. The reasoning behind a recursive structure is oftentimes the prejudice that some variables, e.g., some macroeconomic variables like inflation, tend to move slowly, while other variables, e.g. financial variables like stock prices, react faster. However, in practice this intuitive reasoning oftentimes allows to order only some, but not all variables recursively. This motivates us to consider the block-recursive SVAR, meaning that the structural shocks are ordered in blocks of consecutive shocks and each structural shock can simultaneously affect all variables in the same block and in blocks ordered below but not variables in blocks ordered above.[24] Figure 4.1 shows different block-recursive structures in a SVAR with four variables. The examples show that a block-recursive structure generalizes the unrestricted SVAR and the fully-recursive SVAR and includes both as extreme cases.

We now introduce the notation for the block-recursive SVAR. Suppose that the structural shocks can be ordered into $m \leq n$ blocks of consecutive shocks. Let the indices $p_1 = 1 < p_2 < \ldots < p_m \leq n$ denote the beginning of a new block and for a given block p_i let $\tilde{\varepsilon}_{p_i,t}$ and $\tilde{u}_{p_i,t}$ denote

[24] Zha (1999) derives identifying restrictions for the block-recursive SVAR. The author restricts not only the simultaneous interaction, but also the lagged interaction. Our proposed block-recursive structure affects only the simultaneous interaction, while the lagged interaction remains unrestricted.

Figure 4.1: Examples of different block-recursive SVAR models.



(a) One Block (b) Two Blocks (c) Three Blocks (d) Four Blocks

Note: The figure illustrate how the the block structure can be defined by the structural shocks and our definition of $\tilde{\varepsilon}_{p_i}$ and $\tilde{u}_{p_i}$, $i = 1, \ldots, m$.

the vectors of all structural and reduced form shocks in the ith block, such that

$$\tilde{\varepsilon}_{p_i,t} := \left[\varepsilon_{p_i,t}, \varepsilon_{p_i+1,t}, \ldots, \varepsilon_{p_{i+1}-1,t}\right]' \quad \text{and} \quad \tilde{u}_{p_i,t} := \left[u_{p_i,t}, u_{p_i+1,t}, \ldots, u_{p_{i+1}-1,t}\right]', \tag{4.4}$$

where $p_{m+1} := n + 1$ for ease of notation. Moreover, let l_i denote the number of shocks in block i for $i = 1, ..., m$. The vector of all structural shocks ε_t can then be decomposed into the m blocks $\varepsilon_t = [\tilde{\varepsilon}'_{p_1,t}, \ldots, \tilde{\varepsilon}'_{p_m,t}]'$ and the reduced form shocks can be decomposed analogously into $u_t = [\tilde{u}'_{p_1,t}, \ldots, \tilde{u}'_{p_m,t}]'$. The SVAR is block-recursive with $m \leq n$ blocks with $p_1 = 1 < p_2 < \ldots < p_m \leq n$, if shocks in the ith block have no simultaneous impact on reduced form shocks in blocks j with $j < i$ such that for $i = 1, \ldots, m$

$$b_{ql} = 0, \text{ for } l \geq p_i \text{ and } q < p_i. \tag{4.5}$$

Any block-recursive structure can be described by the following assumption.

Assumption 4.1. *(Block-recursive interaction.)*
For $m \leq n$ blocks with $p_1 = 1 < p_2 < \ldots < p_m \leq n$ and $q, l = 1, ..., n$ let

$B_0 \in \mathbb{B}_{brec} \equiv \mathbb{B}_{brec}(p_1, \dots, p_m) := \{B \in \mathbb{B} | \ b_{ql} = 0 \ if \ \exists p_i \in \{p_1, ..., p_m\} \ with \ l \geq p_i \ and \ q < p_i\}$.

4.3.2 Imposing structure on the stochastic properties of shocks

Imposing structure according to Assumption 4.1 on the interaction is not sufficient to ensure identification and further assumptions on the dependence and potential non-Gaussianity of the shocks are required. In the following, we discuss different structures imposed on the mutual dependencies of the shocks.

Almost all identification approaches at least assume uncorrelated structural shocks such that $E[\varepsilon_{i,t}\varepsilon_{j,t}] = E[\varepsilon_{i,t}] E[\varepsilon_{j,t}]$ for $i \neq j$.[25] Uncorrelated shocks are justified by the idea that a given structural shock contains no information on other structural shocks, e.g., a structural monetary policy shock should not depend on other structural shocks. In general, imposing uncorrelated structural shocks does not rule out that the structural shocks are dependent. If they are dependent, the interpretation of the estimated SVAR via impulse response functions can be misleading. For example, consider the two random variables $\varepsilon_1 \sim \mathcal{N}(0, 1)$ and $\varepsilon_2 = \varepsilon_1^2 - 1$ such that both random variables are uncorrelated, but dependent. Policy analysis based on impulse response functions typically uses the ceteris paribus assumption that only a single shock varies, while the other shocks remain unchanged. In the example above, both shocks are uncorrelated, but nevertheless always move simultaneously. Therefore, uncorrelated structural shocks are not sufficient to guarantee that the ceteris paribus assumption holds.

A more rigorous implementation of the idea that a given shock contains no information on other shocks is to assume independent shocks such that $E[h(\varepsilon_{i,t})g(\varepsilon_{j,t})] = E[h(\varepsilon_{i,t})] E[g(\varepsilon_{j,t})]$ for $i \neq j$ and any bounded, measurable functions $g(\cdot)$ and $h(\cdot)$. If shocks are independent, a structural shock cannot contain any information on any other structural shock. Therefore, independent structural shocks justify the ceteris paribus interpretation used in policy analysis based on impulse response functions. However, several authors argue that the assumption of independent structural shocks is too strong (cf. Kilian and Lütkepohl (2017, Chapter 14), Lanne and Luoto (2021), Lanne et al. (2021), or Olea et al. (2022)). In particular, independence of the shocks implies that also the volatility processes of the shocks are independent, which may

[25] Proxy-variable identification approaches are different and instead assume that structural shocks are uncorrelated with an external proxy variable (see, e.g., Stock and Watson (2012), or Mertens and Ravn (2013)).

be too restrictive for some macroeconomic applications. For example, suppose that $\tilde{\varepsilon}_{1,t}$ and $\tilde{\varepsilon}_{2,t}$ are drawn independently of each other and represent unscaled structural shocks. Moreover, in each period an additional volatility shock v_t is drawn independently of the other shocks and the structural shocks are given by $\varepsilon_{1,t} = \tilde{\varepsilon}_{1,t} v_t$ and $\varepsilon_{2,t} = \tilde{\varepsilon}_{2,t} v_t$. These structural shocks are uncorrelated, but dependent since the variance of one shock contains information on the variance of the other shock.

A compromise between the two extreme cases of uncorrelated and independent shocks is the assumption of mean independent shocks, such that $E[\varepsilon_{i,t} g(\varepsilon_{j,t})] = E[\varepsilon_{i,t}] E[g(\varepsilon_{j,t})]$ for $i \neq j$ with a bounded, measurable function $g(\cdot)$. If shocks are mean independent, a structural shock cannot contain any information about the mean of other structural shocks. Mean independent shocks can justify the ceteris paribus assumption used in impulse response analysis and at the same time allow for dependent volatility processes. In particular, the two shocks $\varepsilon_{1,t} = \tilde{\varepsilon}_{1,t} v_t$ and $\varepsilon_{2,t} = \tilde{\varepsilon}_{2,t} v_t$ defined above are mean independent since a given shock contains no information on the mean of the other shock.

Imposing structure on the dependence of the structural shocks allows to derive moment conditions (see, e.g., Lanne and Luoto (2021), Keweloh (2021b), or Guay (2021)). For $i, j, k, l = 1, ..., n$ we define the following moment conditions:

$$\begin{aligned}
&\text{Variance:} && E[e(B)_{i,t}^2] - 1 = 0 && && (4.6)\\
&\text{Covariance:} && E[e(B)_{i,t} e(B)_{j,t}] = 0, && \text{for } i < j && (4.7)\\
&\text{Coskewness:} && E[e(B)_{i,t}^2 e(B)_{j,t}] = 0, && \text{for } i \neq j && (4.8)\\
& && E[e(B)_{i,t} e(B)_{j,t} e(B)_{k,t}] = 0, && \text{for } i < j < k && (4.9)\\
&\text{Cokurtosis:} && E[e(B)_{i,t}^3 e(B)_{j,t}] = 0, && \text{for } i \neq j && (4.10)\\
& && E[e(B)_{i,t}^2 e(B)_{j,t} e(B)_{k,t}] = 0, && \text{for } i \neq j,\ i \neq k,\ j < k && (4.11)\\
& && E[e(B)_{i,t} e(B)_{j,t} e(B)_{k,t} e(B)_{l,t}] = 0, && \text{for } i < j < k < l && (4.12)\\
& && E[e(B)_{i,t}^2 e(B)_{j,t}^2] - 1 = 0, && \text{for } i < j && (4.13)
\end{aligned}$$

The variance conditions in Equation (4.6) follow from the unit variance normalization. The remaining conditions are derived from different assumptions on the dependence of the structural shocks. In particular, uncorrelated structural shocks only imply the covariance conditions in Equation (4.7). Mean independent shocks additionally imply the coskewness conditions in Equation (4.8) and (4.9) and the cokurtosis conditions in Equation (4.10)-(4.12). In addition, the symmetric cokurtosis conditions in Equation (4.13) follow from independent shocks.

Moreover, note that if all structural shocks are Gaussian, the conditions in Equation (4.8)-(4.13) do not contain information beyond the information contained in the variance and covariance conditions.

4.4 Estimation of a block-recursive SVAR

In this section, we combine identification based on block-recursiveness restrictions and non-Gaussian shocks. First, for a given block-recursive structure we derive corresponding identifying asymmetric cokurtosis conditions based on mean independent shocks within the blocks. Importantly, identification is achieved without many higher-order moment conditions and holds under fairly general conditions on the dependencies of the shocks. Second, we show that additional overidentifying higher-order moment conditions, some of these conditions additionally require the assumption of independent shocks, can decrease the asymptotic variance of the estimator if the overidentifying conditions are valid. Third, we propose to use a LASSO-type GMM estimator to select the valid and relevant overidentifying higher-order moment conditions in a data-driven way. Consistency of the estimator only relies on the identifying moment conditions and, thus, is robust to various kinds of dependencies of the shocks. Furthermore, it can exploit efficiency gains from valid and relevant overidentifying conditions and ignore noise from valid but redundant overidentifying conditions.

4.4.1 Identification

In this section, we show that identification of a block-recursive SVAR can be achieved by the variance and covariance conditions in Equation (4.6) and (4.7) and the asymmetric cokurtosis conditions in Equation (4.10) corresponding to innovations in the same block. The identification result is robust in the sense that it allows for various sorts of dependencies of the shocks. To be clear, shocks in different blocks only need to be uncorrelated and shocks in the same block only need to fulfill the asymmetric cokurtosis conditions.

Let $E[f_{\mathbf{2}}(B, u_t)] = 0$ contain all variance and covariance conditions in Equation (4.6) and (4.7) and let $E[f_{\mathbf{4_{p_k}}}(B, u_t)] = 0$ contain all asymmetric cokurtosis conditions from Equation (4.10) corresponding to shocks in block k, e.g., $E[e(B)_{i,t}^3 e(B)_{j,t}] = 0$ for $i, j = p_k, ..., p_{k+1} - 1$ and $i \neq j$. We define the identifying moment conditions as

$$E[f_{\mathbf{N}}(B, u_t)] := E \begin{bmatrix} f_{\mathbf{2}}(B, u_t) \\ f_{\mathbf{4_{p_1}}}(B, u_t) \\ \vdots \\ f_{\mathbf{4_{p_m}}}(B, u_t) \end{bmatrix} = 0. \tag{4.14}$$

In the following, we simplify the notation for moment conditions, e.g., we write $E[f_{\mathbf{N}}(B, u_t)]$ instead of $E[f_{\mathbf{N}}(B, u_t)] = 0$. Note that the identifying moment conditions do not contain asymmetric cokurtosis conditions of shocks in different blocks, e.g., the moment conditions $E[e(B)_{i,t}^3 e(B)_{j,t}]$ for shocks $e(B)_{i,t}$ and $e(B)_{j,t}$ in different blocks are not contained in $E[f_{\mathbf{N}}(B, u_t)]$. The conditions $E[f_{\mathbf{N}}(B, u_t)]$ can be justified by the following assumption.

Assumption 4.2. *(Block-recursive mean independence.)*
For $m \leq n$ blocks with $p_1 = 1 < p_2 < \ldots < p_m \leq n$,

(i) all shocks are uncorrelated, i.e., $E[\varepsilon_{i,t}\varepsilon_{j,t}] = 0$ for $i \neq j$.

(ii) all shocks within the same block are mean independent, i.e., $E[\varepsilon_{i,t}|\varepsilon_{-i,t}] = 0$ for $i \in \{p_k, p_k + 1, ..., p_{k+1} - 1\}$ and $-i = \{p_k, p_k + 1, ..., p_{k+1} - 1\} \backslash i$ for $k = 1, \ldots, m$.

The identifying moment conditions contain n variance conditions, $n(n-1)/2$ covariance conditions and $\sum_{k=1}^m l_k(l_k - 1)/2$ asymmetric cokurtosis conditions, where $l_k := p_{k+1} - p_k$ denotes the number of shocks in block k. Therefore, each additional specified block refines the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)]$ such that they contain fewer higher-order moment conditions. In the extreme case when the SVAR is specified recursively, meaning each block contains only one variable, the identifying moment conditions contain no higher-order moment conditions. In the other extreme case of a single block containing all variables, the identifying moment conditions contain all $n(n-1)$ asymmetric cokurtosis conditions and are similar to the conditions proposed in Lanne and Luoto (2021).[26]

The following proposition shows that the identifying moment conditions are sufficient to locally identify the block-recursive SVAR.

Proposition 4.1. *(Identification in the block-recursive SVAR.)*
Let $u_t = B_0 \varepsilon_t$ *with* $m \leq n$ *blocks and* $B_0 \in \mathbb{B}_{brec} \equiv \mathbb{B}_{brec}(p_1, \ldots, p_m)$ *such that Assumption 4.1 holds. Moreover, suppose that Assumption 4.2 holds. If at most one structural shock in each block has zero excess kurtosis, the identifying moment conditions* $E[f_{\mathbf{N}}(B, u_t)] = 0$ *locally identify* $B = B_0$ *for* $B \in \mathbb{B}_{brec}$.

Proof. The proof recursively applies the identification result from Lanne and Luoto (2021) and can be found in Appendix 4.8.3. □

In Proposition 4.1 the impact of shocks on variables in different blocks is identified based on covariance conditions. The interaction of shocks on variables within the same block is identified based on asymmetric cokurtosis conditions and the local identification result of Lanne and Luoto (2021). Local identification means that the moment conditions $E[f_{\mathbf{N}}(B, u_t)]$ identify B_0 in a small neighborhood of B_0 (see Hall (2005)). Importantly, the proposition also holds for different higher-order moment conditions ensuring identification within the blocks. For example, the identifying conditions $E[f_{\mathbf{N}}(B, u_t)]$ could contain all variance-covariance, coskewness and cokurtosis conditions implied by independent structural shocks for each block. In this case, global identification up to sign and permutation within each block follows from Keweloh (2021b).

Without further restrictions, data-driven approaches relying on non-Gaussian and independent shocks can only ensure identification up to sign and permutation. This means that the order and sign of the shocks in the impulse response functions is not identified. In practice, the

[26] Lanne and Luoto (2021) propose to select $n(n-1)/2$ asymmetric cokurtosis conditions, which is sufficient for local identification if none of the asymmetric conditions does include the third power of a Gaussian shock. They advocate to rely on a moment selection criterion to avoid including redundant conditions or conditions of Gaussian shocks. Additionally, Lanne and Luoto (2021) note that including all $n(n-1)$ asymmetric cokurtosis conditions ensures local identification even if conditions related to Gaussian shocks are included. We argue that the degree of overidentification remains reasonably small even if we include all asymmetric cokurtosis conditions. Therefore, including redundant conditions can be expected to be rather harmless. For example, in a SVAR with four variables and no restrictions the identifying moment conditions consists of 22 conditions to identify 16 parameters. Thus, we suggest to use all asymmetric cokurtosis conditions in order to avoid the cumbersome process of selecting a subset of the conditions.

researcher has to manually assign labels to the shocks. Restricting the solution to a given block-recursive structure simplifies the permutation or labeling problem. In particular, shocks can only be permuted inside blocks. For instance, in example (b) in Figure 4.1 shocks from the second block cannot be permuted into the first block since this violates the block-recursive structure. Therefore, specifying a finer block-recursive structure simplifies the labeling of the shocks.

Define the block-recursive SVAR GMM estimator which minimizes the variance, covariance and the asymmetric cokurtosis conditions over the set of block-recursive matrices as

$$\hat{B}_{\mathbf{N}} := \underset{B \in \mathbb{B}_{brec}}{\arg\min}\, g_{\mathbf{N}}(B)' W_{\mathbf{N}} g_{\mathbf{N}}(B), \tag{4.15}$$

with a suitable weighting matrix $W_{\mathbf{N}}$ and $g_{\mathbf{N}}(B) := 1/T \sum_{t=1}^{T} f_{\mathbf{N}}(B, u_t)$. Consistency and asymptotic normality follow from the identification result in Proposition 4.1 and standard assumptions including valid moment conditions implied by the dependence structure imposed in Assumption 4.2. That is,

$$\hat{B}_{\mathbf{N}} \xrightarrow{p} B_0 \tag{4.16}$$

$$\sqrt{T}\left(\left(vec\left(\hat{B}_{\mathbf{N}}\right) - vec\left(B_0\right)\right) \xrightarrow{d} \mathcal{N}\left(0, V_{\mathbf{N}}\right), \tag{4.17}$$

where the formula for the asymptotic variance, $V_{\mathbf{N}}$, is standard but lengthy and, therefore, deferred to Appendix 4.8.1. Moreover, under standard assumptions the weighting matrix $W_{\mathbf{N}}^* := S_{\mathbf{N}}^{-1}$ with $S_{\mathbf{N}} := \lim_{T\to\infty} E[g_{\mathbf{N}}(B) g_{\mathbf{N}}(B)']$ leads to the estimator $\hat{B}_{\mathbf{N}}^*$ with lowest possible asymptotic variance (see, e.g., Hall (2005)).

In many applications, the researcher is only interested in some structural shocks. For this case, we derive a partial identification result under weaker assumptions.

Proposition 4.2. *(Partial identification in the block-recursive SVAR.)*
Let $u_t = B_0 \varepsilon_t$ with $m \leq n$ blocks and $B_0 \in \mathbb{B}_{brec} \equiv \mathbb{B}_{brec}(p_1, \ldots, p_m)$ such that Assumption 4.1 holds. Moreover, let $B_{i,0}$ denote the columns of B_0 representing impact of the structural shocks in the ith block. Let $\widetilde{\mathbb{B}}_{brec} := \mathbb{B}_{brec}(\tilde{p}_1, \ldots, \tilde{p}_{\tilde{m}})$ denote a potentially different block-recursive interaction. Assume that there exists a block $\tilde{p}_j$ of $\widetilde{\mathbb{B}}_{brec}$ which contains the shocks of block p_i, i.e., there exits a j, $1 \leq j \leq \tilde{m}$, such that $\tilde{p}_j = p_i$ and $\tilde{p}_{j+1} = p_{i+1}$.

The moment conditions $E\begin{bmatrix} f_{\mathbf{2}}(B,u_t) \\ f_{\mathbf{4}_{\tilde{\mathbf{p}}_\mathbf{j}}}(B,u_t) \end{bmatrix} = 0$ *locally identify* $B_{i,0}$ *for* $B \in \widetilde{\mathbb{B}}_{brec}$ *if the following conditions hold:*

1. *The shocks* ε_t *are uncorrelated.*

2. *The asymmetric cokurtosis conditions of block* $\tilde{p}_j$ *hold.*

3. *At most one shock in block* $\tilde{p}_j$ *has zero excess kurtosis.*

Proof. The proof can be found in Appendix 4.8.3. □

Proposition 4.2 reveals that we can identify a specific block of shocks by using only the second moments of all shocks and the asymmetric cokurtosis conditions of the shocks in the block of interest as long as the block of interest is specified correctly and contains at most one Gaussian shock. To see the advantages of the partial identification result, consider that we are only interested in the last two structural shocks in Figure 4.1 (b). In this example, Proposition 4.2 implies that the impact of the last two shocks is identified even if (i) the first and second shock are both Gaussian, (ii) the first and second shock do not satisfy the asymmetric cokurtosis conditions but are only uncorrelated, or (iii) the block-recursive structure is misspecified as the one displayed in Figure 4.1 (c). Additionally, Proposition 4.2 implies that the moment conditions used in Proposition 4.1 identify the shocks in a block of interest if the block of interest is specified correctly, contains at most one Gaussian shock, and there exists a B such that the moment conditions are fulfilled. However, the B matrix can differ from B_0, except for the columns corresponding to the block of interest.

4.4.2 Overidentification and efficiency gains

In the previous section, we proposed a block-recursive SVAR GMM estimator, which uses only a (small) subset of asymmetric cokurtosis conditions, and provide an identification result which does not require independent shocks. However, the excluded set of coskewness and cokurtosis conditions can decrease the asymptotic variance of the estimator and hence, increase the efficiency of the estimator. In this section, we define the overidentified block-recursive SVAR GMM estimator which contains all coskewness and cokurtosis conditions implied by independent shocks. Additionally, we derive conditions for the redundancy and relevance of the overidentifying coskewness and cokurtosis conditions in a recursive SVAR with independent structural shocks.

Assumption 4.3. *(Independent shocks.)*
All shocks are independent, i.e., $\varepsilon_{i,t}$ is independent of $\varepsilon_{j,t}$ for $i \neq j$.

For a given block-recursive SVAR, define the overidentifying moment conditions as

$$E[f_{\mathbf{D}}(B, u_t)] = E \begin{bmatrix} f_{\mathbf{3} \setminus \mathbf{N}}(B, u_t) \\ f_{\mathbf{4} \setminus \mathbf{N}}(B, u_t) \end{bmatrix}, \tag{4.18}$$

where $E[f_{\mathbf{3} \setminus \mathbf{N}}(B, u_t)]$ contains all coskewness conditions from Equation (4.8)-(4.9), and $E[f_{\mathbf{4} \setminus \mathbf{N}}(B, u_t)]$ contains all cokurtosis conditions from Equation (4.10)-(4.13), implied by independent shocks and not included in the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)]$.

The overidentified block-recursive SVAR GMM estimator is defined as

$$\hat{B}_{\mathbf{N}+\mathbf{D}} := \underset{B \in \mathbb{B}_{brec}}{\arg\min} \begin{bmatrix} g_{\mathbf{N}}(B) \\ g_{\mathbf{D}}(B) \end{bmatrix}' W_{\mathbf{N}+\mathbf{D}} \begin{bmatrix} g_{\mathbf{N}}(B) \\ g_{\mathbf{D}}(B) \end{bmatrix}, \tag{4.19}$$

with a suitable weighting matrix $W_{\mathbf{N}+\mathbf{D}}$ and $g_{\mathbf{D}}(B) := 1/T \sum_{t=1}^{T} f_{\mathbf{D}}(B, u_t)$. Note that the overidentified block-recursive SVAR GMM estimator uses all coskewness and cokurtosis conditions implied by independent shocks. That is, the moment conditions used for estimation are the same as in the SVAR GMM estimator proposed by Keweloh (2021b). However, the latter estimator neither uses restrictions nor distinguishes between identifying and overidentifying moment conditions. In contrast to that, we allow for block-recursive restrictions. These restrictions allow to transform identifying into overidentifying moment conditions.

Consistency and asymptotic normality of the overidentified block-recursive SVAR GMM estimator in Equation (4.19) require that not only the identifying but also the overidentifying moment conditions are valid, which holds if the shocks are independent as assumed in Assumption 4.3.

That is,

$$\hat{B}_{\mathbf{N+D}} \xrightarrow{p} B_0 \tag{4.20}$$

$$\sqrt{T}\left(\left(vec\left(\hat{B}_{\mathbf{N+D}}\right) - vec\left(B_0\right)\right) \xrightarrow{d} \mathcal{N}\left(0, V_{\mathbf{N+D}}\right), \tag{4.21}$$

where the formula for the asymptotic variance, $V_{\mathbf{N+D}}$, is standard and can be found in Appendix 4.8.1. Again, under standard assumptions the weighting matrix $W^*_{\mathbf{N+D}} := S^{-1}_{\mathbf{N+D}}$ with $S_{\mathbf{N+D}} := \lim_{T\to\infty} E[g_{\mathbf{N+D}}(B_0)g_{\mathbf{N+D}}(B_0)']$, where $g_{\mathbf{N+D}}(B_0) := [g_{\mathbf{N}}(B_0)', g_{\mathbf{D}}(B_0)']'$, leads to the estimator $\hat{B}^*_{\mathbf{N+D}}$ with lowest possible asymptotic variance (see, e.g., Hall (2005)).

Adding additional valid moment conditions can never increase the asymptotic variance of the GMM estimator (see, e.g., Breusch et al. (1999)). Therefore, if the structural shocks are independent such that the overidentifying conditions hold, the asymptotic variance of $\hat{B}^*_{\mathbf{N+D}}$ is equal to or smaller than the asymptotic variance of $\hat{B}^*_{\mathbf{N}}$. If including an additional moment condition decreases the asymptotic variance of the estimator, the moment condition is called relevant, otherwise the moment condition is called redundant. A moment condition is called partially relevant for a subset of parameters if it decreases the asymptotic variance of a subset of parameters. If this is not the case, the moment condition is called partially redundant.

In the following proposition, we show that overidentifying higher-order moment conditions in $E[f_{\mathbf{D}}(B, u_t)]$ can decrease the asymptotic variance of the SVAR GMM estimator. To this end, we consider the special case of a recursive SVAR with independent shocks. In this case, the SVAR is identified solely by second-order moment conditions and all coskewness and cokurtosis moment conditions are overidentifying. The proposition highlights that some coskewness and cokurtosis conditions are always (partially) redundant, while other conditions are relevant if certain conditions for the skewness, excess kurtosis, and elements of the inverse of B_0 are fulfilled. The proposition also implies that if at least one shock has a non-zero skewness, at least one higher-order moment condition will be relevant and consequently, the recursive SVAR GMM estimator based solely on second-order moment conditions, which is equal to frequently used estimator obtained by applying the Cholesky decomposition, is inefficient.

Proposition 4.3. *(Redundant and relevant moment conditions in the recursive SVAR.)* *Let $A := B_0^{-1}$ and let a_{ql} denote the element at row q and column l of A. Additionally let $i, j, k, l \in$*

$\{1, ..., n\}$ and $i \neq j \neq k \neq l$. The impact of a shock $\epsilon_{q,t}$ is equal to the unrestricted elements in the q-th row of B_0. In a recursive SVAR with independent structural shocks the following redundancy statements hold w.r.t. the identifying second-order moment conditions $E[f_{\mathbf{2}}(B, u_t)]$.

Coskewness condition:

1. *$E[e(B_0)_i e(B_0)_j e(B_0)_k]$ is redundant.*
2. *$E[e(B_0)_i^2 e(B_0)_j]$ is partially redundant for the impact of the shock $\epsilon_{q,t}$ with $q \neq j$.*
3. *$E[e(B_0)_i^2 e(B_0)_j]$ is partially redundant for the impact of the shock $\epsilon_{j,t}$ if and only if*

for $i < j$	*for $i > j$*
$\frac{2E[\epsilon_{j,t}^3]}{E[\epsilon_{j,t}^4]-1} a_{jj} = 0.$	$\begin{aligned} \frac{2E[\epsilon_{j,t}^3]}{E[\epsilon_{j,t}^4]-1} a_{jj} + E[\epsilon_{i,t}^3] a_{ij} &= 0, \\ E[\epsilon_{i,t}^3] a_{i,z} &= 0, \quad z = j+1, \ldots, i. \end{aligned}$

Cokurtosis condition:

1. *$E[e(B_0)_i e(B_0)_j e(B_0)_k e(B_0)_l]$ and $E[e(B_0)_i^2 e(B_0)_j e(B_0)_k]$ are redundant.*
2. *$E[e(B_0)_i^3 e(B_0)_j]$ is partially redundant for the impact of the shock $\epsilon_{q,t}$ with $q \neq j$.*
3. *$E[e(B_0)_i^3 e(B_0)_j]$ is partially redundant for the impact of the shock $\epsilon_{j,t}$ if and only if*

for $i < j$	*for $i > j$*
$\frac{2E[\epsilon_{j,t}^3]E[\epsilon_{i,t}^3]}{E[\epsilon_{j,t}^4]-1} a_{jj} = 0.$	$\begin{aligned} \frac{2E[\epsilon_{j,t}^3]E[\epsilon_{i,t}^3]}{E[\epsilon_{j,t}^4]-1} a_{jj} + (E[\epsilon_{i,t}^4] - 3) a_{ij} &= 0, \\ (E[\epsilon_{i,t}^4] - 3) a_{i,z} &= 0, \quad z = j+1, \ldots, i. \end{aligned}$

4. *$E[e(B_0)_i^2 e(B_0)_j^2 - 1]$ is partially redundant for the impact of the shock $\epsilon_{q,t}$ with $q \neq i$ and $i < j$.*
5. *$E[e(B_0)_i^2 e(B_0)_j^2 - 1]$ is partially redundant for the impact of the shock $\epsilon_{i,t}$ with $i < j$ if and only if*

$$E[\epsilon_{j,t}^3] E[\epsilon_{i,t}^3] a_{jz} = 0, \quad z = i, \ldots, j.$$

Proof. The proof can be found in Appendix 4.8.4. □

In practice, the conditions in Proposition 4.3 cannot be verified since the matrix B_0, the skewness, and the kurtosis of the structural shocks are unknown a priori. Furthermore, Proposition 4.3 only covers a recursive SVAR with independent shocks, i.e., if the shocks are only mean independent or the SVAR has a different block-recursive structure, we do not have a theoretical result on which moment conditions are relevant and which are not.

4.4.3 Data-driven moment selection

Section 4.4.1 provides an identification result for block-recursive SVARs only requiring a (small) subset of cokurtosis conditions which is robust in the sense that it allows for various kinds of dependencies of the shocks. Section 4.4.2 stresses that there is a trade-off between robustness and efficiency of the estimator. For robustness, we leave out overidentifying conditions, which has the downside that some of these conditions may be valid and relevant, i.e., decrease the asymptotic variance of the estimator. However, an advantage is that one does not include potentially invalid overidentifying conditions, which could lead to an inconsistent overidentified block-recursive SVAR GMM estimator in Equation (4.19). Additionally, valid but redundant overidentifying conditions can lead to a many moment problem and a poor finite sample performance of the overidentified block-recursive SVAR GMM estimator, compare Cheng and Liao (2015), Hall (2005), and Hall (2015). Therefore, we propose to use the pGMM estimator of Cheng and Liao (2015) to detect and include only the relevant and valid overidentifying moment conditions in a data-driven way. By including valid and relevant moment conditions in the estimation, we exploit the asymptotic efficiency gains of relevant moments. By leaving out invalid or redundant moment conditions, we can avoid inconsistent estimates and issues related to many moment conditions.

In general, the overidentifying higher-order moment conditions $E[f_{\mathbf{D}}(B, u_t)]$ can be separated into three sets: $E[f_{\mathbf{A}}(B, u_t)]$ contains valid and relevant moment conditions, $E[f_{\mathbf{R}}(B, u_t)]$ contains valid but redundant conditions, and $E[f_{\mathbf{I}}(B, u_t)]$ contains invalid moment conditions. The goal is to select the moments $E[f_{\mathbf{A}}(B, u_t)]$ and to leave out the moments $E[f_{\mathbf{R}}(B, u_t)]$ and $E[f_{\mathbf{I}}(B, u_t)]$. However, in practice the researcher does not know whether a given moment condition is invalid, redundant, or valid and relevant. Therefore, we propose to detect and select the relevant and valid overidentifying moment conditions in a data-driven way. Based on Cheng and Liao (2015), we define the block-recursive SVAR pGMM estimator

$$\{\widehat{B}^{pGMM}_{\mathbf{N+D}}, \widehat{\beta}\} := \underset{\{B,\beta\}\in\Lambda}{\arg\min} \begin{bmatrix} g_{\mathbf{N}}(B) \\ g_{\mathbf{D}}(B) - \beta \end{bmatrix}' W_{\mathbf{N+D}} \begin{bmatrix} g_{\mathbf{N}}(B) \\ g_{\mathbf{D}}(B) - \beta \end{bmatrix} + \lambda \sum_{j\in\widetilde{D}} \omega_j \, |\beta_j|, \tag{4.22}$$

where $\lambda \geq 0$ is a tuning parameter specified by the researcher, $\beta \in \mathbb{R}^{k_{\mathbf{D}}}$ is the vector of slackness parameters, $\Lambda := \{\mathbb{B}_{brec}, \mathbb{R}^{1\times k_{\mathbf{D}}}\}$ is the parameter space of $\{B, \beta\}$, $\omega \in \mathbb{R}^{k_{\mathbf{D}}}$ is a vector of weights used in the penalty term, and $\widetilde{D} := \{1, \ldots, k_{\mathbf{D}}\}$ with $k_{\mathbf{D}}$ denoting the number of conditions in $E[f_{\mathbf{D}}(B, u_t)]$.

The vector of slackness parameters β allows the moment conditions $E[f_{\mathbf{D}}(B, u_t)]$ to deviate from zero without increasing the first part of the loss function and therefore, to decrease their impact on the estimation. However, each element of β gets penalized in the second part of the loss function and consequently, giving slack to overidentifying moments adds a cost, i.e., increases the loss function. The vector of weights ω and the tuning parameter λ govern the cost of giving slack to moment conditions. In particular, a smaller λ makes it cheaper to give slack to all overidentifying moments and a smaller ω_j makes it less costly to give slack to a specific overidentifying moment j.

The pGMM estimator in Equation (4.22) has two special cases. First, if $\lambda = 0$, adding slack to the overidentifying moments is not penalized. Therefore, the solution of the pGMM estimator is $\widehat{B}^{pGMM}_{\mathbf{N+D}} = \hat{B}_{\mathbf{N}}$ and $\widehat{\beta} = g_{\mathbf{D}}\left(\hat{B}_{\mathbf{N}}\right)$, where $\hat{B}_{\mathbf{N}}$ is the solution of the the block-recursive SVAR GMM estimator in Equation (4.15) using only the identifying moments $E[f_{\mathbf{N}}(B, u_t)]$ and the weighting matrix $W_{\mathbf{N}}$, equal to the block of the weighting matrix $W_{\mathbf{N+D}}$ corresponding to the identifying conditions $E[f_{\mathbf{N}}(B, u_t)]$. Second, if $\lambda = \infty$, deviations of $\widehat{\beta}$ from zero become infinitely costly for overidentifying moments with $\omega_j > 0$. Assuming $\omega > 0$, the pGMM estimator cannot give slack to any overidentifying moment condition. Thus, $\widehat{B}^{pGMM}_{\mathbf{N+D}} = \hat{B}_{\mathbf{N+D}}$ and $\widehat{\beta} = 0$ minimize the loss function of the pGMM estimator, where $\hat{B}_{\mathbf{N+D}}$ is the solution of the the overidentified block-recursive SVAR GMM estimator in Equation (4.19), using the weighting matrix $W_{\mathbf{N+D}}$. Choices of λ other than $\lambda = 0$ or $\lambda = \infty$ lead to solutions which lie between these extreme cases. In practice, we recommend using cross-validation to find the optimal value of λ.

The penalty term uses weights $\omega_j \geq 0$, $\forall j \in \tilde{D}$, to shrink the elements of β differently. Let $E[f_{\mathbf{D}_j}(B, u_t)]$ for $j \in \widetilde{D}$ correspond to one specific moment of $E[f_{\mathbf{D}}(B, u_t)]$. A higher ω_j leads to more shrinkage for β_j and consequently, makes it more likely that β_j becomes zero, meaning that the corresponding moment $E[f_{\mathbf{D}_j}(B, u_t)]$ gets selected. Furthermore, $\omega_j = 0$ implies that even if the tuning parameter λ is large, there is no cost for giving slack to the moment condition $E[f_{\mathbf{D}_j}(B, u_t)]$, implying that those moments do not influence the estimated $\widehat{B}^{pGMM}_{\mathbf{N}+\mathbf{D}}$. Since we aim to select only the relevant and valid moment conditions $E[f_{\mathbf{A}}(B, u_t)$, and not the invalid $E[f_{\mathbf{I}}(B, u_t)]$ or redundant moment conditions $E[f_{\mathbf{R}}(B, u_t)]$, we would specify $\omega_j > 0$ for all valid and relevant conditions, and $\omega_j = 0$ for all invalid or redundant conditions. To achieve this without prior knowledge on $E[f_{\mathbf{A}}(B, u_t)]$, $E[f_{\mathbf{R}}(B, u_t)]$, and $E[f_{\mathbf{I}}(B, u_t)]$, Cheng and Liao (2015) construct ω_j allowing information-based adaptive adjustment for each moment in $E[f_{\mathbf{D}}(B, u_t)]$. More precisely, they use

$$\omega_j = \frac{\mu_j^{r_1}}{|\beta_j^{*r_2}|},\ j \in \tilde{D}, \tag{4.23}$$

where μ_j is a measure for the empirical relevance of the moment condition $E[f_{\mathbf{D}_j}(B, u_t)]$, relative to the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)]$, and β_j^* is a preliminary consistent estimator of $E[f_{\mathbf{D}_j}(B_0, u_t)]$ and $r_1 \geq r_2 \geq 0$ are constants specified by the researcher. The use of $1/|\beta_j^{*r_2}|$ resembles an adaptive LASSO penalty (cf. Zou (2006)) and implies that moments with small β_j^* are subject to more shrinkage. Since β_j^* is a consistent estimator and the true value of β_j^* for a valid moment is zero, the adaptive penalty ensures that valid moments get selected. However, using only the adaptive penalty, we would unintendedly incentivize the estimator to select also redundant moments since, by definition, these are also valid. To avoid selecting redundant moments, Cheng and Liao (2015) suggest to multiply the adaptive penalty with

$$\mu_j = \rho_{\max}\left(\widehat{V}_{\mathbf{N}} - \widehat{V}_{\mathbf{N}+\mathbf{D}_j}\right),\ j \in \tilde{D}, \tag{4.24}$$

where $\rho_{\max}(A)$ is the maximum eigenvalue of a square matrix A and $\widehat{V}_{\mathbf{N}}$ and $\widehat{V}_{\mathbf{N}+\mathbf{D}_j}$ are consistent estimators of the efficient asymptotic variance-covariance matrices $V^*_{\mathbf{N}}$ and $V^*_{\mathbf{N}+\mathbf{D}_j}$, defined in Appendix 4.8.1. If the maximum eigenvalue of $V^*_{\mathbf{N}} - V^*_{\mathbf{N}+\mathbf{D}_j}$ is positive, then adding moment condition $E[f_{\mathbf{D}_j}(B, u_t)]$ to the conditions $E[f_{\mathbf{N}}(B, u_t)]$ decreases the asymptotic variance of the estimator and hence, moment condition $E[f_{\mathbf{D}_j}(B, u_t)]$ is relevant. Therefore, μ_j estimates the empirical relevance of the moment $E[f_{\mathbf{D}_j}(B, u_t)]$.[27]

Cheng and Liao (2015) show that, under conditions, the pGMM estimator consistently selects the valid and relevant moments, i.e., $\lim_{T\to\infty} P(\hat{\beta}_j = 0) = 1$ if the moment condition $E[f_{\mathbf{D}_j}(B, u_t)]$ is in $E[f_{\mathbf{A}}(B, u_t)]$, and does not select the invalid or redundant moments, i.e., $\lim_{T\to\infty} P(\hat{\beta}_j = 0) = 0$ if the moment condition $E[f_{\mathbf{D}_j}(B, u_t)]$ is in $E[f_{\mathbf{R}}(B, u_t)]$ or $E[f_{\mathbf{I}}(B, u_t)]$. They also derive that, under conditions, the pGMM estimator is a consistent estimator of B_0 and asymptotically normal with asymptotic variance $V_{\mathbf{N}+\mathbf{A}}$.[28] In our case, the conditions in particular require that Assumption 4.2 holds. However, consistency and asymptotic normality do not rely on independent shocks, i.e., Assumption 4.3. Even though the SVAR pGMM estimator uses the moment conditions $E[f_{\mathbf{N}}(B, u_t)]$ and $E[f_{\mathbf{D}}(B, u_t)]$ for estimation, its asymptotic variance only depends on the moments conditions $E[f_{\mathbf{D}}(B, u_t)]$ and $E[f_{\mathbf{A}}(B, u_t)]$. That is, the SVAR pGMM estimator successfully ignores the redundant and invalid moments and decreases the asymptotic variance by incorporating the information contained in the relevant and valid moments. The weighting matrix $W^*_{\mathbf{N}+\mathbf{D}} := S^{-1}_{\mathbf{N}+\mathbf{D}}$ leads to the estimator with the lowest possible asymptotic variance (Hall, 2005), corresponding to the asymptotic variance of the oracle estimator. The oracle estimator uses only moment conditions $E[f_{\mathbf{N}}(B, u_t)]$ and $E[f_{\mathbf{A}}(B, u_t)]$ and is infeasible in practice without prior knowledge of $E[f_{\mathbf{D}}(B, u_t)]$ and $E[f_{\mathbf{A}}(B, u_t)]$. However, the SVAR pGMM estimator is as efficient as the oracle estimator asymptotically.

4.5 Finite sample performance

In this section, we conduct two Monte Carlo studies. The first one illustrates that the performance of SVAR estimators can be improved substantially by exploiting the block-recursive structure. This is especially relevant for SVARs with a large number of variables. The second Monte

[27] Cheng and Liao (2015) show that $V^*_{\mathbf{N}} - V^*_{\mathbf{N}+\mathbf{D}_j}$ is positive semidefinite for every $j \in \widetilde{D}$, implying that the maximum eigenvalue of $V^*_{\mathbf{N}} - V^*_{\mathbf{N}+\mathbf{D}_j}$ is nonnegative. Furthermore, note that both $\widehat{V}_{\mathbf{N}} \equiv \widehat{V}_{\mathbf{N}}\left(\hat{B}_{\mathbf{N}}\right)$ and $\widehat{V}_{\mathbf{N}+\mathbf{D}_j} \equiv \widehat{V}_{\mathbf{N}+\mathbf{D}_j}\left(\hat{B}_{\mathbf{N}}\right)$ are evaluated at $\hat{B}_{\mathbf{N}}$, which is obtained from Equation (4.15). Thereby, we do not rely on $\hat{B}_{\mathbf{N}+\mathbf{D}_j}$ to estimate $V^*_{\mathbf{N}+\mathbf{D}_j}$ since the moment associated with $\mathbf{D}_j$ may be invalid and hence, $\widehat{V}_{\mathbf{N}+\mathbf{D}_j}\left(\hat{B}_{\mathbf{N}+\mathbf{D}_j}\right)$ inconsistent for $V^*_{\mathbf{N}+\mathbf{D}_j}$.

[28] This result is not explicitly stated in Cheng and Liao (2015) but follows from their Remark 3.5 using the Cramér-Wold device, an arbitrary weighting matrix W and replacing the variance of the sample GMM estimator with the asymptotic variance. We prove the result in Appendix 4.8.5 under Assumption 4.1 and 4.2.

Carlo study focuses on how to incorporate information in overidentifying higher-order moment conditions. More concretely, we demonstrate that the pGMM estimator selects relevant and does not select redundant moment conditions in a data-driven way and thereby, improves the finite sample performance.

For both Monte Carlo experiments, we consider three different sample sizes $T = \{100,\ 250,\ 1\,000\}$ to analyze the influence of the sample size on the performance of the estimators. We independently and identically draw each structural shock ϵ_{it}, $i = 1, \ldots, n$, $t = 1, \ldots, T$, from the two-component mixture

$$\epsilon_{it} \sim 0.79\, \mathcal{N}(-0.2, 0.7^2) + 0.21\, \mathcal{N}(0.75, 1.5^2),$$

where $\mathcal{N}(\mu, \sigma^2)$ indicates a normal distribution with mean μ and standard deviation σ. The shocks have skewness 0.9 and excess kurtosis 2.4.

We compare the finite sample performance of various SVAR estimators.[29] Based on the simulations presented in Keweloh (2021a), we use continuous updating estimators (CUEs) instead of GMM estimators and estimate the asymptotically efficient weighting matrix based on serially and mutually independent shocks.[30] We refer to the estimators as follows:

- GMM: Continuous updating estimator based on Equation (4.15) using only the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)]$.
- oGMM: Overidentified continuous updating estimator based on Equation (4.19) using the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)]$ and overidentifying moment conditions $E[f_{\mathbf{D}}(B, u_t)]$.
- GMM-Oracle: Overidentified continuous updating estimator based on Equation (4.19) using the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)]$ and the relevant overidentifying moment conditions $E[f_{\mathbf{A}}(B, u_t)]$.

[29] The estimators are implemented in python and the pGMM estimator uses the solvers of Defferrard et al. (2017).

[30] Keweloh (2021a) demonstrates that the inability to precisely estimate S, the long-run covariance matrix of the moment conditions, and as consequence the efficient weighting matrix leads to a poor small sample performance of CUE and two-step GMM estimators. Recognizing this downside, Keweloh (2021a) proposes a novel estimator for S exploiting serially and mutually independent shocks.Keweloh (2021a) illustrates that the estimator for S substantially increases the small sample performance of the two-step CUE and GMM estimator. Additionally, Keweloh (2021a) illustrates that CUE estimators are less biased than GMM estimator in small samples.

- pGMM: Continuous updating LASSO estimator based on Equation (4.22).

We only indicate which block-recursive structure is imposed for estimation, when necessary (e.g., when comparing an GMM estimator without restrictions with a block-recursive GMM estimator).

4.5.1 Block-Recursive Structure

We simulate a SVAR with $n = 2$ and $n = 4$ variables. The mixing matrices B_0 are given by

$$B_0 = \begin{bmatrix} 10 & 5 \\ 5 & 10 \end{bmatrix} \quad \text{and} \quad B_0 = \begin{bmatrix} 10 & 5 & 0 & 0 \\ 5 & 10 & 0 & 0 \\ 5 & 5 & 10 & 5 \\ 5 & 5 & 5 & 10 \end{bmatrix}. \tag{4.25}$$

The Monte Carlo study analyzes the impact of imposing a block-recursive structure for GMM estimators. In the small SVAR with $n = 2$, we impose no restrictions. In the large SVAR with $n = 4$, we estimate the GMM estimator without restrictions and the block-recursive GMM estimator, using the block-recursive structure in Equation (4.25), i.e., we apply zero restrictions for all elements where B_0 is zero.[31]

Table 4.1 summarizes the results of $M = 3,500$ Monte Carlo simulations. The table shows the average of each estimated element $\bar{b}_{ij} = 1/M \sum_{m=1}^{M} \hat{b}_{ij}^m$ and the estimated mean squared error (MSE), $\hat{\sigma}_{i,j}^2 = 1/M \sum_{m=1}^{M} \left(\hat{b}_{ij}^m - b_{ij}\right)^2$, where b_{ij} denotes the element of B_0 in row i and column j and $\hat{b}_{ij}^m$ its estimated value in Monte Carlo run m. Moreover, we calculate the average over the empirical biases, $Bias := \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} \left(\bar{b}_{ij} - b_{ij}\right)$, and the average over the estimated MSEs, $Var := \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} \hat{\sigma}_{i,j}^2$, across estimated elements in $\hat{B}$, i.e., $w_{i,j}$ equals zero if $\hat{b}_{ij}^m$ is restricted to be zero and one over the number of estimated elements in $\hat{B}$ otherwise. Additionally, we report the number of moments used by each estimator. For each estimator, the average bias and MSE decreases with the sample size. Furthermore, the simulation highlights how the performance of the GMM estimators, which are based entirely on non-Gaussianity, decreases with an increasing model size (e.g., the average bias and MSE for each sample size is up to 2.1 and 1.9 times higher for

[31] In this Monte Carlo study, we focus on GMM estimators. We include the oGMM, GMM-Oracle and pGMM estimator in the second Monte Carlo study.

Table 4.1: Finite sample performance - Block-recursive SVAR.

		n=2	n=4	
		GMM	GMM	block-recursive GMM
$T = 100$	$\hat{B}$	$\begin{bmatrix} \underset{(2.26)}{9.78} & \underset{(4.31)}{4.90} \\ \underset{(4.24)}{4.90} & \underset{(2.18)}{9.76} \end{bmatrix}$	$\begin{bmatrix} \underset{(3.24)}{9.28} & \underset{(4.82)}{4.63} & \underset{(5.31)}{0.04} & \underset{(5.27)}{0.07} \\ \underset{(4.87)}{4.70} & \underset{(3.20)}{9.23} & \underset{(5.32)}{0.08} & \underset{(5.14)}{0.05} \\ \underset{(6.54)}{4.68} & \underset{(6.74)}{4.62} & \underset{(5.01)}{9.27} & \underset{(6.54)}{4.74} \\ \underset{(6.67)}{4.67} & \underset{(6.53)}{4.65} & \underset{(6.48)}{4.66} & \underset{(4.93)}{9.33} \end{bmatrix}$	$\begin{bmatrix} \underset{(2.31)}{9.74} & \underset{(4.30)}{4.91} & . & . \\ \underset{(4.43)}{4.87} & \underset{(2.18)}{9.74} & . & . \\ \underset{(2.51)}{4.86} & \underset{(2.44)}{4.89} & \underset{(2.17)}{9.63} & \underset{(4.41)}{4.84} \\ \underset{(2.56)}{4.87} & \underset{(2.45)}{4.91} & \underset{(4.24)}{4.84} & \underset{(2.34)}{9.64} \end{bmatrix}$
	#Mo	5.00	22.00	14.00
	Bias	-0.1649	-0.3314	-0.1878
	MSE	3.25	5.41	3.03
		n=2	n=4	
		GMM	GMM	block-recursive GMM
$T = 250$	$\hat{B}$	$\begin{bmatrix} \underset{(1.10)}{9.88} & \underset{(2.30)}{4.90} \\ \underset{(2.22)}{4.98} & \underset{(1.13)}{9.85} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.64)}{9.56} & \underset{(2.77)}{4.79} & \underset{(3.19)}{0.02} & \underset{(3.21)}{0.06} \\ \underset{(2.69)}{4.77} & \underset{(1.65)}{9.54} & \underset{(3.14)}{-0.01} & \underset{(3.26)}{0.04} \\ \underset{(4.05)}{4.74} & \underset{(3.94)}{4.83} & \underset{(2.76)}{9.56} & \underset{(3.92)}{4.83} \\ \underset{(4.10)}{4.74} & \underset{(3.91)}{4.82} & \underset{(3.86)}{4.79} & \underset{(2.85)}{9.61} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.07)}{9.87} & \underset{(2.41)}{4.91} & . & . \\ \underset{(2.33)}{4.94} & \underset{(1.15)}{9.83} & . & . \\ \underset{(1.16)}{4.93} & \underset{(1.20)}{4.91} & \underset{(1.13)}{9.81} & \underset{(2.30)}{4.92} \\ \underset{(1.14)}{4.94} & \underset{(1.21)}{4.92} & \underset{(2.32)}{4.91} & \underset{(1.09)}{9.84} \end{bmatrix}$
	#Mo	5.00	22.00	14.00
	Bias	-0.0982	-0.2065	-0.1069
	MSE	1.69	3.18	1.54
		n=2	n=4	
		GMM	GMM	block-recursive GMM
$T = 1000$	$\hat{B}$	$\begin{bmatrix} \underset{(0.24)}{9.96} & \underset{(0.46)}{5.00} \\ \underset{(0.48)}{4.97} & \underset{(0.22)}{9.97} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.26)}{9.92} & \underset{(0.53)}{4.99} & \underset{(0.64)}{0.00} & \underset{(0.54)}{0.02} \\ \underset{(0.51)}{4.95} & \underset{(0.29)}{9.94} & \underset{(0.61)}{0.00} & \underset{(0.53)}{0.02} \\ \underset{(0.73)}{4.95} & \underset{(0.72)}{4.99} & \underset{(0.56)}{9.92} & \underset{(0.65)}{4.99} \\ \underset{(0.69)}{4.95} & \underset{(0.66)}{4.99} & \underset{(0.75)}{4.96} & \underset{(0.43)}{9.95} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.22)}{9.97} & \underset{(0.48)}{5.02} & . & . \\ \underset{(0.46)}{4.97} & \underset{(0.24)}{9.99} & . & . \\ \underset{(0.25)}{4.98} & \underset{(0.28)}{5.01} & \underset{(0.21)}{9.96} & \underset{(0.40)}{4.99} \\ \underset{(0.25)}{4.98} & \underset{(0.27)}{5.01} & \underset{(0.41)}{4.98} & \underset{(0.20)}{9.97} \end{bmatrix}$
	#Mo	5.00	22.00	14.00
	Bias	-0.0262	-0.0295	-0.0124
	MSE	0.35	0.57	0.31

The table reports the average $\bar{b}_{ij}$ and the corresponding estimated MSE (in parentheses) of each estimated element in $\hat{B}$ as well as the BIAS and MSE across estimated elements in $\hat{B}$ over $3,500$ Monte Carlo replicates. We estimate the GMM estimator without restrictions for $n = 2$ and $n = 4$, and the block-recursive GMM estimator for $n = 4$, which uses zero restrictions highlighted by the dots.

the GMM estimator with $n = 4$ compared to the GMM estimator with $n = 2$). The Monte Carlo study illustrates how in a typical macroeconomic application, which rarely or if at all contains more than a few hundred observations, data-driven estimates based on non-Gaussianity become less reliable the more variables the SVAR contains. However, the simulation also stresses that exploiting the block-recursive structure annihilates the deterioration of the performance induced by a larger model. That is, the average bias and MSE for each sample size in Table 4.1 is at least 1.8 and 1.8 times higher for the GMM estimator with $n = 4$ compared to the block-recursive GMM estimator with $n = 4$. Using the block-recursive structure allows to identify the four elements on the lower left of B_0 (each with a value of 5) only by covariance moment conditions (which explains why the average MSE of the block-recursive GMM estimator with $n = 4$ even can be lower than or comparable to the GMM estimator with $n = 2$, which relies on higher-order moment conditions).

Our results suggest that if in a given application well-justified restrictions are available, these restrictions should be used as they substantially improve the performance of the estimator.

4.5.2 Recursive Structure

In this subsection, we simulate a recursive SVAR using $n = 4$ variables and

$$B_0 = \begin{pmatrix} 10 & 0 & 0 & 0 \\ 5 & 10 & 0 & 0 \\ 5 & 5 & 10 & 0 \\ 5 & 5 & 5 & 10 \end{pmatrix}. \tag{4.26}$$

For the estimation of B_0, we impose a recursive order for all considered estimators, i.e., we use zero restrictions for all elements where B_0 is zero. In this setup, the pGMM, GMM-Oracle, and the oGMM estimator are efficient estimators and have a smaller asymptotic variance than the GMM estimator, which is equivalent to the estimator obtained by applying a Cholesky decomposition. By using a recursive structure, we can apply Proposition 4.3 to calculate whether an overidentifying moment condition is relevant or redundant. Therefore, we can analyze whether the pGMM estimator selects relevant moment conditions and does not select redundant moment conditions. With the imposed recursive order, the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)]$ contain 10 and the overidentifying conditions $E[f_{\mathbf{D}}(B, u_t)]$ contain 47 conditions. All moment conditions in $E[f_{\mathbf{D}}(B, u_t)]$ are valid. More precisely, 17 of overidentifying conditions are redundant and 30 overidentifying conditions are relevant.

The construction of the weights for the pGMM estimator as in Equation (4.23) requires an initial consistent estimate $\hat{B}$ to estimate β^* and the asymptotic variance in Equation (4.24). To this end, we apply the GMM estimator, which is the Cholesky estimator in this case. Moreover, we again use the assumption of independent shocks to estimate the asymptotic variance, as proposed by Keweloh (2021a). We use $r_1 = 2$ and $r_2 = 1$ in Equation (4.23) and additionally, we normalize the weights such that they sum to one, i.e., we use $\omega_j^* := \omega_j / \sum_{k \in \widetilde{D}} \omega_k$, allowing for straightforward comparison among the weights.

We choose the optimal λ for the pGMM estimator with 5-fold cross-validation from a sequence of 10 potential values. The maximum value of the sequence of λ's depends on the sample size, ensuring that it is large enough to select all moments j for which $\omega_j^* > 10^{-4}$.[32] We also include $\lambda = 0$ in the range of possible values to allow our estimator to simplify to the recursive SVAR. The selection of the optimal tuning parameter is based on the median of the GMM loss of each left-out fold.

Table 4.2 summarizes the results of $M = 3,500$ Monte Carlo simulations. We report the same summary statistics as in Table 4.1. In addition, we calculate the average number of moments selected by the pGMM estimator and the median of the chosen λ's for the pGMM estimator across Monte Carlo runs. In Appendix 4.10, we display results including the Post-pGMM estimator which uses the moments selected by pGMM in a second stage estimation.

The GMM estimator performs well in the smallest sample size in terms of bias and MSE. However, the GMM estimator is asymptotically inefficient and has the largest MSE among all considered estimators for $T = 250$ and $T = 1000$. Due to many moments, the oGMM estimator performs worst in terms of bias and MSE among the considered estimators for $T = 100$. Yet, its performance improves with sample size and it eventually outperforms the GMM estimator in terms of MSE. The bias is highest for the oGMM and GMM-Oracle estimator across sample sizes, which might be explained by the greater number of moments used by these estimators. Note that both estimators are asymptotically efficient. Nevertheless, many moment conditions can still lead to a finite sample bias. The MSE of the GMM-Oracle estimator is already comparable to the GMM

[32] We specify the maximum value of the sequence of λ's in a data-driven way using the subgradient of Equation (4.22) with respect to β. We give more details on how to construct the maximum value of the sequence of λ's in the cross-validation in Appendix 4.8.6.

Table 4.2: Finite sample performance - Recursive SVAR and the pGMM estimator.

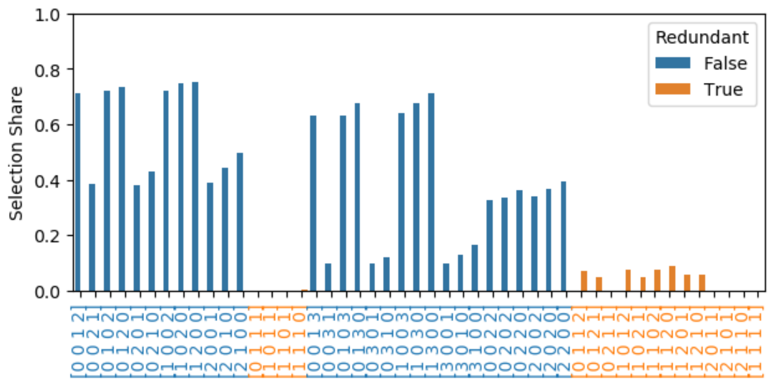
		GMM	oGMM	GMM-Oracle	pGMM
$T = 100$	$\hat{B}$	$\begin{bmatrix} \underset{(1.09)}{9.93} & . & . & . \\ \underset{(1.21)}{4.98} & \underset{(1.02)}{9.86} & . & . \\ \underset{(1.49)}{4.97} & \underset{(1.29)}{4.95} & \underset{(1.12)}{9.83} & . \\ \underset{(1.71)}{4.96} & \underset{(1.46)}{4.93} & \underset{(1.27)}{4.91} & \underset{(1.08)}{9.78} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.07)}{9.77} & . & . & . \\ \underset{(1.31)}{4.90} & \underset{(1.01)}{9.71} & . & . \\ \underset{(1.69)}{4.89} & \underset{(1.43)}{4.88} & \underset{(1.10)}{9.70} & . \\ \underset{(2.07)}{4.90} & \underset{(1.74)}{4.88} & \underset{(1.46)}{4.88} & \underset{(1.09)}{9.69} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.07)}{9.76} & . & . & . \\ \underset{(1.17)}{4.91} & \underset{(1.02)}{9.70} & . & . \\ \underset{(1.50)}{4.91} & \underset{(1.26)}{4.88} & \underset{(1.10)}{9.69} & . \\ \underset{(1.81)}{4.92} & \underset{(1.51)}{4.88} & \underset{(1.25)}{4.88} & \underset{(1.10)}{9.67} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.09)}{9.96} & . & . & . \\ \underset{(1.15)}{5.00} & \underset{(1.01)}{9.88} & . & . \\ \underset{(1.46)}{4.98} & \underset{(1.22)}{4.96} & \underset{(1.11)}{9.85} & . \\ \underset{(1.71)}{4.99} & \underset{(1.42)}{4.96} & \underset{(1.21)}{4.95} & \underset{(1.10)}{9.82} \end{bmatrix}$
	#Mo	10.00	57.00	40.00	24.22
	Bias	−0.0883	−0.1806	−0.1804	−0.0650
	MSE	1.27	1.40	1.28	1.25
	λ	.	.	.	71.08
		GMM	oGMM	GMM-Oracle	pGMM
$T = 250$	$\hat{B}$	$\begin{bmatrix} \underset{(0.43)}{9.97} & . & . & . \\ \underset{(0.51)}{4.99} & \underset{(0.43)}{9.96} & . & . \\ \underset{(0.64)}{4.98} & \underset{(0.52)}{5.00} & \underset{(0.45)}{9.93} & . \\ \underset{(0.72)}{4.98} & \underset{(0.61)}{4.99} & \underset{(0.51)}{4.98} & \underset{(0.45)}{9.91} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.40)}{9.90} & . & . & . \\ \underset{(0.49)}{4.96} & \underset{(0.40)}{9.90} & . & . \\ \underset{(0.65)}{4.96} & \underset{(0.51)}{4.97} & \underset{(0.42)}{9.87} & . \\ \underset{(0.73)}{4.97} & \underset{(0.61)}{4.96} & \underset{(0.49)}{4.97} & \underset{(0.42)}{9.86} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.40)}{9.90} & . & . & . \\ \underset{(0.44)}{4.97} & \underset{(0.40)}{9.90} & . & . \\ \underset{(0.59)}{4.97} & \underset{(0.46)}{4.97} & \underset{(0.42)}{9.87} & . \\ \underset{(0.65)}{4.98} & \underset{(0.54)}{4.97} & \underset{(0.44)}{4.96} & \underset{(0.42)}{9.85} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.42)}{9.99} & . & . & . \\ \underset{(0.45)}{5.01} & \underset{(0.41)}{9.97} & . & . \\ \underset{(0.59)}{5.01} & \underset{(0.46)}{5.02} & \underset{(0.42)}{9.94} & . \\ \underset{(0.66)}{5.02} & \underset{(0.55)}{5.01} & \underset{(0.44)}{5.00} & \underset{(0.43)}{9.92} \end{bmatrix}$
	#Mo	10.00	57.00	40.00	27.20
	Bias	−0.0311	−0.0676	−0.0656	−0.0114
	MSE	0.53	0.51	0.48	0.48
	λ	.	.	.	118.92
		GMM	oGMM	GMM-Oracle	pGMM
$T = 1000$	$\hat{B}$	$\begin{bmatrix} \underset{(0.11)}{10.00} & . & . & . \\ \underset{(0.13)}{5.00} & \underset{(0.11)}{9.99} & . & . \\ \underset{(0.15)}{4.99} & \underset{(0.13)}{4.99} & \underset{(0.11)}{9.99} & . \\ \underset{(0.19)}{4.99} & \underset{(0.15)}{4.99} & \underset{(0.13)}{4.99} & \underset{(0.11)}{9.98} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.10)}{9.98} & . & . & . \\ \underset{(0.12)}{4.99} & \underset{(0.10)}{9.97} & . & . \\ \underset{(0.13)}{4.99} & \underset{(0.11)}{4.99} & \underset{(0.10)}{9.98} & . \\ \underset{(0.16)}{4.99} & \underset{(0.14)}{4.99} & \underset{(0.11)}{4.99} & \underset{(0.10)}{9.97} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.10)}{9.98} & . & . & . \\ \underset{(0.11)}{4.99} & \underset{(0.10)}{9.97} & . & . \\ \underset{(0.13)}{4.99} & \underset{(0.10)}{4.99} & \underset{(0.10)}{9.98} & . \\ \underset{(0.15)}{4.99} & \underset{(0.13)}{4.99} & \underset{(0.11)}{4.99} & \underset{(0.10)}{9.97} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.11)}{10.00} & . & . & . \\ \underset{(0.11)}{5.00} & \underset{(0.10)}{9.99} & . & . \\ \underset{(0.13)}{5.00} & \underset{(0.11)}{5.00} & \underset{(0.10)}{10.00} & . \\ \underset{(0.16)}{5.00} & \underset{(0.13)}{5.00} & \underset{(0.11)}{5.00} & \underset{(0.10)}{9.98} \end{bmatrix}$
	#Mo	10.00	57.00	40.00	29.59
	Bias	−0.0076	−0.0158	−0.0158	−0.0021
	MSE	0.13	0.12	0.11	0.12
	λ	.	.	.	75.34

The table reports the average $\bar{b}_{ij}$ and the corresponding estimated MSE (in parentheses) of each estimated element in $\hat{B}$ as well as the BIAS and MSE across estimated elements in $\hat{B}$ over 3,500 Monte Carlo replicates for the GMM estimator, the oGMM estimator, the GMM-Oracle estimator, and the pGMM estimator. All estimator use zero restrictions which are highlighted by the dots.

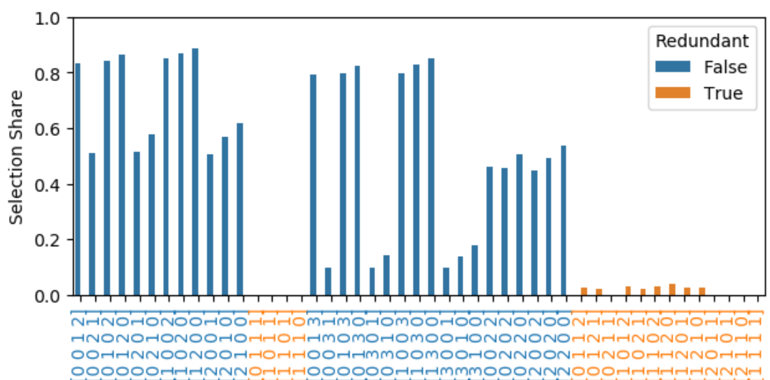
estimator in small samples. Relative to the other estimators, its MSE further decreases with the sample size and it performs best in the largest sample size. In general, the GMM-Oracle estimator is infeasible since the redundant moments are unknown a priori.In contrast to that, the pGMM estimator is feasible and uses a data-driven approach to select the relevant and valid moments. The pGMM estimator performs well across all sample sizes in terms of bias and MSE. For $T = 100$, its bias and MSE is notably smaller than the one of the oGMM and the GMM-Oracle estimator and surprisingly, also smaller than the one of the GMM estimator. In the largest sample, the pGMM estimator performs similar to the oGMM and GMM-Oracle estimator in terms of MSE and best in terms of bias. The Post-pGMM estimator reported in Appendix 4.10 performs similar to the pGMM estimator. The simulation shows that the pGMM estimator can, without prior specification, distinguish informative from non-informative overidentifying moments, which solves the many moments problem of the oGMM estimator and allows to exploit information in overidentifying higher-order moments already in small samples.

Table 4.2 indicates that the average number of selected moments increases only slightly as T increases. Even for $T = 1000$, the pGMM estimator only selects 20 out of 30 valid and relevant overidentifying moments in addition to the 10 identifying moments. That said, the remaining 10 moments would only decrease the MSE from 0.12 to 0.11, indicating that the moments not being selected would not lower the MSE much. Figure 4.2 illustrates that pGMM estimator only selects relevant moments and manages to leave out redundant moments, especially as T increases. Moreover, the share of selections of each moment across all Monte Carlo runs rises with the sample size for the majority of relevant moments. In Figure 4.7, we plot the average weight of each moment across Monte Carlo runs. By comparing Figure 4.2 and Figure 4.7, we argue that there is a clear correlation between the average weight and the number of selections of each moment. More precisely, all redundant moments have an average weight which is very close to zero and hence, they are not selected by the pGMM estimator.

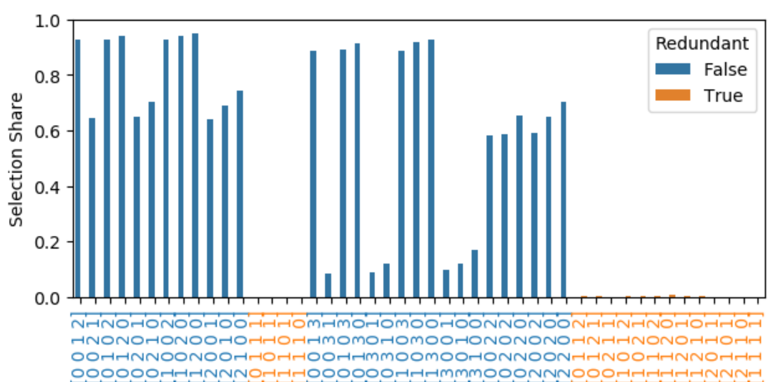
Figure 4.2: Finite sample performance - Share of selected moments by the pGMM estimator.



(a) $T = 100$



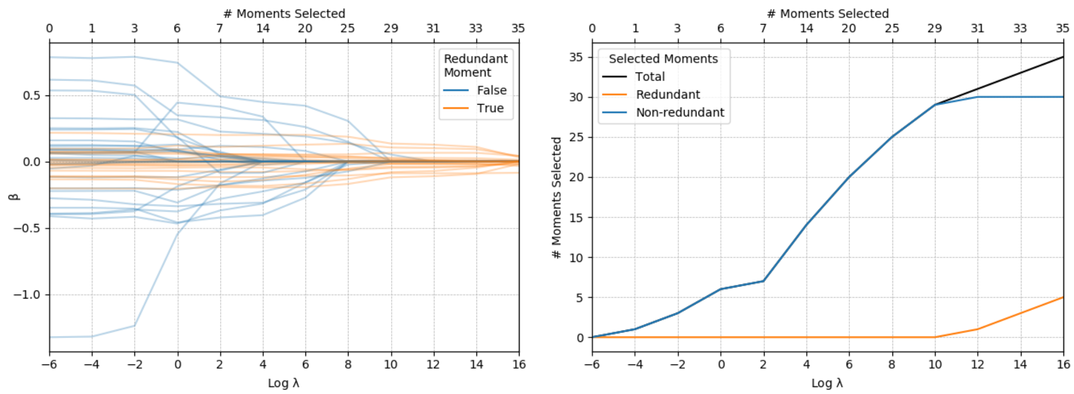
(b) $T = 250$



(c) $T = 1000$

Note: The figure shows how often each moment gets selected. Redundant moments (orange) and relevant moments (blue) are displayed on the x-axis. Each x-axis label abbreviates a moment condition, e.g., $[0, 1, 2, 1]$ corresponds to $E[e(B)^0_{1,t}\, e(B)^1_{2,t}\, e(B)^2_{3,t}\, e(B)^1_{4,t}]$.

Figure 4.3: Finite sample performance - Illustration of influence of λ on β for the pGMM estimator.



(a) Trace Plot

(b) Selected Moments

Note: Panel (a) of the figure shows the values of β in dependence on $log(\lambda)$ for one Monte Carlo run for $T = 100$ and the corresponding number of selected moments in $\widetilde{D}$. Panel (b) of the figure splits the number of selected moments into the number of selected redundant and the number of selected relevant moments for each $log(\lambda)$.

Figure 4.3 highlights the influence of λ on β and hence, on the number of selected moment conditions for one Monte Carlo run. For the purpose of illustration, we use a wider range of of λ values for this plot. For instance, for $\log(\lambda) = -6$ no overidentifying moment conditions are selected and the solution of the pGMM estimator corresponds to the one of the GMM estimator. Further, the number of selected moments increases as λ increases, i.e., the penalty shrinks the elements of β to zero. Furthermore, the relevant moments get selected first when λ increases and we do not select any redundant moment until λ becomes very large.

4.6 Disentangling speculative demand and supply shocks in the oil market

In this section, we propose a SVAR model for the oil market to analyze the impact of flow and speculative supply and of flow and speculative demand shocks on the real oil price. A flow oil supply shock represents an exogenous deviation in the present amount of oil coming out of the ground and a flow oil demand shock represents an exogenous deviation in the present amount of oil being consumed. A speculative oil supply shock represents a shift in the expected future oil supply and a speculative oil demand shock a shift in the expected future oil demand.

We consider a SVAR with monthly data from January 1974 to December 2019 of the form

$$\begin{bmatrix} O_t \\ Y_t \\ OP_t \\ SR_t \end{bmatrix} = \alpha + \sum_{i=1}^{12} A_i \begin{bmatrix} O_{t-i} \\ Y_{t-i} \\ OP_{t-i} \\ SR_{t-i} \end{bmatrix} + \begin{bmatrix} u_t^O \\ u_t^Y \\ u_t^{OP} \\ u_t^{SR} \end{bmatrix}. \tag{4.27}$$

The variable O_t is the log difference of global oil production, Y_t is the log difference of industrial production, measuring economic activity, OP_t is the growth rate of real oil price, and SR_t are real monthly stock returns.[33] We decompose the reduced form shocks u_t into four structural shocks with

$$\begin{bmatrix} u_t^O \\ u_t^Y \\ u_t^{OP} \\ u_t^{SR} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & 0 & 0 \\ b_{21} & b_{22} & 0 & 0 \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{41} \end{bmatrix} \begin{bmatrix} \varepsilon_t^s \\ \varepsilon_t^d \\ \varepsilon_t^{s-exp} \\ \varepsilon_t^{d-exp} \end{bmatrix}, \tag{4.28}$$

where ε_t^s is a flow supply shock for oil, ε_t^d is a flow demand shock for oil, ε_t^{s-exp} is a speculative oil supply shock, and ε_t^{d-exp} is a speculative oil demand shock. The block-recursive restrictions in Equation (4.28) imply that oil production and economic activity behave sluggishly and can contemporaneously only respond to flow supply and demand shocks, whereas oil prices and stock returns can immediately incorporate all available information and contemporaneously respond to flow and speculative supply and demand shocks.

The simultaneous relationship is estimated using the block-recursive SVAR pGMM estimator.[34] In line with the Monte Carlo simulations, we apply continuous updating for the weighting matrix and use the assumption of serially and mutually independent shocks to estimate the

[33] Global oil production is given by the global crude oil including lease condensate production obtained from the U.S. EIA. We obtain industrial production by the monthly industrial production index in the OECD and six major other countries from Baumeister and Hamilton (2019). The real oil price is equal to the refiner's acquisition cost of imported crude oil from the U.S. EIA deflated by the U.S. CPI. Real stock prices correspond to the aggregate U.S. stock index constructed by the OECD deflated by the U.S. CPI.

[34] In Appendix 4.11, we conduct various robustness checks. In particular, we estimate the block-recursive SVAR using the GMM estimator from Equation (4.15) and the overidentified GMM estimator from Equation (4.19). Estimates using the white fast SVAR GMM estimator proposed by Keweloh (2021b) and the PML estimator proposed by Gouriéroux et al. (2017) are qualitatively similar and available on request. Additionally, we report results for different specifications of the variables in the block-recursive SVAR.

asymptotically efficient weighting matrix as proposed by Keweloh (2021a). With the imposed block-recursive structure, we can divide the moment conditions into 14 identifying conditions $E[f_{\mathbf{N}}(B, u_t)]$ and 43 overidentifying conditions $E[f_{\mathbf{D}}(B, u_t)]$. We use the same specifications to construct the weights as in the Monte Carlo simulation, i.e., we use $r_1 = 2$ and $r_2 = 1$ in Equation (4.23). For the cross-validation, we consider a range of 28 values for λ, including $\lambda = 0$. The maximum value of λ is chosen such that all conditions $E[f_{\mathbf{D}}(B, u_t)]$ for which $\omega_j / \sum_{k \in \widetilde{D}} \omega_k > 10^{-7}$ get selected. With the chosen $\lambda = 34679$, which is the 27th value of the considered sequence, 12 coskewness and 12 cokurtosis conditions are selected.[35]

For each estimated structural shock, Table 4.3 shows the estimated skewness, kurtosis and p-value of the Jarque-Bera test. To ensure identification, at most one structural shock in each block may be Gaussian. With our block-recursive structure, each block contains only two shocks and, therefore, it is sufficient for identification to show that at least one structural shock in each block is non-Gaussian. Furthermore, the block-recursive structure implies that each of the two unmixed innovations in the first block is equal to a linear combination of the two structural shocks in the first block, i.e., if both structural shocks are Gaussian, the two unmixed innovations have to be Gaussian as well. However, the skewness, kurtosis, and the Jarque-Bera test suggest that the unmixed innovations in the first block are non-Gaussian and, hence, that at least one structural shock in the first block is non-Gaussian. Consequently, the first block is identified. Moreover, the unmixed innovations in the second block are equal to a linear combination of the structural shocks in the second block (the argument follows from Equation (4.33) in the proof of Proposition 4.2). Again, the skewness, kurtosis and the Jarque-Bera test suggest that the unmixed innovations in the second block are non-Gaussian, implying that at least one structural shock in the second block is non-Gaussian. Thus, the second block is also identified. Consequently, our block-recursive SVAR is identified.

In Figure 4.4, we show impulse response functions (IRFs).

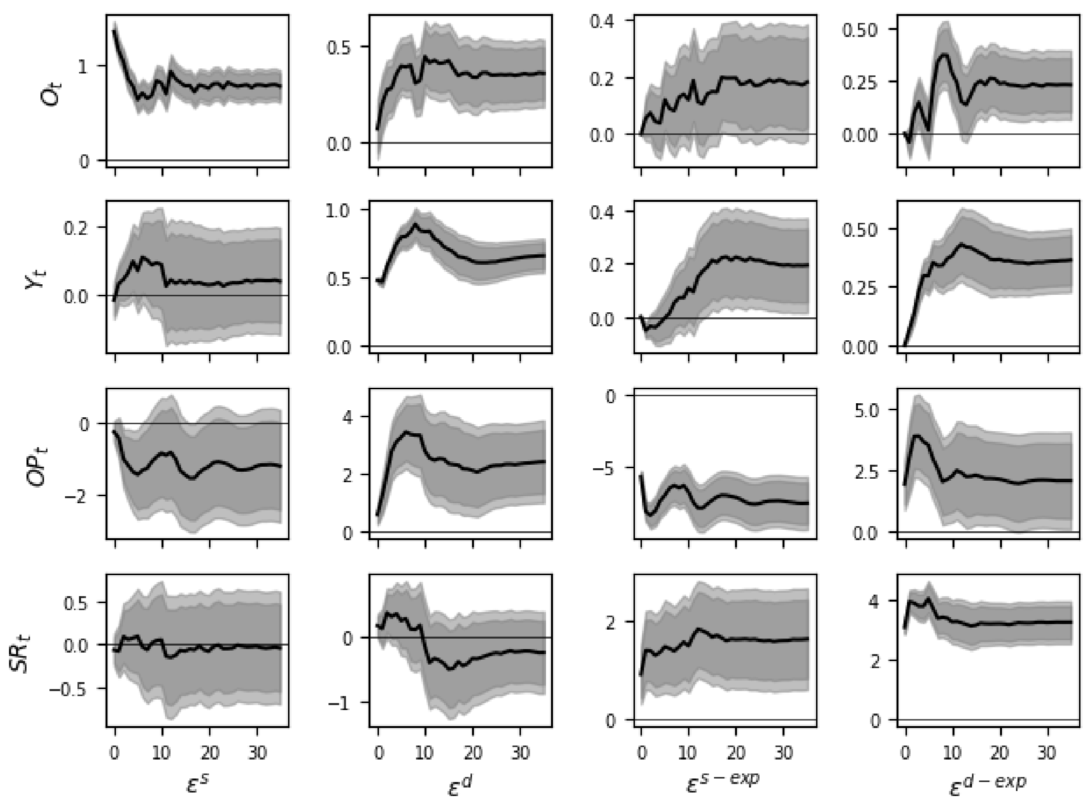
[35] Additionally, we compute the block-recursive SVAR pGMM estimator using the plugin rule $\lambda = k_{\mathbf{D}}^{r_2/4} T^{(-0.5 - r_2/4)}$, where $k_{\mathbf{D}}$ denotes the number of overidentifying moment conditions, see Cheng and Liao (2015). The estimator selects 8 coskewness and 6 cokurtosis conditions.

Table 4.3: Oil market SVAR - Non-Gaussianity of the estimated structural shocks.

	ε_t^s	ε_t^d	ε_t^{s-exp}	ε_t^{d-ext}
Skewness	-0.97	-0.21	0.46	-0.82
Kurtosis	9.92	4.58	6.79	6.88
JB-Test	0.00	0.00	0.00	0.00

Note: Skewness, kurtosis and the p-value of the Jarque-Bera test.

Figure 4.4: Oil market SVAR - Impulse responses.



Note: Impulse responses to the estimated structural shocks for the block-recursive SVAR pGMM estimator. Confidence bands are symmetric 68% and 80% bands based on standard errors and 1000 replications. The rows show the cumulative responses. The x-axis displays monthly lags.

With the block-recursive structure, labeling of the shocks in the plot of the IRFs is straightforward. In the first block, there is only one shock which leads to a significant immediate increase of economic activity and, thus, an immediate increase in demand for oil. We label this shock as the flow demand shock and the remaining shock in the first block as the flow supply shock.

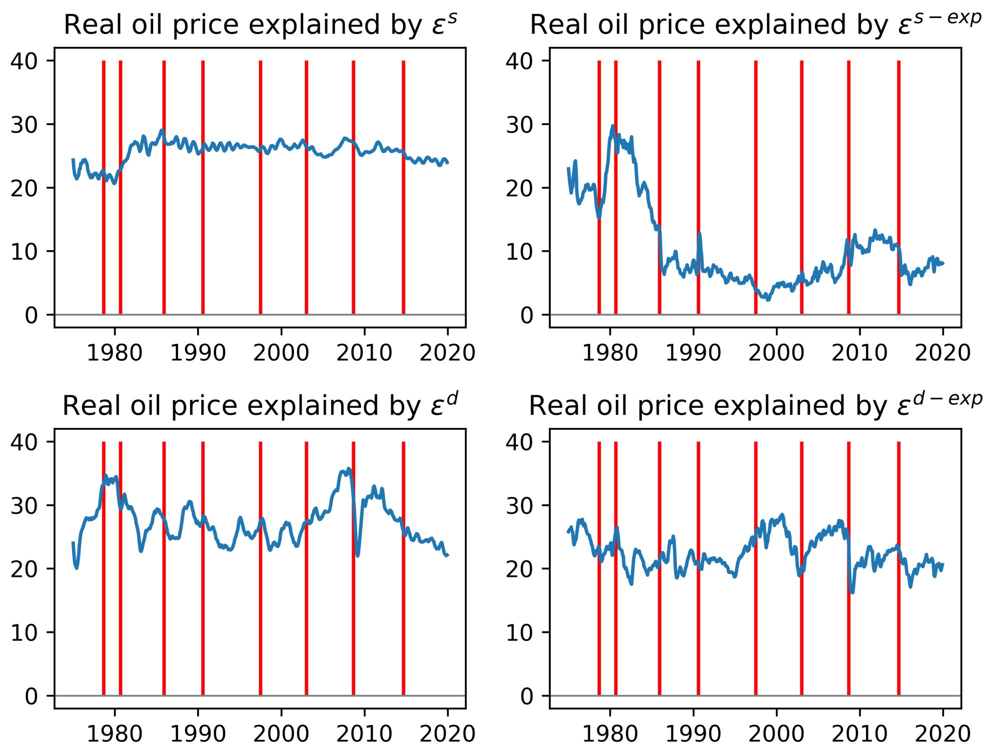
In the second block, one shock leads to an immediate increase of the real oil price and to a long-run increase of economic activity. We label this shock as the speculative oil demand shock. The remaining shock in the second block leads to an immediate decrease of the oil price and to an increase of economic activity and oil production in the long-run, which corresponds to the speculative oil supply shock.

Our results show that flow supply shocks immediately increase oil production and decrease the real oil price and flow demand shocks increase economic activity and the real oil price. Moreover, oil production responds to the demand shock with a lagged increase. Interestingly, we find that real stock returns do not respond significantly to flow demand and supply shocks. With respect to the speculative shocks, we find that a supply expectation shock leads to an increase of oil production and of economic activity after one year. Furthermore, it immediately and permanently decreases the real oil price and increases real stock returns. A speculative demand shock increases oil production and economic activity. Additionally, the speculative demand shocks leads to an immediate increase of the real oil price and of real stock returns.

Figure 4.5 shows the contribution of the estimated structural shocks to the evolution of the real oil price. Figure 4.8 in Appendix 4.11 shows the historical evolution of the real oil price. Figure 4.5 suggests that the increase of the real oil price from 1978 to 1981 is mainly driven by flow supply and speculative supply shocks. Moreover, we find that the decline of the real oil price from 1981 to 1985 is largely explained by speculative supply shocks. Additionally, the decrease in real oil prices after the collapse of OPEC in 1985 and the peak of real oil prices during the Persian Gulf War in 1990 can to a large extent be explained explained by speculative supply shocks. The run-up in the real oil prices from 2003 to 2008 is driven by flow demand, speculative demand, and speculative supply shocks. Flow demand and speculative demand shocks explain the plunge of the real oil price during the financial crisis in 2008. Additionally, most of the recovery of the real oil price after the financial crisis is explained by demand shocks. The collapse of the real oil price since mid 2014 is related to flow demand, speculative demand, and speculative supply shocks.

The IRFs in Figure 4.4 show no evidence against a recursive structure of the shocks in the first block. That said, our results clearly suggest that the second block does not have a recursive structure since the two structural shocks in the second block have an immediate impact on both

Figure 4.5: Oil market SVAR - Real oil price evolution explained by the estimated structural shocks.



Note: In each of the panels, we simulate the real oil price (blue line) by setting all but one of the shocks to zero (and for ease of interpretation, we also set $\alpha = 0$ in Equation (4.27)). The red vertical bars indicate the following events: Iranian Revolution (1978 : 9), Iran Iraq War (1980 : 9), collapse of OPEC (1985 : 12), Persian Gulf War (1990 : 8), Asian Financial Crisis of (1997 : 7), Iraq War (2003 : 1), the collapse of Lehman Brothers (2008 : 9), and the oil price decline in mid 2014.

reduced form shocks in the second block. As a robustness-check and to illustrate the impact of misspecification in the second block, we estimate a recursive specification as proposed in Kilian and Park (2009). That is, we restrict b_{12} and b_{34} in Equation (4.28) to zero. In this case, the interpretation of the shocks changes and we refer to the third and fourth shock as speculative oil price shock and residual stock market shock, respectively.

Figure 4.10 in Appendix 4.11 displays the IRFs of the recursive SVAR. The response of the real oil price to flow supply and demand shocks in the recursive model is similar to the the one in the block-recursive model. The speculative oil price shock leads to an decrease of the real oil price. However, none of the remaining variables shows any significant response to the speculative oil price shock, except for economic activity which shows a small negative reaction in the first seven month. In the recursive SVAR for the oil market, we cannot distinguish between speculative supply and speculative demand shocks. Rather, the speculative oil price shock contains a mixture of the speculative supply and speculative demand shock. However, the impact of the speculative oil price shocks on oil production and the economy should depend on the source of the speculative oil price shock and, thus, it is not surprising that we are unable to find a clear response of oil production, economic activity, and the stock market to the speculative oil price shock in the recursive specification.

As a further robustness-check, we estimate the SVAR without any restrictions on the interaction, i.e., we estimate the model without the zero restrictions given in Equation (4.28). In this case, the labeling of the shocks is the same as in Equation (4.28). However, the difference is that oil production and economic activity can now contemporaneously respond to speculative supply and demand shocks. Figure 4.11 and Figure 4.12 in Appendix 4.11 show the corresponding IRFs. Overall, the unrestricted responses in Figure 4.11 are comparable to the block-recursive responses in Figure 4.4. However, the confidence bands are broader and there is no significant response of the real oil price to flow supply and (almost) no significant response to flow demand shocks.

4.7 Conclusion

For a non-Gaussian block-recursive SVAR, we derive a small set of identifying moment conditions based on the assumption of mean independent shocks. Moreover, we derive overidentifying moment conditions, some of these require mean independent shocks and some of these additionally require independent shocks. We show that the overidentifying conditions can decrease the asymptotic variance of the block-recursive SVAR GMM estimator. In particular, we prove that the frequently applied Cholesky estimator can be inefficient. Since some of the overidentifying moment conditions may be redundant, i.e., may not decrease the asymptotic variance, or be invalid, i.e., may lead to inconsistent estimates, we employ the block-recursive SVAR pGMM estimator to select only the relevant and valid overidentifying moment conditions.

We demonstrate in a Monte Carlo experiment that imposing a block-recursive structure substantially increases the finite sample performance compared to unrestricted estimators. Furthermore, a second Monte Carlo experiment highlights that, for a given block-recursive structure, the block-recursive SVAR pGMM estimator selects only relevant moment conditions and thereby, increases finite sample precision compared to the block-recursive SVAR GMM estimator and overidentified block-recursive SVAR GMM estimator.

Our application analyzes the impact of flow and speculative supply and flow and speculative demand shocks in the oil market. We argue that there are some but not enough well-justified restrictions available to identify the SVAR based on second moments. Traditional approaches would either rely on additional less credible restrictions or refrain from using any restrictions and solely rely on non-Gaussianity. The proposed block-recursive estimator allows to utilize only the well-justified restrictions and, therefore, offers a compromise between both approaches. The application illustrates that by combining data-driven identification with traditional zero restrictions we are able to gain deeper insights into the transmission of demand and supply shocks in the oil market.

4.8 Appendix - Supplementary Notation and Proofs

We include the formulas in Appendix 4.8.1 and 4.8.2 for completeness, even though they are standard textbook results (cf. Hall (2005)).

4.8.1 Appendix - Asymptotic variance of the block-recursive SVAR GMM estimator

The asymptotic variance of the block-recursive SVAR GMM estimator defined in Equation (4.15) is given by

$$V_{\mathbf{N}} := M_{\mathbf{N}} S_{\mathbf{N}} M'_{\mathbf{N}} \tag{4.29}$$

where

$$M_{\mathbf{N}} := \left(G'_{\mathbf{N}} W_{\mathbf{N}} G_{\mathbf{N}}\right)^{-1} G'_{\mathbf{N}} W_{\mathbf{N}}, \qquad S_{\mathbf{N}} := \lim_{T\to\infty} E\left[T g_{\mathbf{N}}(B_0) g_{\mathbf{N}}(B_0)\right],$$
$$G_{\mathbf{N}} := E\left[\frac{\partial f_{\mathbf{N}}(B_0, u_t)}{\partial vec(B)'}\right].$$

Consequently, using the weighting matrix $W^*_{\mathbf{N}} := S^{-1}_{\mathbf{N}}$ leads to the estimator $\hat{B}^*$ with the asymptotic variance

$$V^*_{\mathbf{N}} := (G'_{\mathbf{N}} S^{-1}_{\mathbf{N}} G_{\mathbf{N}})^{-1}, \tag{4.30}$$

which is the lowest possible asymptotic variance (see Hall (2005)).

4.8.2 Appendix - Asymptotic variance of the (overidentified) block-recursive SVAR GMM estimator

The asymptotic variance of the overidentified block-recursive SVAR GMM estimator defined in Equation (4.19) is given by

$$V_{\mathbf{N+D}} := M_{\mathbf{N+D}} S_{\mathbf{N+D}} M'_{\mathbf{N+D}}, \tag{4.31}$$

where

$$M_{\mathbf{N+D}} := \left(G'_{\mathbf{N+D}} W_{\mathbf{N+D}} G\right)^{-1} G'_{\mathbf{N+D}} W_{\mathbf{N+D}}, \qquad S_{\mathbf{N+D}} := \lim_{T\to\infty} E\left[g_{\mathbf{N+D}}(B_0) g_{\mathbf{N+D}}(B_0)'\right],$$

$$G_{\mathbf{N+D}} := \begin{bmatrix} G_{\mathbf{N}} \\ G_{\mathbf{D}} \end{bmatrix}, \qquad g_{\mathbf{N+D}}(B_0) := \begin{bmatrix} g_{\mathbf{N}}(B_0) \\ g_{\mathbf{D}}(B_0) \end{bmatrix},$$

$$G_{\mathbf{D}} := E\left[\frac{\partial f_{\mathbf{D}}(B_0, u_t)}{\partial vec(B)'}\right].$$

Using the weighting matrix $W^*_{\mathbf{N+D}} := S^{-1}_{\mathbf{N+D}}$ leads to the estimator $\hat{B}^*_{\mathbf{N+D}}$ with the asymptotic variance

$$V^*_{\mathbf{N+D}} := (G'_{\mathbf{N+D}} S^{-1}_{\mathbf{N+D}} G_{\mathbf{N+D}})^{-1}, \tag{4.32}$$

which is the lowest possible asymptotic variance (see Hall (2005)). To construct $V_{\mathbf{N+D}_j}$ and $V^*_{\mathbf{N+D}_j}$, $j \in \widetilde{D}$, we replace the moment conditions $f_{\mathbf{D}_j}(B, u_t)$ by moment condition $f_{\mathbf{D}_j}(B, u_t)$, $j \in \widetilde{D}$, in Equation (4.31) and (4.32).

4.8.3 Appendix - Identification in the block-recursive SVAR

Proof of Proposition 4.1.

For ease of notation, we omit the time index t and w.l.o.g., consider an example with two blocks[36]

$$\begin{bmatrix} u_{p_1} \\ u_{p_2} \end{bmatrix} = \begin{bmatrix} B_{11,0} & 0 \\ B_{21,0} & B_{22,0} \end{bmatrix} \begin{bmatrix} \varepsilon_{p_1} \\ \varepsilon_{p_2} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{bmatrix},$$

[36] If the SVAR contains more than two blocks, the procedure outlined in the proof can be repeated multiple times to identify arbitrary many blocks. For example, a SVAR with three blocks

$$\begin{bmatrix} u_{p_1} \\ u_{p_2} \\ u_{p_3} \end{bmatrix} = \begin{bmatrix} B_{11,0} & 0 & 0 \\ B_{21,0} & B_{22,0} & 0 \\ B_{32,0} & B_{32,0} & B_{33,0} \end{bmatrix} \begin{bmatrix} \epsilon_{p_1} \\ \epsilon_{p_2} \\ \epsilon_{p_3} \end{bmatrix} \text{ can be written as } \begin{bmatrix} u_{p_1} \\ \tilde{u}_{p_2} \end{bmatrix} = \begin{bmatrix} B_{11,0} & 0 \\ \tilde{B}_{21,0} & \tilde{B}_{22,0} \end{bmatrix} \begin{bmatrix} \epsilon_{p_1} \\ \tilde{\epsilon}_{p_2} \end{bmatrix},$$

with $\tilde{u}_{p_2} = [u'_{p_2}, u'_{p_3}]'$, $\tilde{B}_{22,0} = \begin{bmatrix} B_{22,0} & 0 \\ B_{32,0} & B_{33,0} \end{bmatrix}$, $\tilde{B}_{21,0} = \begin{bmatrix} B_{21,0} \\ B_{31,0} \end{bmatrix}$,and $\tilde{\epsilon}_{p_2} = [\epsilon'_{p_2}, \epsilon'_{p_3}]'$. Our proof then shows how to identify $B_{11,0}$, $\tilde{B}_{21,0} = \begin{bmatrix} B_{21,0} \\ B_{31,0} \end{bmatrix}$, and ϵ_{p_1}. Defining

$$\begin{bmatrix} z_{p_2} \\ z_{p_3} \end{bmatrix} := \begin{bmatrix} u_{p_2} \\ u_{p_3} \end{bmatrix} - \begin{bmatrix} B_{21,0} \\ B_{31,0} \end{bmatrix} \epsilon_{p_1} \text{ then yields } \begin{bmatrix} z_{p_2} \\ z_{p_3} \end{bmatrix} = \begin{bmatrix} B_{22,0} & 0 \\ B_{32,0} & B_{33,0} \end{bmatrix} \begin{bmatrix} \epsilon_{p_2} \\ \epsilon_{p_3} \end{bmatrix},$$

which is another block-recursive SVAR with two blocks.

where u_{p_1} and u_{p_2} contain the reduced form shocks of the first and second block, ε_{p_1} and ε_{p_2} contain the structural shocks of the first and second block, and $B_{11,0}$, $B_{21,0}$, $B_{22,0}$, B_{11}, B_{21}, and B_{22} are the corresponding blocks of the matrices B_0 and B.

First, let $E[f_{\mathbf{2_{p_1}}}(B,u)] = 0$ contain all (co-)variance conditions of shocks in the first block. The block-recursive structure implies that $u_{p_1} = B_{11,0}\varepsilon_{p_1}$. If at most one structural shock in the first block has zero excess kurtosis, it follows from Lanne and Luoto (2021) that the conditions containing only shocks in the first block

$$E\begin{bmatrix} f_{\mathbf{2_{p_1}}}(B,u) \\ f_{\mathbf{4_{p_1}}}(B,u) \end{bmatrix} = 0$$

locally identify $B_{11} = B_{11,0}$, the impact of the shocks in the first block on the variables in the first block.

Second, let $E\left[f_{\mathbf{2_{p_1 p_2}}}(B,u)\right] = 0$ contain all covariance conditions belonging to shocks in both blocks. At the local solution $B_{11} = B_{11,0}$, the covariance conditions containing shocks of both blocks only hold if $B_{21} = B_{21,0}$. To see this, rewrite the covariance conditions as $E\left[e_{p_2}(B)e_{p_1}(B)'\right] = 0$. With the partitioned inverse of B and the block-recursive structure, it holds that $e_{p_2}(B) = -B_{22}^{-1}B_{21}B_{11}^{-1}B_{11,0}\varepsilon_{p_1} + B_{22}^{-1}\left(B_{21,0}\varepsilon_{p_1} + B_{22,0}\varepsilon_{p_2}\right)$. Therefore, with $B_{11} = B_{11,0}$ it holds that

$$E\left[e_{p_2}(B)e_{p_1}(B)'\right] = -B_{22}^{-1}B_{21}E\left[\varepsilon_{p_1}\varepsilon_{p_1}'\right] + B_{22}^{-1}B_{21,0}E\left[\varepsilon_{p_1}\varepsilon_{p_1}'\right] + B_{22,0}E\left[\varepsilon_{p_2}\varepsilon_{p_1}'\right].$$

With $E\left[\varepsilon_{p_1}\varepsilon_{p_1}'\right] = I$ and $E[\varepsilon_{p_2}\varepsilon_{p_1}'] = 0$, the condition $E\left[e_{p_2}(B)e_{p_1}(B)'\right] = 0$ implies $0 = -B_{22}^{-1}\left(B_{21} - B_{21,0}\right)$ at $B_{11} = B_{11,0}$. Therefore, at the local solution $B_{11} = B_{11,0}$ the covariance conditions $E\left[f_{\mathbf{2_{p_1 p_2}}}(B,u)\right]$, globally identify $B_{21} = B_{21,0}$ the impact of shocks in the first block on variables in the second block.

Finally, let $E[f_{\mathbf{2_{p_2}}}(B,u)] = 0$ contain all (co-)variance conditions of shocks in the second block. At the solution $B_{11} = B_{11,0}$ and $B_{21} = B_{21,0}$ the unmixed innovations of the second block $e_{p_2}(B)$ are mixtures of the structural shocks in the second block and are not influenced by shocks from the first block. This follows from the partitioned inverse of B and the block-recursive structure such that $e_{p_2}(B) = B_{22}^{-1}B_{22,0}\varepsilon_{p_2}$. If at most one structural shock in the second block has zero excess kurtosis, it then again follows from Lanne and Luoto (2021) that at the solution $B_{11} = B_{11,0}$ and $B_{21} = B_{21,0}$ the remaining conditions containing only shocks in the second block

$$E \begin{bmatrix} f_{\mathbf{2_{p_2}}}(B,u) \\ f_{\mathbf{4_{p_2}}}(B,u) \end{bmatrix} = 0$$

locally identify $B_{22} = B_{22,0}$, meaning the impact of shocks in the second block on variables in the second block.

□

Proof of Proposition 4.2.

To simplify the notation let

$$\begin{aligned}
\tilde{u}_1 &:= [u_1, ..., u_{p_i-1}]', & \tilde{e}_1(B) &:= [e_1(B), ..., e_{p_i-1}(B)]', & \tilde{\varepsilon}_1 &:= [\varepsilon_1, ..., \varepsilon_{p_i-1}]', \\
\tilde{u}_2 &:= [u_{p_i}, ..., u_{p_{i+1}-1}]', & \tilde{e}_2(B) &:= [e_{p_i}(B), ..., e_{p_{i+1}-1}(B)]', & \tilde{\varepsilon}_2 &:= [\varepsilon_{p_i}, ..., \varepsilon_{p_{i+1}-1}]', \\
\tilde{u}_3 &:= [u_{p_{i+1}}, ..., u_n]', & \tilde{e}_3(B) &:= [e_{p_{i+1}}(B), ..., e_n(B)]', & \tilde{\varepsilon}_3 &:= [\varepsilon_{p_{i+1}}, ..., \varepsilon_n]',
\end{aligned}$$

such that $\tilde{u}_1$, $\tilde{e}_1(B)$, and $\tilde{\varepsilon}_1$ contain all reduce form shocks, unmixed innovations, and structural shocks in blocks preceding the ith block of $\mathbb{B}_{brec}$, $\tilde{u}_2$, $\tilde{e}_2(B)$, and $\tilde{\varepsilon}_2$ contain the innovations and shocks in the i-th block of $\mathbb{B}_{brec}$, and $\tilde{u}_3$, $\tilde{e}_3(B)$, and $\tilde{\varepsilon}_3$ contain the innovations and shocks following block i of $\mathbb{B}_{brec}$. Moreover, we denote parts of the B_0 matrix as follows

$$\begin{bmatrix} \tilde{u}_1 \\ \tilde{u}_2 \\ \tilde{u}_3 \end{bmatrix} = \begin{bmatrix} B_{11,0} & 0 & 0 \\ B_{21,0} & B_{22,0} & 0 \\ B_{31,0} & B_{32,0} & B_{33,0} \end{bmatrix} \begin{bmatrix} \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \\ \tilde{\varepsilon}_3 \end{bmatrix},$$

and B_{11}, B_{21}, B_{31}, B_{22}, B_{32}, and B_{33} denote the respective parts of a given B matrix.

With the block-recursive structure and the partitioned inverse, it holds that

$$\begin{aligned}
\tilde{e}_1(B) &= B_{11}^{-1} B_{11,0} \tilde{\varepsilon}_1, \\
\tilde{e}_2(B) &= -B_{22}^{-1} B_{21} B_{11}^{-1} B_{11,0} \tilde{\varepsilon}_1 + B_{22}^{-1} \left(B_{21,0} \tilde{\varepsilon}_1 + B_{22,0} \tilde{\varepsilon}_2 \right).
\end{aligned}$$

For any matrix B satisfying $E\left[f_{\mathbf{2}}(B,u_t)\right]=0$ and, therefore, $0=E\left[\tilde{e}_2(B)\tilde{e}_1(B)'\right]$ it holds that $0=-B_{22}^{-1}\left(B_{21,0}-B_{21}B_{11}^{-1}B_{11,0}\right)B_{11,0}'(B_{11}^{-1})'$ and, thus, $B_{21}=B_{21,0}B_{11,0}^{-1}B_{11}$. Any B Matrix satisfying the condition $0=E\left[\tilde{e}_2(B)\tilde{e}_1(B)'\right]$ thus yields innovations of the second block equal to

$$\tilde{e}_2(B)=B_{22}^{-1}B_{22,0}\tilde{\varepsilon}_2, \tag{4.33}$$

meaning the innovations of the second block are equal to a linear combination of the structural shocks in the second block. Applying the identification result from Lanne and Luoto (2021) yields that the conditions $E\left[f_{\mathbf{4}_{\tilde{\mathbf{p}}_{\mathbf{i}}}}(B,u_t)\right]=0$ locally identify $B_{22,0}$.

Analogously, with the block-recursive structure and the partitioned inverse it holds that

$$\begin{aligned}\tilde{e}_3(B)=&-B_{33}^{-1}\begin{bmatrix}B_{31} & B_{32}\end{bmatrix}\begin{bmatrix}B_{11} & 0\\ B_{21} & B_{22}\end{bmatrix}^{-1}\begin{bmatrix}B_{11,0} & 0\\ B_{21,0} & B_{22,0}\end{bmatrix}\begin{bmatrix}\tilde{\varepsilon}_1\\ \tilde{\varepsilon}_2\end{bmatrix}\\ &+B_{33}^{-1}\left(\begin{bmatrix}B_{31,0} & B_{32,0}\end{bmatrix}\begin{bmatrix}\tilde{\varepsilon}_1\\ \tilde{\varepsilon}_2\end{bmatrix}+B_{33,0}\tilde{\varepsilon}_3\right).\end{aligned}$$

With $B_{21}=B_{21,0}B_{11,0}^{-1}B_{11}$ it follows that

$$\begin{aligned}\tilde{e}_3(B)=&-B_{33}^{-1}\begin{bmatrix}B_{31} & B_{32}\end{bmatrix}\begin{bmatrix}B_{11}^{-1} & 0\\ -B_{22}^{-1}B_{21,0}B_{11,0}^{-1}B_{11}B_{11}^{-1} & B_{22}^{-1}\end{bmatrix}\begin{bmatrix}B_{11,0} & 0\\ B_{21,0} & B_{22,0}\end{bmatrix}\begin{bmatrix}\tilde{\varepsilon}_1\\ \tilde{\varepsilon}_2\end{bmatrix}\\ &+B_{33}^{-1}\left(\begin{bmatrix}B_{31,0} & B_{32,0}\end{bmatrix}\begin{bmatrix}\tilde{\varepsilon}_1\\ \tilde{\varepsilon}_2\end{bmatrix}+B_{33,0}\tilde{\varepsilon}_3\right)\\ =&-B_{33}^{-1}\begin{bmatrix}B_{31} & B_{32}\end{bmatrix}\begin{bmatrix}B_{11}^{-1} & 0\\ -B_{22}^{-1}B_{21,0}B_{11,0}^{-1} & B_{22}^{-1}\end{bmatrix}\begin{bmatrix}B_{11,0} & 0\\ B_{21,0} & B_{22,0}\end{bmatrix}\begin{bmatrix}\tilde{\varepsilon}_1\\ \tilde{\varepsilon}_2\end{bmatrix}\\ &+B_{33}^{-1}\left(B_{31,0}\tilde{\varepsilon}_1+B_{32,0}\tilde{\varepsilon}_2+B_{33,0}\tilde{\varepsilon}_3\right)\\ =&-B_{33}^{-1}\begin{bmatrix}B_{31}B_{11}^{-1}-B_{32}B_{22}^{-1}B_{21,0}B_{11,0}^{-1} & B_{32}B_{22}^{-1}\end{bmatrix}\begin{bmatrix}B_{11,0} & 0\\ B_{21,0} & B_{22,0}\end{bmatrix}\begin{bmatrix}\tilde{\varepsilon}_1\\ \tilde{\varepsilon}_2\end{bmatrix}\\ &+B_{33}^{-1}\left(B_{31,0}\tilde{\varepsilon}_1+B_{32,0}\tilde{\varepsilon}_2+B_{33,0}\tilde{\varepsilon}_3\right).\end{aligned}$$

Hence, at $B_{22} = B_{22,0}$ the condition $E\left[f_{\mathbf{2}}(B, u_t)\right] = 0$ implies $0 = E[\tilde{e}_3(B)\tilde{e}_2(B)']$ and therefore,

$$0 = B_{33}^{-1}(-B_{32}B_{22}^{-1}B_{22,0} + B_{32,0})$$

which implies $B_{32} = B_{32,0}$. □

4.8.4 Appendix - Redundant and relevant moment conditions in the recursive SVAR

The proof of Proposition 4.3 requires to verify the redundancy conditions from Breusch et al. (1999). However, verifying these conditions is a lengthy task. We derive analytical expressions for the conditions in Appendix 4.9 and summarize them in Lemma 4.15 in Appendix 4.9. The following proof of Proposition 4.3 uses Lemma 4.9 and 4.15 from Appendix 4.9.

Proof of Proposition 4.3.
In the recursive SVAR, the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)]$ only contain second-order moment conditions and therefore, are referred to as $E[f_{\mathbf{2}}(B, u_t)]$ in this proof.

Breusch et al. (1999) show that overidentifying moment conditions $E[f_{\mathbf{D}}(B, u_t)]$ are redundant w.r.t. the identifying moment conditions $E[f_{\mathbf{2}}(B, u_t)]$ if and only if

$$G_{\mathbf{D}} = S_{\mathbf{D2}}S_{\mathbf{2}}^{-1}G_{\mathbf{2}},$$

where

$$G_{\mathbf{D}} := E\left[\frac{\partial f_{\mathbf{D}}(B_0, u_t)}{\partial vec(B)'}\right], \qquad G_{\mathbf{2}} := E\left[\frac{\partial f_{\mathbf{2}}(B_0, u_t)}{\partial vec(B)'}\right],$$

$$S_{\mathbf{2}} := \lim_{T\to\infty} E\left[g_{\mathbf{2}}(B_0)g_{\mathbf{2}}(B_0)'\right], \qquad S_{\mathbf{D2}} := \lim_{T\to\infty} E\left[g_{\mathbf{D}}(B_0)g_{\mathbf{2}}(B_0)'\right].$$

Moreover, Breusch et al. (1999) show that overidentifying moment conditions $E[f_{\mathbf{D}}(B, u_t)]$ are partially redundant w.r.t. $E[f_{\mathbf{2}}(B, u_t)]$ for a subset of coefficients $b \subset vec(B)$ w.r.t. the moment conditions $E[f_{\mathbf{2}}(B,u_t)]$ if and only if

$$G_{\mathbf{D}}^{b} - S_{\mathbf{D2}}S_{\mathbf{2}}^{-1}G_{\mathbf{2}}^{b} = \left(G_{\mathbf{D}}^{\neg b} - S_{\mathbf{D2}}S_{\mathbf{2}}^{-1}G_{\mathbf{2}}^{\neg b}\right)\left(\left(G_{\mathbf{2}}^{\neg b}\right)' S_{\mathbf{2}}^{-1}G_{\mathbf{2}}^{\neg b}\right)\left(\left(G_{\mathbf{2}}^{\neg b}\right)' S_{\mathbf{2}}^{-1}G_{\mathbf{2}}^{b}\right), \tag{4.34}$$

where

$$G_{\mathbf{2}}^{b} := E\left[\frac{\partial f_{\mathbf{2}}(u_t,B_0)}{\partial b'}\right], \qquad G_{\mathbf{D}}^{b} := E\left[\frac{\partial f_{\mathbf{D}}(u_t,B_0)}{\partial b'}\right],$$
$$G_{\mathbf{2}}^{\neg b} := E\left[\frac{\partial f_{\mathbf{2}}(u_t,B_0)}{\partial(\neg b)'}\right], \qquad G_{\mathbf{D}}^{\neg b} := E\left[\frac{\partial f_{\mathbf{D}}(u_t,B_0)}{\partial(\neg b)'}\right],$$

and where $\neg b$ denotes all unrestricted elements of B not contained in b. With Lemma 4.9 it holds that ${G_{\mathbf{2}}^{b_i}}' S_{\mathbf{2}}^{-1}G_{\mathbf{2}}^{b_j} = 0$ for $i,j \in \{1,\ldots,n\}$ with $i \neq j$. Therefore, for any vector $b_i = [b_{ii},...,b_{ni}]$ representing the impact of the ith structural shock $\epsilon_{i,t}$ it holds that ${G_{\mathbf{2}}^{b_i}}' S_{\mathbf{2}}^{-1}G_{\mathbf{2}}^{\neg b_i}$ is zero. Therefore, for any vector $b_i = [b_{ii},...,b_{ni}]$ the right hand side of Equation (4.34) is zero and hence the partial redundancy condition simplifies to

$$G_{\mathbf{D}}^{b_i} - S_{\mathbf{D2}}S_{\mathbf{2}}^{-1}G_{\mathbf{2}}^{b_i} = 0.$$

The statements then follow from Lemma 4.15. □

4.8.5 Appendix - Asymptotic variance of the block-recursive SVAR pGMM estimator

We show how to derive the asymptotic variance of the pGMM estimator, $V_{\mathbf{N}+\mathbf{A}}$, based on Remark 3.5 of Cheng and Liao (2015). We first show Lemma 4.1 and then apply the result in Remark 3.5 of Cheng and Liao (2015). Recall that $E[f_{\mathbf{I}}(B,u_t)]$ and $E[f_{\mathbf{R}}(B,u_t)]$ denote the sets of invalid and redundant moment conditions, respectively. Denote $E[f_{\mathbf{U}}(B,u_t)]$ as moment conditions either in $E[f_{\mathbf{I}}(B,u_t)]$ or $E[f_{\mathbf{R}}(B,u_t)]$ and the number of moment conditions $E[f_{\mathbf{U}}(B,u_t)]$ by $k_{\mathbf{U}}$. Similarly, we denote $k_{\mathbf{A}}$ as the number of moment conditions in $E[f_{\mathbf{A}}(B,u_t)]$. Further, define the number of unrestricted elements in $vec(B)$ as d_B. In the proof of Lemma 4.1, we use the indices $1 \equiv \mathbf{N}+\mathbf{A}$, $2 \equiv (\mathbf{N}+\mathbf{A},\mathbf{U})$, $3 \equiv (\mathbf{U},\mathbf{N}+\mathbf{A})$, and $4 \equiv \mathbf{U}$ to keep notation uncluttered. Let $\iota^* = (\iota', \mathbf{0}'_{k_{\mathbf{U}}})'$ where $\iota = (1,\ldots,1)'$ is a $d_B \times 1$ vector, i.e., ${\iota^*}' A \iota^*$ gives the leading $d_B \times d_B$-upper west block of an arbitrary $(d_B + k_{\mathbf{U}}) \times (d_B + k_{\mathbf{U}})$ matrix A.

Lemma 4.1.

$$\iota^{*\prime} \left(\Gamma' W \Gamma\right)^{-1} \left(\Gamma' W S_{\mathbf{N+D}} W \Gamma\right) \left(\Gamma' W \Gamma\right)^{-1} \iota^* = V_{\mathbf{N+A}},$$

where

$$\Gamma := \begin{bmatrix} G_{\mathbf{N+A}} & \mathbf{0}_{(k_{\mathbf{N}}+k_{\mathbf{A}})\times k_{\mathbf{U}}} \\ G_{\mathbf{U}} & -I_{k_{\mathbf{U}}} \end{bmatrix}, \qquad V_{\mathbf{N+A}} := M_{\mathbf{N+A}} S_{\mathbf{N+A}} M'_{\mathbf{N+A}}$$

with

$$\begin{aligned}
M_{\mathbf{N+A}} &:= \left(G'_{\mathbf{N+A}} W^{pi}_{\mathbf{N+A}} G_{\mathbf{N+A}}\right)^{-1} G'_{\mathbf{N+A}} W^{pi}_{\mathbf{N+A}}, \\
S_{\mathbf{N+A}} &:= \lim_{T\to\infty} E\left[g_{\mathbf{N+A}}(B_0) g_{\mathbf{N+A}}(B_0)'\right], \\
G_{\mathbf{N+A}} &:= \begin{bmatrix} G_{\mathbf{N}} \\ G_{\mathbf{A}} \end{bmatrix}, \\
W^{pi}_{\mathbf{N+A}} &:= \left(W_{\mathbf{N+A}} - W_{\mathbf{N+A},\mathbf{I}\cup\mathbf{R}} W^{-1}_{\mathbf{I}\cup\mathbf{R}} W_{\mathbf{I}\cup\mathbf{R},\mathbf{N+A}}\right), \\
G_{\mathbf{A}} &:= E\left[\frac{\partial f_{\mathbf{A}}(B_0, u_t)}{\partial vec(B)'}\right], \\
W_{\mathbf{N+D}} &:= \begin{bmatrix} W_{\mathbf{N+A}} & W_{\mathbf{N+A},\mathbf{I}\cup\mathbf{R}}, \\ W_{\mathbf{I}\cup\mathbf{R},\mathbf{N+A}} & W_{\mathbf{I}\cup\mathbf{R}} \end{bmatrix}, \\
W_{\mathbf{N+A}} &\in \mathbb{R}^{(k_{\mathbf{N}}+k_{\mathbf{A}})\times(k_{\mathbf{N}}+k_{\mathbf{A}})}, \\
W_{\mathbf{N+A},\mathbf{I}\cup\mathbf{R}} &\in \mathbb{R}^{(k_{\mathbf{N}}+k_{\mathbf{A}})\times(k_{\mathbf{D}}-k_{\mathbf{A}})}, \\
W_{\mathbf{I}\cup\mathbf{R},\mathbf{N+A}} &= W'_{\mathbf{N+A},\mathbf{I}\cup\mathbf{R}}, \\
W_{\mathbf{I}\cup\mathbf{R}} &\in \mathbb{R}^{(k_{\mathbf{D}}-k_{\mathbf{A}})\times(k_{\mathbf{D}}-k_{\mathbf{A}})}.
\end{aligned}$$

Proof. Recall that $G_{\mathbf{N+A}}$ and $G_{\mathbf{U}}$ have dimension $(k_{\mathbf{N}} + k_{\mathbf{A}}) \times d_B$ and $k_{\mathbf{U}} \times d_B$, respectively. We define

$$L := \begin{bmatrix} L_1 & L_2 \\ L_3 & L_4 \end{bmatrix} := \left(\Gamma' W \Gamma\right)^{-1}.$$

Additionally, let

$$N := \begin{bmatrix} N_1 & N_2 \\ N_3 & N_4 \end{bmatrix} := \left(\Gamma' W S_{\mathbf{N+D}} W \Gamma\right),$$

and denote the inverse of W by

$$W^{ipi} := \begin{bmatrix} W_1^{ipi} & W_2^{ipi} \\ W_3^{ipi} & W_4^{ipi} \end{bmatrix} := W^{-1} = \begin{bmatrix} W_1 & W_2 \\ W_3 & W_4 \end{bmatrix}^{-1}.$$

Let $W_1^{pi} := \left(W_1 - W_2 W_4^{-1} W_3\right)$. Then, by the partitioned inverse, $W_1^{ipi} := \left(W_1^{pi}\right)^{-1}$. By similar arguments as leading to (2.18) in the Online Appendix of Cheng and Liao (2015), we get that

$$L_1 = \left(G_1' \left(W_1 - W_2 W_4^{-1} W_3\right) G_1\right)^{-1} = \left(G_1' W_1^{pi} G_1\right)^{-1}$$

and, by using the partitioned inverse formula again, and similar arguments as leading to (2.10), (2.11) and (2.18) in the Online Appendix of Cheng and Liao (2015), that

$$\begin{aligned} L_3 &= -W_4^{-1} \left(-G_1' W_2 - G_4' W_4\right)' \left(G_1' W_1^{pi} G_1\right)^{-1} \\ &= \left(W_4^{-1} W_3 G_1 + G_4\right) L_1 \\ &= X L_1, \end{aligned} \tag{4.35}$$

where we used that $W_4' = W_4$, $W_3 = W_2'$ and $X := \left(W_4^{-1} W_3 G_1 + G_4\right)$. Further, let

$$H := \begin{bmatrix} H_1 & H_2 \\ H_3 & H_4 \end{bmatrix} := W S_{\mathbf{N+D}} W,$$

where

$$\begin{aligned}
H_1 &:= W_1S_1W_1 + W_2S_3W_1 + W_1S_2W_3 + W_2S_4W_3 \\
H_2 &:= W_1S_1W_2 + W_2S_3W_2 + W_1S_2W_4 + W_2S_4W_4 \\
H_3 &:= W_3S_1W_1 + W_4S_3W_1 + W_3S_2W_3 + W_4S_4W_3 \\
H_4 &:= W_3S_1W_2 + W_4S_3W_2 + W_3S_2W_4 + W_4S_4W_4.
\end{aligned}$$

Note that $H_3 = H_2'$ since $W_3 = W_2'$, $W_1 = W_1'$, $W_4 = W_4'$, $S_3 = S_2'$, $S_1 = S_1'$ and $S_4 = S_4'$. Hence, similar to (2.11) in the online Appendix of Cheng and Liao (2015),

$$\begin{aligned}
N_1 &= G_1'H_1G_1 + G_4'H_3G_1 + G_1'H_2G_4 + G_4'H_4G_4 \\
&= G_1'H_1G_1 + G_4'H_2'G_1 + G_1'H_2G_4 + G_4'H_4G_4 \\
N_2 &= -G_1'H_2 - G_4'H_4 \\
N_3 &= N_2' \\
N_4 &= H_4.
\end{aligned}$$

Then,

$$\begin{aligned}
\iota^{*\prime}\left(\Gamma'W\Gamma\right)^{-1}\left(\Gamma'WS_{\mathbf{N+D}}W\Gamma\right)\left(\Gamma'W\Gamma\right)^{-1}\iota^* &= \iota^{*\prime}LNL\iota^* \\
&= L_1N_1L_1 + L_2N_3L_1 + L_1N_2L_3 + L_2N_4L_3 \\
&= L_1N_1L_1 + L_3'N_3L_1 + L_1N_2L_3 + L_3'N_4L_3 \\
&\overset{(4.35)}{=} L_1N_1L_1 + L_1'X'N_2'L_1 + L_1N_2XL_1 + L_1'X'N_4XL_1 \\
&= L_1\left(N_1 + X'N_2' + N_2X + X'N_4X\right)L_1, \qquad (4.36)
\end{aligned}$$

where we used that $L_1' = L_1$, $L_3' = L_2$, and $N_3' = N_2$.

Next, define $Y := N_1 + X'N_2' + N_2X + X'N_4X$. Then, multiplying out gives

$$\begin{aligned}
Y =& G_1'H_1G_1 + G_4'H_3G_1 + G_1'H_2G_4 + G_4'H_4G_4 + \left(G_1'W_2W_4^{-1} + G_4'\right)\left(-H_2'G_1 - H_4'G_4\right) \\
&+ \left(-G_1'H_2 - G_4'H_4\right)\left(W_4^{-1}W_2'G_1 + G_4\right) + \left(G_1'W_2W_4^{-1} + G_4'\right)H_4\left(W_4^{-1}W_2'G_1 + G_4\right) \\
=& G_1'W_2W_4^{-1}H_4W_4^{-1}W_2'G_1 + G_1'H_1G_1 - G_1'W_2W_4^{-1}H_2'G_1 - G_1'H_2W_4^{-1}W_2'G_1 \\
=& G_1'\left(W_2W_4^{-1}H_4W_4^{-1}W_2' + H_1 - W_2W_4^{-1}H_2' - H_2W_4^{-1}W_2'\right)G_1 \\
=& G_1'\left(W_2W_4^{-1}W_3S_1W_2W_4^{-1}W_3 + W_1S_1W_1 - W_2W_4^{-1}W_3S_1W_1 - W_1S_1W_2W_4^{-1}W_3\right)G_1 \\
=& G_1'\left(W_1 - W_2W_4^{-1}W_3\right)S_1\left(W_1 - W_2W_4^{-1}W_3\right)G_1 \\
=& G_1'W_1^{pi}S_1W_1^{pi}G_1
\end{aligned} \tag{4.37}$$

Plugging (4.37) into (4.36), we obtain

$$\begin{aligned}
&\iota^{*\prime}\left(\Gamma'W\Gamma\right)^{-1}\left(\Gamma'WS_{\mathbf{N+D}}W\Gamma\right)\left(\Gamma'W\Gamma\right)^{-1}\iota^* \\
&\quad = L_1\left(G_1'W_1^{pi}S_1W_1^{pi}\right)G_1L_1 \\
&\quad = \left(G_1'W_1^{pi}G_1\right)^{-1}\left(G_1'W_1^{pi}S_1W_1^{pi}G_1\right)\left(G_1'W_1^{pi}G_1\right)^{-1} \\
&\quad = \left(G_{\mathbf{N+A}}'W_{\mathbf{N+A}}^{pi}G_{\mathbf{N+A}}\right)^{-1}\left(G_{\mathbf{N+A}}'W_{\mathbf{N+A}}^{pi}S_{\mathbf{N+A}}W_{\mathbf{N+A}}^{pi}G_{\mathbf{N+A}}\right)\left(G_{\mathbf{N+A}}'W_{\mathbf{N+A}}^{pi}G_{\mathbf{N+A}}\right)^{-1}
\end{aligned}$$

which was to show. □

Note that in the following proposition, we treat the number of valid and relevant moment conditions, $k_{\mathbf{A}}$, and the number of invalid moment conditions, $k_{\mathbf{I}}$, as fixed constants to keep our asymptotic results for the pGMM estimator in line with the asymptotic results for the block-recursive SVAR GMM estimator in Equation (4.19). Cheng and Liao (2015) allow both $k_{\mathbf{A}}$ and $k_{\mathbf{I}}$ to increase with the sample size. However, their results also hold when the number of moment conditions is fixed.

Proposition 4.4. *Assume that the Assumptions in Theorem 3.3 of Cheng and Liao (2015) hold. Further, assume that* $E\left[\frac{\partial f_{\mathbf{A}}(B_0,u_t)}{\partial vec(B)'}\right] = \frac{\partial E[f_{\mathbf{A}}(B_0,u_t)]}{\partial vec(B)'}$ *and Assumption 4.1 and 4.2 hold. Then,*

$$\sqrt{T}\left(vec(\hat{B}_{\mathbf{N+D}}) - vec(B_0)\right) \xrightarrow{d} \mathcal{N}(0, V_{\mathbf{N+A}})$$

Proof. Define $\Sigma_{CL} := \left(\Gamma' W \Gamma\right)^{-1} \left(\Gamma' W S_{\mathbf{N+D}} W \Gamma\right) \left(\Gamma' W \Gamma\right)^{-1}$ and $\gamma = \left(\nu', \mathbf{0}'_{k_{\mathbf{U}}}\right)'$ where $\nu \in \mathbb{R}^{d_B}$ is an arbitrary vector. Then, by Remark 3.5 of Cheng and Liao (2015),

$$\left|\left|\Sigma_{CL}^{1/2}\, \gamma\right|\right|^{-1} \sqrt{T} \nu' \left(vec(\hat{B}_{\mathbf{N+D}}) - vec(B_0)\right) \xrightarrow{d} \mathcal{N}(0,1),$$

where $||a|| := \sqrt{a'a}$ is the ℓ_2-norm of an arbitrary vector a.

Note that Lemma 4.1 immediately implies $\left|\left|\Sigma_{CL}^{1/2} \gamma\right|\right| = \sqrt{\gamma' \Sigma_{CL} \gamma} = \sqrt{\nu' V_{\mathbf{N+A}}(W) \nu}$. Hence,

$$\left|\left|V_{\mathbf{N+A}}(W)^{1/2} \nu\right|\right|^{-1} \sqrt{T} \nu' \left(vec(\hat{B}_{\mathbf{N+D}}) - vec(B_0)\right) \xrightarrow{d} \mathcal{N}(0,1),$$

where $V_{\mathbf{N+A}}(W)$ is the asymptotic variance of $vec(\hat{B}_{\mathbf{N+D}})$ since it holds that

$$\nu^{*\prime} V_{\mathbf{N+A}}(W) \nu^* = \left|\left|V_{\mathbf{N+A}}(W)^{1/2} \nu\right|\right|^{-2} \nu' V_{\mathbf{N+A}}(W) \nu = 1$$

where $\nu^* := \left|\left|V_{\mathbf{N+A}}(W)^{1/2} \nu\right|\right|^{-1} \nu$.

Consequently, using the Cramér-Wold device, we get

$$\sqrt{T} \left(vec(\hat{B}_{\mathbf{N+D}}) - vec(B_0)\right) \xrightarrow{d} \mathcal{N}(0, V_{\mathbf{N+A}}).$$

□

4.8.6 Appendix - Choice of maximum λ in the cross-validation

In the following, we illustrate how to choose the maximum value of λ in the cross-validation. Define the loss function of the pGMM estimator as

$$L^*(B, \beta) := L(B, \beta) + \lambda \sum_{i \in \widetilde{D}} \omega_i \, |\beta_i|, \tag{4.38}$$

where $L(B, \beta) := \begin{bmatrix} g_{\mathbf{N}}(B) \\ g_{\mathbf{D}}(B, \beta) \end{bmatrix}' W \begin{bmatrix} g_{\mathbf{N}}(B) \\ g_{\mathbf{D}}(B, \beta) \end{bmatrix}$.

Further, let $z \in \partial ||\beta||_1$, where $z \in \mathbb{R}^{k_{\mathbf{D}}}$, denote the subgradient for the ℓ_1-norm evaluated at β, i.e.,

$$z_i = sign(\beta_i), \text{ if } \beta_i \neq 0,$$
$$z_i \in [-1, 1], \quad \text{ if } \beta_i = 0, \tag{4.39}$$

for $i = 1, \ldots, k_{\mathbf{D}}$ (Wainwright, 2009). Then, the first order condition of the pGMM estimator with respect to β_i, $i = 1, \ldots, k_{\mathbf{D}}$, evaluated at β and B is

$$\frac{\partial L^*(B, \beta)}{\partial \beta_i} = \frac{\partial L(B, \beta)}{\partial \beta_i} + \lambda \omega_i \, z_i = 0 \tag{4.40}$$

Note that $\omega_i \geq 0$. However, if $\omega_i = 0$, β_i is not penalized and therefore, we only consider $i \in \widetilde{P} := \{j \in \tilde{D} | \, \omega_j > 0\}$ for which, by definition, $\omega_i > 0$ when choosing the maximum value of λ in the cross-validation. By (4.39) and (4.40), $\beta = \mathbf{0} = (0, \ldots, 0)'$ and $B = B_0$ minimize the loss function in (4.38) only if

$$\frac{1}{\omega_i} \frac{\partial L(B_0, \mathbf{0})}{\partial \beta_i} \in \lambda[-1, 1],$$

for $i \in \widetilde{P}$. Thus,

$$\max_{i \in \widetilde{P}} \left| \frac{1}{\omega_i} \frac{\partial L(B_0, \mathbf{0})}{\partial \beta_i} \right| \leq \lambda.$$

This motivates us to use

$$\lambda_{\max} = \max_{i \in \widetilde{P}} \left| \frac{1}{\omega_i} \frac{\partial L(B_0, \mathbf{0})}{\partial \beta_i} \right|.$$

as the largest value in the cross-validation. Note that any $\lambda > \lambda_{\max}$ would not have an effect on β as $\lambda_{\max}$ already shrinks all elements of β to zero. In practice, we replace B_0 and ω_i by consistent estimators to obtain $\lambda_{\max}$. Furthermore, we consider a weight ω_j to be positive and hence, $j \in \widetilde{P}$, if $\omega_j / \sum_{k \in \widetilde{D}} \omega_k > 10^{-4}$.

4.9 Appendix - Details on Proof of Proposition 4.3

4.9.1 Appendix - Notation and preparations part 1

Consider a recursive SVAR $u = B_0\epsilon$ with independent structural shocks with mean zero and unit variance. Let $A := B_0^{-1}$ and a_{ql} $[b_{ql}]$ denote the element at row q and column l of A $[B_0]$. Moreover, let $\omega_{i^2} := \omega_{ii} := E\left[\epsilon_i^2\right]$, $\omega_{i^3} := \omega_{iii} := E\left[\epsilon_i^3\right]$, and $\omega_{i^4} := \omega_{iiii} := E\left[\epsilon_i^4\right]$ for $i = 1, \ldots, n$. Throughout the appendix, the superscript $^{(*)}$ indicates that the equality follows from $e(B_0) = \epsilon$ with ϵ being mutually independent with mean zero and unit variance. Additionally, the superscript $^{(**)}$ indicates that the equality follows from B_0 and hence A_0 being recursive.

We divide the variance-covariance conditions $E[f_{\mathbf{2}}(B, u_t)]$ into a set of variance conditions

$$E[f_{\mathbf{2_M}}(B, u_t)] := E \begin{bmatrix} e(B)_{1,t}^2 - 1 \\ \vdots \\ e(B)_{n,t}^2 - 1 \end{bmatrix}, \tag{4.41}$$

and $n-1$ sets of covariance conditions $E[f_{\mathbf{2_{C_1}}}(B, u_t)], \ldots, E[f_{\mathbf{2_{C_{n-1}}}}(B, u_t)]$ where

$$E[f_{\mathbf{2_{C_i}}}(B, u_t)] := E \begin{bmatrix} e(B)_{i,t} e(B)_{i+1,t} \\ \vdots \\ e(B)_{i,t} e(B)_{n,t} \end{bmatrix}, \text{ for } i = 1, \ldots, n-1. \tag{4.42}$$

We divide the coskewness conditions $E[f_{\mathbf{3}}(B, u_t)]$ into n subsets

$$E[f_{\mathbf{3_{ii}}}(B, u_t)] := E \begin{bmatrix} e(B)_{1,t} e(B)_{i,t}^2 \\ \vdots \\ e(B)_{i-1,t} e(B)_{i,t}^2 \\ e(B)_{i,t}^2 e(B)_{i+1,t} \\ \vdots \\ e(B)_{i,t}^2 e(B)_{n,t} \end{bmatrix}, \text{ for } i = 1, \ldots, n, \tag{4.43}$$

and one additional subset $E[f_{\mathbf{3_{rest}}}(B, u_t)]$ containing all remaining coskewness conditions of

$E[f_{\mathbf{3}}(B,u_t)]$ not contained in a subset $E[f_{\mathbf{3_{ii}}}(B,u_t)]$, which are all coskewness conditions of the type $E[e(B)_{i,t}e(B)_{j,t}e(B)_{k,t}]$ with $i \neq j \neq k$.

We divide the cokurtosis conditions $E[f_{\mathbf{4}}(B,u_t)]$ into n subsets

$$E[f_{\mathbf{4_{ii}}}(B,u_t)] = E\begin{bmatrix} e(B)_{1,t}e(B)^3_{i,t} \\ \vdots \\ e(B)_{i-1,t}e(B)^3_{i,t} \\ e(B)^3_{i,t}e(B)_{i+1,t} \\ \vdots \\ e(B)^3_{i,t}e(B)_{n,t} \end{bmatrix}, \text{ for } i = 1,\dots,n, \tag{4.44}$$

$n-1$ subsets

$$E[f_{\mathbf{4_{iii}}}(B,u_t)] = E\begin{bmatrix} e(B)^2_{i,t}e(B)^2_{i+1,t} \\ \vdots \\ e(B)^2_{i,t}e(B)^2_{n,t} \end{bmatrix}, \text{ for } i = 1,\dots,n-1, \tag{4.45}$$

and one additional subset $E[f_{\mathbf{4_{rest}}}(B,u_t)]$ containing all remaining cokurtosis conditions of $E[f_{\mathbf{4}}(B,u_t)]$ not contained in a subset $E[f_{\mathbf{4_{ii}}}(B,u_t)]$ or $E[f_{\mathbf{4_{iii}}}(B,u_t)]$, which are all cokurtosis conditions of the type $E[e(B_0)_i e(B_0)_j e(B_0)_k e(B_0)_l]$ and $E[e(B_0)^2_i e(B_0)_j e(B_0)_k]$ with $i \neq j \neq k \neq l$.

Throughout the Appendix, we will use the following lemmata.

Lemma 4.2. *The derivative of the i-th element of the unmixed innovations at B_0 with respect to an element b_{pq} is given by*

$$\frac{\partial e_{it}(B_0)}{\partial b_{pq}} = \begin{cases} -a_{ip}\epsilon_{qt}, & \text{if } i \geq p \\ 0, & \text{else} \end{cases}. \tag{4.46}$$

Proof.

$$\frac{\partial e_t(B_0)}{\partial b_{pq}} = \frac{\partial B_0^{-1}}{\partial b_{pq}} u_t \tag{4.47}$$

$$= \left(-B_0^{-1} \frac{\partial B_0}{\partial b_{pq}} B_0^{-1} \right) u_t \tag{4.48}$$

$$= -B_0^{-1} \frac{\partial B_0}{\partial b_{pq}} B_0^{-1} B_0 \epsilon_t \tag{4.49}$$

$$= -B_0^{-1} \frac{\partial B_0}{\partial b_{pq}} \epsilon_t \tag{4.50}$$

$$= -A_0 \frac{\partial B_0}{\partial b_{pq}} \begin{bmatrix} \epsilon_{1t} \\ \vdots \\ \epsilon_{nt} \end{bmatrix} \tag{4.51}$$

$$= - \begin{bmatrix} a_{1p} \\ \vdots \\ a_{np} \end{bmatrix} \epsilon_{qt} \overset{\text{recursive SVAR}}{=} - \begin{bmatrix} 0 \\ \vdots \\ 0 \\ a_{pp} \\ \vdots \\ a_{np} \end{bmatrix} \epsilon_{qt} \tag{4.52}$$

$\square$

Lemma 4.3. *For $i = 1, \ldots, n-1$ and $j = 1, \ldots, n$ let*

$$G_{\mathbf{2}}^{b_{j:,j}} := \left[G_{\mathbf{2_M}}^{b_{j:,j}}, G_{\mathbf{2_{C_1}}}^{b_{j:,j}}, \ldots, G_{\mathbf{2_{C_{n-1}}}}^{b_{j:,j}} \right]' \tag{4.53}$$

with

$$G_{\mathbf{2_M}}^{b_{j:,j}} := E \left[\frac{\partial f_{\mathbf{2_M}}(u_t, B_0)}{\partial \left(b_{jj}, \ldots, b_{nj} \right)'} \right], \tag{4.54}$$

$$G_{\mathbf{2_{C_i}}}^{b_{j:,j}} := E \left[\frac{\partial f_{\mathbf{2_{C_i}}}(u_t, B_0)}{\partial \left(b_{jj}, \ldots, b_{nj} \right)'} \right]. \tag{4.55}$$

Then

$$G_{\mathbf{2_M}}^{b_{j:,j}} = -2 \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} & \\ a_{jj} & \dots & a_{jn} \\ & 0_{(n-j)\times(n-j+1)} & \end{bmatrix}}_{n\times(n-j+1)} = -2 \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} & \\ a_{jj} & 0_{1\times(n-j)} & \\ & 0_{(n-j)\times(n-j+1)} & \end{bmatrix}}_{n\times(n-j+1)} \tag{4.56}$$

and

$$G_{\mathbf{2_{C_i}}}^{b_{j:,j}} = 0_{(n-i)\times(n-j+1)}, \qquad \textit{for } i \neq j, \tag{4.57}$$

$$G_{\mathbf{2_{C_i}}}^{b_{i:,i}} = -\underbrace{\begin{bmatrix} a_{i+1,i} & a_{i+1,i+1} & & 0 \\ \vdots & & \ddots & \\ a_{ni} & a_{n,i+1} & \dots & a_{nn} \end{bmatrix}}_{(n-i)\times(n-i+1)}. \tag{4.58}$$

Proof. Equation (4.56): The (q,r)-th entry of $G_{\mathbf{2_{C_i}}}^{b_{j:,j}}$ is equal to

$$G_{\mathbf{2_{C_i}}}^{b_{j:,j}}(q,r) = E\left[\frac{\partial(e(B_0)_q^2 - 1)}{\partial b_{j+r-1,j}}\right] \overset{(*)}{=} \begin{cases} -2a_{q,j+r-1}, & \text{if } q = j \\ 0, & \text{else} \end{cases}. \tag{4.59}$$

Equation (4.57) and (4.58): The (q,r)-th entry of $G_{\mathbf{2_{C_i}}}^{b_{j:,j}}$ is equal to

$$G_{\mathbf{2_{C_i}}}^{b_{j:,j}}(q,r) = E\left[\frac{\partial(e(B_0)_i e(B_0)_{i+q})}{\partial b_{j+r-1,j}}\right] \overset{(*)}{=} \begin{cases} -a_{i+q,j+r-1}, & \text{if } j = i \\ -a_{i,j+r-1} \overset{(**)}{=} 0, & \text{if } j = (i+q) \\ 0, & \text{else} \end{cases}. \tag{4.60}$$

□

Lemma 4.4. *For $i, j = 1, \ldots, n$ let*

$$G_{\mathbf{3_{ii}}}^{b_{j:,j}} := E\left[\frac{\partial f_{\mathbf{3_{ii}}}(u_t, B_0)}{\partial\left(b_{jj}, \ldots, b_{nj}\right)'}\right], \tag{4.61}$$

$$G_{\mathbf{3_{rest}}}^{vec(B)} := E\left[\frac{\partial f_{\mathbf{3_{rest}}}(u_t, B_0)}{\partial vec(B)'}\right]. \tag{4.62}$$

Then

$$G_{\mathbf{3_{ii}}}^{b_{j:,j}} = 0_{(n-1)\times(n-j+1)}, \qquad \text{for } i \neq j, \tag{4.63}$$

$$G_{\mathbf{3_{ii}}}^{b_{i:,i}} = -\omega_{iii} \underbrace{\begin{bmatrix} & 0_{(i-1)\times(n-i+1)} & & \\ a_{i+1,i} & a_{i+1,i+1} & & 0 \\ \vdots & & \ddots & \\ a_{ni} & a_{n,i+1} & \ldots & a_{nn} \end{bmatrix}}_{(n-1)\times(n-i+1)}, \tag{4.64}$$

and

$$G_{\mathbf{3_{rest}}}^{vec(B)} = 0. \tag{4.65}$$

Proof. Equation (4.63): The (q,r)-th entry of $G_{\mathbf{3_{ii}}}^{b_{j:,j}}$ with $q < i$ is equal to

$$G_{\mathbf{3_{ii}}}^{b_{j:,j}}(q,r) = E\left[\frac{\partial(e(B_0)_i^2 e(B_0)_q)}{\partial b_{j+r-1,j}}\right] = \begin{cases} -a_{q,j+r-1}\omega_{iii}, & j = i, q \geq r+i-1 \\ -a_{q,j+r-1}\omega_{iii} \overset{(**)}{=} 0, & j = i, q < r+i-1 \\ 0, & \text{else} \end{cases} \tag{4.66}$$

and the (q,l)-th entry of $G_{\mathbf{3_{ii}}}^{b_{j:,j}}$ with $q \geq i$ is equal to

$$G_{\mathbf{3_{ii}}}^{b_{j:,j}}(q,r) = E\left[\frac{\partial(e(B_0)_i^2 e(B_0)_{q+1})}{\partial b_{j+r-1,j}}\right] = \begin{cases} -a_{q+1,j+r-1}\omega_{iii}, & j = i, q \geq r+i-1 \\ -a_{q+1,j+r-1}\omega_{iii} \overset{(**)}{=} 0, & j = i, q < r+i-1 \cdot \\ 0, & \text{else} \end{cases} \tag{4.67}$$

Every element in $G_{\mathbf{3_{rest}}}^{vec(B)}$ can be written as $E[\frac{\partial(e(B)_{a,t}e(B)_{b,t}e(B)_{c,t})}{\partial b_{q,l}}]$ for some $a, b, c \in \{1, \dots, n\}$ with $a \neq b \neq c$. Equation (4.65) follows with Lemma 4.2, $e(B_0) = \epsilon$, and independence and mean zero of ϵ_t.

□

Lemma 4.5. *Let*

$$G_{\mathbf{4_{ii}}}^{b_{j:,j}} := E\left[\frac{\partial f_{\mathbf{4_{ii}}}(u_t, B_0)}{\partial (b_{jj}, \dots, b_{nj})'}\right], \qquad \textit{for } i = 1, \dots, n-1,\ j = 1, \dots, n, \tag{4.68}$$

$$G_{\mathbf{4_{iii}}}^{b_{j:,j}} := E\left[\frac{\partial f_{\mathbf{4_{iii}}}(u_t, B_0)}{\partial (b_{jj}, \dots, b_{nj})'}\right], \qquad \textit{for } i,\ j = 1, \dots, n, \tag{4.69}$$

$$G_{\mathbf{4_{rest}}}^{vec(B)} := E\left[\frac{\partial f_{\mathbf{4_{rest}}}(u_t, B_0)}{\partial vec(B)'}\right]. \tag{4.70}$$

Then

$$G_{\mathbf{4_{ii}}}^{b_{i:,i}} = -2 \underbrace{\begin{bmatrix} a_{ii} & a_{in} \\ & \\ a_{ii} & a_{in} \end{bmatrix}}_{(n-i)\times(n-i+1)} = -2 \underbrace{\begin{bmatrix} a_{ii} & \\ & 0_{(n-i)\times(n-j)} \\ a_{ii} & \end{bmatrix}}_{(n-i)\times(n-i+1)}, \tag{4.71}$$

$$G_{\mathbf{4_{ii}}}^{b_{j:,j}} = 0_{(n-i)\times(n-j+1)}, \qquad \textit{for } i > j, \tag{4.72}$$

$$G_{\mathbf{4_{ii}}}^{b_{j:,j}} = -2 \underbrace{\begin{bmatrix} & 0_{(j-i-1)\times(n-j+1)} & \\ a_{jj} & \dots & a_{jn} \\ & 0_{(n-j)\times(n-j+1)} & \end{bmatrix}}_{(n-i)\times(n-j+1)} = -2 \underbrace{\begin{bmatrix} & 0_{(j-i-1)\times(n-j+1)} \\ a_{jj} & 0_{(1)\times(n-j)} \\ & 0_{(n-j)\times(n-j+1)} \end{bmatrix}}_{(n-i)\times(n-j+1)}, \qquad \textit{for } i < j, \tag{4.73}$$

and

$$G_{\mathbf{4_{iii}}}^{b_{i:,i}} = -\omega_{iiii} \underbrace{\begin{bmatrix} a_{1i} & a_{1n} \\ & \\ a_{i-1,i} & a_{i-1,n} \\ a_{i+1,i} & a_{i+1,n} \\ & \\ a_{ni} & a_{nn} \end{bmatrix}}_{(n-1)\times(n-i+1)} = -\omega_{iiii} \underbrace{\begin{bmatrix} & 0_{(i-1)\times(n-i+1)} & & \\ a_{i+1,i} & a_{i+1,i+1} & & 0 \\ \vdots & & \ddots & \\ a_{ni} & \dots & & a_{nn} \end{bmatrix}}_{(n-1)\times(n-i+1)}, \tag{4.74}$$

$$G_{\mathbf{4_{iii}}}^{b_{j:,j}} = -3 \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} & \\ a_{ij} & \dots & a_{in} \\ & 0_{(n-j-1)\times(n-j+1)} & \end{bmatrix}}_{(n-1)\times(n-j+1)} \tag{4.75}$$

$$= -3 \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} & & \\ a_{ij} & \dots & a_{ii} & 0_{1\times(n-i)} \\ & 0_{(n-j-1)\times(n-j+1)} & & \end{bmatrix}}_{(n-1)\times(n-j+1)}, \qquad \textit{for } i > j, \tag{4.76}$$

$$G_{\mathbf{4_{iii}}}^{b_{j:,j}} = -3 \underbrace{\begin{bmatrix} & 0_{(j-1-1)\times(n-j+1)} & \\ a_{ij} & \dots & a_{in} \\ & 0_{(n-j)\times(n-j+1)} & \end{bmatrix}}_{(n-1)\times(n-j+1)} = 0_{(n-1)\times(n-j+1)}, \qquad \textit{for } i < j, \tag{4.77}$$

and

$$G_{\mathbf{4_{rest}}}^{vec(B)} = 0. \tag{4.78}$$

Proof. Equation (4.71), (4.72), and (4.73): The (q,r)-th entry of $G_{\mathbf{4_{ii}}}^{b_{j:,j}}$ is equal to

$$G_{\mathbf{4_{ii}}}^{b_{j:,j}}(q,r) = E\left[\frac{\partial(e(B_0)_i^2 e(B_0)_{i+q}^2 - 1)}{\partial b_{j+r-1,j}}\right] \overset{(*)}{=} \begin{cases} -2a_{i,j+r-1}, & \text{if } j = i, r = 1 \\ -2a_{i,j+r-1} \overset{(**)}{=} 0, & \text{if } j = i, r \neq 1 \\ -2a_{i+q,j+r-1}, & \text{if } j = i+q, r = 1 \\ -2a_{i+q,j+r-1} \overset{(**)}{=} 0, & \text{if } j = i+q, r \neq 1 \\ 0, & \text{else} \end{cases}. \tag{4.79}$$

Equation (4.74), (4.75), and (4.77): The (q,r)-th entry of $G_{\mathbf{4_{iii}}}^{b_{j:,j}}$ with $q < i$ is equal to

$$G_{\mathbf{4_{iii}}}^{b_{j:,j}}(q,r) = E\left[\frac{\partial(e(B_0)_i^3 e(B_0)_q)}{\partial b_{j+r-1,j}}\right] \overset{(*)}{=} \begin{cases} -a_{q,j+r-1}\omega_{iiii} \overset{(**)}{=} 0, & \text{if } j = i \\ -3a_{i,j+r-1}, & \text{if } j = q, r \leq i-j+1 \\ -3a_{i,j+r-1} \overset{(**)}{=} 0, & \text{if } j = q, r > i-j+1 \\ 0, & \text{else} \end{cases}. \tag{4.80}$$

The (q,r)-th entry of P_i^j with $q \geq i$ is equal to

$$G_{\mathbf{4_{iii}}}^{b_{j:,j}}(q,r) = E\left[\frac{\partial(e(B_0)_i^3 e(B_0)_{q+1})}{\partial b_{j+r-1,j}}\right] \overset{(*)}{=} \begin{cases} -a_{q+1,j+r-1}\omega_{iiii}, & \text{if } j = i, q-r \geq j-2 \\ -a_{q+1,j+r-1}\omega_{iiii} \overset{(**)}{=} 0, & \text{if } j = i, q-r < j-2 \\ -3a_{i,j+r-1}, & \text{if } j = q+1 \\ 0, & \text{else} \end{cases}. \tag{4.81}$$

Equation (4.78) follows analogously to Equation (4.65). ☐

Lemma 4.6. *The matrix $S_{\mathbf{2}} = E\left[f_{\mathbf{2}}(u_t, B_0) f_{\mathbf{2}}(u_t, B_0)'\right]$ can be written as*

$$S_{\mathbf{2}} = \begin{bmatrix} S_{\mathbf{2}_M} & S_{\mathbf{2}_M \mathbf{2}_C} \\ S_{\mathbf{2}_C \mathbf{2}_M} & S_{\mathbf{2}_C} \end{bmatrix}, \text{ with } \qquad \begin{aligned} S_{\mathbf{2}_M} &:= E\left[f_{\mathbf{2}_M}(u_t, B_0) f_{\mathbf{2}_M}(u_t, B_0)'\right], \\ S_{\mathbf{2}_M \mathbf{2}_C} &:= E\left[f_{\mathbf{2}_M}(u_t, B_0) f_{\mathbf{2}_C}(u_t, B_0)'\right], \\ S_{\mathbf{2}_C \mathbf{2}_M} &:= E\left[f_{\mathbf{2}_C}(u_t, B_0) f_{\mathbf{2}_M}(u_t, B_0)'\right], \\ S_{\mathbf{2}_C} &:= E\left[f_{\mathbf{2}_C}(u_t, B_0) f_{\mathbf{2}_C}(u_t, B_0)'\right], \end{aligned} \tag{4.82}$$

and

$$S_{\mathbf{2}_M} = \underbrace{\begin{bmatrix} \omega_{1111} - 1 & & 0 \\ & \ddots & \\ 0 & & \omega_{nnnn} - 1 \end{bmatrix}}_{n \times n}, \tag{4.83}$$

$$S_{\mathbf{2}_M \mathbf{2}_C} = 0_{n \times n(n-1)}, \tag{4.84}$$

$$S_{\mathbf{2}_C \mathbf{2}_M} = 0_{n(n-1) \times n}, \tag{4.85}$$

$$S_{\mathbf{2}_C} = I_{n(n-1) \times n(n-1)}. \tag{4.86}$$

Therefore, $S_{\mathbf{2}}$ is equal to

$$S_{\mathbf{2}} = \begin{bmatrix} \omega_{1111} - 1 & & 0 & \\ & \ddots & & 0_{n \times n(n-1)} \\ 0 & & \omega_{nnnn} - 1 & \\ & 0_{n(n-1) \times n} & & I_{n(n-1) \times n(n-1)} \end{bmatrix}. \tag{4.87}$$

Proof. Equation (4.83), (4.84), (4.86), and (4.87): The (q,r)-th entry of $S_{\mathbf{2}_M}$ is equal to

$$S_{\mathbf{2}_M}(q, r) = E\left[(e(B_0)_q^2 - 1)(e(B_0)_r^2 - 1)\right] \overset{(*)}{=} \begin{cases} \omega_{qqqq} - 1, & \text{if } q = r \\ 0, & \text{else} \end{cases}. \tag{4.88}$$

Any entry of $S_{\mathbf{2}_M \mathbf{2}_C}$ can be written as

$$E\left[(e(B_0)_a^2 - 1)(e(B_0)_b e(B_0)_c)\right] \overset{(*)}{=} 0 \tag{4.89}$$

for some $a, b, c \in \{1, \ldots, n\}$ with $b \neq c$. Any entry of $S_{\mathbf{2}_C \mathbf{2}_C}$ can be written as

$$E\left[(e(B_0)_a e(B_0)_b)(e(B_0)_c e(B_0)_d)\right] \overset{(*)}{=} \begin{cases} 1, & \text{if } a = c, b = d \\ 0, & \text{else} \end{cases}, \tag{4.90}$$

for some $a, b, c, d \in \{1, \ldots, n\}$ with $a \neq b$ and $c \neq d$. The case $a = c$ and $b = d$ occurs at the diagonal of $S_{\mathbf{2}_C \mathbf{2}_C}$.

$\square$

Lemma 4.7. *For $i = 1, \ldots, n$ and $j = 1, \ldots, n-1$ let*

$$S_{\mathbf{3_{ii}2}} = \left[S_{\mathbf{3_{ii}2_M}} \quad S_{\mathbf{3_{ii}2_{C_1}}} \quad \ldots \quad S_{\mathbf{3_{ii}2_{C_{n-1}}}}\right], \tag{4.91}$$

$$S_{\mathbf{3_{rest}2}} = \left[S_{\mathbf{3_{rest}2_M}} \quad S_{\mathbf{3_{rest}2_{C_1}}} \quad \ldots \quad S_{\mathbf{3_{rest}2_{C_{n-1}}}}\right] \tag{4.92}$$

with

$$S_{\mathbf{3_{ii}2_M}} := E\left[f_{\mathbf{3_{ii}}}(u_t, B_0) f_{\mathbf{2_M}}(u_t, B_0)'\right], \tag{4.93}$$

$$S_{\mathbf{3_{ii}2_{C_j}}} := E\left[f_{\mathbf{3_{ii}}}(u_t, B_0) f_{\mathbf{2_{C_j}}}(u_t, B_0)'\right], \tag{4.94}$$

$$S_{\mathbf{3_{rest}2_M}} := E\left[f_{\mathbf{3_{rest}}}(u_t, B_0) f_{\mathbf{2_M}}(u_t, B_0)'\right], \tag{4.95}$$

$$S_{\mathbf{3_{rest}2_{C_j}}} := E\left[f_{\mathbf{3_{rest}}}(u_t, B_0) f_{\mathbf{2_{C_j}}}(u_t, B_0)'\right], \tag{4.96}$$

and

$$S_{\mathbf{3_{ii}2_M}} = \underbrace{\begin{bmatrix} \omega_{111} & & 0 & 0 & & & \\ & \ddots & & 0 & & 0_{(n-i)\times(n-i)} & \\ 0 & & \omega_{(i-1)^3} & 0 & & & \\ & & & 0 & \omega_{(i+1)^3} & & 0 \\ & 0_{(i-1)\times(i-1)} & & 0 & & \ddots & \\ & & & 0 & 0 & & \omega_{nnn} \end{bmatrix}}_{(n-1)\times(n)} \tag{4.97}$$

and

$$S_{\mathbf{3_{ii}2_{C_i}}} = \omega_{iii} \underbrace{\begin{bmatrix} 0_{(i-1)\times(n-i)} \\ I_{(n-i)\times(n-i)} \end{bmatrix}}_{(n-1)\times(n-j)}, \qquad \textit{for } i = 1, \dots, n-1, \tag{4.98}$$

$$S_{\mathbf{3_{ii}2_{C_j}}} = 0_{(n-1)\times(n-j)}, \qquad \textit{for } i < j, \tag{4.99}$$

$$S_{\mathbf{3_{ii}2_{C_j}}} = \omega_{iii} \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j)} & \\ 0_{1\times(i-j-1)} & 1 & 0_{1\times(n-i)} \\ & 0_{(n-j-1)\times(n-j)} & \end{bmatrix}}_{(n-1)\times(n-j)}, \qquad \textit{for } i > j, \tag{4.100}$$

and

$$S_{\mathbf{3_{rest}2}} = 0. \tag{4.101}$$

Proof. Equation (4.97): The (q,r)-th entry of $S_{\mathbf{3_{ii}2_M}}$ with $q < i$ is equal to

$$S_{\mathbf{3_{ii}2_M}}(q,r) = E\left[(e(B_0)_i^2 e(B_0)_q)(e(B_0)_r^2 - 1)\right] \overset{(*)}{=} \begin{cases} \omega_{qqq}, & \text{if } r = q \\ 0, & \text{else} \end{cases} \tag{4.102}$$

and the (q,r)-th entry of $S_{\mathbf{3_{ii}2_M}}$ with $q \geq i$ is equal to

$$S_{\mathbf{3_{ii}2_M}}(q,r) = E\left[(e(B_0)_i^2 e(B_0)_{q+1})(e(B_0)_r^2 - 1)\right] \overset{(*)}{=} \begin{cases} \omega_{(q+1)^3}, & \text{if } r = q+1 \\ 0, & \text{else} \end{cases}. \tag{4.103}$$

Equation (4.98), (4.99) and (4.100): The (q,r)-th entry of $S_{\mathbf{3_{ii}2_{C_j}}}$ with $q < i$ is equal to

$$S_{\mathbf{3_{ii}2_{C_j}}}(q,l) = E\left[(e(B_0)_i^2 e(B_0)_q)(e(B_0)_j e(B_0)_{j+r})\right] \overset{(*)}{=} \begin{cases} \omega_{iii}, & \text{if } i = j + r, j = q \\ 0, & \text{else} \end{cases} \tag{4.104}$$

and the (q,r)-th entry of $S_{\mathbf{3_{ii}2_{C_j}}}$ with $q \geq i$ is equal to

$$S_{\mathbf{3_{ii}2_{C_j}}}(q,l) = E\left[(e(B_0)_i^2 e(B_0)_{q+1})(e(B_0)_j e(B_0)_{j+r})\right] \overset{(*)}{=} \begin{cases} \omega_{iii}, & \text{if } i=j, q+1=j+r \\ 0, & \text{else} \end{cases}. \tag{4.105}$$

Equation (4.101) holds, since every moment condition in $E[f_{\mathbf{3_{rest}}}(u_t,B_0)f_{\mathbf{2}_M}(u_t,B_0)']$ can be written as

$$E\left[(e(B_0)_a e(B_0)_b e(B_0)_c)(e(B_0)_d^2 - 1)\right] \tag{4.106}$$

for some $a,b,c,d = \{1,\dots,n\}$ and $a \neq b \neq c$, which implies $E\left[(e(B_0)_a e(B_0)_b e(B_0)_c)(e(B_0)_d^2 - 1)\right] = 0$ by independence and mean zero of ϵ.

$\square$

Lemma 4.8. *Let*

$$S_{\mathbf{4_{ii}2}} = \begin{bmatrix} S_{\mathbf{4_{ii}2_M}} & S_{\mathbf{4_{ii}2_{C_1}}} & \dots & S_{\mathbf{4_{ii}2_{C_{n-1}}}} \end{bmatrix}, \qquad \textit{for } i = 1,\dots,n-1, \tag{4.107}$$

$$S_{\mathbf{4_{iii}2}} = \begin{bmatrix} S_{\mathbf{4_{iii}2_M}} & S_{\mathbf{4_{iii}2_{C_1}}} & \dots & S_{\mathbf{4_{iii}2_{C_{n-1}}}} \end{bmatrix}, \qquad \textit{for } i = 1,\dots,n, \tag{4.108}$$

$$S_{\mathbf{4_{rest}2}} = \begin{bmatrix} S_{\mathbf{4_{rest}2_M}} & S_{\mathbf{4_{rest}2_{C_1}}} & \dots & S_{\mathbf{4_{rest}2_{C_{n-1}}}} \end{bmatrix} \tag{4.109}$$

with

$$S_{\mathbf{4_{ii}2_M}} := E\left[f_{\mathbf{4_{ii}}}(u_t,B_0)f_{\mathbf{2}_M}(u_t,B_0)'\right], \qquad \textit{for } i = 1,\dots,n-1, \tag{4.110}$$

$$S_{\mathbf{4_{ii}2_{C_j}}} := E\left[f_{\mathbf{4_{ii}}}(u_t,B_0)f_{\mathbf{2}_{C_j}}(u_t,B_0)'\right], \qquad \textit{for } i,j = 1,\dots,n-1, \tag{4.111}$$

$$S_{\mathbf{4_{iii}2_M}} := E\left[f_{\mathbf{4_{iii}}}(u_t,B_0)f_{\mathbf{2}_M}(u_t,B_0)'\right], \qquad \textit{for } i = 1,\dots,n, \tag{4.112}$$

$$S_{\mathbf{4_{iii}2_{C_j}}} := E\left[f_{\mathbf{4_{iii}}}(u_t,B_0)f_{\mathbf{2}_{C_j}}(u_t,B_0)'\right], \qquad \textit{for } j = 1,\dots,n-1,\ i = 1,\dots,n, \tag{4.113}$$

$$S_{\mathbf{4_{rest}2_M}} := E\left[f_{\mathbf{4_{rest}}}(u_t,B_0)f_{\mathbf{2}_M}(u_t,B_0)'\right], \tag{4.114}$$

$$S_{\mathbf{4_{rest}2_{C_j}}} := E\left[f_{\mathbf{4_{rest}}}(u_t,B_0)f_{\mathbf{2}_{C_j}}(u_t,B_0)'\right], \qquad \textit{for } j = 1,\dots,n-1, \tag{4.115}$$

and

$$S_{\mathbf{4_{ii}2_M}} = \underbrace{\begin{bmatrix} & \omega_{iiii}-1 & \omega_{(i+1)^4}-1 & & 0 \\ 0_{(n-i)\times(i-1)} & \vdots & & \ddots & \\ & \omega_{iiii}-1 & 0 & & \omega_{nnnn}-1 \end{bmatrix}}_{(n-i)\times(n)}, \tag{4.116}$$

$$S_{\mathbf{4_{iii}2_M}} = \omega_{iii} \underbrace{\begin{bmatrix} \omega_{111} & & 0 & 0 & & & \\ & \ddots & & 0 & & 0_{(n-i)\times(n-i)} & \\ 0 & & \omega_{(i-1)^3} & 0 & & & \\ & & & 0 & \omega_{(i+1)^3} & & 0 \\ & 0_{(i-1)\times(i-1)} & & 0 & & \ddots & \\ & & & 0 & 0 & & \omega_{nnn} \end{bmatrix}}_{(n-1)\times(n)}, \tag{4.117}$$

and

$$S_{\mathbf{4_{ii}2_{C_i}}} = \omega_{iii} \underbrace{\begin{bmatrix} \omega_{(i+1)^3} & & 0 \\ & \ddots & \\ 0 & & \omega_{nnn} \end{bmatrix}}_{(n-i)\times(n-i)}, \tag{4.118}$$

$$S_{\mathbf{4_{ii}2_{C_j}}} = 0_{(n-i)\times(n-j)}, \qquad \textit{for } i \neq j, \tag{4.119}$$

$$S_{\mathbf{4_{iii}2_{C_i}}} = \omega_{iiii} \underbrace{\begin{bmatrix} 0_{(i-1)\times(n-j)} \\ I_{(n-i)\times(n-j)} \end{bmatrix}}_{(n-1)\times(n-i)}, \tag{4.120}$$

$$S_{\mathbf{4_{iii}2_{C_j}}} = 0_{(n-1)\times(n-j)}, \qquad \textit{for } i < j, \tag{4.121}$$

$$S_{\mathbf{4_{iii}2_{C_j}}} = \omega_{iiii} \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j)} & \\ 0_{1\times(i-j-1)} & 1 & 0_{1\times(n-i)} \\ & 0_{(n-j-1)\times(n-j)} & \end{bmatrix}}_{(n-1)\times(n-j)}, \qquad \textit{for } i > j, \tag{4.122}$$

and

$$S_{\mathbf{4_{rest}2}} = 0. \tag{4.123}$$

Proof. Equation (4.116): The (q,r)-th entry of $S_{\mathbf{4_{ii}2_M}}$ is equal to

$$S_{\mathbf{4_{ii}2_M}}(q,r) = E\left[(e(B_0)_i^2 e(B_0)_{i+q}^2 - 1)(e(B_0)_r^2 - 1)\right] \tag{4.124}$$

$$\overset{(*)}{=} \begin{cases} \omega_{iiii} - 1, & \text{if } r = i \\ \omega_{(i+q)^4} - 1, & \text{if } r = i+q \cdot \\ 0, & \text{else} \end{cases} \tag{4.125}$$

Equation (4.117): The (q,r)-th entry of $S_{\mathbf{4_{iii}2_M}}$ with $q < i$ is equal to

$$S_{\mathbf{4_{iii}2_M}}(q,r) = E\left[(e(B_0)_i^3 e(B_0)_q)(e(B_0)_r^2 - 1)\right] \overset{(*)}{=} \begin{cases} \omega_{iii}\omega_{qqq}, & \text{if } r = q \\ 0, & \text{else} \end{cases} \tag{4.126}$$

and the (q,r)-th entry of $SS_{\mathbf{4_{iii}2_M}}$ with $q \geq i$ is equal to

$$S_{\mathbf{4_{iii}2_M}}(q,r) = E\left[(e(B_0)_i^3 e(B_0)_{q+1})(e(B_0)_r^2 - 1)\right] \overset{(*)}{=} \begin{cases} \omega_{iii}\omega_{(q+1)^3}, & \text{if } r = q+r \\ 0, & \text{else} \end{cases}. \tag{4.127}$$

Equation (4.118) and (4.119): The (q,r)-th entry of $S_{\mathbf{4_{ii}2_{C_j}}}$ is equal to

$$S_{\mathbf{4_{ii}2_{C_j}}}(q,r) = E\left[(e(B_0)_i^2 e(B_0)_{i+q}^2 - 1)(e(B_0)_j e(B_0)_{j+r})\right] \overset{(*)}{=} \begin{cases} 0, & \text{if } i \neq j \\ \omega_{iii}\omega_{(i+q)^3}, & \text{if } i = j, q = r \cdot \\ 0, & \text{if } i = j, q \neq r \end{cases} \tag{4.128}$$

Equation (4.120), (4.121), and (4.122): The (q,r)-th entry of $S_{\mathbf{4_{iii}2_{C_j}}}$ with $q < i$ is equal to

$$S_{\mathbf{4_{iii}2_{C_j}}}(q,r) = E\left[(e(B_0)_i^3 e(B_0)_q)(e(B_0)_j e(B_0)_{j+r})\right] \overset{(*)}{=} \begin{cases} 0, & \text{if } i < j \\ \omega_{iiii}, & \text{if } i = j, q = j + r \ ^{(***)} \\ 0, & \text{if } i = j, q \neq j + r \\ \omega_{iiii}, & \text{if } i > j, q = j \\ 0, & \text{if } i > j, q \neq j \end{cases} \tag{4.129}$$

and note that for $q < i$ the case $^{(***)}$ never occurs since $i > q = j + r = i + r$ implies $r < 0$. Moreover, the (q,r)-th entry of $S_{\mathbf{4_{iii}2_{C_j}}}$ with $q \geq i$ is equal to

$$S_{\mathbf{4_{iii}2_{C_j}}}(q,r) = E\left[(e(B_0)_i^3 e(B_0)_{q+1})(e(B_0)_j e(B_0)_{j+r})\right] \overset{(*)}{=} \begin{cases} 0, & \text{if } i < j \\ \omega_{iiii}, & \text{if } i = j, q + 1 = j + r \\ 0, & \text{if } i = j, q + 1 \neq j + r \\ \omega_{iiii}, & \text{if } i > j, q + 1 = j \ ^{(***)} \\ 0, & \text{if } i > j, q + 1 \neq j \end{cases} \tag{4.130}$$

and note that for $q \geq i$ the case $^{(***)}$ never occurs since $q + 1 = j < i \leq q$ implies $1 < 0$.

Equation (4.123)) holds, since every moment condition in $E[f_{\mathbf{4_{rest}}}(u_t, B) f_{\mathbf{2}_{C_j}}(u_t, B_0)']$ can be written as

$$E\left[(e(B_0)_a e(B_0)_b e(B_0)_c)(e(B_0)_d e(B_0)_f)\right] \tag{4.131}$$

$a, b, c, d, f = \{1, \ldots, n\}$ and $a \neq b \neq c$, which implies $E\left[(e(B_0)_a e(B_0)_b e(B_0)_c)(e(B_0)_d e(B_0)_f)\right] = 0$ by independence and mean zero of ϵ.

□

4.9.2 Appendix - Preparations part 2

Lemma 4.9. *For b_{pq} and $b_{\tilde{p}\tilde{q}}$ with $p, q, \tilde{p}, \tilde{q} \in \{1, \dots, n\}$ and $p \geq q$ and $\tilde{p} \geq \tilde{q}$ and $q \neq \tilde{q}$ it holds that*

$$(G_{\mathbf{2}}^{b_{pq}})' S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b_{\tilde{p}\tilde{q}}} = 0. \tag{4.132}$$

Proof. For b_{pq} and $b_{\tilde{p}\tilde{q}}$ with $p, q, \tilde{p}, \tilde{q} \in \{1, \dots, n\}$ and $p \geq q$ and $\tilde{p} \geq \tilde{q}$ and $q \neq \tilde{q}$ it holds that

$$(G_{\mathbf{2}}^{b_{pq}})' S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b_{\tilde{p}\tilde{q}}} = (G_{\mathbf{2M}}^{b_{pq}})' S_{\mathbf{2}}^{-1} G_{\mathbf{2M}}^{b_{\tilde{p}\tilde{q}}} + \sum_{i=1}^{n-1} (G_{\mathbf{2C_i}}^{b_{pq}})' S_{\mathbf{2}}^{-1} G_{\mathbf{2C_i}}^{b_{\tilde{p}\tilde{q}}}. \tag{4.133}$$

The statement then follows by plugging in Equation (4.56), (4.57), and (4.58). □

With Lemma 4.9 it holds that $G_{\mathbf{2}}^{b_i\prime} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b_j} = 0$ for $i, j \in \{1, \dots, n\}$ with $i \neq j$. Therefore, for any vector $b_i = [b_{ii}, \dots, b_{ni}]$ representing the impact of the i-th structural shock $\epsilon_{i,t}$ it holds that $G_{\mathbf{2}}^{b_i\prime} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{\neg b_i}$ is zero. Therefore, for any vector $b_i = [b_{ii}, \dots, b_{ni}]$ the right hand side of Equation (4.34) is zero and hence Equation (4.34) simplifies to

$$G_{\mathbf{D}}^{b_i} - S_{\mathbf{D2}} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b_i} = 0. \tag{4.134}$$

The following Lemmas yield analytic expressions for $S_{\mathbf{D2}} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b_i}$ in Equation (4.134).

Lemma 4.10. *For $i, j = 1, \dots, n$ it holds that*

$$S_{\mathbf{3_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} = -2 \frac{1}{\omega_{jjjj} - 1} \tag{4.135}$$
$$\times \underbrace{\begin{bmatrix} \omega_{111} & & 0 & 0 & & & \\ & \ddots & & 0 & & 0_{(n-i)\times(n-i)} & \\ 0 & & \omega_{(i-1)^3} & 0 & & & \\ & & & 0 & \omega_{(i+1)^3} & & 0 \\ & 0_{(i-1)\times(i-1)} & & 0 & & \ddots & \\ & & & 0 & 0 & & \omega_{nnn} \end{bmatrix}}_{(n-1)\times(n)}$$
$$\times \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j)\times(n-j+1)} \end{bmatrix}}_{n\times(n-j+1)}.$$

For $i = j$

$$S_{\mathbf{3_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} = O_{n-1\times n-j+1}. \tag{4.136}$$

For $i < j$

$$S_{\mathbf{3_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} = -\frac{2\omega_{jjj}}{\omega_{jjjj} - 1} \underbrace{\begin{bmatrix} & 0_{(j-2)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j)\times(n-j+1)} \end{bmatrix}}_{(n-1)\times(n-j+1)}. \tag{4.137}$$

For $i > j$

$$S_{\mathbf{3_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} = -\frac{2\omega_{jjj}}{\omega_{jjjj} - 1} \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j-1)\times(n-j+1)} \end{bmatrix}}_{(n-1)\times(n-j+1)}. \tag{4.138}$$

Furthermore, for $i = j$

$$S_{\mathbf{3_{ii}2_{C_i}}} G_{\mathbf{2_{C_i}}}^{b_{i:,i}} = -\omega_{iii} \underbrace{\begin{bmatrix} 0_{(i-1)\times(n-i)} \\ I_{(n-i)\times(n-i)} \end{bmatrix}}_{(n-1)\times(n-i)} \underbrace{\begin{bmatrix} a_{i+1,i} & a_{i+1,i+1} & & 0 \\ \vdots & & \ddots & \\ a_{ni} & a_{n,i+1} & \dots & a_{nn} \end{bmatrix}}_{(n-i)\times(n-i+1)} \tag{4.139}$$

$$= -\omega_{iii} \underbrace{\begin{bmatrix} & 0_{(i-1)\times(n-i+1)} & & \\ a_{i+1,i} & a_{i+1,i+1} & & 0 \\ \vdots & & \ddots & \\ a_{ni} & a_{n,i+1} & \dots & a_{nn} \end{bmatrix}}_{(n-1)\times(n-i+1)}. \tag{4.140}$$

For $i < j$

$$S_{\mathbf{3_{ii}2_{C_j}}} G_{\mathbf{2_{C_j}}}^{b_{j:,j}} = 0_{(n-1)\times(n-j)} \underbrace{\begin{bmatrix} a_{j+1,j} & a_{j+1,j+1} & & 0 \\ \vdots & & \ddots & \\ a_{nj} & a_{n,j+1} & \dots & a_{nn} \end{bmatrix}}_{(n-j)\times(n-j+1)} = 0_{(n-1)\times(n-j+1)}. \tag{4.141}$$

For $i > j$

$$S_{\mathbf{3_{ii}2_{C_j}}} G_{\mathbf{2_{C_j}}}^{b_{j:,j}} = -\omega_{iii} \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j)} & \\ 0_{1\times(i-j-1)} & 1 & 0_{1\times(n-i)} \\ & 0_{(n-j-1)\times(n-j)} & \end{bmatrix}}_{(n-1)\times(n-j)} \underbrace{\begin{bmatrix} a_{j+1,j} & a_{j+1,j+1} & & 0 \\ \vdots & & \ddots & \\ a_{nj} & a_{n,j+1} & \dots & a_{nn} \end{bmatrix}}_{(n-j)\times(n-j+1)} \tag{4.142}$$

$$= -\omega_{iii} \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} & & \\ a_{ij} & \dots & a_{ii} & 0_{1\times(n-i)} \\ & 0_{(n-j-1)\times(n-j+1)} & & \end{bmatrix}}_{(n-1)\times(n-j+1)}. \tag{4.143}$$

Proof. Follows from Lemma 4.3, Lemma 4.6, and Lemma 4.7 and simple matrix algebra. □

Lemma 4.11. *For $i = 1, \dots, n$ and $j = 1, \dots, n-1$ it holds that*

$$S_{\mathbf{4_{iii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} = -2\frac{\omega_{iii}}{\omega_{jjjj} - 1} \tag{4.144}$$

$$\times \underbrace{\begin{bmatrix} \omega_{111} & & 0 & 0 & & & \\ & \ddots & & 0 & & 0_{(n-i)\times(n-i)} & \\ 0 & & \omega_{(i-1)^3} & 0 & & & \\ & & & 0 & \omega_{(i+1)^3} & & 0 \\ & 0_{(i-1)\times(i-1)} & & 0 & & \ddots & \\ & & & 0 & 0 & & \omega_{nnn} \end{bmatrix}}_{(n-1)\times(n)}$$

$$\times \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j)\times(n-j+1)} \end{bmatrix}}_{n\times(n-j+1)}.$$

For $i = j$

$$S_{\mathbf{4_{iii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} = O_{n-1\times n-j+1}. \tag{4.145}$$

For $i < j$

$$S_{\mathbf{4_{iii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} = -\frac{2\omega_{jjj}\omega_{iii}}{\omega_{jjjj}-1} \underbrace{\begin{bmatrix} & 0_{(j-2)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j)\times(n-j+1)} \end{bmatrix}}_{(n-1)\times(n-j+1)}. \tag{4.146}$$

For $i > j$

$$S_{\mathbf{4_{iii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} = -\frac{2\omega_{jjj}\omega_{iii}}{\omega_{jjjj}-1} \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j-1)\times(n-j+1)} \end{bmatrix}}_{(n-1)\times(n-j+1)}. \tag{4.147}$$

Furthermore, for $i = j$

$$S_{\mathbf{4_{iii}2_{C_i}}} G_{\mathbf{2_{C_i}}}^{b_{i:,i}} = -\omega_{iiii} \underbrace{\begin{bmatrix} 0_{(i-1)\times(n-i)} \\ I_{(n-i)\times(n-i)} \end{bmatrix}}_{(n-1)\times(n-i)} \underbrace{\begin{bmatrix} a_{i+1,i} & a_{i+1,i+1} & & 0 \\ \vdots & & \ddots & \\ a_{ni} & a_{n,i+1} & \dots & a_{nn} \end{bmatrix}}_{(n-i)\times(n-i+1)} \tag{4.148}$$

$$= -\omega_{iiii} \underbrace{\begin{bmatrix} & 0_{(i-1)\times(n-i+1)} & & \\ a_{i+1,i} & a_{i+1,i+1} & & 0 \\ \vdots & & \ddots & \\ a_{ni} & a_{n,i+1} & \dots & a_{nn} \end{bmatrix}}_{(n-1)\times(n-i+1)}. \tag{4.149}$$

For $i < j$

$$S_{\mathbf{4_{iii}2_{C_j}}} G_{\mathbf{2_{C_j}}}^{b_{j:,j}} = 0_{(n-1)\times(n-j)} \underbrace{\begin{bmatrix} a_{j+1,j} & a_{j+1,j+1} & & 0 \\ \vdots & & \ddots & \\ a_{nj} & a_{n,j+1} & \dots & a_{nn} \end{bmatrix}}_{(n-j)\times(n-j+1)} = 0_{(n-1)\times(n-j+1)}. \tag{4.150}$$

For $i > j$

$$S_{\mathbf{4_{iii}2_{C_j}}} G_{\mathbf{2_{C_j}}}^{b_{j:,j}} = -\omega_{iiii} \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j)} & \\ 0_{1\times(i-j-1)} & 1 & 0_{1\times(n-i)} \\ & 0_{(n-j-1)\times(n-j)} & \end{bmatrix}}_{(n-1)\times(n-j)} \underbrace{\begin{bmatrix} a_{j+1,j} & a_{j+1,j+1} & & 0 \\ \vdots & & \ddots & \\ a_{nj} & a_{n,j+1} & \dots & a_{nn} \end{bmatrix}}_{(n-j)\times(n-j+1)} \tag{4.151}$$

$$= -\omega_{iiii} \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} & & \\ a_{ij} & \dots & a_{ii} & 0_{1\times(n-i)} \\ & 0_{(n-j-1)\times(n-j+1)} & & \end{bmatrix}}_{(n-1)\times(n-j+1)}. \tag{4.152}$$

For $i = 1, \dots, n-1$ and $j = 1, \dots, n-1$

$$S_{\mathbf{4_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} = -2\frac{1}{\omega_{jjjj} - 1} \times \underbrace{\begin{bmatrix} & \omega_{iiii} - 1 & \omega_{(i+1)^4} - 1 & & 0 \\ 0_{(n-i)\times(i-1)} & \vdots & & \ddots & \\ & \omega_{iiii} - 1 & 0 & & \omega_{nnnn} - 1 \end{bmatrix}}_{(n-i)\times(n)} \times \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j)\times(n-j+1)} \end{bmatrix}}_{n\times(n-j+1)}. \tag{4.153}$$

For $i = j$

$$S_{\mathbf{4_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} = -2\frac{\omega_{iiii}-1}{\omega_{iiii}-1} \underbrace{\begin{bmatrix} a_{jj} & \\ & 0_{(n-i)\times(n-j)} \\ a_{jj} & \end{bmatrix}}_{(n-i)\times(n-j+1)} = -2 \underbrace{\begin{bmatrix} a_{jj} & \\ & 0_{(n-i)\times(n-j)} \\ a_{jj} & \end{bmatrix}}_{(n-i)\times(n-j+1)}. \tag{4.154}$$

For $i < j$

$$S_{\mathbf{4_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} = -2\frac{\omega_{jjjj}-1}{\omega_{jjjj}-1} \underbrace{\begin{bmatrix} & 0_{(j-i-1)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j)\times(n-j+1)} \end{bmatrix}}_{(n-i)\times(n-j+1)} = -2 \underbrace{\begin{bmatrix} & 0_{(j-i-1)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j)\times(n-j+1)} \end{bmatrix}}_{(n-i)\times(n-j+1)}. \tag{4.155}$$

For $i > j$

$$S_{\mathbf{4_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} = 0_{(n-i)\times(n-j+1)}. \tag{4.156}$$

Furthermore,

$$S_{\mathbf{4_{ii}2_{C_i}}} G_{\mathbf{2_{C_i}}}^{b_{i:,i}} = -\omega_{iii} \underbrace{\begin{bmatrix} \omega_{(i+1)^3} & & 0 \\ & \ddots & \\ 0 & & \omega_{nnn} \end{bmatrix}}_{(n-i)\times(n-i)} \underbrace{\begin{bmatrix} a_{i+1,i} & a_{i+1,i+1} & & 0 \\ \vdots & & \ddots & \\ a_{ni} & a_{n,i+1} & \dots & a_{nn} \end{bmatrix}}_{(n-i)\times(n-i+1)} \tag{4.157}$$

$$= -\omega_{iii} \underbrace{\begin{bmatrix} a_{i+1,i}\omega_{(i+1)^3} & a_{i+1,i+1}\omega_{(i+1)^3} & & 0 \\ \vdots & & \ddots & \\ a_{ni}\omega_{nnn} & a_{n,i+1}\omega_{nnn} & \dots & a_{nn}\omega_{nnn} \end{bmatrix}}_{(n-i)\times(n-i+1)}. \tag{4.158}$$

Proof. Follows from Lemma 4.3, Lemma 4.6, and Lemma 4.8 and simple matrix algebra. □

Lemma 4.12. *For $i, j = 1, \ldots, n$ and $i = j$*

$$G^{b_{j:,j}}_{\mathbf{3_{ii}}} - S_{\mathbf{3_{ii}}\mathbf{2}} S_{\mathbf{2}}^{-1} G^{b_{j:,j}}_{\mathbf{2}} = -\omega_{iii} \underbrace{\begin{bmatrix} & 0_{(i-1)\times(n-i+1)} & & \\ a_{i+1,i} & a_{i+1,i+1} & & 0 \\ \vdots & & \ddots & \\ a_{ni} & a_{n,i+1} & \ldots & a_{nn} \end{bmatrix}}_{(n-1)\times(n-i+1)} \tag{4.159}$$

$$+ \omega_{iii} \underbrace{\begin{bmatrix} & 0_{(i-1)\times(n-i+1)} & & \\ a_{i+1,i} & a_{i+1,i+1} & & 0 \\ \vdots & & \ddots & \\ a_{ni} & a_{n,i+1} & \ldots & a_{nn} \end{bmatrix}}_{(n-1)\times(n-i+1)}$$

$$= 0_{(n-1)\times(n-i+1)}. \tag{4.160}$$

For $i, j = 1, \ldots, n$ and $i < j$

$$G^{b_{j:,j}}_{\mathbf{3_{ii}}} - S_{\mathbf{3_{ii}}\mathbf{2}} S_{\mathbf{2}}^{-1} G^{b_{j:,j}}_{\mathbf{2}} = \frac{2\omega_{jjj}}{\omega_{jjjj} - 1} \underbrace{\begin{bmatrix} & 0_{(j-2)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j)\times(n-j+1)} \end{bmatrix}}_{(n-1)\times(n-j+1)}. \tag{4.161}$$

For $i, j = 1, \ldots, n$ and $i > j$

$$G^{b_{j:,j}}_{\mathbf{3_{ii}}} - S_{\mathbf{3_{ii}}\mathbf{2}} S_{\mathbf{2}}^{-1} G^{b_{j:,j}}_{\mathbf{2}} = \frac{2\omega_{jjj}}{\omega_{jjjj} - 1} \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j-1)\times(n-j+1)} \end{bmatrix}}_{(n-1)\times(n-j+1)} \tag{4.162}$$

$$+ \omega_{iii} \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} & & \\ a_{ij} & \ldots & a_{ii} & 0_{1\times(n-i)} \\ & 0_{(n-j-1)\times(n-j+1)} & & \end{bmatrix}}_{(n-1)\times(n-j+1)}.$$

Proof. For $i, j = 1, \dots, n$ let

$$W_{\mathbf{3_{ii}}}^{b_{j:,j}} := S_{\mathbf{3_{ii}2}} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b_{j:,j}}. \tag{4.163}$$

Then, for $i = 1, \dots, n-1$

$$W_{\mathbf{3_{ii}}}^{b_{i:,i}} = -\omega_{iii} \underbrace{\begin{bmatrix} & 0_{(i-1)\times(n-i+1)} & & \\ a_{i+1,i} & a_{i+1,i+1} & & 0 \\ \vdots & & \ddots & \\ a_{ni} & a_{n,i+1} & \dots & a_{nn} \end{bmatrix}}_{(n-1)\times(n-i+1)}, \tag{4.164}$$

for $i = n$

$$W_{\mathbf{3_{ii}}}^{b_{i:,i}} = 0_{(n-1)\times 1}, \tag{4.165}$$

for $i < j$

$$W_{\mathbf{3_{ii}}}^{b_{j:,j}} = -\frac{2\omega_{jjj}}{\omega_{jjjj} - 1} \underbrace{\begin{bmatrix} & 0_{(j-2)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j)\times(n-j+1)} \end{bmatrix}}_{(n-1)\times(n-j+1)}, \tag{4.166}$$

for $i > j$

$$W_{\mathbf{3_{ii}}}^{b_{j:,j}} = -\frac{2\omega_{jjj}}{\omega_{jjjj} - 1} \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j-1)\times(n-j+1)} \end{bmatrix}}_{(n-1)\times(n-j+1)} - \omega_{iii} \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} & & \\ a_{ij} & \dots & a_{ii} & 0_{1\times(n-i)} \\ & 0_{(n-j-1)\times(n-j+1)} & & \end{bmatrix}}_{(n-1)\times(n-j+1)}. \tag{4.167}$$

Moreover, it holds that

$$W_{\mathbf{3_{ii}}}^{b_{j:,j}} = \left[S_{\mathbf{3_{ii}2_M}}, S_{\mathbf{3_{ii}2_{C_1}}}, \ldots, S_{\mathbf{3_{ii}2_{C_{n-1}}}}\right] \begin{bmatrix} S_{\mathbf{2_M}} & 0 \\ 0 & I_{n(n-1)\times n(n-1)} \end{bmatrix}^{-1} \left[G_{\mathbf{2_M}}^{b_{j:,j}}, G_{\mathbf{2_{C_1}}}^{b_{j:,j}}, \ldots, G_{\mathbf{2_{C_{n-1}}}}^{b_{j:,j}}\right]' \tag{4.168}$$

$$= S_{\mathbf{3_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} + \sum_{q=1}^{n-1} S_{\mathbf{3_{ii}2_{C_q}}} G_{\mathbf{2_{C_q}}}^{b_{j:,j}}. \tag{4.169}$$

From Lemma 4.3 it follows that $G_{\mathbf{2_{C_q}}}^{b_{j:,j}} = 0$ for $i \neq j$ and, hence, for $i = 1, \ldots, n$, $j = 1, \ldots, n-1$

$$W_{\mathbf{3_{ii}}}^{b_{j:,j}} = S_{\mathbf{3_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} + S_{\mathbf{3_{ii}2_{C_j}}} G_{\mathbf{2_{C_j}}}^{b_{j:,j}} \tag{4.170}$$

and for $j = n$

$$W_{\mathbf{3_{ii}}}^{b_{j:,j}} = S_{\mathbf{3_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}}. \tag{4.171}$$

With Lemma 4.10 (implying $S_{\mathbf{3_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{i:,i}} = 0$ and $S_{\mathbf{3_{ii}2_{C_j}}} G_{\mathbf{2_{C_j}}}^{b_{j:,j}} = 0$ for $i < j$) it follows that

$$W_{\mathbf{3_{ii}}}^{b_{i:,i}} = S_{\mathbf{3_{ii}2_{C_i}}} G_{\mathbf{2_{C_i}}}^{b_{i:,i}}, \qquad \text{for } i < n, \tag{4.172}$$

$$W_{\mathbf{3_{ii}}}^{b_{i:,i}} = 0_{(n-1)\times 1}, \qquad \text{for } i = n, \tag{4.173}$$

$$W_{\mathbf{3_{ii}}}^{b_{j:,j}} = S_{\mathbf{3_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}}, \qquad \text{for } i < j, \tag{4.174}$$

$$W_{\mathbf{3_{ii}}}^{b_{j:,j}} = S_{\mathbf{3_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} + S_{\mathbf{3_{ii}2_{C_j}}} G_{\mathbf{2_{C_j}}}^{b_{j:,j}}, \qquad \text{for } i > j. \tag{4.175}$$

The statements then follow with Lemma 4.10 and Lemma 4.4.

□

Lemma 4.13. *For $i, j = 1, \ldots, n-1$ and $i = j$*

$$G^{b_{j:,j}}_{\mathbf{4_{ii}}} - S_{\mathbf{4_{ii}2}} S^{-1}_{\mathbf{2}} G^{b_{j:,j}}_{\mathbf{2}} = -2 \underbrace{\begin{bmatrix} a_{ii} & \\ \vdots & 0_{(n-i)\times(n-j)} \\ a_{ii} & \end{bmatrix}}_{(n-i)\times(n-i+1)} + 2 \underbrace{\begin{bmatrix} a_{ii} & \\ \vdots & 0_{(n-i)\times(n-i)} \\ a_{ii} & \end{bmatrix}}_{(n-i)\times(n-i+1)} \tag{4.176}$$

$$+ \omega_{iii} \underbrace{\begin{bmatrix} a_{i+1,i}\omega_{(i+1)^3} & a_{i+1,i+1}\omega_{(i+1)^3} & & 0 \\ \vdots & & \ddots & \\ a_{ni}\omega_{nnn} & a_{n,i+1}\omega_{nnn} & \ldots & a_{nn}\omega_{nnn} \end{bmatrix}}_{(n-i)\times(n-i+1)}$$

$$= \omega_{iii} \underbrace{\begin{bmatrix} a_{i+1,i}\omega_{(i+1)^3} & a_{i+1,i+1}\omega_{(i+1)^3} & & 0 \\ \vdots & & \ddots & \\ a_{ni}\omega_{nnn} & a_{n,i+1}\omega_{nnn} & \ldots & a_{nn}\omega_{nnn} \end{bmatrix}}_{(n-i)\times(n-i+1)}. \tag{4.177}$$

For $i = 1, \ldots, n-1$, $j = 1, \ldots, n$, and $i > j$

$$G^{b_{j:,j}}_{\mathbf{4_{ii}}} - S_{\mathbf{4_{ii}2}} S^{-1}_{\mathbf{2}} G^{b_{j:,j}}_{\mathbf{2}} = 0_{(n-i)\times(n-j+1)} - 0_{(n-i)\times(n-j+1)} \tag{4.178}$$

$$= 0_{(n-i)\times(n-j+1)}. \tag{4.179}$$

For $i = 1, \ldots, n-1$, $j = 1, \ldots, n$, and $i < j$

$$G^{b_{j:,j}}_{\mathbf{4_{ii}}} - S_{\mathbf{4_{ii}2}} S^{-1}_{\mathbf{2}} G^{b_{j:,j}}_{\mathbf{2}} = -2 \underbrace{\begin{bmatrix} & 0_{(j-i-1)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j)\times(n-j+1)} \end{bmatrix}}_{(n-i)\times(n-j+1)} + 2 \underbrace{\begin{bmatrix} & 0_{(j-i-1)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j)\times(n-j+1)} \end{bmatrix}}_{(n-i)\times(n-j+1)} \tag{4.180}$$

$$= 0_{(n-i)\times(n-j+1)}. \tag{4.181}$$

Proof. For $i = 1, \ldots, n-1$, $j = 1, \ldots, n$ let

$$W^{b_{j:,j}}_{\mathbf{4_{ii}}} := S_{\mathbf{4_{ii}2}} S^{-1}_{\mathbf{2}} G^{b_{j:,j}}_{\mathbf{2}}. \tag{4.182}$$

Then, for $i = 1, \ldots, n-1$

$$W_{\mathbf{4_{ii}}}^{b_{i:,i}} = -2 \underbrace{\begin{bmatrix} a_{ii} & \\ & 0_{(n-i)\times(n-i)} \\ a_{ii} & \end{bmatrix}}_{(n-i)\times(n-i+1)} - \omega_{iii} \underbrace{\begin{bmatrix} a_{i+1,i}\omega_{(i+1)^3} & a_{i+1,i+1}\omega_{(i+1)^3} & & 0 \\ \vdots & & \ddots & \\ a_{ni}\omega_{nnn} & a_{n,i+1}\omega_{nnn} & \ldots & a_{nn}\omega_{nnn} \end{bmatrix}}_{(n-i)\times(n-i+1)}, \quad (4.183)$$

for $i < j$

$$W_{\mathbf{4_{ii}}}^{b_{j:,j}} = -2 \underbrace{\begin{bmatrix} & 0_{(j-i-1)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j)\times(n-j+1)} \end{bmatrix}}_{(n-i)\times(n-j+1)}, \quad (4.184)$$

and for $i > j$

$$W_{\mathbf{4_{ii}}}^{b_{j:,j}} = 0_{(n-i)\times(n-j+1)}. \quad (4.185)$$

Moreover, it holds that

$$W_{\mathbf{4_{ii}}}^{b_{j:,j}} = \begin{bmatrix} S_{\mathbf{4_{ii}2_M}} & S_{\mathbf{4_{ii}2_{C_1}}} & \ldots & S_{\mathbf{4_{ii}2_{C_{n-1}}}} \end{bmatrix} \begin{bmatrix} S_{\mathbf{2_M}} & 0 \\ 0 & I_{n(n-1)\times n(n-1)} \end{bmatrix}^{-1} \begin{bmatrix} G_{\mathbf{2_M}}^{b_{j:,j}}, G_{\mathbf{2_{C_1}}}^{b_{j:,j}}, \ldots, G_{\mathbf{2_{C_{n-1}}}}^{b_{j:,j}} \end{bmatrix}' \quad (4.186)$$

$$= S_{\mathbf{4_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} + \sum_{q=1}^{n-1} S_{\mathbf{4_{ii}2_{C_q}}} G_{\mathbf{2_{C_q}}}^{b_{j:,j}}. \quad (4.187)$$

From Lemma 4.3 it follows that $G_{\mathbf{2_{C_q}}}^{b_{j:,j}} = 0$ for $i \neq j$ and, hence, for $i = 1, \ldots, n$, $j = 1, \ldots, n-1$

$$W_{\mathbf{4_{ii}}}^{b_{j:,j}} = S_{\mathbf{4_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} + S_{\mathbf{4_{ii}2_{C_j}}} G_{\mathbf{2_{C_j}}}^{b_{j:,j}} \quad (4.188)$$

and for $j = n$

$$W_{\mathbf{4_{ii}}}^{b_{j:,j}} = S_{\mathbf{4_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}}. \quad (4.189)$$

From Lemma 4.8 it follows that $S_{\mathbf{4_{ii}2_{C_j}}} = 0$ for $i \neq j$ and, hence,

$$W_{\mathbf{4_{ii}}}^{b_{i:,i}} = S_{\mathbf{4_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{i:,i}} + S_{\mathbf{4_{ii}2_{C_i}}} G_{\mathbf{2_{C_i}}}^{b_{i:,i}} \tag{4.190}$$

$$W_{\mathbf{4_{ii}}}^{b_{j:,j}} = S_{\mathbf{4_{ii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}}, \qquad \text{for } i \neq j. \tag{4.191}$$

The statements then follow with Lemma 4.11 and Lemma 4.5.

□

Lemma 4.14. *For $i\,, j = 1, \ldots, n$ and $i = j$*

$$G_{\mathbf{4_{iii}}}^{b_{j:,j}} - S_{\mathbf{4_{iii}2}} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b_{j:,j}} = -\omega_{iiii} \underbrace{\begin{bmatrix} & 0_{(i-1)\times(n-i+1)} & & \\ a_{i+1,i} & a_{i+1,i+1} & & 0 \\ \vdots & & \ddots & \\ a_{ni} & \ldots & & a_{nn} \end{bmatrix}}_{(n-1)\times(n-i+1)} \tag{4.192}$$

$$+ \omega_{iiii} \underbrace{\begin{bmatrix} & 0_{(i-1)\times(n-i+1)} & & \\ a_{i+1,i} & a_{i+1,i+1} & & 0 \\ \vdots & & \ddots & \\ a_{ni} & a_{n,i+1} & \ldots & a_{nn} \end{bmatrix}}_{(n-1)\times(n-i+1)}$$

$$= 0_{(n-1)\times(n-i+1)}. \tag{4.193}$$

For $i, j = 1, \ldots, n$ *and* $i > j$

$$G_{\mathbf{4_{iii}}}^{b_{j:,j}} - S_{\mathbf{4_{iii}2}} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b_{j:,j}} = -3 \underbrace{\begin{bmatrix} & & 0_{(j-1)\times(n-j+1)} & & \\ a_{ij} & & \ldots & a_{ii} & 0_{1\times(n-i)} \\ & & 0_{(n-j-1)\times(n-j+1)} & & \end{bmatrix}}_{(n-1)\times(n-j+1)} \tag{4.194}$$

$$+ \frac{2\omega_{jjj}\omega_{iii}}{\omega_{jjjj} - 1} \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j-1)\times(n-j+1)} \end{bmatrix}}_{(n-1)\times(n-j+1)}$$

$$+ \omega_{iiii} \underbrace{\begin{bmatrix} & & 0_{(j-1)\times(n-j+1)} & & \\ a_{ij} & & \ldots & a_{ii} & 0_{1\times(n-i)} \\ & & 0_{(n-j-1)\times(n-j+1)} & & \end{bmatrix}}_{(n-1)\times(n-j+1)} \tag{4.195}$$

$$= (\omega_{iiii} - 3) \underbrace{\begin{bmatrix} & & 0_{(j-1)\times(n-j+1)} & & \\ a_{ij} & & \ldots & a_{ii} & 0_{1\times(n-i)} \\ & & 0_{(n-j-1)\times(n-j+1)} & & \end{bmatrix}}_{(n-1)\times(n-j+1)} \tag{4.196}$$

$$+ \frac{2\omega_{jjj}\omega_{iii}}{\omega_{jjjj} - 1} \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} \\ a_{jj} & 0_{1\times(n-j)} \\ & 0_{(n-j-1)\times(n-j+1)} \end{bmatrix}}_{(n-1)\times(n-j+1)}.$$

For $i, j = 1, \ldots, n$ and $i < j$

$$G_{\mathbf{4_{iii}}}^{b_{j:,j}} - S_{\mathbf{4_{iii}2}} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b_{j:,j}} = 0_{(n-1)\times(n-j+1)} + \frac{2\omega_{jjj}\omega_{iii}}{\omega_{jjjj} - 1} \underbrace{\begin{bmatrix} & 0_{(j-2)\times(n-j+1)} & \\ a_{jj} & 0_{1\times(n-j)} & \\ & 0_{(n-j)\times(n-j+1)} & \end{bmatrix}}_{(n-1)\times(n-j+1)} \tag{4.197}$$

$$= \frac{2\omega_{jjj}\omega_{iii}}{\omega_{jjjj} - 1} \underbrace{\begin{bmatrix} & 0_{(j-2)\times(n-j+1)} & \\ a_{jj} & 0_{1\times(n-j)} & \\ & 0_{(n-j)\times(n-j+1)} & \end{bmatrix}}_{(n-1)\times(n-j+1)}. \tag{4.198}$$

Proof. For $i, j = 1, \ldots, n$ let

$$W_{\mathbf{4_{iii}}}^{b_{j:,j}} := S_{\mathbf{4_{iii}2}} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b_{j:,j}}. \tag{4.199}$$

Then, for $i = 1, \ldots, n-1$

$$W_{\mathbf{4_{iii}}}^{b_{i:,i}} = -\omega_{iiii} \underbrace{\begin{bmatrix} & 0_{(i-1)\times(n-i+1)} & & \\ a_{i+1,i} & a_{i+1,i+1} & & 0 \\ \vdots & & \ddots & \\ a_{ni} & a_{n,i+1} & \ldots & a_{nn} \end{bmatrix}}_{(n-1)\times(n-i+1)}. \tag{4.200}$$

For $i = n$

$$W_{\mathbf{4_{iii}}}^{b_{j:,j}} = 0_{(n-1)\times 1}. \tag{4.201}$$

For $i < j$

$$W_{\mathbf{4_{iii}}}^{b_{j:,j}} = -\frac{2\omega_{jjj}\omega_{iii}}{\omega_{jjjj}-1} \underbrace{\begin{bmatrix} & 0_{(j-2)\times(n-j+1)} & \\ a_{jj} & 0_{1\times(n-j)} & \\ & 0_{(n-j)\times(n-j+1)} & \end{bmatrix}}_{(n-1)\times(n-j+1)}. \tag{4.202}$$

For $i > j$

$$W_{\mathbf{4_{iii}}}^{b_{j:,j}} = -\frac{2\omega_{jjj}\omega_{iii}}{\omega_{jjjj}-1} \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} & \\ a_{jj} & 0_{1\times(n-j)} & \\ & 0_{(n-j-1)\times(n-j+1)} & \end{bmatrix}}_{(n-1)\times(n-j+1)} - \omega_{iiii} \underbrace{\begin{bmatrix} & 0_{(j-1)\times(n-j+1)} & & \\ a_{ij} & \dots & a_{ii} & 0_{1\times(n-i)} \\ & 0_{(n-j-1)\times(n-j+1)} & & \end{bmatrix}}_{(n-1)\times(n-j+1)}. \tag{4.203}$$

Moreover, it holds that

$$W_{\mathbf{4_{iii}}}^{b_{j:,j}} = \left[S_{\mathbf{4_{iii}2_M}}, S_{\mathbf{4_{iii}2_{C_1}}}, \dots, S_{\mathbf{4_{iii}2_{C_{n-1}}}}\right] \begin{bmatrix} S_{\mathbf{2_M}} & 0 \\ 0 & I_{n(n-1)\times n(n-1)} \end{bmatrix}^{-1} \left[G_{\mathbf{2_M}}^{b_{j:,j}}, G_{\mathbf{2_{C_1}}}^{b_{j:,j}}, \dots, G_{\mathbf{2_{C_{n-1}}}}^{b_{j:,j}}\right]' \tag{4.204}$$

$$= S_{\mathbf{4_{iii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} + \sum_{q=1}^{n-1} S_{\mathbf{4_{iii}2_{C_q}}} G_{\mathbf{2_{C_q}}}^{b_{j:,j}}. \tag{4.205}$$

From Lemma 4.3 it follows that $G_{\mathbf{2_{C_q}}}^{b_{j:,j}} = 0$ for $i \neq j$ and, hence, for $i = 1, \dots, n$, $j = 1, \dots, n-1$

$$W_{\mathbf{4_{iii}}}^{b_{j:,j}} = S_{\mathbf{4_{iii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} + S_{\mathbf{4_{iii}2_{C_j}}} G_{\mathbf{2_{C_j}}}^{b_{j:,j}} \tag{4.206}$$

and for $j = n$

$$W_{\mathbf{4_{iii}}}^{b_{j:,j}} = S_{\mathbf{4_{iii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}}. \tag{4.207}$$

With Lemma 4.11 it follows that

$$W_{\mathbf{4_{iii}}}^{b_{i:,i}} = S_{\mathbf{4_{iii}2_{C_i}}} G_{\mathbf{2_{C_i}}}^{b_{i:,i}}, \qquad \text{for } i < n, \tag{4.208}$$

$$W_{\mathbf{4_{iii}}}^{b_{i:,i}} = 0_{(n-1)\times 1}, \qquad \text{for } i = n, \tag{4.209}$$

$$W_{\mathbf{4_{iii}}}^{b_{j:,j}} = S_{\mathbf{4_{iii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}}, \qquad \text{for } i < j, \tag{4.210}$$

$$W_{\mathbf{4_{iii}}}^{b_{j:,j}} = S_{\mathbf{4_{iii}2_M}} S_{\mathbf{2_M}}^{-1} G_{\mathbf{2_M}}^{b_{j:,j}} + S_{\mathbf{4_{iii}2_{C_j}}} G_{\mathbf{2_{C_j}}}^{b_{j:,j}}, \qquad \text{for } i > j. \tag{4.211}$$

The statements then follow with Lemma 4.11 and Lemma 4.5.

□

4.9.3 Appendix - Final Lemma

We now combine the conditions in Lemma 4.12 - 4.14 into conditions for specific moment conditions.

Lemma 4.15. *In a recursive SVAR with independent shocks, it holds that for* $i, j, k, l \in \{1, \ldots, n\}$

1. *coskewness moment condition* $E[f_{\mathbf{D}}(B, u_t)] = E[e(B)_i e(B)_j e(B)_k]$ *with* $i \neq j \neq k$ *satisfy*

$$G_{\mathbf{D}}^{b} - S_{\mathbf{D2}} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b} = 0 \tag{4.212}$$

   for every unrestricted element b *of* B_0.

2. *coskewness moment condition* $E[f_{\mathbf{D}}(B, u_t)] = E[e(B)_i^2 e(B)_j]$ *with* $i \neq j$ *satisfy*

$$G_{\mathbf{D}}^{b_{pq}} - S_{\mathbf{D2}} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b_{pq}} = \begin{cases} \frac{2E[\epsilon_{j,t}^3]}{E[\epsilon_{j,t}^4]-1} a_{jp}, & \text{if} p = j, \quad q = j, \quad i < j, \\ \frac{2E[\epsilon_{j,t}^3]}{E[\epsilon_{j,t}^4]-1} a_{jp} + E[\epsilon_{i,t}^3] a_{ip}, & \text{if } q = j, \quad p = j, \quad i > j, \\ E[\epsilon_{i,t}^3] a_{ip}, & \text{if } q = j, \quad p = j+1, \ldots, i, \quad i > j, \\ 0, & \text{else} \end{cases}. \tag{4.213}$$

3. *cokurtosis conditions* $E[f_{\mathbf{D}}(B, u_t)] = E[e(B_0)_i e(B_0)_j e(B_0)_k e(B_0)_l]$ *and* $E[f_{\mathbf{D}}(B, u_t)] =$

$E[e(B_0)_i^2 e(B_0)_j e(B_0)_k]$ with $i \neq j \neq k \neq l$ satisfy

$$G_{\mathbf{D}}^{b} - S_{\mathbf{D2}} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b} = 0 \tag{4.214}$$

for every unrestricted element b of B_0.

4. *cokurtosis conditions $E[f_{\mathbf{D}}(B, u_t)] = E[e(B_0)_i^2 e(B_0)_j^2 - 1]$ with $i \neq j$ satisfy*

$$G_{\mathbf{D}}^{b_{pq}} - S_{\mathbf{D2}} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b_{pq}} = \begin{cases} E[\epsilon_{i,t}^3] E[\epsilon_{j,t}^3] a_{jp}, & \textit{if } q = i, \quad p = i, \ldots, j, \\ 0, & \textit{else} \end{cases}. \tag{4.215}$$

5. *cokurtosis conditions $E[f_{\mathbf{D}}(B, u_t)] = E[e(B_0)_i^3 e(B_0)_j]$ with $i \neq j$ satisfy*

$$G_{\mathbf{D}}^{b_{pq}} - S_{\mathbf{D2}} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b_{pq}} = \begin{cases} \frac{2E[\epsilon_{j,t}^3] E[\epsilon_{i,t}^3]}{E[\epsilon_{j,t}^4] - 1} a_{jp}, & \textit{if } p = j, \quad q = j, \quad i < j, \\ \frac{2E[\epsilon_{j,t}^3] E[\epsilon_{i,t}^3]}{E[\epsilon_{j,t}^4] - 1} a_{jp} + (E[\epsilon_{i,t}^4] - 3) a_{ip}, & \textit{if } q = j, \quad p = j, \quad i > j, \\ (E[\epsilon_{i,t}^4] - 3) a_{ip}, & \textit{if } q = j, \quad p = j+1, \ldots, i > j, \\ 0, & \textit{else} \end{cases}. \tag{4.216}$$

Proof. The statements directly follow from Lemma 4.12 - 4.14. □

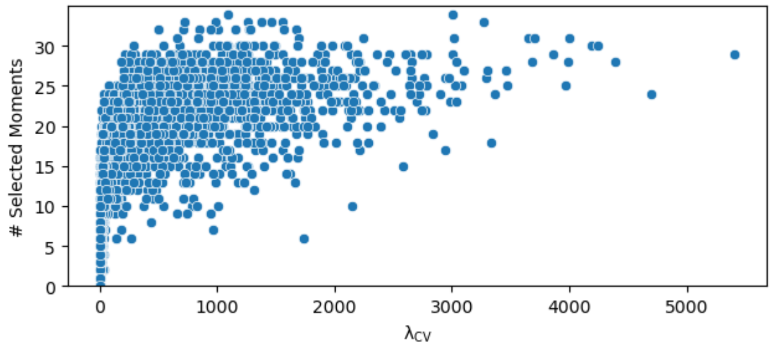
4.10 Appendix - Finite sample performance

Table 4.4: Finite sample performance - Recursive SVAR and the pGMM estimator including Post-LASSO.

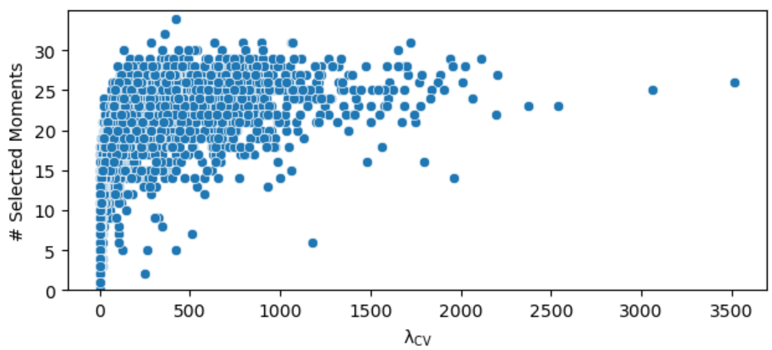
		GMM	oGMM	GMM-Oracle	pGMM	Post-pGMM
$T = 100$	$\hat{B}$	$\begin{bmatrix} \underset{(1.09)}{9.93} & . & . & . \\ \underset{(1.21)}{4.98} & \underset{(1.02)}{9.86} & . & . \\ \underset{(1.49)}{4.97} & \underset{(1.29)}{4.95} & \underset{(1.12)}{9.83} & . \\ \underset{(1.71)}{4.96} & \underset{(1.46)}{4.93} & \underset{(1.27)}{4.91} & \underset{(1.08)}{9.78} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.07)}{9.77} & . & . & . \\ \underset{(1.31)}{4.90} & \underset{(1.01)}{9.71} & . & . \\ \underset{(1.69)}{4.89} & \underset{(1.43)}{4.88} & \underset{(1.10)}{9.70} & . \\ \underset{(2.07)}{4.90} & \underset{(1.74)}{4.88} & \underset{(1.46)}{4.88} & \underset{(1.09)}{9.69} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.07)}{9.76} & . & . & . \\ \underset{(1.17)}{4.91} & \underset{(1.02)}{9.70} & . & . \\ \underset{(1.50)}{4.91} & \underset{(1.26)}{4.88} & \underset{(1.10)}{9.69} & . \\ \underset{(1.81)}{4.92} & \underset{(1.51)}{4.88} & \underset{(1.25)}{4.88} & \underset{(1.10)}{9.67} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.09)}{9.96} & . & . & . \\ \underset{(1.15)}{5.00} & \underset{(1.01)}{9.88} & . & . \\ \underset{(1.46)}{4.98} & \underset{(1.22)}{4.96} & \underset{(1.11)}{9.85} & . \\ \underset{(1.71)}{4.99} & \underset{(1.42)}{4.96} & \underset{(1.21)}{4.95} & \underset{(1.10)}{9.82} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.06)}{9.84} & . & . & . \\ \underset{(1.09)}{4.96} & \underset{(1.01)}{9.79} & . & . \\ \underset{(1.39)}{4.94} & \underset{(1.19)}{4.93} & \underset{(1.11)}{9.77} & . \\ \underset{(1.61)}{4.94} & \underset{(1.38)}{4.93} & \underset{(1.18)}{4.91} & \underset{(1.11)}{9.73} \end{bmatrix}$
	#Mo	10.00	57.00	40.00	24.22	24.22
	Bias	−0.0883	−0.1806	−0.1804	−0.0650	−0.1256
	MSE	1.27	1.40	1.28	1.25	1.21
	λ	.	.	.	71.08	.
		GMM	oGMM	GMM-Oracle	pGMM	Post-pGMM
$T = 250$	$\hat{B}$	$\begin{bmatrix} \underset{(0.43)}{9.97} & . & . & . \\ \underset{(0.51)}{4.99} & \underset{(0.43)}{9.96} & . & . \\ \underset{(0.64)}{4.98} & \underset{(0.52)}{5.00} & \underset{(0.45)}{9.93} & . \\ \underset{(0.72)}{4.98} & \underset{(0.61)}{4.99} & \underset{(0.51)}{4.98} & \underset{(0.45)}{9.91} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.40)}{9.90} & . & . & . \\ \underset{(0.49)}{4.96} & \underset{(0.40)}{9.90} & . & . \\ \underset{(0.65)}{4.96} & \underset{(0.51)}{4.97} & \underset{(0.42)}{9.87} & . \\ \underset{(0.73)}{4.97} & \underset{(0.61)}{4.96} & \underset{(0.49)}{4.97} & \underset{(0.42)}{9.86} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.40)}{9.90} & . & . & . \\ \underset{(0.44)}{4.97} & \underset{(0.40)}{9.90} & . & . \\ \underset{(0.59)}{4.97} & \underset{(0.46)}{4.97} & \underset{(0.42)}{9.87} & . \\ \underset{(0.65)}{4.98} & \underset{(0.54)}{4.97} & \underset{(0.44)}{4.96} & \underset{(0.42)}{9.85} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.42)}{9.99} & . & . & . \\ \underset{(0.45)}{5.01} & \underset{(0.41)}{9.97} & . & . \\ \underset{(0.59)}{5.01} & \underset{(0.46)}{5.02} & \underset{(0.42)}{9.94} & . \\ \underset{(0.66)}{5.02} & \underset{(0.55)}{5.01} & \underset{(0.44)}{5.00} & \underset{(0.43)}{9.92} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.41)}{9.93} & . & . & . \\ \underset{(0.44)}{4.98} & \underset{(0.41)}{9.92} & . & . \\ \underset{(0.57)}{4.98} & \underset{(0.45)}{4.99} & \underset{(0.43)}{9.89} & . \\ \underset{(0.64)}{4.99} & \underset{(0.54)}{4.98} & \underset{(0.44)}{4.97} & \underset{(0.44)}{9.87} \end{bmatrix}$
	#Mo	10.00	57.00	40.00	27.20	27.20
	Bias	−0.0311	−0.0676	−0.0656	−0.0114	−0.0480
	MSE	0.53	0.51	0.48	0.48	0.48
	λ	.	.	.	118.92	.
		GMM	oGMM	GMM-Oracle	pGMM	Post-pGMM
$T = 1000$	$\hat{B}$	$\begin{bmatrix} \underset{(0.11)}{10.00} & . & . & . \\ \underset{(0.13)}{5.00} & \underset{(0.11)}{9.99} & . & . \\ \underset{(0.15)}{4.99} & \underset{(0.13)}{4.99} & \underset{(0.11)}{9.99} & . \\ \underset{(0.19)}{4.99} & \underset{(0.15)}{4.99} & \underset{(0.13)}{4.99} & \underset{(0.11)}{9.98} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.10)}{9.98} & . & . & . \\ \underset{(0.12)}{4.99} & \underset{(0.10)}{9.97} & . & . \\ \underset{(0.13)}{4.99} & \underset{(0.11)}{4.99} & \underset{(0.10)}{9.98} & . \\ \underset{(0.16)}{4.99} & \underset{(0.14)}{4.99} & \underset{(0.11)}{4.99} & \underset{(0.10)}{9.97} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.10)}{9.98} & . & . & . \\ \underset{(0.11)}{4.99} & \underset{(0.10)}{9.97} & . & . \\ \underset{(0.13)}{4.99} & \underset{(0.10)}{4.99} & \underset{(0.10)}{9.98} & . \\ \underset{(0.15)}{4.99} & \underset{(0.13)}{4.99} & \underset{(0.11)}{4.99} & \underset{(0.10)}{9.97} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.11)}{10.00} & . & . & . \\ \underset{(0.11)}{5.00} & \underset{(0.10)}{9.99} & . & . \\ \underset{(0.13)}{5.00} & \underset{(0.11)}{5.00} & \underset{(0.10)}{10.00} & . \\ \underset{(0.16)}{5.00} & \underset{(0.13)}{5.00} & \underset{(0.11)}{5.00} & \underset{(0.10)}{9.98} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.11)}{9.99} & . & . & . \\ \underset{(0.11)}{5.00} & \underset{(0.11)}{9.98} & . & . \\ \underset{(0.13)}{4.99} & \underset{(0.11)}{4.99} & \underset{(0.10)}{9.98} & . \\ \underset{(0.16)}{5.00} & \underset{(0.13)}{4.99} & \underset{(0.11)}{4.99} & \underset{(0.11)}{9.97} \end{bmatrix}$
	#Mo	10.00	57.00	40.00	29.59	29.59
	Bias	−0.0076	−0.0158	−0.0158	−0.0021	−0.0122
	MSE	0.13	0.12	0.11	0.12	0.12
	λ	.	.	.	75.34	.

The table reports the average $\bar{b}_{ij}$ and the corresponding estimated MSE (in parentheses) of each estimated element in $\hat{B}$ as well as the BIAS and MSE across estimated elements in $\hat{B}$ over $3,500$ Monte Carlo replicates for the GMM estimator, the oGMM estimator, the GMM-Oracle estimator, the pGMM estimator, and the Post-pGMM estimator. The Post-pGMM estimator uses only the overidentifying moment conditions selected by the pGMM estimator for the estimation of the block-recursive SVAR. All estimator use zero restrictions which are highlighted by the dots.

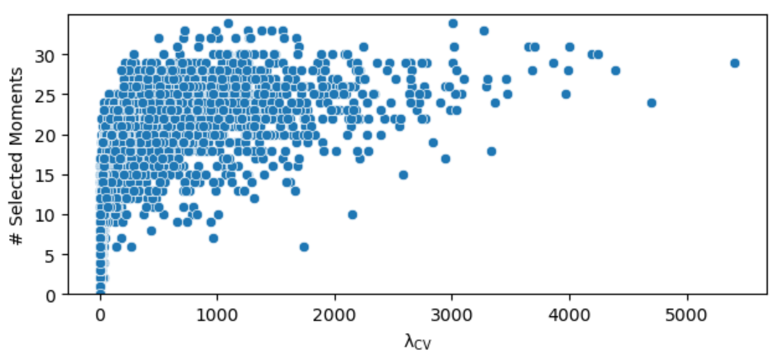
Figure 4.6: Finite sample performance - Relationship of chosen λ_{CV} and the number of selected moments across Monte Carlo runs in the recursive SVAR.



(a) $T = 100$



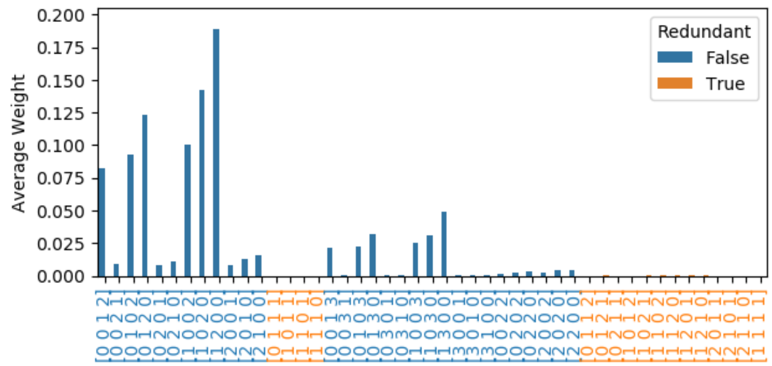
(b) $T = 250$



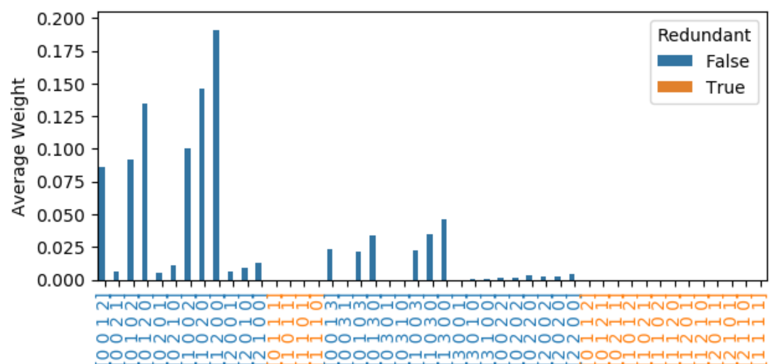
(c) $T = 1000$

Note: The figure shows the chosen λ_{CV} in the cross-validation and the corresponding number of selected moments for each of the $M = 3,500$ Monte Carlo simulations.

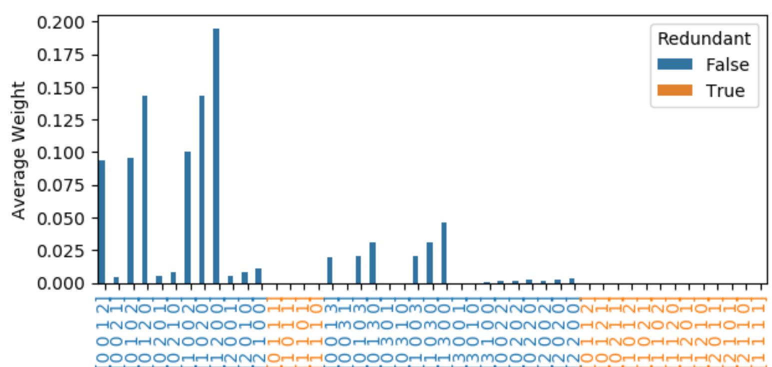
Figure 4.7: Finite sample performance - Average weight of moments across Monte Carlo runs.



(a) $T = 100$



(b) $T = 250$



(c) $T = 1000$

Note: Redundant moments (orange) and relevant moments (blue) are displayed on the x-axis. Each x-axis label abbreviates a moment condition, e.g., $[0, 1, 2, 1]$ corresponds to $E[e(B)^0_{1,t}\, e(B)^1_{2,t}\, e(B)^2_{3,t}\, e(B)^1_{4,t}]$.

4.11 Appendix - Empirical illustration

This section contains supplementary material and robustness checks for the application presented in Section 4.6. Table 4.5 shows descriptive statistics of the variables used in the SVAR. Table 4.6 shows the correlation between the estimated structural shocks from the block-recursive SVAR pGMM estimator and the reduced form shocks. Figure 4.8 shows the historical evolution of the real oil price.

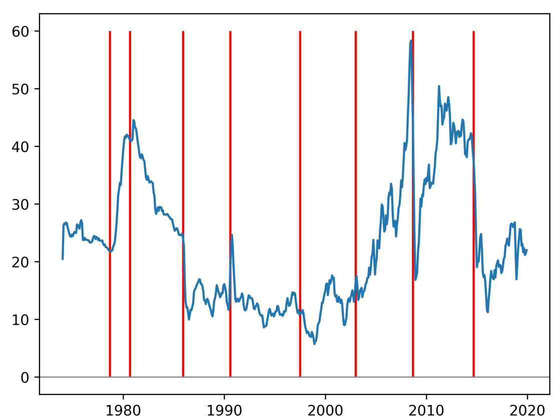
Table 4.5: Oil market SVAR - Descriptive statistics.

	Mean	Median	Std. deviation	Variance	Skewness	Kurtosis
O_t	0.078	0.19	1.5	2.26	-1.66	10.8
Y_t	0.20	0.29	0.60	0.37	-1.2	5.21
OP_t	0.32	0.03	7.31	53.4	0.06	4.46
SR_t	0.34	0.62	3.61	13.03	-0.82	3.67

Table 4.6: Oil market SVAR - Correlation of reduced form and estimated structural shocks.

	u^O	u^Y	u^{OP}	u^{SR}
ε^s	1	-0.03	-0.13	-0.05
ε^d	0.06	1	0.12	0.06
ε^{s-exp}	-0.08	0.02	0.94	-0.27
ε^{d-exp}	-0.05	0.02	0.33	0.96

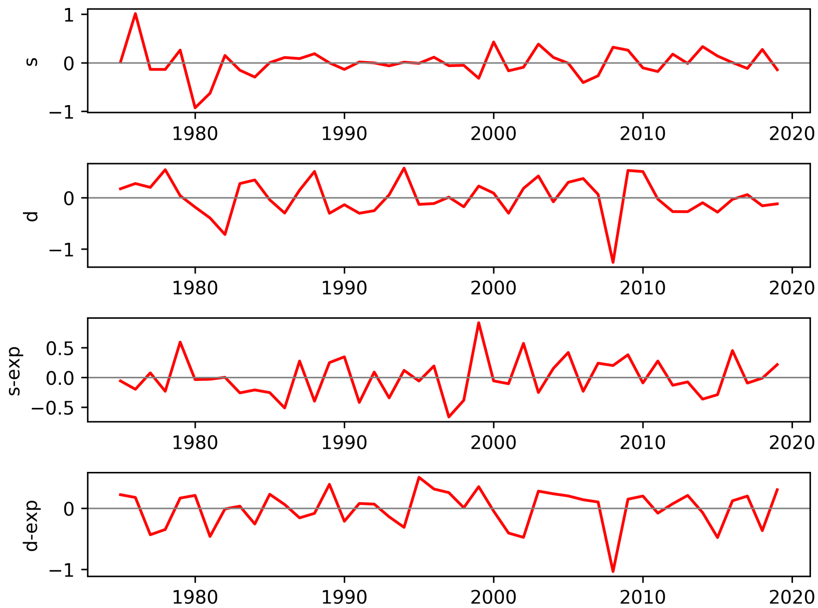
Figure 4.8: Real oil price.



Note: The vertical bars indicate the following events: Iranian Revolution 1978 : 9, Iran Iraq War 1980 : 9, collapse of OPEC 1985 : 12, Persian Gulf War 1990 : 8, Asian Financial Crisis of 1997 :7, Iraq War 2003 : 1, the collapse of Lehman Brothers (2008 : 9), and the oil price decline in mid 2014.

Figure 4.9 shows the estimated structural shocks across years.

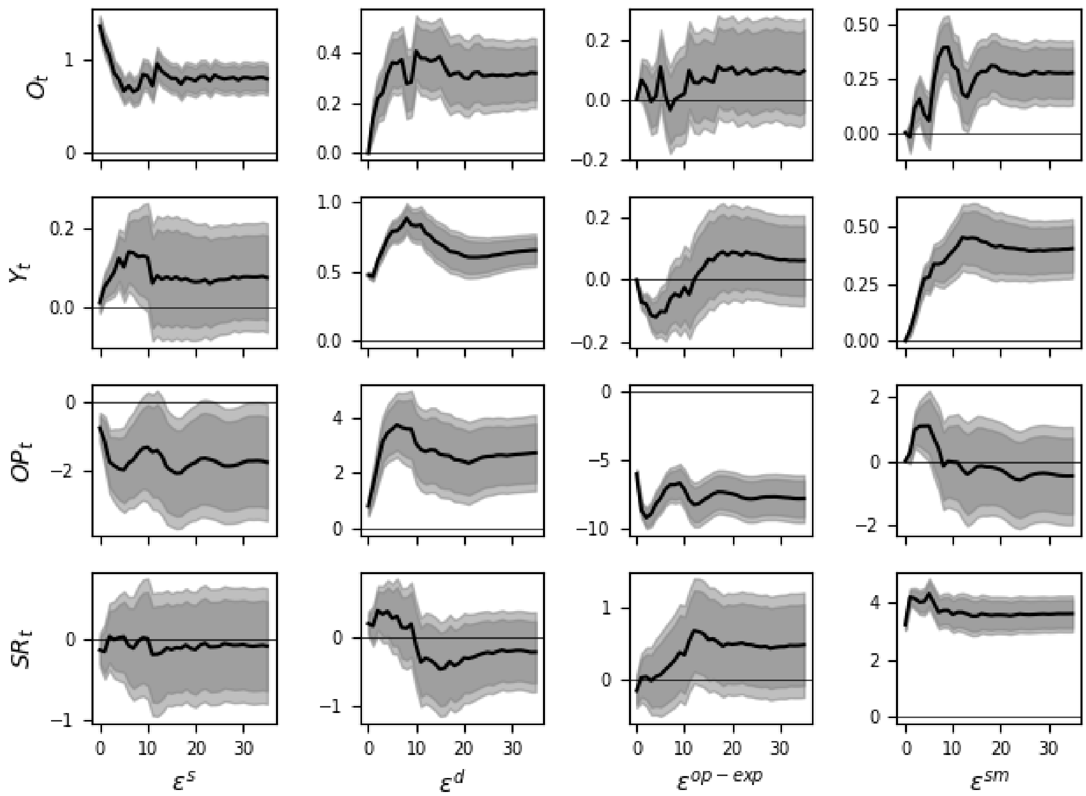
Figure 4.9: Oil market SVAR - Estimated structural shocks, averaged to annual frequency.



Note: The figure shows the average across years for each estimated structural shocks of the block-recursive SVAR pGMM estimator.

Figure 4.10 shows the IRF for the recursive oil market SVAR from Section 4.6 estimated with the SVAR GMM estimator from Equation (4.15). In the recursive SVAR, the SVAR GMM estimator is just identified and equal to the estimator obtained by applying the Cholesky decomposition to the variance-covariance matrix of the reduced form shocks.

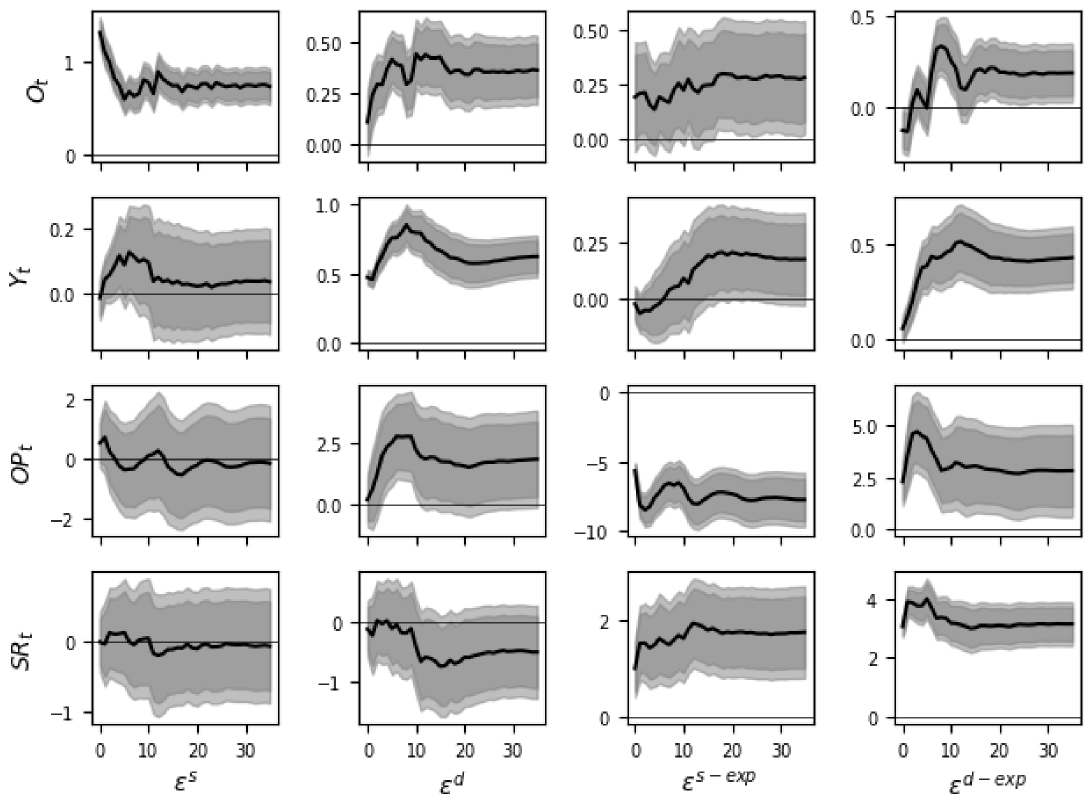
Figure 4.10: Oil market SVAR - Impulse responses (recursive SVAR).



Note: Impulse responses to the recursive oil market SVAR from Section 4.6 estimated with the recursive SVAR GMM estimator from Equation (4.15), equal to the estimator obtained by applying the Cholesky decomposition to the variance-covariance matrix of the reduced form shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows show the cumulative responses. The shock ε^{op-exp} denotes a speculative oil price shock and the shock ε^{sm} represents a residual stock market shock.

Figure 4.11 shows the IRF for the unrestricted oil market SVAR from Section 4.6 estimated with the unrestricted SVAR GMM estimator from Equation (4.15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks.

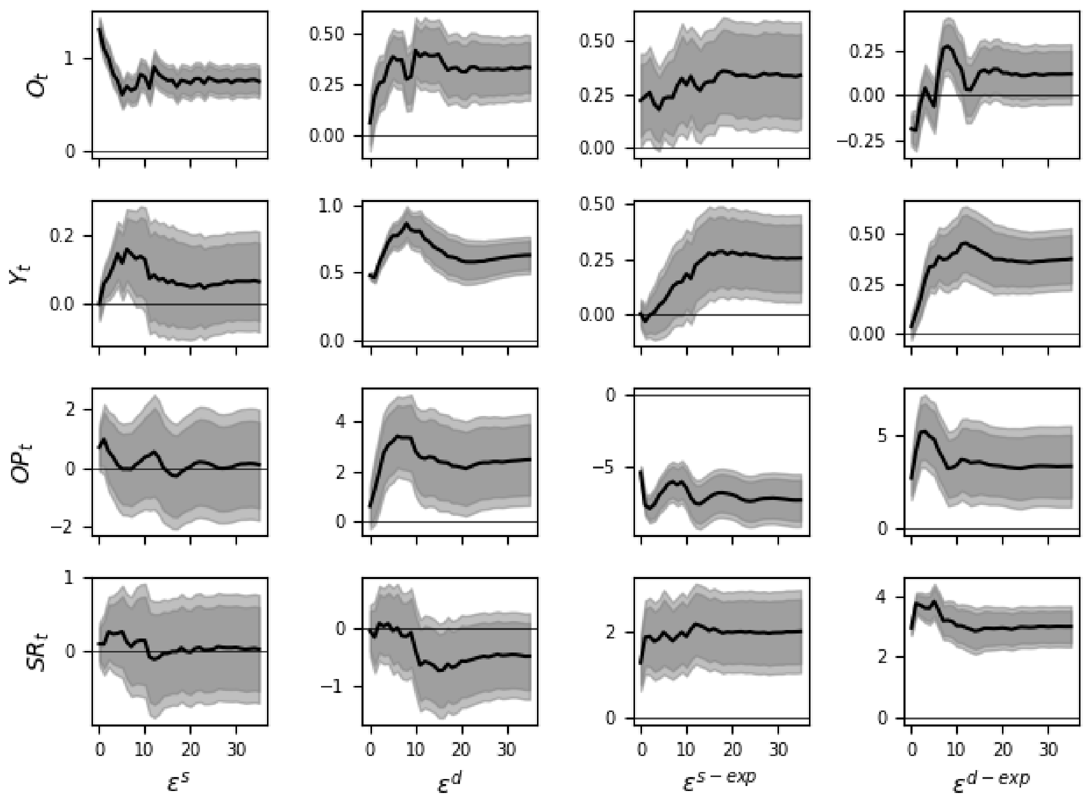
Figure 4.11: Oil market SVAR - Impulse responses (unrestricted SVAR with GMM).



Note: Impulse responses to the estimated structural shocks for the unrestricted oil market SVAR from Section 4.6 estimated with the unrestricted SVAR GMM estimator from Equation (4.15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows show the cumulative responses.

Figure 4.12 shows the IRF for the unrestricted oil market SVAR from Section 4.6 estimated with the overidentified restricted SVAR GMM estimator from Equation (4.19) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks.

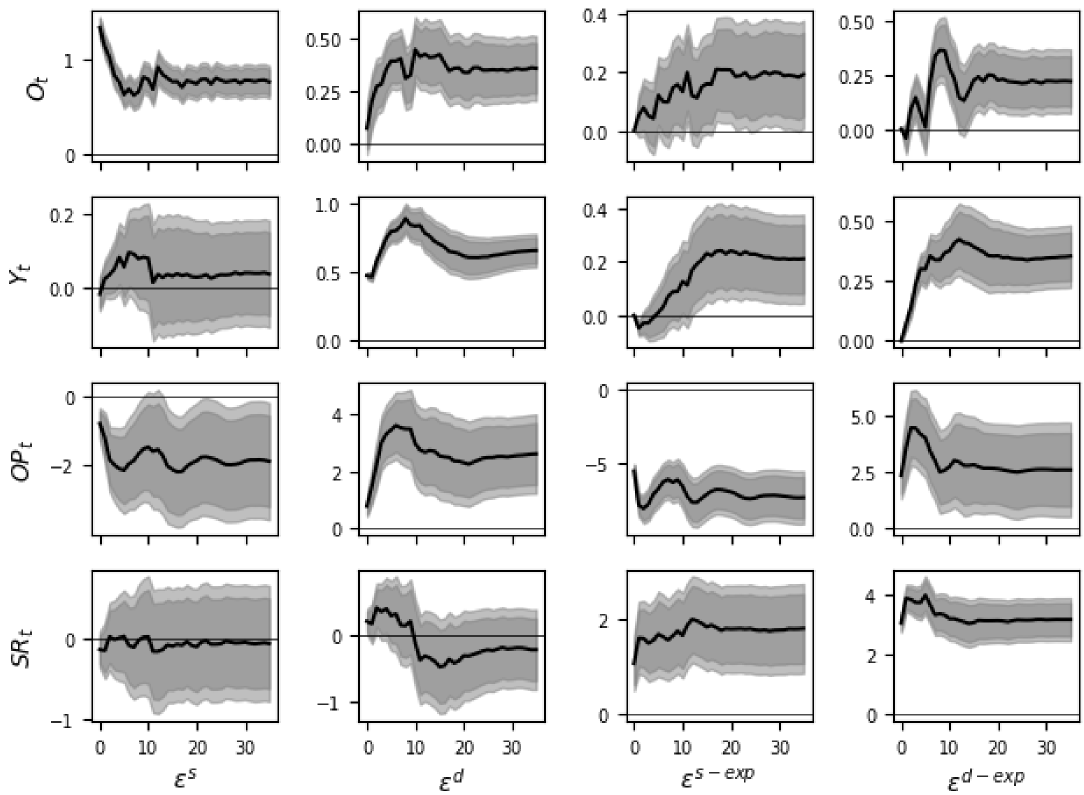
Figure 4.12: Oil market SVAR - Impulse responses (unrestricted SVAR with oGMM).



Note: Impulse responses to the estimated structural shocks for the unrestricted oil market SVAR from Section 4.6 estimated with the overidentified unrestricted SVAR GMM estimator from Equation (4.19) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows show the cumulative responses.

Figure 4.13 shows the IRF for the block-recursive oil market SVAR from Section 4.6 estimated with the block-recursive SVAR GMM estimator from Equation (4.15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks.

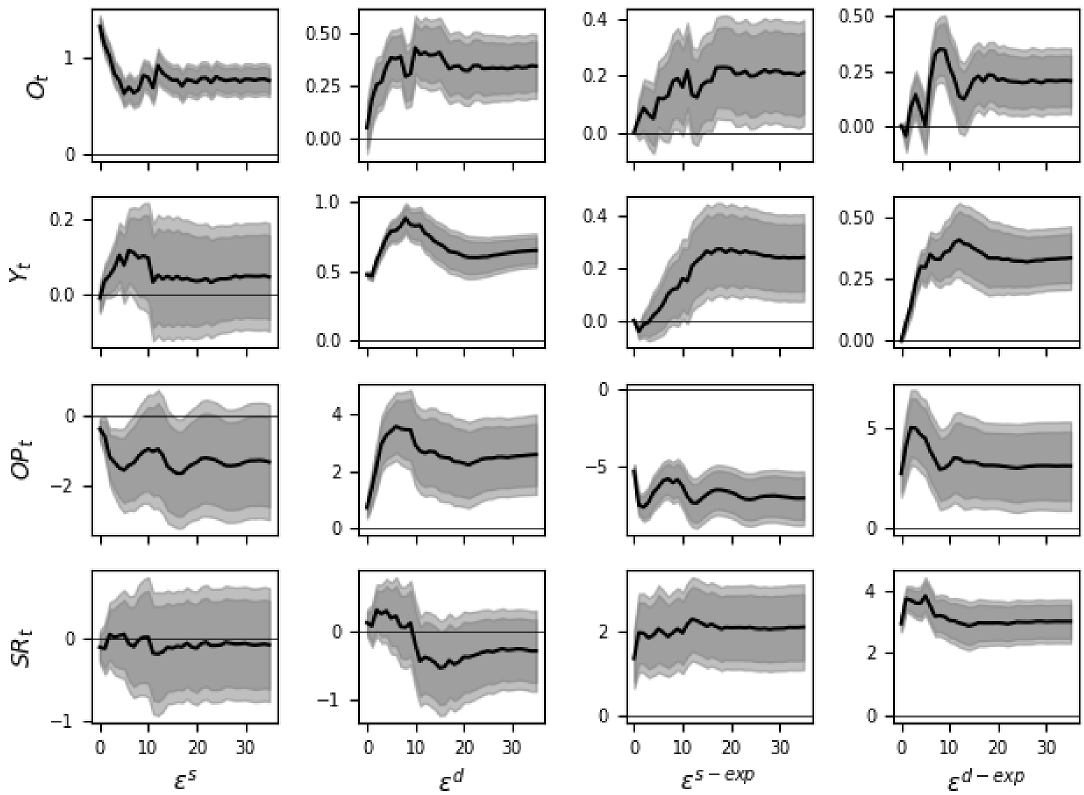
Figure 4.13: Oil market SVAR - Impulse responses (block-recursive SVAR with GMM).



Note: Impulse responses to the estimated structural shocks for the block-recursive oil market SVAR from Section 4.6 estimated with the block-recursive SVAR GMM estimator from Equation (4.15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows show the cumulative responses.

Figure 4.14 shows the IRF for the block-recursive oil market SVAR from Section 4.6 estimated with the overidentified block-recursive SVAR GMM estimator from Equation (4.19) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks.

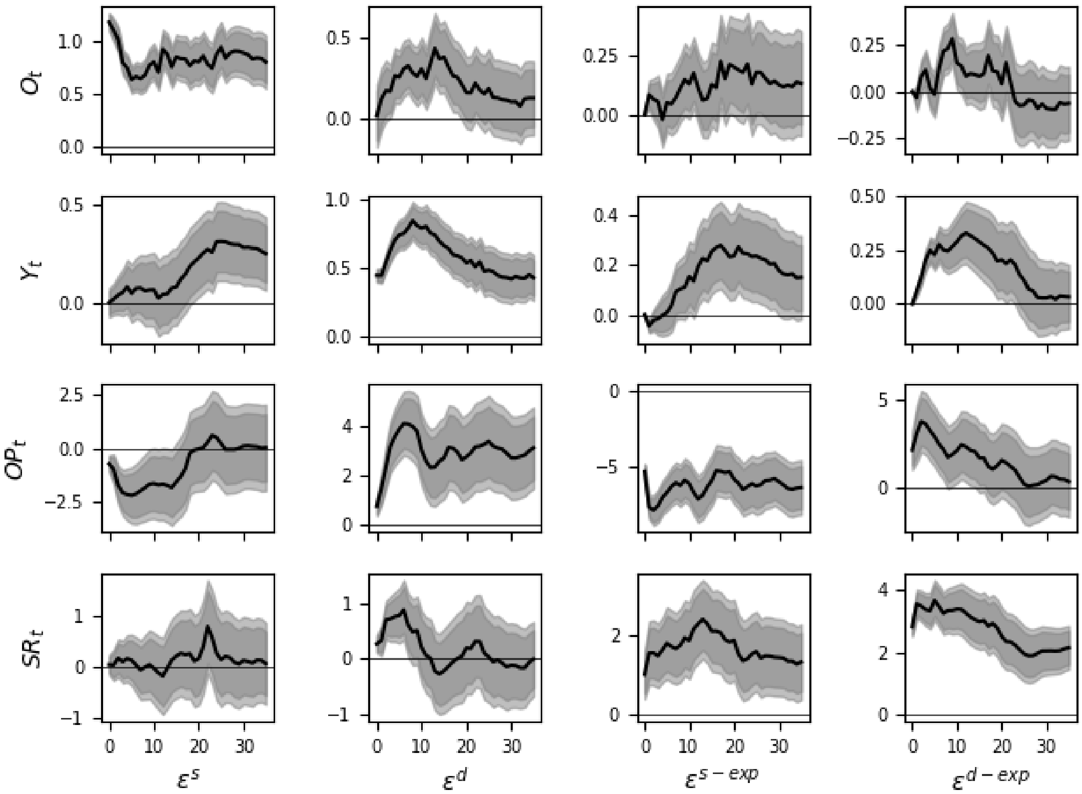
Figure 4.14: Oil market SVAR - Impulse responses (block-recursive SVAR with oGMM).



Note: Impulse responses to the estimated structural shocks for the block-recursive oil market SVAR from Section 4.6 estimated with the overidentified block-recursive SVAR GMM estimator from Equation (4.19) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows show the cumulative responses.

Figure 4.15 shows the IRF for the block-recursive oil market SVAR from Section 4.6 using 24 lags estimated with the block-recursive SVAR GMM estimator from Equation (4.15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks.

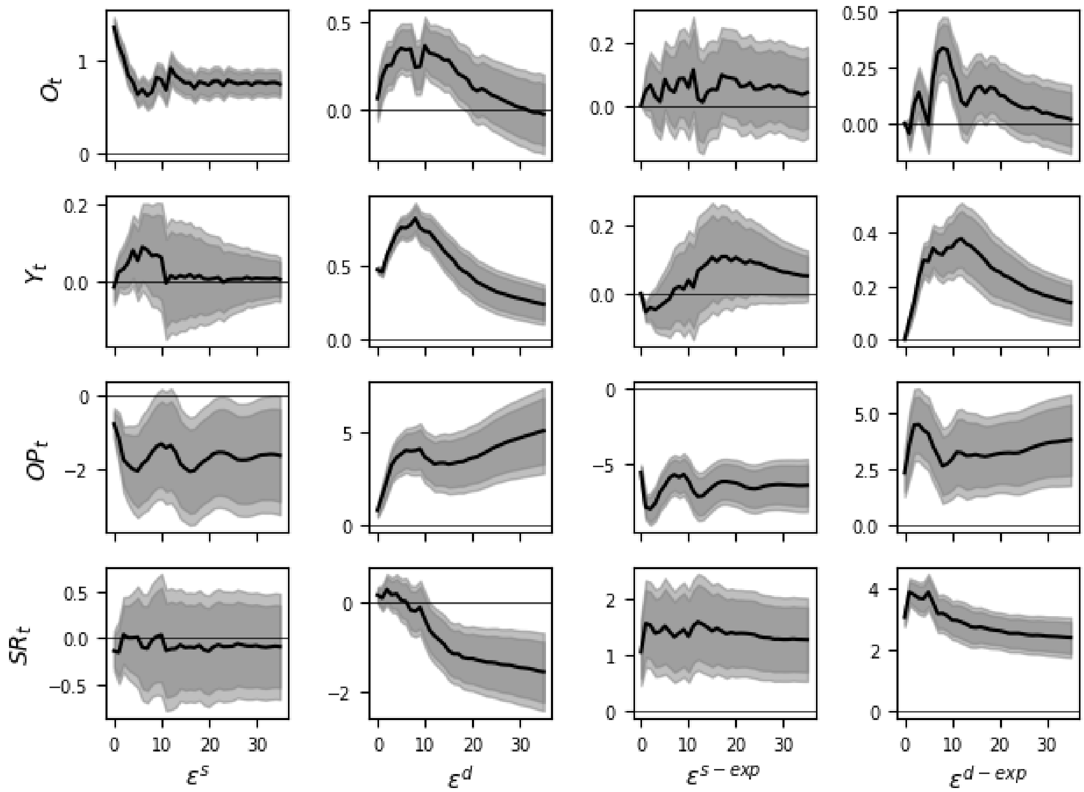
Figure 4.15: Oil market SVAR - Impulse responses (using 24 instead of 12 lags).



Note: Impulse responses to the estimated structural shocks for the block-recursive oil market SVAR from Section 4.6 24 estimated with the block-recursive SVAR GMM estimator from Equation (4.15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows show the cumulative responses.

Figure 4.16 shows the IRF for the block-recursive oil market SVAR from Section 4.6 using the percentage deviation of industrial production from a linear trend instead of the log difference of industrial production. The SVAR is estimated with the block-recursive SVAR GMM estimator from Equation (4.15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks.

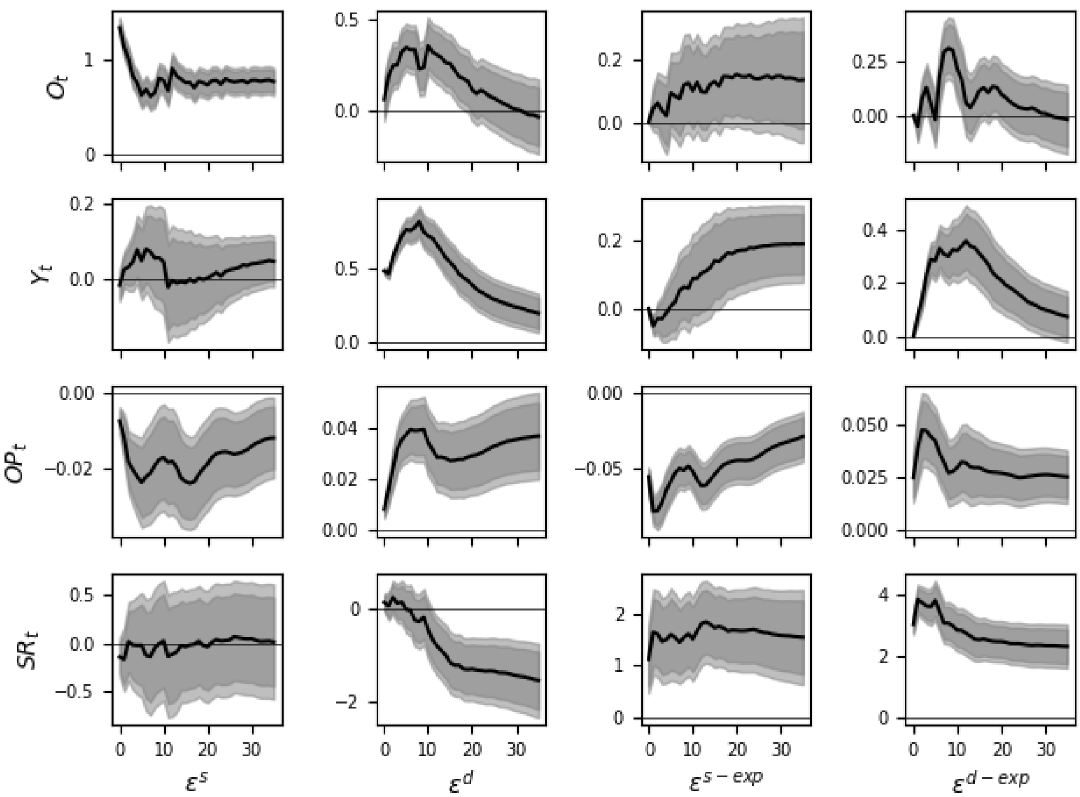
Figure 4.16: Oil market SVAR - Impulse responses (using the percentage deviation of industrial production from a linear trend).



Note: Impulse responses to the estimated structural shocks for the block-recursive oil market SVAR from Section 4.6 using the percentage deviation of industrial production from a linear trend instead of the log difference of industrial production. The SVAR is estimated with the block-recursive SVAR GMM estimator from Equation (4.15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows O_t, OP_t, and SR_t show the cumulative responses.

Figure 4.17 shows the IRF for the block-recursive oil market SVAR from Section 4.6 using log of real oil price instead of real oil price growth and the percentage deviation of industrial production from a linear trend instead of the log difference of industrial production. The SVAR is estimated with the block-recursive SVAR GMM estimator from Equation (4.15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks.

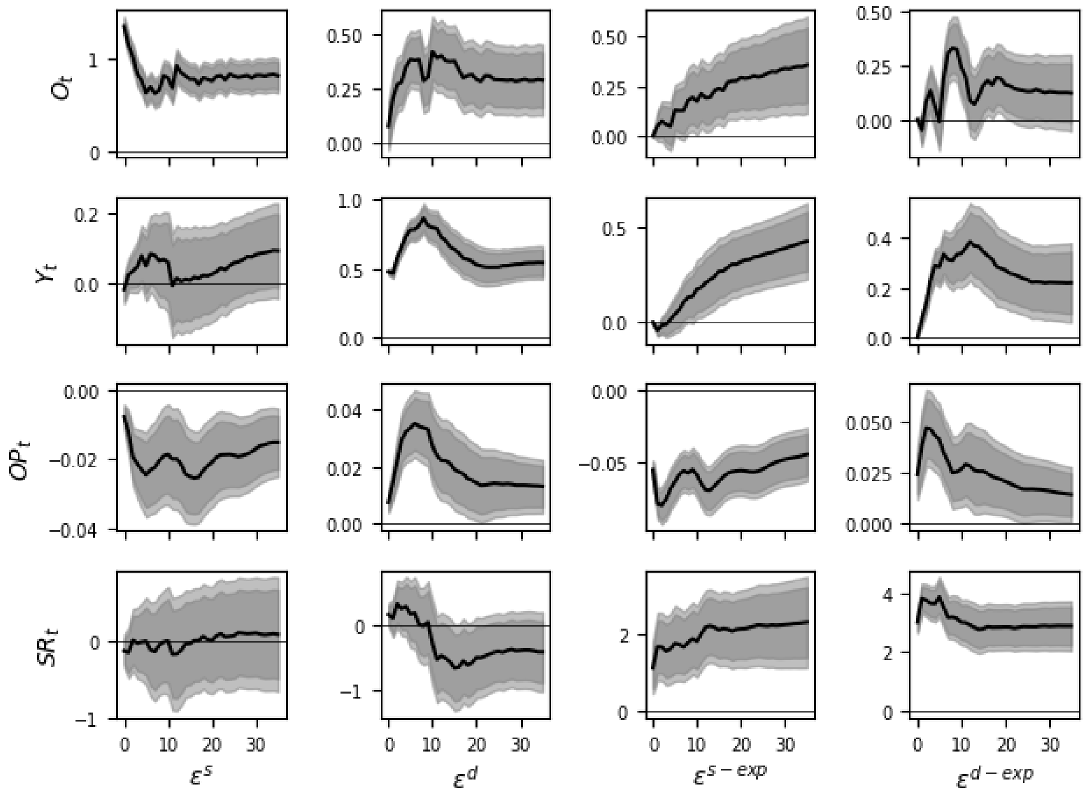
Figure 4.17: Oil market SVAR - Impulse responses (using the percentage deviation of industrial production from a linear trend and the log of the real oil price).



Note: Impulse responses to the estimated structural shocks for the block-recursive oil market SVAR from Section 4.6 using the log of the real oil price instead of real oil price growth and the percentage deviation of industrial production from a linear trend instead of the log difference of industrial production. The SVAR is estimated with the block-recursive SVAR GMM estimator from Equation (4.15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows O_t and SR_t show the cumulative responses.

Figure 4.18 shows the IRF for the block-recursive oil market SVAR from Section 4.6 using log of real oil price instead of real oil price growth. The SVAR is estimated with the block-recursive SVAR GMM estimator from Equation (4.15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks.

Figure 4.18: Oil market SVAR - Impulse responses (using the log of the real oil price).



Note: Impulse responses to the estimated structural shocks for the block-recursive oil market SVAR from Section 4.6 using the log of the real oil price instead of real oil price growth. The SVAR is estimated with the block-recursive SVAR GMM estimator from Equation (4.15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows O_t, Y_t, and SR_t show the cumulative responses.

References

Andrews, D. W. (1999). Consistent Moment Selection Procedures for Generalized Method of Moments Estimation. *Econometrica 67*(3), 543–563.

Anttonen, J., M. Lanne, and J. Luoto (2021). Statistically Identified SVAR Model with Potentially Skewed and Fat-Tailed Errors. *Available at SSRN 3925575*.

Apergis, N. and S. M. Miller (2009). Do Structural Oil-Market Shocks Affect Stock Prices? *Energy Economics 31*(4), 569–575.

Baumeister, C. and J. D. Hamilton (2019). Structural Interpretation of Vector Autoregressions with Incomplete Identification: Revisiting the Role of Oil Supply and Demand Shocks. *American Economic Review 109*(5), 1873–1910.

Baumeister, C. and L. Kilian (2016). Understanding the Decline in the Price of Oil since June 2014. *Journal of the Association of Environmental and resource economists 3*(1), 131–158.

Beaudry, P. and F. Portier (2006). Stock Prices, News, and Economic Fluctuations. *American Economic Review 96*(4), 1293–1307.

Bianchi, F., H. Kung, and G. Morales (2019). Growth, Slowdowns and Recoveries. *Journal of Monetary Economics 101*, 47–63.

Bjørnland, H. C. and K. Leitemo (2009). Identifying the Interdependence between US Monetary Policy and the Stock Market. *Journal of Monetary Economics 56*(2), 275–282.

Blanchard, O. J. (1989). A Traditional Interpretation of Macroeconomic Fluctuations. *The American Economic Review*, 1146–1164.

Blanchard, O. J. and D. Quah (1993). The Dynamic Effects of Aggregate Demand and Supply Disturbances: Reply. *The American Economic Review 83*(3), 653–658.

Bonhomme, S. and J.-M. Robin (2009). Consistent Noisy Independent Component Analysis. *Journal of Econometrics 149*(1), 12–25.

Breusch, T., H. Qian, P. Schmidt, and D. Wyhowski (1999). Redundancy of Moment Conditions. *Journal of Econometrics 91*(1), 89–111.

Burnside, C. and M. Eichenbaum (1996). Small-Sample Properties of GMM-Based Wald Tests. *Journal of Business & Economic Statistics 14*(3), 294–308.

Byrne, J. P., M. Lorusso, and B. Xu (2019). Oil Prices, Fundamentals and Expectations. *Energy Economics 79*, 59–75.

Carvalho, C., A. Ferrero, and F. Nechio (2016). Demographics and Real Interest Rates: Inspecting the Mechanism. *European Economic Review 88*, 208–226.

Cheng, X. and Z. Liao (2015). Select the Valid and Relevant Moments: An Information-Based LASSO for GMM with Many Moments. *Journal of Econometrics 186*(2), 443–464.

Comon, P. (1994). Independent Component Analysis, a New Concept? *Signal processing 36*(3), 287–314.

Defferrard, M., R. Pena, and N. Perraudin (2017, Dec). PyUNLocBoX: Optimization by Proximal Splitting.

Donald, S. G. and W. K. Newey (2000). A Jacknife Interpretation of the Continuous Updating Estimator. *Economics Letters 67*(3), 239–243.

Gouriéroux, C., A. Monfort, and J.-P. Renne (2017). Statistical Inference for Independent Component Analysis: Application to Structural VAR Models. *Journal of Econometrics 196*(1), 111–126.

Guay, A. (2021). Identification of Structural Vector Autoregressions Through Higher Unconditional Moments. *Journal of Econometrics 225*(1), 27–46.

Hall, A. R. (2005). *Generalized Method of Moments.* Oxford University Press.

Hall, A. R. (2015). Econometricians Have Their Moments: GMM at 32. *Economic Record 91*, 1–24.

Hall, A. R., A. Inoue, K. Jana, and C. Shin (2007). Information in Generalized Method of Moments Estimation and Entropy-Based Moment Selection. *Journal of Econometrics 138*(2), 488–512.

Han, C. and P. C. Phillips (2006). GMM with Many Moment Conditions. *Econometrica 74*(1), 147–192.

Hansen, L. P., J. Heaton, and A. Yaron (1996). Finite-Sample Properties of Some Alternative GMM Estimators. *Journal of Business & Economic Statistics 14*(3), 262–280.

Herwartz, H. (2018). Hodges–Lehmann Detection of Structural Shocks – An Analysis of Macroeconomic Dynamics in the Euro Area. *Oxford Bulletin of Economics and Statistics*.

Herwartz, H. and M. Plödt (2016). The Macroeconomic Effects of Oil Price Shocks: Evidence from a Statistical Identification Approach. *Journal of International Money and Finance 61*, 30–44.

Hyvärinen, A., K. Zhang, S. Shimizu, and P. O. Hoyer (2010). Estimation of a Structural Vector Autoregression Model Using Non-Gaussianity. *Journal of Machine Learning Research 11*(May), 1709–1731.

Jordà, Ò., S. R. Singh, and A. M. Taylor (2022). Longer-run economic consequences of pandemics. *Review of Economics and Statistics 104*(1), 166–175.

Juvenal, L. and I. Petrella (2015). Speculation in the Oil Market. *Journal of Applied Econometrics 30*(4), 621–649.

Keweloh, S. A. (2021a). A Feasible Approach to Incorporate Information in Higher Moments in Structural Vector Autoregressions. *Discussion Papers SFB 823*.

Keweloh, S. A. (2021b). A Generalized Method of Moments Estimator for Structural Vector Autoregressions Based on Higher Moments. *Journal of Business & Economic Statistics 39*(3), 772–782.

Keweloh, S. A. and A. Seepe (2020). Monetary Policy and the Stock Market - A Partly Recursive SVAR Estimator. *Discussion Papers SFB 823*.

Kilian, L. (2009). Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market. *American Economic Review 99*(3), 1053–69.

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

Kilian, L. and D. P. Murphy (2014). The Role of Inventories and Speculative Trading in the Global Market for Crude Oil. *Journal of Applied econometrics 29*(3), 454–478.

Kilian, L. and C. Park (2009). The Impact of Oil Price Shocks on the US Stock Market. *International Economic Review 50*(4), 1267–1287.

Kontonikas, A. and Z. Zekaite (2018). Monetary Policy and Stock Valuation: Structural VAR Identification and Size Effects. *Quantitative Finance 18*(5), 837–848.

Lanne, M., K. Liu, and J. Luoto (2021). Identifying Structural Vector Autoregression via Large Economic Shocks. *Available at SSRN 3910532*.

Lanne, M. and J. Luoto (2020). Identification of Economic Shocks by Inequality Constraints in Bayesian Structural Vector Autoregression. *Oxford Bulletin of Economics and Statistics 82*(2), 425–452.

Lanne, M. and J. Luoto (2021). GMM Estimation of Non-Gaussian Structural Vector Autoregression. *Journal of Business & Economic Statistics 39*(1), 69–81.

Lanne, M., H. Lütkepohl, and K. Maciejowska (2010). Structural Vector Autoregressions with Markov Switching. *Journal of Economic Dynamics and Control 34*(2), 121–131.

Lanne, M., M. Meitz, and P. Saikkonen (2017). Identification and Estimation of Non-Gaussian Structural Vector Autoregressions. *Journal of Econometrics 196*(2), 288–304.

Laopodis, N. T. (2013). Monetary Policy and Stock Market Dynamics Across Monetary Regimes. *Journal of International Money and Finance 33*, 381–406.

Lewis, D. J. (2021). Identifying Shocks via Time-Varying Volatility. *The Review of Economic Studies 88*(6), 3086–3124.

Lezcano-Casado, M. and D. Martınez-Rubio (2019). Cheap Orthogonal Constraints in Neural Networks: A Simple Parametrization of the Orthogonal and Unitary Group. In *International Conference on Machine Learning*, pp. 3794–3803. PMLR.

Luenberger, D. G., Y. Ye, et al. (2016). *Linear and Nonlinear Programming*, Volume 4. Springer.

Lütkepohl, H. and A. Netšunajev (2017). Structural Vector Autoregressions with Smooth Transition in Variances. *Journal of Economic Dynamics and Control 84*, 43–57.

Mertens, K. and M. O. Ravn (2013). The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States. *American Economic Review 103*(4), 1212–47.

Moneta, A., D. Entner, P. O. Hoyer, and A. Coad (2013). Causal Inference by Independent Component Analysis: Theory and Applications. *Oxford Bulletin of Economics and Statistics 75*(5), 705–730.

Moran, P. and A. Queralto (2018). Innovation, Productivity, and Monetary Policy. *Journal of Monetary Economics 93*, 24–41.

Moussa, Z. and A. Thomas (2021). Identifying Oil Supply News Shocks and their Effects on the Global Oil Market. *USAEE Working Paper No. 21-490*.

Newey, W. K. and D. McFadden (1994). Large Sample Estimation and Hypothesis Testing. *Handbook of Econometrics 4*, 2111–2245.

Newey, W. K. and K. D. West (1994). Automatic Lag Selection in Covariance Matrix Estimation. *The Review of Economic Studies 61*(4), 631–653.

Newey, W. K. and F. Windmeijer (2009). Generalized Method of Moments with Many Weak Moment Conditions. *Econometrica 77*(3), 687–719.

Olea, M., J. Luis, M. Plagborg-Møller, and E. Qian (2022). SVAR Identification From Higher Moments: Has the Simultaneous Causality Problem Been Solved? *Working Paper*. Prepared for AEA Papers and Proceedings.

Peersman, G. (2005). What Caused the Early Millennium Slowdown? Evidence Based on Vector Autoregressions. *Journal of Applied Econometrics 20*(2), 185–207.

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

Rigobon, R. and B. Sack (2003). Measuring the Reaction of Monetary Policy to the Stock Market. *The Quarterly Journal of Economics 118*(2), 639–669.

Rigobon, R. and B. Sack (2004). The Impact of Monetary Policy on Asset Prices. *Journal of Monetary Economics 51*(8), 1553–1575.

Rubio-Ramírez, J. F., D. F. Waggoner, and T. Zha (2010). Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference. *The Review of Economic Studies 77*(2), 665–696.

Sadorsky, P. (1999). Oil Price Shocks and Stock Market Activity. *Energy Economics 21*(5), 449–469.

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

Smets, F. and R. Wouters (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review 97*(3), 586–606.

Stock, J. H. and M. Watson (2012). Disentangling the Channels of the 2007-09 Recession. *Brookings Papers on Economic Activity 43*(1), 81–156.

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.

Wainwright, M. J. (2009). Sharp Thresholds for High-Dimensional and Noisy Sparsity Recovery Using ℓ_1-Constrained Quadratic Programming (Lasso). *IEEE transactions on information theory 55*(5), 2183–2202.

Wu, J. C. and F. D. Xia (2016). Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound. *Journal of Money, Credit and Banking 48*(2-3), 253–291.

Zha, T. (1999). Block Recursion and Structural Vector Autoregressions. *Journal of Econometrics 90*(2), 291–316.

Zou, H. (2006). The Adaptive Lasso and Its Oracle Properties. *Journal of the American Statistical Association 101*(476), 1418–1429.