

FINANCIAL ASSETS, FISCAL POLICY, AND THE MACROECONOMY

Nils Wittmann*

July 31, 2016

Dissertation zur Erlangung des akademischen Grades Doktor rerum politicarum
der Technischen Universität Dortmund.

*TU Dortmund und Ruhr Graduate School in Economics. Technische Universität Dortmund,
Wirtschafts- und Sozialwissenschaftliche Fakultät, Applied Economics. Vogelpothsweg 87, D-44227
Dortmund. nils@wittmann.us

Contents

1	Introduction	10
2	Abnormal stock returns of military firms and the fiscal multiplier	16
2.1	Introduction	16
2.2	Time series evidence for fiscal multipliers	18
2.3	Stock returns & data definitions	22
2.3.1	Anticipation shocks & excess return	22
2.3.2	Data definitions and sources	29
2.4	Estimation results	30
2.4.1	Identification and estimation	30
2.4.2	Estimation results from abnormal and excess return innovations	31
2.5	Robustness of results & limitations	35
2.5.1	Robustness I: extending the sample	35
2.5.2	Robustness II: five-factor and momentum-model	38
2.5.3	Limitations and discussion	40
2.6	Conclusion	41
2.7	Appendix	43
3	The procyclicality of consumer credit	53
3.1	Introduction	53
3.2	Data, Granger causality and estimation	55
3.2.1	Data	55
3.2.2	Identification, choice of model and estimation	56
3.3	Estimation results	59
3.3.1	Robustness	62
3.4	Discussion of the results	63
3.5	Conclusion	64
3.6	Appendix	66
4	Technology shocks, tax cuts and their impact on household debt: Empirical evidence & theoretical explanation	73
4.1	Introduction	73
4.2	Empirical evidence	75
4.2.1	Data and identification	75
4.2.2	VAR results	78
4.3	Model	80
4.3.1	Households	80

4.3.2	Firms	84
4.3.3	Government	86
4.3.4	Aggregation and market clearing	86
4.3.5	Equilibrium	87
4.4	Parametrization	87
4.4.1	Calibration	87
4.4.2	Estimation	88
4.5	Results	89
4.6	Conclusion	92
4.7	Appendix	94
5	Forced deleveraging, fiscal policy rules and the zero lower bound	101
5.1	Introduction	101
5.2	House prices and devaluation shocks in literature	102
5.3	Empirical evidence for house price shocks	104
5.4	Model	106
5.4.1	Households	106
5.4.2	Aggregate price level and price dispersion	111
5.4.3	The public sector	111
5.4.4	Equilibrium conditions	113
5.5	Calibration and results	113
5.5.1	Calibration	113
5.5.2	Simulation results	116
5.5.3	Binding constraints & the ZLB	120
5.5.4	Valuation shocks at the ZLB	121
5.5.5	Discussion and limitations	127
5.6	Conclusion	127
5.7	Appendix	129
6	Market inefficiencies and forecastability of spot rates in the shipping sector	134
6.1	Introduction	134
6.2	Market properties in the shipping sector	135
6.2.1	Forward freight rates & spot rates	137
6.2.2	The timing structure of freight rates	138
6.2.3	The sign of risk premia	139
6.3	Data and estimation	140
6.3.1	Data	140

6.4	Estimation methods	142
6.4.1	Descriptive data & hypotheses	145
6.5	Estimation results	146
6.5.1	ARIMA estimation	146
6.5.2	VAR estimation	146
6.5.3	VECM estimation	147
6.5.4	ARIMA-TC estimation	148
6.5.5	Impulse responses	148
6.5.6	Forecasts and RMSE	149
6.6	Robustness and trading scheme	153
6.6.1	Trading scheme	154
6.7	Conclusion	158
6.8	Appendix	160
7	Conclusion	178
8	References	180

List of Figures

1	Defense and total fiscal spending	20
2	Correctly measured excess return (case 1)	24
3	Overestimated excess return (case 2).	24
4	Underestimated excess return (case 3)	25
5	Abnormal returns and Fisher and Peters (2010) returns	28
6	IR for abnormal and excess return shocks I	32
7	IR for abnormal and excess return shocks II	34
8	IR baseline and extended sample I	36
9	IR baseline and extended sample II	38
10	Defense and non-defense spending	46
11	IR t-bill in levels I	47
12	IR t-bill in levels II	48
13	Cumulative IR for abnormal and excess returns I	49
14	Cumulative IR for abnormal and excess returns II	50
15	Ramey news and abnormal returns	51
16	Factor models as proxy for future military spending	52
17	Unsecured consumer credit as a ratio to output	53
18	Impulse responses conditional on tax cuts	60
19	Impulse responses conditional on TFP shocks	61
20	IR for tax cuts, controlling for including wages	66
21	IR for TFP shocks, controlling for wages	67
22	IR for tax cuts, controlling for investment	68
23	TFP shock estimation including investment	69
24	IRs comparing lag lengths (Tax shock)	70
25	IRs comparing lag lengths (TFP shock)	71
26	Impulse responses SVAR estimation	79
27	Empirical and matched impulse responses	91
28	IR robustness monetary policy vs. baseline	95
29	IR robustness lag length 3 lags vs. baseline	96
30	IR robustness lag length 5 lags vs. baseline	97
31	IR robustness gov't debt vs. baseline	98
32	IR robustness quadratic trend vs. baseline	99
33	IR robustness unemployment vs. baseline	100
34	Case-Shiller Index	103
35	VAR evidence on house price shocks	105
36	Spending cuts and tax increases after devaluation shocks I	117

37	Spending cuts and tax increases and devaluation shocks II	118
38	IR tax increases after devaluation shocks, ZLB vs. no-ZLB I	122
39	IR tax increases after devaluation shocks, ZLB vs. no-ZLB II	123
40	IR spending cuts after devaluation shocks, ZLB vs. no-ZLB I	125
41	IR spending cuts after devaluation shocks, ZLB vs.no-ZLB II	126
42	Spending vs. tax stabilization of debt after valuation shocks	132
43	Spending vs. tax stabilization of debt after valuation shocks	133
44	Comovement spot and FFA contracts on panamax routes	141
45	One day ahead forecasts S_{t+1}	144
46	Two-day forecasts, using \hat{S}_{t+1}	144
47	Orthogonalized impulse responses VAR estimation	149
48	RMSE change when switching models	153
49	Trading scheme returns for both contracts	157
50	Trading scheme returns for both contracts incl. transaction costs . . .	158
51	Improvement of RMSE compared to random walk	175

List of Tables

1	Data sources & definitions	56
2	Results of Granger causality test	58
3	Full data sources & definitions	72
4	Data sources	76
5	Granger causality test results	77
6	Model calibration	88
7	Estimated model parameters	90
8	Full data sources	94
9	Calibration of parameters	114
10	Estimation of autoregressive parameters	116
11	Data sources for VAR estimation	129
12	Complete sources for parameters calibration	130
13	Details on the estimation of average taxes	131
14	RMSE Panamax	150
15	Reduction in RMSE in % for Capesize contract	151
16	Returns for Panamax and Capesize trading scheme	155
17	Sharpe ratios for trading scheme	156
18	Descriptive data Panamax contract	160
19	Descriptive data Capesize contract	160
20	Cointegration in Panamax market	161
21	Cointegration in Capesize market	161
22	ARIMA(2,1,0) estimation, Capesize market	161
23	ARIMA(2,1,0) estimation, Panamax market	162
24	VAR estimation for spot rate, Panamax (+1Q)	162
25	VAR estimation for FFA rate, Panamax (+1Q)	162
26	VAR estimation for spot rate, Panamax (+2Q)	163
27	VAR estimation for FFA rate, Panamax (+2Q)	163
28	VAR estimation for spot rate, Panamax (+1A)	164
29	VAR estimation results for FFA rate, Panamax (+1A)	164
30	VAR estimation results for spot rates, Capesize (+1Q)	165
31	VAR estimation for FFA, Capesize (+1Q)	165
32	VAR estimation for spot rates, Capesize (+2Q)	166
33	VAR estimation for FFA, Capesize (+2Q)	166
34	VAR estimation for spot rate, Capesize (+1A)	167
35	VAR estimation results for FFA, Capesize (+1A)	167
36	VECM estimation results, Panamax (+1Q)	168

37	VECM estimation results, Panamax (+2Q)	169
38	VECM estimation results, Panamax (+1A)	170
39	VECM estimation results, Capesize (+1Q)	171
40	VECM estimation results, Capesize (+2Q)	172
41	VECM estimation results, Capesize (+1A)	173
42	ARIMA TC, Panamax (+6M)	173
43	ARIMA TC, Panamax (+12M)	174
44	ARIMA TC, Capesize (+6M)	174
45	ARIMA TC, Capesize (+12M)	174

Acknowledgments

My appreciation goes to a number of people who have encouraged and guided me during the process of writing this thesis.

I would like to thank Prof. Dr. Ludger Linnemann for his support and guidance and his admittance into his team, as well as JProf. Dr. Roland Winkler and Prof. Dr. Wolfram Richter.

My gratitude is also extended to the Ruhr Graduate School in Economics, whose members have accepted me into their program and provided an excellent learning environment and financial support. The well-organized structural doctoral program together with its exceptional faculty have laid the foundation for my development.

Further, I would like to thank my co-authors, Christopher Krause and Mathias Klein, also from the Ruhr Graduate School in Economics, for productive collaboration during two projects, as well as Marcus Eppinger.

In addition, my gratitude is extended to researchers who have provided me with academic guidance. Among these are, Michael Stein, Matteo Iacoviello, Ryan Peters, Jonas Fisher, Jeffrey Sheen and Dominik Fuchs.

During the last years, I have attended conferences and workshops resulting in fruitful comments in Dortmund, London, Singapore, Cambridge, Barcelona, Malta, Thessaloníki and Luxembourg, with the financial assistance of the Ruhr Graduate School and the TU Dortmund, for which I am grateful.

Finally, and most importantly, I would like to express my gratitude to my parents Volker and Jutta as well as to my uncle Dietmar, who have always supported and encouraged me to pursue my study goals.

1 Introduction

In this thesis, I provide several essays in support to answering the question about the interaction of different fiscal instruments, macro aggregates, and financial assets.

To discover the impact of fiscal policy and other shocks on macro aggregates, I present five essays on financial assets, fiscal policy and the macroeconomy while estimating empirical models using U.S. data and building theoretical economies.

The second chapter of this thesis contributes to existing literature by estimating the fiscal spending multiplier using stock returns of military contractors in an expectation-augmented vector autoregression (EVAR) setup.

Two main issues need consideration when estimating the impact of fiscal spending on output and other aggregates. First, aggregated spending contains endogenous components, which are adjusted to local and state needs and also in reaction to structural changes, for instance, the demographic structure, and are therefore correlated with the state of the economy. Using such a combination of endogenous and exogenous spending would, therefore, dilute the structural estimate for the fiscal multiplier. To cope with this problem, I instead incorporate defense spending in my estimation, as suggested by Ramey and Shapiro (1998), Fisher and Peters (2010) and Ramey (2011b). This type of measure is likely to be exogenous while reacting to international military challenges rather than to local needs.

Second, agents in the economy have expectations on future fiscal policy. Forward-looking rational agents learn of changes to the fiscal budget beforehand and adjust their economic choices before actual changes appear in the data. Therefore, datasets of the researcher and the one of the agents do not align, which biases the estimate of the impact of fiscal spending shocks on output. Different solutions to this problem have been suggested. Ramey (2011a) creates a narrative series of discounted changes to military spending, whereas Fisher and Peters (2010) use stock portfolio returns of military-focused firms, exceeding the market return.

To address the foresight problem, I follow Fisher and Peters (2010). The authors suggest portfolio returns of military contractors as a proxy for future defense spending increases since financial markets react instantaneously to news in the economy causing higher stock returns of receivers of additional funds. However, the way Fisher and Peters (2010) capture this shock has drawbacks, as they only use the difference of military firms stock returns and the market return, ignoring certain dynamics.

Instead of calculating portfolio returns in excess of the market return, I estimate a measure for abnormal and hence unexpected returns. In a Sharpe (1964), Lintner (1965) and Black (1972) (SLB) type market model regression, I define an anticipated

fiscal spending shock as the residual of the regression, as also suggested by DellaVigna and La Ferrara (2010) for detecting illegal arms trades. Furthermore, Fama and French (1993) and Fama and French (2012) suggest that persistent covariates exist that help to explain stock market returns. Fama and French (1993) suggestions have become known as the “three-factor-model” and include a measure for the size and book-to-market value of firms. Ignoring these factors may overestimate the impact of fiscal spending on output.

When ordering military spending first in an EVAR estimation and incorporating the refined anticipation shock series, the spending multiplier takes a value of 1.2. Therefore, Fisher and Peters (2010), who suggest a multiplier of 1.5, tend to overestimate the true multiplier by relying on their stock market measure.

Extending the sample to include financial crises data yields a decrease in importance of the anticipated military spending shock since the last observations add a period of non-defense spending changes.

The third chapter continues to investigate the impact of fiscal policy on main aggregates and financial measures. In contrast to the paper on fiscal spending shocks and the fiscal multiplier, chapter 3 takes a look at the other side of fiscal policy, taxes. In “The procyclicality of consumer credit” the interaction between tax cuts, TFP shocks and consumer credit is investigated. The former shock is incorporated since it is of importance for the policy-maker, whereas TFP shocks are seen as one of the major drivers of the business cycle.

Recently, consumer credit has more than doubled, also because the Financial Liberalization has led to an easier access to unsecured credit. So far, both shocks were mainly used to quantify their dynamic effects on variables like output, consumption or hours,¹ whereas this paper takes a closer look at how private credit evolves to conditional changes in both measures.

In order to proceed, the exogeneity of the Romer and Romer (2010) tax measure and the Basu et al. (2006) TFP shock is investigated since this is of importance for the choice of the estimation method. Granger causality tests suggest that both shocks can partly be predicted by past observations of the other variables included in the VARs. This paper, therefore, compares impulse responses from exogenous VARs with those estimated with VARs for quarterly U.S. data. The paper suggests that due to the results of the Granger causality tests, the less restrictive method, a VAR, should be estimated when including either shock measure. In such a setup, the shocks are contemporaneously exogenous, rather than strictly exogenous as in

¹For the TFP shock, see, among others, Basu et al. (2006) and Christiano et al. (2004) and for the tax shock, some prominent examples are Romer and Romer (2010), Mertens and Ravn (2011) and Favero and Giavazzi (2012).

the case of the exogenous VAR. Results are then compared to identify the bias from estimating exogenous VARs.

The variables in the estimation include either shock measure, output, non-durable and durable private consumption, hours, the real interest rate and consumer credit. The fiscal policy shock in the form of tax cuts and the TFP innovation create a persistent expansion in the economy. The findings are in line with Ludvigson (1999), who provides descriptive evidence for the TFP shock, and Nakajima and Ríos-Rull (2014), who gives unconditional correlations between consumer credit and output. This paper, however, provides conditional evidence that credit, consumption and output co-move based on structural VAR estimates. As a result, the boom that is initiated in the economy is partly credit-financed. Therefore, in good times, consumers choose to borrow and consume more rather than to save and buffer future adverse shocks.

The bias arising when estimating exogenous VARs is rather small for tax cuts compared to TFP shocks. When using the TFP measure, the exogenous VAR significantly underestimates the impact on the other variables. Responses lie outside the confidence bands so that they statistically differ.

Since there is an empirically documented positive comovement of unsecured consumer credit with other aggregates after tax cuts and TFP shocks, this analysis is extended to investigate the impact of these two shocks on total private debt, which also includes collateralized loans.

In chapter 4, it is shown, how total private debt reacts to tax cuts and TFP shocks. Since both shocks are contemporaneously exogenous and Granger-caused by the other variables, this paper estimates the impact of the fiscal and TFP shock also in a structural VAR approach. We find that tax cuts and innovations in TFP, induce an expansion in the economy, which causes output, durable and non-durable consumption to increase in a humped-shaped manner. Interestingly, it is found that total private debt increases as well, which refutes traditional consumption smoothing arguments, as is observed for unsecured credit as well.

The empirically observed comovement is robust to a number of modifications. Therefore, a theoretical DSGE model is built and brought to the data. The model economy is closely related to the ones in Iacoviello (2005) and Monacelli (2009). It is populated by two types of household who differ in their willingness to postpone consumption into the future, creating lenders and borrowers. Both consume a basket of non-durable and durable consumption goods. In contrast to the lender, the borrowing capacity of the impatient households is limited to a fraction of their non-depreciated stock of durables.

To match theoretical impulse responses with the empirical data, deep model parameters are estimated following Cogley and Nason (1995) and Mertens and Ravn (2013), estimating VARs with actual and model generated data. The results suggest that the model can successfully account for the sizes and the hump-shaped patterns of the empirical dynamics in all variables. In line with the empirical findings, the model produces persistent increases in total private debt, which last for more than 20 quarters. Moreover, the debt responses almost perfectly match the empirical counterparts. The point estimates of deep model parameters are in line with findings in previous studies (Mertens and Ravn, 2013; Iacoviello, 2005).

Because the model finds that a part of the population is borrowing-constrained and hence faces a limit on the ability to borrow, I investigate the impact of shocks to the fraction of collateral that borrowers can use to accumulate new debt as devaluation shocks. Such a shock can be used as a proxy for house price shocks to model the sharp decline in prices at the beginning of the financial crisis.

In chapter 5, I, therefore, investigate the effects of devaluation shocks on lenders, borrowers, and especially on output. I extend the housing DSGE model of Iacoviello (2005) by a fiscal sector and investigate the interaction of fiscal policy and devaluation shocks.

The model economy is similar to the one in chapter 4 and comprises of an extended fiscal sector. The government levies distortionary taxes on borrowers and lenders and consumes a fraction of total output. In contrast to the former model, the economy includes entrepreneurs, who produce intermediate goods by combining labor and capital inputs with housing services.

I model the starting point of the financial crises as a devaluation shock to households' ability to borrow, as suggested by Eggertsson and Krugman (2012) and Iacoviello and Guerrieri (2015). As a consequence, indebted households are forced to deleverage quickly and also to sell their property, which causes house prices to decrease. Since borrowing households feel poorer, they reduce consumption, which causes a recession, while entrepreneurs are hit the hardest due to real estate in their production function.

Governments around the world responded to the financial crises with large stimulus packages which caused government debt to surge, while the nominal interest rate hit zero in the U.S. In response, policy makers suggested reducing public debt. In my model economy, the fiscal authority has two instruments to do so: cut spending or increase taxes. I demonstrate that the choice of the fiscal rule does not matter in times when the economy is sufficiently far away from the Zero Lower Bound on interest rates (ZLB). However, once at the ZLB, the monetary authority is incapable

of reducing the nominal interest rate to help stimulate the economy. I can show that the ZLB greatly amplifies the effects of valuation shocks. Therefore, it is revealed that cutting government spending at the ZLB has a more negative impact on output, as the hands of the monetary authority are tied. Cutting spending compared to increasing taxes leads to a four to five-fold negative effect compared to when the economy is distant to the ZLB.

In conclusion, the devaluation shock has asymmetric effects. Lenders profit from lower house prices while borrowers are severely hit. In the current situation, the results in this model suggest that lowering spending to stabilize debt causes a more severe recession in the economy, compared to raising taxes.

As the financial crisis has shown, financial markets have a crucial impact on the macro economy, which is also found in chapter 2 and 5. In the former, I capture anticipated fiscal spending with stock market measures. However, the validity of such measure crucially depends on the liquidity and hence the efficiency of the market. In chapter 6 I investigate the interaction of financial assets with short and long maturity when markets are illiquid instead. Here, I take a look at the pre-crisis shipping sector asset market as a special case of illiquid markets with good data quality and data availability.

This paper then summarizes the current literature on financial assets in the shipping sector and investigates market efficiency. In contrast to the stock market measures in the second chapter, the market for shipping goods across oceans is characterized by thin trading and the inability to short-sell assets.

Markets are efficient if the best forecast for future spot rates is the current forward price so that the Unbiased Expectations Hypothesis (UEH) holds. Using cointegration tests, the UEH is rejected for shipping rates so that forward prices at maturity do not have to equal spot prices. Then, the term structure of shipping rates matters so that market participants face a risk premia.

In such illiquid markets, forecasts for future financial assets can be generated using the information embedded in forward prices. The paper, therefore, lets different forecasting models compete for their forecastability of future spot rates and compares forecast errors generated from random walks to those obtained from ARIMA processes, VARs and error correction models. It is found that forward rates can explain future spot rates, with only a weak vice versa relationship. Finally, the model with the lowest forecast error (VAR) is used to generate a trading scheme which outperforms the market, even after controlling for transactions costs.

The remainder of this thesis is structured as follows: Chapter 2 gives an estimate for the fiscal multiplier based on stock market measures. In chapter 3 (4) the impact

of tax cuts and TFP shocks on consumer credit (total household debt) is investigated. Chapter 5 shows that cutting spending to stabilize debt at the ZLB can potentially create larger output gap compared to increasing taxes. Then, chapter 6 uses forward shipping rates to forecast spot rates. The last chapter concludes this thesis.

2 Abnormal stock returns of military firms and the fiscal multiplier

2.1 Introduction

Economic literature remains divided about the quantitative and qualitative impact of fiscal spending on macro aggregates and about the size of the fiscal multiplier. However, estimates of the fiscal multiplier are crucial for policy makers when deciding on the fiscal spending budget.

The main contribution of this paper is to find a correctly specified proxy for fiscal spending shocks, based on a stock market measure as originally suggested by Fisher and Peters (2010), to estimate the fiscal multiplier for the U.S.

I extend recent attempts to identify fiscal spending shocks by statistical innovations in defense contractors' stock returns, controlling for firm size and leverage effects in a Fama and French (1993) type of setup. The estimated cumulative multiplier estimated from anticipated spending shocks takes a value of 1.2.

Empirical research provides evidence for a broad range of fiscal multipliers, depending on the dataset, identification strategies, the type of government spending and its persistence (Ramey, 2012). Identification of fiscal spending shocks is necessary, which can be done by Cholesky decomposition to the residual variance-covariance matrix with spending ordered first as suggested by Blanchard and Perotti (2002) and followed by Fatas and Mihov (2001) and Perotti (2002).

However, when estimating fiscal multipliers, two issues must be considered. The endogeneity of government spending and fiscal foresight.

First, a major part of public expenditure decisions are endogenously adapted to local and state needs and likely to bias estimates for the fiscal multiplier (Pappa, 2007). Ramey and Shapiro (1998) and Ramey (2011a) propose the usage of defense spending rather than total expenditures since it is adjusted to meet international conflicts that are exogenous.

Second, Ramey (2011a), Leeper et al. (2012) and Leeper et al. (2013), show that forward-looking agents react quickly to announcements of policy changes by adjusting their economic choices before actual budget changes take place, causing movements in consumption, investment, and output to be earlier than documented by the data. Such misalignment of researchers' and agents' information set causes

fiscal multipliers to be biased, and structural spending shocks may not be recoverable (Perotti, 2011; Ramey, 2011a; Leeper et al., 2013). Different solutions to the foresight issue have been suggested. Ramey and Shapiro (1998) dummy out four war dates for which spending increases were anticipated. In contrast, Ramey (2011a) suggests amending the VAR estimation by a narrative series capturing expectations on future military spending, also called the EVAR approach.

A narrative measure as suggested by Ramey (2011a), however, has the drawback of being subjective since it is based on newspaper articles. In contrast, Fisher and Peters (2010) propose a proxy for anticipated military spending changes based on financial market data. A reaction of rational agents who learn of a possible military intervention would instantly appear in the stock returns of the likely beneficiary of additional military funds. Fisher and Peters (2010) define their proxy for anticipated military spending as the different between the returns of a portfolio consisting of stocks from companies specialized in producing military equipment and the market return, called excess returns.

These returns, however, ignore critical stock market reactions, for instance, the negative market reactions when political tensions increase as documented by Shapiro and Switzer (2011) and hence tend to overestimate anticipated spending increases. Also, excess returns implicitly assume that defense stocks react to market return movements one-for-one. In contrast, Fama and French (1998) show that the reaction of defense firms stocks to market movements is sluggish since these firms mostly receive contracts from the government and are hence relatively immune to stock index movements.

I, therefore, modify the existing Fisher and Peters (2010) stock return series by estimating the proxy for anticipation shocks as *abnormal* returns of stocks of military firms. These returns cannot be explained by reactions to markets movements and are the residuals of a Sharpe (1964), Lintner (1965) and Black (1972) (SLB) market model and are hence exogenous. I then include two additional covariates in the stock return regressions as suggested by Fama and French (1993) that are known to explain stock returns.

Both, the abnormal return approach, and the inclusion of the Fama and French (1993) covariates result in a purification of the proxy for future military spending changes and reduce its volatility compared to the Fisher and Peters (2010) measure. When ordering military spending first and the stock market abnormal returns last in an EVAR, the spending multiplier resulting from a shock to the abnormal returns series is 1.2 which compares to an estimate of 1.5 using the original Fisher and Peters (2010) series.

The remainder of this paper is organized as follows. The next section summarizes recent literature, followed by an introduction of the data. Section 3 describes the creation of the abnormal returns shock series augmented by the Fama and French (1993) factors, followed by a brief summary of VAR methods and estimation results in section 4, while chapter 5 presents robustness checks and a discussion. Finally, chapter 6 summarizes the paper and offers a conclusion.

2.2 Time series evidence for fiscal multipliers

Empirical work contributes to finding the fiscal multiplier by frequently relying on some version of a vector autoregression originally proposed by Sims (1981). Such estimations include different sets of variables, some with total spending, some with military spending along with tax measures, output, hours or wages while treating all variables as being endogenous.

Identification of fiscal spending shocks is necessary since reduced form systems are merely capable of summarizing the data, and shocks are not orthogonal. Identification can, for instance, be achieved through exclusion restrictions from institutional knowledge, economic theory, or Cholesky decompositions of the residual variance-covariance matrix. Such restrictions limit the in-period feedback of variables because of information processing lags or data inertia.

Blanchard and Perotti (2002) identify spending and tax shocks in a VAR including tax revenues, government spending, and output by breaking up the respective components to find their shock-feedback elasticities. They find output and taxes to increase in a hump-shaped manner following spending shocks resulting in a fiscal multiplier of 1.29.

The Blanchard and Perotti (2002) paper has triggered additional fiscal VAR research, for instance, Fatas and Mihov (2001) and Perotti (2002) et al. Multiplier estimates vary between 0.6 and 1.8, with the within range of the multipliers being almost as large as the across-study-range, as Ramey (2011a) notes.

Sign restrictions can alternatively be used in contrast to exclusion restrictions since the ex-ante assumption of zero-feedback can be too restrictive. This estimation method produces a set of impulse responses with an a priori specified sign. As Uhlig (2005), Pappa (2007) and Mountford and Uhlig (2009) demonstrate, sign restrictions give a feedback tendency that can sometimes be derived more credibly from theory than zero restrictions. Sign restrictions use an orthogonal decomposition P of the residual covariance matrix Σ_u while allowing for numerous orthogonalizations. Structural innovations are constructed as $\epsilon_t = P^{-1}u_t$ in which u_t are the reduced form residuals and ϵ are structural shocks. Then, sign restrictions and ac-

tual responses are examined. If no structural shock matching the feedback profile can be detected, eigenvalue/eigenvector matrices are rotated until impulse responses coincide with the expected signs, forming a bandwidth of possible responses rather than point estimates (Pappa, 2007). Sign restrictions and exclusion restrictions can be combined, as in Dungey and Fry (2009), for further flexibility. Uhlig (2005) concludes the sign-restrictions-derived fiscal multipliers to be between 0.5 (deficit financed spending) and two if a surprise deficit-financed tax shock occurs.

Recent papers like Kirchner et al. (2010) and Auerbach and Gorodnichenko (2012) question the symmetry of responses of output after spending shocks in recessions and expansions. Auerbach and Gorodnichenko (2012) implement regime switching models to determine time-varying multipliers over the business cycle, with smooth transition VARs using Blanchard and Perotti (2002) data. The multiplier for total spending is estimated to be 2.5 in a recession and 0.6 in expansions supporting the Keynesian argument of government consumption being less likely to crowd out private consumption if the economy is slack. In contrast, Owyang et al. (2013) demonstrate that there is no significant difference between U.S. multipliers in recessions and booms, whereas for Canada the difference is significantly different from zero.

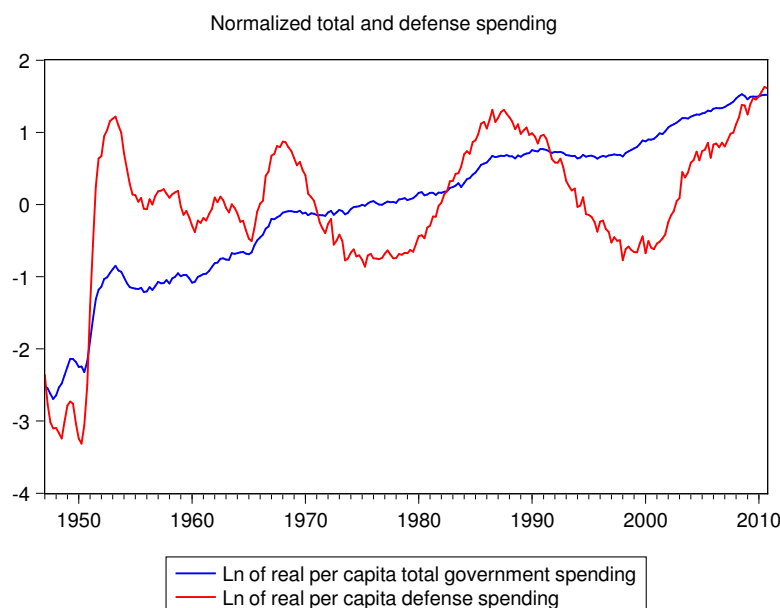
Most of the previously described studies have focused on aggregated spending. However, this measure contains non-defense spending components, which are primarily managed on state and local levels rather than by the federal government and include decisions on education, infrastructure, and community spending (Ramey, 2011a). Non-defense expenditures are adjusted to the demographic change and are therefore endogenously determined. It is doubtful whether non-defense spending components can contribute to the multiplier question due to their endogeneity (Pappa, 2007). Military spending decisions, in contrast, are made on the federal level and are determined in response to external events and are hence independent of business cycle developments.²

Moreover, as Figure 1 suggests, military spending is the major source of volatility of total spending, as also documented by Perotti (2011) and Ramey (2012).³

The annualized standard deviation of military spending is around 7.5% and thus approximately double the standard deviation of total fiscal expenditures. Sharp spending spikes during the Korean and Vietnam War and a misalignment of fiscal expansions and the actual beginning of a war is visible. For the first war, military

²Neglecting possible causality between economic conditions and the probability of military interventions.

³Additionally, Figure 10 on page 46 depicts changes in military spending and non-military spending, spelling out the importance of military spending decisions for total volatility.

Figure 1: Defense and total fiscal spending

Notes: The red line denotes the log of real per defense spending. The blue line is the log of real per total government expenditures. Vertical lines are the Korean War, Vietnam War, the Russian invasion of Afghanistan and the 2001 WTC attacks, according to Ramey and Shapiro (1998). Series are rescaled (zero mean and unit standard deviation) for comparability. See data section for sources.

expenditures more than doubled and remained persistently higher than previously. Policymakers, however, are mostly interested in multipliers from temporary spending boosts, often leading to an exclusion of the Korean War, which I follow (Fisher and Peters, 2010; Ramey, 2011a). Research on fiscal multipliers, incorporating military spending are for instance Ramey and Shapiro (1998); Barro and Redlick (2011); Fisher and Peters (2010) and Ramey (2011a), leading to multipliers between 0.6 and 1.5.

Except for the two latter studies, most papers abstract from fiscal foresight, that is individuals have superior knowledge on future spending changes compared to the econometrician analyzing the data, as is also visible in Figure 1 by a trailing of actual governments expenditures compared to the vertical war dates. It takes time for policy makers to discuss potential spending changes and sign them into law and signing and implementation do most likely not line up. Agents, however, learn of the consultations beforehand, adjust their economic choices, causing a discrepancy between their information and the information contained in the time series (Uhlig, 2005). If announcement effects of military build-ups are ignored, a missing state

variable problem causes VARs to be non-invertible, with the non-existence of a finite order moving average representation (Watson, 1986).

Disregarding the foresight issue generates estimates in which statistical shocks do not span the true information set of the agents and recovering structural shocks becomes impossible as documented by Leeper et al. (2012, 2013) and Perotti (2011). Such VAR innovations then dilute the actual fiscal spending shock and are a linear combination of past, and current innovations, which invalidate inference on the transition mechanism of fiscal spending shocks (Fernández-Villaverde et al., 2011). Foresight on military spending needs to be taken into account, which has been twofold in literature.

First, the dummy variable approach controls for foresight by dummifying out certain dates. Ramey and Shapiro (1998) create four war dates⁴ for unusual increases in the defense budget. As shortcoming, responses to negative and positive spending shocks are treated equally, ignoring their magnitude, since each of the episodes is characterized by the same dummy with a limited number of observations for shock dates. Ramey and Shapiro (1998) conclude the fiscal multiplier estimated when controlling for war dates to be slightly larger than unity.

The second approach to control for foresight is to incorporate an additional series capturing news on future macro fundamentals, the expectation augmented VAR (EVAR), since implementation lags cannot directly be observed in the macro aggregates. By imposing structure upon the data, identification under foresight is achieved, and defense spending in VARs is regarded as an instrument for total government spending (Ramey, 2011a). Ramey (2009, 2011a) estimates defense news EVARs including a series of present discounted values of expected military budget changes. Her series is based on *Business Week* articles reporting on events leading to changes in the military budget.⁵ She finds a stronger increase in output after an anticipated spending shock compared to results when ignoring foresight, concluding the multiplier to be between 1.1 and 1.2. Anticipation shocks crowd out private consumption, whereas if foresight is ignored, consumption tends to increase. Perotti (2011) estimates the multiplier in EVARs to be slightly smaller than unity.

In contrast to Ramey's narrative approach, Fisher and Peters (2010) define an anticipation shock as a shock to excess returns of military contractors' stock prices. In efficient markets, these have the advantage of instantaneously processing expectations appearing in stock returns. Their approach leads to a multiplier of 1.5. In

⁴The North Korean invasion of South Korea (1950q3), the beginning of American attacks in Vietnam (1965q1), the Russian invasion of Afghanistan (1980q1) and the WTC attacks (2001q3), which was amended later. Dates and the corresponding military budget are depicted in Figure 1.

⁵Figure 15 on page 51 displays the fit between the news series and actual changes to the budget.

the remainder of this paper, I extend the Fisher and Peters (2010) methodology to cope with structural patterns in returns, described in detail in the next section.

2.3 Stock returns & data definitions

In this section, I derive a novel measure to capture foresight on military spending. Afterward, I describe the remaining data utilized in the VAR estimation and to derive the fiscal multiplier.

Fiscal spending shock identification with stock prices is useful if the information contained in stock prices is a valid instrument for capturing agents' perception on future changes in fundamentals (Beaudry and Portier, 2006). Market efficiency suggests that the stock price today reflects the perception of probability-weighted discounted future profits of the agents. When investors assign a higher probability weight to military interventions with a greater defense spending need, they will correct their expectations of profits of receivers of additional government funds, purchase the respective stocks and cause instantaneous upward price shifts, as documented by Shapiro and Switzer (2011).⁶ Likely spending cuts, in contrast, lead to lower expected profits, a defense stock sell-off and lower stock returns.

My approach is to improve the Fisher and Peters (2010) shock by controlling for anomalies in asset pricing models, as suggested by Fama and French (1993). Such a measure for expectations of prospective defense contractors' profits needs to be free of market movements and only express expectations on future profits from additional defense contracts, as derived in the following subsections.

2.3.1 Anticipation shocks & excess return

Fisher and Peters (2010) construct cumulative excess returns (CER) from military contractors' stock returns,⁷ of firms ever ranked among top three receivers of the highest amount of government defense funds. Stock returns of any company with a military focus that has received one of the three largest dollar volumes of Department of Defense contracts in a fiscal year are reported. The authors aggregate market-value-weighted mean returns as top three portfolio return r_t^{t3} . Although such aggregation is not immune to criticism since military firms' order books consist of

⁶The assumption of higher profits due to higher revenue is plausible because of a lack of competition in the defense sector.

⁷Specializing in munitions and related equipment according to SIC (Standard Industry Classification) codes. See Fisher and Peters (2010, p 417) for details.

a variety of public and private contracts from home and abroad⁸, it proxies well for future military spending, compared to similar aggregations.⁹ A regression of total government spending on the lags of cumulative abnormal returns yields significant coefficients on the second and third lags, suggesting an anticipation horizon between half a year and three-quarters, in line with Ramey (2011a). Consequently, the information embedded in the aggregated top three portfolio returns is useful as a proxy for anticipated military spending to find the fiscal multiplier.

Finally, the technological progress and the markups of all defense firms are assumed to be identical to the rest of the economy.¹⁰

The Fisher and Peters (2010) anticipation shocks are then calculated as excess returns r_t^{er} given by

$$r_t^{er} = r_t^{t3} - r_t^m, \quad (1)$$

in which r_t^{t3} is the defense portfolio return and r_t^m is the market return. Calculating the proxy for future changes to the military budget, however, has drawbacks, as shown below. My contribution to the question of the size of the fiscal multiplier is twofold:

1. Find unexpected increases in the military budget by using *abnormal* rather than excess returns of firms specialized in ammunitions and arms.
2. Estimate abnormal returns while controlling for covariates that have been known to be persistent when explaining stock returns.

Abnormal stock returns in the market model. A stock return measure for anticipated fiscal spending needs to be free of stock market movements that are unrelated to military expenditures. Excess returns as in Fisher and Peters (2010), do not capture true spending shocks but include all factors that have an impact on stock prices except the market return. Consider the onset of a US military intervention. While learning about additional future military funds dedicated to arms and ammunition manufacturers, forward-looking investors purchase relevant stocks and cause upward price shifts. Excess returns as suggested by Fisher and Peters (2010) can point in three directions, depending on the reaction of the market as also depicted in Figure 2 to 4.

⁸Defense contractors' stock prices should primarily be driven by home sales to be meaningful for US spending shocks identification. Home sales (2010), according to balance sheets: Lockheed Martin: 83,8%; General Dynamics: 71,7%; Raytheon: 88.0%; Grumman: 92%.

⁹Outperforming the Fama and French (1992, 1993) guns portfolio, which includes all companies with military focus.

¹⁰Controlled for by including a deterministic trend, see p [419], [426] in Fisher and Peters (2010) for a discussion of critical issues.

1. The market does not respond. Excess returns would correctly capture anticipated fiscal spending shocks resulting from military interventions.
2. A negative market reaction due to global uncertainty arising from geopolitical tension may cause investors to shift funds into safe assets, which results in an overestimation of the true anticipation shock.
3. A positive market reaction and a positive but weaker response of military contractors' stocks together with only the jump in stock prices based on higher expected sales. Excess returns would underestimate the true shock.

Figure 2: Correctly measured excess return (case 1)

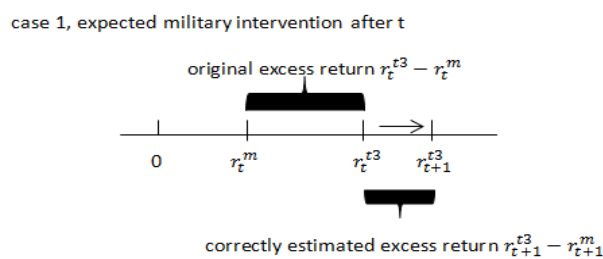
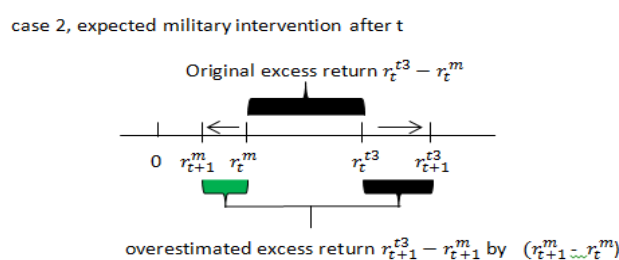
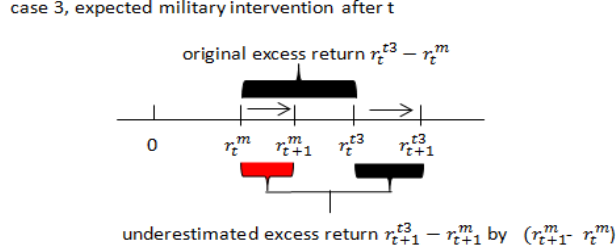


Figure 3: Overestimated excess return (case 2).



It is likely that the first case, as the excess returns in Fisher and Peters (2010), implicitly assume, does not hold. Shapiro and Switzer (2011), for instance, document adverse market reactions following war news, while military contractors experience a positive price shift. In this case, excess returns as a proxy for anticipation shocks would be misleading. A second bias can occur, because excess returns assume the impact of interest rate changes on the market to be identical to the reaction of the stock prices of military firms.

Figure 4: Underestimated excess return (case 3)


The bias arising from the usage of excess returns can be shown in Equation 2 known as the Sharpe (1964), Lintner (1965) and Black (1972) (SLB) market model for asset pricing

$$r_t^{t3} - r_t^f = \alpha + \beta[r_t^m - r_t^f] + \epsilon_t, \quad (2)$$

where r_t^{t3} is the top three portfolio return, r_t^f is the risk free rate, r_t^m is the market return and α identifies market independent top three portfolio-specific returns. β denotes the sensitivity to market movements for which each additional unit of risk is compensated by β . ϵ_t is a stochastic, abnormal return incorporating all unexplained factors, including those attributed to fiscal spending boosts. The information contained in ϵ_t is of interest to capture military spending increases rather than relying on ex-post realized excess returns.

Abnormal returns derived from Equation 2 are well founded in empirical finance. Rotemberg and Woodford (1992) propose shocks from regressing military spending on its lags and military employment and DellaVigna and La Ferrara (2010) propose a method to detect illegal arms trades based on unexpected stock returns of defense contractors around weapons embargoes. In DellaVigna and La Ferrara (2010), abnormal returns are also the residuals of Equation 2 and are hence the part of the SLB market model that cannot be explained by the regression equation. Additionally, Corsetti et al. (2012) estimates fiscal spending rules and also uses the residuals of the regression as spending shock.

ϵ_t is uncorrelated with portfolio returns, representing exogenous shocks, whenever surprise changes in military spending can neither be explained by the difference in the market return and the risk-free rate, with the shock being equal to

$$\epsilon_t = (r_t^{t3} - r_t^f) - \alpha - \beta[r_t^m - r_t^f]. \quad (3)$$

An excess return innovation as measured by Fisher and Peters (2010) is described by $r_t^{t3} - r_t^m$. From Equation 2 the true shock and the excess returns shock only coincide if defense returns move one-for-one with the market ($\beta = 1$) and the portfolio specific α is zero. However, the β of defense firms tends to be smaller than unity, and α to be significantly larger than zero (Fama and French, 1998).¹¹ Military firms primarily receive contracts from the federal government so that their stock returns are largely driven by shocks and less by a reaction to market movements. The tendency of the bias from excess returns shocks rather than using abnormal returns is given by

$$\begin{aligned} \epsilon^{true} &\geq \epsilon^{FP} \\ r_t^{t3} - \beta^t r_t^m + r_t^f(\beta^t - 1) - \alpha^{true} &\geq r_t^{t3} - \beta^{FP} r_t^m + r_t^f(\beta^{FP} - 1) - \alpha^{FP}, \end{aligned}$$

and when assuming $\beta^{FP} = 1$ and $\alpha^{FP} = 0$

$$(r_t^m - r_t^f)(1 - \beta^t) - \alpha^t \geq 0,$$

in which the superscript *true* denotes the true α and β and *FP* refers to the Fisher and Peters (2010) estimate.¹² The direction of the bias depends on time-varying differences between market and risk-free returns. Abnormal returns, in contrast, correctly proxy for the anticipation shock in all market phases and every β and α .

The Fama and French (1993) three-actor model. Fama and French (1993) show that explanatory power of SLB model is weak. Banz (1981), Fama and French (1992), Fama and French (1993) and Fama and French (1998) suggest extending the original SLB model by controlling for persistent structural effects when explaining stock returns. Fama and French (1993) suggest two factors that proxy for common underlying risk factors in stock returns while increasing the explanatory power of the SLB model.

First, Banz (1981) claims the existence of a size effect in stock returns embedded in market equity, *ME*.¹³ Low *ME* values (i.e. small companies) were shown to have too high a return given their market β for several years and vice versa. If stock portfolios are formed on size, investors rate small firms as riskier compared to large companies, compensated for by higher returns. Fama and French (1998) document the return difference, a size premium, between small and large stocks to be around 7.7 percent per annum.

¹¹Fama and French (1998) estimate the average β of military contractors (guns portfolio) to be around 0.9 and the respective α to be 6.3.

¹²See appendix for details on the derivation of the bias.

¹³Product of outstanding shares and current stock price.

Second, as Fama and French (1992) demonstrate, a book-to-market equity value (BE/ME) also contributes in explaining stock returns. Firms with relative high BE/ME values tend to have persistently higher returns compared to firms with low BE/ME values. Large BE/ME values can be seen as investors' punishment because they believe in lower future earnings, sell the respective stocks, and causing a decline in ME . Lower expected earnings proxy for financial distress making the individual stock position riskier. Consequently, investors demand a higher risk premium. According to Fama and French (1993), stock portfolios demonstrated a premium of approximately five percent between 1963 and 1991 due to higher perceived financial distress, called value premium.

The size and value premium persist in the SLB market model regressions, even when competing with other factors so that small firms and companies with a high BE/ME tend to be riskier with higher demanded risk premia (Fama and French, 2012).

Taking these two covariates into account, an anticipated spending shock needs to be free of such known structural anomalies because they can contribute to also explaining top three portfolio returns. Fama and French (1992, 1993) extend the SLB model to a three-factor model, which leaves the anticipated spending shock free of movements in the relevant factors:

1. A size factor, determined by ordering stocks into three groups according to ME each year. Small stocks are those in the bottom 10%, large stocks those in the top 10% value of market capitalization.
2. Stocks are assigned to three portfolios according to book-to-market equity. Breakpoints are in the lowest and highest 30%. Stocks in the upper portfolio are referred to as value stocks, those in the lowest 30% as growth stocks.

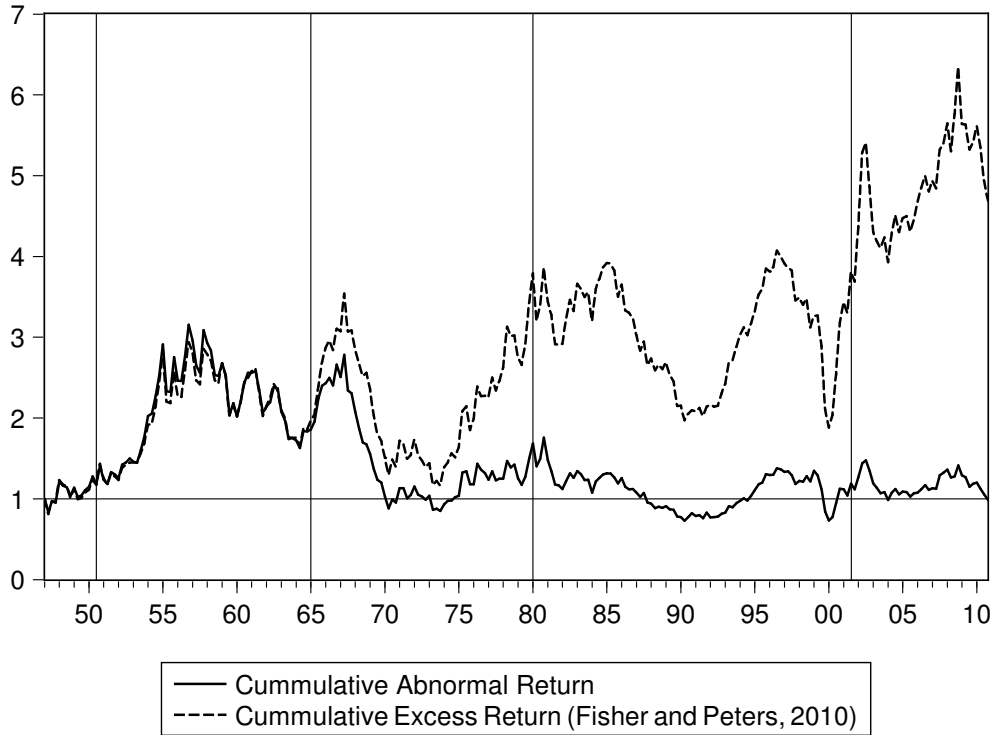
The return difference between the first Fama and French factor, SMB (Small Minus Big), is captured as the equally weighted average of three small stock portfolio combinations and the returns of three big stock portfolios combinations: $SMB_t = 1/3(\text{Small Value} + \text{Small Neutral} + \text{Small Growth})_t - 1/3(\text{Big Value} + \text{Big Neutral} + \text{Big Growth})_t$. Secondly, HML_t (High Minus Low), is the equally weighted average of return differences of the value portfolios (high BE/ME) and the growth portfolio (low BE/ME) so that $HML_t = 1/2(\text{Small Value} + \text{Big Value})_t - 1/2(\text{Small Growth} + \text{Big Growth})_t$. The third factor is the market-specific reaction as in the standard SLB regression, making up the Fama and French three-factor model.¹⁴

¹⁴Factors are from mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html and continuously updated.

$$r_t^{t3} - r_t^f = \alpha + \beta[r_t^m - r_t^f] + hHML_t + sSMB_t + \epsilon_t \quad (4)$$

In this equation, the first two factors are defined as above and h and s are the two regression coefficients of the two additional Fama and French (1993) factors. The estimation leaves the anticipation shock ϵ_t to be free of movements in the three factors and thus pure compared to the Fisher and Peters (2010) excess returns shock. Having captured the innovation as ϵ_t , I create cumulative abnormal returns, with $1957q3 = 1$. Figure 5 depicts cumulative abnormal returns and the Fisher and Peters (2010) series, with a comovement of both series until the onset of the Vietnam War.

Figure 5: Abnormal returns and Fisher and Peters (2010) returns



Notes: Cumulative abnormal returns (solid line, estimated) and cumulative excess return (dashed line, Fisher and Peters (2010)) of the top three receivers of government military funding, 1957q3 to 2010q4. Vertical lines denote the Ramey and Shapiro (1998) war dates.

Abnormal returns experience a mean of zero with a standard deviation of $\sigma_{ar} = 17.4\%$ whereas excess returns exhibited an annualized mean return of $\mu_{ar} = 2.4\%$ with higher standard deviation of $\sigma_{er} = 18.6\%$ for the excess return series.¹⁵ On

¹⁵An Augmented Dicky Fuller test on the stationarity of the abnormal returns series is borderline significance on the 10% significance level while it can be rejected for the Fisher and Peters (2010) excess returns series.

average, however, expected changes in defense expenditures are zero for the sample since wars and spending cuts are exogenous, ruling out ever-increasing military spending. The new shock measure enters fiscal VARs, together with the variables described in the next subsection and is placed last.

2.3.2 Data definitions and sources

I include quarterly U.S. FRED and Ramey (2011a) data from 1947q1 to 2013q4, upper limited by the availability of the Fisher and Peters (2010) series.

Output, its components, the three month t-bill, the Barro and Redlick average marginal tax rate, total population, wages, and hours are obtained from Ramey (2011a),¹⁶ and FRED¹⁷.

I denote the variables as gdp , the log of real per capita GDP;¹⁸ g^m , the log of real per capita military expenditures; g , the log of total per capita government spending; $tb3$, the nominal three-month treasury-bill rate; w , the log of real wages; c , the log of real private per capita consumption expenditures; h , the log of total hours worked; mtr , the Barro-Redlick average marginal tax rate and t the log of real per capita government revenues.

Francis and Ramey (2009) note that the reaction of hours depends on the way they enter the VAR system. If hours are to be included in first differences, hours may fall after a spending shock. In contrast, hours are naturally bounded, which is why I include hours in levels.

Finally, CAR (Cumulative Abnormal Returns) are then formed as described above, whereas CER (Cumulative Excess Returns) denotes the original Fisher and Peters (2010) series. Additional information on the data can be found in the appendix in section 2.7.

¹⁶Available as GDP (GDP) total nominal defense spending (rdef), total government consumption (rgov), nominal three-month treasury-bill rate (tb3), total household based hours worked (tothours), the Barro-Redlick marginal tax (amtbr), total population (pop) and cumulative abnormal military contractors' returns (top3xsret) at weber.ucsd.edu/vramey/research.html.

¹⁷Private consumption expenditures (pcecc96), nonfarm business per hour compensation (comprnfb), total government current receipts (fgrecept) and GDP Implicit Price Deflator in \$2005 (gdpdef) available at research.stlouisfed.org/

¹⁸Deflated with GDP implicit deflator and CPI respectively in \$2005. Ramey (2011a) uses a weighted average of CPI inflation and manufacturing inflation based on ratios of the nominal values of defense and investment to GDP, and the component series weights on each type of inflation for GDP, consumption and spending (Ramey, 2011a). Nevertheless, my results remain robust if Ramey's aggregation is used.

2.4 Estimation results

I estimate impulse responses from fiscal VARs conditional on innovations in abnormal returns and then compare impulse responses to the Fisher and Peters (2010) excess return shocks. I conclude the fiscal multiplier derived from abnormal return shocks to be 1.2, lower compared to a multiplier using innovations in excess returns shocks. To compare my results with the ones previously found, I initially estimate VARs from 1957q3 to 2007q4, before extending the sample to include financial crisis data to check for influential data points.

My results are qualitatively in line with Ramey and Shapiro (1998); Fisher and Peters (2010) and Ramey (2011a), but differ quantitatively. An anticipation shock, constructed as abnormal returns and controlling for the Fama and French (1993) factors embedded in stock prices, raises output more profoundly and crowds out private consumption by more compared to the Fisher and Peters (2010) shock for six-quarter before rising, compared to impulse responses estimated using Fisher and Peters (2010) returns.

2.4.1 Identification and estimation

The benchmark VAR estimation includes a column vector X_t' of endogenous variables, ordered as military spending, output, t-bill, aggregated spending and either the original Fisher and Peters (2010) series or the abnormal return series. Total spending is then substituted for additional variables of interest like hours, wages, consumption and tax measures as also done in Fisher and Peters (2010). The reduced form VAR can be expressed as

$$X_t = B(L)X_{t-1} + u_t, \tag{5}$$

in which $B(L)$ is an autoregressive lag polynomial, L is the lag operator and u_t are reduced form innovations. The VARs include six lags,¹⁹ a constant and also a linear deterministic time trend and is estimated using ordinary least squares.

Since only the reduced form innovations u_t with limited economic meaning can directly be observed, structural shocks need to be recovered.

The system is identified by Cholesky decomposition to Σ_u , the reduced form variance-covariance matrix, with defense spending ordered first and one of the two return measures last. Innovations are then orthogonal to variables representing the state of the economy. This ordering implies defense spending to be contemporane-

¹⁹Suggested by *AIC* and *SC* and qualitatively robust to estimations with two up to five lags.

ously exogenous to the remaining variables. Additionally, I allow the return series to react contemporaneously upon all other variables acknowledging stock returns' instantaneous information processing nature. The ordering of variables in X_t in this setup can be essential for the outcome of the impulse responses, which I address in the section on robustness. Results remain similar, however, if I order the return series, e_t^{car} and e_t^{cer} , first, as also noted by Fisher and Peters (2010).

I also substitute wages, hours, private consumption, tax revenue and the Barro-Redlick marginal tax rate for total fiscal expenditures, since they are of relevance for the policy maker, as also suggested in Fisher and Peters (2010).

2.4.2 Estimation results from abnormal and excess return innovations

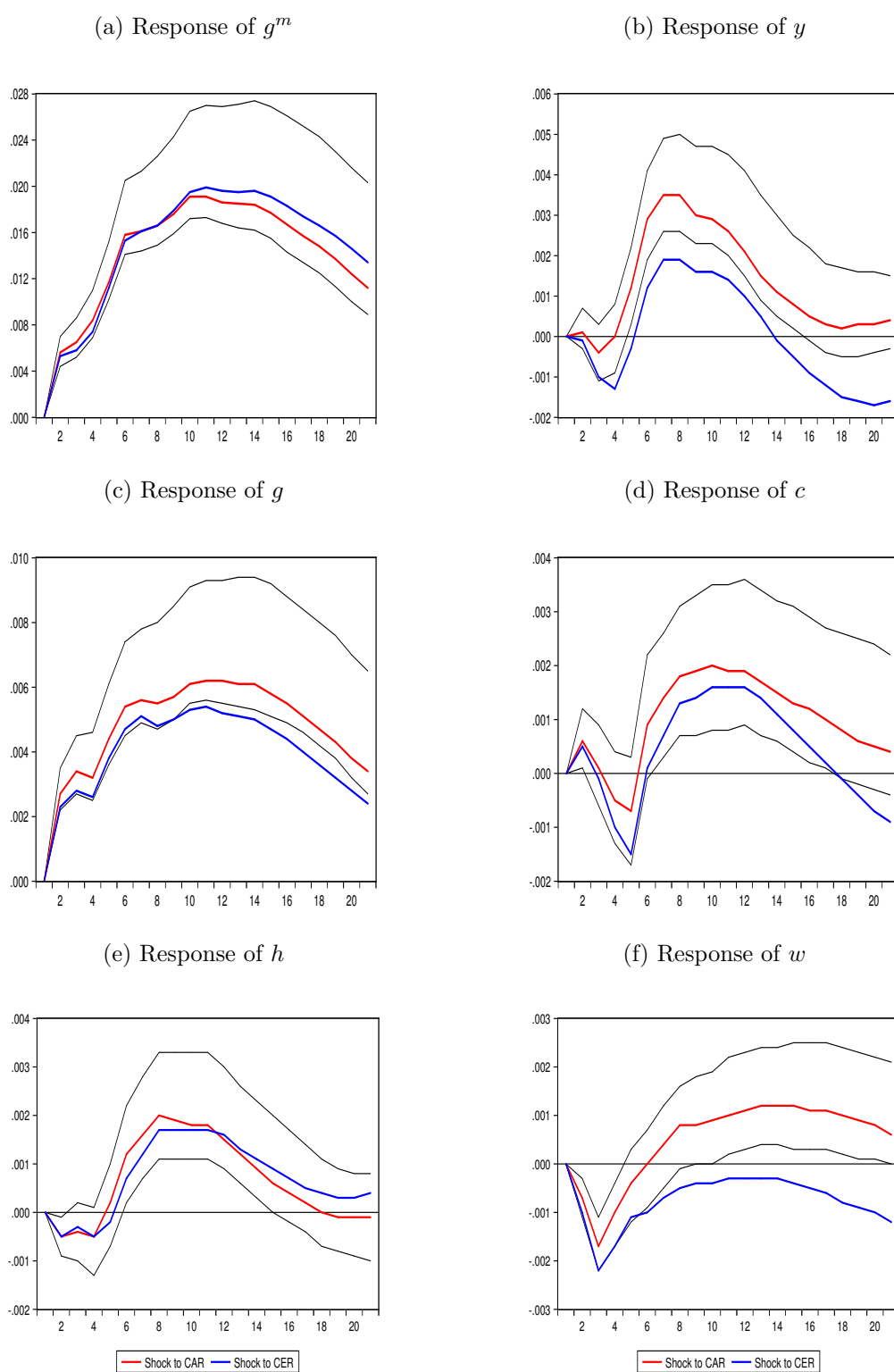
Fisher and Peters (2010) find that their constructed cumulative excess returns, initiate an expansion in the economy with an increase in output, hours, wages and a crowding in of private and government consumption with a resulting five-year multiplier of 1.5, which compares to my estimate of 1.2²⁰. Private consumption is initially crowded out before it increases in a hump-shaped manner.

Figure 6 represents structural impulse responses of military and total government expenditures, output, consumption, hours and wages recovered after a one standard deviation increase in the return indices. Figure 13 in the appendix displays cumulative IR. Red lines denote responses to an innovation in abnormal returns, with respective 68% bootstrapped confidence bands (thin dashed line).²¹ Blue lines are point estimates of the responses to the excess return innovations.

²⁰ $\frac{0.0267}{0.2943} \times 1/0.0736 = 1.2\%$ with 0.0267 being the cumulative effect of the CAR shock on output, 0.2943 the effect after five years on g^m and 0.0736 is the average ratio of military spending to gdp .

²¹If 90% confidence bands are used, the reaction of the respective variables are all non-zero as well, except the impulse response of the three-month t-bill, which would not be different from zero.

Figure 6: IR for abnormal and excess return shocks I



Notes: red: IRs of the logs of real capita a) military spending; b) GDP; c) total government spending; d) personal consumption, and e) total hours and f) real per hour wages, to a one std deviation shock to CAR. Black: respective 68% confidence intervals (10.000 replications). Blue: point estimates IR given CER shocks, 1957q3-2007q4.

A shock to accumulated returns displays a hump-shaped response pattern in all six variables, whereas the strength of the impact depends on the choice of return series.

Output, private consumption and hours remain statistically not different from zero for six quarters. As visible, the response of output after the Fisher and Peters (2010) shock lies outside the confidence bands of the abnormal returns shock that I construct in this paper. Total and military spending exhibit an instantaneous increase, whereas the reaction of defense spending is quantitatively more than twice as large in the peak as the one of total fiscal expenditures. Military spending peaks at ten quarters. The cumulative difference after 20 quarters is approximately 1.5 percentage points, as shown in the appendix in Figure 13.

The impulse response of output suggests a three to four quarters delayed hump-shaped response, consistent with Ramey (2011a). Figure 6b also indicates that output reacts 20% stronger after the shock to abnormal returns compared to a shock to excess returns.²² The discrepancy between responses after both shocks takes a 20 quarter value of 1.2 percentage points, as Figure 13 suggests. For the 9/11 spending increase, the difference accumulates to approximately to USD 1.3bn over 20 quarters, whereas the difference for the Korean War spending growth would have been roughly USD 5bn.

IRs of private consumption reveal a quantitative difference of both impulse responses, by initially being crowded out twice as strong following the CER shock. Five to six quarters later, consumption rises more profoundly in the case of the CAR innovation. The gap after 20 quarters takes a value of 1.2 percentage points, as also depicted by cumulative impulse responses in Figure 13.

Impulse responses of hours worked are similar for both shocks, with a period of inactivity followed by an expansion after five quarters consistent with, for instance, in Ramey (2011b).

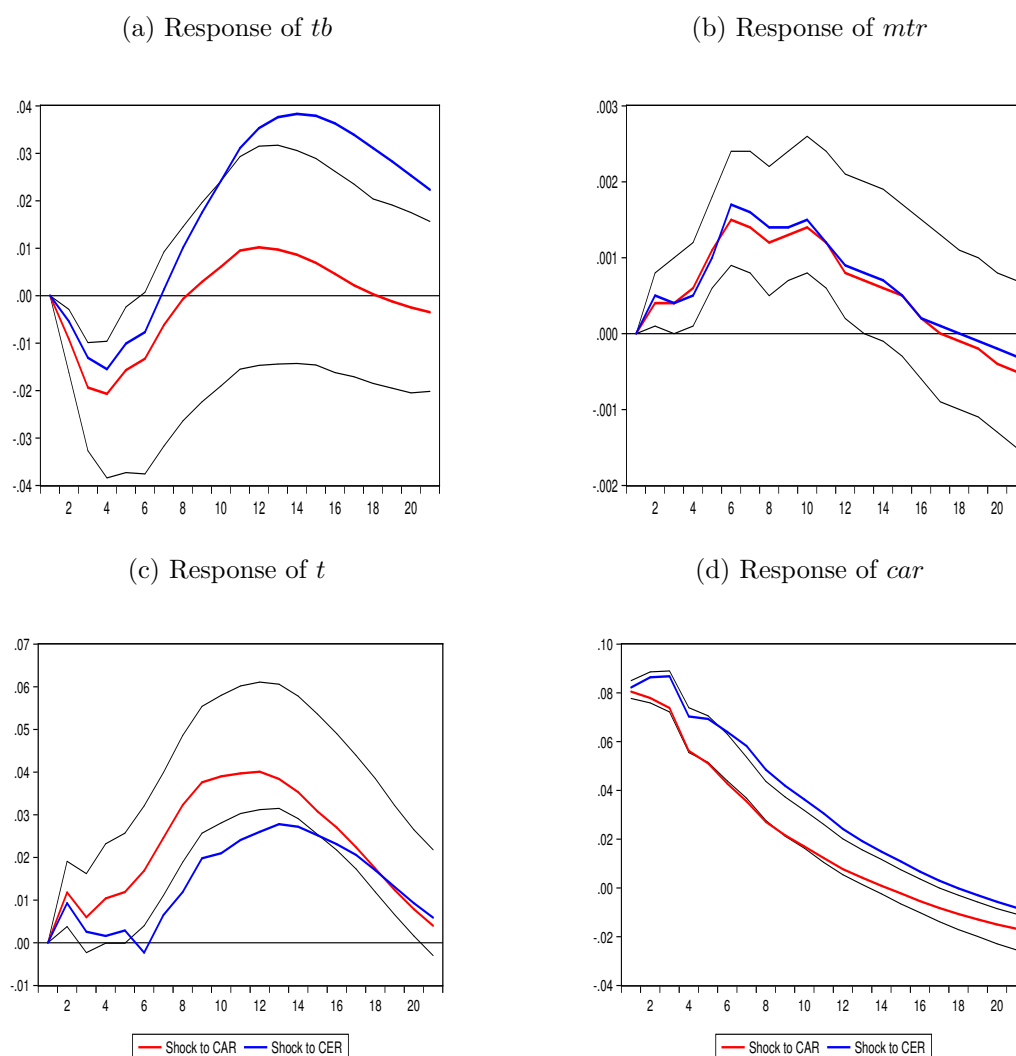
Impulse responses of wages demonstrate a significant decline after the two shocks. After an impulse to CAR, real wages respond positively five quarters after the shock, whereas the deviation from pre-shock states after CER shock is persistently negative for the post-shock periods. The impulse after the CAR shock exhibits a hump-shaped pattern, with a maximum increase after 14 quarters, taking a value of approximately 1%. The cumulative difference between the two shocks after 20 quarters is 2.6 percentage points.

²²0.35% compared to 0.25% at the peak of the response.

In conclusion, the reaction of consumption to CER shocks demonstrates stronger effects, whereas CAR shocks have a higher impact on output and wages. The reactions of total spending, military spending, and hours are similar.

Figure 7 summarizes the response of the policy instruments to analyze whether the response of output is based on interest rate effects. I also add the Barro-Redlick average marginal tax rate to the system.²³

Figure 7: IR for abnormal and excess return shocks II



Notes: red: IRs of a) nominal three-month t-bill b) Barrow-Redlick average marginal tax rate c) logs of tax revenues d) cumulative returns, to one std deviation CAR shock. black: respective 68% bootstrapped confidence bands (10.000 reps). Blue: point estimates IR to one standard deviation CER shock, 1957q3-1907q4.

IRs of the t-bill show the period of inactivity, but with a high degree of uncertainty, expressed in wider confidence bands. The policy instrument initially drops

²³The Barro-Redlick average marginal tax is available on an annual base.

by 2 percentage points after the CAR innovation and by only 1% after the CER shock before rebounding. After four quarters, the t-bill starts to rise, peaking at 1% after 11 quarters following a CAR shock and 3.8% following the CER shock after eight quarters.

A non-zero response of the t-bill might, in contrast to Fisher and Peters (2010), has an impact on the other variables. When looking at conditional responses after abnormal returns shocks, output or consumption may react upon the instrument directly. However, taking the variance decomposition of GDP into account, the interest rate effect accounts for only five percent of the variance in output.

The marginal tax rate increases for approximately four years, before declining as indicated by Figure 7b. As is evident from Figure 7c, tax revenues rise by 1.5 percentage points more after a shock to abnormal returns when comparing with the Fisher and Peters (2010) shock. Such an increase is a combination of higher revenues owing greater economic activity and higher tax rates, as Figure 7b suggests so that additional spending seems taxed-financed.

Variance decompositions can be used to check how relevant the difference in impulse responses between the two shocks is. CAR and CER both take a value of 28% of explained variance of g^m after 20 periods. However, the fraction of explained variance is higher when using the CAR shock for GDP (5% for CER vs. 8%, total spending g (14% and 19%), consumption (3% and 4%) and hours worked (4% and 5%). Thus, CAR innovations drive the macroeconomic variables in the system and explain additional of their variances compared to CER shocks.

2.5 Robustness of results & limitations

This section discusses main findings and limitations.

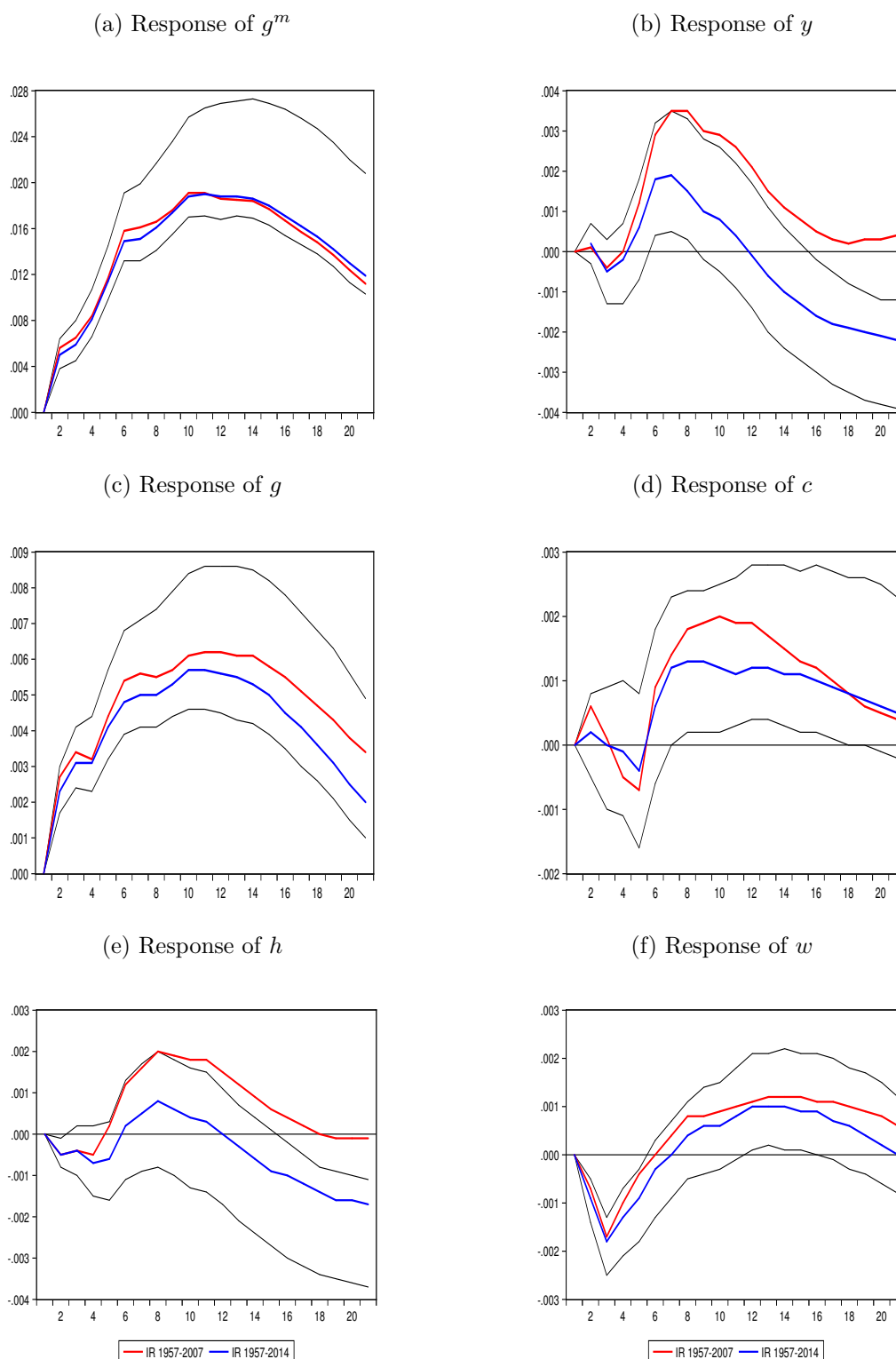
Results are robust to ordering the abnormal returns variable first and also to extending the sample.

When substituting the log of the three-month t-bill for the levels variables which leads to an elasticity representation, the results remain similar, as suggested in Figure 12 in the appendix.

2.5.1 Robustness I: extending the sample

Because the financial crisis might substantially have an impact on the estimated coefficients of government spending on output, I extend the sample to incorporate recent data. Hence, I estimate two different sub-samples, displayed in Figure 8, the original one with 1957q3-2007q4 data and one sample with data from 1957q3 to

Figure 8: IR baseline and extended sample I



Notes: red: IRs the logs of a) real per capita military spending b) real per capita GDP c) real per capita total government spending d) real private per capita consumption, e) total hours worked and f) real per hour wage, to a one std deviation shock to CAR, baseline (57-07). Blue: point estimates after one std deviation CAR shock for subsample 1957q3-2013q4, with respective 68% bootstrapped confidence interval (10.000 replications), black lines.

2013q4 (blue lines). The extended sample is of interest since large amounts of funds were injected into the economy as part of fiscal stimulus packages, and interest rates reached the Zero Lower Bound. Therefore, extending the sample by six years adds interesting observation points. I compare impulse responses conditional on CAR shocks for both subsamples.

Figure 8 depicts differences for the two subsamples in which blue line denotes the extended sample (up to 2013q4) and the red lines indicate the responses estimated with data up to 2007q4, both for the shock resulting from the estimation in this paper.

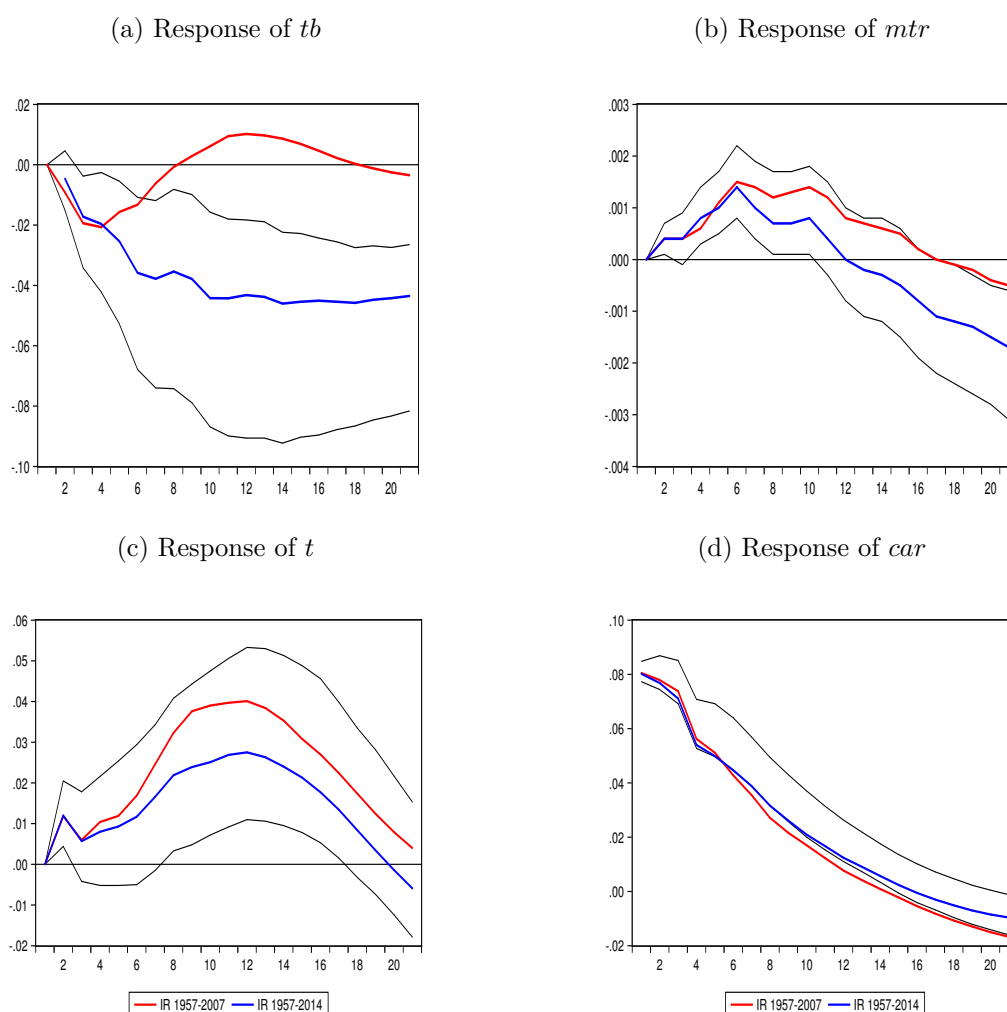
An extension of the sample leads to a less profound increase in output. Quantitatively, the peak response of output of 1.9% compares to a 3.5% response for the pre-financial crisis sample. This result stands in contrast with Auerbach and Gorodnichenko (2012) who estimate a stronger reaction of output to fiscal policy shocks if the economy is slack. However, the full sample is a blend of both subsamples. During the financial crisis, non-defense spending was increased in contrast to earlier periods which reduces the importance for anticipated military spending shocks. Therefore, the effect of the CAR shock is not as important a driver as in times of war. The fraction of explained variance of CAR shocks for output after 20 quarters falls from 7% to only 3% if the sample is extended.

When extending the sample, hours only increase half as strong compared to the shorter baseline. Finally, Figure 8d suggests that including financial crisis data causes the peak of the IRs to be almost twice as high, compared with the original sample. The remaining IRs are displayed in Figure 9.

A CAR innovation, for the baseline dataset, causes the t-bill to decrease initially, converging to zero deviation after 30 quarters. The last added observations between 2008 and 2013 introduce higher uncertainty, visible in the upper left corner of Figure 9 and given by wider confidence bands compared to the shorter subsample. In the full sample, the impulse response after the shock is significantly negative.

The marginal tax rate shows an earlier decline, almost two years earlier compared to the full sample. A similar effect is present when looking at the impulse responses of tax revenues. The peak of the response of tax revenue in the dataset including the financial crisis is two-thirds of the peak of the subsample.

The qualitative estimation results are also robust to substituting the tax variable with the fraction of taxes revenues over GDP. The impulse responses remain almost identical.

Figure 9: IR baseline and extended sample II


Notes: red: IRs of the a) nominal three-month t-bill b) Barrow-Redlick average marginal tax rate c) logs of tax revenues d) cumulative returns, to a one std deviation shock to CAR, baseline (57-07). Blue: point estimates after one standard deviation CAR shock for subsample 1957q3-2013q4 with respective 68% bootstrapped confidence intervals (10.000 reps), black lines.

2.5.2 Robustness II: five-factor and momentum-model

Fama and French (2015) extend the three-factor-model and find two additional covariates. These factors capture profitability and the investment patterns of firms and explain stock returns as do the previously described factors. The authors argue that some variation in stock returns, related to profitability and investment decisions cannot be explained by the three-factor-model. The authors hence amend the three-factor model for profitability and for the degree of investment appetite of firms. Fama and French (2008) and Novy-Marx (2013) show that profitable companies, as well as firms with conservative investment decisions, realize too high stock returns given their market beta.

Hence, I extend the baseline estimation and include these two additional factors in the estimation of the SLB regression for abnormal returns series as

$$r_t^{t3} - r_t^f = \alpha + \beta[r_t^m - r_t^f] + hHML_t + sSMB_t + rRMW_t + cCMA_t + \epsilon_t, \quad (6)$$

in which the first ($[r_t^m - r_t^f]$), second (HML) and third factor (SMB) are the ones from the data section of this paper. RMW_t denotes a portfolio of firms, generated from robust-minus-weak profitability firms,²⁴ and CMA_t is the difference between stocks of conservative and aggressive investment firms.²⁵ The respective regression coefficients are r and c .

Estimation results of abnormal stock returns as a proxy for future military spending changes, however, are similar to the previously found results. If all the factor exposures, HML , SMB and $[r_t^m - r_t^f]$ explained all variations in stock returns, α would be zero. In the estimation in Equation 6, this is the case since the constant is not statistically different from zero. The explained variance of the model slightly increases compared to the three-factor-model. However, impulse responses estimated after anticipation shocks remain similar, with the cumulative multiplier being close to the one estimated using the three-factor-model.

Additionally, Carhart (1997) and Fama and French (2012) argue that regression betas are time-varying, also expressed by a momentum effect so that well-performing stocks remain well-performing for a particular time and vice versa. However, Fama and French (2015) show that the regression slope of such is zero in the five-factor-model²⁶. Moreover, amending the three-factor-model with the "momentum effect" in the estimation of my abnormal returns series does not have an impact on the size of the fiscal multiplier either.

Figure 16 in the appendix depicts the proxy for changes in military spending using the three-factor, five-factor, and momentum model. Quantitatively, the series only marginally differ so that the multiplier estimated with either of the models is

²⁴Measured with accounting data of the previous year, as the difference between annual revenues minus the cost of goods sold, interests, administrative expenses and sorted from high to low, called "operating profitability". Low profitability is defined as the bottom quintile, high profitability as the upper quintile of all firms since the three middle quintiles are very similar according to Fama and French (2015). RMW is calculated as $RMW = \frac{1}{2}(\text{small robust} + \text{big robust}) - \frac{1}{2}(\text{small weak} + \text{big weak})$, as also done for the size factor.

²⁵Investment is defined as growth of total assets for each year, divided by total assets that year. Then again, conservative investment is defined as firms in the bottom quintile, whereas aggressive investment is defined as firms in the upper quintile. CMA_t then calculated as shown above as $CMA = \frac{1}{2}(\text{small conservative} + \text{big conservative}) - \frac{1}{2}(\text{small aggressive} + \text{big aggressive})$.

²⁶All data on CMA_t , RMW_t and MOM_t is obtained from Kenneth French's homepage at [mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data Library](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data%20Library).

almost identical. Therefore, the estimate for the multiplier is robust to changing the underlying factor model.

2.5.3 Limitations and discussion

A fiscal multiplier derived from abnormal returns shocks is smaller when adding financial crisis data. If multipliers are estimated to be higher than true multipliers, they are misleading, and a larger spending increase is needed to overcome a slump. As Ramey (2011a) notes, VAR multiplier estimates can be time-varying and decreasing as government expenditure hits its peak, which can potentially explain why the response of output in the extended sample is weaker.

The financial crisis caused non-defense components to increase substantially, while introducing the first observations with such an increase in non-military spending, ever. Endogenously determined spending decisions create biased estimates for the fiscal multiplier as shown above.

Perotti (2011) criticizes VAR approaches with military expenditure to produce multipliers smaller than one if World War 2 and the Korean War are included.²⁷ Hence, I check the results if the Korean War is included with 1947 onwards data. The reactions of total spending, the t-bill, and private consumption remain similar and the response of output changes in the same manner as if financial crisis data is included. The response of output becomes negative after several quarters. Since the series begins in 1947, spending cuts as a consequence of World War 2 causes military spending to fall after the anticipation shock. Wages increase more profoundly, without a quick convergence as in the original sample. The reaction of total hours becomes negative and stays below zero before converging back to zero in a hump-shaped manner.

Finally, Barro and Redlick (2011) criticize that spending multipliers derived from military expenditures shocks can be misleading if command and control techniques like rationing private expenditures on goods and services or drafting the population to serve are present. Spending multipliers then can be lower than in peace times. However, this effect could be offset by a mandated increase in production and labor, which in turn would raise the multipliers. In conclusion, the direction of the bias from both effects connected with military spending is ambiguous. Therefore, using military spending seems more appropriate compared to using total spending to avoid the endogeneity problem.

²⁷He suggests, in contrast, to Ramey to dummy out two quarters of World War 2 and the Korean War, which would be enough to control for announcement effects.

2.6 Conclusion

The estimated size of the fiscal multiplier using stock prices in an EVAR approach depends on how anticipated spending shocks are measured. Two common problems when estimating fiscal multipliers are solved in this paper. I address the endogeneity problem by ordering defense spending first in an orthogonal VAR since this measure is likely to be exogenous.

Also, fiscal foresight needs to be taken into account. I do so by relying on a proxy for future military spending changes based on stock returns of military contractors as suggested by Fisher and Peters (2010). However, as I can show, excess returns do not capture the true anticipated spending increase.

I, therefore, introduce a novel measure of anticipated spending changes by re-defining the externally created measure by Fisher and Peters (2010). Instead of calculating excess returns, as the authors do, I suggest estimating spending increases by regression the Fisher and Peters (2010) on a number of factors that are known to explain stock returns as indicated by Fama and French (1993) in an SLB market model regression. The resulting anticipation shock is then captured as abnormal returns, that is, everything that cannot be explained by the regression model.

I include the novel measure as a proxy for anticipation shocks to answer the question of how the U.S. economy responds to government spending shocks.

Estimation results from Cholesky-decomposed reduced form residuals variance-covariance matrix with defense spending ordered first, and the return series last, are qualitatively in line with Fisher and Peters (2010) and their measure of excess returns. Impulse responses recovered conditional on abnormal returns, however, show that there is a significant difference in impulse responses compared to a shock to the original Fisher and Peters (2010) excess returns series. After an anticipation shock, output and consumption, total hours and private consumption react with a delay after a shock to cumulative abnormal returns instead of excess returns.

Results are important for researchers and policymakers alike since I create a purer shock to control for foresight and give an improved estimate for the multiplier, taking a value of 1.2. However, I cannot rule out interdependency of a reaction of the t-bill since the impact is stronger in my sample compared to the original Fisher and Peters (2010) series.

I also extend the Fisher and Peters (2010) sample and check for robustness by adding financial crisis data, resulting in significant lower impulse responses of output. The relationship between the anticipation shocks and military spending becomes weaker since non-defense spending measure increased during the current

crisis. Extending the sample with post-Great Recession data, once it becomes available, may add additional insight on the multiplier question.

2.7 Appendix

Bias from using excess returns. This subsection isolates the shock in the SLB model and compares the Fisher and Peters (2010) excess return shock (FP) with the new anticipation shock from abnormal stock returns.

Fisher and Peters (2010) imply that the difference between the defense portfolio and the market to be the anticipation shock. However, this difference assumes the elasticity of the defense portfolio with respect to the risk-free rate to be identical to the interest rate elasticity of the market. The following relation holds.

$$\begin{aligned} r_t &= \alpha + r_t^f + \beta[r_t^m - r_t^f] + \epsilon_t \\ r_t - r_t^f - r_t^m - r_t^f\beta &= \alpha + \epsilon_t \\ [r_t - \beta \times r_t^m] + r_t^f(\beta - 1) - \alpha &= \epsilon_t, \end{aligned}$$

where α is the firm-specific return, and ϵ is the spending shock, defined by Fisher and Peters (2010). For $\epsilon = r_t - r_t^m$ to hold, β has to be equal to unity and $\alpha = 0$ or $\beta \times r_t^m + r_t^f(\beta - 1) - \alpha = r_t^m$. Both is questionable to hold since the returns are highly time-varying and not only depending on the risk-free rate. Additionally, if shocks are exogenous, the stock returns of defense contractors are mainly driven by fiscal spending and not by a reaction to market movements. Since β compensates for bearing additional units of risks, if shocks were truly exogenous, β is smaller than one and close to zero. Excess returns, however, imply a β of unity. Thus, from the equation above follows

$$\begin{aligned} \epsilon^{true} &\geq \epsilon^{FP} \\ r_t - \beta^{true}r_t^m + r_t^f(\beta^{true} - 1) - \alpha &\geq r_t - \beta^{FP}r_t^m + r_t^f(\beta^{FP} - 1) - \alpha^{FP} \\ r_t - \beta^{true}r_t^m + r_t^f(\beta^{true} - 1) &\geq r_t - r_t^m \\ -\beta^{true} \times r_t^m + r_t^f(\beta^{true} - 1) &\geq -r_t^m \\ r_t^m(1 - \beta^{true}) - r_t^f(1 - \beta^{true}) - \alpha^{true} &\geq 0 \\ (r_t^m - r_t^f)(1 - \beta^{true}) - \alpha^{true} &\geq 0 \end{aligned}$$

Data description. pcecc96: This variable denotes real (in terms of billions chained 2005 USD) personal consumption expenditures and is available from the Federal Reserve Bank of St. Louis Economic Research at [research.stlouisfed.org/fred2/series / PCECC96](https://research.stlouisfed.org/fred2/series/PCECC96) with identifier BEA Account Code DPCERX1, based on the U.S. Bureau of Economic Analysis data. Quarterly data is available from 1947q1 onwards and is the annualized seasonal adjusted rate. All data is in logs. It is intended to be used in fiscal VARs in chapter 2.4.1 to proxy for the reaction of households spending behavior after an anticipation shock. I denote this series as c .

comprnfb: This variable denotes Nonfarm Business Sector Real Compensation Per Hour and hence proxies for wages, while excluding farm business sector dynamics. It is available from the Federal Reserve Bank of St. Louis Economic Research at research.stlouisfed.org/fred2/series/COMPRNFB/. Data is available as an index with 2009 being equal to 100 as the seasonally adjusted rate. All data is in logs. It is intended to be used in fiscal VARs in chapter 2.4.1 as a proxy for the reaction of real wages after an anticipation shock. I denote this series as w .

fgrecpt: This variable denotes Federal Government Current Receipts (in terms of billions chained 2005 USD) and is available from the Federal Reserve Bank of St. Louis Economic Research at [research.stlouisfed.org/fred2/series/ fgrecpt](https://research.stlouisfed.org/fred2/series/fgrecpt) with BEA Account Code W005RC1. Quarterly data is available from 1947q1 onwards and is the annualized seasonal adjusted rate. All data is in logs. It is intended to be used in fiscal VARs in chapter 2.4.1 to proxy for the reaction government revenues to changes after an anticipation shock. I denote this series as t .

gdpdef: This variable denotes the Gross Domestic Product Implicit Price Deflator in 2005 USD, available from the Federal Reserve Bank of St. Louis Economic Research at research.stlouisfed.org/fred2/series/gdpdef with BEA Account Code A191RD3. Quarterly data is available from 1947q1 onwards and is seasonally adjusted. It is intended to deflate Ramey data for use in fiscal VARs in chapter 2.4.1. This series only indirectly appears in the estimation and does therefore not receive an identifier.

tb3ms: This variable denotes the three-month t-bill: Secondary Market Rate, available from the Federal Reserve Bank of St. Louis Economic Research at research.stlouisfed.org/fred2/series/tb3ms. Quarterly data is created from monthly data with the end of period data, available from 1934m1 onwards and is not seasonally adjusted. It is intended for use in fiscal VARs in chapter 2.4.1 to capture issues related to monetary policy. I denote this series as $tb3$.

rgdp: This variable denotes real (in terms of billions chained 2005 USD) gross domestic product and is available for download at Valery Ramey's website at [econ-](https://www.valeryramey.com/)

web.ucsd.edu/vramey/research.html data and based on Ramey (2011a). Quarterly data is available from 1947q1 and seasonally adjusted. All data is in logs. It is intended to be used in fiscal VARs in chapter 2.4.1 to proxy for the reaction of real gross domestic product to changes after an anticipation shock. I denote this series as gdp .

rdef: This variable denotes real (in terms of billions chained 2005 USD) government military spending and is available for download at Valery Ramey's website at econweb.ucsd.edu/vramey/research.html data and based on Ramey (2011a). Quarterly data is available from 1947q1 and seasonally adjusted. All data is in logs. It is intended to be used in fiscal VARs in chapter 2.4.1 to proxy for the reaction of real defense spending to changes after an anticipation shock. I denote this series as g^m .

rgov: This variable denotes real (in terms of billions chained 2005 USD) government total consumption expenditures and is available for download at Valery Ramey's website at econweb.ucsd.edu/vramey/research.html data and based on Ramey (2011a). Quarterly data is available from 1947q1 and seasonally adjusted. All data is in logs. It is intended to be used in fiscal VARs in chapter 2.4.1 to proxy for the reaction of defense spending to changes after an anticipation shock. I denote this series as g .

tothours: This variable denotes total household based hours worked and is available for download at Valery Ramey's website at econweb.ucsd.edu/vramey/research.html data and based on Ramey (2011a). Quarterly data is available from 1947q1 and it is a Household based measure and intended to be used in fiscal VARs in chapter 2.4.1 to proxy for the reaction of hours worked to changes after an anticipation shock. I denote this series as h .

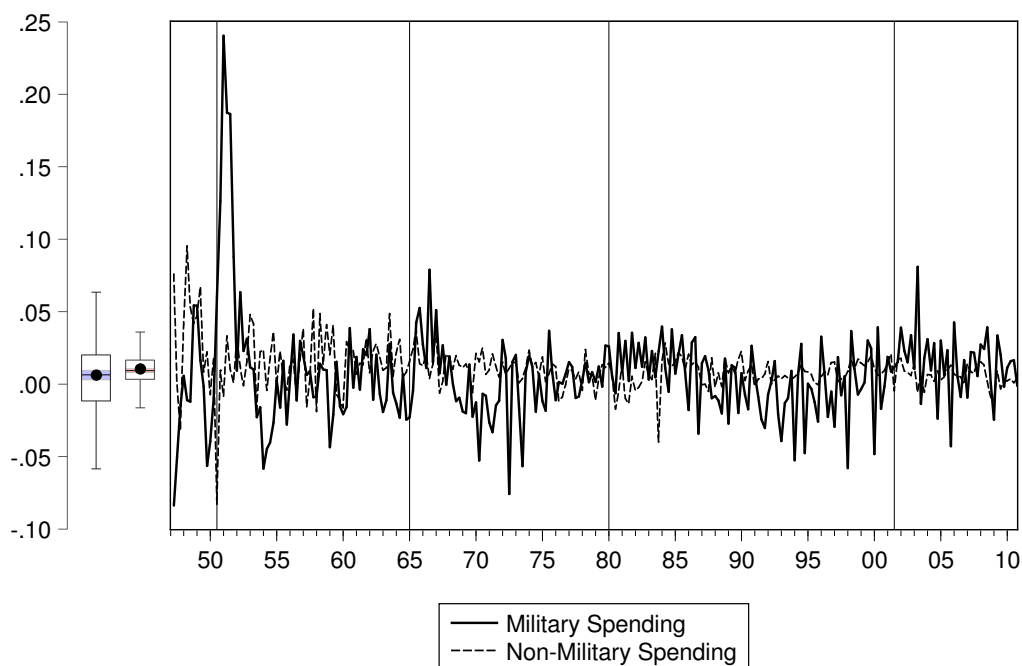
amtbr: This variable denotes Barro - Redlick Average Marginal Tax Rate and is available for download at Valery Ramey's website at econweb.ucsd.edu/vramey. Quarterly data is available from 1947q1 and seasonally adjusted. It is intended to be used in fiscal VARs in chapter 2.4.1 to proxy for the reaction of the Marginal Tax Rate to changes after an anticipation shock. I denote this series as mtr .

totpop: This variable denotes total population and is available for download at Valery Ramey's website at econweb.ucsd.edu/vramey/research.html data and based on Ramey (2011a). Quarterly data is available from 1947q1 and seasonally adjusted. It is intended to be used to create per capita series for some of the variables in chapter 2.4.1. This variable does not receive an identifier.

top3xsret: This variable denotes cumulative abnormal military contractors' returns, available for download at Valery Ramey's website at econweb.ucsd.edu/vramey/research.html data and based on Ramey (2011a). Quarterly data is available

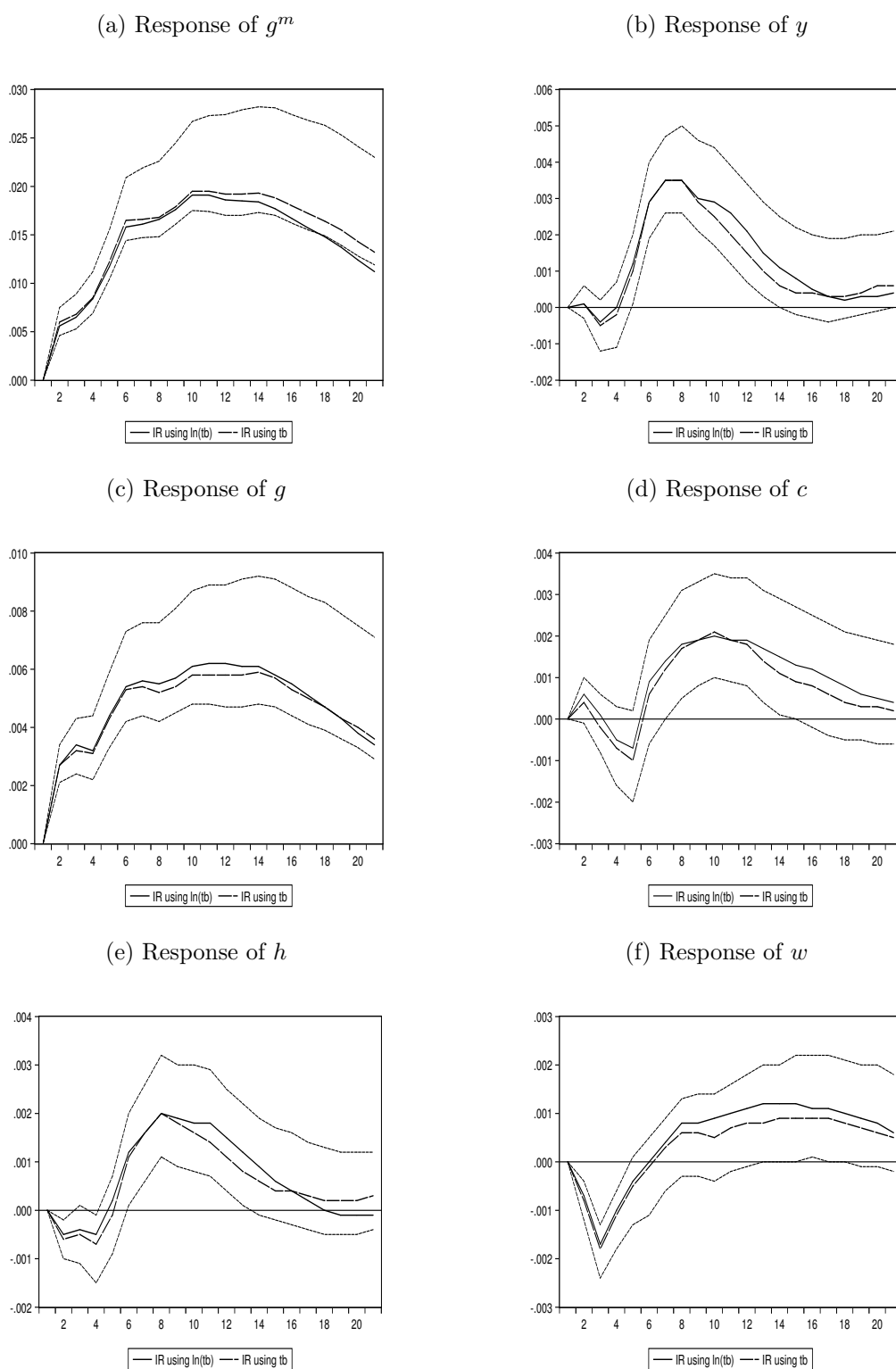
from 1947q1. This measure is the exact series, used in Fisher and Peters (2010) and constructed as described in the data section of my paper. I create a cumulative returns series for the reasons described in the data section as well. It is intended to be used as anticipation shock in chapter 2.4.1. This variable is denoted as *CER*.

Figure 10: Defense and non-defense spending



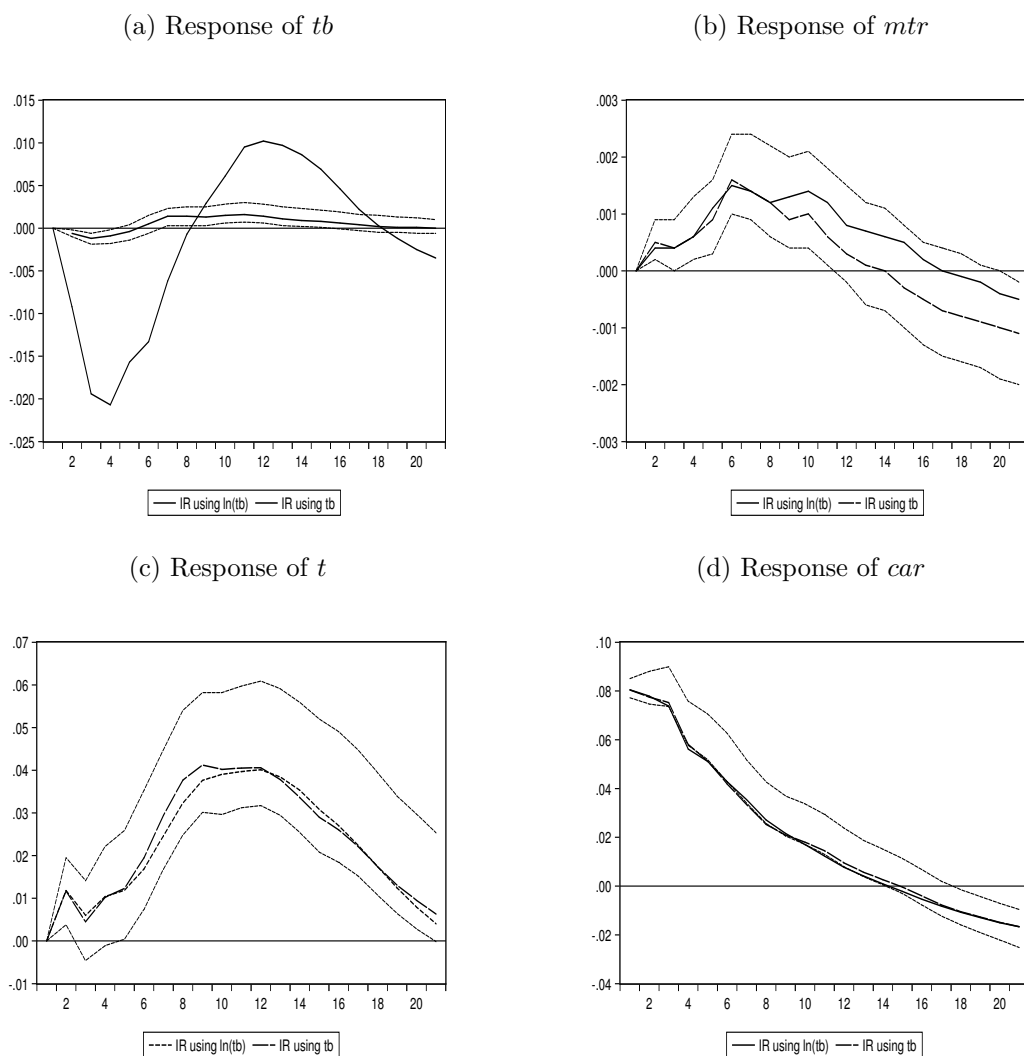
Notes: The solid line denotes log changes in real military per capita spending, whereas the dotted line is the log change in real non-military per capita spending with the respective box plots. All data 1947q1 to 2010q4. Vertical lines are the Korean War, Vietnam War, Russian invasion of Afghanistan and the 2001 WTC attacks, according to Ramey and Shapiro (1998).

Figure 11: IR t-bill in levels I



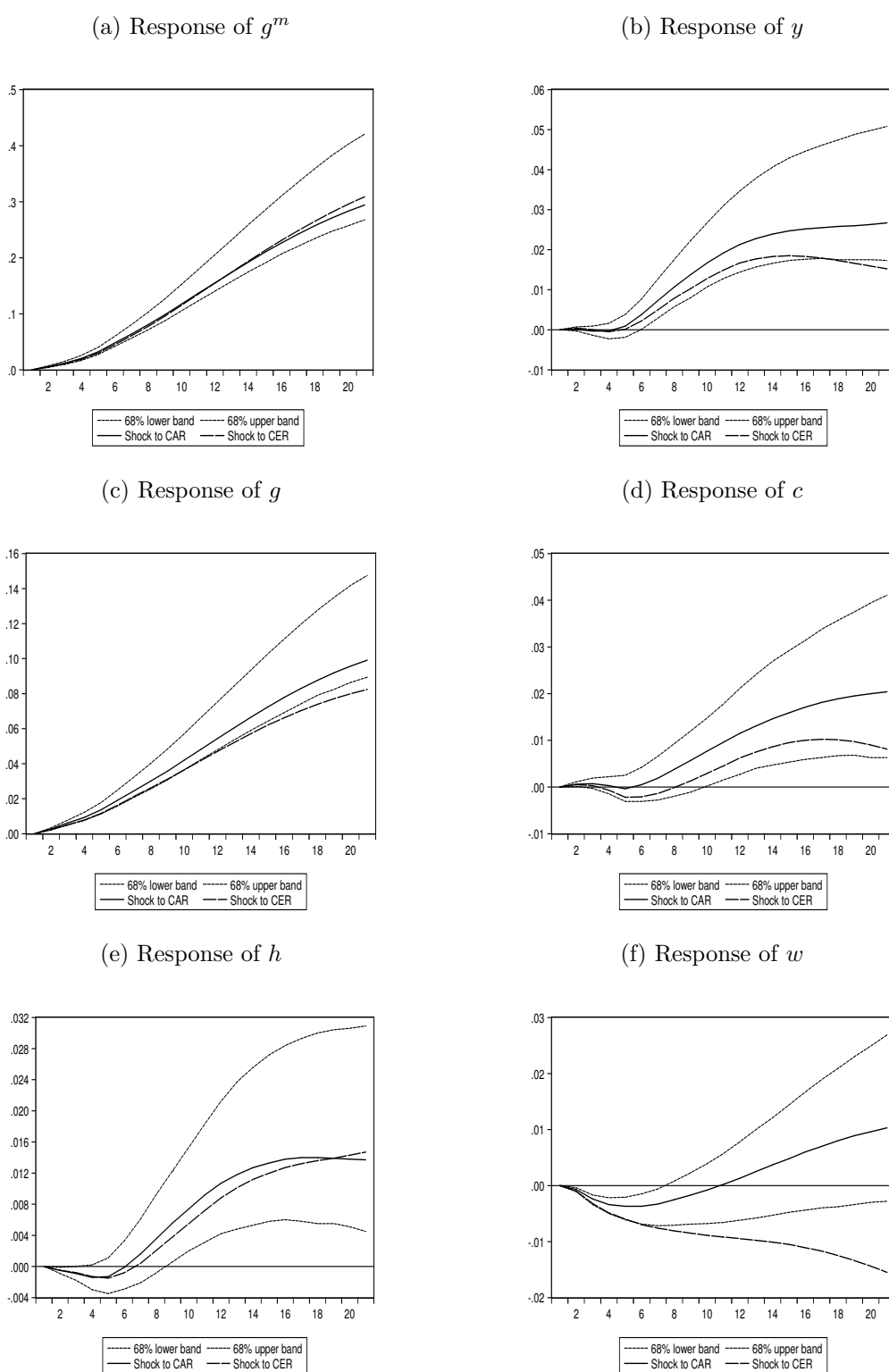
Notes: Solid: IRs the logs of real per capita; a) military spending; b) GDP; c) total government spending; d) personal consumption, and e) total hours and f) real per hour wages, to a one std deviation shock to CAR. Short dashed: respective 68% Hall CI (10000 reps), both t-bill in logs. Bold dashed: point estimates after one std deviation CAR shock, t-bill in levels. All 1957q3-2007q4

Figure 12: IR t-bill in levels II



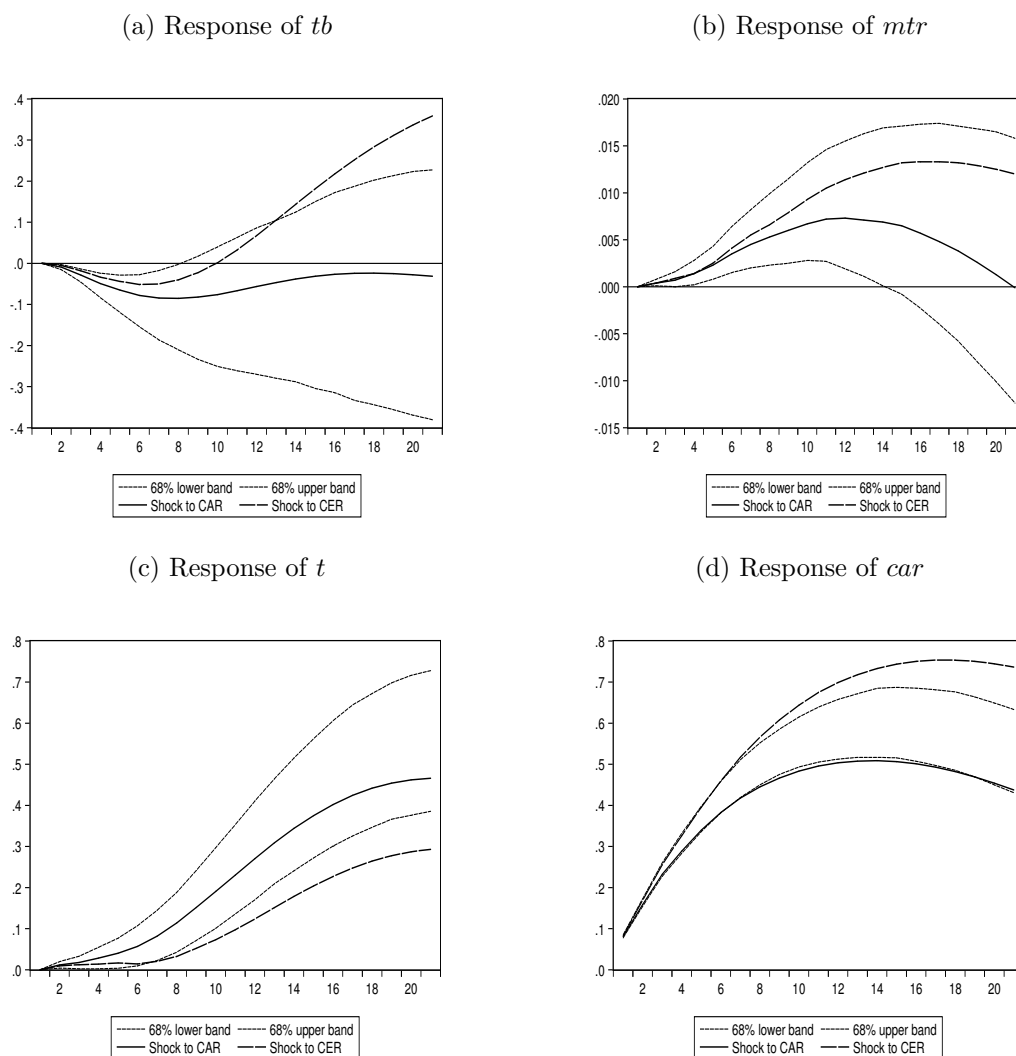
Notes: Solid: IRs of the a) nominal three month t-bill b) Barrow-Redlick average marginal tax rate c) logs of tax revenues d) cumulative returns, to a one std deviation shock to CAR. Short dashed: respective 68% Hall CI (10.000 reps). Bold dashed: point estimates after one std deviation CAR shock, t-bill in levels. All 1957q3-2007q4

Figure 13: Cumulative IR for abnormal and excess returns I



Notes: Solid: Cumulative IRs the logs of real per capita a) military spending; b) GDP; c) total government spending; d) personal consumption, and e) total hours and f) real per hour wages, to a one std deviation shock to CAR. Short dashed: respective 68% Hall CI (10000 reps). Bold dashed: point estimates after one std deviation CER shock, 1957q3-2007q4.

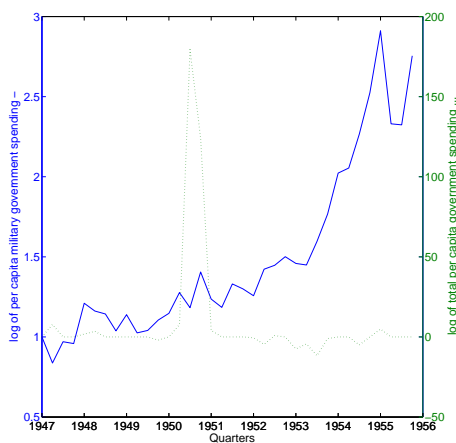
Figure 14: Cumulative IR for abnormal and excess returns II



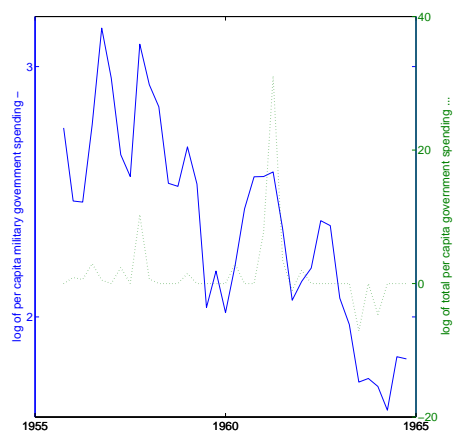
Notes: Solid: Cumulative IRs of the a) nominal three month t-bill b) Barrow-Redlick average marginal tax rate c) logs of tax revenues d) cumulative returns, to a one std deviation shock to CAR. Short dashed: respective 68% Hall CI (10000 reps). Bold dashed: point estimates of to a one std deviation shock to CER, 1957q3-2007q4.

Figure 15: Ramey news and abnormal returns

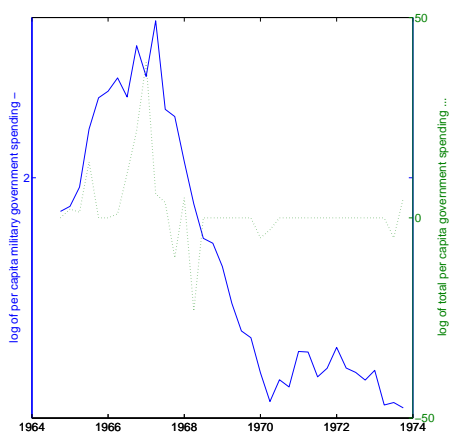
(a) Response of g^m



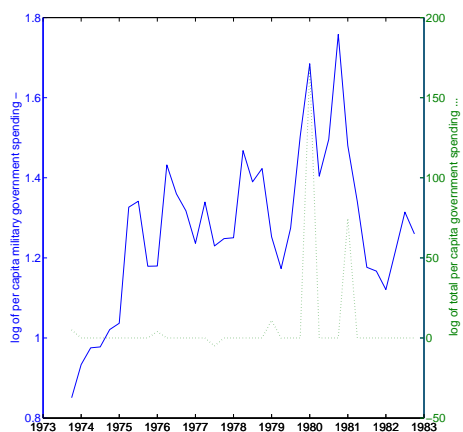
(b) Response of y



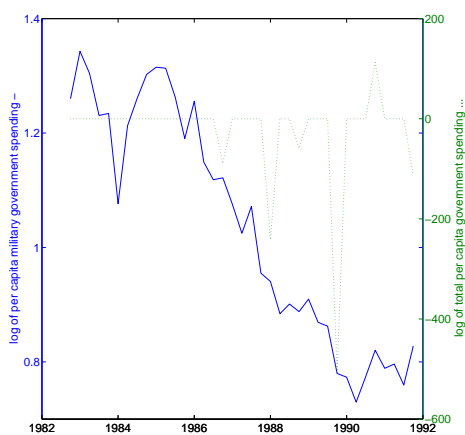
(c) Response of g



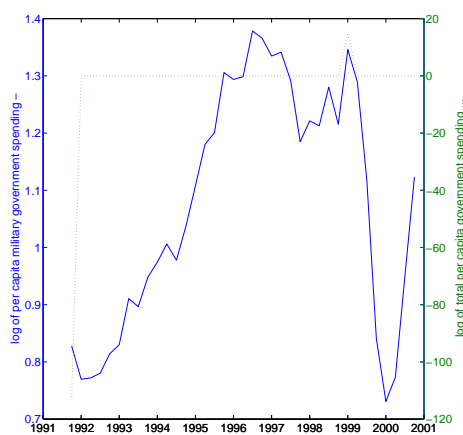
(d) Response of c



(e) Response of h

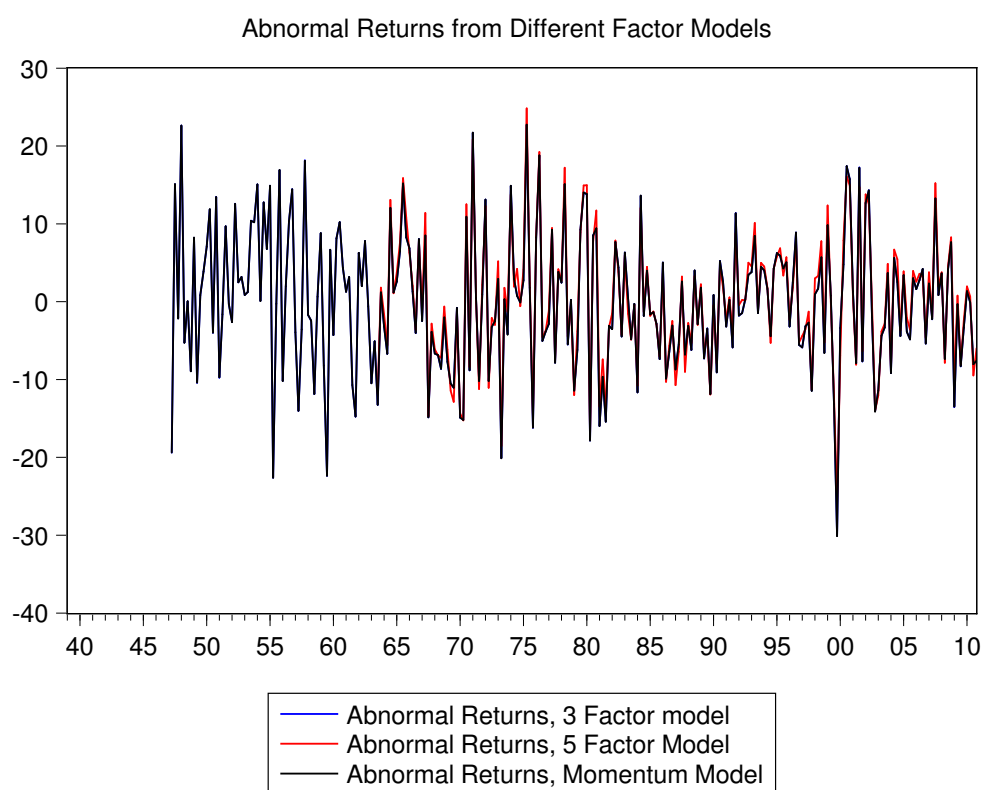


(f) Response of w



Notes: Ramey's news variable and abnormal returns shock, separated in six periods. The dotted line denotes the Ramey defense news variable, whereas the solid line denotes abnormal returns.

Figure 16: Factor models as proxy for future military spending



Notes: This figure compares abnormal returns shocks derived from the three-factor model used in the paper to a five-factor and momentum model as suggested by Fama and French (2012). The blue line denotes abnormal returns estimating in a three-factor model, the red in a five-factor, and the black line in a momentum model regression.

3 The procyclicality of consumer credit

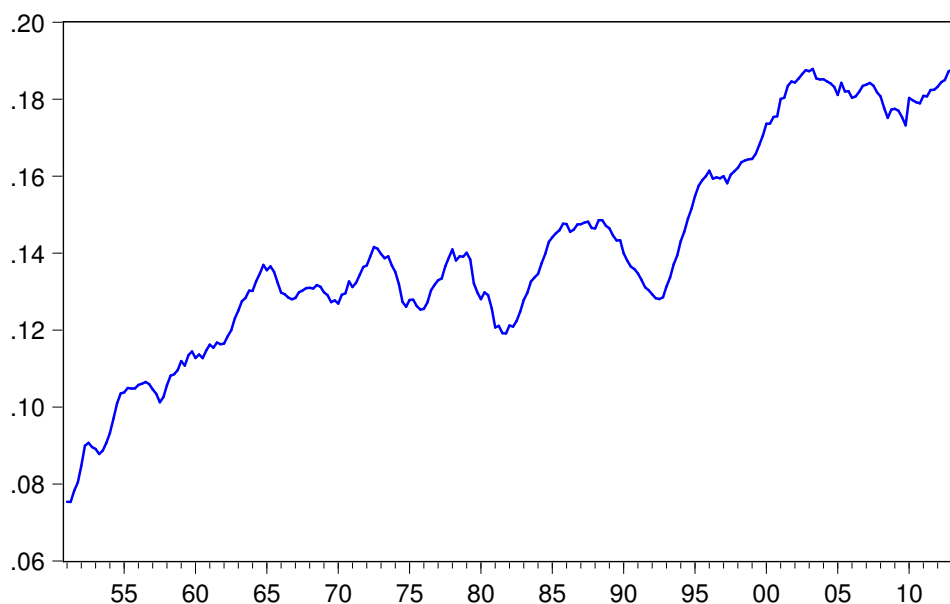
Coauthor: Mathias Klein

3.1 Introduction

In this paper, we study the business cycle properties of unsecured consumer credit conditional on tax cuts as defined by Romer and Romer (2010) and on total factor productivity shocks, as suggested by Basu et al. (2006).

The financial liberalization starting in the 1980s led to easier access to financial markets for U.S. consumers and the demand for unsecured consumer credit strongly increased thereafter. Between 1980 and 2015, the ratio of total consumer credit-to-output increased from 12% to 18%, as Figure 17 illustrates and also more than doubled as a ratio to consumption.

Figure 17: Unsecured consumer credit as a ratio to output



Notes: Real private consumer credit over real GDP for U.S. data from 1950 to 2015.

This financial integration prompted researchers to investigate the interactions between indebtedness of households and its effect on consumer expenditures and other macroeconomic variables (Andrés et al., 2015; Mian and Sufi, 2010).

Surprisingly, the business cycle properties of consumer credit are widely ignored as empirical, and theoretical models studying the relation between financial markets and the macroeconomy mainly focus on the dynamic effects of mortgages (e.g. Iacoviello (2005) or corporate debt e.g. (Bernanke et al. (1999)).

Household theory suggests that easier access to credit allows households to buffer against business cycle fluctuations and to keep their stream of consumption constant by saving in good times and borrowing in bad times. Consumption smoothing arguments hence suggest countercyclical credit movements after distortions. However, Ludvigson (1999) provides evidence using U.S. data that consumption growth is positively correlated with consumer credit growth. In addition to this descriptive view, Nakajima and Ríos-Rull (2014) find evidence that the unconditional correlation between credit, output and consumption is significantly positive. Their findings hence suggest that households use unsecured credit to finance additional consumption in expansions and do not save.

While these studies take unconditional correlations between credit and aggregates as empirical motivation, this paper focuses on responses conditional on two types of shocks having occurred, which is more informative. One reason is that unconditional correlations can result from the simultaneous interaction of several shocks. An unconditional correlation could be close to zero because two shocks hit the economy, one of which produces a positive and the other one, a negative correlation. Second, correlations, conditioned on a certain shock having occurred are more helpful in making inferences about the economic structure.

Analyzing conditional correlations, therefore, requires identification of shocks. We rely on the narrative tax measure taken from Romer and Romer (2010) and also on a measure of total factor productivity (TFP), identified by Basu et al. (2006). The choice for the first shock is motivated by its importance for the policy-making process, whereas the latter is one of the main drivers of the business cycle.

Given these measures of economic shocks, the question then arises how they should be econometrically treated to analyze the behavior of consumer credit. This paper compares two types of specifications, a standard VAR, and an exogenous VAR (VARX). Studies incorporating the TFP and the narrative tax measure often use a single equations approach or VARX method. However, in this paper we argue that a VAR, identified by Cholesky decomposition to the residual variance-covariance matrix and ordering either of the two shocks first to be more appropriate for the two shocks since Granger causality tests suggest neither of the shocks to be strictly exogenous. In a VAR setup, however, both shocks are treated as contemporaneously exogenous so that the system can consistently be estimated with OLS.

We find a conditional comovement of consumer credit, output, and consumption, given the two innovations. This suggests that agents finance additional consumption by borrowing via unsecured credit, which refutes the consumption smoothing argument.

The remainder of this paper is structured as follows: The next section introduces our data, describes estimation methods and presents the properties of the main macro aggregates, and consumer credit. We then discuss Granger causality of the two externally specified shocks. The following section then shows our results and includes a brief discussion. A final chapter concludes.

3.2 Data, Granger causality and estimation

This section presents our data and estimation approach to finding the conditional response of consumer credit on tax and TFP innovations.

3.2.1 Data

We use quarterly U.S. data from 1966q1 to 2014q4 for the TFP shock and up to 2007q4 for the tax shock, limited by data availability of the tax series. Our sample hence contains 168 data points for the tax shock and 196 observations in total for the TFP innovation.

The choice of variables and their ordering is based on Mertens and Ravn (2011), extended to control for monetary policy. The vector of variables includes, in this particular ordering, real per capita output (y_t), total personal per capita consumption expenditures on non-durables (c_t^n) and durables (c_t^d), hours worked (h_t), the real interest rate (r_t) and real per capita unsecured consumer credit (cc_t). Except for the real interest rate, all data is seasonally adjusted and in logs.

Actual tax changes are a linear combination of exogenous and endogenous tax changes that co-move with the business cycle which biases the estimation of the conditional response of consumer credit given tax changes or TFP shocks. To cope with this issue, we use the Romer and Romer (2010) measure for present discounted values of exogenous tax liability changes, denoted as τ_t . The authors take a narrative approach by reviewing the economic reports of the president and reports of congressional committees to record the timing and the size of major tax policy changes. The authors incorporate only those events that are motivated by past actions, beliefs about fairness or philosophy and that are unlikely to be correlated with other factors that co-move with the business cycle. Romer and Romer (2010) scale the tax shock by dividing the present discounted value of tax changes by output.

To obtain the conditional response of consumer credit, given TFP innovations we incorporate the Basu et al. (2006) TFP measure, which we denote as z_t . The authors base their estimation on weighted, industry specific Solow residuals, controlling for utilization, non-constant returns and attaching a weight to industry-specific technology, in an IV estimation approach. This procedure ensures the shock to be contemporaneously uncorrelated with other shocks in the economy.

We then create an index with 1966q1=100 of the TFP measure and use the log of this series.

Data, except the shocks and hours, can be obtained from FRED database. Data definitions and sources are given in Table 1, and further information can be found in the appendix in Table 3.

Table 1: Data sources & definitions

y	output	Log of per capita nominal gross domestic product divided by the GDP deflator
c^n	non durable consumption	Log of per capita (personal consumption expenditures on non durables plus personal expenditures on services) divided by each individual price deflator
c^d	durable cons purchases	Log of per capita personal purchases of durable consumption goods divided by its deflator
r	real interest rate	Difference between three-month month t-bill and annualized CPI quarter to quarter inflation
cc	consumer credit	Log of per capita total unsecured consumer debt, divided by the consumer price index
N	population	Total population aged 16 and over, used to create per capita values
h	hours	Product of hours per worker and civilian non-farm employment divided by population
τ	tax Shock	Romer and Romer (2010) exogenous tax shock
z	technology shock	Basu et al. (2006) exogenous technology shock

Notes: Sources and definition of variables for estimation in VARs and exogenous VARs. Full time series-specific information and sources can be found in the appendix in Table 3. The first column denotes the code for the respective variable. FRED data available at research.stlouisfed.org.

3.2.2 Identification, choice of model and estimation

The estimation method depends on the nature of the externally identified shock measures. Estimating VARs or VARX are possible. The VARX takes the following form

$$X_t = A(L)X_{t-1} + B(L)s_t + u_t, \quad (7)$$

in which $X_t = [y_t, c_t^n, c_t^d, h_t, r_t, cc_t]'$ with the shock being either the tax shock or the TFP innovation so that $s_t \in [\tau_t, z_t]$. $A(L)$ is a lag polynomial of order four and $B(L)$ a lag polynomial of order five (suggested by Schwarz information criterion, Schwarz (1978)). u_t denotes all other shocks that are uncorrelated with either the tax shock or the TFP measure. The VARX estimation includes a constant term and deterministic trend and is estimated with OLS.

Estimating VARXs assumes that the two shock series are strictly exogenous so that they are not influenced by past observations in X .

Furthermore, the VAR takes the following form and is closely related to the VARX estimation

$$X_t = C(L)X_{t-1} + u_t, \quad (8)$$

where variable definitions are identical to the ones given for Equation 7, and $X_t = [s_t, y_t, c_t^n, c_t^d, h_t, r_t, cc_t]'$, s_t is the respective shock and $C(L)$ is a lag polynomial of order four. The residual variance-covariance is identified by Cholesky decomposition. The VAR and VARX can consistently be estimating with OLS.

In the estimation in Equation 8, the shock which is ordered first is contemporaneously unaffected by the other variables, but not vice versa so that the VAR is recursive. After the parameters are estimated, impulse responses can be created to find the impact of the two shocks on consumer credit.

The crucial requirement when estimating both methods by OLS is that the Romer and Romer (2010) tax measure and the Basu et al. (2006) TFP series are contemporaneously uncorrelated with other shocks in the economy. By construction via the IV approach for the Basu et al. (2006) series, this is the case.²⁸

Contemporaneous exogeneity holds true for both estimation methods, either in the VARX by restricting the feedback, or due to the recursivity of the VAR estimation.

Which method should be preferred when including the two externally specified shocks? Since both estimation methods are economically possible, we prefer the one which impose less restriction on the estimation. We, therefore, chose to estimate the VAR and test whether both shocks are not only contemporaneously, but also strictly exogenous. If the latter is the case, estimating a VARX and VAR should lead to the same results.

Granger causality. Basu et al. (2006) report that their measure of technology is not Granger caused by past observations in macroaggregates, so it cannot be predicted using a sample ranging from 1949 to 1996. Following these authors would

²⁸We do not question the validity of the IV approach taken by Basu et al. (2006) here.

suggest estimating VARX rather than VARs since the TFP measure would be strictly exogenous.

Concerning the tax measure, Romer and Romer (2010) together with Mertens and Ravn (2011) provide evidence on the strict exogeneity of the tax measure. The latter study provides evidence for this by recoding the original series measuring tax revenue changes over GDP, by assigning these changes to three groups: A -1 is assigned to negative tax innovations, a 1 is assigned to positive innovations and periods without any tax changes receive a zero. Mertens and Ravn (2011) show that when rescaling the tax innovation like this, it is not Granger-caused by the lags of macroeconomic variables, visible in the last two columns in the Table 2. This would suggest taxes to be strictly exogenous as well and therefore to estimate a VARX.

To check if the series are strictly exogenous in our sample, we test for Granger causality of the original Romer and Romer (2010) measure. We then compare the results, with the findings of Mertens and Ravn (2011) for the tax shock and in Basu et al. (2006) for the TFP measure. Our results of this test are given in Table 2.

Table 2: Results of Granger causality test

	Index TFP				RR10				MR11 recoded	
	sample	obs	F-stat	p-value	sample	obs	F-stat	p-value	F-stat	p-value
y	66–14	191	2.395	0.042	66–07	164	1.386	0.870	1.041	0.387
c^n	66–14	191	1.723	0.146	66–07	164	1.858	0.073	1.078	0.368
c^d	66–14	191	3.673	0.006	66–07	164	1.243	0.292	0.280	0.890
h	66–14	191	1.661	0.160	66–07	164	1.151	0.189	1.354	0.252
r	66–14	191	0.967	0.426	66–07	164	2.148	0.062	0.186	0.945
cc	66–14	191	2.395	0.042	66–07	164	1.386	0.870	1.041	0.387

Notes: Null hypothesis: The variable does not Granger cause the technology / tax measure. Specification: four lags, linearly detrended data for output (y), non durable consumption (c^n), durable consumption (c^d), hours worked (h), real interest rate (r), consumer credit (cc) and TFP index as used in VARs. RR10 denotes the tax series created by Romer and Romer (2010) and Mertens and Ravn (2011) is the redefined Mertens and Ravn (2011) taxes series.

The results in Table 2 suggest that the strict exogeneity of the two shock series is not given. The lags of output and durable consumption have a significant impact on the TFP series. Additionally, the tax measure is Granger caused by the real interest rate and non-durable consumption. These estimates are contradictory to the results reported in Basu et al. (2006) and the Mertens and Ravn (2011) transformation. As we are interested in the actual size of the tax changes and not simply the direction of the tax innovation, as redefined by Mertens and Ravn (2011), using the original Romer and Romer (2010) tax series seems appropriate for our set-up.

Given these results, treating both shock measures as strictly exogenous, impulse responses recovered from and the inference drawn from estimating VARXs may

be misleading. We take this fact into account by ordering the shocks first in two separate VARs (one for taxes, ones for the TFP measure) and by applying a Cholesky decomposition to the reduced form residual variance-covariance matrix. In this case, the tax and the TFP measure are contemporaneously uncorrelated with other variables. To quantify the bias when treating both shocks as strictly exogenous, we also estimate VARXs and compare impulse responses to the ones obtained from the VARs estimation.

3.3 Estimation results

We produce impulse responses conditional on the two shocks for the VARs and also for the VARXs. The size of the technology shock is equivalent to an increase of one percent in the Basu et al. (2006) Solow residual. The Romer and Romer (2010) structural tax shock is equivalent to a reduction of discounted total tax revenues over GDP of one percentage point. We report impulse responses together with 68 percent confidence bands. The dotted blue lines denote responses estimated with VAR, whereas triangled red responses result from the estimation of the VARX.

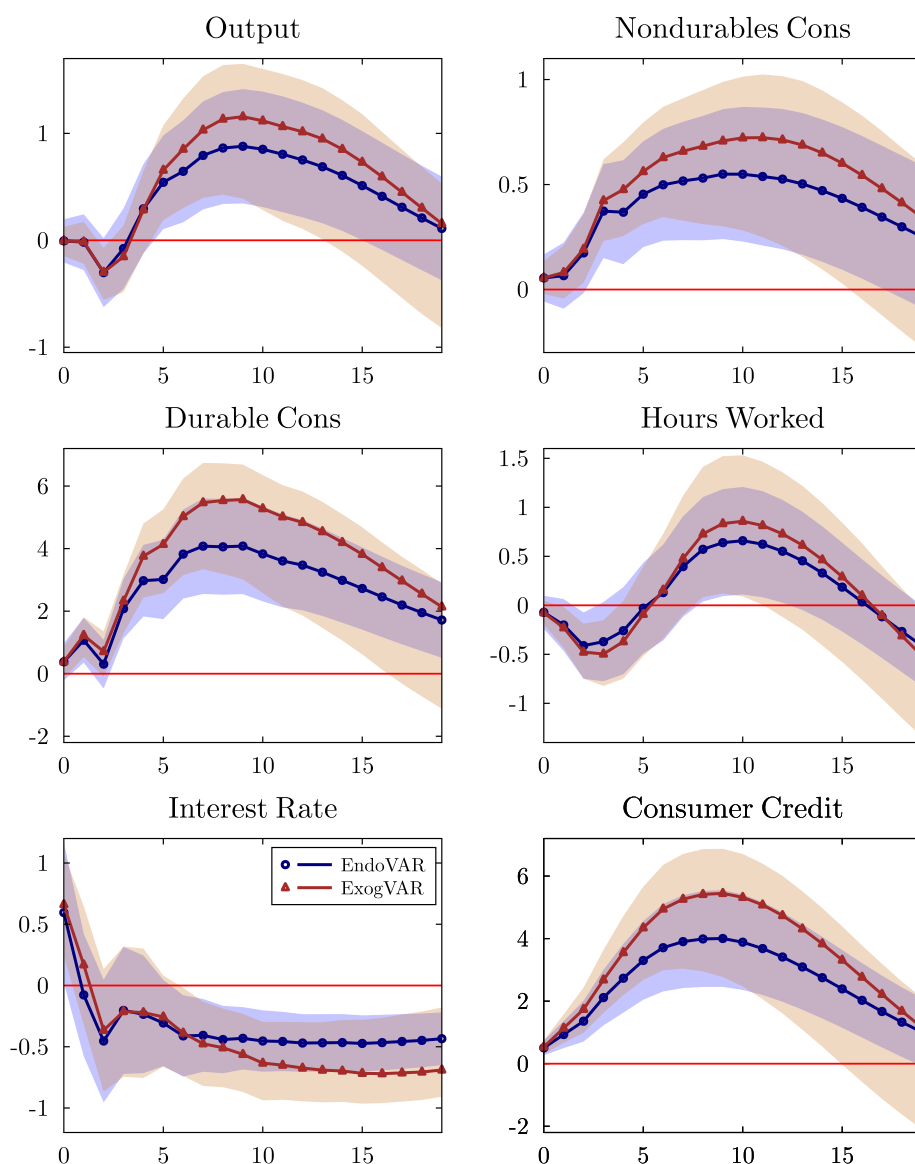
The response to a one percentage point decrease in discounted tax revenues over GDP are given in Figure 18, and the responses to a one percent increase in technology are reported in Figure 19.

Figure 18 shows that a tax cut induces an expansion in the economy. The economic expansion is persistent, lasting for roughly four years, with a peak increase of 1% of output before the economy returns to pre-shock equilibrium. Output, non-durable and durable consumption, as well as consumer credit significantly, increase in a hump-shaped manner, peaking around ten quarters after the shock. The strongest reactions are visible in the response of durable consumption and consumer credit so that credit increases between four and five percent, similar to the increase in durable consumption. The comovement of consumer credit with the real variables refutes consumption smoothing so that households choose to finance additional consumption by accumulating more credit. Therefore, a significant part of the boom is credit-financed.

The smallest reaction is visible in non-durable consumption and the real interest rate, the latter with a decrease of around 0.5% percentage points. In contrast to the significant increase of all variables, hours drop slightly on impact before rising in a hump-shaped manner as well.

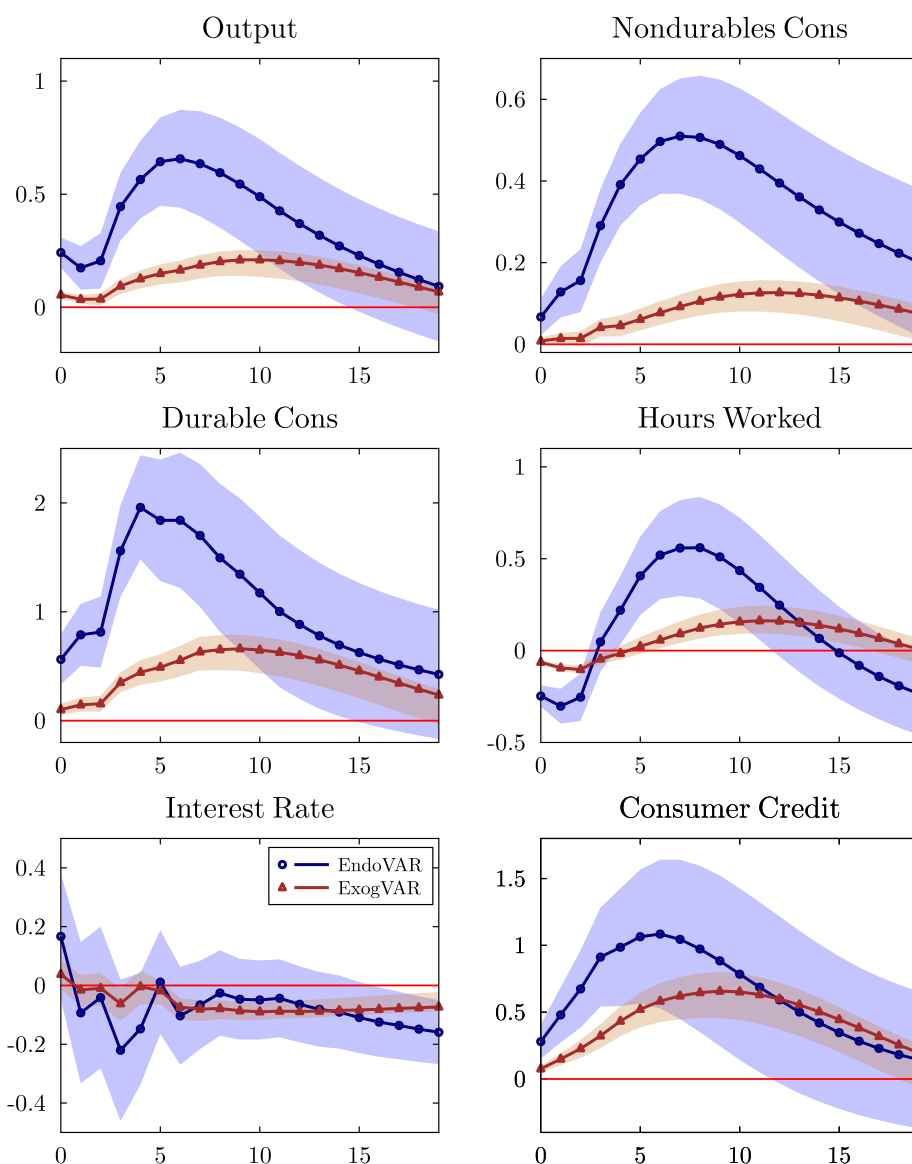
Impulse responses estimated from VARs and VARXs conditional on the tax shock are similar. However, reactions triggered by the shock in VARXs are slightly higher,

Figure 18: Impulse responses conditional on tax cuts



Notes: This figure depicts VAR and VARX impulse responses for a one percentage point decrease in discounted change in tax revenues over GDP, as measured by Romer and Romer (2010) $X_t = [\tau_t, y_t, c_t^p, c_t^d, h_t, r_t, cc_t]'$. Dotted blue lines denote responses estimated with VAR, whereas triangled red responses result from the estimation of the VARX. Reduced form residual variance-covariance matrices are Cholesky decomposed. Red and blue shaded areas are 68% bootstrapped confidence intervals with 10.000 replications, and the dark area denotes the overlapping of both confidence bands.

Figure 19: Impulse responses conditional on TFP shocks



Notes: This figure depicts VAR and VARX impulse responses for a one percent increase in TFP, as measured by the Basu et al. (2006). $X_t = [z_t, y_t, c_t^n, c_t^d, h_t, r_t, cc_t]'$. Dotted blue lines denote responses estimated with VARs, whereas triangled red responses result from the estimation of the VARX. Reduced form residual variance-covariance matrices are Cholesky decomposed. Red and blue shaded areas are 68% bootstrapped confidence intervals with 10,000 replications, and the dark area denotes the overlapping of both confidence bands.

but within each others confidence bands. Therefore, the choice of VAR method does not seem critical as confidence bands overlap in almost all post-shock periods.

Figure 19 depicts impulse responses for an increase in TFP. The shock induces a similar pattern of comovement between the variables as for the tax shock, inducing a temporary economic expansion. All variables, except hours, increase in a hump-shaped manner and peak around ten quarters after the shock. The strongest reaction can be seen for durable consumption with a rise of 2% percent and consumer credit with an increase between 0.5% and 1%. Following the TFP shock, households significantly decrease hours worked on impact and increase them after that for 8 quarters before returning to pre-shock values.

The choice of VAR estimation method here, however, is crucial, since confidence bands do not overlap for more than half of the periods. Estimating VARXs hence significantly underestimate the impact of TFP shocks on the other variables. When estimating impulse responses from VARXs, output only increases half as strong with a peak of 0.2% compared to the VAR estimation. Both types of consumption increase four and five times stronger, when estimating VARs, so that confidence bands of the VAR and VARX generated impulse responses do not overlap.

The initiated expansion, however, still causes consumer credit to increase between 0.6% and 1%, depending on the choice of VAR. The reaction is 1/3 lower when estimating VARXs compared to the VAR. The response of the real interest rate is not significantly different from zero.

In conclusion, both shocks initiate a boom that causes a significant increase in all variables, except hours. Impulse responses estimated from VARX and VARS are similar for tax shocks while deviating significantly for the TFP measure. Therefore the dynamic lag structure should not be ignored and hence estimating VARs seem more suitable using the two externally specified shock series.

3.3.1 Robustness

This subsection shows that our results are robust to a number of modifications.

First, the number of lags in our estimation does not have an impact on the qualitative responses of the variables, so that the results remain similar as visible in Figure 24 and Figure 25 in the appendix. The comovement of consumer credit and the other variables remains intact so that impulse responses when incorporating 3 and 5 lags are within the confidence bands of the baseline estimation. Using additional lags in the estimation causes impulse responses to be slightly higher.

Second, we include wages²⁹ ordered after consumption in our estimations. Results for both shocks can be found in Figure 20 and 21. Tax reductions cause wages to increase persistently for more than 20 quarters, peaking at 0.6 % increase, before converging to the pre-shock level. The responses of the other variables are robust to including wages in the estimation so that the results are slightly larger, but close to our baseline results.

The TFP shock causes wages to increase on impact while hours show a significantly negative effect, as can be seen in Figure 21. However, the qualitative comovement between the variables and the strong increase in consumer credit remains.

Third, we substitute wages for investment.³⁰ These results can be found in Figure 22 and Figure 23. After the tax cut, investment drops on impact, before increasing and peaking at 2.1%. Including investment in the estimation causes impulse responses to be slightly lower. Nevertheless, responses are still within the confidence bands of the baseline estimation.

Finally, we check whether the economic expansion is characterized by lower unemployment. We, therefore, include unemployment in our estimation ordered before the interest rate. As a result, unemployment decreases after both shocks. Consequently, unemployment is lower conditional on tax cuts and TFP innovations.

Changing the order of the system for the variables described does not alter our results so that the comovement of all variables remains obvious.

In conclusion, the result of a positive response of consumer credit conditional on tax cuts and TFP shocks is robust to a number of modifications. A comovement with output and consumption is visible when including wages or investment in the estimation or extending or shortening the number of lags in the estimation.

3.4 Discussion of the results

Our result of a strong conditional positive comovement between output, consumption, and credit stands in contrast with consumption smoothing. Although the financial liberalization led to an easier access to consumer credit, our results suggest that consumers do not use consumer credit to insure against adverse shocks.

In order to use credit as an instrument to insure, households' need to have unlimited access to these credit instruments. Yet, in reality, some households are borrowing constrained.

²⁹Compensation of employees, per capita values. FRED data: W209RC1Q027SBEA.

³⁰Sum of Gross Private Domestic Investment and Gross government investment, seasonally adjusted, per capita values, deflated with individual deflators. Source: FRED GPDI and A782RC1Q027SBEA.

An explanation for our results, therefore, could be the presence of borrowing constrained households as shown by Iacoviello (2005) and also in the next chapter. Borrowing then would be limited by the amount of households' stock of collateral. The same could hold true for unsecured borrowing when households use their income to secure credit as shown by Ludvigson (1999). Thus, borrowing in terms of consumer credit does not lead to consumption smoothing one would expect if access to credit is constrained.

Conditional impulse responses after TFP shocks, estimated from VAR estimation are significantly larger than those obtained from the VARX estimation. Both confidence bands do not overlap for most of the periods. The TFP measure is Granger-caused by the other variables in our system, and thus we suggest to estimate VARs since the two measure are not strictly exogenous. Therefore, one would ignore dynamics and the impact of the other variables, when assuming TFP to be strictly exogenous.

Finally, our estimation could suffer from fiscal foresight when using taxes in the VAR estimation. Since our sample contains quarterly data, observing actual tax changes only in the period when they are implemented, would result in a misalignment of the movements in macroeconomic variables due to agents' prior knowledge. In this case, the information set would be smaller than the one of the agents. A regression of total private debt on tax income could then suffer from an omitted variable bias, and the resulting structural coefficient would be inconsistent. However, Romer and Romer (2010) discount all future tax changes to the date, when the bill was passed, so that the issue of fiscal foresight is reduced to a minimum.

3.5 Conclusion

In this paper, we investigate the conditional impact of tax cuts, as measured by Romer and Romer (2010) and TFP innovations, suggested by Basu et al. (2006), on unsecured consumer credit. Instead of investigating unconditional correlations between consumer credit and aggregates only, we proceed by estimating VARXs and VARs given the two shocks.

We find impulse responses conditional on tax innovations are similar for both estimation methods. However, when looking at the impact of TFP shocks, the VARX estimations underestimates impulse responses.

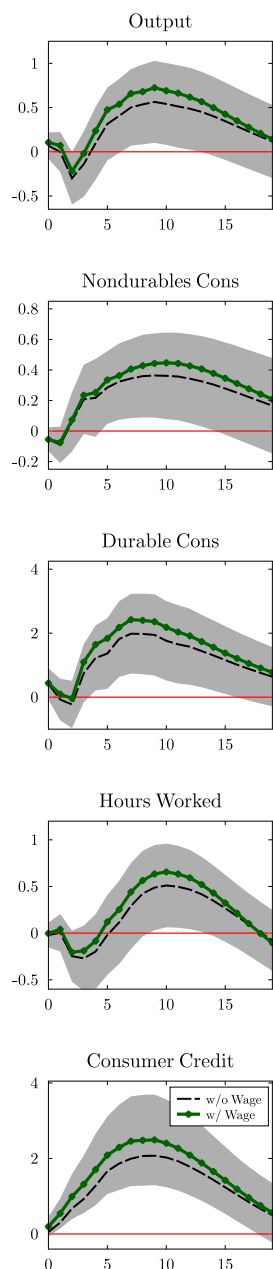
As we can show, the TFP measure and the tax innovation are Granger-caused by the other variables so that both shocks are not strictly exogenous. We, therefore, believe that VARs are the more general way to estimate the impact of TFP shocks and tax shocks, when using the Basu et al. (2006) and Romer and Romer (2010)

series. In such a recursive estimation, both shocks are contemporaneously exogenous so that estimating the system with OLS is consistent.

A consumption smoothing argument would suggest households to prefer a constant stream of consumption and to save in good times to insure against a drop in consumption in bad times. From our results, however, it is obvious that given TFP innovations and tax cuts, agents do not use consumer credit to smooth consumption and to insure against negative shocks. We show that a conditional comovement of all variables after the two shocks exist, and a boom in the economy is initiated as consumer credit, Output, durable purchases and non-durable consumption increase and seem to co-move after both shocks. Therefore, the increase in GDP is partly credit financed, given tax cuts and TFP shocks.

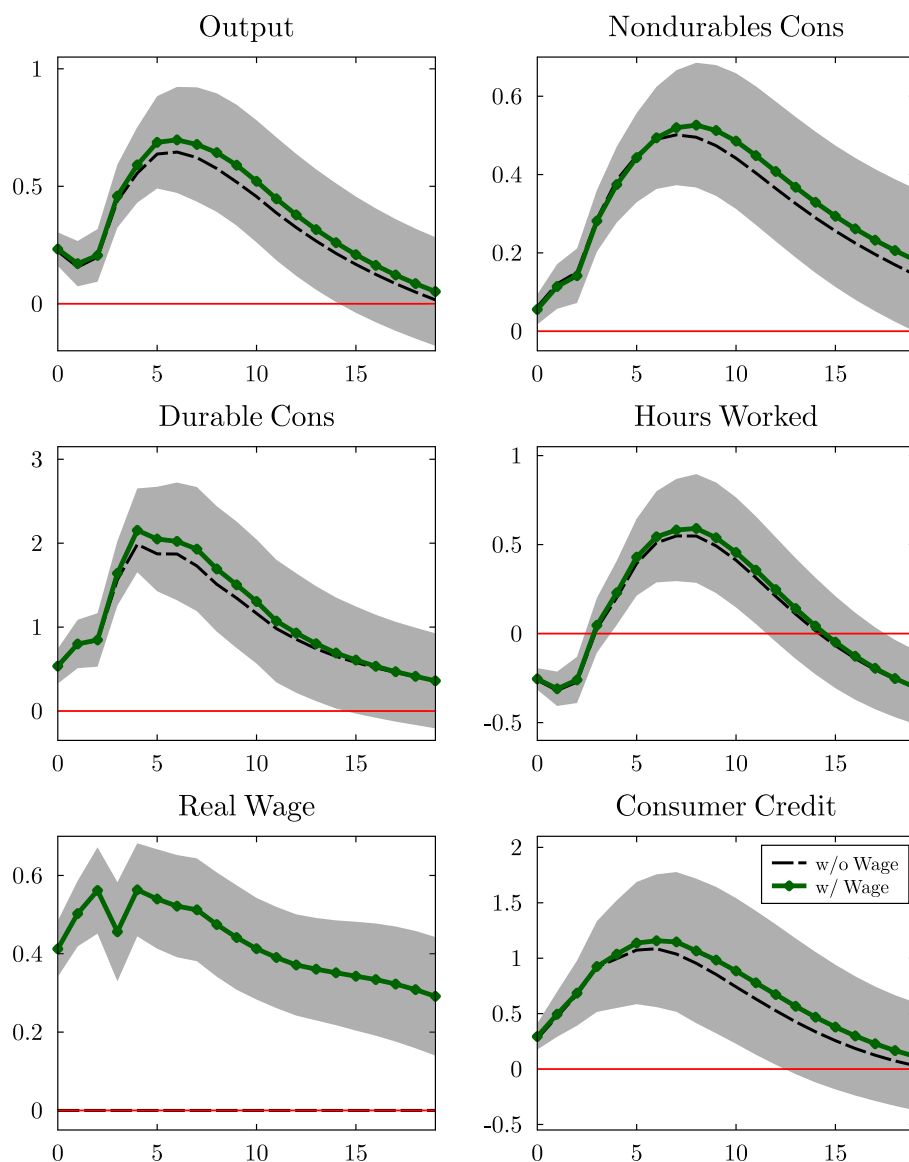
3.6 Appendix

Figure 20: IR for tax cuts, controlling for including wages



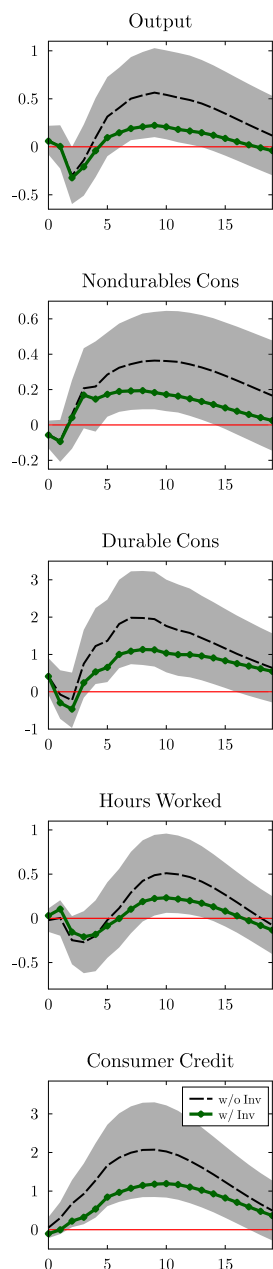
Notes: This figure depicts VAR impulse responses to a one percentage point decrease in discounted change in tax revenues over GDP, as measured by the Romer and Romer (2010) and $X_t = [\tau_t, y_t, c_t^n, c_t^d, w_t, h_t, r_t, cc_t]'$. The dashed line denotes our baseline estimation resulting from VARS along with 68% bootstrapped confidence intervals with 10,000 replications. Dotted lines denote responses estimated with the VAR when wages is included after the two consumption measures. For all estimations, reduced form residual variance-covariance matrices are Cholesky decomposed.

Figure 21: IR for TFP shocks, controlling for wages

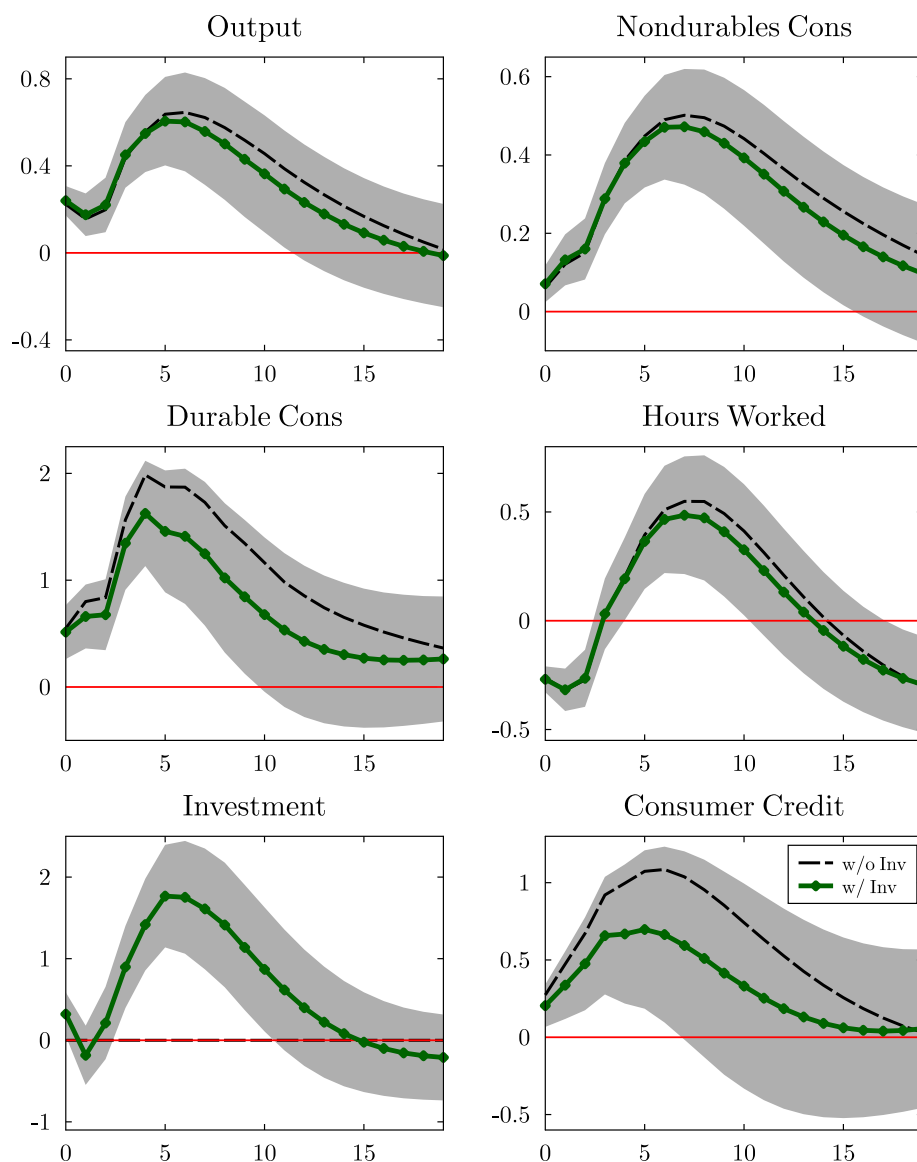


Notes: This figure depicts VAR impulse responses to a one percent increase in TFP, as measured by Basu et al. (2006) and $X_t = [\tau_t, y_t, c_t^n, c_t^d, w_t, h_t, r_t, cc_t]$. The dashed line denotes our baseline estimation resulting from VARS along with 68% bootstrapped confidence intervals with 10,000 replications. Dotted lines denote responses estimated with the VAR when wages is included wages after the two consumption measures. For all estimations, reduced form residual variance-covariance matrices are Cholesky decomposed.

Figure 22: IR for tax cuts, controlling for investment

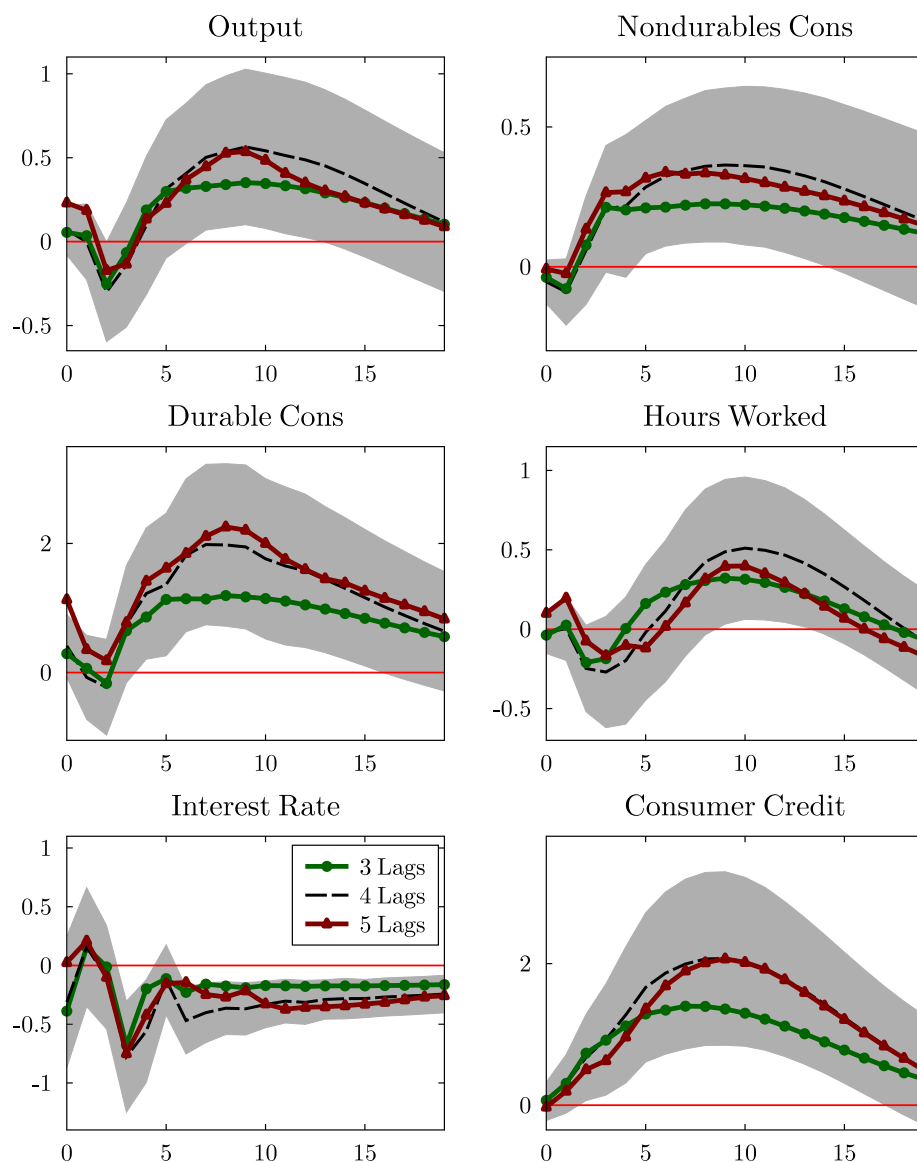


Notes: This figure depicts VAR impulse responses to a one percent decrease in discounted change in tax revenues over GDP, as measured by the Romer and Romer (2010) and $x_t = [\tau_t, y_t, c_t^n, c_t^d, h_t, I_t, r_t, cc_t]'$. The dashed line denotes our baseline estimation resulting from VARS along with 68% bootstrapped confidence intervals with 10,000 replications. Dotted lines denote responses estimated with the VAR when investment is included after hours. For all estimations, reduced form residual variance-covariance matrices are Cholesky decomposed.

Figure 23: TFP shock estimation including investment


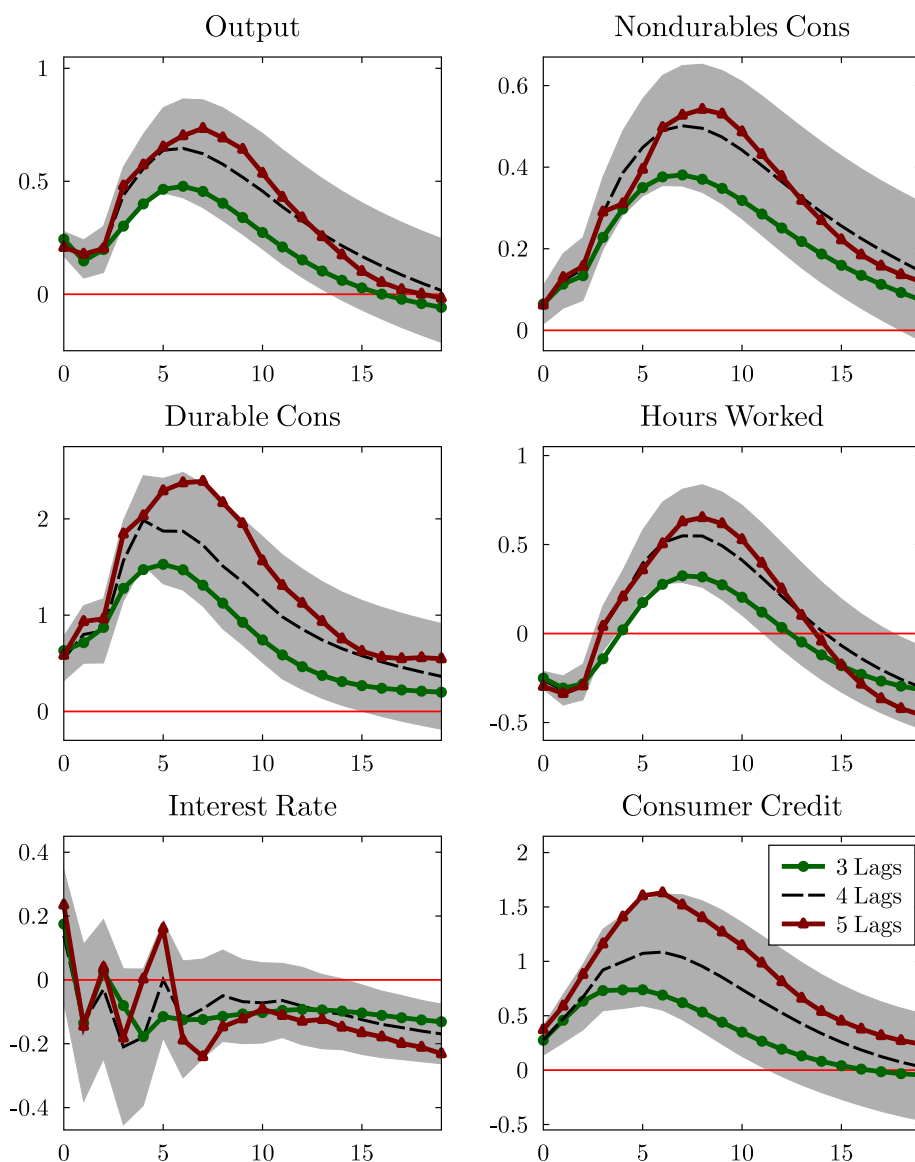
This figure depicts VAR impulse responses to a one percent increase in TFP, as measured by Basu et al. (2006) and $X_t = [\tau_t, y_t, c_t^n, I_t, c_t^d, h_t, r_t, cc_t]'$. The dashed line denotes our baseline estimation resulting from VARS along with 68% bootstrapped confidence intervals with 10,000 replications. Dotted lines denote responses estimated with the VAR when investment is included after hours. For all estimations, reduced form residual variance-covariance matrices are Cholesky decomposed.

Figure 24: IRs comparing lag lengths (Tax shock)



Notes: This figure depicts VAR impulse responses for a one percent decrease in discounted change in tax revenues over GDP, as measured by the Romer and Romer (2010) and $X_t = [z_t, y_t, c_t^n, c_t^d, h_t, r_t, cc_t]'$. Reduced form residuals' variance-covariance matrices are Cholesky decomposed. The shaded areas are 68% bootstrapped confidence intervals with 10,000 replications belonging to the baseline estimation with four lags.

Figure 25: IRs comparing lag lengths (TFP shock)



Notes: This figure depicts VAR impulse responses for a one percent increase in TFP, as measured by Basu et al. (2006) and $X_t = [z_t, y_t, c_t^n, c_t^d, h_t, r_t, cc_t]'$. Reduced form residuals' variance-covariance matrices are Cholesky decomposed. The shaded areas are 68% bootstrapped confidence intervals with 10,000 replications belonging to the baseline estimation with four lags.

Table 3: Full data sources & definitions

Var	Definition	code	Manipulation
y	Output	GDP	Gross Domestic Product, Seasonally Adjusted Annual Rate in Billions of Dollars
c^n	Non durable consumption	PCND	Personal Consumption Expenditures: nondurable goods, seasonally adjusted annual rate in Billions of Dollars
		PCESV	Personal Consumption Expenditures: Services, Seasonally Adjusted Annual Rate in Billions of Dollars
c^d	Durable purchases	PCDG	Personal Consumption Expenditures: Durable Goods, Seasonally Adjusted Annual Rate in Billions of Dollars
h	Hours worked	h	product of hours per worker and civilian non-farm employment divided by population, see Mertens and Ravn (2011), extended.
d	total private debt	CMDEBT	Households and Nonprofit Organizations; Credit Market Instruments; Liability, Level, seasonally adjusted in Billions of Dollars
N	Population	POP	Civilian Non institutional Population, Not Seasonally Adjusted, Thousands of Persons
tb	t-bill	TB3MS	3-Month Treasury Bill: Secondary Market Rate, Not Seasonally Adjusted, Percent
	GDP deflator	GDPDEF	Gross Domestic Product: Implicit Price Deflator, Index 2009=100
	non durables deflator	DNDGRG3 Q086SBEA	Personal consumption expenditures: Nondurable goods, Index 2009=100
		DSERRG3 Q086SBEA	Personal consumption expenditures: Services, Index 2009=100
	Durables deflator	DDURRG3 Q086SBEA	Personal consumption expenditures: Durable goods, Index 2009=100
	Investment deflator	GPDICTPI	Gross Private Domestic Investment: Chain-type Price Index, Index 2009=100
		B79RG3 Q086SBEA	Government gross investment: Federal: Nondefense (chain-type price index), Index 2009=100
	CPI deflator	CPIAUCSL	Consumer Price Index for All Urban Consumers: All Items

Notes: Sources and definition of variables for estimation in VAR and VARX, with respective deflators and FRED codes.

4 Technology shocks, tax cuts and their impact on household debt: Empirical evidence & theoretical explanation

Coauthors: Christopher Krause and Mathias Klein

4.1 Introduction

Since the beginning of the 1980s, total household debt increased substantially and almost doubled relative to GDP or private personal income in the US economy. This significant rise in household leveraging has led to a strand of literature studying the interaction between financial markets and the macroeconomy.³¹ This study empirically investigates the impact of total factor productivity (TFP) shocks and tax innovations on household debt for the US economy and proposes a model with financial frictions that is capable of explaining the empirical observations.

It is widely agreed that introducing financial frictions into stochastic general equilibrium (DSGE) models changes the economic dynamics to shocks not just quantitatively but also qualitatively. Monacelli (2009) demonstrates that financial frictions are needed to account for the non-durable and durable consumption responses to a monetary policy shock as observed in the data. Based on vector autoregressions (VARs), Andrés et al. (2015) find that an expansionary government spending shock is followed by a significant and persistent increase in household debt. The authors propose a model in which private borrowing is limited to the value of the households' collateral.³² Based on these findings, this paper empirically shows that household debt moves procyclically in response to TFP shocks and tax innovations. Additionally, it is demonstrated that a DSGE model in which borrowing is limited by a collateral constraint can successfully account for these empirical results.

To study the impact on household debt, (i) the TFP series from Basu et al. (2006) and (ii) the Romer and Romer (2010) tax measure is incorporated into recursive SVARs. We select the TFP shock because technology improvements are one, among others, of the major drivers of the business cycle (e.g. Fisher (2006), Justiniano et al.

³¹Some influential studies are Kiyotaki and Moore (1997), Bernanke et al. (1999), and Iacoviello (2005).

³²In a similar vein, Eggertsson and Krugman (2012) theoretical show that the size of the government spending multiplier crucially depends on the degree of financial market imperfections.

(2010)). Moreover, Mertens and Ravn (2013) empirically show that tax changes induce important impulses to U.S. output fluctuations. Also, tax changes represent an important instrument for the fiscal authority to stimulate the economy.

So far, both of these shock series are mainly used to quantify their dynamic effects on variables like output, consumption or hours worked³³ whereas this paper takes a closer look at how household debt evolves to changes in both measures.

Our empirical results suggest that increases in total factor productivity, as well as tax cuts, lead to a significant and persistent increase in household debt. Moreover, this paper finds that both shocks have expansionary effects on output, durable and non-durable consumption. These findings indicate that the rise in economic activity in response to both shocks is partly financed by an increase in private borrowing. From a theoretical perspective, a positive debt response refutes consumption smoothing which assumes households to save in good times and borrow in bad. However, it is demonstrated that a theoretical model in which borrowing is limited by a collateral constraint as suggested by Monacelli (2009) produces such positive debt responses following both shocks. By applying impulse response matching it is then shown that this approach is capable of successfully explaining the empirical results.

Our proposed DSGE model is closely related to those used in the housing literature Iacoviello (2005) and in the literature on durable goods Monacelli (2009). The model economy is populated by two types of households, different in their willingness to postpone consumption into the future, which creates borrowers and lenders. Both agents earn after-tax labor income and receive utility from leisure and consuming a basket of durable and non-durable goods. The government purchases a stream of goods which is financed by distortionary labor income taxes and balances its budget every period by paying out lump-sum transfers. As the central building block of the model, borrowers face a collateral constraint so that the amount of newly issued private debt is restricted to a fraction of the value of their durable stock following Monacelli (2009). Both innovations lead to an expansion in the modeled economy characterized by increases in output, non-durable consumption, and durable consumption. By assuming that the borrowing constraint holds with equality in the neighborhood of the steady-state, discount factors of the two types of households have to differ, as Iacoviello (2005) and Monacelli (2009) show.

To bring theoretical impulse responses as close as possible to the empirical data, deep model parameters are estimated. Instead of comparing the impulse responses from structural VARs to the theoretical responses from a model, this approach

³³For the TFP shock see, among others, Basu et al. (2006), Christiano et al. (2004) and for the tax shock some prominent examples are Romer and Romer (2010), Mertens and Ravn (2013), Favero and Giavazzi (2012).

minimizes the distance between structural VAR responses run on the data and identical VAR responses run on simulated model data. Thus, the U.S. data and the model simulations are treated equally so that problems like small-sample biases or lag-truncation biases are avoided as Cogley and Nason (1995) and Kehoe (2006) show.

Our results from the matching procedure suggest that the model can successfully account for the sizes and the hump-shaped patterns of the empirical dynamics in all four variables. In line with the empirical findings, the model produces persistent increases in household debt which last for more than 20 quarters. Moreover, the models' debt responses almost perfectly match the empirical counterparts. The point estimates of deep model parameters are in line with findings by previous studies (Mertens and Ravn, 2013; Iacoviello, 2005). We find that almost 50% of households are faced with a collateral constraint so that their ability to borrow to finance consumption is limited.

This study is a contribution to the existing literature in two dimensions. It is the first study giving a estimate for households' debt responses to technology improvements and tax cuts based on SVARs. Additionally, this paper contributes to the literature by showing that an estimated DSGE model with financial frictions matches the empirical responses of the major variable of interest, household debt, but also output, non-durable consumption, durable consumption, to these shocks quantitatively.

The remainder of the paper is organized as follows. Section 2 presents the results from the SVAR estimation. section 3 lays out the theoretical model. Section 4 describes the models' calibration and estimation strategy. Section 5 presents the results of the impulse response matching approach. Finally, the last section concludes.

4.2 Empirical evidence

In this section, we present our data, estimation method and SVAR results on the impact of technology shocks and tax cuts on total household debt and other main aggregates of interest.

4.2.1 Data and identification

Our benchmark VAR consists of five variables. Apart from the main variable of interest, total household debt (d_t), we include output (y_t), non-durable consumption (c_t^n), consumption expenditures on durables (c_t^d), as well as one of the two shock

measures, technology (z_t) or the tax innovation (τ_t). All variables are linearly detrended before estimation and enter the VAR in logs of real per capita, seasonally adjusted values.

Precise definitions and data sources are summarized in Table 4. We measure the

Table 4: Data sources

Variable	Definition
y Output	Log of per capita Nominal gross domestic product divided by the GDP deflator
c^n Non-durable consumption	Log of per capita (personal consumption expenditures on non durables plus personal expenditures on services) divided by each individual price deflator
c^d Durable purchases	Log of per capita personal purchases of durable consumption goods divided by its deflator
d Total private debt	Log of per capita total private debt, divided by the consumer price index
τ Tax Shock	Romer and Romer (2010) exogenous tax shock
z Technology Shock	Basu et al. (2006) exogenous technology Shock

Notes: All data are linearly detrended and logs of real per capita, seasonally adjusted values and are obtained from FRED database. Full time series specific information and sources can be found in the appendix.

impact of technology shocks using the TFP series computed in Basu et al. (2006). This series is a Solow residual-based measure of technology corrected for labor and capital utilization, non-constant returns to scale and imperfect competition.

To cope with the issue of endogenous and exogenous tax changes, we utilize the Romer and Romer (2010) tax measure.³⁴ The authors take a narrative approach to disentangling exogenous and endogenous tax change effects by analyzing presidential speeches, the Economic Reports of the President and reports of Congressional committees. Their resulting shock series is measured in changes in tax revenues relative to GDP, discounted to the day when the bill was signed to avoid a misalignment of the data set and agents' economic choices, also called fiscal foresight.

Since the identification of the empirical model depends on the nature of the two shock series, i.e. if they are (strictly) exogenous, we perform Granger causality tests, as also done in the former chapter, but with a different set of variables. In particular, we use these tests to find the suited VAR estimation method.

³⁴Actual changes in tax rates or tax revenues are a linear combination of exogenous and endogenous tax changes, which would dilute the structural effect of tax innovations on total private debt. We hence seek to use a measure for exogenous tax changes only, rather than including automatic tax adjustments that co-move with the business cycle.

The results are summarized in Table 5. We find that the lags of output and durable consumption Granger cause TFP at the 95% significant level, which is in contrast to results reported in Basu et al. (2006)³⁵. When testing whether the Romer and Romer (2010) tax measure can be predicted by past observations of our main aggregates, Granger causality cannot be rejected either. Lagged values of output and durable consumption, include information which predicts the specific size of future tax changes.

Table 5: Granger causality test results

	Basu et al. TFP series				Romer & Romer tax series			
	Sample	Obs	F-stat	p-value	Sample	Obs	F-stat	p-value
Output	1966–2014	191	2.395	0.042	1966–2007	164	2.487	0.025
Non-durable consumption	1966–2014	191	1.723	0.146	1966–2007	164	1.486	0.186
Durable expenses	1966–2014	191	3.673	0.006	1966–2007	164	2.871	0.011
Household debt	1966–2014	191	0.967	0.426	1966–2007	164	1.121	0.303

Notes: Null hypothesis: The variable does not Granger cause the TFP / tax measure. Specification: six lags, linearly detrended data for output, non-durable consumption, durable consumption, hours worked, household debt (d) and TFP index as used in VARs. The Romer and Romer tax series is recoded as in chapter 3 and in Mertens and Ravn (2013)

Given these results, treating both measures as strictly exogenous series seems misleading and estimating exogenous VARs, as done in Basu et al. (2006) and Mertens and Ravn (2013), will not reveal the true impact of technology shocks and tax innovations on the variables of interest.

We acknowledge the fact of contemporaneous exogeneity of the two shock series by estimating VARs, in which the specific shock is ordered first. This identification approach implies, that the TFP and the tax measure are contemporaneously unaffected by the other variables in the system while the subsequent variables have an impact through the lag structures.³⁶ Our baseline SVAR takes the following form

$$X_t = A(L)X_{t-1} + u_t, \quad (9)$$

in which $X_t = [s_t, y_t, c_t^n, c_t^d, d_t]'$, $s_t \in \{z_t, \tau_t\}$. $A(L)$ is a lag polynomial of order 4 and the estimation includes a constant term. Finally, u_t denotes reduced form residuals, and their variance-covariance matrix is orthogonalized by Cholesky decomposition, and the VAR is estimating using ordinary least squares.

³⁵One explanation for the different results may be the different and shorter period considered by Basu et al. (2006), as also suggested in chapter 3.

³⁶Note that estimating VARX systems does not have an impact on the qualitative results, whereas the strength of the impact of the two shocks is quantitatively larger for the tax shock and smaller for the TFP shock, while still being for the majority of the periods within 68% confidence bands, and always within 90% confidence intervals.

Both shocks enter our SVARs along with quarterly US data from 1966q1 to 2007q4 for the tax shock and 1966q1 to 2014q4 for the TFP shock. Due to data limitations of the tax measure the two samples have different lengths.

4.2.2 VAR results

We produce one set of impulse responses for the TFP shock and one for the tax cut. The size of the technology shock is the equivalent to an increase of one percent in the Basu et al. (2006) measure for total factor productivity. The Romer and Romer (2010) structural tax shock is equivalent to a reduction of total tax revenues relative to GDP of one percentage point. We report impulse responses together with 68% (dark gray), and 90% (light gray) bootstrapped confidence bands, computed with 10.000 bootstrap replications. Figure 26 (a) depicts the results for the TFP shock and Figure 26 (b) those for the tax cut.

Both, the tax reduction and the increase in technological progress initiate an expansion characterized by hump-shaped dynamics in output, non-durable consumption, and durable expenditures. This boom is persistent, lasting for more than five years before the economy returns to its pre-shock level. While most of the variables do not change on impact when the economy faces a tax cut, the TFP shock influences the aggregates already on impact.

With respect to our primary variable of interest, household debt, we find that both shocks lead to a significant and persistent increase in private borrowing. This result indicates that the expansion in the economy is partly financed by a rise in household debt. For the TFP shock, household debt peaks after around 5 quarters, while for the tax cuts it converges back to pre-shock levels later. From a theoretical perspective, a sharp rise in household debt following both shocks is in contrast to consumption smoothing of households. This assumption predicts a fall in private borrowing in expansionary times as a buffer against future negative shocks.³⁷

Concerning the volatility of our endogenous variables, we can detect a clear pattern in responses to both exogenous innovations. Durable purchases and household debt react the strongest following both shocks.³⁸

Non-durable consumption shows the smallest increases of all endogenous variables included in our VAR estimations.

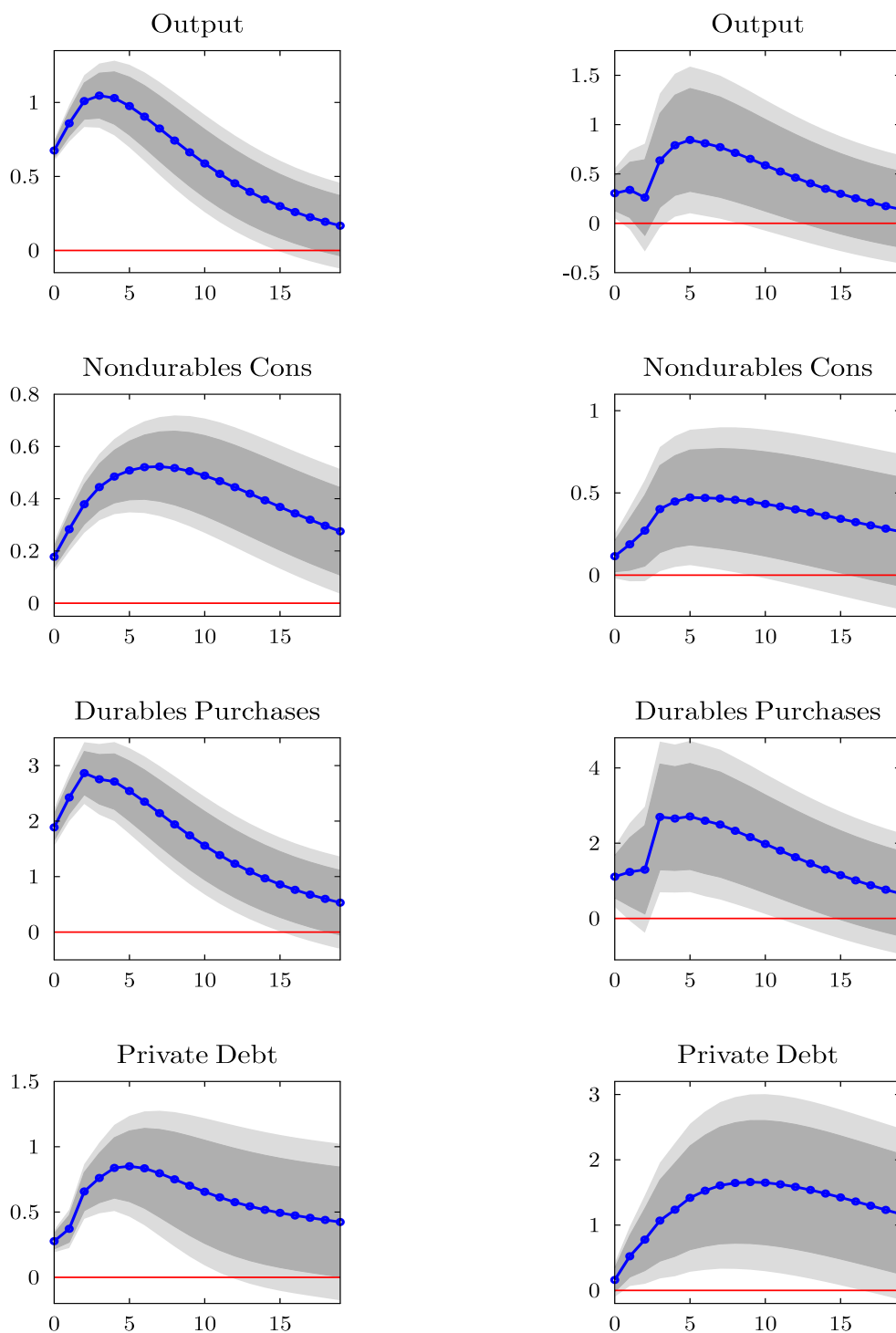
³⁷If we include the unemployment rate in our estimation, we find that unemployment is reduced after the two shocks. Thus, the rise in household debt is not caused by a decrease in household income.

³⁸We interpret this strong comovement between durable purchases and household debt as justification for the borrowing constraint in our model which we describe in the next section.

Figure 26: Impulse responses SVAR estimation

(a) TFP Shock

(b) Tax Shock



Notes: SVAR impulse responses to a one percent increase in TFP (a) and a one percentage point decrease in tax revenues over GDP along with 68% (dark gray) and 90% (light gray) bootstrapped centered confidence intervals with 10.000 bootstrapped replications. Reduced form residual variance-covariance matrices are Cholesky decomposed.

These empirical findings are robust to the specification of alternative orderings, less or additional lags, including hours in the estimation and to the introduction of alternative variables. We find that the initially observed comovement of the variables remains intact. Figures for robustness checks can be found in the appendix.

4.3 Model

This section presents a DSGE model with financial frictions that we use to explain our empirical findings. The model consists of two types of households, a representative final goods firm, a monopolistically competitive intermediate goods sector, and a government sector.

4.3.1 Households

Our model is based on the ones in Mertens and Ravn (2011), for the durable consumption sector, government structure and utility function and also on Monacelli (2009) for the borrowing constraint.

The model economy is populated by a continuum of infinitely-lived households that are heterogeneous in terms of their desire to save. Hence, a fraction χ of households becomes *lenders* (subscript l), while the remaining fraction $1-\chi$ becomes *borrowers* (subscript b). Borrowing households face a collateral constraint which ensures that private borrowing is restricted to a certain amount of their stock of durables.

Lenders. Lending households' preferences are given by

$$E_0 \sum_{t=0}^{\infty} \beta_l^t \left(\frac{\gamma_{l,t}^{1-\sigma_l} - 1}{1 - \sigma_l} - \gamma_l \frac{n_{l,t}^{1+\eta_l}}{1 + \eta_l} \right), \quad (10)$$

where E_0 denotes the expectation operator conditional on all information available at time 0. $0 < \beta_l < 1$ is the lenders' specific discount factor, $\sigma_l > 0$ is a curvature parameter, $\gamma_l > 0$ is the preference weight that measures disutility of labor, $n_{l,t}$, and $\eta_l \geq 0$ is the lenders' specific inverse Frisch elasticity.

$\mathcal{Y}_{l,t}$ denotes a consumption basket defined as

$$\mathcal{Y}_{l,t} = c_{l,t}^{\vartheta_l} v_{l,t-1}^{1-\vartheta_l} - \psi_l c_{l,t-1}^{\vartheta_l} v_{l,t-2}^{1-\vartheta_l}, \quad (11)$$

where $c_{l,t}$ is consumption of non-durable goods and $v_{l,t-1}$ denotes the stock of durable goods held at the beginning of period t . $\vartheta_l \in [0, 1]$ measures the elasticity of substitution between non-durable and durable consumption of lending households, and

$\psi_l \in [0, 1]$ governs the lenders' degree of habit persistence. We follow Mertens and Ravn (2011) by assuming that non-durable and durable consumption are complementary goods for households, which is assured by the specific functional form of (11).

Lending households maximize (10) with respect to their budget constraint given by

$$c_{l,t} + x_{l,t} + b_{l,t} + d_{l,t} \leq (1 - \tau_t)w_t n_{l,t} + (1 + r_{g,t-1})\frac{b_{l,t-1}}{\pi_t} + (1 + r_{d,t-1})\frac{d_{l,t-1}}{\pi_t} + \frac{\Pi_t}{\chi} + tr_t, \quad (12)$$

where $\pi_t = p_t/p_{t-1}$ is the gross inflation rate and $b_{l,t}$ are the lender's holdings of one-period government bonds with interest $r_{g,t}$. Lenders receive after-tax labor income, $(1 - \tau_t)w_t n_{l,t}$, where τ_t is the labor income tax rate and w_t is the real wage rate which households take as given. $x_{l,t}$ represents purchases of new durable goods. In addition, lending households earn financial income, $(1 + r_{d,t-1})d_{l,t-1}$, from offering one-period private debt to borrowers at interest $r_{d,t-1}$ which is guaranteed to be repaid in the next period. tr_t denotes lump-sum transfers paid by the fiscal authority and Π_t/χ are the individual profits from owning intermediate goods firms.

The law of motion for the durable stock is given by

$$v_{l,t} = \left(1 - \frac{\phi_v}{2} \left(\frac{x_{l,t}}{x_{l,t-1}} - 1\right)^2\right) x_{l,t} + (1 - \delta_v)v_{l,t-1}, \quad (13)$$

in which δ_v denotes a constant depreciation rate and the parameter ϕ_v captures costs of adjusting the stock of durable goods. We choose this quadratic and convex functional form since it satisfies the properties generally imposed on adjustment costs (see, for example, Christiano et al., 2005).³⁹

Letting $\lambda_{l,t}$ be the lenders' Lagrange multiplier corresponding to their budget constraint, the first-order conditions (FOCs) for non-durable consumption, govern-

³⁹Let $\Phi(x_t/x_{t-1})$ be the general adjustment cost function. Then, convexity implies $\Phi(1) = \Phi'(1) = 0$ and $\Phi''(1) = \phi_v > 0$ which is assured by the functional form in (13).

ment bond holdings, hours worked, durable consumption, debt supply and durable purchases are given by

$$c_{l,t} : \lambda_{l,t} = \vartheta_l \left(\Upsilon_{l,t}^{-\sigma_l} - \psi_l \beta_l E_t \Upsilon_{l,t+1}^{-\sigma_l} \right) \left(\frac{v_{l,t-1}}{c_{l,t}} \right)^{1-\vartheta_l}, \quad (14)$$

$$b_{l,t} : \lambda_{l,t} = \beta_l E_t \left\{ \lambda_{l,t+1} \frac{1+r_{g,t}}{\pi_{t+1}} \right\}, \quad (15)$$

$$n_{l,t} : \lambda_{l,t} (1-\tau_t) w_t = \gamma_l n_{l,t}^{\eta_l}, \quad (16)$$

$$v_{l,t} : \lambda_{l,t} q_{v,t} = \beta_l E_t \left\{ \lambda_{l,t+1} \left[\left(\frac{1-\vartheta_l}{\vartheta_l} \right) \frac{c_{l,t+1}}{v_{l,t}} + q_{v,t+1} (1-\delta_v) \right] \right\}, \quad (17)$$

$$d_{l,t} : \lambda_{l,t} = \beta_l E_t \left\{ \lambda_{l,t+1} \frac{1+r_{d,t}}{\pi_{t+1}} \right\}, \quad (18)$$

$$\begin{aligned} x_{l,t} : 1 - q_{v,t} \left(1 - \frac{\phi_v}{2} \left(\frac{x_{l,t}}{x_{l,t-1}} - 1 \right)^2 - \phi_v \left(\frac{x_{l,t}}{x_{l,t-1}} - 1 \right) \frac{x_{l,t}}{x_{l,t-1}} \right) \\ = \beta_l E_t \left\{ \frac{\lambda_{l,t+1}}{\lambda_{l,t}} q_{v,t+1} \phi_v \left(\frac{x_{l,t+1}}{x_{l,t}} - 1 \right) \left(\frac{x_{l,t+1}}{x_{l,t}} \right)^2 \right\}, \end{aligned} \quad (19)$$

where $q_{v,t}$ denotes the lenders' shadow value of new consumer durables. Equation (14) states that $\lambda_{l,t}$ equals the marginal utility of non-durable consumption. Equation (15) is the standard Euler equation for government bond holdings. Equation (16) sets the marginal rate of substitution between consumption and leisure equal to the after-tax real wage rate. Equation (17) shows that the shadow value of new consumer durables is equal to the expected discounted utility stream received from the durable stock (net of depreciation). Equation (18) sets $\lambda_{l,t}$ equal to the expected discounted utility stream of future debt interest rate payments. Equation (19) states that the change in consumer durables is related to the expected discounted value of current and future levels of $q_{v,t}$.

Borrowers. Preferences of borrowing households are defined as

$$E_0 \sum_{t=0}^{\infty} \beta_b^t \left(\frac{\Upsilon_{b,t}^{1-\sigma_b} - 1}{1-\sigma_b} - \gamma_b \frac{n_{b,t}^{1+\eta_b}}{1+\eta_b} \right), \quad (20)$$

in which $0 < \beta_b < 1$ is the specific discount factor of borrowers, $\sigma_b > 0$ is a curvature parameter, $\gamma_b > 0$ is a scaling parameter measuring the borrowers disutility of labor,

$n_{b,t}$, and $\eta_b \geq 0$ is their specific inverse Frisch elasticity. Again, $\Upsilon_{b,t}$ denotes a consumption basket defined as

$$\Upsilon_{b,t} = c_{b,t}^{\vartheta_b} v_{b,t-1}^{1-\vartheta_b} - \psi_b c_{b,t-1}^{\vartheta_b} v_{b,t-2}^{1-\vartheta_b}. \quad (21)$$

Here, $c_{b,t}$ denotes borrowers' consumption of non-durable goods and $v_{b,t-1}$ is the stock of durable goods held at the beginning of period t . $\vartheta_b \in [0, 1]$ measures borrowers' substitution elasticity between non-durable and durable consumption and $\psi_b \in [0, 1]$ measures the degree of habit persistence.

The budget constraint of borrowing households is given by

$$c_{b,t} + x_{b,t} + (1 + r_{d,t-1}) \frac{d_{b,t-1}}{\pi_t} \leq (1 - \tau_t) w_t n_{b,t} + d_{b,t} + tr_t. \quad (22)$$

$x_{b,t}$ denotes borrowers' purchases of new consumer durables and $d_{b,t}$ is the amount of one-period-debt received from lenders which has to be repaid plus interest $r_{d,t-1}$ in the subsequent period. $(1 - \tau_t) w_t n_{b,t}$ denotes borrowers' after-tax labor income.

The borrowers' stock of durables accumulates according to

$$v_{b,t} = \left(1 - \frac{\phi_v}{2} \left(\frac{x_{b,t}}{x_{b,t-1}} - 1 \right)^2 \right) x_{b,t} + (1 - \delta_v) v_{b,t-1}. \quad (23)$$

As a central building block of our model, borrowing is endogenously determined by a collateral constraint, similar to the one used in Iacoviello (2005) and Monacelli (2009). The amount of debt that has to be repaid by borrowers in the following period, $d_{b,t}$, is the net-of-depreciation durable stock

$$d_{b,t} \leq \kappa(1 - \delta_v) v_{b,t}, \quad (24)$$

where $\kappa > 0$ denotes the share of borrowers' durable stock that can be used as collateral. This borrowing constraint implies two noteworthy points. First, by assuming that (24) holds with equality, β_b has to be smaller than β_l , and thus, borrowers hold a positive steady state amount of debt. Second, changes in the stock of durable goods affect borrowing but also spending (of constrained households). The magnitude of this effect crucially depends on the size of κ .

The borrowers FOCs take the following expressions

$$c_{b,t} : \lambda_{l,t} = \vartheta_b \left(\Upsilon_{b,t}^{-\sigma_b} - \psi_b \beta_b E_t \Upsilon_{b,t+1}^{-\sigma_b} \right) \left(\frac{v_{b,t-1}}{c_{b,t}} \right)^{1-\vartheta_b}, \quad (25)$$

$$n_{b,t} : \lambda_{b,t} (1 - \tau_t) w_t = \gamma_b n_{b,t}^{\eta_b}, \quad (26)$$

$$v_{b,t} : \lambda_{b,t} q_{x,t} = \beta_b E_t \left\{ \lambda_{b,t+1} \left[\left(\frac{1 - \vartheta_b}{\vartheta_b} \right) \frac{c_{b,t+1}}{v_{b,t}} + q_{x,t+1} (1 - \delta_v) \right] \right\} + \mu_t (1 - \delta_v) \kappa, \quad (27)$$

$$d_{b,t} : \lambda_{b,t} = \beta_b E_t \left\{ \lambda_{b,t+1} \frac{1 + r_{d,t}}{\pi_{t+1}} \right\} + \mu_t, \quad (28)$$

$$\begin{aligned} x_{b,t} : 1 - q_{x,t} \left(1 - \frac{\phi_v}{2} \left(\frac{x_{b,t}}{x_{b,t-1}} - 1 \right)^2 - \phi_v \left(\frac{x_{b,t}}{x_{b,t-1}} - 1 \right) \frac{x_{b,t}}{x_{b,t-1}} \right) \\ = \beta_b E_t \left\{ \frac{\lambda_{b,t+1}}{\lambda_{b,t}} q_{x,t+1} \phi_v \left(\frac{x_{b,t+1}}{x_{b,t}} - 1 \right) \left(\frac{x_{b,t+1}}{x_{b,t}} \right)^2 \right\}, \end{aligned} \quad (29)$$

where $\lambda_{b,t}$ is the Lagrange multiplier on the borrowers budget constraint, μ_t denotes the Lagrange multiplier of collateral constraint (24), and $q_{x,t}$ denotes the borrowers shadow value of new consumer durable purchases. Interpretations of equations (25), (26), and (29) are identical to those of the lending households. The last term of (27) governs that the shadow value of new consumer durables is related to the marginal utility of relaxing the collateral constrained measured though the time-varying Lagrange multiplier μ_t . (28) shows that for positive values of μ_t the marginal utility of current consumption is larger than the marginal value of shifting one unit of consumption intertemporally. A higher value for μ_t induces a larger marginal benefit of increasing the stock of durable consumption goods which leads to a loosening of the collateral constraint to purchase additional current consumption.

4.3.2 Firms

The firm sector consists of a perfectly competitive final goods firm and a continuum of monopolistically competitive intermediate goods firms. Each intermediate goods firm $i \in [0, 1]$ produces a differentiated good $y_t(i)$ according to the production function

$$y_t(i) = z_t n_t(i), \quad (30)$$

where $n_t(i)$ denotes the quantity of labor services used by firm i . The technology level z_t is common across all operating firms and is assumed to follow an AR(1) process around its non-stochastic steady state value \bar{z} ,

$$\log(z_t) = (1 - \rho_z) \log(\bar{z}) + \rho_z \log(z_{t-1}) + \varepsilon_{z,t}, \quad (31)$$

in which $\varepsilon_{z,t}$ is *i.i.d* and $|\rho_z| < 1$. The representative final goods firm produces the final consumption good y_t , combining $y_t(i)$ units of each intermediate good, using the technology

$$y_t = \left(\int_0^1 y_t(i)^{\frac{\xi-1}{\xi}} di \right)^{\frac{\xi}{\xi-1}}, \quad (32)$$

where $\xi > 1$ is the elasticity of substitution between different intermediate goods. Profit maximization subject to (32) yields the demand function for intermediate good i ,

$$y_t(i) = y_t \left(\frac{p_t(i)}{p_t} \right)^{-\xi}, \quad (33)$$

where

$$p_t = \left(\int_0^1 p_t(i)^{1-\xi} di \right)^{\frac{1}{1-\xi}} \quad (34)$$

is the price index of the final good.

Each firm in the intermediate goods sector chooses its price level $p_t(i)$ to maximize the expected present value of real profits. Following Rotemberg (1982), each firm faces quadratic adjustment costs which are assumed to take the functional form of Ireland (1997). Thus, real profits of firm i are given by

$$\Pi_t(i) = \left[\left(\frac{p_t(i)}{p_t} \right)^{1-\xi} - \frac{w_t}{z_t} \left(\frac{p_t(i)}{p_t} \right)^{-\xi} - \frac{\varphi}{2} \left(\frac{p_t(i)}{\bar{\pi} p_{t-1}(i)} - 1 \right)^2 \right] y_t, \quad (35)$$

where $\varphi > 0$ determines the adjustment costs and $\bar{\pi}$ is the steady state inflation rate.

Assuming symmetry in equilibrium, the optimality condition becomes

$$\varphi \left(\frac{\pi_t}{\bar{\pi}} - 1 \right) \frac{\pi_t}{\bar{\pi}} = (1 - \xi) + \xi \frac{w_t}{z_t} + E_t \left[\beta \frac{\lambda_{l,t+1}}{\lambda_{l,t}} \left(\frac{\pi_{t+1}}{\bar{\pi}} - 1 \right) \frac{\pi_{t+1}}{\bar{\pi}} \frac{y_{t+1}}{y_t} \right]. \quad (36)$$

In case of fully flexible prices, i.e. $\varphi = 0$, real marginal costs equal $(\xi - 1)/\xi$, which is the inverse of the firm's price markup.

4.3.3 Government

The government collects distortionary labor income taxes and issues new bonds to finance public spending, to service debt from last period and to pay out lump-sum transfers to households. Hence, the government's budget constraint reads

$$g_t + tr_t + (1 + r_{g,t-1}) \frac{b_{t-1}}{\pi_t} = \tau_t w_t n_t + b_t, \quad (37)$$

where government spending g_t is a fixed fraction of aggregate output, and transfers tr_t adjust to balance the budget in every period. We estimate an AR(2) process for the tax rate around its non-stochastic steady-state value $\bar{\tau}$ which gives the best empirical fit as also shown in Mertens and Ravn (2011). The process is given by

$$\log(\tau_t) = (1 - \rho_{\tau,1} - \rho_{\tau,2}) \log(\bar{\tau}) + \rho_{\tau,1} \log(\tau_{t-1}) + \rho_{\tau,2} \log(\tau_{t-2}) - \varepsilon_{\tau,t}, \quad (38)$$

where $\varepsilon_{\tau,t}$ is *i.i.d.*, and $|\rho_{\tau,1} + \rho_{\tau,2}| < 1$.

Monetary policy is determined by a Taylor-type rule of the form

$$r_{g,t} = \bar{r}_g \left(\frac{\pi_t}{\pi^*} \right)^{\phi_\pi}, \quad (39)$$

where $\pi^* = \bar{\pi}$ is the inflation rate target and ϕ_π is the policy response to inflation deviations from its target.

4.3.4 Aggregation and market clearing

Aggregate variables are defined as the weighted average of the respective measures for each household type. Thus, we get

$$c_t = \chi c_{l,t} + (1 - \chi) c_{b,t}, \quad (40)$$

$$v_t = \chi v_{l,t} + (1 - \chi) v_{b,t}, \quad (41)$$

$$x_t = \chi x_{l,t} + (1 - \chi) x_{b,t}, \quad (42)$$

$$n_t = \chi n_{l,t} + (1 - \chi) n_{b,t}. \quad (43)$$

Credit and bond market clearing requires

$$\chi d_{l,t} = (1 - \chi) d_{b,t}, \quad (44)$$

$$b_t = \chi b_{l,t} \quad (45)$$

while the aggregate resource constraint is given by

$$c_t + x_t + g_t = \left[1 - \frac{\varphi}{2} \left(\frac{\pi_t}{\bar{\pi}} - 1 \right)^2 \right] y_t. \quad (46)$$

4.3.5 Equilibrium

A competitive equilibrium is given by the sequence of endogenous variables $\{y_t, c_t, c_{l,t}, c_{b,t}, v_t, v_{l,t}, v_{b,t}, x_t, x_{l,t}, x_{b,t}, \Upsilon_{l,t}, \Upsilon_{b,t}, n_t, n_{l,t}, n_{b,t}, d_{l,t}, d_{b,t}, b_t, b_{l,t}, \pi_t, tr_t, g_t, \lambda_{l,t}, \lambda_{b,t}, q_{v,t}, q_{x,t}, \mu_t, w_t, r_{d,t}, r_{g,t}\}_{t=0}^{\infty}$ that satisfy the households' first-order conditions, the firms' optimality conditions, the production function, the government budget constraint, the monetary policy rule, the stochastic processes, credit and bond market clearing, the aggregation identities and the aggregate resource constraint, given the exogenous realizations of $\{z_t, \tau_t\}_{t=0}^{\infty}$.

To solve the model by a log-linear approximation around its deterministic steady state, we assume that all inequalities hold with equality in equilibrium.

4.4 Parametrization

To study whether our proposed model can account for the empirical findings, we estimate deep model parameters by applying an impulse-response matching approach as suggested by Cogley and Nason (1995). The set of parameters is partitioned into two subsets, $\Theta \equiv [\theta_1, \theta_2]$, where θ_2 contains the parameters to be estimated and θ_1 contains the parameters that are calibrated prior to estimation. The elements of the latter subset are fixed because they are either difficult to identify in model estimation procedures or are chosen to match certain steady-state targets observed in the data.

4.4.1 Calibration

One model period is set to be a quarter. We choose the lenders' discount factor to be 0.993 implying an annual steady-state interest rate of 3% and follow Iacoviello and Neri (2010) by setting the borrowers' discount factor to 0.97 to induce a significant degree of impatience. We include this parameter in θ_1 because our set of moments is not able to jointly identify σ_l , σ_b and the inverse Frisch elasticities η_l and η_b . The preference parameters determining disutility of work, γ_l and γ_b , are calibrated so that steady state hours worked equal 33% of individual time endowment. ϑ_l and ϑ_b equal 0.75 which implies an aggregate steady state durable-to-non-durable-consumption ratio of 20%, in line with the corresponding number in the US during our sample period. The elasticity of substitution between intermediate goods, ξ ,

equals 11 implying a steady-state markup of $\xi/(\xi - 1) = 1.1$. For the debt-to-value ratio, κ , we again follow Iacoviello (2010) and choose 0.85 so that borrowing households can use 85% of their durable stock as collateral. Following Mertens and Ravn (2011), the depreciation rate of durable goods is set to 0.025 implying a steady state annual depreciation of 10%. We set the policy parameter in the Taylor rule, ϕ_π , to 1.5 as in Monacelli (2009). The steady state government-bonds-to-GDP and government-spending-to-GDP ratios and the labor income tax rate equal 0.60, 0.18, and 0.28, respectively, as found by Trabandt and Uhlig (2011). Table 6 summarizes the calibration of θ_1 .

Table 6: Model calibration

Param	Value	Description	Target
β_l	0.99	Discount factor lenders	Ann. interest rate 3%
β_b	0.97	Discount factor borrowers	Small degree of impatience
γ_l		Preference parameter	SS hours of lenders to 0.33
γ_b		Preference parameter	SS hours of borrowers to 0.33
ξ	11	Elasticity of substitution	SS markup of 11
δ_v	0.025	Depreciation rate durable goods	Mertens and Ravn (2011)
κ	0.85	Debt-to-value ratio	Andrés et al. (2013)
ϑ	0.85	Preference parameter	$\bar{X}/(\bar{C} + \bar{X}) = 0.20$
ϕ_π	1.5	Taylor rule parameter	Monacelli (2009)
\bar{B}/\bar{Y}	0.60	Government debt to GDP	
\bar{G}/\bar{Y}	0.18	Government spending to GDP	Trabandt/Uhlig (2011)
$\bar{\tau}_n$	0.28	SS tax rate	Trabandt/Uhlig (2011)

4.4.2 Estimation

We estimate $\theta_2 = [\eta, \sigma, \psi_l, \psi_b, \phi_v, \chi, \rho_z, \varphi, \rho_{\tau,1}, \rho_{\tau,2}]$, by matching the impulse responses generated by the model to the empirical responses derived in section 2. Estimating these parameters does not have an impact on the set of calibrated parameters in θ_1 , except on the Frisch elasticity η . We target $\bar{n}_l = \bar{n}_b = 0.33$ so that γ_l and γ_b are endogenously determined.⁴⁰ We follow Cogley and Nason (1995) and Mertens and Ravn (2011), and treat model and data symmetrically. This implies that we use our model to simulate artificial samples and estimate impulse responses in exactly the same way as the empirical ones are obtained.⁴¹

⁴⁰The two habit parameters have an impact on the steady-state value of \mathcal{T} and λ , but not on the first set of calibrated parameters θ_1 .

⁴¹Using this approach avoids certain pitfalls of the theoretical impulse responses to the empirical ones, applied by e.g. Christiano et al. (2005) or Altig et al. (2011) as argued by Kehoe (2006) and Dupaigne et al. (2007).

In particular, the model-generated impulse responses are constructed according to the following algorithm.

Algorithm 1 (Construction of model-generated IRFs) *For each of the two shocks, we take three steps:*

1. *Draw 100 sequences of innovations from the original shock series (with replacement) with a length of 168 periods for the tax shock and 195 for the TFP shock. Simulate the model for each draw so that there are 100 artificial samples. Each of these simulated datasets consists of the model counterparts to the SVAR time series.*
2. *Add a small (1e-6) white noise measurement error to each artificial time series to avoid stochastic singularity.*
3. *Estimate IRFs and take mean responses over all 100 replications for each artificial dataset by estimating (9).*

Let $\widehat{\Omega}_d$ be the vector of empirical moments and let $\widehat{\Omega}_m(\theta_2|\theta_1)$ be the vector of simulated moments estimated from the same SVAR as their empirical counterparts conditioned on θ_1 . Vector θ_2 then solves the following minimization problem,

$$\widehat{\theta}_2 = \arg \min_{\theta_2} \left[\left(\widehat{\Omega}_d - \widehat{\Omega}_m(\theta_2|\theta_1) \right)' W^{-1} \left(\widehat{\Omega}_d - \widehat{\Omega}_m(\theta_2|\theta_1) \right) \right], \quad (47)$$

where W is a positive-definite weighting matrix which we find by the following procedure. First, we approximate the covariance matrix of the empirical IRFs by bootstrapping. Instead of the full matrix, we only use its diagonal which displays the variances of the IRFs and set all off-diagonal elements to zero. Hence, we only put weight on moments that are observed in the data and force the estimation to exclude moments that are off-diagonal (see Cochrane, 2005, chap. 11). Finally, we use an estimate of the weighting matrix's asymptotic covariance matrix as proposed by Hall et al. (2012) to compute standard errors for θ_2 .

4.5 Results

Table 7 shows the parameter estimates of our model estimation. We observe values for the inverse Frisch elasticities, η 0.227. These estimates are lower than those typically assumed in the macroeconomic literature, whereas Iacoviello (2010) estimate similar values in a similar model set-up. Our point estimates imply that labor supply of both agents reacts quite elastically to changes in the real wage rate.

Table 7: Estimated model parameters

Parameter	Estimated Value	Standard Error	Description
σ	0.6664		Utility curvature
η	0.2269		Inverse Frisch elasticity
ψ_l	0.5672		Habit parameter lenders
ψ_b	0.8394		Habit parameter borrowers
ϕ_v	0.0921		Durables adjustment cost
χ	0.5398		Share of lending households
$\rho_{\tau,1}$	1.8611		AR coefficient tax shock
$\rho_{\tau,2}$	-0.8745		AR coefficient tax shock
ρ_z	0.9415		AR coefficient tfp shock
φ	8.1276		Rotenberg price adjustment

Notes: Standard errors of vector θ are computed from an estimate of its asymptotic covariance matrix following Hall et al. (2012) and Mertens and Ravn (2011).

The degree of habit formation is larger for borrowing households than for lending, where the specific point estimates are in the range of values typically estimated (e.g. Christiano et al., 2005).

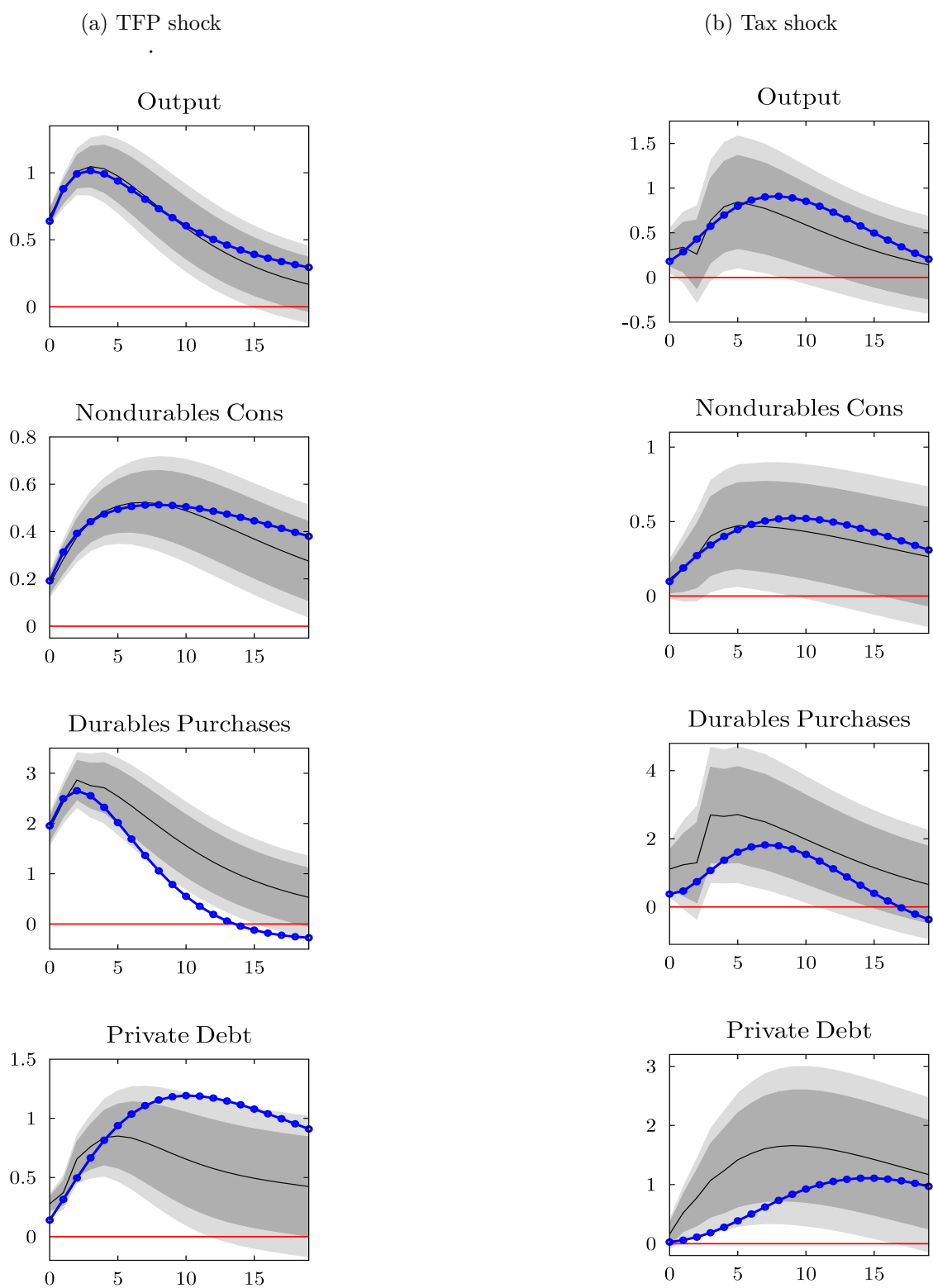
The estimate of the durable adjustment cost parameter of 0.092 is lower, compared to other studies (e.g. Mertens and Ravn, 2011). The estimated Rotenberg price adjustment coefficient, $AC(rot)$, takes a value of 8.128. Our estimate suggests that roughly 80% of firms do not adjust their prices in a given quarter, in line with Gavin et al. (2015). The share of lending households in the economy is estimated to be 54%, consistent with estimates of the proportion of unconstrained consumers by Jappelli (1990), Kiley (2010), and Hall et al. (2012).

Our estimates for the autoregressive parameters of the TFP process, $\rho_z = 0.9415$, and the tax process, $\rho_{\tau,1} = 1.8611$, $\rho_{\tau,2} = -0.8745$, which sum up to 0.987 indicate a higher degree of persistence for the tax process. The degree of persistence of the tax process is similar to the one obtained by Mertens and Ravn (2011).

Figure 27 depicts the model dynamics to a one percent increase in total factor productivity (left panel) and to a one percentage point decrease in total tax revenues over output (right column) given the parameter estimates reported in Table 7 (dotted lines) along with the empirical estimates and its confidence bands from section 4.2.2.

As visible, the model can successfully account for the sizes and hump-shaped responses of the empirical counterparts. For almost all periods, the theoretical responses lie within the empirical confidence intervals. In line with the data, the strongest model responses can be observed for durable purchases and household debt, whereas nondurable consumption shows the smallest relative deviations following both innovations.

Figure 27: Empirical and matched impulse responses



Notes: This figure depicts VAR estimated impulse responses with actual data (solid line) along with 68% bootstrapped confidence bands (dark grey) and 90% confidence bands (light grey). The dotted lines denote matched impulse responses using our model.

In panel (a) we observe that the model produces impact responses close to the empirical ones for the TFP innovation. The technology improvement leads to an expansion in the theoretical economy lasting for almost 20 quarters. The model's output, non-durable and durable consumption responses reach its peaks slightly before the empirical counterparts. As durable consumption rise, the collateral constraint becomes less binding, such that borrowing households increase their private debt holdings. The model is able to capture the persistent increase in household debt as found in the data while the response lies within 90% confidence bands for all periods. However, the model to some extent overestimates the debt response for the quarters 3-14.

In line with our empirical findings, the model does not show any impact response for most of the variables after a tax reduction as can be observed in panel (b). The limited model responses can be explained by the estimated strong habits in consumption and positive durable adjustment cost which reduce the impact effects. For non-durable consumption, the model implied response reaches its peak after around seven periods, similar to the maximum empirical response. The specific maximum of the theoretical responses for output and durable consumption peak some quarters later than found for the empirical counterparts. The model underestimates the effect of tax cuts on durable purchases. Similar to the TFP shock, the model matches the households' empirical debt response quite well. The theoretical response falls within the empirical 68% confidence bands for most of the 20 periods. The increase in private borrowing following a tax reduction can be explained by the similar mechanism as described for the TFP shock. The expansionary effects of the tax innovation rise the stock of durables held by constrained agents such that though the collateral constraint private borrowing rises in response.

Our analysis suggests, that an estimated version of the model as described in section 4.3 is able to explain the empirical dynamics following technology improvements and tax reductions. When studying in more detail how private borrowing reacts, we find that the differences between theoretical and empirical responses are almost negligible.

4.6 Conclusion

The interrelation between financial market imperfections and macroeconomic outcomes is at the core of recent research. In this paper, we study the effects of TFP shocks and tax cuts on main aggregates for the US economy while taking a closer look at how households' borrowing decisions are affected by both innovations. We select these specific shocks because of their importance for business cycle fluctuations

and as an important instrument for the fiscal authority to stimulate the economy in the short run.

By estimating SVARs, we find that both shocks lead to an expansion in the economy, characterized by significant increases in output, non-durable consumption, and durable consumption. Moreover, our results suggest that household debt rises substantially and in a hump-shaped manner in response to technology improvements and tax reductions.

In order to account for the empirically estimated comovement between economic activity and private borrowing, we propose a theoretically model with financial frictions similar to the one in Monacelli (2009). The model economy is populated by two household types, savers, and borrowers, which differ in their willingness to postpone consumption into the future. Borrowers face a collateral constraint so that the amount of newly issued private debt is restricted to a fraction of their stock of durables.

We estimate deep model parameters by matching the theoretically implied impulse response function to the empirical ones in response to both shocks. Our findings suggest that the proposed model successfully accounts for the sizes and hump-shaped patterns of the empirical dynamics. With respect to our major variable of interest, household debt, the estimated model matches the empirical responses almost perfectly. Estimated parameters are in line with findings in previous studies. We find that almost 50% of private households do face a collateral constraint that restricts their optimal borrowing decision.

4.7 Appendix

Data Source. Our data is obtained from FRED database and includes the following data series

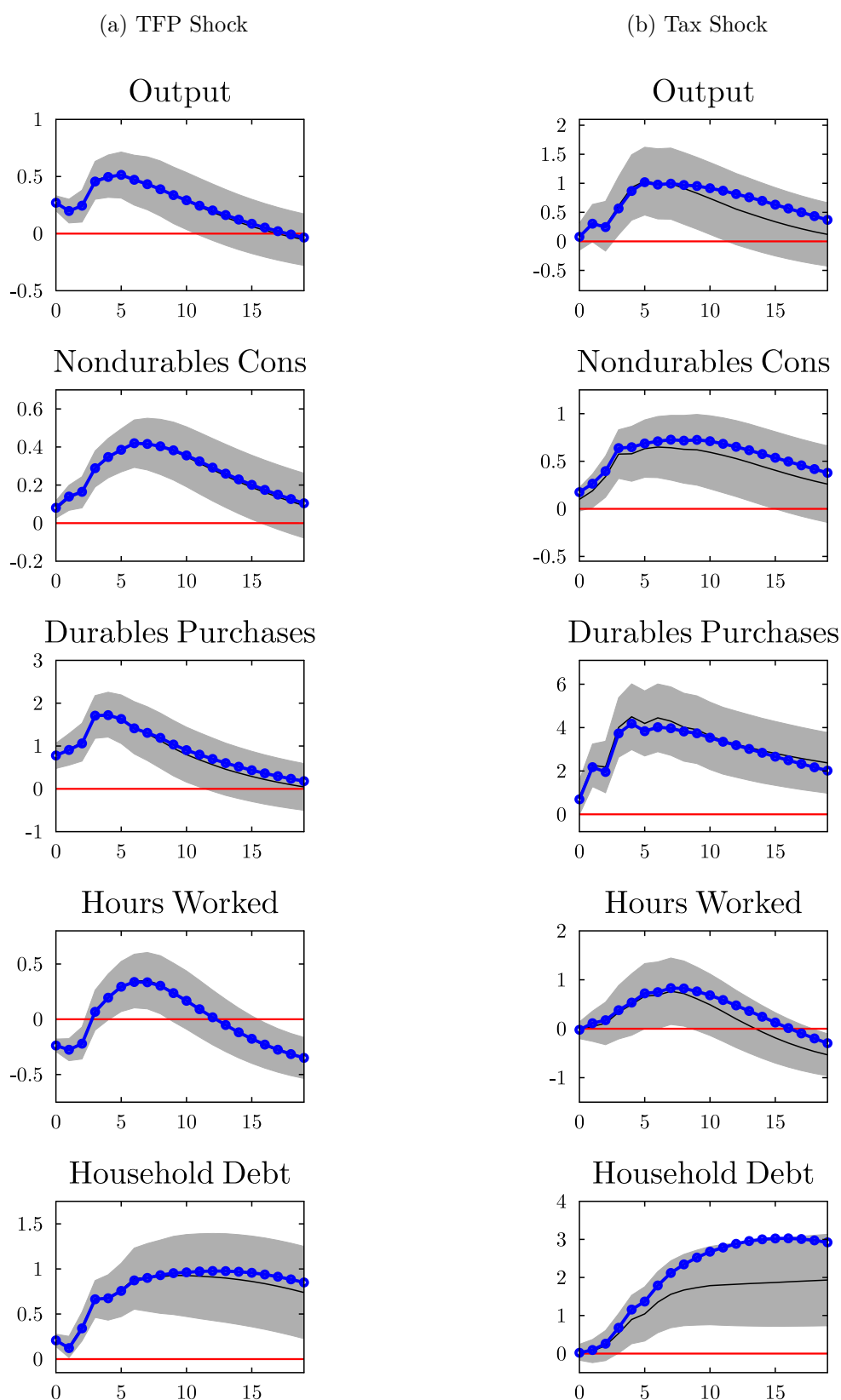
Table 8: Full data sources

Var	Definition	code	Manipulation
y	Output	GDP	Gross Domestic Product, Seasonally Adjusted Annual Rate in Billions of Dollars
c^n	Non durable consumption	PCND	Personal Consumption Expenditures: nondurable goods, seasonally adjusted annual rate in Billions of Dollars
		PCECV	Personal Consumption Expenditures: Services, Seasonally Adjusted Annual Rate in Billions of Dollars
c^d	Durable purchases	PCDG	Personal Consumption Expenditures: Durable Goods, Seasonally Adjusted Annual Rate in Billions of Dollars
h	Hours worked	h	Product of hours per worker and civilian non-farm employment divided by population combined with Francis and Ramey (2009) hours worked series, see Mertens and Ravn (2013), extended.
d	total private debt	CMDEBT	Households and Nonprofit Organizations; Credit Market Instruments; Liability, Level, seasonally adjusted in Billions of Dollars
N	Population	POP	Civilian Non institutional Population, Not Seasonally Adjusted, Thousands of Persons
Robustness checks			
tb	t-bill	TB3MS	3-Month Treasury Bill: Secondary Market Rate, Not Seasonally Adjusted, Percent
g	government consumption	GCEC96	Real Government Consumption Expenditures and Gross Investment, Seasonally Adjusted Annual rate in Billions of Dollars
Price Deflators			
	GDP deflator	GDPDEF	Gross Domestic Product: Implicit Price Deflator, Index 2009=100
	non durables deflator	DNDGRG3	Personal consumption expenditures: Nondurable goods, Index 2009=100
		Q086SBEA	
		DSERRG3	
	Durables deflator	Q086SBEA	Personal consumption expenditures: Durable goods, Index 2009=100
		DDURRG3	
	Investment deflator	Q086SBEA	Gross Private Domestic Investment: Chain-type Price Index, Index 2009=100
		GPDICTPI	
	CPI deflator	B798RG3	Government gross investment: Federal: Non-defense (chain-type price index), Index 2009=100
		Q086SBEA	
		CPIAUCSL	Consumer Price Index for All Urban Consumers: All Items

Notes: This table gives FRED codes for the variables used in our estimation.

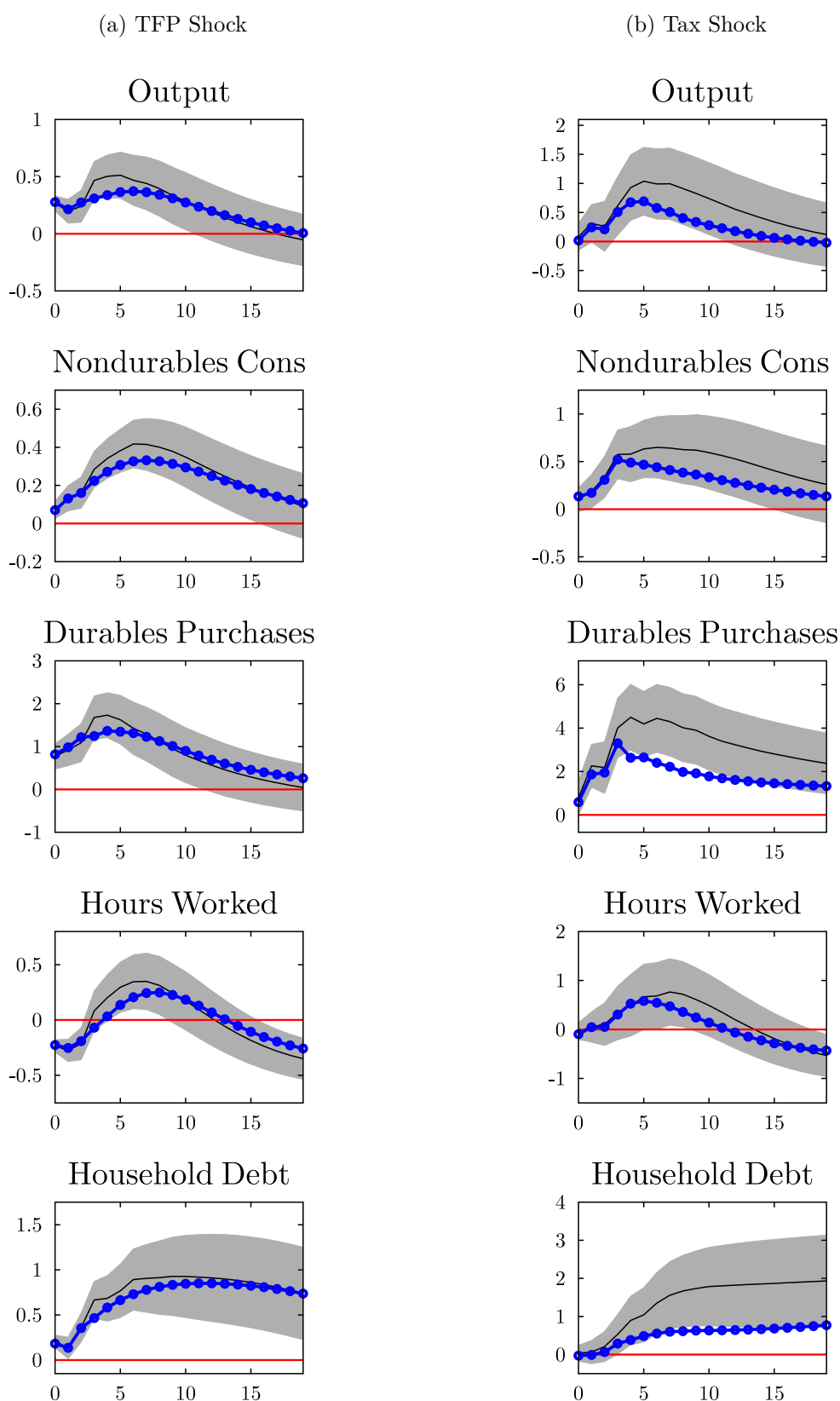
In addition to the FRED data series, we include the Romer and Romer (Romer and Romer, 2010) tax series as τ , available at eml.berkeley.edu/~dromer/. In total, this includes 54 observations of quarterly tax changes.

Figure 28: IR robustness monetary policy vs. baseline



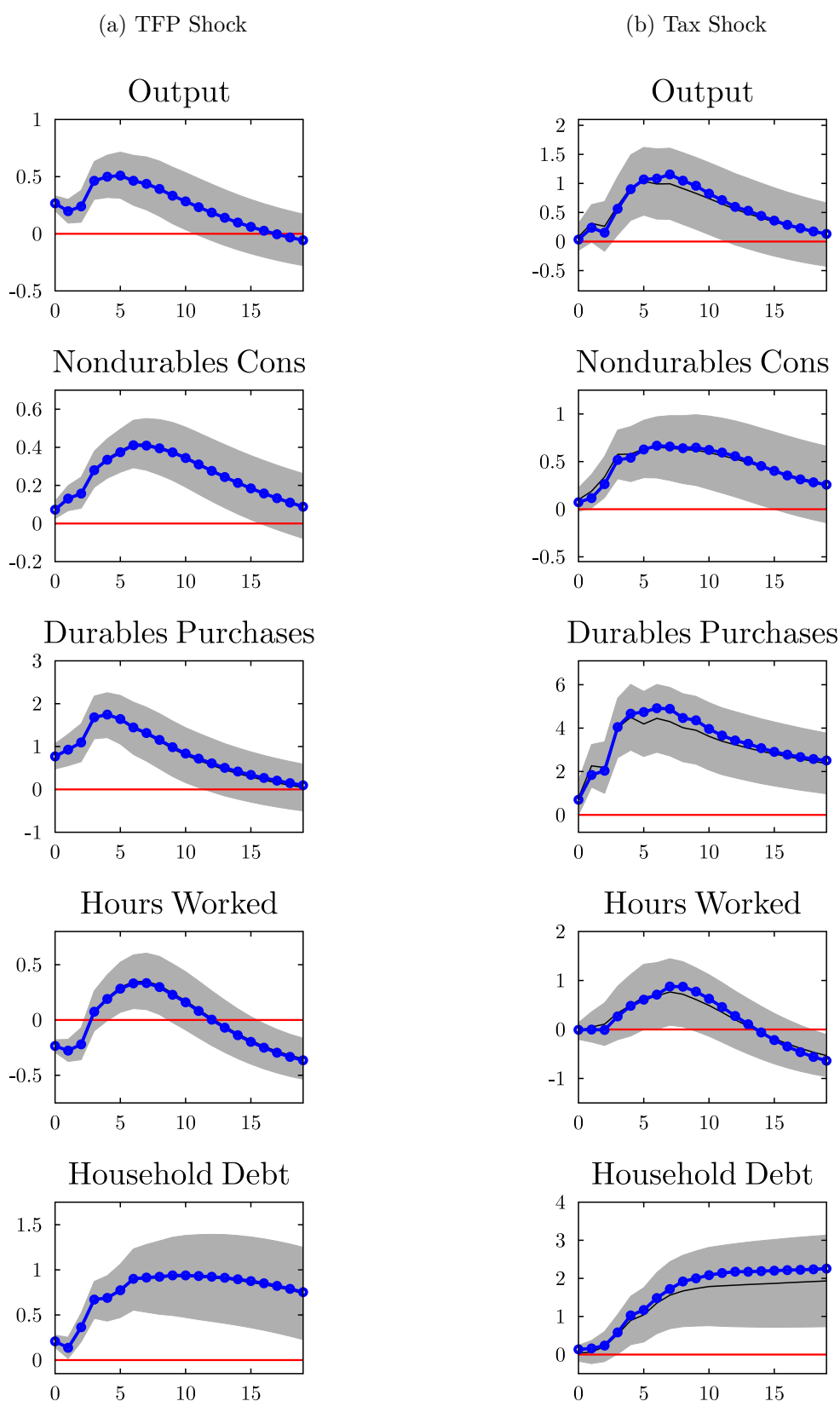
Notes: SVAR impulse responses (black line=baseline) to a one percent increase in TFP (a) and a one percentage point decrease of tax revenues over GDP along with 68% (dark grey) bootstrapped centered confidence intervals with 10.000 replications. The dotted line denotes the impulse responses when controlling for monetary policy (federal funds rate ordered second last).

Figure 29: IR robustness lag length 3 lags vs. baseline



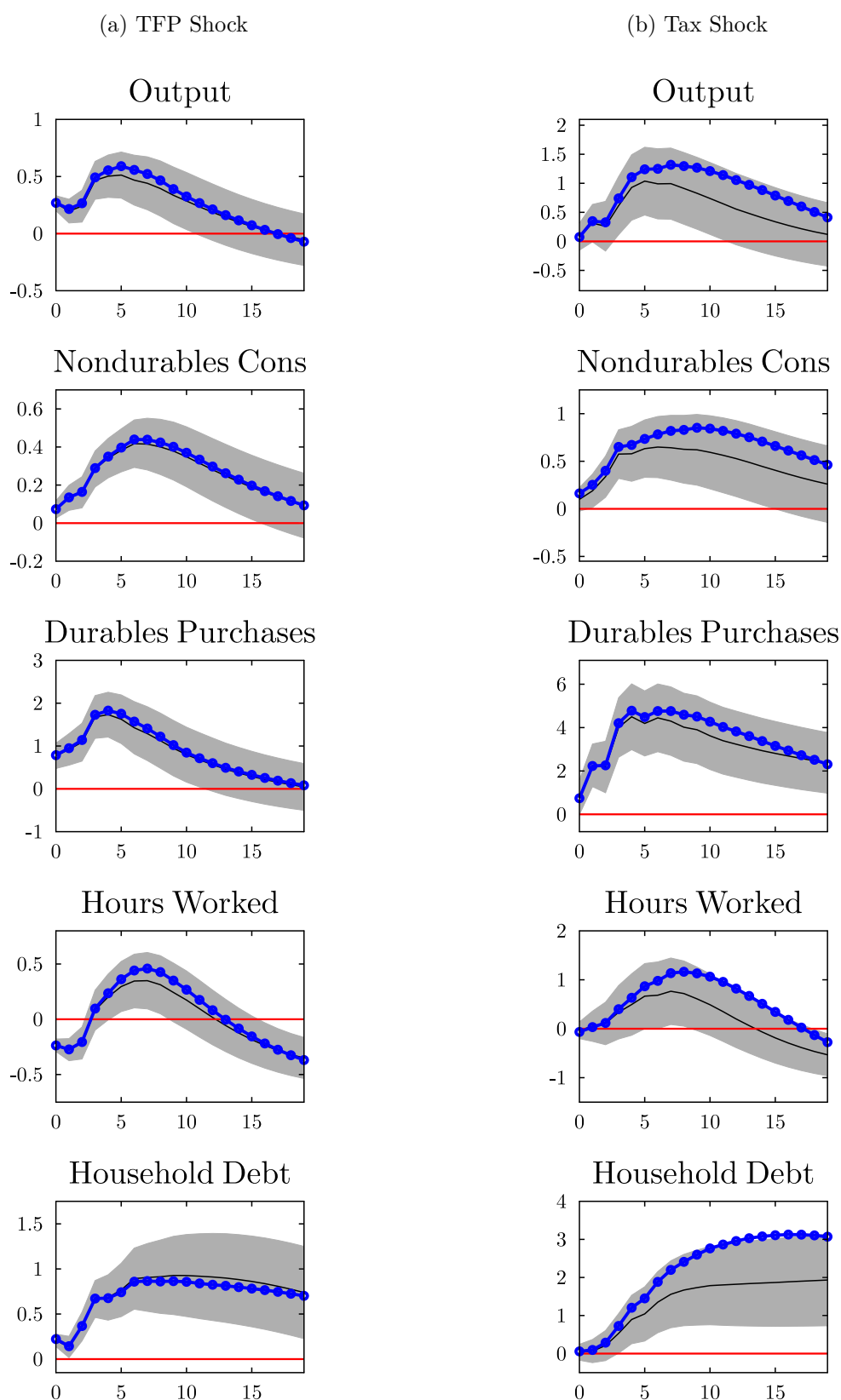
Notes: SVAR impulse responses (black line=baseline) to a one percent increase in TFP (a) and a one percentage point decrease of tax revenues over GDP along with 68% (dark grey) bootstrapped centered confidence intervals with 10.000 replications. The dotted line denotes the impulse responses when the lag length is reduced to 3.

Figure 30: IR robustness lag length 5 lags vs. baseline



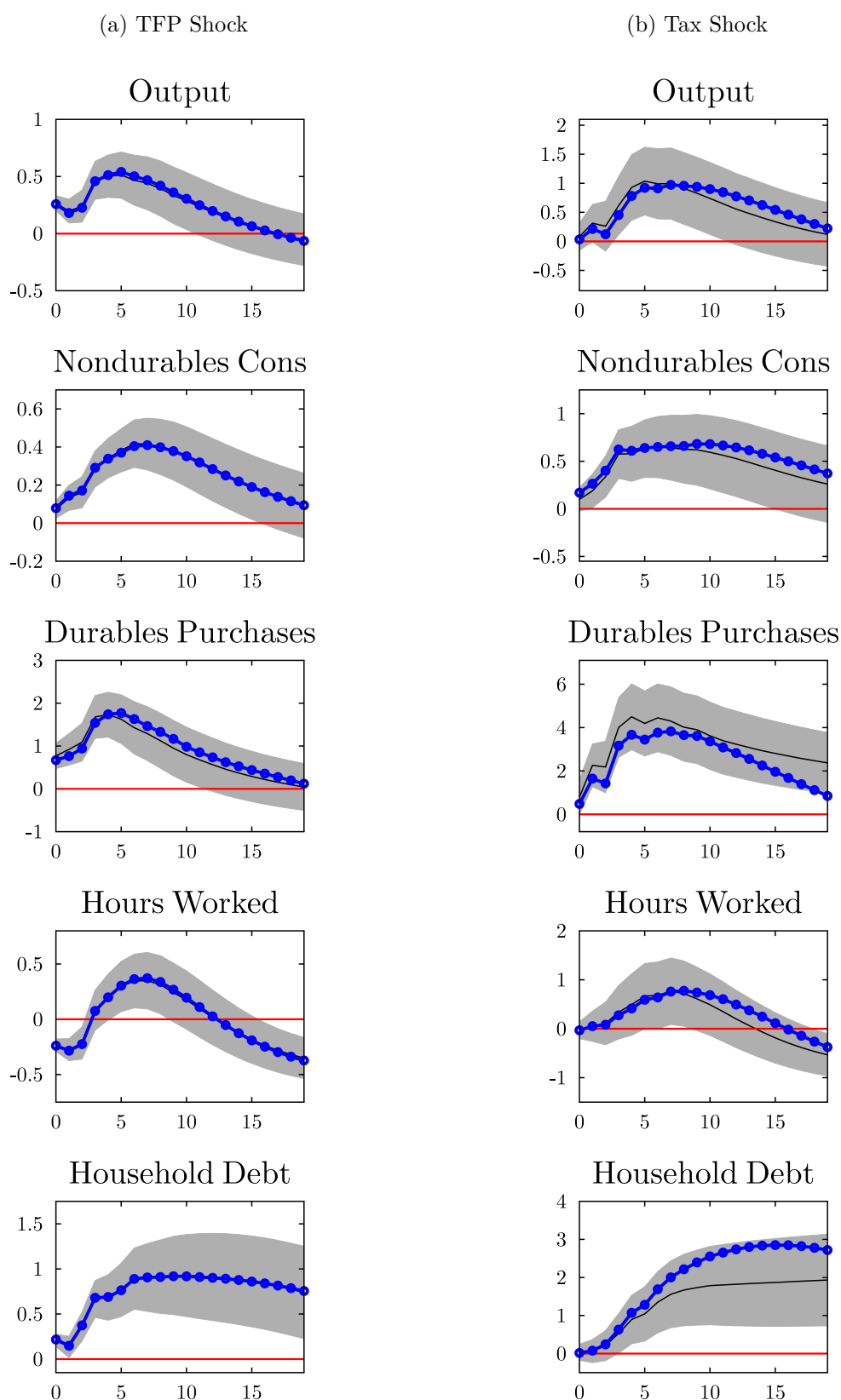
Notes: SVAR impulse responses (black line=baseline) to a one percent increase in TFP (a) and a one percentage point decrease of tax revenues over GDP along with 68% (dark grey) bootstrapped centered confidence intervals with 10.000 replications. The dotted line denotes the impulse responses when the lag length is set to 5.

Figure 31: IR robustness gov't debt vs. baseline



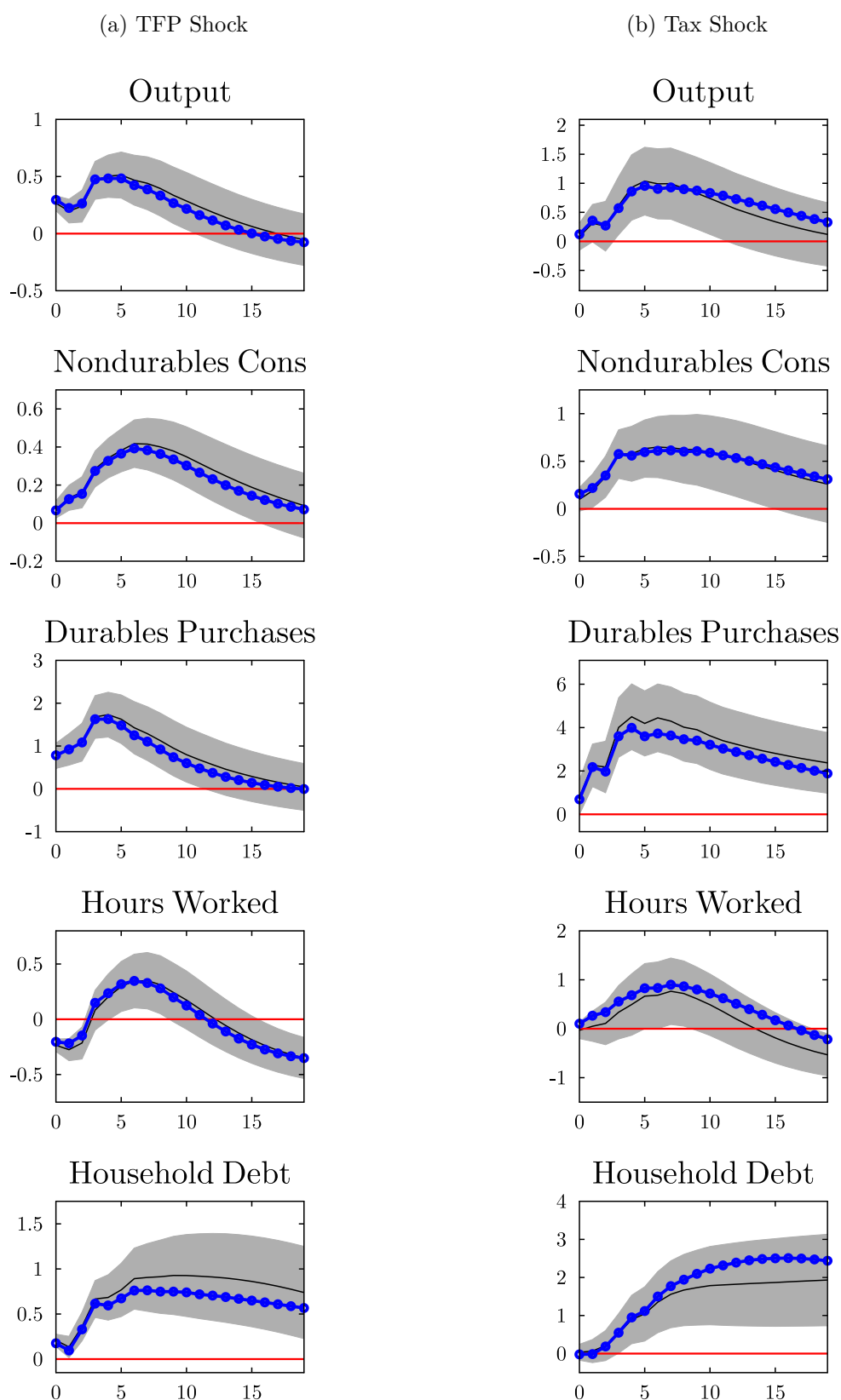
Notes: SVAR impulse responses (black line=baseline) to a one percent increase in TFP (a) and a one percentage point decrease of tax revenues over GDP along with 68% (dark grey) bootstrapped centered confidence intervals with 10.000 replications. The dotted line denotes the impulse responses when controlling for government debt (debt-over-GDP) ordered third.

Figure 32: IR robustness quadratic trend vs. baseline



Notes: SVAR impulse responses (black line=baseline) to a one percent increase in TFP (a) and a one percentage point decrease of tax revenues over GDP along with 68% (dark grey) bootstrapped centered confidence intervals with 10.000 replications. The dotted line denotes the impulse responses when estimating including a quadratic trend instead of a linear one.

Figure 33: IR robustness unemployment vs. baseline



Notes: SVAR impulse responses (black line=baseline) to a one percent increase in TFP (a) and a one percentage point decrease of tax revenues over GDP along with 68% (dark grey) bootstrapped centered confidence intervals with 10.000 replications. The dotted line denotes the impulse responses when estimating controlling for unemployment, ordered third.

5 Forced deleveraging, fiscal policy rules and the zero lower bound

5.1 Introduction

The financial crisis was preceded by a housing boom while interest rates remained low. Miss-incentives to borrow against low values of collateral and seemingly ever-rising real estate values promised high returns. However, the real estate market in the U.S. collapsed, causing a severe recession.

In this paper, I extend the Iacoviello (2005) housing sector DSGE model by a standard fiscal sector and including fiscal rules. I then investigate the impact of house price shocks at the Zero Lower Bound on interest rates (ZLB) and the interaction with fiscal policy.

Recent research has underlined the importance of financial frictions for modeling this crisis since such declines in house prices are difficult to model and cannot sufficiently be explained by preference shocks alone, as Iacoviello (2005) and Iacoviello and Guerrieri (2015) note. Therefore, a shock that is capable of matching the empirically marked drop in prices is required, involving housing debt and forced deleveraging.

I model the starting point of the financial crises as a devaluation of the stock of housing. Eggertsson and Krugman (2012) and Iacoviello (2015) model this shock as a deterioration in the ability to borrow. Such a sudden downward revaluation of real estate leads to a decline in households' ability to use their housing position as collateral for borrowing. Since households borrow to fund additional property and consumption, a devaluation shock causes a recession.

In the literature, the interaction between government debt, devaluation shocks and fiscal policy has been ignored so far. In reaction to the recent economic downturn, governments borrowed to establish stimulus packages, leading to a surge in public debt. In the U.S., the stimulus package at the end of 2007 caused government debt to rise from 63% debt-to-GDP to over 100% at the end of 2014. To reduce public debt, taxes can be increased, or the fiscal authority can lower government spending. In this paper, I ask which of the two has the milder impact on output.

In a calibrated DSGE model, I combine the housing sector suggested in Iacoviello (2005) amended for fiscal rules, as indicated in Fernández-Villaverde et al. (2011),

to investigate the impact of devaluation shocks on borrowers and lenders. These rules propose to raise taxes or reduce spending to lower public debt.

I proceed with an economy in which the fiscal authority does not have access to lump-sum taxes so that it can only levy distortionary taxes on consumption and labor. I find that when the economy is sufficiently far away from the ZLB, the policy of how to reduce government debt is irrelevant for the impact on output. Raising taxes or cutting spending leads to almost the same reduction in output.

I then investigate if this result still holds, when interest rates cannot be lowered further. I proceed by applying the algorithm suggested by Iacoviello and Guerrieri (2015) and find that the ZLB acts as an amplifier of the devaluation shock since the monetary authority is incapable of supporting the process of recovery with conventional monetary policy. In such a situation, cutting spending leads to a more severe recession, compared to raising taxes. Thus, variation in real estate prices and households' ability to borrow has a substantial impact on business cycle dynamics at the ZLB.

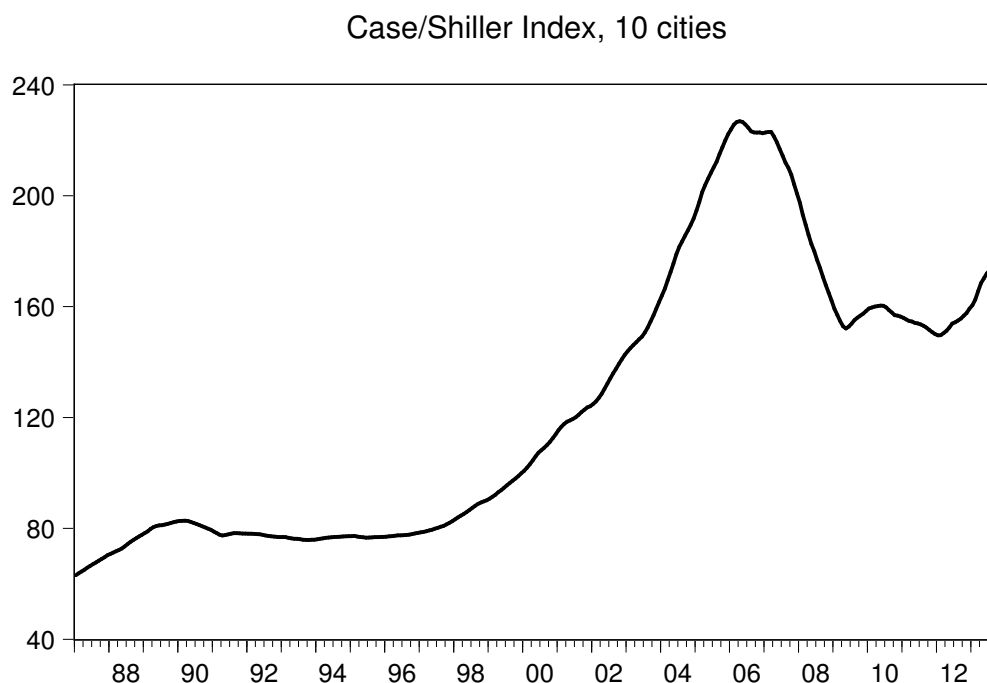
The remainder of this paper is structured as follows: The next section introduces the housing literature followed by empirical evidence for house price shocks. Section 3 introduces the model followed by the simulation results abstracting from the ZLB. I then present the Iacoviello and Guerrieri (2015) algorithm which can be used to solve the DSGE model with occasionally binding constraints. The next section, then discusses the results, followed by concluding remarks.

5.2 House prices and devaluation shocks in literature

Residential property in the United States in the late 90s accounted for 25% of household wealth, about 35% in the United Kingdom and 70% in Germany (Justiniano et al., 2010). As Iacoviello (2005) and Iacoviello (2011) show, it is a major component of wealth and also positively correlated with consumption.

Owning real estate can be seen as an asset, collateralizable subject to borrowing constraints. However, owning property is risky, and adverse price changes force households to rebalance their portfolios, which can cause recessions as Iacoviello (2015) shows.

The price of housing has been a leading indicator for eight post-WW2 U.S. recessions since it co-moves with output and consumption (Kydland et al., 2012). However, shocks to house prices do not simply precede recessions, but can also be responsible for them, as Leamer (2007) notes. Figure 34 depicts house prices having increased by 70% between 1990 and 2007. By mid-2009, house prices in the ten largest cities in the U.S. plummeted by almost 40 percent.

Figure 34: Case-Shiller Index

Notes: Development of house prices between 1987 and 2014 for the Case-Shiller house price index for ten city homes.

Mortgage contracts make it possible to translate the value of the stock of housing as collateral into current availability of mortgages for households (Calza et al., 2013). New debt can then be used to invest in new housing or to consume. Several authors have argued that collateral constraints amplify the response to aggregate shocks. For instance, Zeldes (1989) provides panel data evidence for the presence of liquidity-constrained households. Leamer (2007) regards housing as the primary source of variation in business cycle dynamics and shows that financial frictions like borrowing constraints are essential elements to amplify and propagate business cycle fluctuations. Kiyotaki and Moore (1997) and Hubbard (1998) argue that on firms' side, investment decisions depend on firms' measures of their net worth. Therefore, the impact of devaluation shocks is crucial for investment decisions, which in turn cause business cycle fluctuations (Iacoviello, 2005). Also, in the previous chapter it is found that a fraction of households in the economy are borrowing-constrained, which I follow in this chapter.

After devaluation shocks, the borrowing constraint becomes tighter. Forced deleveraging then leads to a reduction in wealth and consumption, lower invest-

ment in housing and eventually to a real estate sell-off that lowers house prices. In the case of falling asset prices together with collateralized property, devaluation shocks are not only a consequence of economic downturns but also causal for them (Iacoviello, 2005). Iacoviello (2015) shows that of the output reduction of 13 percent during the Great Recession, 2/3 can be explained by a combination of devaluation shocks, housing preference shocks, and default shocks.

Forced deleveraging has several implications for different types of households, depending on their time preference. On the one hand, borrowers are forced to deleverage and hence face a negative wealth effect, though, on the contrary, lending households benefit from lower house prices and can acquire additional real estate since they are not borrowing-constrained. Iacoviello (2005) therefore distinguishes into two types of households, patient and impatient ones.

As a consequence of the downturn at the end of 2007, the Federal Reserve reduced the nominal interest rate until it hit almost zero by December 2008. An instrument to react to a recession with lower interest rates to stimulate the economy and encourage borrowing was removed. Therefore, the transition of devaluation shocks at the ZLB and the impact of debt stabilization is crucial. As it is well known, the ZLB can amplify devaluation shocks, which is also found in this paper. To investigate business cycle properties at the ZLB, Iacoviello and Guerrieri (2015) introduce an algorithm to piecewise solve DSGE models with occasionally binding constraints so that I can investigate the impact of fiscal policy rules at the ZLB.

5.3 Empirical evidence for house price shocks

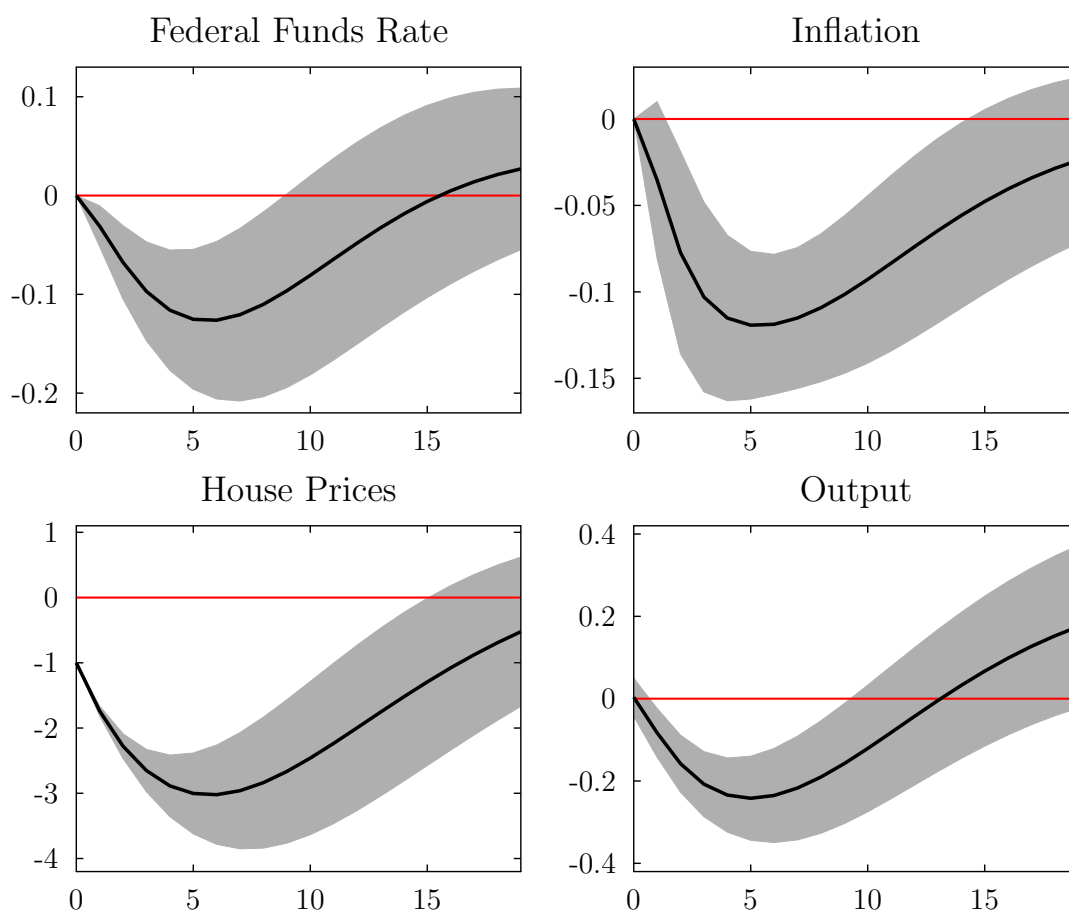
In this section, I examine the impact of a house price shock in a four variable vector autoregression. Since devaluation shocks are not directly observable, I use an orthogonal shock to the Case-Shiller price index as a proxy for house price shocks. The vector autoregression is ordered as in Iacoviello (2005) and data sources can be found in the appendix in section 5.7 so that $X_t = [i, \pi, q, Y]'_t$,⁴² in which i denotes the level of the Federal Funds rate, π is the first difference in the log of the CPI index from the respective quarter one year ago, q denotes the log of the S&P Case-Shiller home price index for 10 cities and Y is the log of real Gross Domestic Product. The reduced form equation is given by

$$X_t = A(L)X_{t-1} + u_t, \quad (48)$$

⁴²The ordering did only marginally affect the results, and qualitative responses remain similar.

in which X_t contains the endogenous variables in the order specified above. $A(L)$ is a lag polynomial of order two and u_t denotes reduced form residuals. The estimation includes two lags, suggested by the Schwarz information criteria (Schwarz, 1978), a constant term and a linear time trend and is estimated with ordinary least squares.⁴³

Figure 35: VAR evidence on house price shocks



Notes: This Figure indicates impulse responses conditional on a negative one percent structural shock to house prices, in a vector autoregression with the following ordering: $X = [i, \pi, q, Y]'_t$. i denotes the end of quarter Federal Funds rate, π is the quarter to four quarters ago log difference in the consumer price index, q denotes the log of the Case-Shiller index, and Y is real per capita gross domestic output. The reduced form residual variance-covariance matrix is Cholesky decomposed. The estimation includes a linear time trend and a constant. The shaded area denotes 68% bootstrapped confidence bands with 10,000 bootstrapped replications.

Since reduced form VARs are known for generating correlated innovations, which restrict disentangling structural shocks from each other, structural shocks are recovered by Cholesky decomposition to the variance-covariance matrix of the reduced form residuals in Equation 48, with the order as specified above. The results of a

⁴³Results are robust to including a quadratic trend as well.

negative structural shock to house prices of the size of one percent are depicted in Figure 35.

It depicts that a one percent decline in house prices induces a recession in the economy. All variables react in a hump-shaped manner with a trough at seven quarters after the shock, before returning to pre-shock values. The Federal Funds rate, inflation and output co-move. Although house prices decline, households cannot accommodate the reduction in their real estate value by expanding their real estate position. Lower wealth then causes output to fall by roughly 0.2%. Due to the recession, house prices continue to fall, with a maximum absolute value of decline of 3%.

In the next section, I model a decrease to house prices, as a devaluation shock and compare policy rules of the fiscal authority to stabilize government debt.

5.4 Model

In this section, I present a DSGE model with financial frictions and a fiscal sector. The model is populated by a continuum of infinitely-lived households who differ in their willingness to postpone consumption into the future. It is based on the model in Iacoviello (2005), extended to include a fiscal sector with policy rules as in Fernández-Villaverde et al. (2011) and amended for devaluation shocks as suggested in Eggertsson and Krugman (2012) and Iacoviello (2015). The model is also solved when the ZLB holds by applying an algorithm suggested by Iacoviello and Guerrieri (2015)

The fiscal authority has access to distortionary taxes only and uses the generated funds for fiscal spending. Spending or taxes are adjusted according to a rule while the other fiscal measure follows an AR(1) process. The government engages in accumulating debt, which only patient households hold. This debt is safe in the sense that the government will always repay it in the subsequent period, plus interest. Calvo price setting based on monopolistic competition at the retailer level creates sticky prices (Calvo, 1983).

5.4.1 Households

In a discrete-time setup, there exists a continuum of patient and impatient households with identical utility functions, as well as entrepreneurs. Preferences are separable in consumption, real estate and disutility from labor. These different types of agents differ in their time preference, which generates borrowers and lenders. House-

holds earn after-tax labor income, consume and may invest or borrow via private debt.

In contrast to the theoretical model in chapter 4, I build a model with three types of households so that the private sector and also the firm sector can be indebted.

Impatient households discount future consumption more heavily with their discount factor β_b being smaller than the one of the patient households (β_l), which creates borrowers and lenders. Both types of agents invest or borrow in one-period bonds that pay an interest in the subsequent period. They provide labor services at their individual real wage rate w_t .⁴⁴ Impatient households maximize utility with respect to their flow of funds constraint and subject to a borrowing constraint, which is always binding (see Iacoviello (2005, p 744) on the derivation for this).

Patient households: Patient households (subscripted with l) derive utility from consumption, real estate, leisure, and supply labor $n_{l,t}$ to entrepreneurs for which they receive the real wage rate $w_{l,t}$, subject to a time varying labor tax τ_t^n .

They may invest in three types of assets: physical real estate $h_{l,t}$ at price q_t and financial assets $b_{l,t}$ (private debt) and gb_t (government bonds). Patient households lend to impatient households and entrepreneurs so that their steady-state private debt position is negative, while it is positive for impatient households. A negative steady-state bond position can be seen as negative debt, through which this type of household saves. Lending to the government is solely possible for patient households as done in, for instance, Eggertsson and Krugman (2012). Private bonds yield a real return of $\frac{R_{l,t-1}b_{l,t-1}}{\pi_t}$, whereas government bonds yield a return of $\frac{i_{l,t-1}gb_{t-1}}{\pi_t}$ and $\pi_t = \frac{P_t}{P_{t-1}}$. Both interest rates are interconnected through a no-arbitrage relation.

Consumption $c_{l,t}$ is taxed also at a time-varying tax rate τ_t^c . Households' preferences are given by

$$E_0 \sum_{t=0}^{\infty} \beta_l^t \left(\ln c_{l,t} + j \ln h_{l,t} - \frac{n_{l,t}^\eta}{\eta} + \varphi \ln \frac{M_{l,t}}{P_t} \right), \quad (49)$$

where E_0 denotes the expectation operator conditional on information available at $t = 0$ and j is a preference parameter for housing. η is the weight on labor disutility and φ denotes the weight on utility received from real money holdings $\frac{M_{l,t}}{P_t}$.

⁴⁴In steady-state, patient labor supply is more productive so that their equilibrium wage is higher than the one of the impatient labor input as in, for instance, (Justiniano et al., 2015).

The discount factor of the household is bounded between zero and unity. Patient households maximize discounted lifetime utility, given their budget constraint

$$\begin{aligned}
 c_{l,t}(1 + \tau_t^c) + q_t \Delta h_{l,t} + \frac{b_{l,t-1} R_{t-1}}{\pi_t} + gb_t &= (1 - \tau_t^n) w_{l,t} n_{l,t} + b_{l,t} \\
 + \frac{gb_{t-1} i_{t-1}}{\pi_t} + F_t - \frac{\Delta M_{l,t}}{P_t}.
 \end{aligned} \tag{50}$$

Lenders receive after-tax labor income $(1 - \tau_t^n) w_{l,t} n_{l,t}$, interest from lending, $b_{l,t}$, to patient households plus last period's amount $\frac{b_{l,t-1} R_{t-1}}{\pi_t}$ and lending to the government $\frac{gb_{t-1} i_{t-1}}{\pi_t}$ plus lump-sum profits F_t from owning the retailer, as well as changes in the real money holdings $\frac{\Delta M_{l,t}}{P_t}$. q_t is the real price of property.

Note that the interest received for private bonds shows up on the left hand side of the constraint, is in fact an income since $b_{l,t}$ is negative in steady-state. Δ is the difference operator. Patient households do not face a borrowing constraint so they can choose their optimal lending level, which is negative for both debt positions in steady-state. First order conditions with respect to $c_{l,t}$, $b_{l,t}$, $n_{l,t}$, $h_{l,t}$, and gb_t are given by

$$\lambda_{l,t} = \frac{1}{c_{l,t}(1 + \tau_t^c)} \tag{51}$$

$$\lambda_{l,t} = \beta_l E_t \lambda_{l,t+1} \frac{R_t}{\pi_{t+1}} \tag{52}$$

$$w_{l,t} = n_{l,t}^{\eta-1} c_{l,t} \frac{1 + \tau_t^c}{1 - \tau_t^n} \tag{53}$$

$$\lambda_{l,t} q_t = \frac{j}{h_{l,t}} + \beta_l E_t (\lambda_{l,t+1} q_{t+1}) \tag{54}$$

$$\lambda_{l,t} = \beta_l E_t \lambda_{l,t+1} \frac{i_t}{\pi_{t+1}}, \tag{55}$$

and a standard FOC with respect to real money holdings, which is irrelevant for the equilibrium since utility is separable in money balances. $\lambda_{l,t}$ is the Lagrangian multiplier on the budget constraint.

Impatient households: Impatient households face the same additive separable utility function as patient households with their respective choice variables. They also maximize utility with variable definition as given for the lender.

$$E_0 \sum_{t=0}^{\infty} \beta_b^t \left(\ln c_{b,t} + j \ln h_{b,t} - \frac{n_{b,t}^\eta}{\eta} + \varphi \ln \frac{M_{b,t}}{P_t} \right) \tag{56}$$

with definitions as above, but subscripted with b , for borrowers. They maximize utility with respect to their budget constraint

$$c_{b,t}(1 + \tau_t^c) + \frac{b_{b,t-1}R_{t-1}}{\pi_t} + q_t\Delta h_{b,t} = b_{b,t} + (1 - \tau_t^n)w_{b,t}n_{bt}, - \Delta \frac{M_{b,t}}{P_t}. \quad (57)$$

Impatient households face a lower discount rate, $\beta_b \in (0, 1)$, than patient households, which makes them borrowers so that their steady-state private debt position is positive. A borrowing constraint limits the amount households can borrow defined by

$$b_{b,t} \leq E_t \frac{\theta_t m_b q_{t+1} h_{b,t} \pi_{t+1}}{R_t}. \quad (58)$$

Equation 58 denotes an upper limit for borrowing, restricted by a fraction m_b (loan-to-value) of the real value of their stock of housing. Setting the value of the discount factor to be smaller than the one of the patient households, ensures that impatient households are borrowers in the steady-state and for all shock paths, as shown in Iacoviello (2005, p. 744). For the simulation below, I assume that the constraint holds with equality, so that households borrow up to their limit.

θ is the parameter which I will use as the devaluation shock described below. In steady state, this parameter is set to one.

FOCs with respect to $c_{b,t}$, $b_{b,t}$, $h_{b,t}$ and $n_{b,t}$ are given by

$$\lambda_{b,t} = \frac{1}{c_{b,t}(1 + \tau_t^c)} \quad (59)$$

$$\lambda_{b,t} = \beta_b E_t \lambda_{b,t+1} \frac{R_t}{\pi_{t+1}} + l m_{b,t} R_t \quad (60)$$

$$\lambda_{b,t} q_t = \frac{j}{h_{b,t}} + \beta_b E_t (q_{t+1} \lambda_{b,t+1} + \theta_t l m_{b,t} m_b q_{t+1} \pi_{t+1}) \quad (61)$$

$$w_{b,t} = n_{b,t}^{\eta-1} c_{b,t} \frac{1 + \tau_t^c}{1 - \tau_t^n} \quad (62)$$

in which $l m_{b,t}$ is the Lagrange multiplier on the borrowing constraint.

Entrepreneurs: Entrepreneurs own the capital stock and produce a homogeneous intermediate good by combining two types of labor inputs, capital and real estate. Entrepreneurs are less willing to postpone consumption compared to patient households, but more ready to do so compared to impatient households, so that their discount factor γ , is set to lie between the discount factors of the impatient households and the patient one so that $\gamma \in (\beta_b, \beta_l)$. Entrepreneurs maximize utility and their production technology is given by

$$Y_t = A_t k_{t-1}^\mu h_{t-1}^\nu n_{l,t}^{\alpha(1-\mu-\nu)} n_{b,t}^{(1-\alpha)(1-\mu-\nu)}. \quad (63)$$

α , μ and ν (all larger than zero) are the capital share in production and the elasticities of output to real estate, respectively. Capital k depreciates at a constant rate δ and entrepreneurs also have access to private debt b_t with also a positive position in steady-state. However, entrepreneurs do not receive utility from housing, unlike the two types of households and use housing instead in producing the intermediate good. They maximize discounted lifetime utility as

$$E_0 \sum_{t=0}^{\infty} \gamma^t \ln c_t, \quad (64)$$

subject to their budget constraint

$$\frac{Y_t}{X_t} + b_t = c_t(1 + \tau_t^c) + q_t \Delta h_t + \frac{R_{t-1} b_{t-1}}{\pi_t} + n_{b,t} w_{b,t} + n_{l,t} w_{l,t} + I_t + \xi_t, \quad (65)$$

in which $\xi_{k,t}$ denotes capital adjustment costs evolving according to $\xi_t = \psi_k (I_t/k_{t-1} - \delta)^2 k_{t-1} / 2\delta$.

Retailers then purchase the intermediate good at the wholesale price P_t^w , transform it into final goods and sell it at price P_t so that the markup is given by $X_t = \frac{P_t}{P_t^w}$, as in Iacoviello (2005).

Entrepreneurs' ability to accumulate debt is also limited by a borrowing constraint reading

$$b_t \leq E_t \frac{\theta_t q_{t+1} h_t \pi_{t+1} m}{R_t}. \quad (66)$$

In this equation, m is the entrepreneur specific loan-to-value parameter. First order conditions for entrepreneurs with respect to c_t , b_t , k_{t+1} and h_t are

$$\lambda_t = \frac{1}{c_t(1 + \tau_t^c)} \quad (67)$$

$$\lambda_t = \gamma E_t \lambda_{t+1} \frac{R_t}{\pi_{t+1}} + l m_t R_t \quad (68)$$

$$\begin{aligned} \lambda_t \left(1 + \frac{\psi}{\delta} \left(\frac{I_t}{k_{t-1}} - \delta\right)\right) &= \gamma E_t \lambda_{t+1} \left(\frac{\mu Y_{t+1}}{X_{t+1} k_t} + 1 - \delta\right) \\ &+ \frac{\psi}{\delta} \left(\frac{I_{t+1}}{k_t} - \delta\right) \left(1/2 \left(\frac{I_{t+1}}{k_t} + \delta\right) + 1 - \delta\right) \end{aligned} \quad (69)$$

$$\lambda_t q_t = \gamma E_t (\lambda_{t+1} q_{t+1} + \theta_t l m_t m q_{t+1} \pi_{t+1}) \quad (70)$$

with lm_t being the Lagrange multiplier on the borrowing constraint for the entrepreneur. FOCs with respect to patient and impatient labor inputs are given by

$$w_{l,t} = \alpha(1 - \mu - \nu) \frac{Y_t}{X_t n_{l,t}} \quad (71)$$

$$w_{b,t} = (1 - \alpha)(1 - \mu - \nu) \frac{Y_t}{X_t n_{b,t}}. \quad (72)$$

5.4.2 Aggregate price level and price dispersion

Prices in this model are sticky, which is introduced through retailers, indexed with z , following Iacoviello (2005). These represent a continuum normalized to unity, buy intermediate goods from entrepreneurs at a competitive wholesale price P_t^w and differentiate the goods without costs into $Y_t(z)$ selling at price $P_t(z)$. I follow Iacoviello (2005) so that the retailers choose the sale prices, given the wholesale price and their demand curve. The Calvo price setting mechanism introduces the stickiness so that retailers can only re-optimize prices with probability $(1 - calvo)$ in every period (Calvo, 1983). In contrast, with probability $calvo$, the retailer has to stick with its current period price. $P_t^*(z)$ is denoted as the reset price with the respective demand given by $Y_{t+k}^*(z) = \frac{P_t^*(z)^\epsilon}{P_{t+k}^\epsilon} Y_{t+k}$. An optimal reset price then solves the following equation

$$\sum_{k=0}^{\infty} calvo^k E_t \left\{ \Lambda_{t,k} \left(\frac{P_t^*(z)}{P_{t+k}} - \frac{X_t}{X_{t+k}} \right) Y_{t+k}^*(z) \right\} = 0, \quad (73)$$

in which $\Lambda_{t,k}$ is the patient households' discount factor, since they own the retailers and X_t again is the markup. Additionally, as in Iacoviello (2005) and Schmitt-Grohe and Uribe (2004) the aggregate price level evolves according to

$$P_t = ((1 - calvo) P_t^{*1-\epsilon} + calvo P_{t-1}^{1-\epsilon})^{\frac{1}{1-\epsilon}}, \quad (74)$$

which is a combination of the optimal reset price P_t^* and the aggregate price level from the previous period, P_{t-1} .

5.4.3 The public sector

The government collects distortionary taxes, issues bonds and consumes a fraction of final output. The monetary authority sets the nominal interest rate i according to

a Taylor rule, targeting the steady state real interest rate (\bar{r}) and output deviations as

$$i_t = i_{t-1}^{rR} \left(\pi_{t-1}^{1+r\pi} \frac{Y_{t-1}^{rY}}{Y_t} \bar{r} \right)^{1-rR}. \quad (75)$$

$r\pi$, rR and rY are the reaction to the inflation rate, interest rate and output, from the previous quarter, and \bar{r} is the steady-state real interest rate. The fiscal authority collects distortionary labor and consumption taxes, τ_t . Both types of taxes are time-varying and deviate from zero in steady-state. The government's budget constraint reads as follow

$$g_t + \frac{gb_{t-1} i_{t-1}}{\pi_t} = gb_t + \tau_t^c (c_t + c_{l,t} + c_{b,t}) + \tau_t^n (n_{l,t} w_{l,t} + n_{b,t} w_{b,t}), \quad (76)$$

in which g_t denotes government spending and τ is the respective tax rate. The law of motion for the devaluation shock is as follows

$$\log \theta_t = (1 - \rho^\theta) \log \bar{\theta} + \rho^\theta \log \theta_{t-1} - e_t^\theta \quad (77)$$

and e_t^θ is the devaluation and shock, and an upper bar denotes the steady-state value of the respective variable. The rules for taxes and spending depend on the fiscal regime when government debt deviates from its steady state value as described in Fernández-Villaverde et al. (2011). Under one, taxes are automatically increased (TI) and government spending is exogenous, given by

$$\log \tau_t = \rho^\tau \log \tau_{t-1} + (1 - \rho^\tau) \log \bar{\tau} + \phi_\tau^{b,y} \left(\frac{gb_{t-1}}{Y_{t-1}} - \frac{\bar{gb}}{\bar{Y}} \right) + \epsilon^\tau \quad (78)$$

$$\log g_t = (1 - \rho^g) \log \bar{g} + \rho_t^g \log g_{t-1} + \epsilon^g. \quad (79)$$

Under the second regime, spending is cut (SC) according to a fiscal policy rule in order to lower government so that both tax rates are exogenous, given by

$$\log g_t = \rho^g \log g_{t-1} + (1 - \rho^g) \log \bar{g} + \phi_g^{b,y} \left(\frac{gb_{t-1}}{Y_{t-1}} - \frac{\bar{gb}}{\bar{Y}} \right) + \epsilon^g \quad (80)$$

$$\log \tau_t = (1 - \rho^\tau) \log \bar{\tau} + \rho_t^\tau \log \tau_{t-1} + \epsilon^\tau. \quad (81)$$

In both regimes, the reaction of spending to deviation from steady state bond-over-output level is estimated to be negative, as seen below, whereas taxes react positively to stabilize debt. The estimation of the parameters is given below. I compare both regimes to investigate, which alternative stabilizes debt with the smaller decrease

in output. Therefore, either spending is cut (SC) according to Equation 80 or taxes are increased in accordance with (78).

5.4.4 Equilibrium conditions

The equilibrium is an allocation $(h_{l,t}, h_{b,t}, Y_t, c_t, c_{l,t}, c_{b,t}, b_t, b_{l,t}, b_{b,t}, gb_t, I_t, n_{l,t}, n_{b,t}, w_{l,t}, w_{b,t}, X_t, q_t)_{t=0}^{\infty}$, satisfying the first order conditions above. The following then are market clearing conditions

$$0 = b_t + b_{b,t} + b_{l,t} \quad (82)$$

$$\bar{H} = h_t + h_{b,t} + h_{l,t} \quad (83)$$

$$Y_t = c_t + c_{l,t} + c_{b,t} + I_t + \psi_t^k + g_t \quad (84)$$

given the sequence of the shocks as specified together with the relevant transversality conditions. The first equation states that the sum of debt held by entrepreneurs and borrowers equals the position held by the lender. The stock of housing \bar{H} is constant and normalized to one so that no new real estate is created, and real estate does not depreciate⁴⁵. Total output is equal to the sum of individual consumption, investment, government spending and capital adjustment costs. In steady-state, labor inputs are set to equal.

The next subsection gives the simulation results of the responses given the devaluation shock, preceded by the calibration.

5.5 Calibration and results

5.5.1 Calibration

The model is calibrated for quarterly frequencies and parameters are estimated where possible. A summary and description of the data is given in Table 9 and data sources can be found in the appendix in Table 12. I follow Iacoviello (2005) for standard parameters and Fernández-Villaverde et al. (2011) for the coefficients of the automatic stabilizers.

I set β_l , the discount factor of the patient household to 0.993, which matches an average annual interest rate between 1990 and 2015 of 3%. The value of the discount factor of the impatient household (β_b) is set according to Iacoviello (2005), who considers a range between 0.91 and 0.98, based on wealth positions of the households. The discount factor of the entrepreneurs matches then twice the real interest rate and is hence 0.9802. The labor share is an average of the last 25 years and given as

⁴⁵Iacoviello (2005) also considers real estate adjustment costs, which he finds to be close to zero.

Table 9: Calibration of parameters

Param	Value	Description	Target/Description
α	0.640	Wage share patient household	Historical labor share
β_l	0.993	Discount factor patient hh	Match SS annual interest of 3%
β_b	0.950	Discount factor impatient hh	Iacoviello (2005)
γ	0.980	Discount factor entrepreneurs	Iacoviello (2005)
δ	0.025	Depreciation rate	Mertens and Ravn (2011)
ν	0.030	Elasticity of output to housing	Iacoviello (2005)
μ	0.300	Capital share in production	Standard parameter
calvo	0.750	Calvo price parameter	Prices constant for one year on average
\bar{X}	1.050	Markup	5% markup in steady state
m	0.950	Loan to value entrepreneur	2013 Federal Housing Finance Agency data
m_b	0.550	Loan to value borrowers	Iacoviello (2005)
η	1.010	Labor disutility	Iacoviello (2005)
rR	0.790	Interest rate coef. in TR	TR regression coefficient
$r\pi$	0.210	Inflation coef. in TR	TR regression coefficient
rY	0.130	Output coef. in TR	TR regression coefficient
ψ	2.000	Capital adjustment costs weight	Iacoviello (2005)
j	0.100	Housing preference parameter	Match 140% residential housing over output
ϕ^K	2.000	Capital adjustment costs.	Iacoviello (2005)
ρ^θ	0.839	AR parameter valuation shock	Iacoviello (2015)
σ^A	0.029	Std of technology shock	Std of Basu et al. (2006) TFP measure
σ^θ	0.024	Std of θ shock	Iacoviello (2015)
$\bar{\tau}^n$	0.200	SS labor tax rate	Estimate, see below
$\bar{\tau}^c$	0.051	SS consumption tax rate	Estimated, see below
ρ^{τ^n}	0.966	AR coefficient fiscal rule τ^n	Estimated with FRED data
ρ^{τ^c}	0.861	AR coefficient fiscal rule τ^c	Estimated with FRED data
ρ^g	0.98	AR coefficient fiscal rule	Estimated with FRED data
ρ^g	0.957	AR coefficient AR(1)	Estimated with FRED data
$\phi_{\tau^n}^y$	0.049	Output gap stabilization, τ^n	Estimated with FRED data
$\phi_{\tau^c}^y$	0.021	Output gap stabilization, τ^c	Estimated with FRED data
ϕ_g^y	-0.192	Output gap stabilization, g	Estimated with FRED data
$\phi_{\tau^n}^b$	0.089	Debt stabilization, τ^n	Estimated with FRED data
$\phi_{\tau^c}^b$	0.002	Debt stabilization, τ^c	Estimated with FRED data
ϕ_g^b	-0.004	Debt stabilization, g	Estimated with FRED data
\bar{g}/\bar{Y}	0.220	SS gov't spending-to-GDP	FRED: GCEC96 / GDPC1
\bar{b}/\bar{Y}	0.410	SS gov't debt-to-GDP	Estimated with FRED data

Notes: This table denotes the calibration for the parameters of the model. Parameters are estimated with FRED data. Table 12 gives sources for each of the series used. Tax rates are constructed as described in Table 13.

the share of labor compensation in GDP at current prices. The depreciation rate for capital is set to 0.025, implying an annual steady-state depreciation of 10 percent, following Mertens and Ravn (2011). The capital share in production, μ , is set to 0.3, which is a standard parameter in real business cycle models and also given in Iacoviello (2005). Then, the elasticity of output to real estate is set to 0.03 as also done in Iacoviello (2005). The Calvo parameter is 0.75 so that prices stay constant on average for one year.

According to Federal Housing Finance Agency data, the average an-to-value (for borrowers) in 2013 was around 80%. As (Iacoviello, 2011) argues, impatient households' borrowing exceeds the average loan-to-value ratio. In 2013, 45% of new loans exceeded 80% of loan-to-value and 25% exceeded 95%, which is higher than what Iacoviello (2005) and Iacoviello (2011) observe for 2004 values. Hence, I set m , the loan-to-value parameter for the entrepreneur to 0.90, which is in the range of what Iacoviello (2005) estimates with impulse responses matching. I follow Iacoviello (2005) and set m_b to 0.55, which is not a critical parameter for my results.

I estimate a Taylor rule to find the response of monetary policy to output fluctuations, past inflation and past interest rate. A regression of the Federal Funds rate on its lag, the output gap and the first lag of inflation from 1980q1 to 2014q4 leads to coefficients of $r^R = 0.79$, $r^Y = 0.13$ and $r^\pi = 0.21$ ⁴⁶

Capital adjustment costs are set to 2, as in Iacoviello (2005). The housing preference parameters are fixed to equal 0.1 for all three types of households. η is fixed to 1.01, which implies a flat labor supply curve as in Iacoviello (2005). The steady-state markup X equals 5% in steady-state. Government spending is set to the average between 19980 and 2015 which is equal to 0.22.

Next, I calibrate the autoregressive parameters. The AR parameter of government expenditures is estimated using fiscal data taking a value of 0.98. I assume both AR(1) coefficients of the tax rates to be identical to the ones obtained from estimating the fiscal rules. These are estimated by regressing the respective tax rate on deviations from potential output and steady-state government debt, as described in Equation 78 and 80. I extend the data suggested in Fernández-Villaverde et al. (2011) so that sample contains data from 1980 to 2015 data. I use real government consumption expenditures and gross investment, real output and federal debt held by the public to estimate averages. Table 10 expresses descriptive observations and

⁴⁶I use yearly quarter-to-quarter log changes in the CPI as inflation measure and the output gap as linearly detrended output. Using HP-detrended output leads to a coefficient of 0.24 for r^Y , which is double the size of the value by Iacoviello (2005), who does not document the estimation for the output gap. I hence chose the output gap constructed with a linear trend in GDP as regressor in the Taylor rule.

estimation results. Complete sources can be found in the appendix in Table 13, which also incorporates a subsection on the estimation of average consumption and labor taxes.

Table 10: Estimation of autoregressive parameters

	labor tax	consumption tax	government spending
Average	20.04%	5.05%	22% (to y)
ρ	0.9662 (0.0000)	0.8614 (0.0000)	0.9563 (0.0000)
ϕ^y	0.0486 (0.006)	0.0213 (0.0070)	-0.1921 (0.0000)
ϕ^b	0.0031 (0.0885)	0.0009 (0.0024)	-0.0040 (0.0290)

Notes: Descriptive data for government spending, debt and taxes. Averages of the observations and regression results from estimating the autoregressive parameters. P-Values are in parenthesis. U.S. data at quarterly frequencies 1980q1 to 2015q2. Estimating using Equation 80 and 78.

The table shows that government spending averages 22% over output in the sample, whereas the mean of the labor tax rate 20%, and the one of the sales tax is 5%, slightly lower compared to estimates in Fernández-Villaverde et al. (2011). Moreover, Table 10 depicts the estimates of the reaction function of the two regimes, spending cuts and tax increases to stabilize government debt. As the table suggests, all coefficients are significantly different from zero on at least the 5% level.⁴⁷

5.5.2 Simulation results

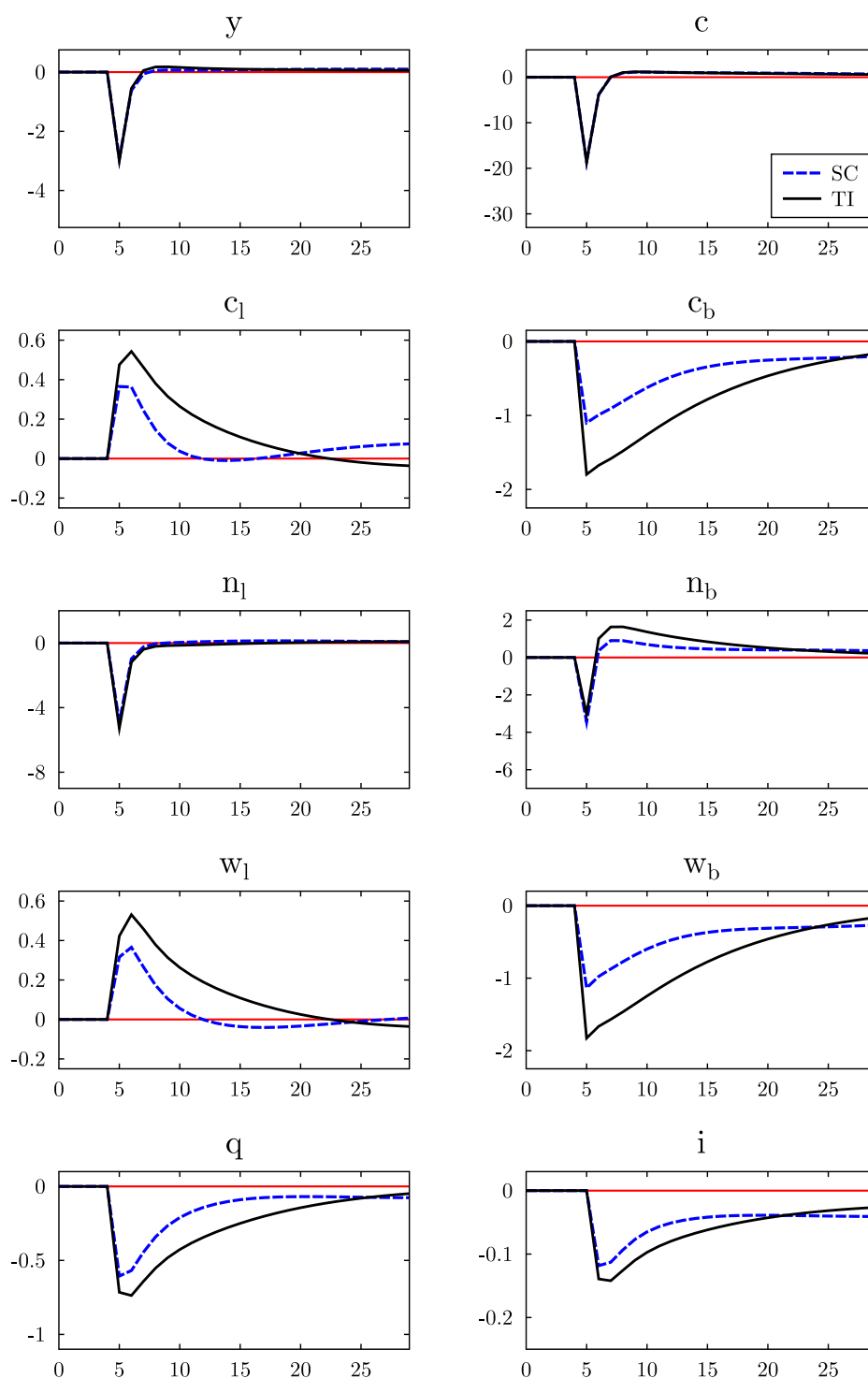
I solve the model using first order perturbation around the non-stochastic steady-state assuming certainty equivalence. I compute impulse responses given a devaluation shock and compare the two different fiscal policy regimes, spending cuts (SC) or tax increases (TI). I follow Iacoviello (2015) where the standard deviation of the devaluation shock is estimated to be equal to 0.024.

Figure 36 and 37 depict the simulation of devaluation shocks in period five and distinguishes between debt stabilization via cutting spending (TC) and increasing taxes (TI). Impulse responses are in percentage deviations from steady-state.

As is obvious, lending and borrowing households are hit asymmetrically. For indebted households, the shock acts like a negative preference shock since the borrowing constraint becomes tighter, leading to forced deleveraging so that debt is

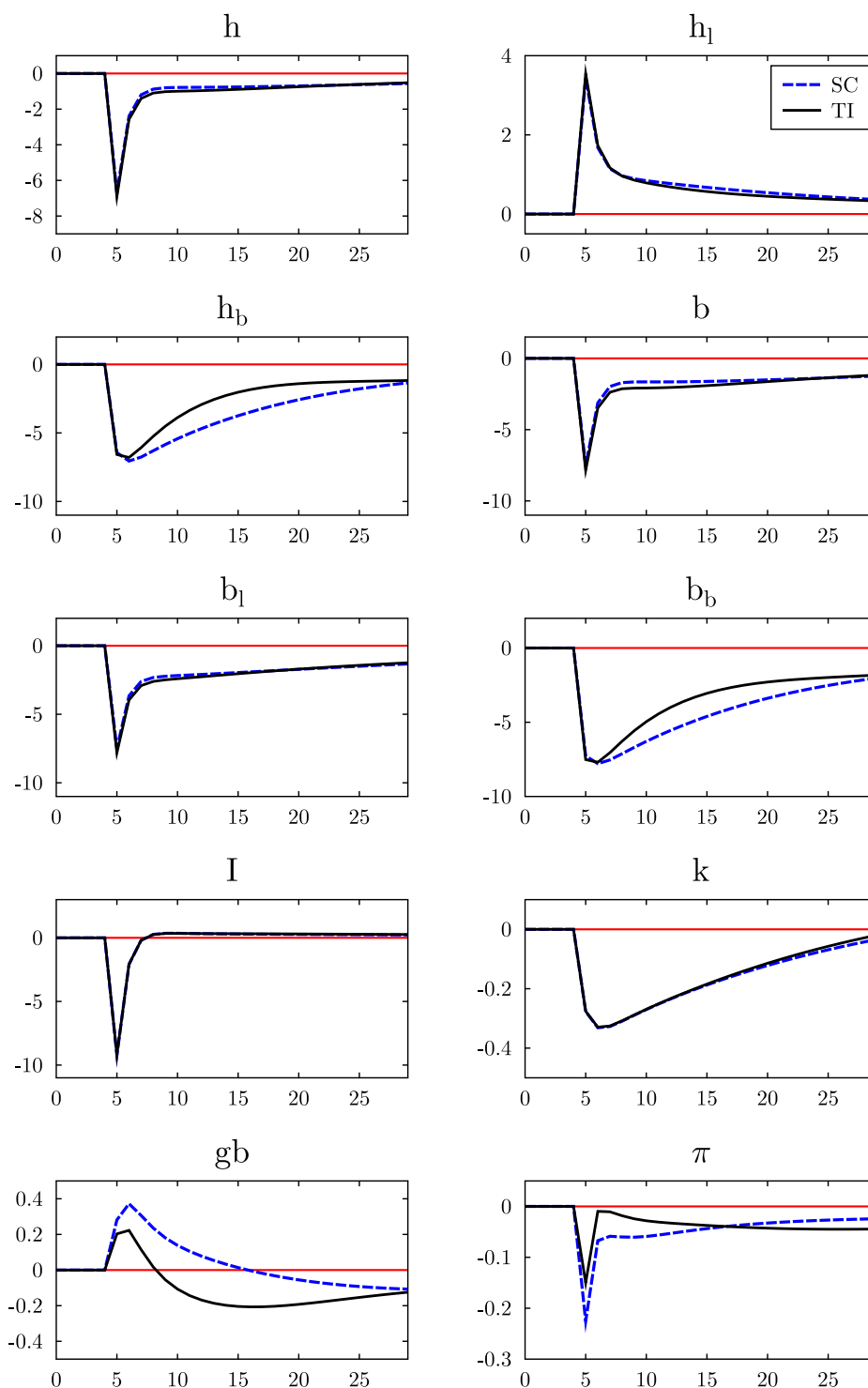
⁴⁷These values above could also be estimated from a VAR approach, as done in the section on total private debt. However, identification of all parameters becomes difficult, as the dimension of the system increases.

Figure 36: Spending cuts and tax increases after devaluation shocks I



Notes: This figure denotes the simulated impulse responses from a one standard deviation shock to the constrained households ability to borrow (devaluation) shock. The black line denotes the impulse response when taxes are increased in response to deviations from steady-state debt-over-output level, whereas the blue line denotes the spending cuts regime. First set of responses, percentage deviations from steady-state.

Figure 37: Spending cuts and tax increases and devaluation shocks II



Notes: This figure denotes the simulated impulse responses from a one standard deviation shock to the constrained households ability to borrow (devaluation) shock. The black line denotes the impulse response when taxes are increased in response to deviations from steady-state debt-over-output level, whereas the blue line denotes the spending cuts regime. Second set of responses, percentage deviations from steady-state.

reduced by 7%. Consequently, a real estate sell-off leads to a decrease of borrowers' property position by 6% and 7% for the entrepreneurs. In contrast, patient households invest in housing, since the real estate price falls and they are not borrowing constrained so that their real estate position increases by about 4%. Since private bond trading is reduced, patient households instead invest in government bonds until interest rates equal out.

Consumption co-moves with housing for indebted agents consistent with Iacoviello (2005) and Iacoviello (2011). Forced deleveraging and a lower stock of housing causes borrowing households to feel poorer so that they reduce their consumption level. This shock has a strong impact on entrepreneurs so that the entrepreneur decreases consumption by 18%, whereas the impatient household reduces consumption by only 1-2%, depending on the regime. The strong reaction observed for the entrepreneur is due to his production function as he uses housing services to produce intermediate goods. A deterioration of real estate used in production and the inability to borrow leads to a reduction in wealth also causing investment to fall. Since utility for these types of households is drawn from consumption only, they are hit the hardest compared to the other types of households. This is in line with Eerola and Määttänen (2012) who suggests that even small shocks can have a severe impact on the net worth of indebted households.

Since total consumption decreases, consumer prices do so as well. Like in most low inflation countries, debt is not indexed to inflation so that falling inflation and sticky prices increases real debt service, which amplifies the negative wealth effect.

The monetary authority responds to lower inflation and a larger output gap with a lower nominal interest rate. Furthermore, on impact lenders and borrowers decrease hours ($n_{l,t}$ and $n_{b,t}$) by 4.5% and 3.8% respectively.

Looking at the difference between the two debt stabilization regimes, there is only a slight difference between spending cuts (SC) and tax increases (TI). In both regimes, output decreases by 3% as a consequence of the reduction in consumption and investment. Therefore, in times when the economy is sufficiently far away from the ZLB, it is irrelevant if taxes are increased or spending is reduced to stabilize deviations from steady-state bond-over-output level regarding output reduction.

The results of the theoretical simulation fit to the VAR estimation results, which also show that house price shocks lead to a reduction in output prices and interest rates.

Next, I compare both debt stabilization regimes once the economy reaches the ZLB.

5.5.3 Binding constraints & the ZLB

The Great Recession led to a period of zero interest rates and high government debt. In this section, I compare the linear baseline simulation above to a version of the model, in which the ZLB is binding only when it is reached after a shock.

Iacoviello and Guerrieri (2015) provide an algorithm for a piecewise solution for dynamic models with occasionally binding constraints. Piecewise, since it links impulse responses when the ZLB holds to the no-ZLB case when it is slack.

As the authors show, occasionally binding constraints can be seen as two regimes of the same model. In one regime, the ZLB is slack, while it binds in the second. The algorithm produces two sets of derivatives used for impulse response calculation and links them. The solution method links the first-order approximation of the model around the steady-state for each regime (Iacoviello and Guerrieri, 2015).

In the first regime, which I refer to as R1, conditions for the existence of a rational expectations solution (Blanchard-Kahn conditions) have to hold, and the interest rate evolves according to the Taylor rule so that

$$i_t = i_{t-1}^{rR} (\pi_{t-1}^{1+r\pi} \frac{Y_{t-1}^{rY}}{\bar{Y}} \bar{r}\bar{r})^{1-rR}, \quad (85)$$

holds as specified above. If a shock hits the economy so that the Taylor rule would suggest to set the interest rate below zero, the algorithm switches to the ZLB regime. Thus, a shock causing a reduction of the interest rate of more than its steady state level, would cause the result in a regime-switch. This happens, when $\hat{i}_t \leq -(1/\beta - 1)$, which is equivalent to a reduction of $1/\beta - 1 = 0.007$ in my model, equivalent to an annualized interest rate reduction of 3 percentage points, since the model is calibrated for quarterly frequencies. The devaluation shock described, is large enough in size so that the interest rate reduction after the shock touches the ZLB.

In the ZLB regime, the interest rate is fixed to 0 ($i = 1$ for the gross interest rate) given by

$$i_t = 1 \forall t \in [t, t_{max}], \quad (86)$$

in which t_{max} ensures that the ZLB regime converges back to the baseline regime (R1) and i is the gross interest rate. In every post shock period, the switching condition is checked. When output is below trend and inflation remains lower than steady state, the Taylor rule would suggest to reduce the interest rate. As long as such a reduction would cause the interest rate to fall by more than its steady level, the model remains at the fixed interest rate regime. Once output has recovered and the interest rate reduction would lead to interest rates above the ZLB, the regime

switches back. Therefore, the algorithm is a guess-and-verify approach with the assumption of i , being a proxy for the ZLB.

The solution algorithm suggested by Iacoviello and Guerrieri (2015) is as follows

- Algorithm 2 (The Iacoviello and Guerrieri (2015) algorithm)**
1. *Specify two separate models in Dynare with identical parameters, endogenous and exogenous variables.*
 2. *In the baseline model, the policy instrument evolves according to a Taylor rule. In the alternative regime, it is zero.*
 3. *Pick the number of periods, t_{max} under which the ZLB is assumed to bind.*
 4. *Include a switching condition to indicate that the model needs to move from baseline to the ZLB case and find a shock that reaches this threshold.*
 5. *Solve the model for baseline (R1) and calculate first order conditions and simulate a devaluation shock.*
 6. *Solve the model with the binding ZLB and recalculate impulse responses for devaluation shocks.*
 7. *Link impulse responses of ZLB case to non-ZLB at t_{max} .*
 8. *Compare both impulse responses.*

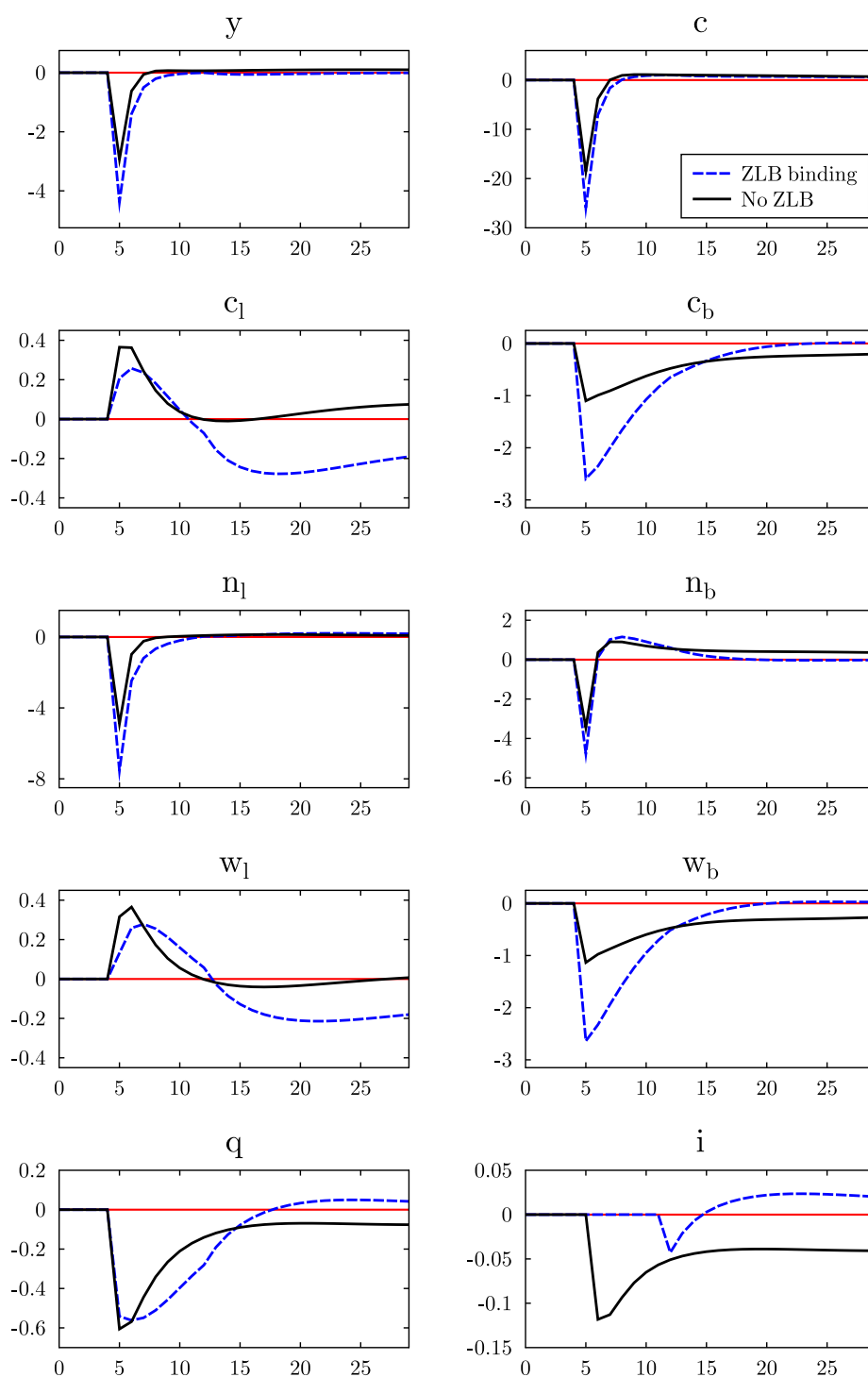
Solutions are time varying decision rule in a guess and verify approach so that each period initial conditions are verified, and the regime is updated if necessary.

5.5.4 Valuation shocks at the ZLB

This section presents the results of the devaluation shock at the ZLB and compares the two debt stabilization policies. Figure 38 and the subsequent one depict impulse responses for the fiscal policy regime when adjusting taxes (TI). Impulse responses are given as percentage deviations from steady state.

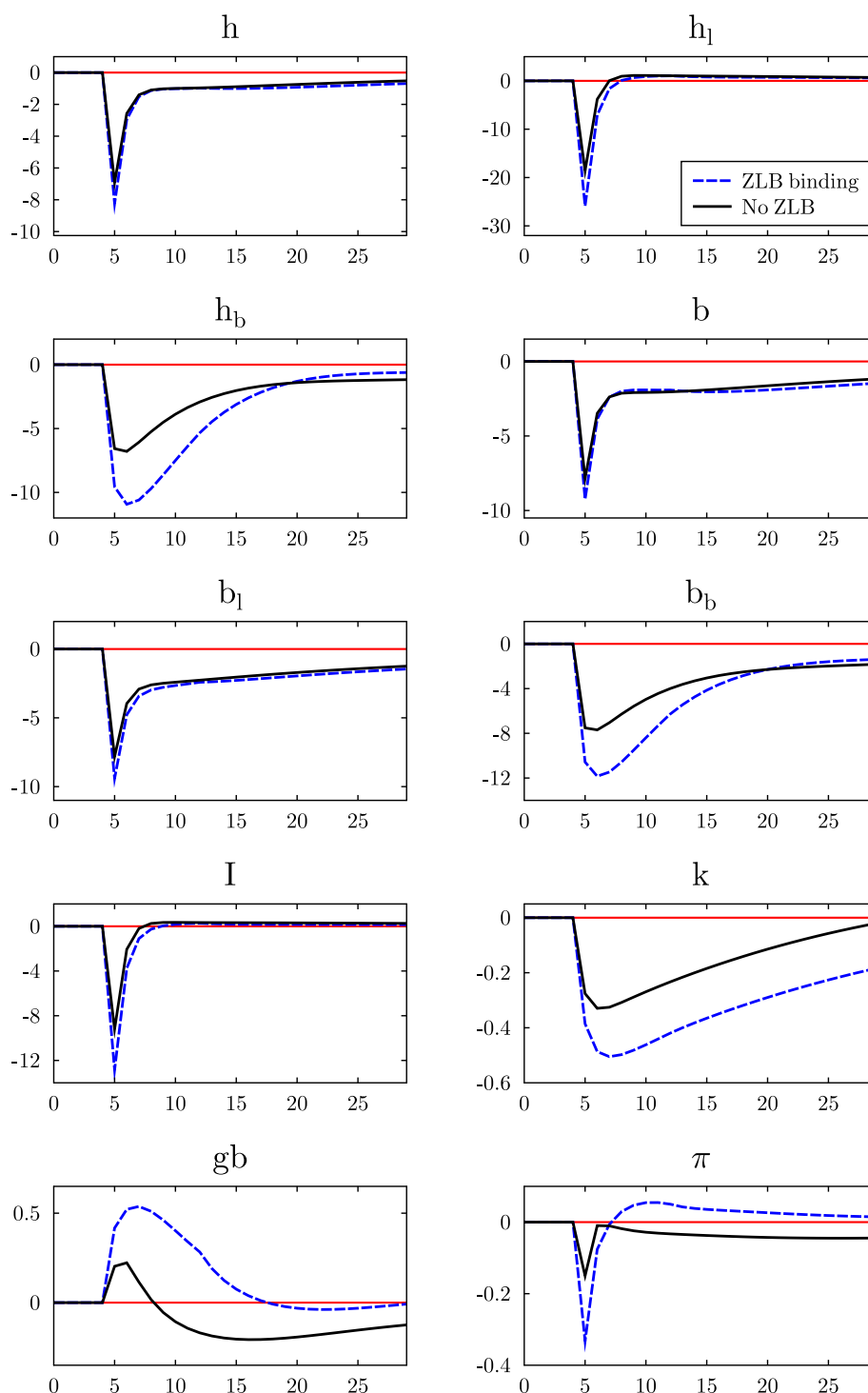
The figures denote the responses of the non-ZLB case, R1 (black line), and the ZLB case (blue dashed line). Impulse responses at the ZLB are linked to the non-ZLB case at t_{max} . As is visible, the ZLB acts as an amplifier of the devaluation shock compared to when the ZLB is slack.

The reaction of the interest in the non-ZLB case, shows a reduction of roughly 1% per quarter, larger its steady state level. Hence, the model switches to the ZLB regime after a devaluation shock so that \hat{i} is zero, denoting a proxy for the ZLB.

Figure 38: IR tax increases after devaluation shocks, ZLB vs. no-ZLB I


Notes: Simulated impulse responses after devaluation shock. The solid black line denotes the simulated responses if the ZLB is slack, whereas the dashed blue line is the impulse responses when the ZLB holds. Impulse responses result from the piecewise solution algorithm suggested by Iacoviello and Guerrieri (2015). The fiscal authority increases consumption and labor taxes to stabilize debt (TI), and spending is adjusted according to an AR(1) process. First set of variables, percentage deviations from steady-state.

Figure 39: IR tax increases after devaluation shocks, ZLB vs. no-ZLB II



Notes: Simulated impulse responses after devaluation shock. The solid black line denotes the simulated responses if the ZLB is slack, whereas the dashed blue line is the impulse responses when the ZLB holds. Impulse responses result from the piecewise solution algorithm suggested by Iacoviello and Guerrieri (2015). The fiscal authority uses consumption and labor taxes to stabilize debt, and spending is adjusted according to an AR(1) process. Second set of variables, percentage deviations from steady-state.

The model switches back in period 13, even though output is below trend since the reduction of the interest rate according to the Taylor rule would not be larger than 3%. Then, impulse responses are linked.

The negative impact of forced deleveraging has a two times stronger effect on output compared to the no-ZLB case. Once the regime switches back to the baseline regime R1, the impulse responses from the ZLB case are linked to the ones in the non-ZLB case, visible by a kink in the impulse responses in period 13. Thus, the ZLB holds for eight periods, and the regime returns to steady state thereafter.

Forced deleveraging leads to lower private bond holdings. Real estate then declines by 10% and 8% for entrepreneurs and impatient households, respectively. House prices decline in a similar way compared to the baseline estimation and patient households respond by building up their real estate position. As entrepreneurs are hit through their production function, the drop in investment is amplified when the ZLB holds. Together with lower consumption, output declines by 5%.

Since the monetary authority is incapable of lowering interest rates until t_{max} , falling prices cause the real debt burden to increase by more compared to the situation when the ZLB is slack, which amplifies the devaluation shock.

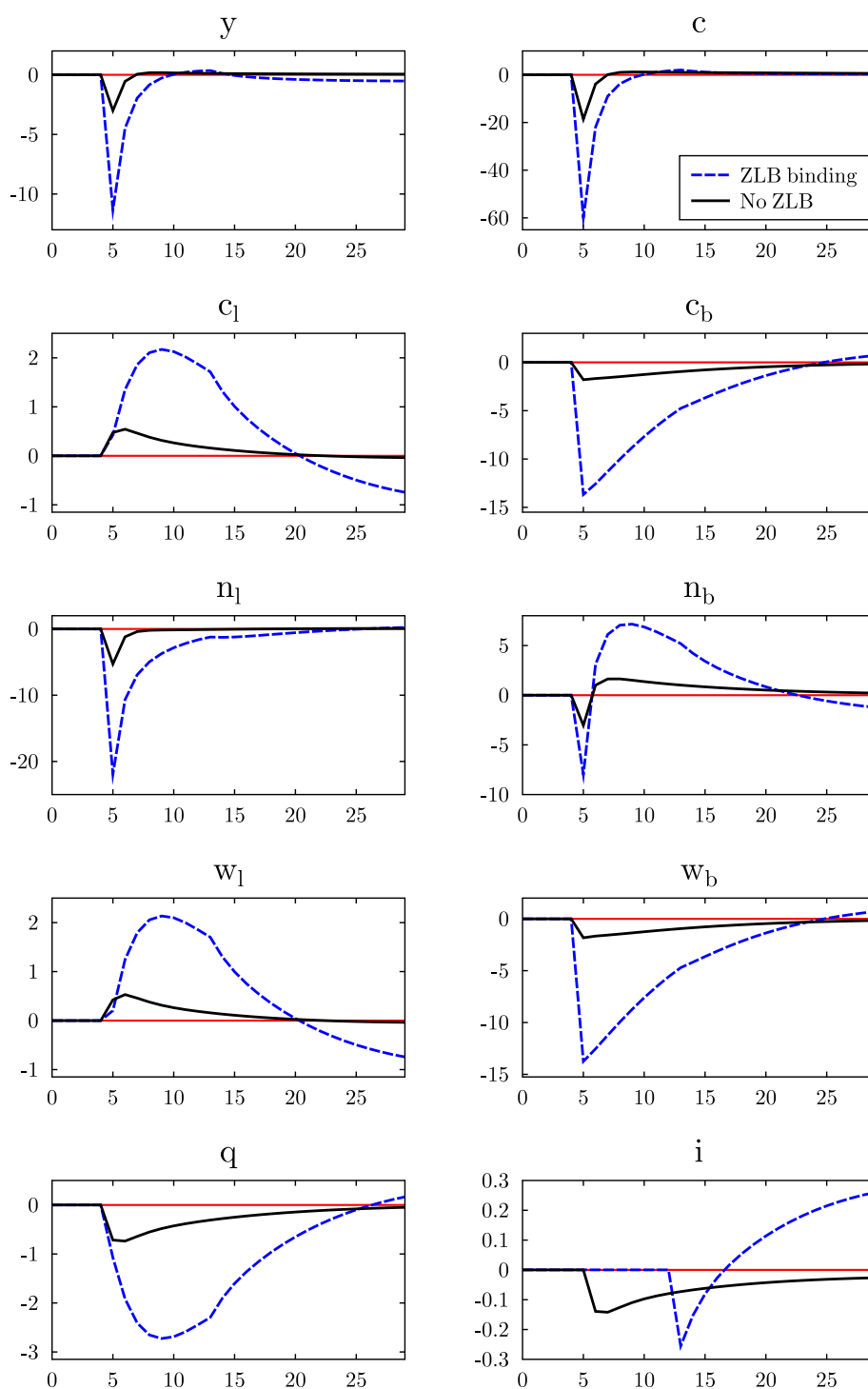
Figure 40 depicts simulated impulse responses for devaluation shocks if spending is cut (SC) at the ZLB after devaluation shocks as deviations from steady-state. Additionally, Figure 42 in the appendix compares both fiscal rules at the ZLB.

A spending cut causes the recession following devaluation shocks to be more severe at the ZLB holds compared to when the ZLB does not hold. Borrowers' and entrepreneurs' real estate position declines by 30% and 15%, which exceeds the impact when interest rates can be adjusted by the central bank. In turn, the real estate price decrease is stronger, and patient households can take more advantage of lower house prices and increase their position by almost 8%. As in the case of a slack ZLB, reduced wealth due to a lower value of real estate causes borrowing households to reduce consumption while the patient household increases it. Output declines by more than 12% due to the relative sizes of respective households.⁴⁸

Figure 42 in the appendix compares spending cuts and tax increases at the ZLB following house price shocks. Using spending cuts to stabilize debt reduces output by more compared to a non-binding ZLB, since the monetary authority is incapable of lowering interest rates. Therefore, if a devaluation shock occurs when debt-over-output deviates from steady-state, tax increases to stabilize debt, rather than government spending reduction leads to a milder impact on output.

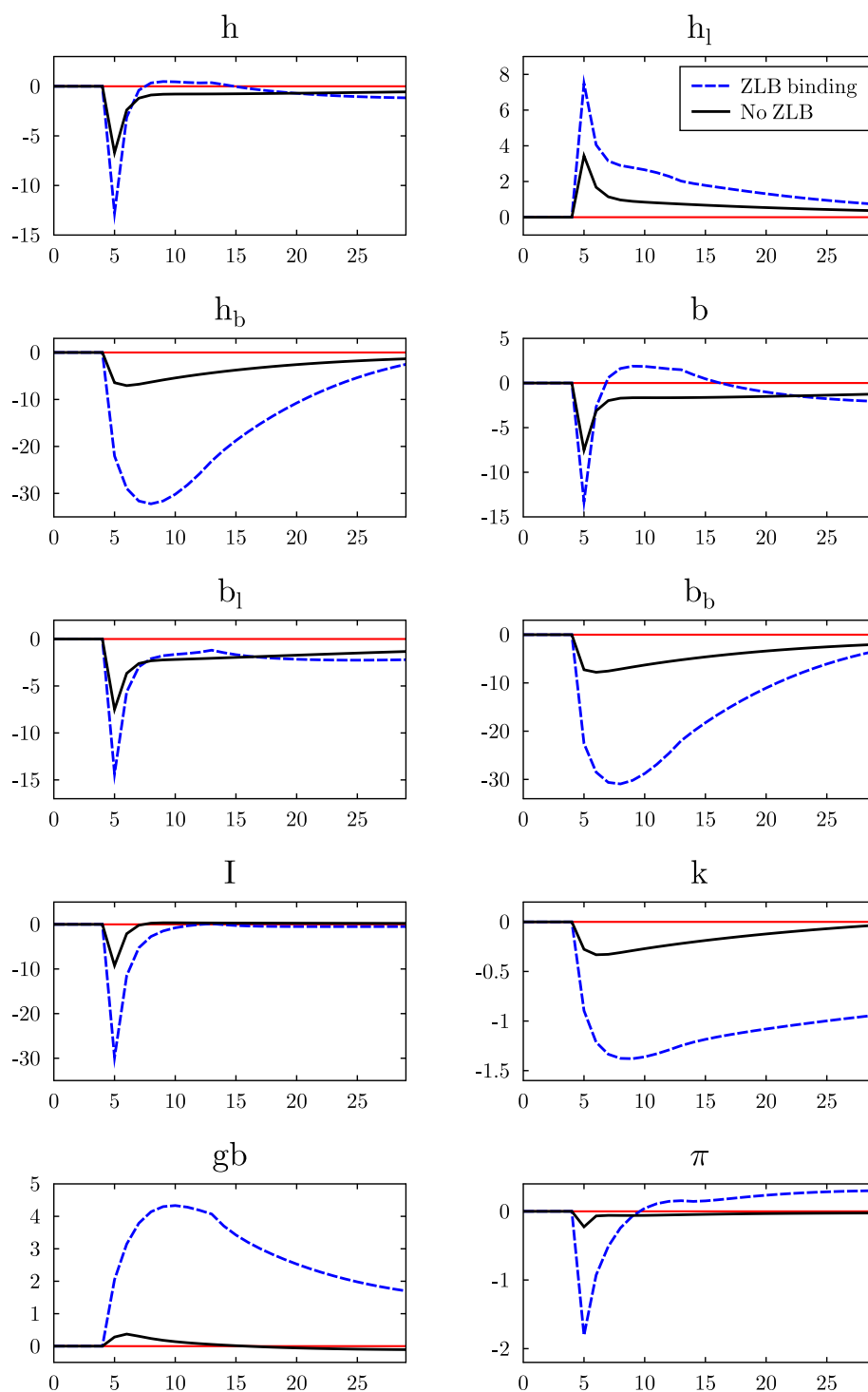
⁴⁸Assuming that each household is represented by a third of the total number of households.

Figure 40: IR spending cuts after devaluation shocks, ZLB vs. no-ZLB I



Notes: Simulated impulse responses for devaluation shock. The solid black line denotes the simulated responses if the ZLB is slack, whereas the dashed blue line is the impulse responses when the ZLB holds. Impulse responses result from the piecewise solution algorithm suggested by Iacoviello and Guerrieri (2015). The fiscal authority spending cuts (SC) to stabilize government debt and consumption and labor taxes are adjusted according to an AR(1) process. First set of variables, percentage deviations from steady-state.

Figure 41: IR spending cuts after devaluation shocks, ZLB vs.no-ZLB II



Notes: Simulated impulse responses for devaluation shock. The solid line black denotes the simulated responses if the ZLB is slack, whereas the dashed blue line is the impulse responses when the ZLB holds. Impulse responses result from the piecewise solution algorithm suggested by Iacoviello and Guerrieri (2015). The fiscal authority uses spending cuts (SC) to stabilize government debt and consumption and labor taxes are adjusted according to an AR(1) process. Second set of variables, deviations from steady-state.

Modifying the model to include housing adjustment costs does not alter this result, as such costs were estimated to be very low Iacoviello (2005).

5.5.5 Discussion and limitations

Devaluation shocks cause a recession in the model economy since some agents are forced to deleverage in line with the debt-driven slump described by Eggertsson and Krugman (2012). In a time when government debt is high, the way the fiscal authority responds to stabilize debt by adjusting fiscal policy is irrelevant for the impact on output only if the ZLB is slack. In contrast, if the ZLB binds, increasing taxes to stabilize debt-over-GDP to return to steady-state causes a less severe recession compared to cutting government spending.

However, the nature of devaluation shocks can be questionable, since its source is not clear cut, as Eggertsson and Krugman (2012) also state. There may be an underlying mechanism for which the devaluation shock is only a proxy.

Further, assuming risk-free borrowing for private and government debt abstracts from the current issues of public debt defaults, or private borrowers filing bankruptcy. In a more sophisticated model, this could be taken into account, and one could investigate if the results still hold.

Concerning the solution algorithm, Iacoviello and Guerrieri (2015) note that it does not account for precautionary behavior due to expected switches from one to the other regime. Households are assumed to either believe R1 will endure forever, or R2 holds so that a regime change occurs unexpectedly to households. Once a transition from the reference to the alternative regime has taken place and back to the reference regime, households do not expect this to happen again. Precautionary savings due to the possibility of a binding zero lower bound in the future do therefore not take place since the shocks are not realized yet. Further, the assumption of a number of periods in which the constraint holds with equality is ad-hoc.

As any linear approximation to the model, the model's solution ignores higher order shocks. Households' optimal allocation, therefore, does not account for uncertainty on the realization of shocks.

5.6 Conclusion

Households receive utility from owning real estate that co-moves with consumption and is an asset that can be used as collateral to borrow additional funds. However, it is also risky.

In this paper, I investigate the impact of devaluation shocks to households' ability to borrow, which have a strong negative effect on consumption, real estate prices and induce a recession in the model economy since such shocks force constrained households to deleverage.

The direction of the impact of devaluation shocks on different types households is asymmetric: They reduce the ability to borrow, which has a severe impact on borrowing constrained households since they are forced to deleverage by selling real estate that in turn reduces their wealth and thus consumption. On the other hand, patient households can take advantage of lower house prices and build up their real estate position and increase consumption because they feel richer.

Devaluation shocks played a significant role at the end of 2007 when house prices collapsed in the U.S. leading to a severe recession while government debt surged. The fiscal authority has taxes and fiscal spending as an instrument lower public debt. As it turns out, in times, when the economy is sufficiently far away from the Zero Lower Bound on interest rates, the choice of whether cutting spending or increasing spending does not have an impact on the severeness of the recession after devaluation shocks.

I then investigate whether this results holds at the ZLB and find that the ZLB acts as an additional amplifier for devaluation shocks. When government debt is high, and a devaluation shock hits, cutting spending to reduce government debt reduces output by more compared to increasing taxes. If taxes are increased with interest rates at the ZLB, it causes output to decrease by 4.5% that compares to 3% if the ZLB is slack. In contrast, if the fiscal authority cuts spending the impact of devaluation shocks on output is four to five times as strong compared to the no-ZBL case.

This paper adds to the current literature on devaluation shocks and amends the Iacoviello (2005) housing sector model with a fiscal sector and the possibility of different fiscal rules. It then shows that the choice of fiscal instruments to stabilize debt is irrelevant only if the economy is away from the ZLB.

5.7 Appendix

Table 11: Data sources for VAR estimation

Variable	description	source
ffr	Federal Funds rate, end of quarter	FRED data: FEDFUNDS
π	4 quarter log differences of CPI	FRED data: CPIAUCSL
C-S Index	log of 10 city Case-Shiller Index	FRED data: SPCS10RSA
y	log of real per per capita GDP	FRED data: GDPC1/POP

Notes: This Table gives data sources for the VAR estimation using FRED data and available at research.stlouisfed.org

Table 12: Complete sources for parameters calibration

Param	Description	Manipulation	Source
α	wage share patient household	average LS 1990-2014	FRED: LABSHIPUSA156NRUG
β_l	discount factor patient	average federal funds rate 1980q1-2014q4	FRED: FEDFUNDS
β_b	discount factor impatient		Iacoviello (2005)
γ	discount factor entrepreneurs		Iacoviello (2005)
δ	depreciation rate	annual depreciation of 10%	Depreciation rate as in Mertens and Ravn (2011)
ν	elasticity of output to real estate		Iacoviello (2005)
μ	capital share in production	prices constant on average for one year	for instance Iacoviello (2005)
calvo	calvo price rigidity	single-Family mortgages average in 2013	- Iacoviello (2005)
m	loan-to-value entrepreneur		Federal Housing Finance Agency, fhfa.gov
m_b	loan-to-value borrower		Iacoviello (2005)
η	labor disutility		Iacoviello (2005)
rR	weight of interest rate reaction in TR	estimate from TR, 1980q1 to 2014q4	FRED: FEDFUNDS, CPIAUCNS, GDPC1
$R\pi$	weight of inflation rate reaction in TR	estimate from TR, 1980q1 to 2014q4	FRED: FEDFUNDS, CPIAUCNS, GDPC1
Ry	weight of output gap reaction in TR	estimate from TR, 1980q1 to 2014q4	FRED: FEDFUNDS, CPIAUCNS, GDPC1
ρ^A	persistence of technology shock	regression on TFP index	Basu et al. (2006)
ψ	capital adjustment costs weight		Iacoviello (2005)
j	housing preference parameter patient HH	fix residential housing over GDP at 140%	Financial Accounts the United State, 2015 table B1.101, B.103, B.104
ϕ^K	capital adjustment costs.		Iacoviello (2005)
ρ^θ	autoregressive parameter in LOM quality shock		Iacoviello (2015)
g/Y	annual gov't spending to GDP	government consumption and investment over GDP	FRED: GCEC96/GDPC1
b/Y	debt-to-GDP	federal debt held by the public over GDP	FRED: FYGFDUN/GDPC1
τ^n	0.28	steady state tax rate	Estimated using FRED data (see Table 13)
τ^c	0.05	steady state tax rate	Estimated using FRED data (see Table 13)

Tax rate calculation. For the calculation of the tax rates, I follow Fernández-Villaverde et al. (2011). Data can be found in NIPA tables and is obtained from FRED. The average consumption tax is

$$\tau_c = \frac{TPI - PRT}{PCE - (TPI - PRT)} \quad (87)$$

In which TPI is taxes on production and imports, PRT denotes local property taxes and PCE are personal consumption expenditures. Additionally, average personal income taxes τ_p and labor taxes τ_n are calculated as follows

$$\tau_p = \frac{PIT}{WSA + PRI/2 + CI} \quad (88)$$

$$\tau_n = \frac{\tau_p(WSA + PRI/2) + CSI}{CEM + PRI/2}. \quad (89)$$

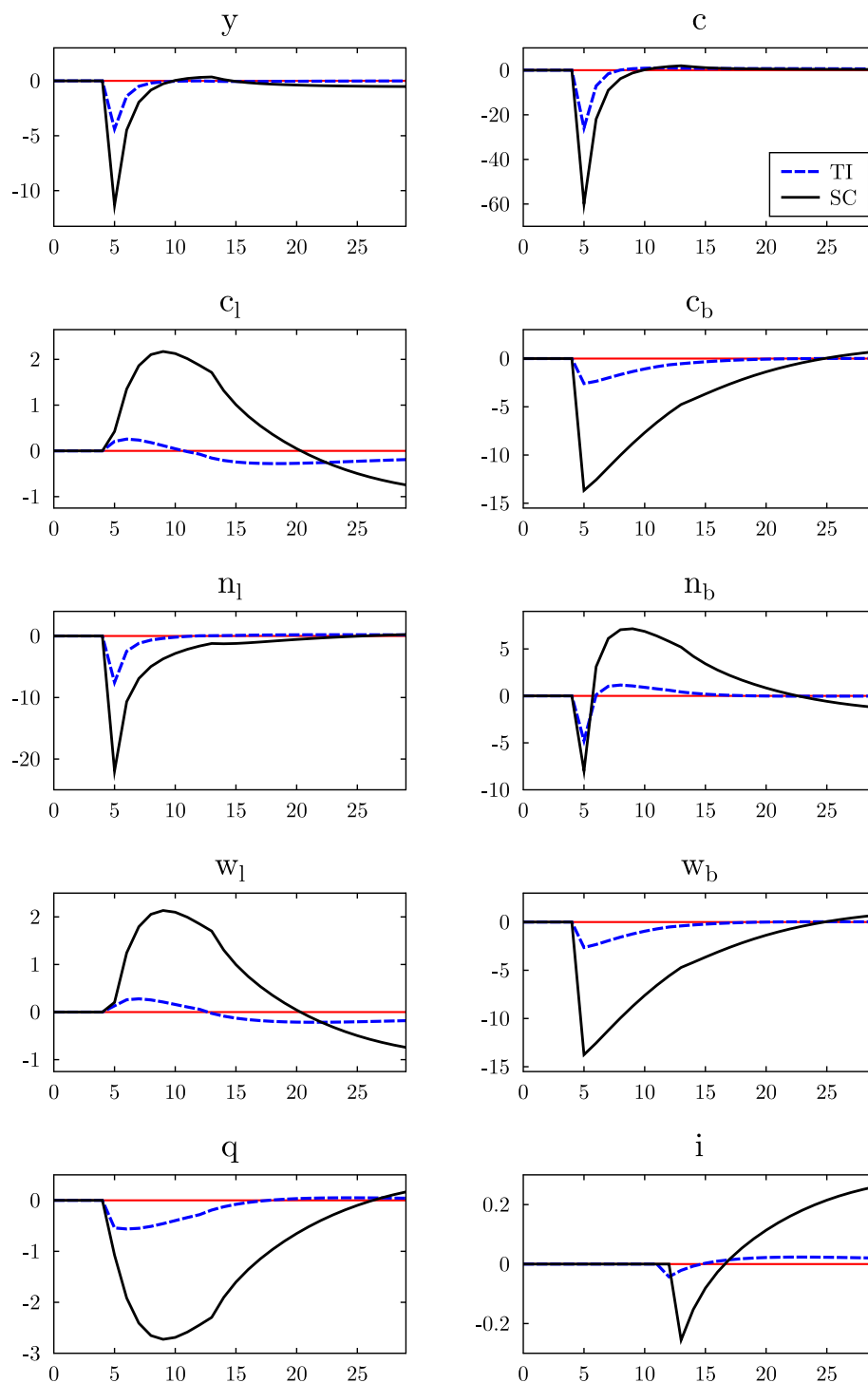
definitions and sources can be taken from Table 13, which also gives FRED data sources.

Table 13: Details on the estimation of average taxes

	Description	FRED Source	
PIT	Fed+local+state taxes on PI	A074RC1A027NBEA B245RC1Q027SBEA	+
WSA	wage and salary accruals	WASCUR	
PRI	proprietor's income	A043RC1Q027SBEA B179RC1A027NBEA	+
CI	capital income = PRI/2+RI+CP+NIS		
RI	rental income	B049RC1A027NBEA	
CP	corporate profits	A551RC1A027NBEA A054RC0A144NBEA	+
NI	interest income	W255RC1A027NBEA	
CSI	contributions social sec	W780RU1Q027NBEA	
CEM	compensation of employees	W209RC1A027NBEA	
TPI	taxes on prod and imports	B234RC1A027NBEA B248RC1Q027SBEA	+
PCE	personal cons expenditures	PCE	
PRT	property tax revenue	S210400	
g	government spending	GCECA	

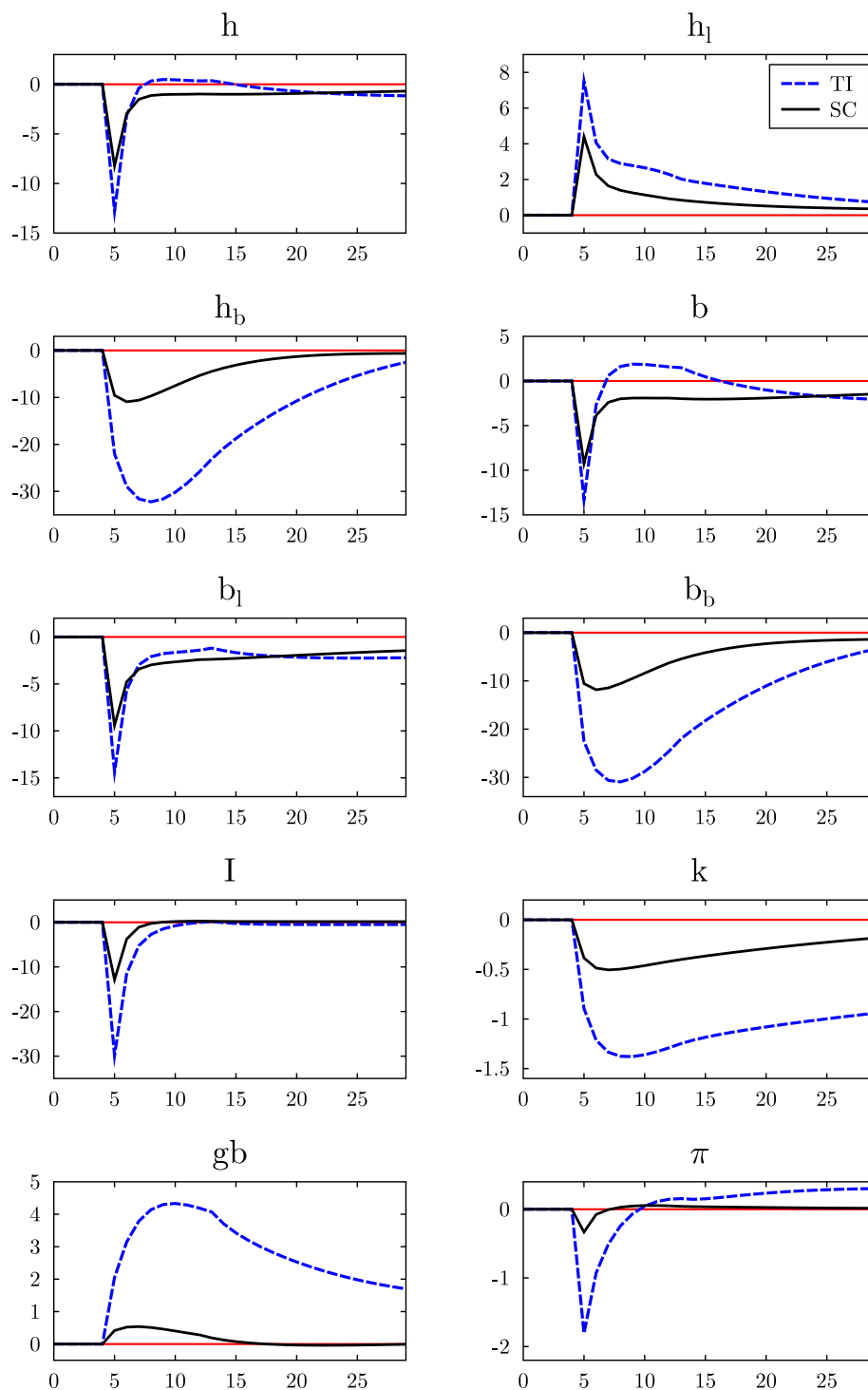
Notes: This table gives data definitions and sources for the estimation of fiscal feedback rules.

Figure 42: Spending vs. tax stabilization of debt after valuation shocks



Notes: This figure denotes the simulated impulse responses from a one standard deviation shock to the constrained households ability to borrow (devaluation) shock at the ZLB and compares both fiscal rules: spending reduction (SC), denoted by the black line and tax increases (TI) denoted by the dashed blue line. First set of responses, percentage deviations from steady-state.

Figure 43: Spending vs. tax stabilization of debt after valuation shocks



Notes: This figure denotes the simulated impulse responses from a one standard deviation shock to the constrained households ability to borrow (devaluation) shock at the ZLB and compares both fiscal rules: spending reduction (SC), denoted by the black line and tax increases (TI) denoted by the dashed blue line. Second set of responses, percentage deviations from steady-state.

6 Market inefficiencies and forecastability of spot rates in the shipping sector

Coauthor: Marcus Eppinger

6.1 Introduction

Market participants can use forward rates to fix today the delivery of an asset, commodity or service in the future. The price of forwards depends, unlike the spot (current) price, on expectations on future events until maturity. The Unbiased Estimator Hypothesis (UEH) connects these forwards and spot rates by a no-arbitrage relationship. The UEH states that the forward rate is a conditionally unbiased predictor for future spot rates (Engel, 1996; Hodrick, 2014). If the UEH holds, the best estimate for the spot rate at maturity is the price suggested by the forward contract today. As maturity of the forward approaches, its price converges towards the spot price. If this is the case, markets are fully efficient and beating the market is impossible.

However, this is not the case in the market for cargo shipping. The statistical properties of cargo spot shipping rates and its derivatives, like forwards, are unique compared to other asset classes. The shipping asset market is dominated by charterers and shipowners who are mainly interested in smooth cash flows rather than in speculation. In the absence of speculative noise trading, which would provide additional liquidity, trading can be thin so that prices stay constant for days. Since the underlying (base value) of shipping rates is a non-storable service, not an asset or commodity, rates cannot be short-sold in a secondary market which reduces liquidity as well and stands in contrast with currency, commodity or other paper markets.

In this paper, we show that these particular properties lead to inefficient markets so that forward prices, do not converge towards spot rates at maturity. We provide evidence that the UEH does not hold, by performing cointegration tests of spot rates and forwards for the two largest types of vessels using which has not been shown before. The cointegration relationship between the two rates is only weak so that the portfolio composition of short and long-run contracts matters for charterers and ship owners, unlike inefficient markets.

We empirically take advantage of inefficient markets and find that Forward Freight Agreements (FFA) have a significant impact on spot prices during the following days. The shipping market, therefore, does not process information contained in forwards quickly enough so that profitable forecasts for the next days can be generated.

We let ARIMA, VAR, and VECM models compete when creating forecasts, and compare the models' forecast errors (RMSE) to those errors resulting from estimating random walks. The model with the lowest RMSEs is a vector autoregression, which is then used to create a trading scheme, capable of outperforming a benchmark index, even after controlling for transaction costs.

In the remainder of the paper, we proceed as follows. We review recent literature on market efficiency in the shipping sector and the sign of risk premia for holding contracts with long or short maturity in the next chapter. In chapter 6.3 we introduce the data and estimation methods, which compete for creating forecasts. Chapter 6.4 compares spot rate forecasts in terms of RMSE. The last section presents a trading scheme based on VAR estimation, which outperforms the benchmark index, and a final chapter concludes.

6.2 Market properties in the shipping sector

The shipping sector experiences high entry costs, time to build and convex operating costs of ships with high idiosyncratic risk and high market entry barriers (Kalouptsi, 2014). The process of cargo shipping is a service rather than a storable product. Unlike in asset trading, where the underlying value is a physical commodity or share, short positions of shipping goods cannot be taken. A cost-of-carry (COC) relationship which states that the price of a derivative, like a forward, at $t - n$ for delivery at t equals the price of the underlying at $t - n$ plus costs associated with purchasing, storing and holding the asset up to t , does not hold (Kavussanos et al., 2005).⁴⁹ Therefore, only supply and demand influence both rates so that the timing structure of investors' portfolios matters and speculative trading then does not avoid arbitrage trades.

Demand for shipping services and hence current spot rates are driven by aggregate seaborne trade and thus subject to world business cycle fluctuations and seasonal effects (Stopford, 1997; Kavussanos and Alizadeh-M., 2002; Kalouptsi, 2014). Supply of shipping services in the short-run, on the other hand, is inflexible and determined by the number of voyages carried out by shipowners which introduces unique characteristics of derivatives on cargo rates (Kalouptsi, 2014).

⁴⁹So that the relationship $E_t(F_{t,t-n}) = S_{t-n}(1 + C)$ does not hold.

These rigidities cause shipping rates to be highly volatile so that the price of *renting* one of the largest types of vessels decreased from 233.000 USD/day on June 5, 2008, to 2316 USD/day on December 2, 2008. *Renting* refers the market price to ship a previously defined quantity of goods on a particular route at a certain point in time and can be negotiated freely, typically involving three parties; a carrier accepts to ship the goods and receives the spot rate from the charterer. Then, brokers match supply and demand for services. Two types of such transactions are common. First, charterer and shipowners can fix a voyage charter for a certain quantity and route. The shipowner is then responsible for the port, fuel, crew, brokerage and canal costs. In contrast, Time Charter (TC) contracts involve the reception a ship for a certain time, while the carrier sends the charterers' goods, and Voyage Costs are borne by the charterer and rates for both obviously differ.

These market participants are interested in estimating future shipping prices to minimize cash flow volatility and to hedge their portfolios (Kavussanos and Visvikis, 2004). They can either engage in forward and future trading⁵⁰ or, wait until maturity and purchase the respective spot price. Forward Freight Agreements (FFAs) replaced futures since trading was too thin and market prices could not be determined. FFAs, which are principal-to-principal agreements traded over-the-counter, usually between shipowners and charterers are based on difference payment between realized spot rates at maturity and rates initially negotiated. In contrast to options trading, honoring FFA contracts is mandatory. Reliable time series data on FFAs as an instrument to insure against price shifts is only available from 2004 onwards so that we take this as a starting point for our estimations.

The decision whether to buy FFA or spot contracts depends on the efficiency of the market. A market is called efficient if all information is processed instantly so that forward rates and spot prices co-move, and profitable forecasts cannot be generated, given by the following

$$\Delta S_{t+1} = \alpha + \beta(F_t - S_t) + \epsilon_{t+1}, \quad (90)$$

in which ΔS_{t+1} is the change in the spot rate from t to $t+1$, F_t is the forward rate at t and S_t is the spot rate. If markets are efficient, the UEH holds so that $\beta = 1$ and $\alpha = 0$ (Hull, 2006). However, the hypothesis tends to be rejected, in, for instance, the currency market as Frankel and Poonawala (2010) note. The UEH, which states that the forward price is equal to its discounted expected future spot price, does not need to hold here, as Kavussanos and Visvikis (2004) note. FFA prices then do not

⁵⁰Forwards are traded over-the-counter, whereas futures are standardized contracts that are exchange-traded and party-to-party negotiation does not take place.

converge to their cost of carry price as maturity approaches. Short-run forecasts for spot rates are therefore possible.

6.2.1 Forward freight rates & spot rates

Time series of freight rates exhibit a high degree of autocorrelation, which can be exploited to create forecasts. Cullinane (1992) shows that short-run forecasts for spot rates can be constructed using futures. The author uses an ARIMA model with daily data from 1985 to 1988 to generate estimates for the Baltic Freight Index (BFI).

Cullinane et al. (1999) confirm the robustness of the earlier results with post-1993 data after crafts were added to the index. They conclude that the shift in the index has not caused the quantitative relationships to break down.

Kavussanos and Nomikos (2003), in contrast, also control for long- and short-run dependencies and produce forecasts for spot rates. In a 1988-1998 sample, the authors demonstrate futures and spot prices to follow a common trend using vector error correction models (VECMs). They compare forecasts' quality in terms of Root Mean Square Error of VECMs and ARIMA models to a random walk. They then find that futures Granger cause spot rates, but the vice versa relationship is weak.

Batchelor et al. (2007) explain this evidence with asymmetrical transaction costs. Such costs in spot markets are typically higher than in future markets because in the former, transaction costs contain the physical good of shipment at that time, unlike the forward market that is based on difference payment. Lower transaction costs in the latter suggest that FFA prices process new information more quickly than spot markets, and hence have a price finding role.

Kavussanos and Visvikis (2004) test the UEH with VECM models. The authors demonstrate a cointegration relationship between futures and spot rates for maturities up to two months, which suggests futures at t to be unbiased estimators for spot rates at maturity with a cointegration vector of $(1, -1)$. Thus, the difference between the spot and future rate is zero in equilibrium. A deviation in the short-run from the cointegration relationship may occur, since Future rates contain more information than spot prices.

Batchelor et al. (2007) compare VARs and ARIMA models for their short-run forecast performance. VECMs have the lowest RMSE among the models, especially for long-run forecast, with only half the RMSE of the second best model.

However, it remains unsolved which of the two rates primarily causes a convergence back to the common trend, which we check in the subsequent chapters.

6.2.2 The timing structure of freight rates

In the market for cargo shipping, the UEH does not need to hold due to the absence of a COC relation, so that the composition of freight rate portfolio has an impact on expected profits from holding contracts with different maturities.

Expectations on future freight rates take the term structure of shipping rates into account. Adland and Cullinane (2005) define freight spreads, or premia, as the difference between the price for long-run contracts and today's spot rates.

According to Stopford (1997), today's supply and demand for shipping goods primarily have an impact on current spot rate, while forward rates reflect long-run expectations on future freight rates, interest rates, and risk premia. This has been called the Expectations Hypothesis of the Term Structure (EHTS) and can be broken down into two versions, according to, for instance, Glen et al. (1981) and Strandenes (1984).

The classic EHTS states that risk premia for holding portfolios containing different maturities are time-invariant but different from zero. A violation of the hypothesis could be due to market inefficiencies, irrational expectations of agents or incorrect underlying models (Kavussanos and Alizadeh-M., 2002). The existence of a risk premium is crucial when estimating forecasts and investigating market efficiency.

Future spot rates tend to increase step by step if market participants expect events which cause the price to ship goods at a future point in time with payment in that moment to increase. Expectations influencing different maturities might be asymmetric, and risk premia can be time-varying (Adland and Cullinane, 2005). Economic expansions mostly cause short-run and long-run rates to rise simultaneously, consistent with the classic EHTS. Recessions, however, mainly cause high maturity rates to decrease, lowering the correlation with short maturity rates so that the classic EHTS cannot be observed in the data and risk premia can be negative in downturns. Adland and Cullinane (2005) thus suggest to reject the EHTS, postulating time-varying risk premia.

The Pure Expectations Hypothesis of the Term Structure (PEHTS), on the other hand, states that the price of long-run contracts incorporates all alternative investments, as well. Risk premia for freight rates are then constant and zero.

The validity of these two hypotheses depends on the choice of the market, the model applied as well on as the sample. Glen et al. (1981) and Strandenes (1984) show that the PEHTS holds, but their results are rejected by Hale and Vanags (1989) and Veenstra (1999) for the Dry-Bulk markets. The latter study suggests that shipowners demand a liquidity premium for long-run freight rates due to the

non-substitutability of specific routes and the inability to sell charter contracts in a secondary market (Veenstra, 1999). Such a premium violates the PEHTS, as also demonstrated for the oil tanker market by Wright (2000) and Adland and Cullinane (2005).

However, most tests for the PEHTS in Wright (2000) check for short-run relationships, but are incapable of testing the hypothesis for long-run contracts (Kavussanos and Alizadeh-M., 2002).

Kavussanos and Alizadeh-M. (2002) demonstrate a long run relationship between spot and Time Charter rates of approximately one to three years. A test for PEHTS and classical EHTS is rejected at the 5% level.

6.2.3 The sign of risk premia

If the timing structure of cargo rates matters and forecasts can be generated, the sign of the risk premium is important since market participants would realize profits for holding the respective contract.

Time series evidence for the sign of the risk premium is mixed, depending on whether the PEHTS holds so that the premium $\mu_t = S_t - E_t\beta F_T$ can either be positive, negative or zero.

Volatile voyage costs and the risk of being forced to take ships to different starting ports in the future create negative risk premia. Technology shocks or stricter safety laws lower the value of vessels, which are assigned to specific routes in long-run contracts, also postulating a negative risk premium which is shown by Kavussanos and Alizadeh-M. (2002) in EGARCH models controlling for cycles in a 1988 to 1997 sample.

Adland and Cullinane (2005) demonstrate spot rate volatility to be higher than forward rates' volatility so that risk-averse ship owners may be willing to sell long-run contracts with a discount to reduce portfolio variance which causes risk premia to be negative.

In contrast, long-run contracts can be less risky than short-run contracts since they keep income and costs for shipowner and charterers constant and predictable. Such security postulates positive risk premia according to Tashman (1996). Moreover, securing shipment for some routes can be difficult for charterers in expansions when excess demand exist. Charterers then would be willing to pay a positive risk premium to secure their shipments. Finally, there is an exit risk of the charterer. The higher the maturity of the contract negotiated, the higher the exit risk, leading to an increase with t in the risk premium of the shipowner.

In conclusion, risk premia can be time-varying with evidence for positive or negative estimates. In the subsequent sections, we investigate the sign of the risk premium by performing cointegration tests.

6.3 Data and estimation

This section describes our data and estimation models that compete in terms of forecastability of future spot rates, based on FFA and Time Charter rates.

6.3.1 Data

Our sample contains daily data from 1 October 2004 to 9 December 2008, the maximum of continuously available data. We exclude the crisis data because we are interested in within-cycle dynamics, as shipping cycles reoccur approximately every six to nine years (Kalouptsi, 2014).

FFAs reflect daily closing values from Baltic Exchange and are based on individual routes or on route baskets. Spot and Time Charter rates are based on Clarksons data⁵¹. Spot rates are daily-recorded, whereas Time Charter rates are published weekly only.

Capesize and Panamax ships have the highest trading volume and most contracts are sold on the 4TC route⁵². Panamax (Capesize) routes account for 50% (35%) of total FFA volume. Panamax ships are those, still capable to transversing the Panama canal, whereas Capesize ships are too large to do so.

FFA contracts for the Capesize Index (BCI) or Panamax Index (BPI) are based on weighted averages of the preceding seven business days each respective month, to avoid a thin trading bias (Dimson and Marsh, 1983).

Because FFAs on baskets of routes (for instance 4TC) are used to hedge monthly income, these are calculated as averages of the month to make sure monthly costs and revenues can be hedged.

FFA+1Q, FFA+2Q and FFA+1A are FFA contracts, maturing at the end of the next quarter, two quarters and at the end of next year. We create first log differences of FFAs to avoid jumps resulting from rollovers from one to the next day, as the last day before maturity of the contract approaches. Until the penultimate trading day, we use changes of the FFAs+1Q. On the last trading day, we use the difference

⁵¹www.clarksons.net/ created from shipping brokers surveys for the rate on a specific route for the respective contract. If there was no sealing of a contract due to thin trading, estimated values of the brokers are used.

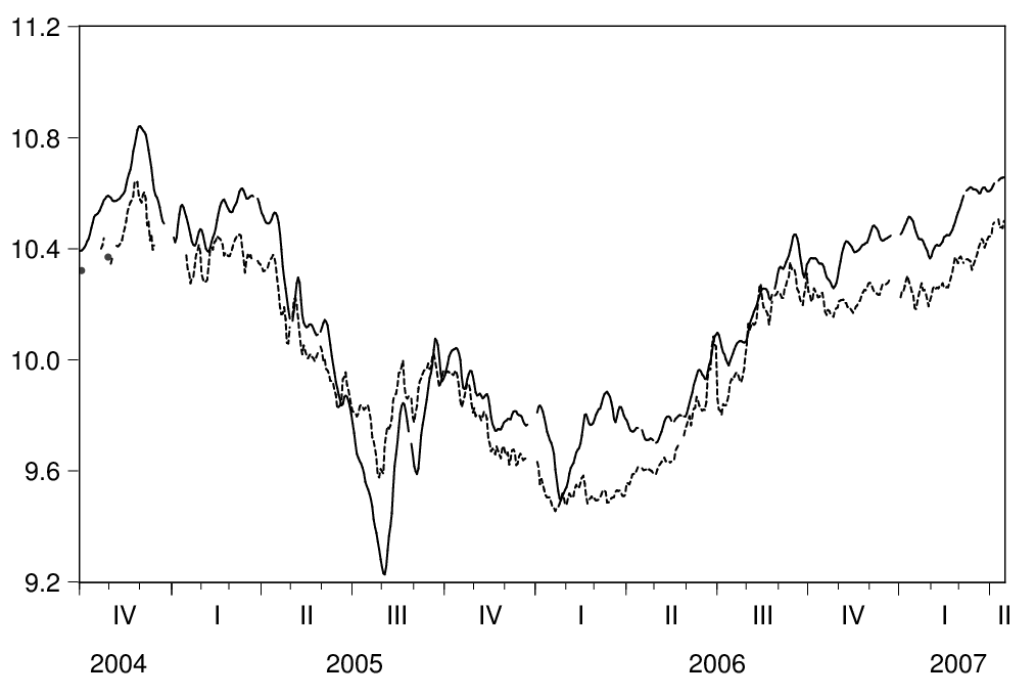
⁵²Weighted average of the busiest four routes of the respective index.

between the FFA+1Q on the last trading day and the FFA+2Q on the penultimate trading day so that only differences of contracts with the same maturity are used.

Time Charter rates are grouped according to ship sizes and are averages of particular routes, with Capesize and Panamax showing the highest trading volume and with maturities of six months, one year and two years.

Finally, we use 4TC spot rates in the BPI and BCI, so that the underlying non-storable service is identical to the ones in the FFA contracts. All data is in USD/day or month or year for Time Charter and USD/ton for spot and FFA contracts. Figure 44 depicts FFA and spot rates on Panamax routes. As is visible, both rates seem to co-move, which we are testing for. Descriptive data can be found in the appendix in Table 18.

Figure 44: Comovement spot and FFA contracts on panamax routes



Notes: This figure shows the comovement of spot and FFA contracts on Panamax routes. The black solid line denotes the log of spot rates and the dashed line is the log of FFA+2Q contract. Both time series are missing some observations due to thin trading.

6.4 Estimation methods

We compare forecasting models to a random walk that postulates tomorrow's spot rate to be the sum of the value of yesterday and white noise, ϵ , given an autoregressive parameter ρ of unity, given by Equation 91

$$E_t S_{t+1} = \rho S_t + \epsilon_{t+1}. \quad (91)$$

Estimating future spot rates based on this equation should produce the highest forecast error compared to the other models.

Additionally, we test if the UEH holds so that forward prices are unbiased estimators for future spot rates. We do this by performing cointegration tests for the FFA rates and spot rates so that $S_t = \alpha + \beta F_{t,t-n} + \epsilon_t$ with F_{t-n} being the forward rate at $t - n$ maturing at t , ϵ_t is white noise and $\alpha = 0$ and $\beta = 1$.

Stationarity tests suggest F_t and S_t to be non-stationary.⁵³ FFA and spot rates are cointegrated, if there exists a vector V so that $V = [S \ F]'$ is $I(d - b)$ and $b > 0$. According to Engle and Granger (1987), Error Correction models (VECMs) exist for every cointegration model. Equation 92 denotes the reduced form VECM

$$\Delta X_t = \Pi X_{t-p} + \Gamma_p \Delta X_{t-p} + u_t. \quad (92)$$

VECMs describe changes in the logs of spot and forward rates by own and the other variable's past observations. X_t is a column vector with $X = [S \ F]_t'$ and u_t a two-dimensional reduced form white noise term with a time-invariant covariance matrix Σ_u . Γ_p are square coefficient matrices of dimensions 2 for short-run adjustments. Π is the long run effect with $\Pi = \alpha \times \beta'$. α is known as the loading coefficients and β as the cointegrating vector capturing the speed of convergence to the common trend after shocks. Each of the right-hand side Γ s are error correction terms, restoring equilibrium after stochastic disturbances by restricting the variables to converge to the pre-shock path. The system is estimated by OLS and also includes a constant.

If the UEH holds, the cointegrating vector V is equal to $[1, -1]'$ so that the difference between FFAs and spot rates is zero in the long run. In this case, markets would be efficient, and the UEH would hold.

We also estimate VAR models, ignoring the cointegration relationship as follows

$$\Delta X_t = A(L)\Delta X_{t-1} + u_t, \quad (93)$$

⁵³Using an Augmented Dicky Fuller test (Dickey and Fuller, 1979), the Kwiatkowski - Phillips - Schmidt - Shin - Test (KPSS-Test) and the Phillip-Perron-Test Phillips and Perron (1988).

with identical variable definitions as in Equation 92 and $A(L)$ is a lag polynomial of order two and the estimation includes a constant. We additionally estimate an univariate ARIMA process as

$$\Delta S_t = \sum_{i=1}^p \alpha_p \Delta S_{t-p} + \sum_{j=1}^q \beta_q \epsilon_{t-q} + \epsilon_t. \quad (94)$$

Equation 94 predicts future values for the spot rate as a linear combination of past residuals and past changes in the spot rate. We also extend the ARIMA model to include Time Charter rates and refer to this model as ARIMA-TC, as described in Equation 95

$$\Delta S_t = \sum_{i=1}^p \alpha_p \Delta S_{t-p} + \sum_{j=1}^q \beta_q \epsilon_{t-q} + \sum_{k=1}^r \gamma_r \Delta TC_{t-r} + \epsilon_t. \quad (95)$$

ΔTC_{t-q} denotes change in the Time Charter rate from q Fridays ago.

We document the quality of forecasts as RMSE as $\sqrt{(\hat{S}_{t+1} - S_{t+1})^2}$, which is the difference between the fitted value and the actual spot rate. Determining the order of the system is essential since spot rates experience a high degree of autocorrelation. SC and AIC (Schwarz, 1978; Akaike, 1974) suggest an ARIMA(2,1,0), VAR(2), VECM(2) and ARIMA-TC(2,1,0,1).

We divide the full sample into an estimation and a forecast period, following Tashman (2000). We then estimate constant rolling windows by estimating one-day-ahead forecasts for the spot rate (\hat{S}_{t+1}) with information known at t . The estimation window is then pushed one day ahead, and the oldest observation drops out since otherwise new observation weights would decline and forecast errors would not be comparable. The next one-day-ahead forecast is then produced using the updated, rolled estimation window with slightly new regression coefficients. This procedure is repeated until the window reaches the last observation of the sample and one-day-ahead forecasts are produced at every point in time, depicted in Figure 45.

To produce two-day out-of-sample forecasts for the spot rate \hat{S}_{t+2} , we rely on one-day-ahead forecasts and proceed as follows:

1. Produce one-day-ahead forecasts for spot and FFA rates for $t + 1$, according to the procedure described above in Figure 45.

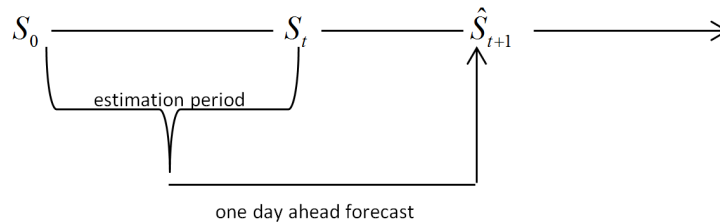
2. Use these one-day-ahead forecasts to create a forecast for the next day with an updated estimation regression which includes \hat{F}_{t+1} and \hat{S}_{t+1} . This, then, is the two-day forecast as depicted in Figure 46

Alternatively to letting the window move one step ahead, we could also let the window roll for two periods, which would create non-overlapping forecast intervals. Such procedure, however, would result in a smaller number of forecasts. We, therefore, move the estimation window only by one step and create overlapping forecast intervals for $t + 2$.

The procedure is repeated recursively to create forecasts for $t + n$, requiring forecasts for spot and FFA rates up to $t + n - 1$. In conclusion, we have nine forecasts for every observation point in our sample, excluding the final 9 ones.

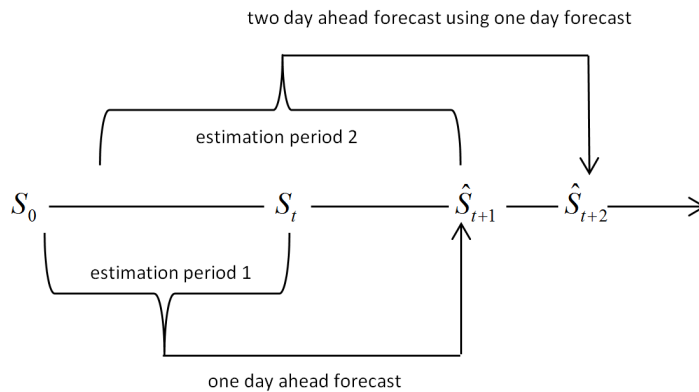
We then take Time Charter rates into account and extend the ARIMA model to provide evidence if TC rates carry additional information into the market. Time

Figure 45: One day ahead forecasts S_{t+1}



Notes: One day ahead forecasts of spot price S_{t+1} , using realized information available up to t . Values of spot and FFA rates are known up to t .

Figure 46: Two-day forecasts, using \hat{S}_{t+1}



Notes: Two-day forecasts for $t + 2$ with overlapping forecast intervals. For VARs, ARIMA models and VECMs, we also require forecasted values for S_{t+1} . After each prediction, the window moves one period ahead and we create the next two-day forecast.

Charter contracts are negotiated for maturing with six months, one, three or five years.

6.4.1 Descriptive data & hypotheses

We test the data for stationarity using following hypothesis with an Augmented Dicky Fuller test (Dickey and Fuller, 1979).

H_1 : Spot rates, FFA rates, and TC rates are integrated of order 1.

If rates are non-stationary, we test for cointegration of spot and FFA rates to find the cointegration vector, which provides the second hypothesis.

H_2 : Spot rates and FFA rates are cointegrated with a cointegration vector of $(1, -1)$.

If the cointegration vector holds, the UEH holds and markets are efficient.

Forecasts based on random walks have low predictive power and are expected to have the highest RMSE. We expect the ARIMA-TC model to outperform the ARIMA model, which in turn are expected to produce higher forecast errors than the VAR models which take multivariate relationships into account. Finally, if the rates are cointegrated, we expect the VECMs to create the highest quality forecast, which is our last hypothesis.

H_3 : $\text{RMSE}(\text{RW-Walk}) > \text{RMSE}(\text{ARIMA}) > \text{RMSE}(\text{ARIMA-TC}) > \text{RMSE}(\text{VAR}) > \text{RMSE}(\text{VECM})$

Table 18 in the appendix reports descriptive data for Panamax and Capesize contracts. Standard deviations for spot and FFA rates for Panamax and Capesize contracts are similar. Observations are varying because of missing observations due to thin trading. The unit root can be rejected for the return series of all contracts. Logs are $I(1)$, except for the Capesize $TC + 6M$ contracts⁵⁴. We, therefore, cannot reject H_1 , and consequently we use log returns in our estimations.

To investigate whether H_2 can be rejected, we also test the data for a cointegration relationship, following Engle and Granger (1987) before estimating VECMs. By doing so, we test the UEH and Table 20 and 21 report test results, which show that evidence for cointegration and therefore market efficiency is mixed and route-dependending. The strength of the relationships also depends on the maturity of

⁵⁴For which the Null can be rejected at the 10% level. The KPSS and Phillips-Peron test confirms the results.

the FFA contracts. For the FFA Panamax rates, the null hypothesis is rejected,⁵⁵ whereas it cannot be rejected for both Time Charter rates.

Hence, evidence for Hypothesis 2 is thus mixed. We do not expect VECMs to outperform forecasts from VAR models for all markets, so we correct H_3 to acknowledge the results of the cointegration analysis as

$$H3': \text{RMSE(RW-Walk)} > \text{RMSE(ARIMA)} > \text{RMSE(ARIMA-TC)} > \\ \text{RMSE(VECM)} > \text{RMSE(VAR)}.$$

In the next section, we report results for the forecastability of future spot rates to check whether H_3 holds.

6.5 Estimation results

The estimation window contains data from October 1, 2004, to December 12, 2007, comprising of 800 observations which creates stable coefficients when moving windows ahead as suggested in Tashman (2000). The results of the ARIMA, VAR, VECM and ARIMA-TC estimations are as follows, and regression outputs are depicted in the appendix. All models are estimated for Capesize and Panamax markets.

6.5.1 ARIMA estimation

We estimate ARIMA(2,1,0) models so that no moving average part is included in the estimation, as suggested by information criteria. Results for spot rates on Panamax routes suggest an adjusted R^2 of 0.75 as depicted in Table 22 in the appendix. Both AR coefficients are significant on the 1% level. Results for Capesize contracts are presented in the appendix from Table 23. The explained variance here is less compared to Panamax routes, taking a value of 0.65. Both AR coefficients are significant and compare to the ones from the Panamax routes. The Durbin-Watson statistic suggests that the AR(2) sufficiently controls for autocorrelation as the value is close to 2.

6.5.2 VAR estimation

VAR estimation results are given in the appendix in Table 24, estimated by OLS for FFA and spot contracts.

Adjusted R^2 s for the different contracts on Panamax routes are around 0.80 for the spot rates. The coefficients for the first two lags of the spot rates and the first

⁵⁵For FFA+1Q the 5% level, for the FFA+2Q and FFA+1A on the 10% level.

lag of the FFAs are significant at least at the 5% level. The evidence for the impact of the second lag for the FFAs is mixed. Therefore, only the first lag of the FFA seem to have a significant impact on current spot rates. The information contained in FFAs is hence processed within a day. FFA+1A have the weakest impact on future spot rates.

The explanatory power of the FFA regression is low and between 3-10%, so less of the variance is explained compared to the spot regressions. Therefore, FFA lags have a price finding role for spot rates, but the vice versa effect seems weak.

Results for the Capesize contracts for the spot rate estimation suggest adjusted R^2 s around 0.67, being slightly lower compared to Panamax contracts, as in the ARIMA estimation. Coefficients for all spot rates FFAs are significant on the 1% level. The FFA regressions again explain less of the variance of the regression model.

Using the VAR, significant coefficients and high R^2 s point towards the forecastability of future spot rates. FFA rates do not seem to be forecastable, and Granger causality tests point towards forecastability of future spot rates only so that FFAs are not Granger caused by spot rates.

6.5.3 VECM estimation

VECM estimates are reported in Table 36, which can be found in the appendix. Significance and explained variances are similar to the VAR estimation results.

The error correction terms show that the spot rate is responsible for adjustments back to common trend after disturbances, and coefficients take a significant negative sign. Disequilibrium is corrected the fastest in the most liquid forward (FFA+1Q) so that roughly 2% are corrected for Capesize contracts and 1% for Panamax contracts. The error correction term of the FFAs, however, is insignificant, so mainly exogenous events have an impact on the forward rate and not on the spot rate. The effects of the lags of spot rates on FFA rates is only partly significant, depending on the route and FFA contract type.

The coefficients for the long-run impact, the cointegration relationship, are all significant on the 1% level so that the cointegration vector causes short-run deviations from the common trend to die out over time. However, this vector deviates from the cointegration vector of $(1, -1)$. Therefore, both rates do not need to converge as maturity approaches. We can, therefore, reject the UEH. However, a risk premium of zero which as suggest by the PEHTS can only be rejected for some of the constants in the VAR estimation, whereas it cannot be rejected for the VECM estimation (Veenstra and Franses, 1997). The constants suggest that evidence for risk premia is mixed. For some of the contracts, holding FFA contracts yield a

positive risk premium, whereas it is the opposite for other contracts. The EHTS can therefore neither be confirmed nor rejected. Investigating the stability of the risk premium, however, would require to a time-varying model structure.

6.5.4 ARIMA-TC estimation

We now extend the ARIMA spot rate model for the last known TC rate which could provide additional explanatory power. Results can be found in Table 42 in the appendix.

We estimate ARIMA-TC(2,1,0; 1), suggested by AIC and SC information criterion so that only the last value of the Time Charter rate is added to the ARIMA estimation. Only one TC lag is added, because the lag two Fridays ago has low explanatory power since information had already been processed by the market through the spot rates, which are on daily basis.

The adjusted R^2 s for the ARIMA-TC estimation for both routes, and all contracts compare to the ARIMA models, but the TC coefficients are insignificant for all contracts. Moreover, model selection criteria recommend to prefer the ARIMA over the ARIMA-TC model. Our results hence show that TC rates do not add significant explanatory power to forecast spot rates.

In conclusion, VAR estimations seem suited to create forecasts for future spot rates since the cointegration relationship is ambiguous. Therefore, markets are partly inefficient so that we reject the UEH since the cointegration vector of $(1, -1)$ could not be confirmed for all contracts and FFA maturities.

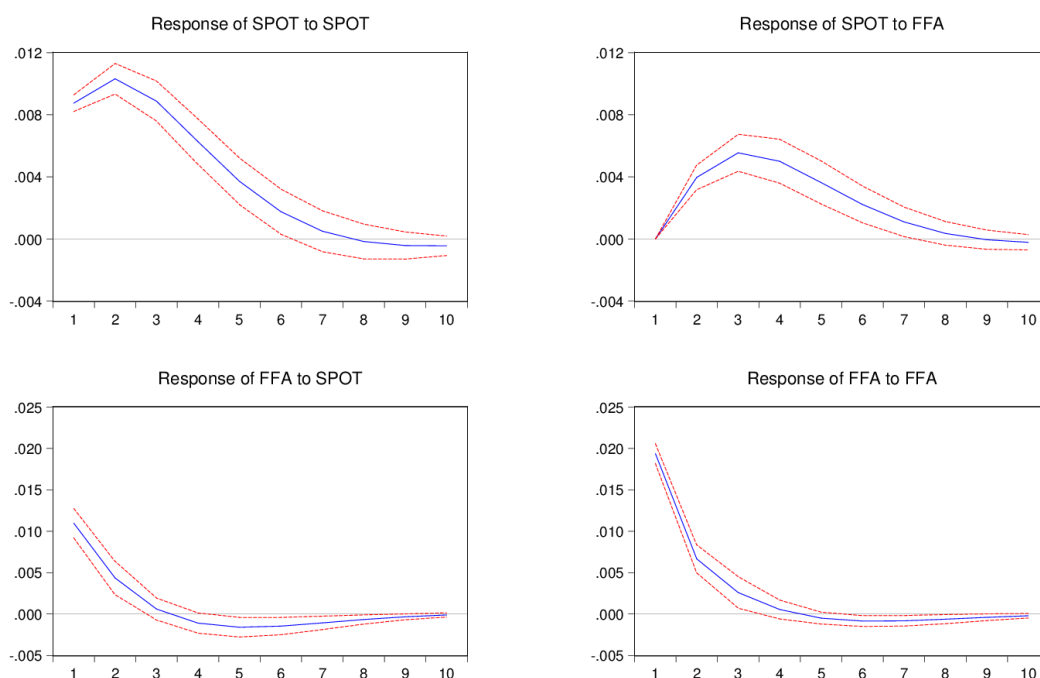
6.5.5 Impulse responses

VAR models are capable of capturing the relationship between spot and FFA rates. We consequently create orthogonalized⁵⁶ impulse responses for the most liquid market, Panamax FFA+1Q contracts, to investigate the impact of shocks on the two rates, ordering spot rates first.⁵⁷ Results of the Capesize contracts are almost identical. Figure 47 displays impulse responses to a one standard deviation shock to both rates.

Figure 47 suggests that the impact of spot rate shocks on FFAs is as weak as suggested by the low explanatory ability of the regression. An increase in spot rates does not necessarily induce FFA rates to rise as well since FFAs mature at least one quarter in the future. As is evident, the influence of spot rate shock on the FFAs

⁵⁶Cholesky decomposed variance-covariance matrix of the reduced form residuals while the order of the variables being relatively robust for the qualitative response direction.

⁵⁷The effects of shocks is robustness to reversing the order.

Figure 47: Orthogonalized impulse responses VAR estimation


Notes: Orthogonalized impulse responses for VAR estimation on Panamax contracts, FFA+1Q. Impulse responses are orthogonalized with Cholesky decomposition, data is in growth rates, $X_t = (S_t, FFA_t)$. The estimation includes two lags and a constant. Impulse responses display daily changes in the rates. 0.02 hence denotes a 2 percentage point increase.

dies out already after three days with a maximum response of a one percentage point increase.

In contrast, a shock to FFAs is more persistent, causing spot rates to increase in a hump-shaped manner, peaking at a 0.5 percentage point increase per day. Spot rates converge to pre-shock levels after seven days.

6.5.6 Forecasts and RMSE

We compare the quality of forecasts to pick a model that can be exploited to create a trading scheme. We compare forecasts regarding RMSE in Table 14 (Panamax) and Table 15 (Capesize) for different FFA maturities and forecast horizons, as well as an improvement in RMSE when switching forecast models.

Negative signs denote a decrease of RMSE, equivalent to an improvement in the accuracy of the forecast.

The random walk has the highest RMSE while increasing as the forecast horizon is extended. It is outperformed by all other models in Panamax and Capesize

Table 14: Reduction in RMSE in % for Panamax contracts

horizon	RW	FFA+1Q		FFA+2Q		FFA+1A		TC_{6M}		TC_{1A}	
		ARIMA v RW	VAR v ARIMA	VEC v VAR	VAR v ARIMA	VEC v VAR	VAR v ARIMA	VEC v VAR	ARIMA-TC v ARIMA	VEC v VAR	ARIMA-TC v ARIMA
+1d	4.40	-69.24	-9.60	-1.01	-8.22	-0.85	-7.41	-0.81	-0.68	-0.81	-0.53
+2d	8.43	-58.17	-12.95	-0.77	-11.61	-0.74	-10.58	-0.89	-0.35	-0.89	-0.10
+3d	12.04	-34.33	-12.81	-0.21	-12.10	-0.38	-11.18	-0.75	-0.27	-0.75	0.12
+4d	15.28	-22.38	-10.80	0.37	-10.23	0.02	-9.77	-0.57	-0.37	-0.57	0.12
+5d	18.19	-13.64	-8.93	0.88	-8.47	0.35	-8.33	-0.44	-0.44	-0.44	-0.07
+6d	20.79	-7.81	-7.67	1.44	-7.08	0.68	-7.31	-0.29	-0.40	-0.29	-0.29
+7d	23.05	-4.41	-5.87	1.96	-5.41	1.01	-5.91	-0.18	-0.22	-0.18	-0.46
+8d	25.11	-2.57	-4.81	2.46	-4.42	1.32	-4.98	-0.06	0.08	-0.06	-0.51
+9d	27.11	-1.87	-4.03	2.59	-3.67	1.40	-4.23	-0.05	0.44	-0.05	-0.37

Notes: RMSE and model comparison for Panamax contracts. FFA+x denotes the FFA contract with maturity $x \in [1Q, 2Q, 1A]$. The first column is the forecast horizon; the second column represents RMS errors in '00 for the log of the spot rate and forecast. The third column depicts the percentage reduction in RMSE using the ARIMA model, followed by comparing the models with each other in the subsequent columns. The percentage improvement is measured as $(RMSE_{ARIMA} - RMSE_{RW})/RMSE_{ARIMA}$. Negative values denote improvements in forecastability in % relative to a comparable model.

Table 15: Reduction in RMSE in % for Capesize contract

horizon	RW	ARIMA v RW		FFA+1Q		FFA+2Q		FFA+1A		TC _{6M}		TC _{1A}	
		ARIMA	VAR	ARIMA	VEC v VAR	ARIMA	VEC v VAR	ARIMA	VEC v VAR	ARIMA-TC v ARIMA	VAR	ARIMA-TC v ARIMA	ARIMA-TC v ARIMA
+1d	6.15	-64.36	-0.88	-12.75	-0.90	-10.97	-0.91	-10.97	-0.91	-0.28	-0.91	-0.28	-0.55
+2d	11.58	-39.73	-0.72	-13.71	-0.90	-12.85	-1.04	-12.85	-1.04	-0.09	-1.04	-0.09	-0.26
+3d	16.44	-24.99	-0.21	-11.22	-0.60	-11.15	-0.86	-11.15	-0.86	0.02	-0.86	0.02	-0.06
+4d	20.70	-17.42	0.58	-8.22	0.05	-8.51	-0.44	-8.51	-0.44	0.31	-0.44	0.31	0.46
+5d	24.54	-12.17	1.27	-6.00	0.56	-6.39	-0.07	-6.39	-0.07	0.41	-0.07	0.41	0.81
+6d	27.93	-8.48	1.91	-4.60	1.03	-4.93	0.27	-4.93	0.27	0.50	0.27	0.50	1.04
+7d	31.09	-6.47	2.47	-2.95	1.53	-3.22	0.62	-3.22	0.62	0.38	0.62	0.38	1.14
+8d	34.08	-5.69	3.02	-2.08	2.02	-2.26	0.99	-2.26	0.99	0.07	0.99	0.07	1.08
+9d	37.00	-6.15	3.22	-1.08	2.27	-1.10	1.19	-1.10	1.19	-0.45	1.19	-0.45	0.76

Notes: RRMSE and model comparison for Capesize spot rate contracts. FFA+x denotes the FFA contract with maturity $x \in [1Q, 2Q, 1A]$. The first column is the forecast horizon; the second column presents RMS errors in '00 for the log of the spot rate and forecast. The third column depicts the percentage reduction in RMSE using the ARIMA model, followed by comparing the models with each other in the subsequent columns. The percentage improvement is measured as $(RMSE_{ARIMA} - RMSE_{RW})/RMSE_{ARIMA}$. Negative values denote improvements in forecastability in % relative to a comparable model.

markets. Improvements by switching from ARIMA to VECM/VAR also increases the accuracy of the predictions. Forecasts for more than one day require the forecast for the previous days, which cause RMSEs to grow in the long-run since forecast errors are transmitted.

ARIMA vs. random walk: ARIMA models, generate lower forecast errors compared to the random walk. The model is capable of forecasting short-run changes the spot rates while accuracy decreases as the forecasting horizon is extended, indicated by only a small improvement in the RMSE compared to the random walk. For nine-day forecasts, the improvement in RMSE is only 1.9% (Capesize) and 6.1% (Panamax).

VAR vs. ARIMA: Since we showed that FFAs have predictive power for future spot rates, we prefer a multivariate model over an univariate one. VARs outperform ARIMA models for all three types of FFA maturities on Capesize and Panamax markets. FFA+1Q generates forecasts with the smallest RMSE. A decline in RMSE can be observed for short-run forecasts and decrease as the forecast horizon is extended.

VECM vs. VAR: VECMs marginally outperform forecasts generated from VARs in the short-run, but with an increase in RMSE for extended forecast periods, owed to the weak cointegration relationship.

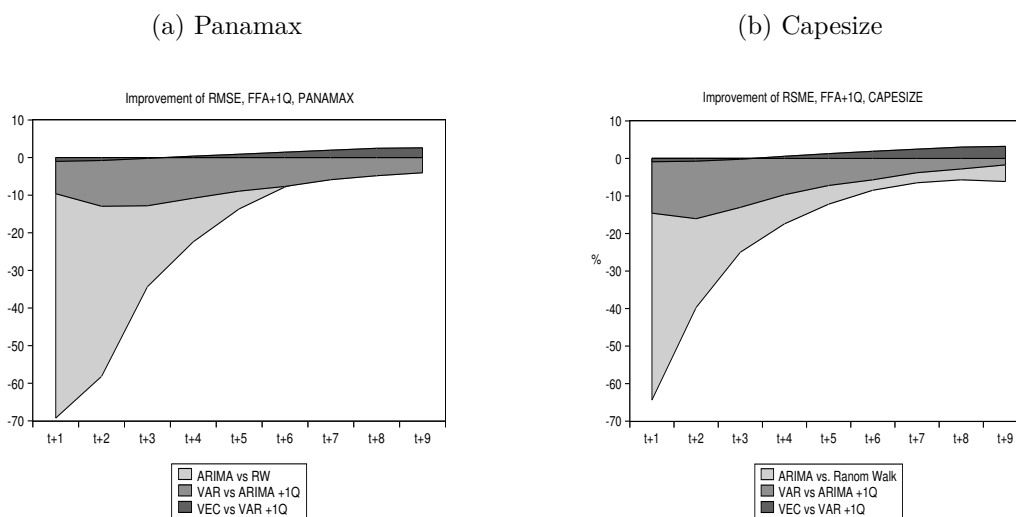
ARIMA vs. ARIMA-TC: Adding TC rates to ARIMAs marginally decreases RMSE for Panamax contracts. Nevertheless, ARIMA models outperform the ARIMA-TC for the Capesize routes for the majority of the contracts, while $TC + 2Q$ and $TC + 1A$ are similar in terms of accuracy. Therefore, Time Charter rates cannot add explanatory power to producing spot rate forecasts which can result from a misalignment between TC and the underlying routes. These routes are not substitutable, so rates do follow a common trend.⁵⁸ Moreover, TC rates are weekly-reported each Friday so that spot rates can process this information directly. Therefore, changes in the TC rates on Friday have an impact on the spot rates the following Monday. Furthermore, in ARIMA-TC models we assume an even impact of TC rates on all daily spot rates and not only on Monday's rate. This assumption may be too simplistic and a reason TC are incapable of enhancing the ARIMA estimation to reduce RMSE.

Figure 48 graphically summarizes improvements and deteriorations in RMSE by switching models for the two most liquid contracts. As is obvious, the strongest decrease in RMSE is archived by switching from the RW to the ARIMA models and then to the VAR. Figures for less liquid markets can be found in the appendix in

⁵⁸However, we minimize this effect by selecting similar routes.

Figure 51, for which results are similar. As obvious, the strongest decrease in RMSE can be archived in during the first days of forecast, while the pattern is similar for Capesize and Panamax markets.

Figure 48: RMSE change when switching models



Notes: The figure reports improvement of RMSE by switching from a random walk to the respective model using FFA+1Q, for different forecast horizons and Capesize and Panamax markets. Negative figures denote a decrease of RMSE error when switching to the respective model. VAR vs. ARIMA denotes, therefore, the change in RMSE for each forecast horizon when switching from ARIMA model to a VAR model. Percentage improvement in RMSE.

In conclusion, VAR models are best suited for generating spot rate forecasts. ARIMA-TC do not to outperform ARIMA models in terms of RMSE. Therefore, we cannot reject Hypothesis 3.

6.6 Robustness and trading scheme

In this section, we use VAR estimates to create a trading scheme. A profit can be generated since the UEH does not hold so that the timing structure of cargo shipping contracts matters. Nevertheless, long-short strategies are not possible due to the inability to short-sell contracts in a secondary market.

Shipowners' and charterers' interests do not align; shipowners are interested in higher future prices while charterers reduce costs if prices decrease. Since market entry barriers are high in the shipping sector, we assume that intermediates or brokers exist matching supply and demand.

If the estimated models are known to such intermediates who are capable of estimating tomorrow's spot rate, they can apply this information to create short-run forwards with the charterers and shipowners. However, forecasting spot rates

to match the maturity of the FFAs is impossible since this would imply estimating a forecast for 90 days. As seen, VAR-generated forecasts become unreliable after 5-6 days, and the reduction in RMSE decreases compared to the random walk. Therefore, we create short-run forecasts for up to 9 days as described in earlier sections.

6.6.1 Trading scheme

We create signals for a market-maker to create an excess return. This market-maker then engages in fixing a forward agreement for the next days and matches both parties by engaging in over-the-counter trading.

We assume that neither charterers nor shipowners have market power so that they do not influence the current or future spot and FFA rate by engaging in more or less trading. Moreover, for charterers it is assumed that shipping on a particular day is necessary so that a wait-and-see strategy is undesirable.

The decision to engage in trading, then, is only determined by the model, so the intermediate cannot deviate from this strategy because he believes he has prime knowledge of the market compared to other market participants.

In detail, we proceed as follows:

1. Estimate VARs for Capesize and Panamax markets, using FFA+1Q, FFA+2Q and FFA+1A from 6 June 2004 to 12 December 2007.
2. Create forecasts of the spot rate for one to nine days with constant rolling windows.
3. If the estimated change in the spot rate is negative (decline in spot rates), offer the charterer an one to nine days FFA contract to ship at today's rate S_t . Evenly share the profit with the owner, charterer, and broker. At maturity, buy shipping service at S_T from owners and receive profit $\frac{1}{3}(S_t - S_T)$.
4. If the estimated change in the spot rate is positive (increase in prices), offer shipowners an one to nine days FFA contract to find a charterer at T at price S_t . At maturity, find charterers that ships at market price S_T and receive a profit of one-third of the increase in S as the profit is evenly shared.
5. Estimate forecast until the rolling window reaches the final data point.
6. Create an index with December 2007 = 100 for each of the FFAs on both markets for one to nine days.

7. Finally, compare trading scheme to MSCI world index as benchmark indicator.

Note that we assume that a bid/ask spread exists. The broker receives only 1/3 of the profit so that the owner and the charterer both have an incentive to engage in forward trading. Otherwise, the market participants would be indifferent. Table 16 denotes the results from this strategy, as it compares mean and aggregated returns using daily data for both markets, all three FFA maturities and one to nine-day forecasts.

Table 16: Returns for Panamax and Capesize trading scheme

	+1d	+2d	+3d	+4d	+5d	+6d	+7d	+8d	+9d
Capesize Market									
Mean +1Q	0.8190%	0.4734%	0.1310%	-0.0410%	-0.0538%	-0.0440%	-0.2342%	-0.0850%	-0.2250
Std	1.0580%	1.2652%	1.3737%	1.3989%	1.4053%	1.4076%	1.3919%	1.4068%	1.3783
correct	82.8402%	65.0888%	49.7041%	47.3373%	44.9704%	49.1124%	44.9704%	45.5621%	42.0118
Mean +2Q	0.8218%	0.4734%	0.1239%	-0.0337%	-0.0751%	-0.0018%	-0.1530%	-0.0615%	-0.1806
Std	1.0558%	1.2651%	1.3744%	1.3991%	1.4043%	1.4083%	1.4032%	1.4081%	1.3848
correct	83.4320%	65.6805%	50.8876%	47.9290%	44.9704%	52.0710%	45.5621%	43.7870%	43.1953
Mean +1A	0.8071%	0.2706%	0.2690%	0.4548%	0.1386%	-0.0254%	-0.0832%	-0.0869%	-0.0856
Std	1.0672%	0.3562%	0.3557%	1.2720%	1.3730%	1.3993%	1.4038%	1.4056%	1.4090
correct	84.0237%	84.0237%	84.0237%	65.0888%	49.1124%	47.3373%	44.9704%	47.9290%	44.3787
Panamax Market									
Mean +1Q	0.5575%	0.4056%	0.2397%	0.1122%	0.0461%	0.0261%	-0.0405%	0.0006%	-0.0642
Std	0.5172%	0.6439%	0.7212%	0.7521%	0.7580%	0.7624%	0.7759%	0.7910%	0.7897
correct	88.1657%	76.9231%	64.4970%	57.3964%	54.4379%	50.2959%	44.3787%	53.2544%	44.9704
Mean +2Q	0.5503%	0.4074%	0.2750%	0.1089%	0.0347%	0.0166%	-0.0275%	0.0175%	-0.0429
Std	0.5249%	0.6428%	0.7084%	0.7526%	0.7586%	0.7626%	0.7765%	0.7908%	0.7912
correct	87.5740%	77.5148%	66.8639%	57.3964%	53.8462%	50.2959%	46.1538%	54.4379%	44.9704
Mean +1A%	0.5489%	0.3971%	0.2688%	0.1032%	0.0293%	-0.0021%	-0.0014%	0.0086%	-0.0133
Std	0.5263%	0.6492%	0.7108%	0.7534%	0.7588%	0.7628%	0.7770%	0.7910%	0.7922
correct	87.5740%	76.9231%	67.4556%	57.3964%	52.6627%	50.2959%	46.1538%	53.2544%	46.7456
Sum MSCI W									-0.1021%
Std MSCI W									1.0163%

Notes: The table shows mean, standard deviation and ratio of correctly estimated forecasts of the intermediation strategy described above, for Capesize markets (upper) and Panamax markets (lower), for forecasts for one day (+1d) for up to nine days (+9d) and using FFA maturities for the respective market ranging from one quarter (+1Q) to one year (+1A). The last two lines denote the mean return and standard deviation of the MSCI World index (a market equity-weighted benchmark of world-wide stock-performances), during the same time frame (13 December 2007 to 1 August 2008). Mean refers to the arithmetic mean of the log returns.

Table 16 implies that the trading strategy for the Capesize markets outperforms, in the absence of trading costs, the ones on Panamax routes using one-day forecasts. The choice of maturity for the type of FFA contracts is not essential for the profit that can be generated following the trading scheme, although short maturities (FFA+1A) generate slightly higher returns.

Capesize trading schemes for the sample period would have created a mean return of 0.82% per day for one-day forecasts. In aggregation, this strategy would

have led to an almost quadrupling of the value invested, which compares to an almost tripling for the Panamax routes. The profit from following this strategy decreases as the forecast horizon is extended. For Capesize contracts, the mean return becomes negative after the third/fourth forecast day, suggesting that only short-run forecasts are capable of generating excess returns which overlap with our results in the estimation section. In contrast, using Panamax contracts, it is possible to generate profits for up to six days, while following the nine-day strategy leads to a loss. In addition to looking at mean returns, we calculate Sharpe ratios (Sharpe, 1966) that link risk and return as follows.

$$SR_i = \frac{\bar{r}_i - \bar{r}^{msci}}{\sigma_i} \quad (96)$$

in which \bar{r}_i denotes the mean return generated by the respective strategy i , \bar{r}^{msci} denotes the mean return by the benchmark and σ_i is the standard deviation of the respective strategy. Table 17 gives the results.

Table 17: Sharpe ratios for trading scheme

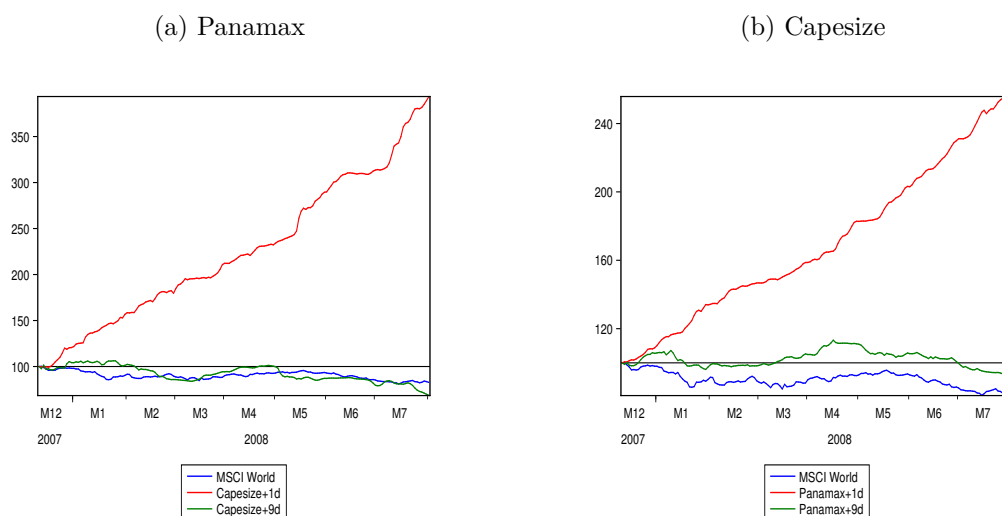
	+1d	+2d	+3d	+4d	+5d	+6d	+7d	+8d	+9d
	Capesize Market								
FFA+1q	0.871	0.455	0.170	0.044	0.034	0.041	-0.095	0.012	-0.089
FFA+2q	0.875	0.455	0.164	0.049	0.019	0.071	-0.036	0.029	-0.057
FFA+1a	0.852	1.046	1.043	0.438	0.175	0.055	0.013	0.011	0.012
	Panamax Market								
FFA+1q	1.275	0.788	0.474	0.285	0.196	0.168	0.079	0.130	0.048
FFA+2q	1.243	0.793	0.532	0.280	0.180	0.156	0.096	0.151	0.075
FFA+1a	1.237	0.769	0.522	0.272	0.173	0.131	0.130	0.140	0.112

Notes: This table gives Sharpe ratios for the individual trading schemes for one to nine days for FFAs with three different maturities for the estimation. The Sharpe ratio is calculated as $SR_i = \frac{\bar{r}_i - \bar{r}^{msci}}{\sigma_i}$.

As Table 17 shows, the Sharpe ratios are higher for Panamax markets compared to Capesize markets so that an investor may choose this market since returns given a unit of risk is higher.

Using the signals we have generated from our VAR estimations, we create an index for each respective contract, which we compare with the benchmark which we rescale to 12 December 2007 = 100. Figure 49 compares the MSCI world index to the ones generated by the two trading scheme.

Figure 49 depicts that using one-day forecasts in Panamax and Capesize markets would have outperformed the benchmark, which decreased by 18% during the observation period.

Figure 49: Trading scheme returns for both contracts


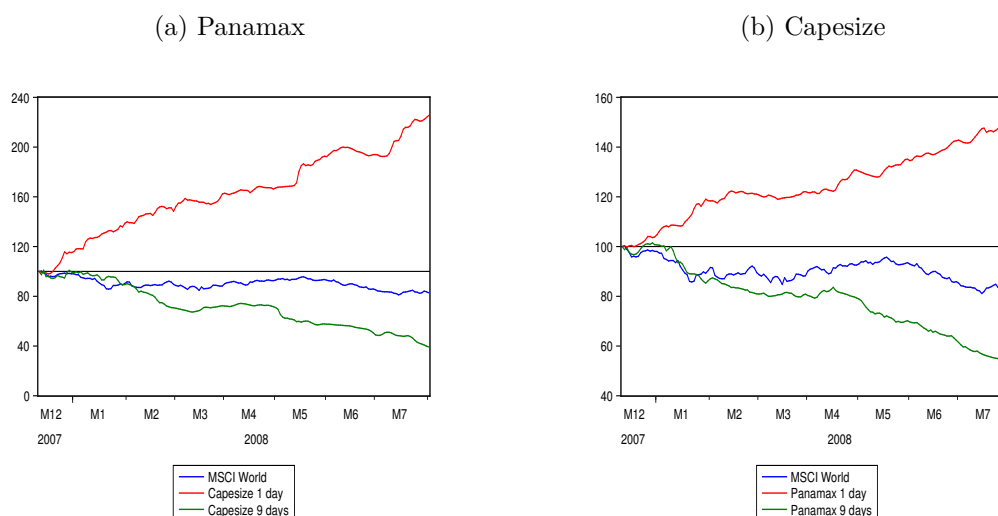
Notes: Development of index created by trading scheme on Capesize and Panamax routes. The benchmark index is the MSCI world index.

Creating FFAs, however, requires negotiation with the counter-party, which is costly. We assume that such negotiations account for approximately 1% as transaction costs. Therefore, any profit generated by the broker is reduced by 1% per trade. Again, using the VAR above, we estimate rolling windows and perform forecasts for up to nine and compare the development with the index, now containing the transaction costs. Figure 50 depicts the development of the indices created over time if we account for trading costs⁵⁹.

Incorporating transaction costs reduce the excess return that can be generated from the trading strategies. Profits, using the one-day strategy is reduced to the double (1.4 fold) in Capesize (Panamax) markets of the investment. When estimating forecasts and following the trading strategy for up to nine days, the value of the amount invested is reduced by half for Panamax contracts and reduced to 40% of the original quantity in the case of the Capesize contracts.

In conclusion, the short-run forecasts are capable of generating an excess profit over the market return. However, these strategies might be limited since trading can be thin at times so that the broker would not be able to find counterparts which reduce excess returns. Therefore, it is unlikely that the maximum profit can be generated. However, since the majority of the forecasts is correct, the information embedded in the VAR model should be taken into account by charterers and shipowners alike when deciding on engaging in an over-the-counter business with a counterpart.

⁵⁹For simplicity reasons, we ignore tax issues.

Figure 50: Trading scheme returns for both contracts incl. transaction costs

Notes: Development of index created by trading scheme Capesize routes, assuming transaction costs of 1% per trade. The benchmark index is the MSCI world index.

6.7 Conclusion

This paper finds that spot and future markets in the cargo shipping sector are inefficient due to a physical restriction on short sales and the lack of speculative trading in the market. Unlike in paper or commodity markets, rates are not connected through a Cost-Of-Carry relationship.

In this paper, we use data on Panamax and Capesize shipping vessels on the most trafficked routes. We investigate the relationship between spot and Forward Freight Agreement rates (FFA). The evidence for cointegration between both rates is mixed: Results show that cointegration depends on the specific contract and route. We can empirically reject the theoretical cointegration vector (1,-1) using a VEC model for most markets and contracts so that the forward price is not an unbiased estimator for future spot rates so that the UEH does not hold. Charterers and shipowners, therefore, have to take the timing structure of freight rates into account when deciding on their rate portfolio. The evidence for the EHTS is mixed so that we cannot demonstrate if holding spot rates or FFA rates is more profitable. Since we do not build a model with a time-varying structure, we cannot determine the time-varying structure of the risk premium, while the PEHTS for a zero risk premia can be rejected for some markets in the VAR estimation.

We, therefore, benefit from inefficient markets and compare the quality of forecasts for spot rates in the shipping sector, by using current spot and FFA rates. We proceed to compare RSME of forecasts derived from five different models: A

VAR VEC ARIMA, ARIMA-TC, and random walk model. We find that FFA rates Granger cause spot rates with only a weak vice versa relationship.

VARs outperform ARIMA and random walk in terms of RMSE. High-quality forecasts of rates with different maturities in the freight market are only observable in the short-run and models are incapable of producing long-run forecasts for spot rates.

The comparison of RMSE shows that VARs can be used to explain future spot rates. We, therefore, use these models to create signals for brokers who match supply and demand to create excess returns. After controlling for transaction costs, our trading scheme would have created a return within one year, equivalent to roughly 40% for Capesize markets and a more than double return on Panamax routes.

Our results take a look at within-cycle dynamics because cycles in the shipping sector tend to reoccur approximately every six to nine years. Therefore, our findings are limited to within-cycle inferences. As soon as additional non-crisis data is available, the sample should be extended to include shipping cycles.

Finally, for future projects we suggest an EGARCH model to generate forecasts of future spot rates accounting for time-variable risk premia.

6.8 Appendix

Table 18: Descriptive data Panamax contract

Panamax	Spot	FFA+1Q	FFA+2Q	FFA+1A	TC 6M	TC 1A
Std deviation	0.0305	0.0426	0.0363	0.0309	0.0925	0.0706
Skewness	-0.5306	0.0058	-0.7065	-0.9204	-0.5847	-2.1081
Kurtosis	9.4431	9.8549	7.2469	11.922	11.182	16.918
Jarque-Bera	2121.292***	2349.474***	969.905***	4149.714***	722.987***	2238.225***
ADF (Level)	-1.8687	-1.3754	-0.9282	-0.8506	-2.2827	-2.0488
ADF (Diff)	-6.7696***	-18.496***	-16.232***	-16.010***	-4.6324***	-4.1507***
PP (Level)	-1.5148	-0.9636	-0.6880	-0.5722	-1.7351	-1.4697
PP (Diff)	-10.073***	-23.891***	-23.745***	-22.886***	-12.425***	-14.131***
KPSS (Level)	0.4900**	0.6166**	0.7264**	0.7543***	0.2380	0.2752
KPSS (Diff)	0.1054	0.1712	0.2347	0.2675	0.0881	0.1266
Q	1541.09***	161.96***	186.23***	199.89***	81.95***	82.57***

Notes: Descriptive data on Panamax contracts, all in log. Spot and FFA rates are daily data, and TC rates are on a weekly basis (Friday). The table also denotes the Jarque Bera statistic for normality and the Dicky Fuller statistic with respective significance (* denotes significance on 10%, ** on the 5% and *** on the 1% level), as well as the Phillips-Peron and KPSS test statistic. Ljung/Box test statistic for autocorrelation is also presented.

Table 19: Descriptive data Capesize contract

Capesize	Spot	FFA+1Q	FFA+2Q	FFA+1A	TC 6M	TC 1A
Std deviation	0.0415	0.0486	0.0409	0.0370	0.1236	0.1050
Skewness	0.2255	-0.4782	-0.6953	-3.2035	-0.7976	-0.9699
Kurtosis	13.871	8.8989	9.6780	44.9197	13.8851	14.1280
Jarque-Bera	5919.413***	1776.631***	2272.221***	89873.670***	1280.900***	1350.394***
ADF (Level)	0.03704	-1.1541	-0.7968	-1.1537	-2.6138*	-2.4428
ADF (Diff)	-6.3909***	-5.8113***	-6.0546***	-5.5513***	-5.3261***	-4.8318***
PP (Level)	-1.9115	-1.2240	-0.9828	-0.8123	-2.0849	-1.8348
PP (Diff)	-11.934***	-23.7029***	-23.5087***	-23.6411***	-14.1397***	-15.4428***
KPSS (Level)	0.3920*	0.5491**	0.6470**	0.6672**	0.2250	0.2597
KPSS (Diff)	0.0737	0.1391	0.1925	0.1815	0.0627	0.0928
Q	1400.20***	254.86***	234.73***	240.77***	66.67***	49.92*

Notes: Descriptive data on Capesize contracts, all in logs. Spot and FFA rates are daily data, and TC rates are on a weekly basis (Friday). The table also denotes the Jarque Bera statistic for normality and the Dicky Fuller statistic with respective significance (* denotes significance on 10%, ** on the 5% and *** on the 1% level), as well as the Phillips-Peron and KPSS test statistic. Ljung/Box test statistic for autocorrelation is also presented.

Table 20: Cointegration in Panamax market

	FFA+1Q	FFA+2Q	FFA+1A	TC 6M	TC 1A
C	-1.105325***	-1.167736***	-1.585646***	-1.221835***	-1.431696***
Ln(FFA / TC)	1.104891***	1.116421***	1.170882***	1.106553***	1.134231***
Adj. R2	0.892014	0.832096	0.779608	0.966652	0.883485
ADF p-value	0.0402	0.0794	0.0701	0.7970	0.7480
ADF p-value (1, -1)	0.0422	0.0930	0.1312	0.8880	0.7480

Notes: Cointegration relationship between FFAs and Time Charter rates of logs of Panamax data and results of the Augmented Dicky Fuller test (ADF). * denotes significance on 10%, ** on the 5% and *** on the 1% level.

Table 21: Cointegration in Capesize market

Capesize	FFA+1Q	FFA+2Q	FFA+1A	TC 6M	TC 1A
C	-1.413937***	-1.591170***	-2.150650***	-1.454579***	-2.164821***
Ln(FFA / TC)	1.130967***	1.153516***	1.219452***	1.123540***	1.193952***
Adj. R2	0.906328	0.832255	0.774076	0.933550	0.864065
ADF p-value	0.1420	0.0411	0.0537	0.9775	0.0102
ADF p-value (1, -1)	0.0973	0.0760	0.1205	0.3463	0.0077

Notes: Cointegration relationship between FFAs Time Charter rates and Spot rates of logs of Capesize data and results of the Augmented Dicky Fuller test (ADF). * denotes significance on 10%, ** on the 5% and *** on the 1% level.

Table 22: ARIMA(2,1,0) estimation, Capesize market

Variable	Coefficient	Std. Error	t Statistic	Prob.
Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 797 after adjustments				
Dependent Variable: DLOG(SPOTRATE)				
DLOG(SPOTRATE(-1))	1.024	0.0333	30.199	0.0000
DLOG(SPOTRATE(-2))	-0.293	0.0338	-8.686	0.0000
R-squared	0.657	Mean dependent var	0.001	
Adjusted R-squared	0.657	S.D. dependent var	0.024	
S.E. of regression	0.014	Akaike info criterion	-5.695	
Sum squared resid	0.156	Schwarz criterion	-5.683	
Log likelihood	22271.432	Durbin-Watson stat	2.04	0

Notes: Results for ARIMA(2,1,0) estimation for Capesize spot rates as univariate setup

Table 23: ARIMA(2,1,0) estimation, Panamax market

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 797 after adjustments				
Dependent Variable: DLOG(SPOTRATE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	1.191	0.032	37.000	0.0000
DLOG(SPOTRATE(-2))	-0.420	0.032	-13.048	0.0000
R-squared	0.755	Mean dependent var	0.001	
Adjusted R-squared	0.755	S.D. dependent var	0.019	
S.E. of regression	0.009	Akaike info criterion	-6.478	
Sum squared resid	0.071	Schwarz criterion	-6.466	
Log likelihood	2583.549	Durbin-Watson stat	1.985	

Notes: Results from ARIMA(2,1,0) estimation on Panamax spot rates as univariate setup

Table 24: VAR estimation for spot rate, Panamax (+1Q)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 797 after adjustments				
Dependent Variable: DLOG(SPOTRATE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	0.944	0.0039	24.508	0.000
DLOG(SPOTRATE(-2))	-0.232	0.034	-6.806	0.000
DLOG(FFARATE(-1))	0.132	0.011	11.580	0.000
DLOG(FFARATE(-2))	0.026	0.012	2.242	0.025
c	0.002	0.001	2.166	0.030
R2	0.792	Mean dependent var	0.001	
Adjusted R-squared	0.791	S.D. dependent var	0.0191	
S.E. of regression	0.009	Akaike info criterion	-6.634	
Sum squared resid	0.061	Schwarz criterion	-6.605	
Log likelihood	2648.797	Durbin-Watson stat	2.013	
F-stat	753.352			

Notes: VAR estimation results for Panamax using the spot rate and FFA+1Q. The table depicts the line-wise OLS results for the spot rate.

Table 25: VAR estimation for FFA rate, Panamax (+1Q)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 797 after adjustments				
Dependent Variable: DLOG(FFARATE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	0.049	0.134	0.361	0.3270
DLOG(SPOTRATE(-2))	-0.190	0.119	-1.596	0.0111
DLOG(FFARATE(-1))	0.255	0.039	6.402	0.0000
DLOG(FFARATE(-2))	0.0239	0.041	0.0579	0.0364
c	0.000	0.001	0.908	0.364
R-squared	0.080	Mean dependent var	0.001	
Adjusted R-squared	0.075	S.D. dependent var	0.032	
S.E. of regression	0.030	Akaike info criterion	-4.134	
Sum squared resid	0.738	Schwarz criterion	-4.105	
Log likelihood	1652.471	Durbin-Watson stat	2.012	
F-stat	17.126			

Notes: VAR estimation results for Panamax using the spot rate and FFA+1Q. The table depicts the line-wise OLS results for the FFA.

Table 26: VAR estimation for spot rate, Panamax (+2Q)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 752 after adjustments				
Dependent Variable: DLOG(SPOTRATE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	0.975	0.0.039	25.166	0.0000
DLOG(SPOTRATE(-2))	-0.255	0.0.035	-7.365	0.0000
DLOG(FFARATE(-1))	0.145	0.0.013	11.200	0.000
DLOG(FFARATE(-2))	0.025	0.0.014	1.866	0.062
c	0.000	0.003	0.333	0.740
R-squared	0.800	Mean dependent var	0.001	
Adjusted R-squared	0.799	S.D. dependent var	0.019	
S.E. of regression	0.009	Akaike info criterion	-6.665	
Sum squared resid	0.055	Schwarz criterion	-6.624	
Log likelihood	2510.975	Durbin-Watson stat	1.992	
F-stat	746.656			

Notes: VAR estimation results for Panamax using the spot rate and FFA+2Q. The table depicts the line-wise OLS results for the spot rate.

Table 27: VAR estimation for FFA rate, Panamax (+2Q)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 752 after adjustments				
Dependent Variable: DLOG(FFARATE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	0.164	0.120	1.368	0.171
DLOG(SPOTRATE(-2))	-0.315	0.107	-2.936	0.003
DLOG(FFARATE(-1))	0.264	0.040	6.574	0.000
DLOG(FFARATE(-2))	0.036	0.042	0.851	0.395
c	-0.006	0.001	-4.307	0.000
R-squared	0.108	Mean dependent var	0.001	
Adjusted R-squared	0.103	S.D. dependent var	0.028	
S.E. of regression	0.027	Akaike info criterion	-4.401	
Sum squared resid	0.532	Schwarz criterion	-4.371	
Log likelihood	1659.9428	Durbin-Watson stat	1.967	
F-stat	746.656			

Notes: VAR estimation results for Panamax using the spot rate and FFA+2Q. The table depicts the line-wise OLS results for the FFA+2q.

Table 28: VAR estimation for spot rate, Panamax (+1A)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 797 after adjustments				
Dependent Variable: DLOG(SPOTRATE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	1.031	0.037	30.089	0.000
DLOG(SPOTRATE(-2))	-0.349	0.034	-10.184	0.000
DLOG(FFARATE(-1))	0.069	0.012	5.668	0.000
DLOG(FFARATE(-2))	-0.001	0.012	-0.019	0.969
c	0.000	0.003	0.538	0.591
R-squared	0.765	Mean dependent var	0.0011	
Adjusted R-squared	0.763	S.D. dependent var	0.019	
S.E. of regression	0.009	Akaike info criterion	-6.511	
Sum squared resid	0.069	Schwarz criterion	-6.482	
Log likelihood	2599.660	Durbin-Watson stat	2.011	
F-stat	642.989			

Notes: VAR estimation results for Panamax using the spot rate and FFA+1A. The table depicts the line-wise OLS results for the spot rate.

Table 29: VAR estimation results for FFA rate, Panamax (+1A)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 797 after adjustments				
Dependent Variable: DLOG(FFARATE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	0.201	0.117	1.723	0.085
DLOG(SPOTRATE(-2))	-0.257	0.109	-2.345	0.019
DLOG(FFARATE(-1))	0.142	0.039	3.675	0.000
DLOG(FFARATE(-2))	-0.029	0.038	-0.752	0.452
c	0.001	0.001	1.742	0.082
R-squared	0.036	Mean dependent var	0.001	
Adjusted R-squared	0.031	S.D. dependent var	0.030	
S.E. of regression	0.030	Akaike info criterion	-4.192	
Sum squared resid	0.069	Schwarz criterion	-4.163	
Log likelihood	1675.656	Durbin-Watson stat	1.831	
F-stat	7.367			

Notes: VAR estimation results for Panamax using the spot rate and FFA+1A. The table depicts the line-wise OLS results for the FFA+1A.

Table 30: VAR estimation results for spot rates, Capesize (+1Q)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 787 after adjustments				
Dependent Variable: DLOG(SPOTRATE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	0.821	0.040	20.300	0.001
DLOG(SPOTRATE(-2))	-0.192	0.035	-5.485	0.000
DLOG(FFA KURS(-1))	0.146	0.017	8.211	0.000
DLOG(FFA KURS(-2))	0.047	0.018	2.618	0.327
c	0.000	0.000	0.392	0.401
R-squared	0.687	Mean dependent var	0.001	
Adjusted R-squared	0.685	S.D. dependent var	0.024	
S.E. of regression	0.013	Akaike info criterion	-4.111	
Sum squared resid	0.141	Schwarz criterion	-4.081	
Log likelihood	2275.503	Durbin-Watson stat	2.059	
F-stat	428.417			

Notes: VAR estimation results for Capesize using the spot rate and FFA+1Q. The table depicts the line-wise OLS results for the spot rate

Table 31: VAR estimation for FFA, Capesize (+1Q)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 787 after adjustments				
Dependent Variable: DLOG(FFARATE)				
DLOG(SPOTRATE(-1))	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	0.308	0.092	3.321	0.001
DLOG(SPOTRATE(-2))	-0.309	0.080	-3.833	0.000
DLOG(FFARATE(-1))	0.169	0.041	4.132	0.000
DLOG(FFARATE(-2))	0.040	0.0981	0.001	0.327
c	0.000	0.001	0.839	0.401
R-squared	0.080	Mean dependent var	0.001	
Adjusted R-squared	0.076	S.D. dependent var	0.032	
S.E. of regression	0.030	Akaike info criterion	-4.111	
Sum squared resid	0.746	Schwarz criterion	-4.081	
Log likelihood	1662.795	Durbin-Watson stat	2.018	
F-stat	17.179			

Notes: VAR estimation results for Capesize using the spot rate and FFA+1Q. The table depicts the line-wise OLS results for the FFA+1Q.

Table 32: VAR estimation for spot rates, Capesize (+2Q)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 766 after adjustments				
Dependent Variable: DLOG(SPOTRATE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	0.846	0.040	21.168	0.000
DLOG(SPOTRATE(-2))	-0.210	0.036	-5.938	0.000
DLOG(FFA KURS(-1))	0.161	0.021	7.461	0.000
DLOG(FFA KURS(-2))	0.042	0.021	1.962	0.050
c	0.000	0.000	0.513	0.608
R-squared	0.679	Mean dependent var	0.001	
Adjusted R-squared	0.677	S.D. dependent var	0.024	
S.E. of regression	0.013	Akaike info criterion	-5.768	
Sum squared resid	0.138	Schwarz criterion	-5.738	
Log likelihood	2214.309	Durbin-Watson stat	2.027	

Notes: VAR estimation results for Capesize using the spot rate and FFA+2Q. The table depicts the line-wise OLS results for the spot rate.

Table 33: VAR estimation for FFA, Capesize (+2Q)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 766 after adjustments				
Dependent Variable: DLOG(FFARATE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	0.227	0.074	3.030	0.002
DLOG(SPOTRATE(-2))	-0.282	-0.282	-4.247	0.000
DLOG(FFARATE(-1))	0.241	0.241	5.077	0.000
DLOG(FFARATE(-2))	0.064	0.030	1.589	0.112
c	-0.001	0.000	-1.568	0.117
R-squared	0.120	Mean dependent var	0.120	
Adjusted R-squared	0.116	S.D. dependent var	0.116	
S.E. of regression	0.025	Akaike info criterion	-4.516	
Sum squared resid	0.484	Schwarz criterion	-4.458	
Log likelihood	1734.437	Durbin-Watson stat	2.012	
F-stat	26.044			

Notes: VAR estimation results for Capesize using the spot rate and FFA+2Q. The table depicts the line-wise OLS results for the FFA+2Q.

Table 34: VAR estimation for spot rate, Capesize (+1A)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 797 after adjustments				
Dependent Variable: DLOG(SPOTRATE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	0.931	0.037	25.400	0.000
DLOG(SPOTRATE(-2))	-0.251	0.034	-7.329	0.000
DLOG(FFARATE(-1))	0.086	0.017	5.137	0.000
DLOG(FFARATE(-2))	0.043	0.017	2.577	0.010
c	0.000	0.001	0.520	0.682
R-squared	0.672	Mean dependent var	0.001	
Adjusted R-squared	0.670	S.D. dependent var	0.024	
S.E. of regression	0.013	Akaike info criterion	-5.732	
Sum squared resid	0.149	Schwarz criterion	-5.702	
Log likelihood	2289.109	Durbin-Watson stat	2.046	
F-stat	405.723			

Notes: VAR estimation results for Capesize using the spot rate and FFA+1A. The table depicts the line-wise OLS results for the spot rate.

Table 35: VAR estimation results for FFA, Capesize (+1A)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 797 after adjustments				
Dependent Variable: DLOG(FFARATE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	0.331	0.083	4.015	0.000
DLOG(SPOTRATE(-2))	-0.343	0.077	-4.446	0.000
DLOG(FFARATE(-1))	0.129	0.038	3.427	0.000
DLOG(FFARATE(-2))	0.041	0.038	1.111	0.266
c	0.001	0.001	0.301	
R-squared	0.062	Mean dependent var	0.001	
Adjusted R-squared	0.057	S.D. dependent var	0.032	
S.E. of regression	0.031	Akaike info criterion	-4.109	
Sum squared resid	0.757	Schwarz criterion	-4.079	
Log likelihood	1642.523	Durbin-Watson stat	1.909	
F-stat	13.059			

Notes: VAR estimation results for Capesize using the spot rate and FFA+1A. The table depicts the line-wise OLS results for the FFA+1A.

Table 36: VECM estimation results, Panamax (+1Q)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 797 after adjustments				
Cointegration Equation				
		coefficient	SE	t-value
spot		1.000		
FFA		-1.038	0.061	-16.786
c		0.349	0.638	0.549
Variable	DLOG(spot)	t-value	dlog(FFA)	t-value.
DLOG(spot(-1))	0.933	24.318	0.049	0.362
DLOG(spot(-2))	-0.229	-6.751	-0.187	-1.570
DLOG(FFA(-1))	0.123	10.871	0.257	6.338
DLOG(FFA(-2))	0.021	1.774	0.249	0.598
EC	-0.009	-3.32	0.001	0.104
R-squared	0.795		0.0789	
Adjusted R-squared	0.793		0.0740	
S.E. of regression	0.009		0.030	
Sum squared resid	0.598		0.739	
Log likelihood	2654.162		1652.061	
Mean dependent var	0.001		0.001	
S.D. dependent var	0.019		0.032	
Akaike info criterion	-6.648		-4.133	
Schwarz criterion	-6.618		-4.104	
F-stat	766.245		1652.061	

Notes: Estimation results for VECM for Panamax and $X_t = [S_t, FFA + 1Q_t]'$. the upper part denotes the cointegration vector and the lower part the short run adjustments including the error correction term (EC) for which the left part denotes the regression results for the spot regression and the right part the results of the FFA regression, including the respective t-value.

Table 37: VECM estimation results, Panamax (+2Q)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 797 after adjustments				
Cointegration Equation				
		coefficient	SE	t-value
spot		1.000		
FFA		-1.041	0.068	-15.037
c		0.0212	0699	0.303
Variable	DLOG(spot)	t-value	dlog(FFA)	t-value.
DLOG(spot(-1))	0.965	24.945	0.169	1.403
DLOG(spot(-2))	-0.247	-7.162	-0.316	-2.936
DLOG(FFA(-1))	0.140	10.761	0.269	6.628
DLOG(FFA(-2))	0.020	1.153	0.039	0.932
EC	-0.008	-2.945	0.005	0.604
R-squared	0.802		0.108	
Adjusted R-squared	0.801		0.103	
S.E. of regression	0.009		0.027	
Sum squared resid	0.055		0.533	
Log likelihood	2515.339		1659.168	
Mean dependent var	0.001		0.001	
S.D. dependent var	0.019		0.028	
Akaike info criterion	-6.676		-4.301	
Schwarz criterion	-6.646		-4.340	
F-stat	757.553		22.494	

Notes: Estimation results for VECM for Panamax and $X_t = [spot_t, FFA + 2Q_t]'$. the upper part denotes the co integration vector and the lower part the short run adjustments including the error correction term (EC) for which the left part denotes the regression results for the spot regression and the right part the results of the FFA regression, including the respective t-value.

Table 38: VECM estimation results, Panamax (+1A)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 797 after adjustments				
Dependent Variable: DLOG(SPOTRATE)				
Cointegration Equation				
		coefficient	SE	t-value
spot		1.000		
FFA		-1.05	0.102	-10.199
c		0.194	1.033	0.190
Variable	DLOG(spot)	t-value	dlog(FFA)	t-value.
DLOG(spot(-1))	1.095	29.873	0.210	1.792
DLOG(spot(-2))	-0.340	-9.939	-0.262	-2.393
DLOG(FFA(-1))	0.0664	5.442	0.148	3.3783
DLOG(FFA(-2))	-0.004	-0.292	-0.024	-0.630
EC	-0.006	-2.526	0.006	0.945
R-squared	0.766		0.036	
Adjusted R-squared	0.765		0.031	
S.E. of regression	0.009		0.029	
Sum squared resid	0.068		0.696	
Log likelihood	2602.452		1675.559	
Mean dependent var	0.002		0.002	
S.D. dependent var	0.019		0.030	
Akaike info criterion	-6.519		-4.192	
Schwarz criterion	-6.489		-4.162	
F-stat	649.451		1675.559	

Notes: Estimation results for VECM for Panamax and $X_t = [spot_t, FFA + 1A_t]'$. the upper part denotes the co integration vector and the lower part the short run adjustments including the error correction term (EC) for which the left part denotes the regression results for the spot regression and the right part the results of the FFA regression, including the respective t-value.

Table 39: VECM estimation results, Capesize (+1Q)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 797 after adjustments				
Cointegration Equation				
		coefficient	SE	t-value
spot		1.000		
FFA		-1.057	0.049	-21.581
c		0.542	0.536	1.010
Variable	DLOG(spot)	t-value	dlog(FFA)	t-value.
DLOG(spot(-1))	0.818	20.256	0.302	3.250
DLOG(spot(-2))	-0.188	-5.439	-0.304	-3.787
DLOG(FFA(-1))	0.130	7.198	0.160	3.825
DLOG(FFA(-2))	0.034	1.883	0.033	0.790
EC	-0.019	-4.326	-0.011	-1.100
R-squared	0.693		0.081	
Adjusted R-squared	0.139		0.077	
S.E. of regression	0.013		0.030	
Sum squared resid	0.598		0.745	
Log likelihood	2284.731		1623.048	
Mean dependent var	0.001		0.001	
S.D. dependent var	0.024		0.032	
Akaike info criterion	-5.793		-4.112	
Schwarz criterion	-6.763		-4.082	
F-stat	445.221		1623.048	

Notes: Estimation results for VECM for Capesize and $X_t = [spot_t, FFA + 1Q_t]'$. the upper part denotes the co integration vector and the lower part the short run adjustments including the error correction term (EC) for which the left part denotes the regression results for the spot regression and the right part the results of the FFA regression, including the respective t-value.

Table 40: VECM estimation results, Capesize (+2Q)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 766 after adjustments				
Cointegration Equation				
		coefficient	SE	t-value
spot		1.000		
FFA		-1.060	0.082	-12.964
c		0.483	0.888	0.545
Variable	DLOG(spot)	t-value	dlog(FFA)	t-value.
DLOG(spot(-1))	0.083	21.114	0.226	3.015
DLOG(spot(-2))	-0.201	-6.697	-0.279	-4.187
DLOG(FFA(-1))	0.153	7.068	0.243	5.5952
DLOG(FFA(-2))	0.034	1.568	0.065	1.582
EC	-0.011	-2.945	-0.000	-0.063
R-squared	0.683		0.118	
Adjusted R-squared	0.681		0.114	
S.E. of regression	0.013		0.025	
Sum squared resid	0.136		0.485	
Log likelihood	2219.143		1733.753	
Mean dependent var	0.002		0.001	
S.D. dependent var	0.023		0.026	
Akaike info criterion	-5.781		-4.513	
Schwarz criterion	-5.750		-4.483	
F-stat	410.332		25.658	

Notes: Estimation results for VECM for Capesize and $X_t = [spot_t, FFA + 2q_t]'$. the upper part denotes the co integration vector and the lower part the short run adjustments including the error correction term (EC) for which the left part denotes the regression results for the spot regression and the right part the results of the FFA regression, including the respective t-value.

Table 41: VECM estimation results, Capesize (+1A)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 797 after adjustments				
Cointegration Equation				
		coefficient	SE	t-value
spot		1.000		
FFA		-1.063	0.016	-10.015
c		0.392	0.344	1.135
Variable	DLOG(spot)	t-value	dlog(FFA)	t-value.
DLOG(spot(-1))	0.925	25.290	0.336	4.060
DLOG(spot(-2))	-0.243	-7.098	-0.346	-4.472
DLOG(FFA(-1))	0.080	4.808	0.134	3.534
DLOG(FFA(-2))	0.037	2.237	0.046	1.225
EC	-0.008	-2.526	0.006	0.850
R-squared	0.675		0.061	
Adjusted R-squared	0.673		0.056	
S.E. of regression	0.013		0.030	
Sum squared resid	0.147		0.757	
Log likelihood	2293.155		1642.349	
Mean dependent var	0.001		0.001	
S.D. dependent var	0.024		0.032	
Akaike info criterion	-5.741		-4.108	
Schwarz criterion	-5.712		-4.079	
F-stat	411.883		12.967	

Notes: Estimation results for VECM for Capesize and $X_t = [spot_t, FFA + 1A_t]'$. the upper part denotes the co integration vector and the lower part the short run adjustments including the error correction term (EC) for which the left part denotes the regression results for the spot regression and the right part the results of the FFA regression, including the respective t-value.

Table 42: ARIMA TC, Panamax (+6M)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 797 after adjustments				
Dependent Variable: DLOG(SPOTRATE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	1.206	0.034	35.055	0.000
DLOG(SPOTRATE(-2))	-0.428	0.037	-11.395	0.000
DLOG TC 6M(-1)	-0.000	0.006	-0.062	0.949
R-squared	0.769	Mean dependent var	0.000	
Adjusted R-squared	0.769	S.D. dependent var	0.019	
S.E. of regression	0.009	Akaike info criterion	-6.532	
Sum squared resid	0.061	Schwarz criterion	-6.513	
Log likelihood	2380.862	Durbin-Watson stat	1.992	

Notes: Results for ARIMA-TC(2,1,0;1) estimation on Panamax +6 Month.

Table 43: ARIMA TC, Panamax (+12M)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 797 after adjustments				
Dependent Variable: DLOG(SPOTRATE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	1.201453	0.034284	35.04413	0.0000
DLOG(SPOTRATE(-2))	-0.419791	0.036968	-11.35565	0.0000
DLOG TC 1A(-1)	-0.005088	0.008137	-0.625267	0.5320
R-squared	0.770080	Mean dependent var	0.000987	
Adjusted R-squared	0.769445	S.D. dependent var	0.019179	
S.E. of regression	0.009209	Akaike info criterion	-6.533122	
Sum squared resid	0.061486	Schwarz criterion	-6.514206	
Log likelihood	2381.056	Durbin-Watson stat	1.983124	

Notes: Results for ARIMA-TC(2,1,0;1) estimation on Panamax +12 Months.

Table 44: ARIMA TC, Capesize (+6M)

Method: Least Squares				
Sample (adjusted): 10/06/2004 11/23/2007				
Included observations: 728 after adjustments				
Dependent Variable: DLOG(SPOTRATE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	1.036	0.034	29.786	0.000
DLOG(SPOTRATE(-2))	-0.319	0.036	-8.807	0.000
DLOG TC 6M(-1)	0.005	0.007	0.789	0.430
R-squared	0.666	Mean dependent var	0.001	
Adjusted R-squared	0.665	S.D. dependent var	0.026	
S.E. of regression	0.013	Akaike info criterion	-5.701	
Sum squared resid	0.142	Schwarz criterion	-5.682	
Log likelihood	2100.994	Durbin-Watson stat	2.073	

Notes: Results for ARIMA-TC(2,1,0;1) estimation, Capesize +6 Months.

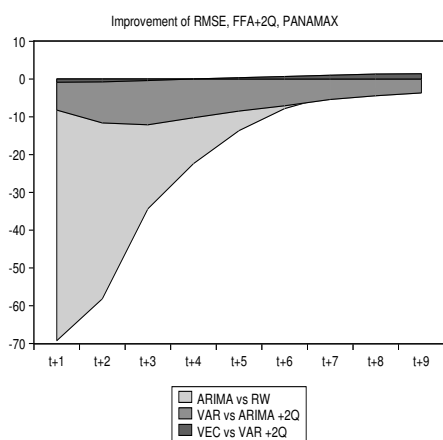
Table 45: ARIMA TC, Capesize (+12M)

Sample (adjusted): 10/06/2004 12/12/2007				
Included observations: 797 after adjustments				
Dependent Variable: DLOG(SPOTRATE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(SPOTRATE(-1))	1.034	0.034838	29.691	0.000
DLOG(SPOTRATE(-2))	-0.311	0.035928	-8.674	0.000
DLOG TC 1A(-1)	0.000	0.008404	0.081	0.935
R-squared	0.665	Mean dependent var	0.001	
Adjusted R-squared	0.664	S.D. dependent var	0.024	
S.E. of regression	0.013	Akaike info criterion	-5.700	
Sum squared resid	0.142	Schwarz criterion	-5.681	
Log likelihood	2100.685	Durbin-Watson stat	2.066	

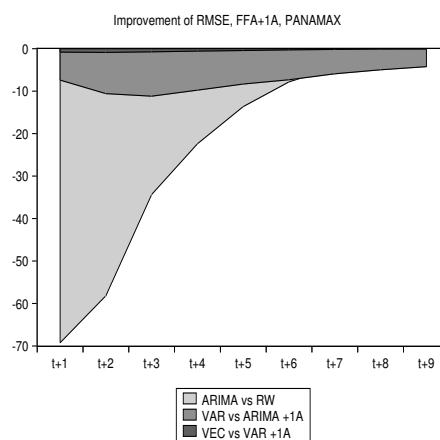
Notes: Results for ARIMA-TC(2,1,0;1) estimation, Capesize +12 Months.

Figure 51: Improvement of RMSE compared to random walk

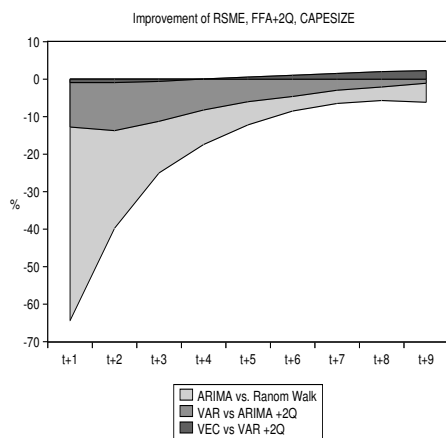
(a) Panamax, FFA+2Q



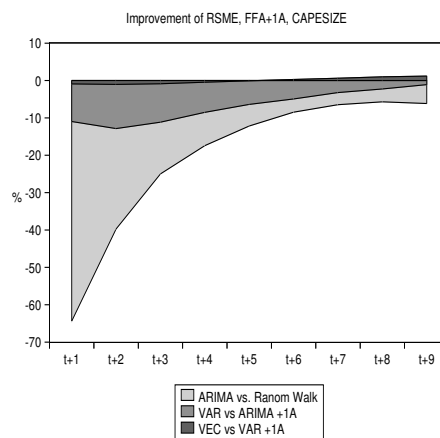
(b) Panamax, FFA+2A



(c) Capesize, FFA+2Q



(d) Capesize, FFA+1A



Notes: Improvement in RMSE by switching from the random walk model to the respective models on Capesize and Panamax contracts.

Acronyms

- FFA: Forward Freight Agreement. Principal-to-principal forward agreement, OTC traded. Payment after settlement between spot rate at T and future rate at $F_{t,T}$.
- Voyage Charter: Under this Agreement, ship owners pay all Voyage Costs, not the charterer
- Voyage Costs: All Costs associated with the trip; port costs, fuel costs, crew and canal costs
- TC: Time Charter Contract. Charterer receives a ship for specified time and route but pays all Voyage costs
- Time Charter: See TC
- BFI: Baltic Freight Index
- BPI: Baltic Panamax Index
- OTC: Over-the-Counter trade
- BHSI: Baltic Handysize Index
- BSI: Baltic Supramax Index
- BCI: Baltic Capesize Index

- Panamax: Largest ships capable of traversing the Panama Canal, 60,000-100,000 DWT
- Capesize: Ships not capable of traversing neither the Sues- nor the Panama Canal, larger than 100,000 DWT
- Handysize: Bulk carrier ships, designed to transport bulk goods, 10,000-40,000 DWT
- Period Charter: Same as time charter, TC
- Spot Rate: Shipping rate at t , freely negotiated for carriage of specified quantity of goods at t
- Cost of Carry: Relationship between Forward price at t , T and future Spot rate at T , $F_{t,T} = E_t \beta S_T$
- Dry Bulk: Unpacked goods in large quantities
- UEH: Unbiased Expectations Hypothesis. In efficient markets, forward price is equal to discounted expected future Spot rate
- PEHTS: Pure EHTS In addition to EHTS, all alternative investments are also taken into account for long-run contracts. The risk premium is constant and zero. If this is the case, markets are efficient
- EHTS: Weak version of PEHTS; risk premium is constant, but not equal to zero
- 4TC: Weighted average of the busiest four Time Charter routes of the respective index
- DWT: Dead weight tonnage. Vessel's weight

7 Conclusion

In my dissertation, I present five essays on the impact of fiscal policy on the macroeconomy while investigating the interaction with financial assets.

In chapter 2, I create a novel measure to capture future military spending changes based on stock market returns of military contractors, based on Fisher and Peters (2010). By relying on defense spending and the stock market measure, I solve the endogeneity problem and foresight issue when estimating fiscal VARs. However, the externally specified shock of Fisher and Peters (2010) has some drawbacks so that I modify the shock to be free of known market anomalies as shown in Fama and French (1993). The spending shock is then measured as abnormal returns in a SLB market model regression.

I then estimate fiscal VARs and derive the fiscal multiplier. When using my new shock and ordering it last in a recursive VAR setup, I can show that the fiscal multiplier takes a value of 1.2. However, when extending the sample to include financial crisis data, the impact of military spending decreases and is only half as strong as in the limited sample.

The next chapter investigates, the second side of fiscal policy, taxes. In a vector autoregression setup, the impact of tax cuts and TFP innovations as suggested by Romer and Romer (2010) and Basu et al. (2006) on consumer credit is investigated. The paper compares estimates from exogenous VARs and standard VARs since the strict exogeneity of these two shocks is debatable. Granger causality tests suggest that both shocks are Granger caused by the other variables in the system. The paper hence suggests estimating the less restrictive method, a standard VAR rather than an exogenous VAR.

Both shocks trigger an expansion in the economy. The paper finds a positive comovement between output, private consumption, hours, and consumer credit conditional on both shocks. Such comovement refutes standard consumption arguments. Thus, households do not save in good times to buffer negative shocks. The initiated boom in the economy after the two shocks is thus partly debt-financed.

Because that chapter provides insight into the interaction between tax cuts, TFP shocks and unsecured consumer credit, the next chapter estimates the impact of TFP innovations and tax cuts on total household debt, which includes also collateralized debt. That paper is able to find the same comovement as in the case of consumer credit. Since this result is robust to a number of modifications, a DSGE model with financial frictions is constructed bring the model to the data and hence to match empirical impulse responses. The resulting parameters are in line with earlier estimates so that parts of the households in the economy are borrowing-constrained.

The model is capable of matching the hump-shaped responses in all variables in the system.

In the fourth essay, I investigate the interaction between house price shocks and fiscal policy rules, if the economy is distant or at the Zero Lower Bound on interest rates (ZLB). A shock to house prices has contributed to starting the financial crisis, as Iacoviello (2015) notes. I can show that house price shocks which I model as devaluation of the stock of housing have a severe negative impact on the economy since borrowing households are forced to quickly deleverage by selling off property and decrease consumption.

In response to the crisis, government debt has surged in many countries. To reduce government debt, taxes can be increased, or spending can be cut. I can show that in times when the economy is distant to the ZLB, it is irrelevant if taxes or spending is used to reduce government debt. However, if the economy is at the ZLB, cutting spending amplifies the recession in contrast to increasing taxes when a deleveraging shock hits. Therefore, in a situation with high government debt, cutting spending has a severe negative impact on output compared to increasing taxes.

Finally, in contrast to the assumption in the chapter on the fiscal multiplier, the last chapter investigates how asset returns are influenced by inefficient markets. I choose the market for shipping goods across oceans which has unique statistical characteristics compared to other asset classes with high data frequency and quality.

Due to high-frequency data and availability, the paper then takes advantage of inefficient asset markets in the shipping sector and compares forecasting models in terms of RMSE. As it turns out, vector autoregression outperforms all other models. With this information, a trading scheme is created which would have outperformed the market, even after controlling for transaction costs.

In conclusion, these five chapters contribute to shed light on the interaction between financial assets, fiscal policy and the macroeconomy using empirical models and also theoretical ones.

8 References

References

- Roar Adland and Kevin P. B. Cullinane. A time-varying risk premium in the term structure of bulk shipping freight rates. *Journal of Transport Economics and Policy*, 39(2):191–208, 2005.
- Hirotsugu Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.
- David Altig, Lawrence J Christiano, Martin Eichenbaum, and Jesper Linde. Firm-specific capital, nominal rigidities and the business cycle. *Review of Economic Dynamics*, 14(2):225–247, 2011.
- J Andrés, JE Boscá, and J Ferri. Household debt and fiscal multipliers. *Documento de Trabajo*, pages 01–48, 2015.
- Alan J. Auerbach and Yuriy Gorodnichenko. Fiscal multipliers in recession and expansion. In *Fiscal Policy after the Financial crisis*, pages 63–98. University of Chicago press, 2012.
- Rolf W. Banz. The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1):3–18, 1981.
- Robert J. Barro and Charles J. Redlick. Macroeconomic effects from government purchases and taxes. *Quarterly Journal of Economics*, 126(1):51–102, 2011.
- Susanto Basu, John G. Fernald, and Miles S. Kimball. Are technology improvements contractionary? *American Economic Review*, 96(5):1418–1448, 2006.
- Roy Batchelor, Amir Alizadeh, and Ilias Visvikis. Forecasting spot and forward prices in the international freight market. *International Journal of Forecasting*, 23(1):101–114, 2007.
- Paul Beaudry and Franck Portier. Stock prices, news and economic fluctuations. *American Economic Review*, 96(4):1293–1307, 2006.
- Ben S. Bernanke, Mark Gertler, and Simon Gilchrist. Chapter 21 the financial accelerator in a quantitative business cycle framework. volume 1, Part C of *Handbook of Macroeconomics*, pages 1341 – 1393. Elsevier, 1999.

- Fischer Black. Capital market equilibrium with restricted borrowing. *Journal of Business*, 3(45):444–455, 1972.
- Olivier J. Blanchard and Roberto Perotti. An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *Quarterly Journal of Economics*, 117(4):1329–1368, 2002.
- Guillermo A. Calvo. Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, 12(3):383–398, 1983.
- Alessandro Calza, Tommaso Monacelli, and Livio Stracca. Housing finance and monetary policy. *Journal of the European Economic Association*, 11(s1):101–122, 2013.
- Marm M. Carhart. On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57–82, 1997.
- Lawrence J. Christiano, Martin Eichenbaum, and Robert Vigfusson. The response of hours to a technology shock: evidence based on direct measures of technology. *Journal of the European Economic Association*, 2(2-3):381–395, 2004.
- Lawrence J Christiano, Martin Eichenbaum, and Charles L Evans. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of political Economy*, 113(1):1–45, 2005.
- John H. Cochrane. Time series for macroeconomics and finance. *Manuscript, University of Chicago*, 2005.
- Timothy Cogley and James M Nason. Output dynamics in real-business-cycle models. *The American Economic Review*, 85(3):492–511, 1995.
- Giancarlo Corsetti, Andre Meier, and Gernot J Müller. What determines government spending multipliers? *Economic Policy*, 27(72):521–565, 2012.
- Kevin P. B. Cullinane. A short-term adaptive forecasting model for biffex speculation: a box-jenkins approach. *Maritime Policy and Management*, 19(2):91–114, 1992.
- Kevin. P. B. Cullinane, Kevin J. Mason, and Mark Cape. A comparison of models for forecasting the baltic freight index: Box-jenkins revisited. *International Journal of Maritime Economics*, 1(2):15–39, 1999.

- Stefano DellaVigna and Eliana La Ferrara. Detecting illegal arms trade. *American Economic Journal: Economic Policy*, 2(4):26–57, 2010.
- David A. Dickey and Wayne A. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a):427–431, 1979.
- Elroy Dimson and Paul R Marsh. The stability of uk risk measures and the problem of thin trading. *The Journal of Finance*, 38(3):753–783, 1983.
- Mardi Dungey and Renee Fry. The identification of fiscal and monetary policy in a structural var. *Economic Modeling*, 26(6):1147–1160, 2009.
- Martial Dupaigne, Patrick Fève, and Julien Matheron. Avoiding pitfalls in using structural vars to estimate economic models. *Review of Economic Dynamics*, 10(2):238–255, 2007.
- Essi Eerola and Niku Määttänen. Borrowing constraints and house price dynamics: the case of large shocks. *Studies in Nonlinear Dynamics & Econometrics*, 16(3):1–31, 2012.
- Gauti B. Eggertsson and Paul Krugman. Debt, deleveraging, and the liquidity trap: A fisher-minsky-koo approach. *The Quarterly Journal of Economics*, 127(3):1469–1513, 2012.
- Charles Engel. The forward discount anomaly and the risk premium: A survey of recent evidence. *Journal of empirical finance*, 3(2):123–192, 1996.
- Robert F. Engle and Clive W. J. Granger. Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, pages 251–276, 1987.
- Eugene F. Fama and Kenneth R. French. The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465, 1992.
- Eugene F. Fama and Kenneth R. French. Common risk factors in the returns on stocks and bonds. *The Journal of Financial Economics*, 1993(33):3–56, 1993.
- Eugene F. Fama and Kenneth R. French. Industry costs of equity. *Journal of Financial Economics*, 1997(43):153–193, 1998.
- Eugene F. Fama and Kenneth R. French. Dissecting anomalies. *The Journal of Finance*, 63(4):1653–1678, 2008.

- Eugene F. Fama and Kenneth R. French. Size, value, and momentum in international stock returns. *Journal of Financial Economics*, 105(3):457–472, 2012.
- Eugene F Fama and Kenneth R French. A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22, 2015.
- Aantonia Fatas and Ilian Mihov. The effects of fiscal policy on consumption and employment: Theory and evidence. *C.E.P.R. Discussion Papers*, (2760):1–32, 2001.
- Carlo Favero and Francesco Giavazzi. Measuring tax multipliers: The narrative method in fiscal vars. *American Economic Journal: Economic Policy*, 4(2):69–94, 2012.
- Jesús Fernández-Villaverde, Pablo A Guerrón-Quintana, Keith Kuester, and Juan Rubio-Ramírez. Fiscal volatility shocks and economic activity. *National Bureau of Economic Research No. w17317*, 2011.
- Jonas D. M. Fisher. The dynamic effects of neutral and investment-specific technology shocks. *Journal of political Economy*, 114(3):413–451, 2006.
- Jonas D.M. Fisher and Ryan Peters. Using stock returns to identify government spending shocks. *Economic Journal*, 120(544):414–436, 2010.
- Neville Francis and Valerie A. Ramey. Measures of per capita hours and their implications for the technology-hours debate. *Journal of Money, Credit and Banking.*, 41(6):1071–1097, 2009.
- Jeffrey Frankel and Jumana Poonawala. The forward market in emerging currencies: Less biased than in major currencies. *Journal of International Money and Finance*, 29(3):585–598, 2010.
- William T. Gavin, Benjamin D. Keen, Alexander W. Richter, and Nathaniel A. Throckmorton. The zero lower bound, the dual mandate, and unconventional dynamics. *Journal of Economic Dynamics and Control*, 55:14–38, 2015.
- Daniel Glen, M. Owen, and R. Van der Meer. Spot and time charter rates for tankers, 1970-1977. *Journal of Transport Economics and Policy*, 1:45–58, 1981.
- Charles Hale and Andrew Vanags. Spot and period rates in the dry bulk market: Some tests for the period 1980-1986. *Journal of Transport Economics and Policy*, 23:281–291, 1989.

- Alastair R Hall, Atsushi Inoue, James M Nason, and Barbara Rossi. Information criteria for impulse response function matching estimation of dsge models. *Journal of Econometrics*, 170(2):499–518, 2012.
- Robert Hodrick. *The empirical evidence on the efficiency of forward and futures foreign exchange markets*, volume 1. Routledge, 2014.
- R. Glenn Hubbard. Capital market imperfections and investment. *Journal of Economic Literature*, 36(1):193–225, 1998.
- John Hull. *Options, Futures, and Other Derivatives*. Prentice-Hall: Englewood Cliffs, NJ, 2006.
- Matteo Iacoviello. House prices, borrowing constraints, and monetary policy in the business cycle. *The American Economic Review*, 95(3):739–764, 2005.
- Matteo Iacoviello. Housing in dsge models: Findings and new directions. In *Housing Markets in Europe*, pages 3–16. Springer, 2010.
- Matteo Iacoviello. Housing wealth and consumption. *FRB International Finance Discussion Paper*, (1027), 2011.
- Matteo Iacoviello. Financial business cycles. *Review of Economic Dynamics*, 18(1): 140–163, 2015.
- Matteo Iacoviello and Luca Guerrieri. Occbin: A toolkit for solving dynamic models with occasionally binding constraints easily. *Journal of Monetary Economics*, 70 (2015):22–38, 2015.
- Matteo Iacoviello and Stefano Neri. Housing market spillovers: Evidence from an estimated dsge model. *American Economic Journal: Macroeconomics*, 2(2):125–64, 2010.
- Peter N. Ireland. A small, structural, quarterly model for monetary policy evaluation. In *Carnegie-Rochester Conference Series on Public Policy*, volume 47, pages 83–108. Elsevier, 1997.
- Tullio Jappelli. Who is credit constrained in the u.s. economy? *The Quarterly Journal of Economics*, 105(1):219–34, 1990.
- Alejandro Justiniano, Giorgio E. Primiceri, and Andrea Tambalotti. Investment shocks and business cycles. *Journal of Monetary Economics*, 57(2):132–145, 2010.

- Alejandro Justiniano, Giorgio E. Primiceri, and Andrea Tambalotti. Household leveraging and deleveraging. *Review of Economic Dynamics*, 18(1):3–20, 2015.
- Myrto Kalouptsi. Time to build and fluctuations in bulk shipping. *The American Economic Review*, 104(2):564–608, 2014.
- Manolis Kavussanos, Ilias Visvikis, and David Menachof. The unbiasedness hypothesis in the freight forward market: Evidence from cointegration tests. *Review of Derivatives Research*, 7(3):241–266, 2005.
- Manolis G. Kavussanos and Amir H. Alizadeh-M. The expectations hypothesis of the term structure and risk premiums in dry bulk shipping freight markets. *Journal of Transport Economics and Policy*, 36:267–304, 2002.
- Manolis G. Kavussanos and Nikos K. Nomikos. Price discovery, causality and forecasting in the freight futures market. *Review of Derivatives Research*, 6(3):203–230, 2003.
- Manolis G. Kavussanos and Ilias D. Visvikis. Market interactions in returns and volatilities between spot and forward shipping freight markets. *Journal of Banking and Finance*, 28(8):2015–2049, 2004.
- Patrick J Kehoe. How to advance theory with structural vars: Use the sims-cogley-nason approach. *National Bureau of Economic Research No. w12575*, 2006.
- Michael T. Kiley. Habit persistence, nonseparability between consumption and leisure, or rule-of-thumb consumers: Which accounts for the predictability of consumption growth? *The Review of Economics and Statistics*, 92(3):679–683, 2010.
- Markus Kirchner, Jacopo Cimadomo, and Sebastian Hauptmeier. Transmission of government spending shocks in the euro area: Time variation and driving forces. *ECB Working Paper*, 2010.
- Nobuhiro Kiyotaki and John Moore. Credit Cycles. *Journal of Political Economy*, 105(2):211–48, April 1997.
- Finn E. Kydland, Peter Rupert, and Roman Sustek. Housing dynamics over the business cycle. *National Bureau of Economic Research No. w18432*, 2012.
- Edward E Leamer. Housing is the business cycle. *National Bureau of Economic Research No. w13428*, 2007.

- Eric M. Leeper, Andrew W. Richter, and Todd B. Walker. Quantitative effects of fiscal foresight. *American Economic Journal: Economic Policy*, 4(2):115–44, 2012.
- Eric M. Leeper, Todd B Walker, and Shu-Chun Susan Yang. Fiscal foresight and information flows. *Econometrica*, 81(3):1115–1145, 2013.
- John Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47(1): 13–37, 1965.
- Sydney Ludvigson. Consumption and credit: a model of time-varying liquidity constraints. *Review of Economics and Statistics*, 81(3):434–447, 1999.
- Karel Mertens and Morten O Ravn. Understanding the aggregate effects of anticipated and unanticipated tax policy shocks. *Review of Economic Dynamics*, 14 (1):27–54, 2011.
- Karel Mertens and Morten O. Ravn. The dynamic effects of personal and corporate income tax changes in the united states. *American Economic Review*, 103(4): 1212–47, 2013.
- Atif Mian and Amir Sufi. The great recession: Lessons from microeconomic data. *The American Economic Review*, 100(2):51–56, 2010.
- Tommaso Monacelli. New keynesian models, durable goods, and collateral constraints. *Journal of Monetary Economics*, 56(2):242–254, 2009.
- Andrew Mountford and Harald. Uhlig. What are the effects of fiscal policy shocks? *Journal of Applied Econometrics*, 24(6):960–992, 2009.
- Makoto Nakajima and José-Víctor Ríos-Rull. Credit, bankruptcy, and aggregate fluctuations. *National Bureau of Economic Research No. w20617*, 2014.
- Robert Novy-Marx. The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1):1–28, 2013.
- Michael T. Owyang, Valerie A. Ramey, and Sarah Zubairy. Are government spending multipliers greater during periods of slack? evidence from 20th century historical data. *National Bureau of Economic Research No. w18769*, 2013.
- Evi P. Pappa. The effects of fiscal shocks on employment and the real wage. *International Economic Review*, 50(1):217–244, 02 2007.

- Roberto Perotti. Estimating the effects of fiscal policy in oecd countries. *ECB Working Papers*, 1(168):33–55, 2002.
- Roberto Perotti. Expectations and fiscal policy: an empirical investigation. *Growth and Sustainability Policies for Europe Working Paper*, 19, 2011.
- Peter C. B. Phillips and Pierre Perron. Testing for a unit root in time series regression. *Biometrika*, 75(2):335–346, 1988.
- Valaery A. Ramey. Identifying government spending shocks: It’s all in the timing. *The Quarterly Journal of Economics*, 126(1):1–50, 2011a.
- Valerie A. Ramey. Can government purchases stimulate the economy? *Prepared for the Journal of Economic Literature Forum on the Multiplier*, pages 1–28, 2011b.
- Valerie. A. Ramey. Can government purchases stimulate the economy? *Prepared for the National Bureau of Economic Research Fiscal Policy After the Financial Crisis*, pages 1–51, 2012.
- Valerie A. Ramey and Matthew D. Shapiro. Costly capital reallocation and the effects of government spending. *Carnegie-Rochester Conference Series on Public Policy*, 1(48):145–194, 1998.
- Christina D. Romer and David H. Romer. The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *American Economic Review*, 100(3):763–801, 2010.
- Julio J. Rotemberg. Sticky prices in the united states. *The Journal of Political Economy*, pages 1187–1211, 1982.
- Julio J. Rotemberg and Michael Woodford. Oligopolistic pricing and the effects of aggregate demand on economic activity. *Journal of Political Economy*, 100(6): 1153–1297, 1992.
- Stephanie Schmitt-Grohe and Martin Uribe. Optimal fiscal and monetary policy under sticky prices. *Journal of Economic Theory*, 114(2):198 – 230, 2004.
- Gideon Schwarz. Estimating the dimension of a model. *The annals of statistics*, 6 (2):461–464, 1978.
- Daniel M. Shapiro and Lorne N. Switzer. War and peace: The reaction of defense stocks. *Journal of Applied Business Research*, 15(3):21–36, 2011.

- William F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, 1964.
- William F. Sharpe. Mutual fund performance. *Journal of business*, pages 119–138, 1966.
- Christopher A. Sims. *An autoregressive index model for the U.S. 1948-1975 in: J. Kmenta and J. B. Ramsey. Large-scale Macro-Econometric Models*, Amsterdam, 1981.
- Martin Stopford. *Maritime Economics*. Routledge: London, 1997.
- SP Strandenes. Price determination in the time charter and second hand markets. *Center for Applied Research, Norwegian School of Economics and Business Administration, Working Paper MU, 6*, 1984.
- Leonard J. Tashman. Market structures, freight rates and assets in bulk shipping. *Norwegian School of Economics and Business Administration, Bergen*, 1996.
- Leonard J. Tashman. Out-of-sample tests of forecasting accuracy: an analysis and review. *International Journal of Forecasting*, 16(4):437–450, 2000.
- Mathias Trabandt and Harald Uhlig. The laffer curve revisited. *Journal of Monetary Economics*, 58(4):305–327, 2011.
- Harald Uhlig. What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics*, 52(2):381–419, 2005.
- Albert W. Veenstra. The term structure of ocean freight rates. *Maritime Policy and Management*, 26(3):279–293, 1999.
- Albert Willem Veenstra and Philip Hans Franses. A co-integration approach to forecasting freight rates in the dry bulk shipping sector. *Transportation Research Part A: policy and practice*, 31(6):447–458, 1997.
- Mark W. Watson. *Vector autoregressions and cointegration. In: Engle, R. Handbook of Econometrics*. Elsevier, Amsterdam, 1986.
- Graham Wright. Spot and period rates in the wet bulk shipping market: testing for long-run parity. *Journal of Transport Economics and Policy*, pages 291–299, 2000.

Stephen P. Zeldes. Consumption and liquidity constraints: An empirical investigation. *Journal of Political Economy*, 97(2):305–346, 1989.

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit selbständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt entnommenen Gedanken sind als solche kenntlich gemacht.

Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch nicht veröffentlicht.

Dortmund, 20.10.2015