Essays in Applied Macroeconomics

Andrea Giovanni Gazzani

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Florence, 19 May 2017

European University Institute
Department of Economics

Essays in Applied Macroeconomics

Andrea Giovanni Gazzani

Thesis submitted for assessment with a view to obtaining the degree of
Doctor of Economics of the European University Institute

Examining Board

Prof. Evi Pappa, EUI, Supervisor
Prof. Alessia Campolmi, EUI & University of Verona
Prof. Luca Gambetti, Universitat Autonoma de Barcelona
Dr. Matteo Iacoviello, Federal Reserve Board

European University Institute

DEPARTMENT OF ECONOMICS

Researcher declaration to accompany the submission of written work

I, Andrea Giovanni Gazzani, certify that I am the author of the work *Essays in Applied Macroeconomics.* I have presented for examination for the PhD thesis at the European University Institute. I also certify that this is solely my own original work, other than where I have clearly indicated, in this declaration and in the thesis, that it is the work of others.

I warrant that I have obtained all the permissions required for using any material from other copyrighted publications.

I certify that this work complies with the *Code of Ethics in Academic Research* issued by the European University Institute (IUE 332/2/10 (CA 297).

Statement of inclusion of previous work (if applicable):

I confirm that chapter 2 was jointly co-authored with Alejandro Vicondoa and I contributed for 50% of the work.

I confirm that chapter 3 was jointly co-authored with Alejandro Vicondoa and I contributed for 50% of the work.

Signature and Date:

Monday, March the 27th, 2017

Abstract

This thesis studies the interaction between the real economy and assets like housing and bonds, and provide a new methodology to assess more accurately the spillovers from financial markets to the real economy.

The first chapter analyses of the role of expectations of future fundamentals in the housing market and their macroeconomic implications. News represent the component of expectations that proves to be correct in the future. Noise constitutes the component of expectations that does not materialize in the future. I find that fundamentals in the housing market are aligned with the real economy and that news shocks are the dominant driver of the housing market in the long run. However, the bulk of fluctuations in housing prices at high-medium frequencies is generated by noise. Notably, the latest housing cycle of the 2000s is entirely driven by expectations unrelated with fundamentals.

The second chapter, jointly written with Alejandro Vicondoa, develops a novel methodology, called Bridge Proxy-SVAR, to study the relationship between time series sampled at different frequencies. Instead of using a joint system, we rely on two systems at different frequencies and bridge them through an instrumental variable approach. We carry out identification at the highest available frequency and study the responses of the macroeconomic aggregates in a second stage. Our analytical, simulation and empirical results show that the Bridge Proxy-SVAR significantly mitigates temporal aggregation biases and it is particularly appealing to study the financial spillovers to the real economy.

In fact, in the third chapter, jointly written with Alejandro Vicondoa, we provide novel evidence on the large macroeconomic spillovers from changes in the liquidity of bonds. In particular, we analyze Italian sovereigns and find that liquidity shocks, orthogonal to changes in default risk, generate strong recessionary effect. Liquidity and default risk affect the real economy through different channels. By analyzing survey data, we find that liquidity shocks, differently from spikes in yields, do no lead to an increase in the rate requested by banks for loans. On the other hand, banks make their deadlines tighter and reduce the amount available for loan because they report problems with the liquidity and asset position.

To all those who fall to the bottom, because they can rise renewed.

To all those who can see beyond and are willing to give a second chance.

Acknowledgements

My gratitude goes to my supervisor Evi Pappa for her support and guidance during the writing of this thesis and the job market. My research has greatly benefited from her supervision during these five years. I would also like to thank Prof. Fabio Canova and Prof. Juan Dolado for many useful suggestions.

The European University Institute allow me to do a great PhD experience in a fascinating location. Thanks to all the fellow PhD students who have shared with me this journey. In particular, working with my coauthor and friend Alejandro Vicondoa was very enjoyable and profitable.

Thanks also to Jessica Spataro, Lucia Vigna and Sarah Simonsen for their prompt help with all administrative issues.

This experience would not have been the same without the heroic Squadra Fantastica and especially my co-misters.

Thanks to my mother for all her teachings and sacrifices.

Finally, I would like to dedicate this thesis to my fiancée Alice, who shared with me the academic track, from the beginning to the end, and much more.

Contents

List of Tables

List of Figures

Chapter 1

News and Noise Bubbles in the Housing Market

"[...] Long-term expectations [...] are arguably the more important determinants of housing demand. [...] Long-term expectations have been consistently more optimistic than short-term expectations across both time and location. [...] It is from these nebulous and relatively slow-moving expectations that the bubble took much of its impetus, and that future home price movements will as well."

from *Case, Shiller and Thompson (2014) "What they have been thinking? Home Buyer Behavior in Hot and Cold Markets"*

1.1 Introduction

The recent boom-bust in the US housing market is a crucial event in contemporaneous economic history. Similar episodes have lately occurred in Spain, Ireland and China. In the US, housing prices rose between 40% and 70% between 2000 and 2006 according to different measures. Then they fell, even more steeply, by similar spectacular amounts after 2006. Housing starts dropped by roughly 80% while the mortgage industry and financial system were stricken very hard beginning of the Great Recession.

To understand this phenomenon, economists have explored important explanatory factors like excessive lending, global imbalances, loosen monetary policy, financial innovations, etc. Case et al. (2015) highlight that, one the other hand,

the role of expectations in the housing market has been significantly underestimated. They analyze the behavior of homebuyers expectations of housing prices from 2003 through 2014 through surveys and highlight some key findings. First, short-term (1 year) expectations have not been over-optimistic but, if anything, have under-reacted to new available information. Second, the roots of the housing boom lie in the long-term (10 year) expectations that were abnormally optimistic.

Understanding the housing market and its drivers is valuable because, due to the its strong ties with the mortgage and the banking industry, housing is particularly important from a macrofinancial stability perspective.[1] While the dot-com bubble lead only to a mild recession, boom-bust episodes in the housing market endanger the stability of the financial system and macroeconomic growth (Crowe et al., 2013; Dell'Ariccia et al., 2016). For example, by using data from 1870 on 17 countries, Jorda et al. (2015, forth) find strong evidence on the predictive power of the housing cycle for financial crisis (especially after WWII). In particular, they compare the consequence of asset price bubbles (equity and housing) and find that the most harmful macroeconomic consequences are generated from leveraged housing bubbles. Finally, Guerrieri and Iacoviello (2016) show the asymmetric effect of housing cycles due to occasionally binding constraint in a dynamic stochastic general equilibrium (DSGE) model. Housing cycles lead to output losses because busts (e.g. 2007-2009) produce larger spillovers than booms (e.g. 2001-2006).

This paper shares the view of Case et al. (2015) on the determinants of housing price. Since homebuyers own their home for many years, purchasing decisions are driven by long-terms expectations. But what determines long-term expectations? Purchase a house means buying a flow of future services, i.e. rents. The fundamental role of rents for housing prices has been studied widely. Among recent contributions, Gallin (2008), Campbell et al. (2009) stand out for studying US data. On the other hand, Ambrose et al. (2013) and Eichholtz et al. (2012) use

[1]Housing is more closely linked to the real economy than other assets because of its unique features. First, housing is the main asset of households and changes in housing wealth have much stronger wealth effect than other assets, e.g. stocks (Case et al. (2005, 2012)). Second, housing is employed as collateral in the mortgage industry. Third, the construction sector, that is mostly labor intensive, comprises an important part of the industrial sector in every economy.

By taking a pure accounting view on US data, housing contributes to GDP in two basic ways: through private residential investment, 5% of GDP, and consumption spending on housing services, 12-13% of GDP, for a total 17-18%. In 2013, the housing stock owned by households and non-profit organization was valued $21.6 trillions, whereas the capitalization of the stock market was $20.3 trillions.

historical data on the housing market of Amsterdam. Gallin (2008) analyze MSAs in the US from 1978 to 2011 and find that pricing error account for half of housing prices volatility. Campbell et al. (2009) study the US housing market and find that expected rent growth is a relevant driver of the rent-price ratio from 1975 to 1997 (together with the expected premia) and by far the main driver from 1997 and 2007. Ambrose et al. (2013) studies 355 years on the housing market in Amsterdam and report two main findings. First, real housing prices and rents are cointegrated and share common fundamentals. Second, deviations from the fundamental housing prices can occur over long periods. Eichholtz et al. (2012) find that rents link the housing market to the real economy by analyzing the housing market of Amsterdam over the period 1550-1850.

Therefore, I take an asset pricing view of housing as the value of housing has to be aligned with the present discount value of rents in the long run. However, because the future is uncertain, expected future rents can be different from the actual ones. In other words, I assume that agents dispose of noisy information on future rents. In this setup, bubbles can arise in the housing market due to imperfect information on future fundamentals. The paper contributes to the literature by analyzing boom-bust episodes in the housing market from a new perspective, with special emphasis on the macroeconomic implications.

I decompose housing price into the correct information (news) and the wrong information (noise) on future rents and analyze their macroeconomic consequences. The identification exploits the non-standard structural Vector Autoregression (SVAR) technique proposed by Forni et al. (2016, forth) (FGLS henceforth). Contrary to DSGE models, this SVAR methodology relies on a minimal set of assumptions and consists of two steps.[2] First, common SVAR estimation and identification procedure are used to recover shocks to expectations about future fundamentals from agents' information set. Second, shocks to expectations are decomposed in news and noise by employing future data on fundamentals, outside agents' information set. I draw the identification assumptions from a simple present value model of housing prices under imperfect information.

News driven business cycle has recently return in vogue thanks to the work of Cochrane (1994) and Beaudry and Portier (2006). Some authors have claimed that news provide noisy information about the future. For instance, Lorenzoni

[2]These assumptions can lead to significantly different results as showed by the opposite conclusions reached by Barsky and Sims (2012) and Blanchard et al. (2013).

(2009), Angeletos and La'O (2013), and Blanchard et al. (2013) study this informational structure in theoretical models. Building upon the work of Lippi and Reichlin (1994), and by simplifying the information structure, Forni et al. (2016) have developed an identification scheme based on dynamic rotations of reduced form residuals to empirically recover news and noise shocks in the stock market. In a companion paper, Forni et al. (forth) study how news and noise drive the business cycle in the US.[3]

My empirical results suggest that fundamentals in the housing market are aligned with the macroeconomy in the long-run. In fact, news (anticipated and realized information about fundamentals) are a major source of fluctuations for rents, housing prices, GDP and stock prices at low frequencies. On the other hand, noise (anticipated but not materialized information about fundamentals) is the most relevant component for short-term fluctuations of housing prices, GDP and residential investment. The historical decomposition suggests that noisy bubbles were a main driver of housing market since the '70s. In particular, the boom-bust occurred in the 2000s is entirely driven by noise, with estimated deviations from the fundamental value in the order of 45%.

My approach is consistent with excess volatility that housing prices exhibit compared to fundamentals (Glaeser et al., 2014). Moreover, I do not take any stance on how agents expectations are formed, but only that information on rents, anticipated or unanticipated, matters for housing prices. Thus both rational and irrational interpretation are compatible with my analysis. In fact, the literature has proposed different way to explain housing price cycles. Zhao (2015) employs an overlapping generation model within a rational framework. Departing from full rationality, Glaeser et al. (2014) propose an extrapolative model of housing prices formation. Adam et al. (2012) and Caines (2015) resort to adaptive learning to explain the dynamic of housing prices. Gelain and Lansing (2014) and Granziera and Kozicki (2015) compare different models of expectations formation on housing prices. While both find that fully rational expectations model under-predict the volatility in housing prices, near rational solutions are instead able to replication the empirical patterns. Engsted et al. (2016) study the explosive behavior of housing prices in OECD countries between 1970 to 2013 and find evidence in favor of the bubble hypothesis for all countries but Germany and Italy. Finally,

[3]Mertens and Ravn (2010) show an application to fiscal policy.

the explanatory power of the model build by Garriga et al. (2012) increases dramatically through shocks to expectations.

The paper is organized as follows. Section 1.2 presents a simple present value model of housing prices. Section 1.3 illustrates the identification strategy. Section 1.4 describes the data, the empirical results and their historical interpretation. Section 1.5 concludes.

1.2 A Present Value Model of Housing under Imperfect Information

In this section, I describe a simple partial equilibrium model of housing. The model can be characterized as a present value model of housing under imperfect information. Housing is an asset that provides a flow of housing services as stocks provide a flow of dividends. Housing services may be traded on the market and produce rental income or they may be directly enjoyed by the owner.[4] The present value model implies that housing prices are the sum of the expected discounted flow of future rents.

Formally, the relationship between prices and rents is determined as:

$$p_t = \mathbb{E}_t \left[\beta_{t,t+1} \left(p_{t+1} + r_{t+1} \right) \right] \tag{1.1}$$

where $\beta_{t,t+1}$ is the stochastic discount discount factor between t and $t+1$ that depends on expected returns.

By iterating forward we obtain:

$$\begin{aligned} p_t &= \mathbb{E}_t \left[\sum_{i=1}^{\infty} \beta_{t,t+i} r_{t+i} \right] + \mathbb{E}_t \left[\lim_{T \to \infty} \sum_{T=0}^{\infty} \beta_{t,T} \mathbb{E}_t \left[p_T \right] \right] \\ &= \mathbb{E}_t \left[\sum_{i=1}^{\infty} \beta_{t,t+i} r_{t+i} \right] \end{aligned} \tag{1.2}$$

[4] Notice that in both cases those housing services have a market value: an actual value in the former and an imputed value computed by public authorities for accounting/taxation purposes in the latter.

with the second term canceling out from the standard transversality condition $\mathbb{E}_t \left[\lim_{T\to\infty} \sum_{T=0}^{\infty} \beta_{t,T} \mathbb{E}_t \left[p_T \right] \right] \rightarrow 0$ that ensures a stable path.[5] Eq.(1.2) states that changes in housing prices may be driven by changes in expected rents and by variations in the discount factor. The discount factor depends on expected returns that can be further decomposed into two components: risk free rate R^{rf} and a risk premium over the risk free rate φ

$$\beta_{t,t+i} = \frac{1}{R^{rf}_{t,t+i} + \varphi_{t,t+i}} \tag{1.3}$$

While the stochastic discount factor cannot be measure exactly, we can roughly account for it with the appropriate long term interest rates. A rate of horizon k is captures fluctuations in the stochastic discount factor from 1 to k period ahead.[6] Moreover, this assumption is necessary to maintain the analysis based on observables.

I introduce a crucial novelty with respect to the previous analyses on housing prices. Purchasing a house means buying a flow of future service whose value is uncertain. Consequently, I assume that agents receive a noisy signal of future rents. This assumption implies that agents' expectations are in part correct and in part wrong. Under this assumption, the process of rents is subject to the *news shock* f_t, anticipated according to the lag polynomial $C(L)$ with L the lag operator:

$$r_t = r_{t-1} + C(L) f_t \tag{1.4}$$

For the sake of clarity, let us simplify the process for rents by considering an innovation anticipated only one period ahead:[7]

$$r_t = r_{t-1} + f_{t-1} \tag{1.5}$$

However, agents cannot directly observe this anticipated shock, but only a noisy signal. The noisy component is labeled *noise shock* n_t. Consequently, agents' expectation can be decomposed as the sum of two gaussian white noise components orthogonal at all lags and leads:

[5] Giglio et al. (2016) provide evidence against violations in the transversality condition in the housing markets of UK and Singapore, even during periods of sizable bubbles.

[6] See Campbell and Shiller (1988)

[7] Notice that the identification strategy will use instead a wide horizon (40 quarters, i.e. 10 years)

$$s_t = f_t + n_t \qquad f_t \perp n_t \tag{1.6}$$

$$\sigma_s^2 = \sigma_f^2 + \sigma_n^2 \tag{1.7}$$

Due to the delayed impact of the news shock f_t on rents, agents cannot disentangle the two components of the signal contemporaneously at time t. Therefore, agents discount the signal according to ψ, which represents the variance of the news relative to the variance of the signal. They will only be able to draw information about the composition of the past signal s_{t-1} by observing the level of current rents r_t. In fact, by construction $\Delta r_t = f_{t-1}$ because the dynamics of rents is affected only by news but not by noise. Only by observing current rents, agents (and so the econometrician) can learn about past news and noise. Importantly, in this setup learning does not regard the knowledge of agents on the true structure of the economy but only the past realizations of news and noise.[8]

In other words, agents learn their past noise contained in the signal by observing their prediction error on rents u_t:

$$\begin{aligned} u_t &= \Delta r_t - \mathbb{E}_{t-1}\left[\Delta r_t\right] \\ &= \Delta r_t - \psi s_{t-1} \end{aligned} \tag{1.8}$$

u_t represents the unanticipated shock to rents and can be rewritten in terms of f_{t-1} and n_{t-1} as $u_t = f_{t-1}(1-\psi) + \psi n_{t-1}$. Equations 1.5 to 1.8 provide the theoretical interpretations of the shocks that are recovered through the identification strategy in Section 1.3. In sum, housing prices are affect by three components that are fundamental: news, risk free rate and risk premia. The fourth component, noise, can be interpreted as a bubble component driven by imperfect information.

In a first approach, I impose directly, even if loosely, the present value relationship by assuming that housing prices are themselves a signal of future rents. However, compared to the equity market and to the bond market, the housing

[8] Agents discount the signal based on its reliability. The more the signal is driven by news, the more agents respond to changes in the signal: $\mathbb{E}_t\left[r_{t+1}\right] = \frac{\sigma_f^2}{\sigma_s^2} s_t$.

Consequently, noise shocks do not have any effect in two cases. First, when the noise shock does not exist $\sigma_n = 0$. Second, when the signal is completely unreliable $\sigma_f = \sigma_s$. On the other hand, the maximum impact of noise shocks corresponds to the case $\sigma_f \approx \sigma_n$ as noise is a significant driver of the signal and the signal contains useful information on future fundamentals.

market features unique characteristics that make it regulated and relatively inefficient. While a vast part of the asset pricing literature on housing as neglected these issues, some key features may lead to departures from the present value relationship. Asset pricing model builds upon the no arbitrage assumption. However, the no arbitrage assumption might not be satisfied due to some special features of housing: I) Houses are sold only entirely and cannot be divided in smaller pieces (indivisibility). Consequently, it is hard to short them; II) The housing market is characterized by borrowing constraints, illiquidity and transaction costs; III) Housing is a collateralizable asset; IV) The housing market is heavily regulated both for purchasing and for renting; III) In the short run, renting and purchasing are substitutes and the two markets are segmented; IV) The time to build in residential investment.[9]

Due to these features, housing prices might not reflect instantaneously the presented discounted value of future rents. To tackle this concerns, I show that I obtain the similar results by relaxing the present value relationship. In different robustness exercises, I exploit other variables as measures of expectation in the housing market and I check ex-post whether the present value relationship is supported by the data. I employ variables that arguably take into account new information immediately as the consumer and home builders surveys and stock prices of home builder companies.

Other potentially relevant issues are fiscal treatment of owning, renting and the tax deductibility of mortgage interests. While those points are impossible to address properly as the data on the subject is poor, I check ex-post whether the shocks that I identify are correlated with changes in the tax rate.[10]

1.3 Identification Strategy

The following moving average (MA) representation expresses the relationship between the variables introduced in Section 1.2: the change in rents Δr_t, housing

[9] See Piazzesi and Schneider (2016).

[10] Property tax data are available only annually and, consequently, the following exercise is performed at this frequency. More in detail, I take the OECD data on property taxes an build a measure of implicit tax rate dividing the revenues by housing prices. I regress the news and noise shock that I identify on the property tax and I do not find a significant relationship. For both news and noise shocks, the coefficient is negative but not even significant at the 10% level. These results are available upon request.

prices p_t,[11] and other relevant variables X_t (in particular the risk free rate and risk premia)

$$\begin{bmatrix} \Delta r_t \\ p_t - r_t \\ X_t \end{bmatrix} = \begin{bmatrix} c(L) & 0 & 0 \\ 1 & 1 & 0 \\ \theta(L) & \gamma(L) & 1 \end{bmatrix} \begin{bmatrix} f_t \\ n_t \\ \varepsilon^x \end{bmatrix} \tag{1.9}$$

where $c(L)$, $\theta(L)$, $\gamma(L)$ are lag polynomials. By construction the news shock f_t has a lagged impact on r_t, $C(0) = 0$ while further lags are different from 0. For example, the representation in eq.(1.5) entails $C(L) = L$. Unfortunately, $C(0) = 0$ implies that the three variables at time t do not provide useful information to recover f_t and n_t. In other words, the VAR estimated using $\begin{bmatrix} \Delta r_t & p_t - r_t & X_t \end{bmatrix}'$ cannot be inverted to recover the MA representation in eq.(1.9) as the determinant of the matrix is 0.

At time t, there are two sources of information. First, we observe housing prices p_t that measure the available information on future rents. Second, the difference between the current realization of rents and agents' expectations yield the past forecast error of agents. In the stylized model presented in Section 1.2, this means $u_t = \Delta r_t - \mathbb{E}_{t-1}[\Delta r_t] = \Delta r_t - \psi s_{t-1}$.

However, only by employing future values of the surprise shock u_t it is possible to recover f_t and n_t. Intuitively, by checking whether agents' expectations of future rents were correct or wrong we can infer whether changes in the signal were driven by news or noise. This is possible because the noise shock does not affect rents at all. Therefore, observing future rents provide perfect information on past news.

The identification strategy consists of two steps.

First, estimate a reduce form VAR using $\begin{bmatrix} \Delta r_t & p_t & X_t \end{bmatrix}'$ and recover $\begin{bmatrix} u_t & s_t & \varepsilon^x \end{bmatrix}'$ through a simple Choleski decomposition. In fact, the corresponding MA representation

$$\begin{bmatrix} \Delta r_t \\ p_t - r_t \\ X_t \end{bmatrix} = \begin{bmatrix} d(L) & c(L)\psi & 0 \\ 0 & 1 & 0 \\ \tilde{\theta}(L) & \tilde{\gamma}(L) & 1 \end{bmatrix} \begin{bmatrix} u_t \\ s_t \\ \varepsilon^x \end{bmatrix} \tag{1.10}$$

[11] Or anyway a signal of future rents that may be different from housing prices. I test the robustness of relaxing the implicit present value relationship in the empirical analysis.

is invertible. The ordering in eq.(1.10) implies that changes in discounting, e.g. risk premia, do not affect contemporaneously housing prices or alternative signals. In the empirical analysis, I will also clean signal shocks from changes in risk premia by ordering X_t before p_t.

Second, we want to exploit future realization of r_t to decompose the signal into news and noise. At this aim, we have to decompose $c(L)$ into two components: $c(L) = b(L)d(L)$. The decomposition is achieved by means of Blaschke matrices as shown in Lippi and Reichlin (1994) and FGLS.[12] $b(L)$ incorporates the roots of $c(L)$ lying inside the unit circle, while $d(L)$ collects the roots of $c(L)$ inside the unit circle. Intuitively, $b(L)$ extracts the information that is available only from the future realizations of r_t, while $d(L)$ represents the new information that agents receive by observing the current value of r_t. This decomposition allows us to express the surprise shocks in the general case as $u_t = \frac{\Delta r_t}{d(L)} - \psi c(L) s_t = b(L) f_t - \psi b(L) s_t$.[13]

Finally, we have derived the relationship

$$\begin{bmatrix} u_t \\ s_t \\ \varepsilon^x \end{bmatrix} = \begin{bmatrix} (1-\psi)\, b(L) & -\psi b(L) & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_t \\ n_t \\ \varepsilon^x \end{bmatrix} \tag{1.11}$$

that links the shocks that can be recovered through standard SVAR tools, belonging to the present information set, $\begin{bmatrix} u_t & s_t \end{bmatrix}'$ and the shocks that we are truly interested in $\begin{bmatrix} f_t & n_t \end{bmatrix}'$, which belong to the future information set. By construction, $b(L)$ contains the roots of $c(L)$ lying inside the unit circle and, consequently, cannot be inverted in the past but only in the future: $b(L)^{-1} = b(L^{-1}) = b(F)$ where F is the forward operator:

$$\begin{bmatrix} f_t \\ n_t \\ \varepsilon^x \end{bmatrix} = \begin{bmatrix} b(F) & \psi & 0 \\ -b(F) & 1-\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_t \\ n_t \\ \varepsilon^x \end{bmatrix} \tag{1.12}$$

Another important point arises from the derivations presented hereby. In fact, after the shocks u_t and s_t are identified with standard VAR tools, the other shocks ε_t^x are not employed in the identification of news and noise shocks. Basically, X_t

[12] Blaschke matrices are complex value operators that conserve the orthonormality of the vectors to which they are applied.

[13] The stylized case mentioned above corresponds to the case $d(L) = 1$ and $c(L) = L$.

is directly relevant only for the identification of u_t and s_t. On the other hand, X_t enters only indirectly the identification of f_t and n_t, but there is no direct link with news and noise shocks.

Finally, the present value model in Section 1.2 is a useful tool to guide the analysis and interpret the empirical results. However, in the empirical analysis I depart from a strict interpretation of the the present value model. The economic framework that guides the empirical investigation can be summarized by: (a) expectations of futures rents are relevant for housing prices such that agents respond to new information on rents coming from news shocks f_t and surprise shocks u_t; (b) the logarithms of rents and housing prices are cointegrated.

Starting from these premises, the econometric assumptions are: I) The news shock f_t produce a permanent effect on rents; II) The noise n_t shock does not affect rents at any lead. III) The signal shock[14] is the sum of the news shock f_t and of the noise shock n_t; IV) The only shock affecting rents on impact is the surprise shock u_t. Additional shocks affect rents only with a lag and are observed.[15]

Some of the assumptions enumerated above are actually not imposed but used for testing the validity of the theoretical framework. In particular, assumption (b) is not imposed. In this way, I can test whether the news shock f_t, which generates a permanent effect on rents by (I), also produces a permanent effect on housing prices. In the same spirit, it can be checked ex-post whether noise shocks, which following (II) have no effect on rents, do not generate permanent effects in housing prices. Moreover, assumption (II) is imposed only on impact and in the long run (40 quarters), but intermediate horizons are used to test whether such an assumption is supported by the data. To conclude, the identification imposes restrictions on the effect of the news shock f_t and of the noise shock n_t on rents while leaving the implications of those shocks on housing prices completely unrestricted.[16]

[14] A shock to housing prices in the baseline analysis.

[15] Rents are a slow moving variable and usually rental contracts last at least for one year.

[16] For a Monte Carlo test of the methodology, see FGLS.

1.4 Empirical Analysis

1.4.1 The Data

I employ US quarterly data from 1963:Q1 - 2016:Q3. I build a measures of total rents multiplying the rents personal consumption expenditures (PCE) price index (*FRED*: DHUTRC1Q027SBEA) by the rents PCE quantity index (*FRED*: DTENRA3). Then, I divide this series by the PCE implicit price deflator (*FRED*: DPCERD3Q086SBEA) and population aged 16 years or more (*FRED*: NP16OV_NBD19480101). In the end, I obtain a measure of real per capita rents that I transform in logs.[17] As a measure of housing prices, I use the log of the Average Price of Houses Sold (*FRED*: ASPUS) deflated by the GDP implicit price deflator (*FRED*: GDPDEF).[18] As first alternative signal of future rents, I use survey variables. In order to measure demand side expectations, I employ the log of Buying Conditions for Houses (Table 41 in the Michigan Survey of Consumers). For the supply side, I take the National Home Builders Association market survey concerning the expected new single family home sales in next 6 months. (Datastream: USNAHB1E). Moreover, I also combine these two variables through a principal component analysis. As second alternative signal, I exploit information on home builders from the stock market through the Datastream Home Builders Stock Price Index (Datastream: HOMESUS). I deflate this variable by the GDP implicit price deflator and take logs.

The risk free interest rate employed in my baseline specification is the 3-Month Treasury Bill, Secondary Market Rate (*FRED*: TB3MS). The interest rate that incorporates a risk premium component is the Moody's Seasoned Aaa Corporate Bond Yield (*FRED*: AAA). As a robustness exercises, I also use interest rates more closely related to the housing market. In particular, I take the 30 years Fixed Rate Mortgage (*FRED*: MORTGAGE30US) as measure including a risk premium and the corresponding maturity of the Treasury Yield (*FRED*: USTYCO30R) as risk free measure.[19]

In the macroeconomic analysis, I include additional variables as the log of the Real Gross Domestic Product (*FRED*: GDPC1), the log of Real Private Residential

[17] Notice that this measure is equivalent to the total per capita dividends employed in FGLS.

[18] Similar results hold by using the Case Shiller S&P Corelogic Home Price Index (*FRED*: CSUSHPISA) or the series available on the website of R. Shiller.

[19] Or the equivalent 20 years maturity that spans a longer sample.

Fixed Investment (*FRED*: PRFIC1) deflated by the GDP implicit price deflator, and the log of the Standard & Poor's Index of 500 Common Stocks (Datastream: US500STK) deflated by the GDP implicit price deflator.

I use the variables in log-levels to avoid cointegration problems following Stock et al. (1990). I include 2 lags as consistently suggested by the three information criteria AIC, BIC, HQC. Variables are downloaded as seasonally adjusted if possible. Otherwise, I employ the *Census X13* to remove the seasonality (if necessary).

1.4.2 Decomposition of Housing Prices

This first stage concerns the decomposition of housing prices in the four components highlighted in Section 1.2. I include rents, housing prices, a risk free rate and rate including risk premia and identify four corresponding shocks in the current information set. The shocks are a surprise shock to rents, a shock to housing prices (signal shock), a risk free rate shock (3 months T-Bill rate) and finally a risk premia shock (Aaa Moody's Corporate Bond Yield). While I present the results using short-term rates, the same results hold using longer rates (e.g. rates closely related to the housing market as the mortgage rate).[20] The only role of variables other than rents and housing prices concerns the identification of surprise and signal shocks. Figure 1.1 display the impulse response functions (IRFs) to the surprise and signal shock.[21]

Next, I use the procedure explained in Section 1.3 to identify news and noise shocks as a linear combination of future surprise and signal shocks. The other shocks identified in the VAR, i.e. risk free rate shocks and risk premia shocks do not enter directly this additional computation. The IRFs are reported in Figure 1.2. The news shock is identified as a shock to housing prices that produces a delayed but permanent increase in (future) rents. On the other hand, the noise shock is a shock to housing prices without impact and long-run effect on rents. Recall that the response of housing prices to news and noise shocks is left completely unconstrained. Figure 1.2 suggests that the news shock generates a permanent effect on

[20] Including the risk premia explicitly as spread or as implicit component in the mortgage rate yields the same results.

[21] The effects of risk free and risk premia shocks are reported in Figure A.1 Appendix. These shock produce negligible effects on rents and a negative response of housing prices in line with economic theory.

housing while the noise shock produce a strong response that dies out after about 20 quarters. Even if not statistically different, the estimate impact responses of housing prices to the noise shock is greater than the response to the news shock.

Finally, the forecast error variance decomposition (FEVD) in Figure 1.3 highlights that, consistently with the identifying assumptions, the noise shock explains a marginal share of the variability in rents, while news shock explains the great majority of it. For what concerns housing prices, the role of the noise shock is dominant up to 15 quarters (ranging from 68% to 20%) and gradually dies out after this horizon. The news shock contributes to the volatility of housing prices for 38% on impact and 82% after 40 quarters.

The information set employed in this exercise poses a potential concern. While Forni et al. (2016) find that the equivalent four variable VAR for the stock market is informationally sufficient, the housing market is more closely related to the real economy for multiple reasons (see also Section 1.2). In fact, the Granger test proposed by Forni and Gambetti (2014) reject the orthogonality of the shocks to lags of potentially informative variables excluded from the VAR (Table A.2).[22] This is problematic because the shocks might be not correctly identified. To tackle this concern, in the next Section I append the VAR with key macroeconomic aggregates.

1.4.3 Macroeconomic Analysis

In this second stage, I include macroeconomic aggregates like GDP, residential investment and the stock price index for multiple reasons. I repeat the Granger test already applied in Section 1.4.2. The test now suggests that the shocks are orthogonal to agents' information set (Table A.3). Thus, the appended VAR is informationally sufficient. There are other two key reasons to extend the analysis. First, the goal of the paper is to assess the macroeconomic effects on news and noise shocks. In particular, we can assess the impact of housing bubbles on the macroeconomy as the identification isolates the bubble component in the housing market. Second, I investigate whether the fundamentals in the housing market are aligned with the macroeconomy in the long-run. I analyze whether a news shock that has a permanent impact on rents has a similar effect on GDP and stock

[22]The variables are reported in Table A.1. Notice that Canova and Sahneh (2016) propose a different test. They claim that Granger tests may detect omitted variables but are not strictly related to non-fundamentalness.

prices. In this specification, I include only the Aaa Moody's Corporate Bond Yield and remove the 3-month T-Bill rate to avoid inflating the dimensionality of the system.

I employ three different measures of expectations in the housing market. The most natural consists of housing prices themselves (as in Section 1.4.2). In this way, housing prices are assumed to incorporate all the available information about future rents.

However, there may be departures from this framework because of the various reasons discussed in Section 1.2. Therefore, I show that similar results hold by employing two alternative variables as signal of future rents. First, I employ the Buying Conditions for Houses from the Michigan Survey of Consumers. I also include supply side expectations by using surveys from the Home Builders National Association. Second, I assume that expectations in the housing market can be (partially) proxied by the Home Builders Stock Prices Index. The three different signals produce very similar results. In particular, it is worth highlighting that across all cases, a news shock produces a permanent effect both on rents and on housing prices. Moreover, a noise shock generates significant responses on housing prices and on macroeconomic aggregates.

1.4.4 Housing Prices as Signal

This section present the results obtained by using housing prices as signal of future rents as implied by the present value relationship in equation 1.2. Notice that the present value relationship is not imposed in a strict way, but I only assume that new information on rent affects housing prices.

In this case, I include the following variables [RENTS, HOUSING PRICES, GDP, RESIDENTIAL INVESTMENT, AAA MOODY'S CORPORATE YIELD, S&P INDEX] as described in Section 1.4.1. The ordering matters for the identification of the surprise shock and the signal shock as they are recovered through recursive zero short-run restrictions (Choleski). I explore the sensitivity of the analysis to two alternative orderings finding comparable results. I order housing prices second after rents, but I also check the robustness of the results by placing them second to last (before stock prices).

Figure 1.4 display the IRFs to a surprise shock and to a signal shocks. The surprise shock generates a permanent effect on rents, housing prices, GDP and

stock prices. There is no significant permanent effect residential investment as for this series the trend is dominated by huge cyclical fluctuations. The signal shock is a mixture of news and noise as represented in equation 1.6. It predicts an increase in rents, GDP and stock prices because it incorporates news. However, for GDP and housing prices this permanent effect is not statistically significant because the signal shock is also incorporating noise.

Figure 1.5 reports the IRFs to news and noise shocks.[23] First, the noise shock does not have any significant effect on rents. I am imposing that the noise shock does not have an impact and long-run (40 quarters) effect on rents, but at intermediate horizon the response is unconstrained. Therefore, we can use the IRF at intermediate horizon as a diagnostic of the identification strategy. In fact, the key assumption is that rents allow to infer the past values of news and noise as they are not affected by noise. Conversely, the news shock has a lagged but persistent effect on rents. The lagged response of rents after the news shock is another good indication of the identification. The news shock is constrained to have a delayed effect on rents. The gradual increase of rent after a news shock suggests that the identification is supported by the data.[24] Housing prices react in a stronger way to the noise shock than to the news shock on impact and in the short-run. This means that the estimated variance of the noise shock is bigger than the variance of the news shock. On the other hand, while the news shock lead to a permanent change in the response of housing prices, the effect of a noise shock gradually dies out and reaches zero after 15 quarters.

Regarding the macroeconomic effects, the news shock has the same permanent effect on rents, housing prices GDP and stock prices. This means that the fundamentals in the housing market are in line with the macroeconomy. The noise shock produces sizable but temporary effects on GDP, residential investment and stock prices.

[23]The empirical results are consistent with DSGE model similar to Iacoviello and Neri (2010) that includes also a rental market. News and noise shock to housing preference, or simply as a reduced form shocks to rents, produce similar response of the corresponding VAR variables.

[24]The analysis may be biased if we think that the economy may be hit by transitory fundamental shocks. These kind of shocks are are neglected in this identification strategy but, if relevant, they would still be captured by the noise shock. The reason lies in the fact that also transitory fundamental shocks have zero long run effect on rents, but, differently from the noise shock itself, they should affect rents at intermediate horizons. Given that noise shocks have no relevant effect on rents at any horizon, we can conclude that identified noise shock does not contain also transitory fundamental shocks.

The quantitative implications of news and noise shocks are reported through the FEVD in Figure 1.6. Rents are explained entirely by the news shock while the noise shock has a marginal effect. The decomposition suggests that the bubble component is the main drive of housing prices at high frequency (75% on impact) but the news is the major driven in the long run, consistently with theory. In fact, after 15 quarters the effect of the noise shock on housing prices dies out. A similar picture holds for GDP. News is the main driver in the long-run, but in the first five quarters noise dominates by explaining about 10% of the volatility of GDP. Finally, residential investment is strongly affected by noise both at short and medium horizons (up to 20%) while in the long run news plays a major role. Due to the huge cyclical fluctuations, there is no significant permanent effect of any shocks on residential investment.

1.4.4.1 Historical Decomposition

Figure 1.7 conveys the historical implication of the analysis. In this figure, I decompose the series of housing prices in a two orthogonal components: the fundamental component and bubble-noise component. The latter is expressed in percentage deviations from the fundamental component. The first important feature is the increasing role of the noise component over the sample. In particular, the noise component turns out to be more important after the deregulation of the 1980s.

In the 1970s, the housing market experienced a boom after the first *oil price shock* in 1973 that is associated with a 25% noisy component.[25] The downturn in the US housing market of the 1990s was driven by noise (32%).

The largest housing cycle in the sample, which spans from 2000 to 2012, is driven by a huge bubble, not related with future rents. The noisy component accounts for sizable deviations of housing prices from fundamentals both during the build-up and the collapse of the real estate market. The peak occurs in 2003 and accounts for 45% deviation from fundamentals. The through explains about 40%

[25]Poterba (1984, 1991) relates this boom to demographics and the relationship between high expected inflation and mortgage tax deductions. Piazzesi and Schneider (2009a) add the role of inter-generational heterogeneity (baby boomers) to the high inflation explanation Demographics factors are characterized as news as they are arguably expected to move future rents. This explanatory factor might be erroneously captured as a noise component. However, the estimation on the sample 1985-2016 yield the same IRFs. Moreover, I control for inflation expectations by including the inflation expectations from the Michigan Survey of consumers (Table 32) finding similar results to the ones reported. These results are available upon request.

of deviations from fundamentals and occurs in 2009. The same historical decomposition of the Case & Shiller Corelogic Home Price Index points at an exceptional role of noise during the 2000s. Therefore, expectations unrelated with fundamentals are a key driver both of the housing boom that preceded the Great Recession and of the following bust. Investigating the underlying reasons for those changes in expectations is beyond the scope of this analysis, but this perspective should receive more attention as already mentioned in Case et al. (2015).

These results are consistent with other existing works. For example, Agnello and Schuknecht (2011) estimate a housing bust phase over 1990-97 (24% magnitude) and a boom from 1998-2005 (42% magnitude).[26]

Another important feature is the higher volatility of the noisy component with respect to the actual housing price series. The excess volatility steams from the highly volatile housing price series employed in the baseline analysis (which is much more volatile than rents). This excess volatility disappears when using the Case & Shiller Corelogic Home Price Index or other signals, whereas the key findings from the historical decomposition hold still.

1.4.5 Alternative Signals

In two robustness exercises, I proxy the expectations in the housing market using two alternative variables. This allows me to relax the assumption that housing prices capture instantaneously all the available information on future rents. In this way, we are able to test whether the results in Section 1.4.4 are driven by this assumption.

First, I use survey information from the Buying Conditions for Houses from the Michigan Survey of Consumer as signal. This variable captures demand side expectations without a specific horizon.[27] In order to include a proxy for the supply side expectations, I employ the National Home Builders Association market survey (new single family home sales in next 6 months). Using this variable, available only on a restricted sample, I find similar results. Finally, employing together demand side and supply side expectations through a principal component analysis yields consistent results. In this case, the VAR features the following

[26] They use the turning point methodology developed by Harding and Pagan (2002).

[27] Ideally, we would like to use a measure of exclusively long-term expectations that, unfortunately, is not available

variables [RENTS, SURVEY, GDP, RESIDENTIAL INVESTMENT, HOUSING PRICES, AAA MOODY'S CORPORATE YIELD, S&P 500].

Second, I exploit the information contained in Home Builders Stock Price Index. This variable contains the stock price of quoted home builders firms. This variable arguably captures instantaneously all the available information. On the other hand, there are also some disadvantages: i) stocks are an alternative asset, with some dynamics that may be unrelated to the housing market at high frequencies;[28] ii) home builders are related exclusively with new houses built. In this case, the VAR includes [RENTS, HOME BUILDERS STOCK PRICE INDEX, GDP, RESIDENTIAL INVESTMENT, HOUSING PRICES, AAA MOODY'S CORPORATE YIELD, S&P 500]

The signal shock is now identified from these two alternative variables and not directly from housing prices. Nonetheless, the results are comparable to the findings presented in Section 1.4.4 (Figure A.3-A.7). In particular, the news shock generates a permanent increase in rents and housing prices, GDP and stock prices. The noise shock produces significant transitory responses in housing prices, GDP, residential investment and stock prices. On the other, rents do not respond significantly to a noise shock.

These two alternative variables captures only partially expectations in the housing market by construction. They convey information exclusively on consumers' expectations or home builders stock prices. Thus, while the results are qualitatively very similar across all the specifications, the FEVD display a smaller role for the noise shock compared to Section 1.4.4 (Figure A.4-A.8). Noise accounts for a maximum of about 10% of the volatility of housing prices using the demand side survey or the HBSPI. However, noise shocks still produce sizable macroeconomic implications as it is reflected in 15% (5%) of variability in GDP and in 20% (5%) residential investment using the survey (HBSPI). On the other hand, the principal component of the demand and supply surveys captures expectations in a more complete fashion. In this case, the quantitative role of noise raises considerably for housing prices (25%), GDP (17%) and residential investment (30%).

Finally, the historical decomposition of housing prices (Figure A.5-A.9) displays some phase shift in the noise component comparing to Figure 1.7. However, the main episode in the early 2000s is consistently decomposed across the three

[28] I obtain comparable results if I use as signal the HBSPI orthogonalized to the general stock price index.

different approaches. Using these two alternative signals, the noise component explains 20% deviation of housing prices from the fundamental value.

1.4.6 Additional Robustness Exercises

I have performed additional robustness exercises: i) The most important concerns the interest rate used which account for changes in discounting. Employing the long-term interest rate, both risk free and including a risk premium, produces very similar results to the baseline analysis; ii) I have included the amount of mortgages (or total loans) to control for the quantity side of credit conditions; iii) There might be concerns that the results are driven by the bust in the housing market after 2007. Similar results, with lower statistical significance, are obtained by running the same exercise on the sample cut in 2006; iv) Finally, I have employed also the Case & Shiller Corelogic Home Price Index instead of the Census measure of housing prices on a restricted sample obtaining consistent results.

1.5 Conclusions

In this paper, I explore the role of long-term expectations in the housing market in line with Case et al. (2015). I argue that long-term expectations have to be aligned with future rents that represent the dividends from housing. However, since the future is uncertain, expectations of future rents are noisy. Therefore, I introduce imperfect information in a present value model of housing prices. Agents receive noisy signals about future fundamentals and, as a result, bubbles can arise due to this informational incompleteness. The approach is compatible both with rational and irrational agents.

From a stylized present value model, I have derived identifying restrictions to apply the non-standard structural VAR procedure developed by Forni et al. (2016, forth). This methodology employs dynamic rotations of the reduced form residuals to recover shocks related (news) and not related (noise) to future rents. The identification exploits future rents to determine whether shocks to housing prices are fundamental or noisy. At the same time, I roughly control for changes in discounting by including different measures of interest rates (risk free and containing a risk premium).

The empirical results provide different insights on the dynamics of the housing market in the US. First, fundamentals in the housing market are aligned with the macroeconomy. In fact, a news shock produces a permanent increase in rents, housing prices, GDP and stock prices. Second, expectations not related to fundamentals are a relevant source of fluctuations for the housing market and macroeconomy. Noise shocks generate sizable responses in housing prices, GDP and residential investment at the business cycle frequency but without permanent effects. In terms of forecast error variance, the noise shock is dominant in explaining housing prices, GDP and residential investment at high frequency, but dies out after about 15 quarters. The historical decomposition shows that the role of noisy bubbles, threatening macroeconomic stability, has increases over time. In particular, the boom-bust in the housing market during 2000s was entirely due to a noisy bubble. Moreover, also the boom in 1970s and the depression in 1990s was driven by a significant noise component.

The results are robust to different specifications. In particular, I relax the present value hypothesis and rely on other variables to proxy expectations in the housing market. On the one hand, I employ survey variables from consumers (demand side) and home builders (supply side). On the other hand, I rely on the stock prices of home builder companies. In both cases, the results are comparable to the baseline specification.

The analysis sheds light on a relevant source of macrofinancial instability that arises from the housing market. This finding is particularly meaningful in light of the asymmetric spillovers from the housing market to the macroeconomy. Guerrieri and Iacoviello (2016) show the implications of the occasionally binding collateral constraint on housing. The contribution of the housing boom (2001-2006) to consumption growth was small. On the other hand, the (negative) implications for consumption became much largest during the bust as the collateral constraint become binding (2007-2009). Consequently, macro-prudential policies with a stabilizing role may have the potential to reduce the harmful consequences of busts and the excess volatility introduced by informational incompleteness. Exploring this venue remains open for future research.

1.6 Figures

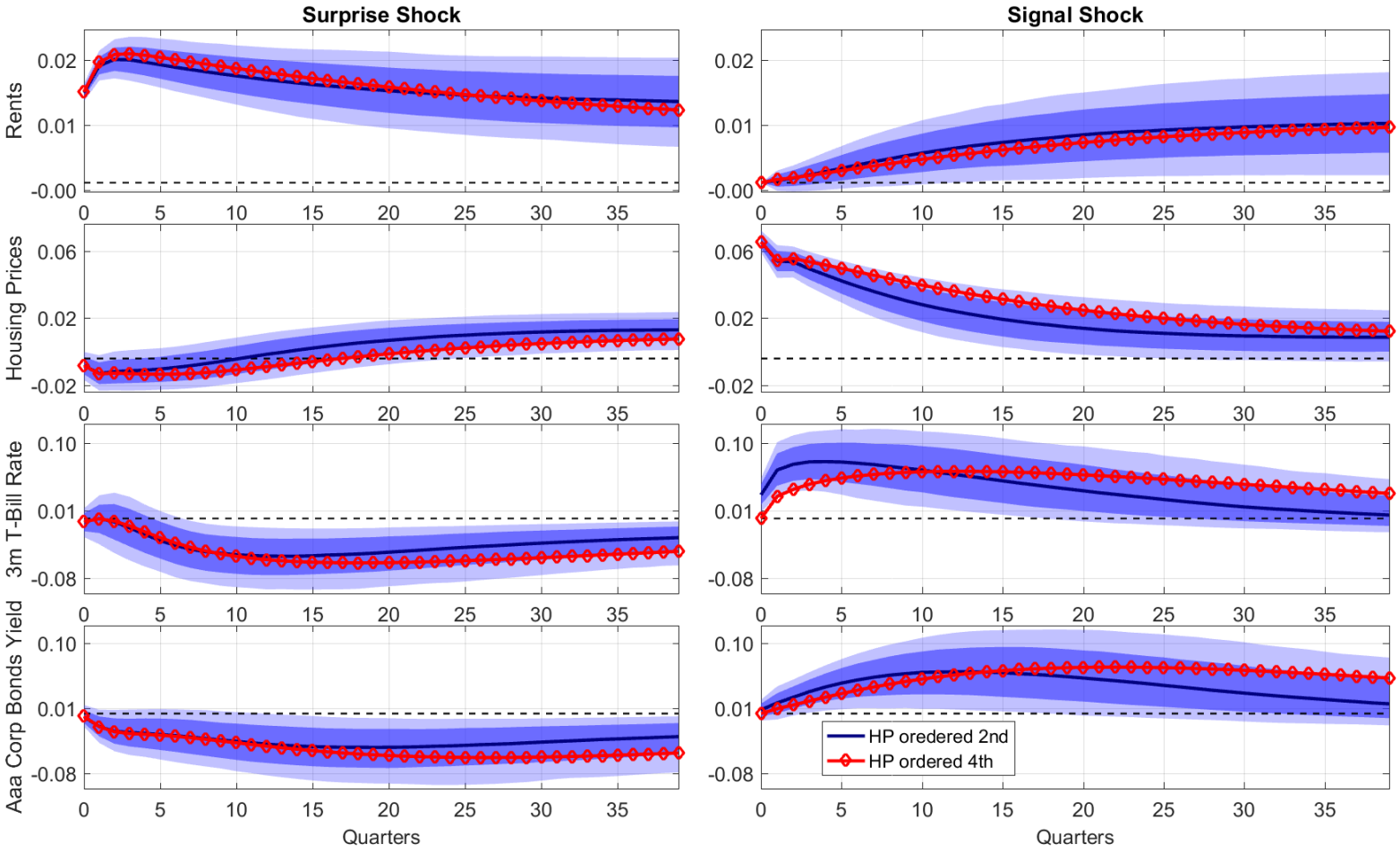

Figure 1.1: IRFs to Surprise and Signal Shocks - Decomposition

IRFs to a surprise shock to rents and to a signal shock. The responses are reported in terms of the standard deviation of the variables in the system. The solid blue line is the median, the dark and light blue shaded areas represents 68% and 90% confidence bands respectively (2000 bootstrap replications). The shocks are identified through the following ordering: [Rents, Housing Prices, 3 Months Bill Rate, Aaa Moody's Corporate Bond Yield]. The red line reports the median IRFs obtained by the recursive ordering [Rents, 3 Months Bill Rate, Aaa Moody's Corporate Bond Yield, Housing Prices]. Sample: 1963:Q1 - 2016:Q3.

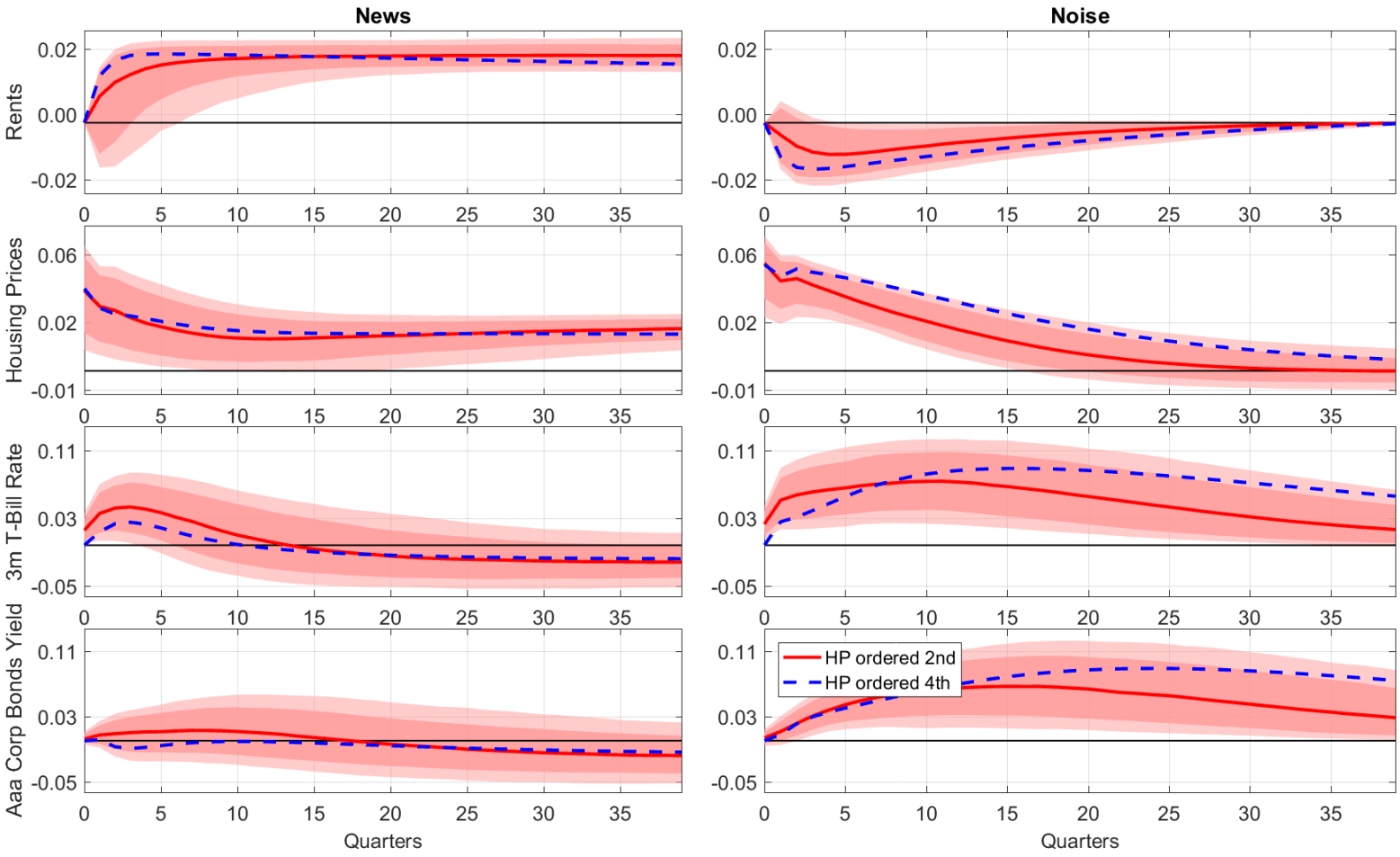

Figure 1.2: IRFs to News and Noise Shocks - Decomposition

IRFs to news and noise shocks. The responses are reported in terms of the standard deviation of the variables in the system. The solid red line is the median, the dark and light red shaded areas represents 68% and 90% confidence bands respectively (2000 bootstrap replications). The shocks are identified through the following ordering: [Rents, Housing Prices, 3 Months Bill Rate, Aaa Moody's Corporate Bond Yield]. The blue dotted line reports the median IRFs obtained by the recursive ordering [Rents, 3 Months Bill Rate, Aaa Moody's Corporate Bond Yield, Housing Prices]. Sample: 1963:Q1 - 2016:Q3.

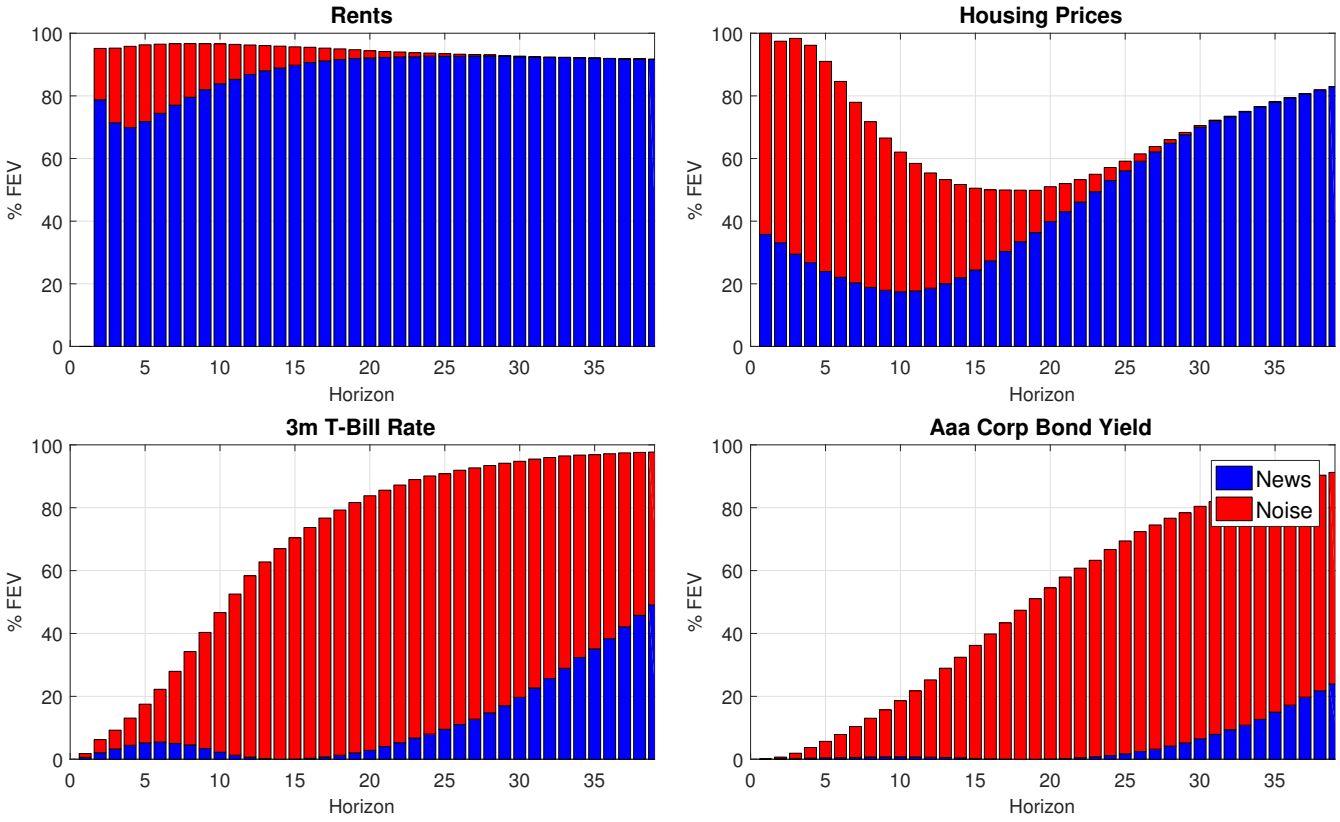

Figure 1.3: Forecast Error Variance Decomposition

Forecast error variance decomposition of the variables in the system. The plot display the share of the variance explained by news and noise at each horizon (not cumulatively). The shocks are identified through the following ordering: [Rents, Housing Prices, 3 Months Bill Rate, Aaa Moody's Corporate Bond Yield]. Sample: 1963:Q1 - 2016Q3.

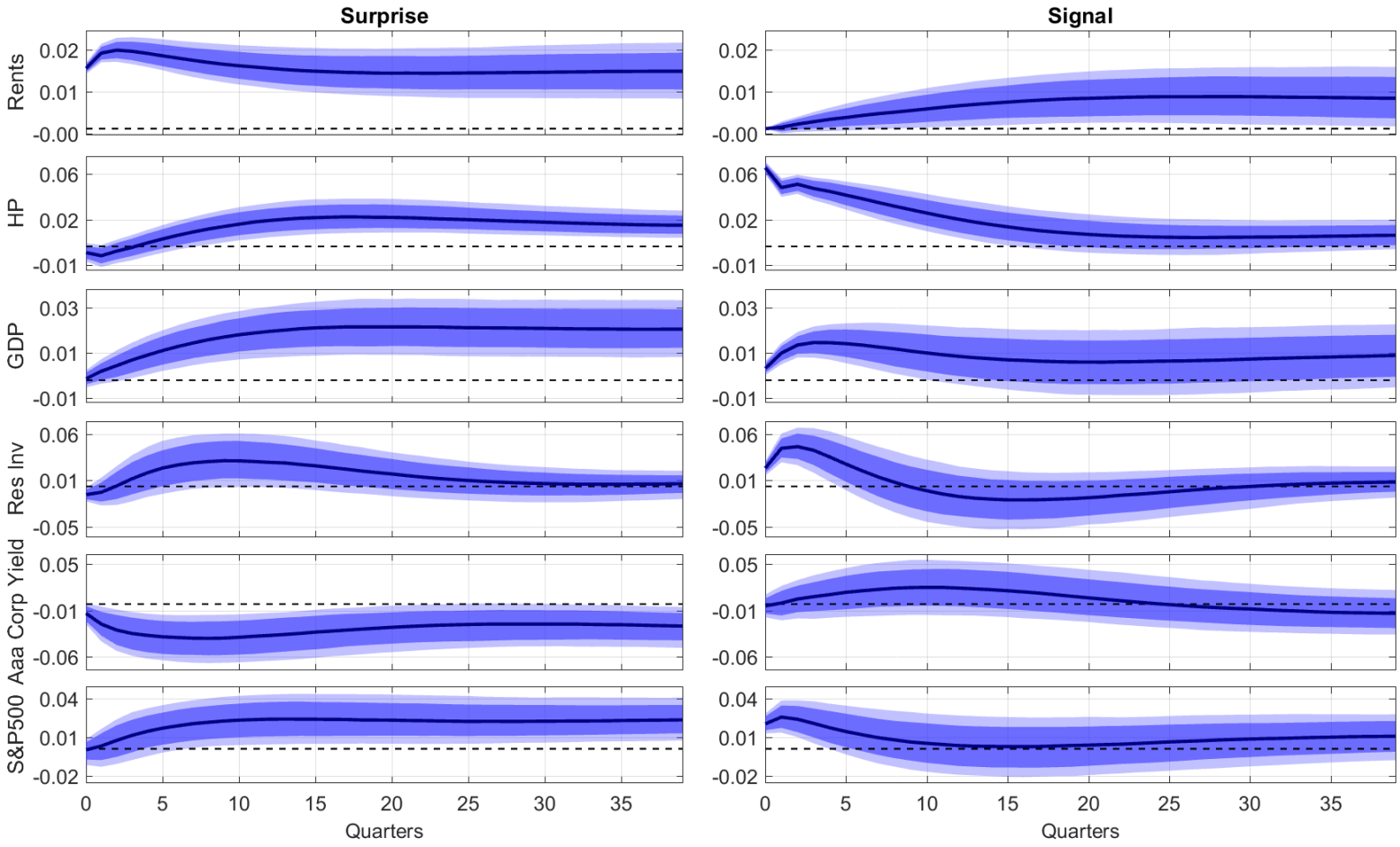

Figure 1.4: IRFs Surprise and Signal Shocks - Macro Analysis

IRFs to a surprise shock to rents and to a signal shock. The responses are reported in terms of the standard deviation of the variables in the system. The solid blue line is the median, the dark and light blue shaded areas represents 68% and 90% confidence bands respectively (2000 bootstrap replications). The shocks are identified through the following ordering: [Rents, Housing Prices, GDP, Residential Investment, Aaa Moody's Corporate Bond Yield, S&P Index]. Sample: 1963:Q1 - 2016:Q3.

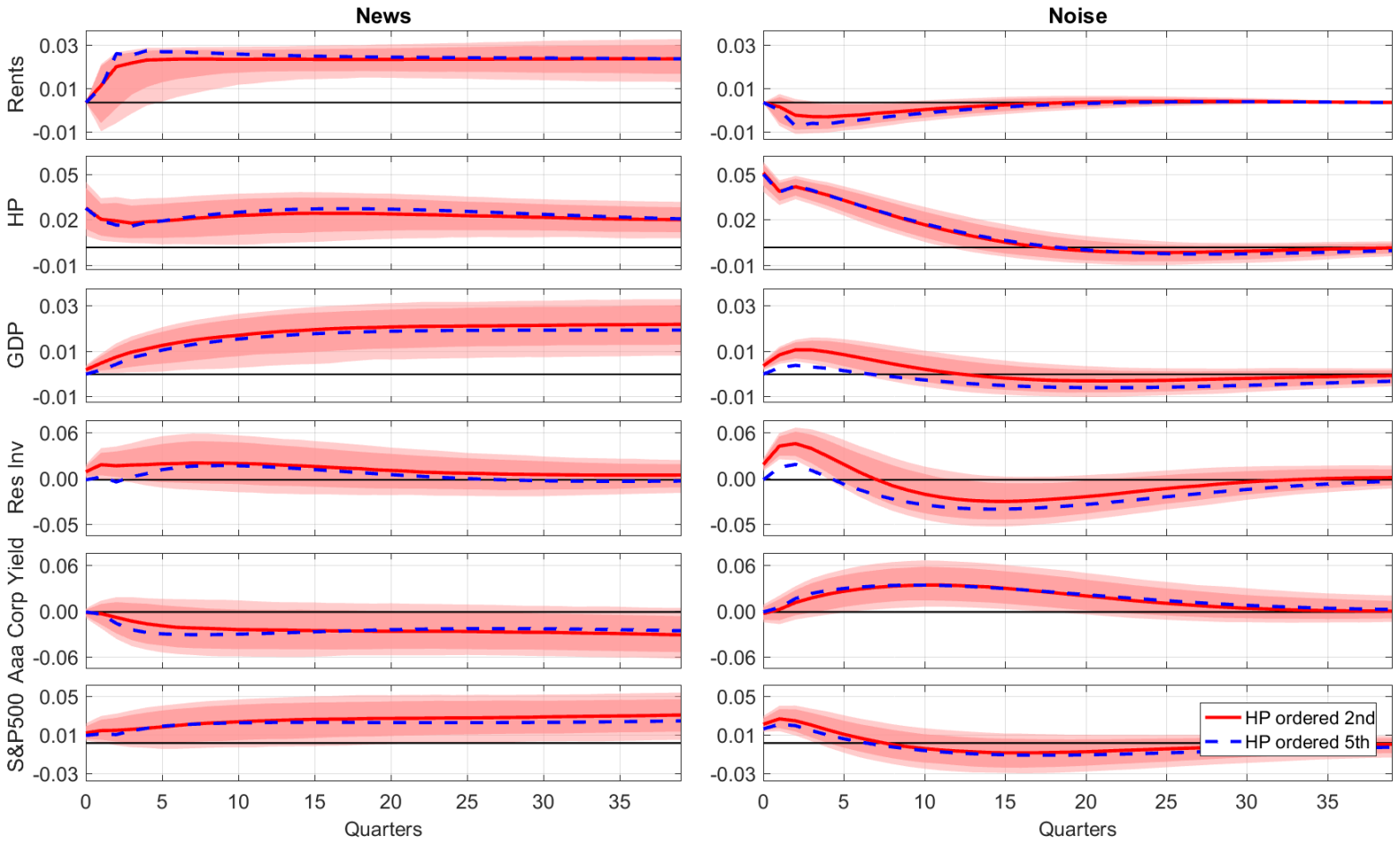

Figure 1.5: IRFs News and Noise Shocks - Macro Analysis

IRFs to news and noise shocks. The responses are reported in terms of the standard deviation of the variables in the system. The solid red line is the median, the dark and light red shaded areas represents 68% and 90% confidence bands respectively (2000 bootstrap replications). The shocks are identified through the following ordering: [Rents, Housing Prices, GDP, Residential Investment, Aaa Moody's Corporate Bond Yield, S&P Index]. The blue dotted line reports the median IRFs obtained by the recursive ordering [Rents, GDP, Residential Investment, Aaa Moody's Corporate Bond Yield, Housing Prices, S&P Index]. Sample: 1963:Q1 - 2016:Q3.

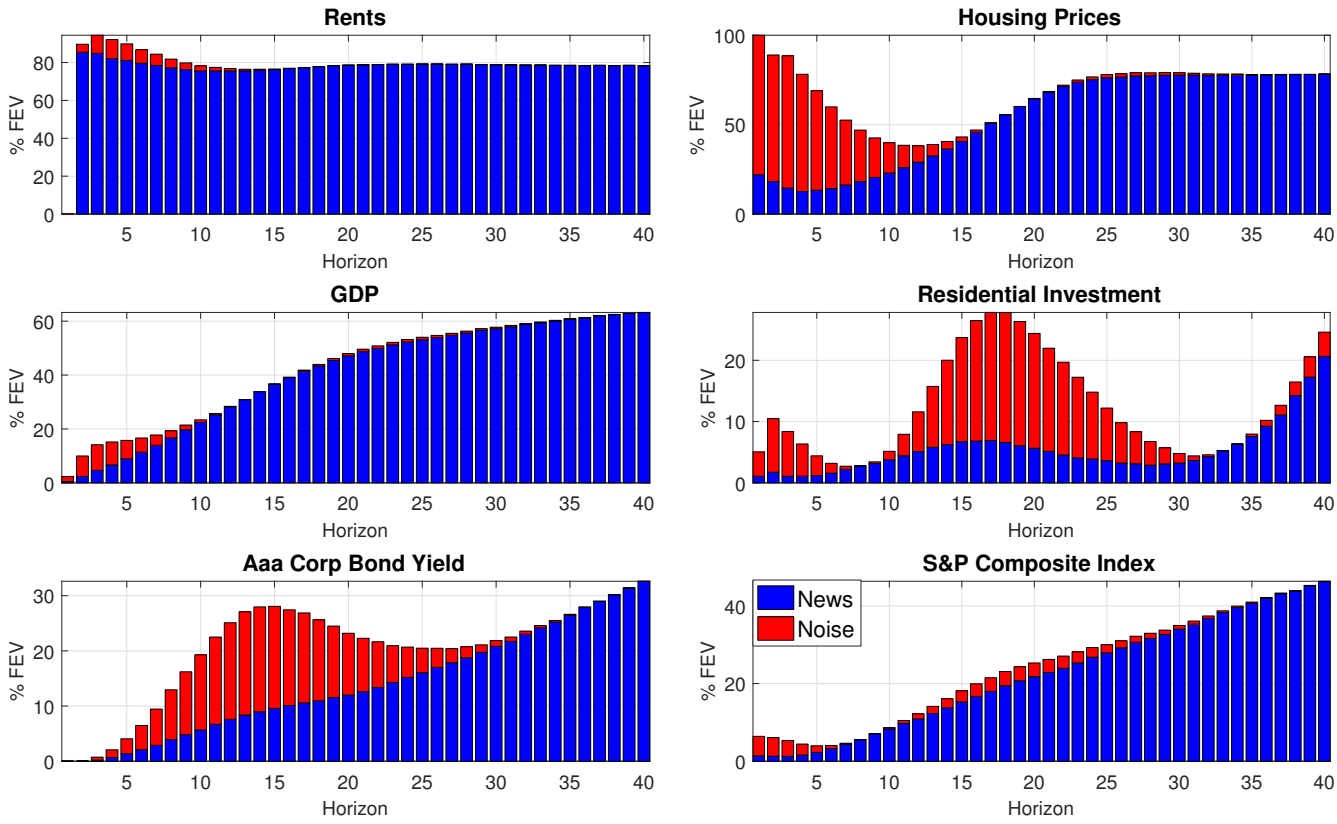

Figure 1.6: Forecast Error Variance Decomposition - Macro Analysis

Forecast error variance decomposition of the variables in the system. The plot display the share of the variance explained by news and noise at each horizon (not cumulatively). The shocks are identified through the following ordering: [Rents, Housing Prices, GDP, Residential Investment, Aaa Moody's Corporate Bond Yield, S&P Index]. Sample: 1963:Q1 - 2016Q3.

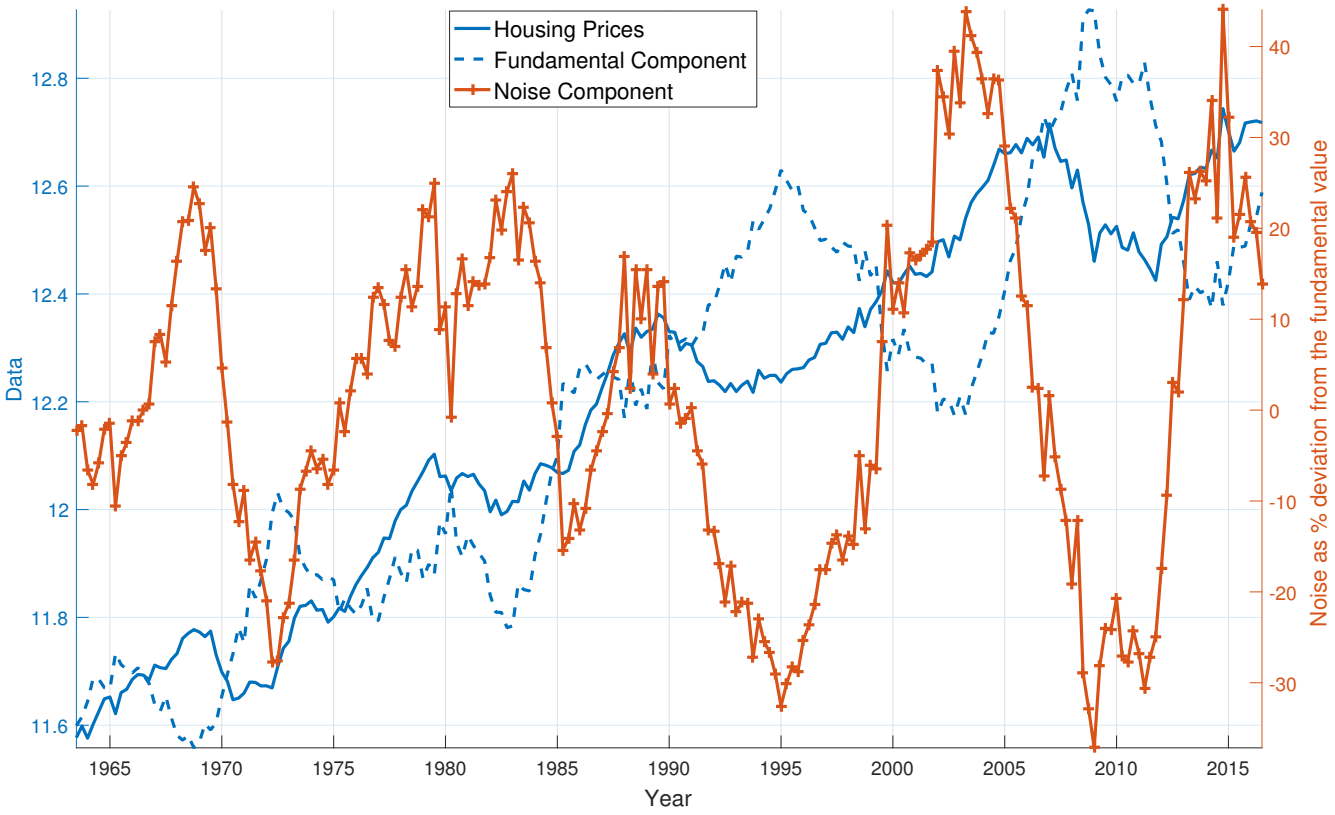

Figure 1.7: Historical Decomposition - Macro Analysis

Historical decomposition of housing prices (dotted blue) into a fundamental component (blue) and noisy component (orange). The shocks are identified through the following ordering: [Rents, Housing Prices, GDP, Residential Investment, Aaa Moody's Corporate Bond Yield, S&P Index]. Sample: 1963:Q1 - 2016Q3.

Chapter 2

Proxy-SVAR as a Bridge between Mixed Frequencies

Joint with Alejandro Vicondoa

2.1 Introduction

Macroeconomists increasingly incorporate information from financial markets, media, and the Web in their empirical analysis and models. The availability of this type of data, in particular from financial markets, allows researchers to draw information that was not available some years ago. Futures markets, for example, provide real-time information on expected policy decisions. Additionally, financial variables attract more attention due to the importance of recent financial-related events like the Great Recession or the European Sovereign debt crisis.

However, while macroeconomic aggregates are available only at the monthly or quarterly frequency, information from financial markets, media and Web is collected in real time or on a daily basis. When facing data sampled at different frequencies, the dominant approach still relies on temporal aggregation. The

variables sampled at higher frequencies are converted to the lowest sampling frequency.[1] In this procedure, many properties of the original series are lost. Of particular interest for macroeconomists, temporal aggregation exacerbates the simultaneity problem that generates identification challenges in structural Vector Autoregressions (SVARs). More specifically, impulse response functions are not invariant to time aggregation as both the contemporaneous covariance of the residuals and the parameters of the Wold representation change. Therefore, analyses which rely on temporal aggregation can be strongly biased (see Marcellino, 1999).[2]

Mixed frequency techniques have consequently attracted a growing interest in recent years. Mixed Data Sampling (MIDAS) and Mixed-Frequency Vector Autoregressions (MF-VARs) are two popular tools designed to deal with mixed frequency data (for a survey on the topic see Foroni, Ghysels, and Marcellino, 2013). Both, however, exhibit some shortcomings due to feasibility and computational constraints. For example, the mismatch in frequencies cannot be too wide and/or the number of high/low frequency variables cannot be too large. An alternative approach, originally developed to overcome identification challenges in VARs, actually constitutes a remedy for temporal aggregation biases. This methodology, called high frequency identification in Proxy-SVAR (HFI-PSVAR), identifies exogenous variations in high frequency variables around particular events and uses them as proxies for the structural shocks of interest (e.g. Gertler and Karadi, 2015). Essentially, the researcher exploits the proxy together with the reduced form residuals of a VAR to identify a shock of interest.[3] However, selecting key events for the phenomenon of interest is seldom straightforward and always arbitrary to a certain degree. Moreover, the Proxy-SVAR assumes that the proxy is orthogonal to the other structural shocks driving the system. Violations of this exclusion restriction would bias the analysis.

[1] This aggregation usually follows either skip-sampling or averaging. Skip-sampling, or point-in-time sampling, is usually applied to stocks. In this case, the variables available at the higher frequency are converted to the lower frequency simply by taking the last value within the low frequency period (for example the last monthly observation within a quarter). In the averaging case, the variables are averaged over the lower frequency period and then observed only once for each of those low frequency periods (for example, the quarterly average of monthly data).

[2] Intuitively, the severity of the simultaneity problem that we face in time series analysis is decreasing with the sampling frequency. At the extreme, temporal aggregation can introduce simultaneity where there is none. Consider for example a monetary policy setup. By aggregating the daily interest rate to the monthly frequency, the interest rate series will incorporate the endogenous reaction of the central bank to the daily changes in (for example) inflation expectations, which occurred within the month.

[3] This identification can be intuitively interpreted as an instrumental variable approach to VARs.

In this paper, we propose a new methodology, labeled "*Bridge Proxy-SVAR*", that links data sampled at different frequencies, i.e. high frequency and low frequency variables, through the Proxy-SVAR.[4] First, we identify the structural shock of interest in high frequency (HF) systems which are not subject to time aggregation and so characterized by less severe identification challenges (simultaneity). Second, we aggregate the series of shocks at the lower frequency, e.g. monthly or quarterly for macro variables. Third, we use the aggregated series of shocks as a proxy for the corresponding structural shock at this lower frequency (LF). Namely, we draw identifying restrictions for the LF representation from HF information. Our methodology builds upon a crucial proposition: identification prior to temporal aggregation is superior to identification post temporal aggregation. We illustrate that this proposition holds analytically in a tractable case. In a bivariate setup where the frequency mismatch is two, we prove that, if the HF shocks are correctly identified, our methodology recovers the correct impact matrix. Monte Carlo experiments generalize the test of the methodology to a variety of cases and data generating processes (DGPs). In evaluating the performances, we focus on the Impulse Response Functions (IRFs) that summarize the relevant information from the estimation of VARs. Importantly, the Monte Carlo simulations also allow us to compare the *Bridge Proxy-SVAR* with the common naive practice of time aggregation (LF-VAR) and with the best possible estimation (HF-VAR). In the LF-VAR, HF variables are introduced as time aggregated so all the available information is compressed at LF. The HF-VAR, instead, is a counter-factual estimation where the LF variables are observable at HF. As such, the HF-VAR also provides the upper bound for the performances of the MF-VAR.

Our results show that the *Bridge Proxy-SVAR* (*Bridge*) is a suitable method for approximating the true underlying responses under different data generating processes. First, the *Bridge* greatly outperforms the LF-VAR in all case and yields similar but less precise estimates to the HF-VAR. Second, our procedure can be applied in a simple manner, without computational burdens, even when the dimensionality of the system is large and when the frequency mismatch is wide. Third, we apply our methodology to assess the effect of monetary policy shocks in the US. Our benchmark is Gertler and Karadi (2015) as they apply the Proxy-SVAR.

[4] In what follows, we consider a standard VAR for the high frequency estimation but the analysis can apply any econometric model more suitable for high frequency data. What matters is the identification of an unpredictable shock, orthogonal to other components.

Their proxy consists of the series of monetary policy surprises built by Gurkaynak, Sack, and Swanson (2005). While this identification exploits key events for monetary policy, i.e. Federal Open Market Committee (FOMC) meeting days, we do not impose a priori any special role for these dates. Nonetheless, we find ex-post that the *Bridge* identifies shocks that are abnormally sizable on FOMC meeting days *vis-à-vis* non-FOMC days. Our series of shocks produces similar macroeconomic effects to those found in Gertler and Karadi (2015). Moreover, the monetary policy shocks we identify are immune to some criticisms posed in the literature on Gertler and Karadi (2015). This is related to the structural identification we employ and to the wide information set included in our HF-VAR. Finally, within our framework we can naturally take a further step consistent with the most recent works on monetary policy. In particular, Gertler and Karadi (2015) capture two distinct components on the path of interest rates, current and future, in their measure of monetary policy surprises, with opposite macroeconomic effect in the pre-crisis sample due to a strong informational content associated with shocks to the future rate.

This paper originated from a companion (applied) paper: Gazzani and Vicondoa (2016), which disentangles the macroeconomic effects of liquidity shocks in the Italian sovereign debt market. Two main identification challenges characterized that setup: first, the daily sampling frequency of financial variables opposed to the monthly frequency of macroeconomic aggregates; second, a severe simultaneity characterizes liquidity and default risk. Moreover, other shocks can contemporaneously hit indicators of liquidity and default risk, further complicating the analysis. In this framework, MF techniques proved to be unfeasible due to frequency and dimensionality reasons. In particular, the daily frequency is crucial for financial variables and we did not want to rely on aggregation to the weekly frequency.[5] Additionally, the inclusion of other financial indicators was necessary to define a sufficient information set. In that setup, we could not find convincing events that could be exploited for identification through the Proxy-SVAR.[6]

[5]This data transformation is usually applied to employ financial and macroeconomic data in a MF setup.

[6]Obviously, relevant events are available for the European Sovereign debt crisis (e.g. narrative events). However, they are not convincing for identification because there are no events that are mainly related with liquidity but not with default risk. A proxy build from this type of events, which is practically used as an instrument for the identification, would not satisfy the exclusion restriction being correlated with other structural shock (in particular default risk shocks).

Instead, we developed the *Bridge Proxy-SVAR*: estimate a daily VAR and identify liquidity shocks in the Italian sovereign debt market, aggregate this series of shocks to the monthly frequency, and use this monthly series as a proxy for the liquidity shocks in a monthly VAR including macroeconomic variables.

The severity of temporal aggregation biases in VAR models is illustrated in Marcellino (1999) and Foroni and Marcellino (2016). MF-VARs are the standard tools to handle data sampled at different frequencies. There are two main approaches to estimating VARs with mixed frequency data. The most popular one, developed by Zadrozny (1988), is based on a state space representation (a dynamic linear model). The system is driven by latent shocks whose economic interpretation is not straightforward. The presence of latent shocks implies that the Forecast Error Variance Decomposition (FEVD) of the system cannot be computed. Some examples of this approach include Mariano and Murasawa (2010), Schofheide and Song (2013), and Foroni, Ghysels, and Marcellino (2013). From a Bayesian perspective, Eraker et al. (2015) and Bluwstein and Canova (2015) estimate the state space representation via Gibbs sampler.[7] The second approach, proposed by Ghysels (2016), is more similar to standard VARs in being driven only by observable shocks. Contrary to model based on a state space representation, all the usual VAR tools are at the researcher's disposal. This particular VAR deals with series sampled at different frequencies through stacking: a HF variable is decomposed into several LF variables and directly employed in the VAR. For example, a monthly variable is introduced as three stacked series in a quarterly model. The shortcoming consists of the curse of dimensionality, i.e. parameters proliferation. Moreover, recovering the HF structural shocks from those in the stacked LF-VAR is not necessarily straightforward. Importantly for structural analyses, Anderson et al. (2016a) and Anderson et al. (2016b) study conditions for identifiability of the HF representation of VARs from mixed frequency data.

Although MF-VARs are powerful tools that suit many analyses, they may not be applicable in some cases. For example, the MF-VAR may not be a feasible approach when the mismatch between high and low frequency variables is large (e.g. 30 in the case of monthly-daily data). Additionally, also the dimensionality of the system can be problematic. In fact, the stacked MF-VAR presents parameter proliferation problems when the researcher has to include many HF variables.

[7] Some work as Angelini, Banbura, and Runstler (2010) have extended the mixed frequency state space representation to Factor models.

Computational problems may arise in the state space MF-VAR when there are many unobservable states (LF variables).

The *Bridge Proxy-SVAR* is a useful alternative in these cases, since it provides relevant computational advantages over the MF-VAR in terms of frequency mismatch and dimensionality. On the other hand, the MF-VAR is a different econometric model that improves, over a LF-VAR, the VAR estimates of both the autoregressive matrix and the impact matrix of the shocks.[8] The *Bridge Proxy-SVAR* only improves the impact matrix through information external to the LF-VAR, but still relies on the same autoregressive matrix of the LF-VAR. Additionally, the MF-VAR can assess the response of a HF variable on a LF variable, while the *Bridge* focuses exclusively on the reversal. Finally, the *Bridge Proxy-SVAR*, as the method developed by Ghysels (2016), relies purely on observables and not on latent variables and shocks as opposed to the state space MF-VAR.

The Proxy-SVAR methodology, developed by Stock and Watson (2012) and Mertens and Ravn (2013), is a very recent development in the identification of SVAR. This method employs exogenous variations in one variable, which is included in the VAR system, as a proxy for the structural shock of interest. The proxy is assumed to be correlated with a structural shock of interest but orthogonal to other structural shocks. In practice, the proxy constitutes an instrument for the reduced form residuals of the VAR and is used for (partial) identification of the covariance matrix of the structural shocks. The clear advantage of this technique is that, as long as the proxy is a relevant and valid instrument, the identification relies on a much weaker set of assumptions than other identification schemes. For example, no assumptions are made on the contemporaneous relationship among the variables in the system. Moreover, Carriero et al. (2015) have shown through Monte Carlo experiments that the PSVAR is robust to measurement errors. Lunsford (2015) provides a characterization of the asymptotic statistical properties of the Proxy-SVAR estimator.[9] When the proxy is a strong (weak) instrument, the estimator for the impact of structural shocks is consistent (inconsistent and biased towards zero). Ludvigson, Ma, and Ng (2015) employ an iterative projection IV to jointly build multiple external instruments. Proxies are

[8] Respectively, the A and B matrices in eq.(2.3).

[9] In Jentsch and Lunsford (2016) the performances of different bootstrapping techniques are compared for the Proxy-SVAR. The suggest that the moving block bootstrap is the best option.

usually built from a narrative description of policy decisions[10] or exploiting high frequency identification around some key events as in the already mentioned case of Gurkaynak, Sack, and Swanson (2005) and Gertler and Karadi (2015).

The *Bridge Proxy-SVAR* generalizes the HFI-PSVAR to those cases where there are no key events or when their selection is troublesome and arbitrary. The advantage of this methodology lies in the high frequency identification that may be cumbersome at low frequencies. At the same time, the high frequency shocks are used to instrument the reduced form residuals (prediction errors) of a LF-VAR. Intuitively, the *Bridge* always employs more information than a naive LF-VAR. Our approach remotely resembles the bridging equations which link data available at different frequencies through linear regression to produce nowcast and short-term forecast; e.g. Baffigi, Golinelli, and Parigi (2004) and Diron (2008). However, we exclusively focus on structural analysis and employ an instrumental variable approach.

After weighing pros and cons of our methodology versus the existing alternatives, we regard the *Bridge* as a particularly suitable tool for structural analysis on macro-financial linkages.

The remainder of this paper is organized as follows. Section 2.2 describes the *Bridge Proxy-SVAR* methodology. Section 2.3 presents the Monte Carlo experiments employed for testing. In Section 2.4, we apply the *Bridge* to study monetary policy in the US. Finally, Section 2.5 concludes.

2.2 Methodology

We introduce our methodology by summarizing the Proxy-SVAR identification (Section 2.2.1). In Section 2.2.2, we explain the steps that constitute the *Bridge Proxy-SVAR* methodology. First, we provide a general description of the identification. Second, an illustrative example shows how the *Bridge* can recover the correct impact matrix B in the VAR representation. On the other hand, when working with temporally aggregated data (LF-VAR) even the correct identification scheme cannot recover the true B matrix.

[10] See for example Stock and Watson (2012), Mertens and Ravn (2013) and Mertens and Ravn (2014).

2.2.1 Proxy-SVAR

Consider the simplest possible VAR representation:

$$Y_t = AY_{t-1} + u_t \qquad u_t \sim \mathcal{N}(0, \Sigma_u) \tag{2.1}$$

where Y_t and u_t are respectively n-dimensional vectors of endogenous variables and reduced form residuals with variance-covariance matrix Σ_u. The objective is to recover the structural form of the VAR, characterized by the vector of structural shocks $\varepsilon_t = B^{-1}u_t$:

$$Y_t = AY_{t-1} + B\varepsilon_t \qquad \varepsilon_t \sim \mathcal{N}(0, \mathcal{I}) \tag{2.2}$$

Let us consider a bivariate VAR system, where X may represent a collection of variable and not necessarily a single variable:

$$\begin{bmatrix} X_t \\ y_t \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^X \\ \varepsilon_t^y \end{bmatrix} \tag{2.3}$$

The Proxy-SVAR is an identification strategy that partially identifies the unknown B matrix. Namely, $\begin{bmatrix} B_{12} \\ B_{22} \end{bmatrix}$ represent the impact response (IRFs) of the system to a structural innovation in the variable y. The Proxy-SVAR exploits the external information to the VAR system contained in z_t. z_t is assumed to be a proxy for, at least, a component of the true ε_t^y with the following (instrumental variable) properties:

$$\begin{aligned} \mathbb{E}\left[\varepsilon_t^y z_t\right] &= \mu \neq 0 \\ \mathbb{E}\left[\varepsilon_t^X z_t\right] &= 0 \end{aligned} \tag{2.4}$$

From the conditions in eq.(2.4), it directly follows that B_{11} is identified up to a scale-sign factor:

$$\mathbb{E}\left[u_t^y z_t\right] = \mathbb{E}\left[\left(B_{22}\varepsilon_t^y + B_{21}\varepsilon_t^X\right) z_t\right] = B_{22}\mu \tag{2.5}$$

In a similar fashion,

$$\mathbb{E}\left[u_t^X z_t\right] = \mathbb{E}\left[\left(B_{12}\varepsilon_t^y + B_{11}\varepsilon_t^X\right) z_t\right] = B_{12}\mu \tag{2.6}$$

The unknown parameter μ represents the share of the information in ε^y captured by z_t. B_{22} can be recovered only if μ is known, which in practice reflects the assumption $\mu = 1 \Rightarrow z_t = \varepsilon^y$. Otherwise, we cannot uniquely identify B_{22} and, as a consequence, B_{12} either. However, μ does not affect the ratio

$$\frac{B_{12}\mu}{B_{22}\mu} = \frac{B_{12}}{B_{22}} \tag{2.7}$$

meaning that B_{12} is identified up to B_{22}. We can interpret this procedure through an instrumental variable approach, in particular as two stages least squares (2SLS):

First Stage: regress u_t^y on z_t that yields $\hat{\beta}_I = B_{22}\mu$ and $\hat{u}_t^y = \widehat{B_{22}\mu} z_t$

Second Stage: regress u_t^X on $\hat{u}_t^y$ where $\hat{\beta}_{II} = \frac{B_{12}}{B_{22}}$ by applying the definition of OLS.

The IRFs to ε^y are then computed across different horizons as:

$$\mathcal{IRF}_0^X = \frac{B_{12}}{B_{22}} \tag{2.8}$$

$$\mathcal{IRF}_n^X = A^{n-1}\mathcal{IRF}_{n-1}^X \quad \forall n > 0 \tag{2.9}$$

2.2.2 Bridge Proxy-SVAR

Traditionally, studies on monetary and fiscal policy have exploited narrative series or key events for identification. Such a strategy is hardly extendable to other areas of research. We therefore propose a more general and structural approach that employs HF information and, in this way, attenuates the time aggregation bias (see Section 2.2.3). Unlike the literature on mixed frequency, we do not model jointly the relationship between HF and LF variables, instead we exploit HF information to draw identification restrictions for the LF-VAR. As we show in Section 2.2.3.1, our approach exploits the superiority of identification prior to temporal aggregation over identification post temporal aggregation. First of all, we describe the steps in the *Bridge Proxy-SVAR* identification.

1. Define two VARs:

   (a) The first VAR, labeled High Frequency VAR (HF-VAR), incorporates the high frequency variables relevant for the analysis (e.g. financial daily).

It includes the variable of interest y and all the other variables necessary for the identification of the shocks. We define this collection of other variables as the information set Ψ. Potentially, the researcher can use other (more appropriate, depending on the case) econometric models for HF data. Moreover, the applied identification scheme should follow from economic theory.[11] If these conditions are satisfied, then $\hat{\varepsilon}_t^y \approx \varepsilon_t^y$.

(b) The second VAR, defined Low Frequency VAR (LF-VAR), includes variables at lower frequency. It features presumably macroeconomic aggregates and the variable y_t aggregated at lower frequency y_τ either by skip-sampling or averaging. The estimation of the LF-VAR yields the reduced form residuals $u_\tau = \begin{bmatrix} u_\tau^X & u_\tau^y \end{bmatrix}'$.

2. Aggregate the shocks estimated at HF to the LF:

$$\begin{aligned} z_\tau &= \frac{1}{m} \sum_{i=t}^{t+m} \hat{\varepsilon}_i^y \quad \text{averaging time aggregation} \\ z_\tau &= \hat{\varepsilon}_{mt}^y \quad \text{skip-sampling time aggregation} \end{aligned}$$

where m is the number of HF periods contained in a LF frequency period. If all sub-periods are the same then, in the averaging case, the correct aggregation scheme is actually given by $z_\tau = \hat{\varepsilon}_t^y$ (the shock in the first HF sub-period). If the assumptions in (1a) are satisfied, then, by construction, the proxy is exogenous $\mathbb{E}\left[\varepsilon_t^X z_t\right] = 0$ and relevant $\mathbb{E}\left[\varepsilon_t^y z_t\right] \neq 0$.

3. Use z_τ as a proxy for the structural shock of interest: instrument u_τ^y with z_τ and estimate the impact effect of a shock in y. This means that we are identifying the second column in the B matrix in eq.(2.2). We can see this procedure as 2SLS or directly as IV:

$$\begin{aligned} B_2 &= \left(z_\tau' u_\tau^y\right)^{-1} z_\tau' u_\tau \\ &= \begin{bmatrix} \mu B_{22} & \mu B_{12} \end{bmatrix} \\ &= \begin{bmatrix} 1 & B_{22}^{-1} B_{12} \end{bmatrix} \end{aligned} \tag{2.10}$$

[11] The higher the frequency at which they are imposed, the less identifying restrictions constrain the data and the more they are likely to hold.

so that the impact response to ε_τ^y is identified up to the impact effect on y itself. If we are confident that $\hat{\varepsilon}_t = \varepsilon_t$, then $\mu = 1$ and we can estimate the size of the shock from the standard deviation of the series obtained from the first stage regression.

Notice that the assumption in point 1, $\hat{\varepsilon}_t^y \approx \varepsilon_t^y$, is far more stringent than what we actually need. In fact, assume that the structural shock of interest can be decomposed as a sum of two orthogonal *iid* components, weighted by the scalars μ_1, μ_2:

$$\varepsilon_t^y = \mu_1 \varsigma_t + \mu_2 \phi_t \qquad \varsigma_t \perp \phi_t \tag{2.11}$$

As explained in Section 2.2.1, the PSVAR partially identifies the B matrix and consequently we need to recover only a component of the HF shock ε_t^y, for example ς_t. Once again, this feature resembles a standard IV case where we exploit an exogenous variation in a variable of interest and not the whole exogenous variation. Recall indeed that eq. (2.4) does not assume the correlation being equal to 1, but only different from 0.

Next, we analyze how the Bridge Proxy-SVAR deals with data sampled at mixed frequencies. Starting from a general case, we move to a tractable example where, if a component of the structural shocks is correctly identified at HF, our proxy recovers the correct true impact matrix B.

2.2.3 Time Aggregation

As a first step, following Foroni and Marcellino (2016), we illustrate the most general formulation. The objective of the analysis is to recover the IRF of the VAR system to a shock in the HF variable. The common practice consists of transforming the HF (indexed by t) at LF (indexed by τ) and running a VAR on time aggregated data. For the sake of simplicity, we consider a stationary case without deterministic components:

$$\begin{aligned} Y_t &= A(L)Y_t + B\varepsilon_t \qquad && \varepsilon_t \sim \mathcal{N}(0, \mathcal{I}),\ t = 1, 2, ..., T \\ [I - A(L)]\, Y_t &= B\varepsilon_t && \varepsilon_t \sim \mathcal{N}(0, \mathcal{I}) \end{aligned} \tag{2.12}$$

Time aggregation is generally a two-step filter. First, the data is transformed through the filter $w(L)$ and, second, the series is made observable only every

m periods through the filter $D(L)$. We consider the time aggregated representation under skip-sampling (or point-in-time sampling) since average sampling introduces an higher order MA component that further complicates the analysis. Nonetheless, we report in Appendix B.3 the same derivations for the averaging scheme and show that similar results hold in our Monte Carlo simulations. In the skip-sampling case, the filter $w(L) = 1$ does not produce any change. We apply the filter $D(L) = I + AL + ... + A^m L^m$ so that the researcher can observe certain variables only once every m periods:

$$\begin{aligned} D(L)\left[I - A(L)\right] Y_t &= D(L)B\varepsilon_t \\ Y_\tau &= C(L)Y_\tau + Q(L)\varepsilon_t \qquad \varepsilon_t \sim \mathcal{N}(0, \mathcal{I}),\ \tau = mt, 2mt, ..., T \\ Y_\tau &= C(L)Y_\tau + \xi_\tau \qquad \xi_\tau \sim \mathcal{N}(0, \Omega) \end{aligned} \tag{2.13}$$

where $C(L) = D(L)A(L)$ and $Q(L) = D(L)B$. Ω is given by the squared contemporaneous elements in the $Q(L)$ matrix as the structural shocks are not autocorrelated. Time aggregation mixes different structural shocks at different times in ξ_τ.

2.2.3.1 An Illustrative Example

We focus now on a more specific case. We aim at assessing the effect of the shock in y, observable at HF, on x, available only at LF. x is time aggregated through skip-sampling. We consider a $VAR(1)$ representation and a mismatch between HF and LF equal to two, such that we can illustrate the methodology through simple algebra:

$$\begin{aligned} Y_t &= AY_{t-1} + B\varepsilon_t \qquad \varepsilon_t \sim \mathcal{N}(0, \mathcal{I}) \\ (I - AL)\, Y_t &= B\varepsilon_t \qquad \varepsilon_t \sim \mathcal{N}(0, \mathcal{I}) \end{aligned} \tag{2.14}$$

To move to the time aggregated representation (under skip-sampling), we apply the filter $D(L) = I + AL$:

$$\begin{aligned} D(L)\left(I - AL\right) Y_t &= D(L)B\varepsilon_t \\ \left(I - A^2L^2\right) Y_t &= (I + AL)B\varepsilon_t \\ Y_\tau &= CY_{\tau-1} + \xi_\tau \qquad \xi_\tau \sim \mathcal{N}(0, BB' + ABB'A') \\ Y_\tau &= CY_{\tau-1} + Q(L)\varepsilon_t \qquad \varepsilon_t \sim \mathcal{N}(0, \mathcal{I}) \end{aligned} \tag{2.15}$$

where $C = A^2$ and $Q(L) = (B + ABL)$. Let us consider the system in extended notation in terms of the reduced form residuals u_t:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} u_t^x \\ u_t^y \end{bmatrix} \tag{2.16}$$

In particular, assume that $B = \begin{pmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{pmatrix}$ so that we are in the standard Cholesky case, as in Foroni and Marcellino (2016):

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^x \\ \varepsilon_t^y \end{bmatrix} \tag{2.17}$$

The temporally aggregated system is given by:

$$\begin{bmatrix} x_\tau \\ y_\tau \end{bmatrix} = \begin{bmatrix} a_{11}^2 + a_{12}a_{21} & a_{11}a_{12} + a_{12}a_{22} \\ a_{11}a_{21} + a_{21}a_{22} & a_{12}a_{21} + a_{22}^2 \end{bmatrix} \begin{bmatrix} x_{\tau-1} \\ y_{\tau-1} \end{bmatrix} + \begin{bmatrix} \xi_\tau^x \\ \xi_\tau^y \end{bmatrix} \tag{2.18}$$

where

$$\begin{bmatrix} \xi_\tau^x \\ \xi_\tau^y \end{bmatrix} = \begin{bmatrix} b_{11}\varepsilon_t^x + (a_{11}b_{11} + a_{12}b_{21})\,\varepsilon_{t-1}^x + a_{12}b_{22}\varepsilon_{t-1}^y \\ b_{21}\varepsilon_t^x + b_{22}\varepsilon_t^y + (a_{21}b_{11} + a_{22}b_{21})\,\varepsilon_{t-1}^x + a_{22}b_{22}\varepsilon_{t-1}^y \end{bmatrix} \tag{2.19}$$

In the temporal aggregation case, biases arise even if the identification exploits the correct Cholesky decomposition of the variance-covariance matrix of the reduced form residuals. The problem originates from the variance-covariance matrix observable at LF: $\Omega = BB' + ABB'A'$ which is different from the true BB'. Intuitively, in the LF-VAR the zero restriction constrains ε_t^y to have a zero effect over x for m periods instead of one (in this simple case $m = 2$). An analytical illustration of the time aggregation bias is reported in Appendix B.2.

Instead of imposing identification restrictions directly on the LF representation, we suggest identifying structural shocks from a HF system, which is not subject to temporal aggregation biases. The (temporally aggregated) structural shocks can be then employed to draw identifying assumptions in the LF-VAR representation. As the variable x is not directly observable at HF, the goodness of the identification is increasing in the amount of information included in the HF-VAR (Ψ). Moreover, Ψ should contain all the variables necessary to achieve a

correct identification at this HF stage, which depends on the specific cases under examination.

In this stylized example, the HF system in the observables, assumed to be again $VAR(1)$, can be express in blocks as:

$$\begin{bmatrix} \Psi_t \\ y_t \end{bmatrix} = \Gamma \begin{bmatrix} \Psi_{t-1} \\ y_{t-1} \end{bmatrix} + \Phi \begin{bmatrix} \varepsilon_t^{\Psi} \\ \varepsilon_t^{y} \end{bmatrix} \tag{2.20}$$

The correct identification is fully achieved if x_t is spanned by the collection of variables that constitute the HF system and the LF-VAR (lagged):[12]

$$x_t \in \text{span}\left\{\Psi_t, \Psi_{t-1}, y_{t-1}, x_{\tau-1}\right\} \tag{2.21}$$

Intuitively, the Proxy-SVAR uses information contained both in the HF system and the LF-VAR. It is the union of these two information sets that has to provide enough information on the unobservable x_t to achieve the correct identification. For simplicity, assume that Ψ_t perfectly incorporates the information contained in x_t. In applied research, if the HF system consists of financial variables, such an assumption is motivated by financial markets incorporating all available information. Moreover, a wide literature studies the reaction of financial markets to macroeconomic data releases. Imposing a recursive structure where y_t is ordered after Ψ_t yields the correct impact matrix B. In this way, identification restrictions do not rely on the temporally aggregated system but are drawn at HF.

Notice that, actually, we do not need to fully capture ε_t^y but only a component of it. In what follows, we assume that the proxy is given by a component of the true structural shock as defined in eq. (2.11). In order to be consistent with the skip-sampling temporal aggregation, we take the last HF shock within the LF interval:

$$z_\tau \;=\; \varsigma_t$$

[12] Notice that those are the necessary requirements to achieve the correct identification. In order to improve over the temporal aggregation practice, i.e. imposing restrictions directly on the LF-VAR representation, the conditions are much milder.

We can express the last stage in the *Bridge* either as a two stage least square ($2SLS$) estimation or directly as IV. In the $2SLS$ case, we use z_τ it in the first stage regression

$$\xi_\tau^y = \beta_{1s} z_\tau + \eta_\tau \qquad \eta_\tau \sim WN$$

where η is the error term, assumed to follow the distribution $iid\,\mathcal{N}(0,\sigma^2)$.[13]

The estimated coefficient from the first stage is:

$$\begin{aligned}
\hat{\beta}_{1s} &= \mathbb{E}\left[z_\tau' z_\tau\right]^{-1} \mathbb{E}\left[z_\tau' \xi_\tau^y\right] \\
&= \frac{\mathbb{E}\left[\varsigma_t \left(b_{21}\varepsilon_t^x + b_{22}\varepsilon_t^y + (a_{21}b_{11} + a_{22}b_{21})\,\varepsilon_{t-1}^x + a_{22}b_{22}\varepsilon_{t-1}^y\right)\right]}{\mathbb{E}\left[\varsigma_t \varsigma_t\right]} \\
&= \mu_1 b_{22}
\end{aligned} \tag{2.22}$$

If we employ the whole shock ε_t^y, then $\hat{\beta}_{1s} = b_{22}$ which is the true parameter in the HF representation. Notice that both requirement for a proxy are satisfied:

$$\begin{aligned}
\mathbb{E}\left[\xi_\tau^y z_\tau\right] = \hat{\beta}_{1s} = \mu_1 b_{22} &\neq 0 \qquad && \textbf{IV relevance} \\
\mathbb{E}\left[\varepsilon_t^x z_\tau\right] &= 0 && \textbf{IV validity} \text{ (by construction)}
\end{aligned} \tag{2.23}$$

The fitted value from the first stage are given by:

$$\hat{\beta}_{1s} z_\tau = \mu_1 b_{22} \varsigma_t \tag{2.24}$$

The second stage regression reads

$$\xi_\tau^x = \beta_{2s}\left(\hat{\beta}_{1s} z_\tau\right) + \varphi_\tau \qquad \varphi_\tau \sim WN \tag{2.25}$$

$$\begin{aligned}
\hat{\beta}_{2s} &= \mathbb{E}\left[\left(\hat{\beta}_{1s} z_\tau\right)\hat{\beta}_{1s} z_\tau\right]^{-1} \mathbb{E}\left[\hat{\beta}_{1s} z_\tau \xi_\tau^x\right] \\
&= \mathbb{E}\left[\mu_1 b_{22} \varsigma_t^2\right]^{-1} \mathbb{E}\left[\varsigma_t \xi_\tau^x\right] \\
&= \left(\mu_1 b_{22}\right)^{-1} \mathbb{E}\left[\varsigma_t \left(b_{11}\varepsilon_t^x + (a_{11}b_{11} + a_{12}b_{21})\,\varepsilon_{t-1}^x + a_{12}b_{22}\varepsilon_{t-1}^y\right)\right] \\
&= 0
\end{aligned} \tag{2.26}$$

[13] Henceforth, white noise (WN) will point at the error term in simple OLS equations, assumed to be distributed as $iid\,\mathcal{N}(0,\sigma^2)$ and uncorrelated with the independent variables.

meaning that the *Bridge* correctly recovers the Cholesky structure of the innovations. We obtain an equivalent result if we apply straight the definition of IV estimator:

$$\begin{aligned}\hat{\beta}_{Proxy} &= \mathbb{E}\left[z_\tau \xi_\tau^y\right]^{-1} \mathbb{E}\left[z_\tau \xi_\tau^x\right] \\ &= \frac{\mathbb{E}\left[\varsigma_t \left(b_{11}\varepsilon_t^x + \left(a_{11}b_{11} + a_{12}b_{21}\right)\varepsilon_{t-1}^x + a_{12}b_{22}\varepsilon_{t-1}^y\right)\right]}{\mathbb{E}\left[\varsigma_t \left(b_{21}\varepsilon_t^x + b_{22}\varepsilon_t^y + \left(a_{21}b_{11} + a_{22}b_{21}\right)\varepsilon_{t-1}^x + a_{22}b_{22}\varepsilon_{t-1}^y\right)\right]} \\ &= 0 \end{aligned} \tag{2.27}$$

Through this tractable case, we have shown analytically that the *Bridge* recovers the true impact matrix, whereas the correct Cholesky ordering imposed at LF introduces biases. The magnitude of these differences in a more general setup can only quantified through Monte Carlo experiments, presented in Section 2.3. Furthermore, we also test the robustness of the methodology to misspecifications and to limited information in the HF system and LF system employed by the *Bridge* (omitted variables).

2.3 Monte Carlo Experiments

Our design is similar to Foroni and Marcellino (2016) who compare the finite sample performances of the HF-VAR, LF-VAR (time aggregated), and the MF-VAR. In the latter, one variable is unobservable at high frequency but the econometrician only observes one out of three observations. We run the same experiment but we substitute the MF-VAR with the *Bridge*. Notice that the HF-VAR constitutes a "counter-factual" first best and an upper bound for the performances of the MF-VAR. Temporal aggregation follows skip-sampling, while in Appendix B.3 we report the main results under the averaging temporal aggregation scheme. We focus on the IRFs that summarize the relevant information on the estimation of the system. To be able to compare the IRFs under HF and LF representation, the IRFs at HF have to be treated in a consistent manner with the temporal aggregation scheme applied to the data.

The benchmark outline of the experiment is the following: we consider a VAR(1) DGPs and, for thirteen representative parametrizations, generate 1000 replications of 3000 HF observations. In a first step, the frequency mismatch is three, so that at LF we dispose of 1000 observations. For the sake of synthesis, we

evaluate the performances of the three identifications through the lens of the *Mean Absolute Distance* (MAD) which measures the distance between the estimated and the true IRFs (cumulated over 8 horizons). For each replication, we compute the MAD and then we average over the whole set of replications.

The analysis begins with a stylized case that highlights the time aggregation bias alone. Then, one step at the time, we include further elements resembling the identification challenges that economists face in applied research.

2.3.1 Pure Time Aggregation

The LF-VAR and the *Bridge* temporally aggregate information in antithetical ways. In a LF-VAR, the aggregation occurs before identification while the *Bridge* identifies structural shocks at HF and then compresses them at LF. We are implicitly comparing the performances under these two temporal aggregation schemes.

The DGP follows the structure:

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \rho & \delta_l \\ \delta_h & \rho \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} e_t^y \\ e_t^x \end{pmatrix} \tag{2.28}$$

where $\begin{pmatrix} e_t^x \\ e_t^y \end{pmatrix} \sim \mathcal{N}(0, I_2)$. Basically, the innovations follow a recursive ordering structure that we correctly apply with the HF, LF and *Bridge*. We test 13 combinations of $\{\rho, \delta_l, \delta_h\}$ that represent different possible structures of the DGP.[14]

Figure 2.1-2.2 display an example of IRFs recovered with the three identifications. The HF-VAR and the *Bridge* perfectly recover the true IRFs, while the LF-VAR overestimates the size of the shock. Not surprisingly, Figure 2.3 points out that the HF identification is the best possible identification. An infinitesimal bias comes from the finite sample estimation of the HF-VAR system. The *Bridge*, which is by construction a second best option, performs very closely to the HF-VAR. Even if the *Bridge* and HF-VAR apply the same identification at HF, the *Bridge* is inefficient due to the two stages in the estimation. The comparison resembles the efficiency loss of the IV estimation with respect to OLS.

For nearly all cases, the *Bridge* recovers the IRFs with a smaller bias than the LF-VAR. Under few DGPs, the exception consists of the shock to the second variable y with zero impact on the first variable x. The zero restriction is imposed in the

[14] The parametrizations are reported in Appendix B.2.

case of the HF-VAR and LF-VAR, while it is estimated from the first stage in the case of the *Bridge*. Even if the median IRF is zero, the IRFs generated by the *Bridge* across the 1000 replications may slightly differ from 0 due to finite sample bias. As a result, when the MAD is generally very low, the *Bridge* may perform worse than the LF-VAR.

While we present the main results of the Monte Carlo under averaging time aggregation in Appendix B.1, Figure 2.4 provides an intuitive portrait of the biases arising from this alternative time aggregation scheme. Even if the correct recursive structure is imposed at LF on the variance-covariance matrix of the reduced form residuals, the restriction constraints three HF periods instead of one. As a result, the LF-VAR estimates strongly biased IRFs, whereas the *Bridge* correctly recover them.

2.3.2 Time Aggregation and Misspecification

In applied research, the econometrician does not know the true DGP and, indeed, the analysis aims at recovering information on it. In this light, the interaction between temporal aggregation and misspecification deserves attention. The DGP deviates from the recursive structure which, on the contrary, is still employed as identifying restriction by HF-VAR, LF-VAR and *Bridge*. Additionally, we consider two further issues: wider frequency mismatch and measurement error.

2.3.2.1 Contemporaneous Effects

The impact matrix features now all non-empty entries:

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \rho & \delta_l \\ \delta_h & \rho \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} 1 & c_1 \\ c_2 & 1 \end{pmatrix} \begin{pmatrix} e_t^x \\ e_t^y \end{pmatrix} \tag{2.29}$$

We present the results under $\{c_1, c_2\} = \{-0.3, 0.1\}$, but we have tested different combinations obtaining similar results. In this case, the *Bridge* closely resembles the performance of the HF-VAR whereas the LF-VAR leads to sizable biases (Figure 2.5).

2.3.2.2 Wider Frequency Mismatch and Measurement Error

First, we now turn to a case in which the mismatch between HF and LF is significantly wider, i.e. $m = 30$, which represent the monthly-daily case. Fig. B.7 compares the identifications over the 13 DGPs through the lens of MAD. The LF-VAR induces a much larger bias with respect to the HF-VAR and *Bridge*.[15]

Second, we test the impact of measurement errors without finding any severe effect for the *Bridge*, while LF-VAR suffers the most. The results reported in Fig. B.8 refer to a case in which the first variable in the system is affected by a sizable measurement error with standard error 0.3 (30% of the actual standard deviation of the structural shocks).

2.3.3 A Practical Case - One LF and Two HF variables

Let us turn now to a more practical case: we consider a situation in which the researcher observes two HF variables and one LF variable. x is observable only at LF, whereas y and z are available at HF. We are interested in studying how the shocks to the HF variables affect x (e.g. how financial shock affect the macroeconomy). Contrary to the previous MC exercises, in the first stage of the *Bridge* we use only the two HF variables. In the second stage, we will include all three variables (time aggregated). Once again, we compare the *Bridge* with the HF-VAR (counter-factual) and LF-VAR.[16]

$$\begin{pmatrix} x_t \\ z_t \\ y_t \end{pmatrix} \begin{pmatrix} \rho & \delta_{l,h_1} & \delta_{l,h_2} \\ \delta_{h_1,l} & \rho & \delta_{h_1,h_2} \\ \delta_{h_2,h_l} & \delta_{h_2,h_1} & \rho \end{pmatrix} \begin{pmatrix} x_{t-1} \\ z_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} 1 & c_{12} & c_{13} \\ c_{21} & 1 & c_{23} \\ c_{31} & c_{32} & 1 \end{pmatrix} \begin{pmatrix} e_t^z \\ e_t^z \\ e_t^y \end{pmatrix} \tag{2.30}$$

Again, under the many parametrization tried, we choose to present the results with

$$\begin{pmatrix} 1 & c_{12} & c_{13} \\ c_{21} & 1 & c_{23} \\ c_{31} & c_{32} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0.65 & 0.8 \\ 0.4 & 1 & 1 \\ 0.5 & 0.8 & 1 \end{pmatrix}$$

[15] Notice that the *Bridge* easily accommodates the daily-quarterly mismatch without relevant computational costs.

[16] In this case, we rely on the conservative identification that is described in Appendix B.1.

This parametrization represents the strong simultaneity among the variables observed at HF (financial variables). The same pattern of the previous exercises emerges also in this practical case (Figure 2.6). The HF identification of the *Bridge*, not subject to temporal aggregation biases, employs only a subset of the actual information. However, the missing variable is included in the LF-VAR representation whose reduced form residuals are instrumented in the second stage of the *Bridge*. Consequently, we are using a richer information set than the LF-VAR. Moreover, economists usually assume that financial markets incorporate with a negligible lag all available information. In empirical implementations, the *Bridge* is therefore unlikely to suffer from a problem of limited information at HF.

2.3.3.1 High Frequency not High enough?

A potential concern arises if the HF identification of the *Bridge* is implemented at the wrong frequency. For example, the correct analysis for financial phenomena could be though as intra-daily and not daily.[17] To address this concern, we test whether, by relying on a HF, which is not high enough, we can still mitigate time aggregation biases. We repeat the same exercise of Section 2.3.3 but, while the HF-VAR employs the correct frequency, the *Bridge* relies on mildly time aggregated data ($m = 3$). The LF-VAR estimation is based on aggregation over nine periods ($m = 9$). Figure B.9 depicts that the *Bridge* still attenuates the biases with respect to the LF-VAR.

2.3.4 Large Systems

Until now, we have studied the performances of different identifications in small systems with ad hoc parametrizations of the DGP. However, we know that many events (shocks) hit economies at the same time and financial markets take this information nearly instantly into account. To represent this situation, we consider a nine variables VAR as DGP. Moreover, in order to tackle any possible suspicion of DGP "self-selection", we randomly parametrized both the autoregressive matrix A and the impact matrix B. The only constraints that we impose ensure the stationary of the system and a mapping between variables and shocks.[18] From

[17] In case relevant data is available at intra-daily frequency, the econometrician can recover shocks at this frequency and link them with macro-variables through the *Bridge*. However, this procedure may induce noise coming from the micro-structure of the market

[18] Each shock impacts the corresponding variable more than other variables.

100 random parametrizations of the system, we generate 1000 data-points at LF across 1000 simulations.[19]

We run this large experiment over three dimensions:

1. the time aggregation scheme: (a) skip-sampling
   (b) averaging
2. information employed by the *Bridge*:
   (a) partial information at LF: the HF stage of the *Bridge* employs full information but the LF stage (and LF-VAR) do not include the last two variables in the system
   (b) full information at LF:
      i. full information at HF: all information is included both in the HF-VAR and in the LF-VAR employed by the *Bridge*
      ii. partial information at HF: the HF stage of the *Bridge* does not include the last two variables in the system
3. frequency mismatch: (a) quarterly-monthly ($m = 3$)
   (b) monthly-daily ($m = 30$)

The case (2b) is a robustness check similar to the practical case presented in Section 2.3.3. However, we do not expect it to be particularly severe if the HF system employs financial data.

The *Bridge* improves over the performances of the LF-VAR across all the cases (Table 2.1). MAD percentage gains over the LF-VAR vary between 10% and 73%. The gains are higher when the *Bridge* employs full information and under the averaging scheme. In the latter case the biases from time aggregation are larger. Figure 2.7 displays examples of a heat-map of the MAD over the three identifications for one of the 100 systems for all combinations of shocks and variables. The similar results of the *Bridge* compared to the HF-VAR stand out immediately. At the same time, the LF-VAR produce much worse estimates than the alternative methods. Figure 2.8 presents an example of IRFs. Even in this large system, the *Bridge* performs very closely to the HF-VAR and it subject only to a loss in precision. In conclusion, the *Bridge* greatly improves the performances of the analysis over the

[19] Similar results hold for 500 observations at LF.

naive practice of time aggregation and it is often close to the performances of a counter-factual HF-VAR. The more complete the information set is at HF in the *Bridge* identification, the closer the results of the *Bridge* to the HF-VAR. On the other hand, employing only partial information in the LF-VAR of the *Bridge* does not produce too severe losses in performances. In fact, the information omitted from the LF system does not affect the estimated B matrix but only the transmission of the shocks.

2.4 Application - Monetary Policy in the US

This section is devoted to an empirical application of our methodology. We choose a popular empirical question in order to have benchmarks for comparison: the macroeconomic effects of monetary policy shocks in the US. The related identification poses great problems due to various reasons and, in particular, due to two challenges. First, the *Federal Reserve* (FED) often changes the policy rate in response to current and expected economic conditions. Such responses cannot obviously be considered exogenous. Second, agents anticipate a large component of the changes in the policy rate (e.g. Vicondoa, 2016) and this anticipation can lead to VAR failures. Romer and Romer (2004) and Gertler and Karadi (2015) (RR and GK henceforth) employ two popular identification strategies and, consequently, constitute our reference points. RR mainly tackle the first challenge, analyzing US monetary policy through a narrative approach that takes into account the information contained in the Greenbook (FED forecasts). Their series of monetary shocks have been updated, among others, by Coibion et al. (2012). GK focus mainly on the second identification threat, using the series of monetary policy surprise built by Gurkaynak, Sack, and Swanson (2005) as a proxy to reach identification in a monthly VAR. Since GK employ the Proxy-SVAR, they are the most natural comparison for the *Bridge*. Gurkaynak, Sack, and Swanson (2005) measure monetary policy surprises as the change in the price of Fed Funds (FF) future contracts around FOMC meetings days. While they exploit these key events for monetary policy, we do not impose a priori any special role for these dates. Nonetheless, we find ex-post that the *Bridge* identifies shocks that are abnormally sizable on FOMC meeting days *vis-à-vis* non-FOMC days. Our series of shocks produces similar macroeconomic effects to those found in Gertler and Karadi (2015). Moreover, the monetary policy shocks we identify are immune to

some criticisms posed in the literature on Gertler and Karadi (2015). For example, our measure of monetary policy shock is orthogonal to changes in risk premia that may be captured by the FF futures. Finally, within our framework, we can easily decompose two components captured by GK and defined in Gurkaynak, Sack, and Swanson (2005) as two orthogonal factors: a "current federal funds rate target" factor and a "future path of policy" factor. The future component is not strictly a monetary policy shock since incorporates significant informational content. This finding is consistent with recent papers by Campbell et al. (2012) and Campbell et al. (2016) who introduced the distinction between Delphic and Odyssean forward guidance.[20]

Instead of focusing on particular events, we estimate a daily VAR on the sample 1991m1-2008m6 to avoid any issue related with the zero lower bound. The optimal number of lags based on the three most popular information criteria is 22. Notice that we may employ more refined econometric models, suitable for financial data, ranging from a VAR featuring stochastic volatility to a SVAR-GARCH.[21] Nonetheless, as we show in a few lines, a standard VAR suffices in this case. A daily analysis over such a long horizon offers vast degrees of freedom allowing us to include a large amount of variables to widen as much as possible the information set.

2.4.1 Romer & Romer and Gertler & Karadi

The Target Fed Fund Rate (TFFR) and the price of the FF Future contract 3 months ahead (FF4) constitute our monetary policy indicators. The TFFR allows us to resemble the analysis of RR while the FF4 corresponds to the analysis of GK. In the latter case, it is necessary to remove the TFFR from the HF-VAR in order to capture a mixture of shocks to the current and future path. We identify monetary policy shocks through a recursive ordering, placing our measure of monetary policy last. In other words, we regress the TFFR (FF4) on the lags and contemporaneous values of the other financial variables (plus the TFFR-FF4 own lags). This procedure orthogonalizes the reduced form residual in the TFFR (FF4) equation

[20] Lakdawala (2016) also studies the macroeconomic effect of current and future factors. However, in our case the decomposition does not exploit FOMC meetings explicitly and it is applied directly within our daily VAR. Moreover, the current factor identified in Lakdawala (2016) leads to a positive reaction of CPI that we find puzzling.

[21] See for example Lutkepohl and Milunovich (2015). Additionally, when using financial variables, identification itself can exploit changes in volatility.

from all innovations in other financial variables. In this way, we ensure that we clean our measure of monetary policy shocks from other innovations in the system occurring in the same day. In an intuitive fashion, we define as monetary policy shocks the new information that enters the system at time t uniquely through our measures of monetary policy (TFFR-FF4). In this way, we clean the residuals in the FF4 from non-monetary disturbances like the endogenous market reaction to macroeconomic news.[22] We are aware that a wide literature studies the reaction of financial markets to monetary policy shocks. For example, stock prices, bond yields and exchange rates respond to the decisions of the central banks in the same day. However, as explained in Section 2.2.1-2.2.3.1, the *Bridge* only requires a component of the true structural shock and not the whole structural shock to yield unbiased estimates of the impact matrix.[23] On the other hand, a more relaxed identification scheme would incorporate other structural shocks in our measure of monetary policy shocks, violating the exclusion restriction and biasing our analysis. This issue is particularly relevant for the FF4 as the price of the FF Futures may incorporate information not strictly related with monetary policy which can nonetheless affect the conduct of monetary policy in the future.[24] Notice that the procedure applied in this case corresponds to the conservative *Bridge* identification formally illustrated and tested in the Monte Carlo experiments (Appendix B.1). Our results are robust to two alternative high frequency identifications that do not rely on timing assumptions. First, we follow Rigobon (2003) in applying the identification through heteroskedasticity: we exploit the change in the volatility of monetary policy shocks between FOMC meeting days and non-FOMC meeting day. Second, independent component analysis allow us to identify structural shock by exploiting the non-normality of the reduced form residuals.[25]

[22] If the information set in our HF system is wide enough, a common unobservable factor may affect all the financial variables at the same time. The available information on macroeconomic aggregates is a good and important candidate. By ordering our monetary policy indicator last, we clean our measure of monetary policy shocks from this unobservable factor, i.e. from all available information captured by financial markets, in particular related with macroeconomic aggregates.

[23] An exogenous variation and not the whole exogenous variation with an instrumental variable terminology.

[24] For the TFFR, the ordering does not matter: the correlation between shocks identified placed the TFFR last or first is 0.97, while repeating the same exercise for the FF4 the correlation falls to 0.7.

[25] Further details on these two alternative identifications are available in Section B.4.1.2 (Appendix).

The full list of variables reads:

VAR: *[Fed Fund Future 3 months; S&P500; VXO; Bid-Cover Ratio in Treasury Auctions; Brent Crude Oil; Eurodollar Exchange Rate; Commodity Price Index; Gold Price Index; BBA Corporate Spread; FED Cleveland Financial Stress Index; Asset Backed Securities (price); 10y Treasure Spread; 5y Treasure Spread; 1y Treasure Spread, Fixed Mortgage Rate; Oil Futures; Dollar-Pound Exchange Rate; Eurodollar Futures; Target Fed Fund Rate]*

We label the shocks identified from our daily VAR as Bridge Target FFR and Bridge FF4 respectively. As a first diagnostic of our identification at HF, we study the relationship between the identified shocks and FOMC meeting days. FOMC meeting days prove to be special day for the size and volatility of the shocks *vis-à-vis* a "normal" day. Quite reasonably, such a special role is more relevant for the shorter horizon contracts, with the maximum for the TFFR. In Appendix B.4, we provide a detailed account through descriptive statistics and regression analysis.

In a second diagnostic, we compare our shocks with RR and GK by restricting our series to the FOMC meeting days only. Table 2.2 reports the contemporaneous correlation across the four series of shocks during FOMC meeting days, while Table 2.3 refers to the monthly aggregates.[26]

Notably, the Bridge TFFR shocks are highly correlated with both the RR series (0.77) and the GK series (0.41). The FF4 shocks are correlated mainly with the GK series (0.61) and less with the RR series (0.27). These correlations decrease once we move from the FOMC dates to the monthly aggregates as we consider all available days in our sample. However, the correlations remain positive and statistically significant also at the monthly frequency.

The lack of correlation around FOMC meetings between the Bridge TFFR shocks and the Bridge FF4 shocks follows by construction from the two estimations at daily frequency. When identifying shocks in the TFFR, we include the FF4 in the VAR and order the TFFR after the FF4. Consequently, a shock to the TFFR does not produce any change in the FF4 in the same day. On the other hand, our

[26]Figures B.17-B.18 display the comparison in monthly terms (in Appendix B.4). The predictive power of our Bridge Target for the RR shocks is reported in Figure B.19. In Figure B.20 we show that both the Bridge Target FFR and the Bridge FF4 contain relevant information to fit the GK shock series.

alternative identification exploits the unexpected daily changes in the FF4 (uncorrelated with the forecast errors of all other variables). As a consequence, the two series of shocks are uncorrelated.

Finally, Table B.9 reports anecdotal evidence of the largest daily shocks from our daily VAR.

We check some properties of our TFFR (FF4) series of shocks that Ramey (2016) and others has found problematic in GK:

- **zero mean**: we test the null hypothesis that our monthly aggregated shocks are drawn from $N(0, \sigma)$ through the Kolmorogov-Smirnov test. We cannot reject the null at any significance level (the sample mean is 0.0007 (0.0016))
- **autocorrelation**: we regress our proxy on its own previous lag and we do not find a significance coefficient. Moreover, the R^2 accounts for 0.02 (0.008)
- **predictability**: we regress our daily proxy, around FOMC meetings, on the Greenbook variables used by the RR.
  - When we perform this exercise on the TFFR shocks, we find some evidence of predictability with the private FED information. Nonetheless, the only significant coefficient related to the current level of output. In our analysis using the sport FF (FF1), this predictability vanishes and the adjusted R^2 turns negative. This discordance is most likely due to the discrete nature of the TFFR.
  - For the FF4, we do not find any significant coefficient and the R^2 is in the order of 0.06, while the adjusted R^2 is negative.

  This dissonance between the predictability in GK versus our series maybe due to the event-study approach of the former study. Using the change in the FF futures around a tight window might not include enough information as all the events across two FOMC meetings (and in other financial markets) are completely discarded.[27]

The third diagnostic refers to the macroeconomic effect of monetary policy shocks: we aggregate by averaging the TFFR and FF4 shocks at monthly frequency and we use the Proxy-SVAR. Using both the shocks in FOMC meeting days only and

[27] These results hold both for the daily and monthly series.

all the monthly shocks, our results are similar to the small scale VAR of GK as reported in Figure 2.9-2.10.[28] Figure 2.9 displays the IRFs computed by employing the identified monetary policy shocks in all the days of the sample. Due to the larger informational content included in the instrument, the confidence bands are narrower than in GK who exploit exclusively monetary policy surprises in FOMC meeting days. On the other hand, in Figure 2.10 we repeat the same exercise but we use only the shocks identified on FOMC meeting days. Furthermore, we do not impute a 30 days window to such shock as in GK and, consequently, the confidence bands are wider than in GK.[29]

If we move to the medium scale system, we find comparable results (Figure B.23).[30] The major difference concerns the response of the excess bond premium: the response is weaker and less persistent in our case. A possible explanation of this finding relies on the risk component. In fact, while GK take the raw change in the price of FF4 contracts, our identification cleans the proxy of the risk component by including many measures of risk in the daily VAR.

Another relevant issue is the informational (Delphic) component that GK include in their measure of monetary policy shocks. Once we include the current rate and the FF future contracts together, we are able to disentangle shocks to the current and future path of interest rates (Figure 2.11). As exemplified by the response of industrial production, a shock to the current rate produces the opposite effects to a shock to the future rate. Moreover, the IRF in GK is exactly the mean between the IRF generate by the two components. We believe that further research should disentangle Odyssean and Delphic components for a better understanding of monetary policy. However, this task goes beyond the scopes of this methodological paper.

[28] In Appendix B.4, we report the same exercise that employs all the daily shocks within a month. We find very similar results.

[29] This implies that our instrument displays many zero entries in the months without FOMC meetings.

[30] Lunsford (2015) provides the correct critical value of the F statistic for the Proxy-SVAR and our first stage result always satisfy his criteria.

2.5 Conclusions

Temporal aggregation is a severe issue in time series analysis, largely ignored in the macroeconomic literature. To alleviate temporal aggregation biases, this paper proposes a new methodology, the *Bridge Proxy-SVAR*, which deals with mixed frequency data. Structural shocks are recovered in high frequency systems, aggregated at the lower frequency, and used as a proxy for a structural shock of interest in lower frequency VARs. By instrumenting the reduced form residuals of a VAR at the macroeconomic frequency, the proxy provides identification restrictions. Our methodology relies on the superiority of identification prior to temporal aggregation over identification post temporal aggregation. In other words, our procedure exploits high frequency data for identification by controlling for the correct information set of policy makers and agents when making announcements or decisions.

The properties of the *Bridge Proxy-SVAR* are studied analytically and its performances are tested through Monte Carlo simulations. Our methodology largely outperforms a LF-VAR using temporally aggregated data, which is the common naive practice in applied macroeconomics. The *Bridge* is also close to the performances of a counter-factual HF-VAR, which constitutes the best possible estimation. In particular, if the amount of information employed is large enough, the *Bridge* replicates the estimation of a HF-VAR with lower precision. The biases introduced by temporal aggregation and the potential gain from the *Bridge* increase with the complexity of the stochastic process under examination. Unlike existing mixed frequency techniques, our methodology can exploit daily data in large dimensional systems to improve the identification of SVARs. At the same time, the MF-VAR is a different econometric model that also improves the autoregressive matrix over a LF-VAR.

As an empirical application, we study the macroeconomic effects of monetary policy in the US. Monetary policy shocks are identified from a large scale daily VAR over the sample 1991m8-2008m6. Although we do not impose any special role for FOMC meeting days, the *Bridge* neatly captures FOMC meeting days as crucial dates. After aggregating the daily shocks at monthly frequency, we use them to instrument the reduced form residuals of the Fed Fund Rate in the monthly VAR of Gertler and Karadi (2015). Our analysis produces very similar IRFs to theirs. Consistently with recent findings in the literature, we show

that Gertler and Karadi (2015) identify a mixture of shocks to the current path and future path of interest rate, where the latter includes relevant informational content.[31]

Importantly for future research, the *Bridge Proxy-SVAR* exploits high frequency information for the identification of SVARs without relying on a definite set of events. The higher the frequency at which they are imposed, the less identifying restrictions constrain the data and the more they are likely to hold. The *Bridge* is particularly promising to improve structural analyses on macro-financial linkages, which are characterized by a wide frequency mismatch and need to take into account a wide information set.

[31] A significant example of the potential of the *Bridge Proxy-SVAR* can be found in the companion paper Gazzani and Vicondoa (2016) where we apply the methodology to identify liquidity shocks in the Italian sovereign debt market.

2.6 Tables

MAD GAINS OVER LF-VAR		
Identification	Temporal Aggregation Scheme	
	SKIP-SAMPLING	AVERAGING
Quarterly-Monthly Frequency Mismatch		
Full information at LF		
HF-VAR	21.2%	41.4%
Bridge	20%	36.7%
Bridge - Partial Information at HF	10.3%	10.8%
Partial Information at LF		
HF-VAR	32%	48.7%
Bridge	31.5%	28%
Monthly-Daily Frequency Mismatch		
Full information at LF		
HF-VAR	70%	81.2%
Bridge	65.6%	72.6%
Bridge - Partial Information at HF	33.2%	47.5%
Partial Information at LF		
HF-VAR	72.9%	84.7%
Bridge	58.7%	64%

Table 2.1: Performance comparison in Monte Carlo simulations

Performance comparison across the counter-factual HF-VAR, the LF-VAR and the Bridge Proxy-SVAR. Performances are evaluated in terms of the Mean Absolute Distance (MAD) between the true IRFs and the estimated IRFs in 100 randomly parametrized DGPs. One summary statistic is computed based all the combinations of shocks-variables in the system. The gains are expressed as percentage MAD gains over the LF-VAR. We analyze different cases for a VAR(1) DGP: I) the frequency mismatch between HF and LF is 3: monthly-quarterly case. II) the frequency mismatch between HF and LF is 30: monthly-daily case. For both I) and II) we study two sub-cases: a) The Bridge employs full information at HF; b) The Bridge employs only partial information at HF (7 out of 9 variables). In this latter case, the Bridge employs the conservative identification discussed in Appendix B.1. For case a) we also analyze: a.1) the LF stage of the Bridge and the LF-VAR use all available information; a.2) the LF stage and the LF-VAR do not include all the variables in the system (only 7 out of 9 variables).

	Bridge Target FFR	Bridge FF4	Romer & Romer	Gertler & Karadi
Bridge Target FFR	1	*	*	*
Bridge FF4	0	1	*	*
Romer & Romer	0.77	0.27	1	*
Gertler & Karadi	0.49	0.61	0.32	1

Table 2.2: Correlation across different monetary policy shocks in FOMC meeting days

Correlations among different monetary policy shocks in FOMC meetings days: 1) Shocks to the Target FFR identified from our daily VAR; 2) Shocks to the Fed Future contracts (3 months ahead) identified from our daily VAR; 3) Monetary policy shocks as in Romer and Romer (2004) shocks extended by Coibion et al. (2012); 4) Monetary policy shocks as in Gertler and Karadi (2015). All coefficients different from 0 are statistically significant at the 1% level.

	Bridge Target FFR	Bridge FF4	Romer & Romer	Gertler & Karadi
Bridge Target FFR	1	*	*	*
Bridge FF4	0.1	1	*	*
Romer & Romer	0.34*	0.18*	1	*
Gertler & Karadi	0.27*	0.23*	0.2*	1

Table 2.3: Correlation across different monetary policy shocks at monthly frequency

*Correlations among monthly measures of different monetary policy shocks: 1) Shocks to the Target FFR identified from our daily VAR; 2) Shocks to the Fed Future contracts (3 months ahead) identified from our daily VAR; 3) Monetary policy shocks as in Romer and Romer (2004) shocks extended by Coibion et al. (2012); 4) Monetary policy shocks as in Gertler and Karadi (2015). * denotes statistical significance at the 1% level.*

2.7 Figures

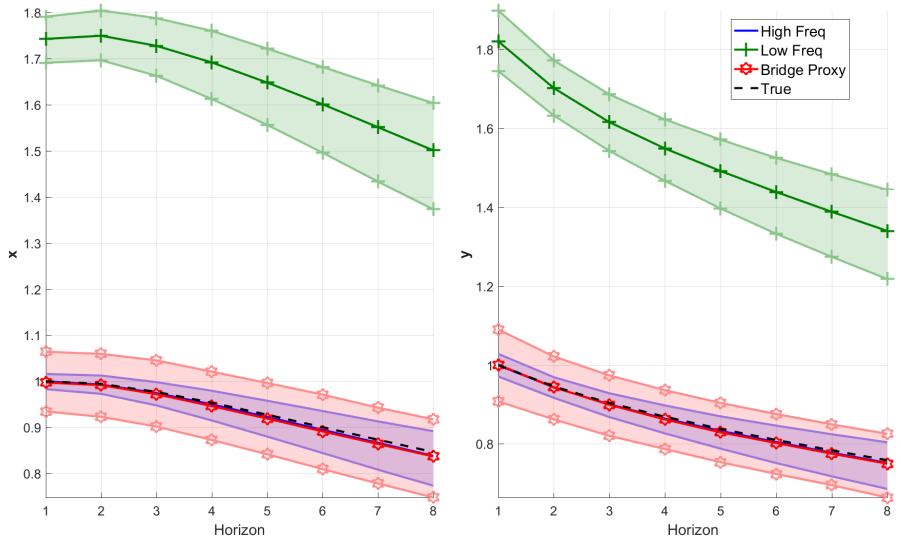

Figure 2.1: IRFs1 in the two variable case - skip sampling

IRFs to a shock in the first variable (x) in the bivariate system. The true IRF is represented by the dotted black line. The shock is identified through the correct recursive structure in the HF system (blue), LF system (green) and Bridge Proxy (red). Shaded areas correspond to the 90% confidence bands across 1000 replications. Time aggregation follows a skip-sampling scheme.

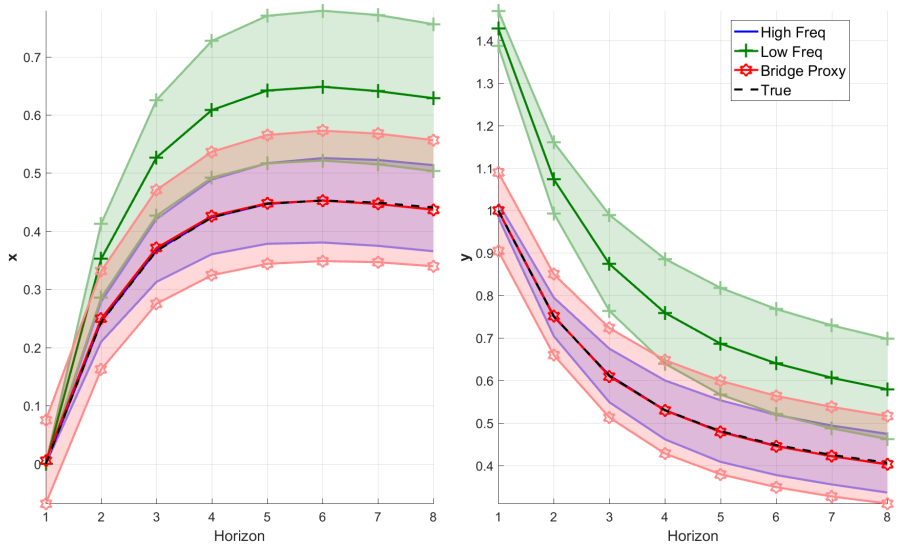

Figure 2.2: IRFs2 in the two variable case - skip sampling

IRFs to a shock in the second variable (y) in the bivariate system. The true IRF is represented by the dotted black line. The shock is identified through the correct recursive structure in the HF system (blue), LF system (green) and Bridge Proxy (red). Shaded areas correspond to the 90% confidence bands across 1000 replications. Time aggregation follows a skip-sampling scheme.

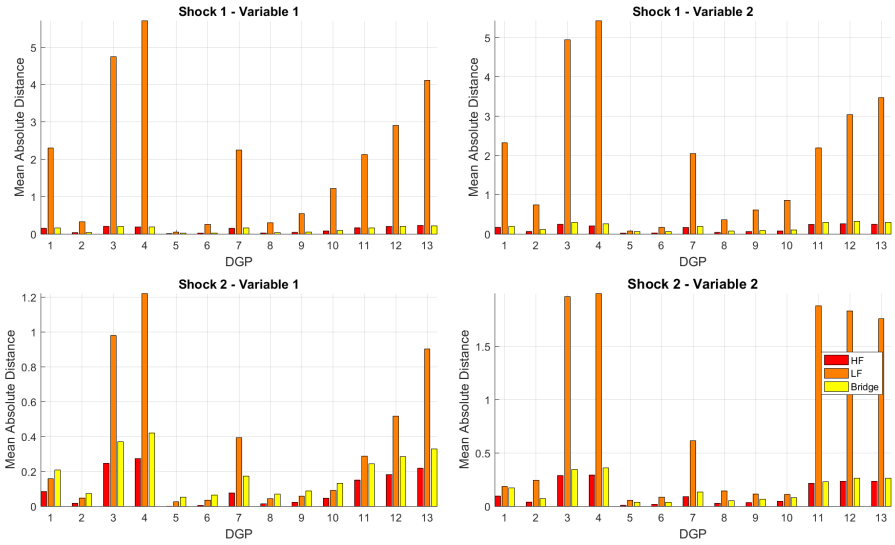

Figure 2.3: MAD comparison in the two variable case - skip sampling

Mean Absolute Distance (MAD) between the true IRFs and the IRFs estimated by the HF-VAR, LF-VAR and Bridge Proxy-SVAR (through the correct recursive scheme). Results are reported for 13 parametrization of the DGP. The MAD is computed by averaging the MAD over the 1000 replications. Time aggregation follows a skip-sampling scheme.

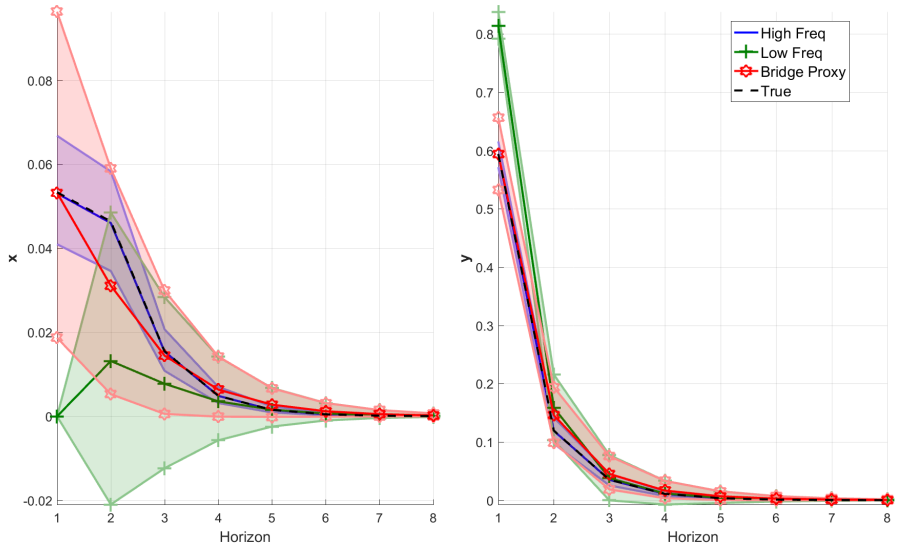

Figure 2.4: IRFs2 in the two variable case - averaging

IRFs to a shock in the second variable (y) in the bivariate system. The true IRF is represented by the dotted black line. The shock is identified through the correct recursive structure in the HF system (blue), LF system (green) and Bridge Proxy (red). Shaded areas correspond to the 90% confidence bands across 1000 replications. Time aggregation follows an averaging scheme.

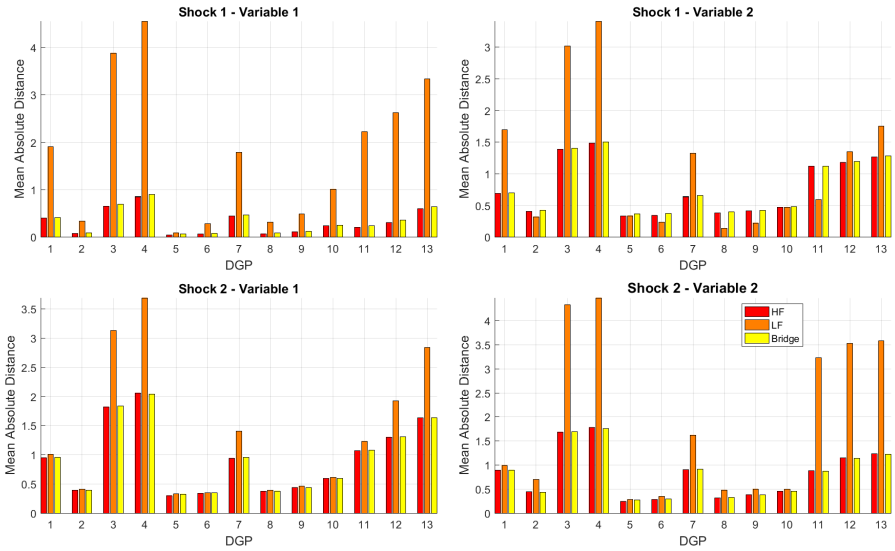

Figure 2.5: MAD comparison in the two variable case - averaging

Mean Absolute Distance (MAD) between the true IRFs and the IRFs estimated by the HF-VAR, LF-VAR and Bridge Proxy-SVAR (through the non-correct recursive scheme). Results are reported for 13 parametrization of the DGP. The MAD is computed by averaging the MAD over the 1000 replications. Time aggregation follows an averaging scheme.

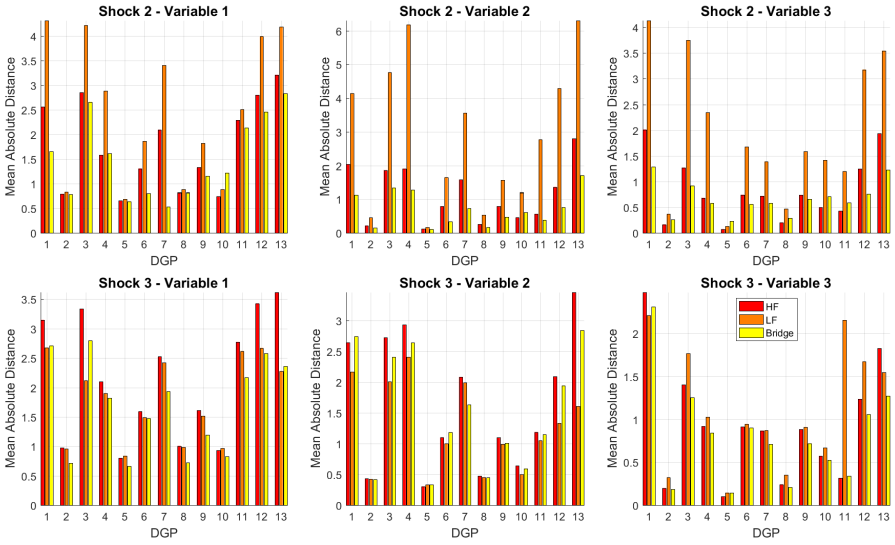

Figure 2.6: MAD comparison in the practical case

Mean Absolute Distance (MAD) between the true IRFs and the IRFs estimated by the HF-VAR, LF-VAR and Bridge Proxy-SVAR (through the correct recursive scheme). Results are reported for 13 parametrization of the DGP. The MAD is computed by averaging the MAD over the 1000 replications. Time aggregation follows a skip-sampling scheme.

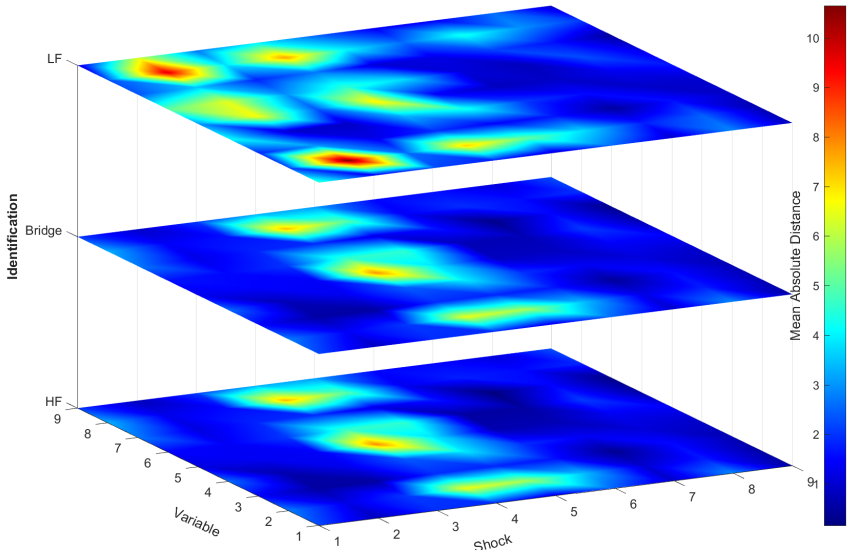

Figure 2.7: MAD heatmap from large randomized Monte Carlo experiment

Mean Absolute Distance (MAD) between the true IRFs and the IRFs estimated by the HF-VAR, LF-VAR and Bridge Proxy-SVAR in one of the 100 randomly parametrized DGPs. Results are reported for each combination of shocks-variables in the system (81). The MAD is computed by averaging the MAD over the 1000 replications. Time aggregation follows a skip-sampling scheme.

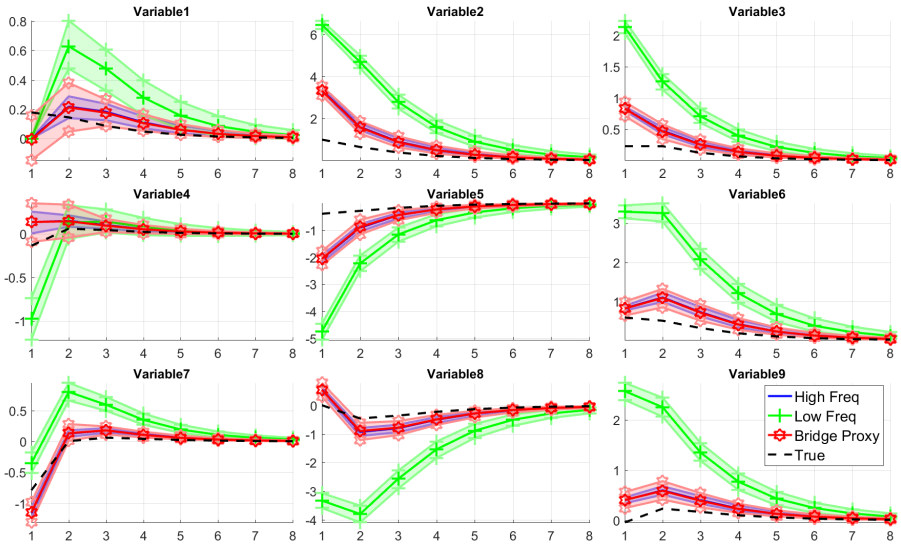

Figure 2.8: IRFs from large randomized Monte Carlo experiment

Example of the IRFs of the system to a shock in the first variable in the system, estimated by the HF-VAR, LF-VAR and Bridge Proxy-SVAR in one of the 100 randomly parametrized DGPs. Shaded areas correspond to the 90% confidence bands across 1000 replications. The true IRF is represented by the dotted black line. Time aggregation follows a skip-sampling scheme.

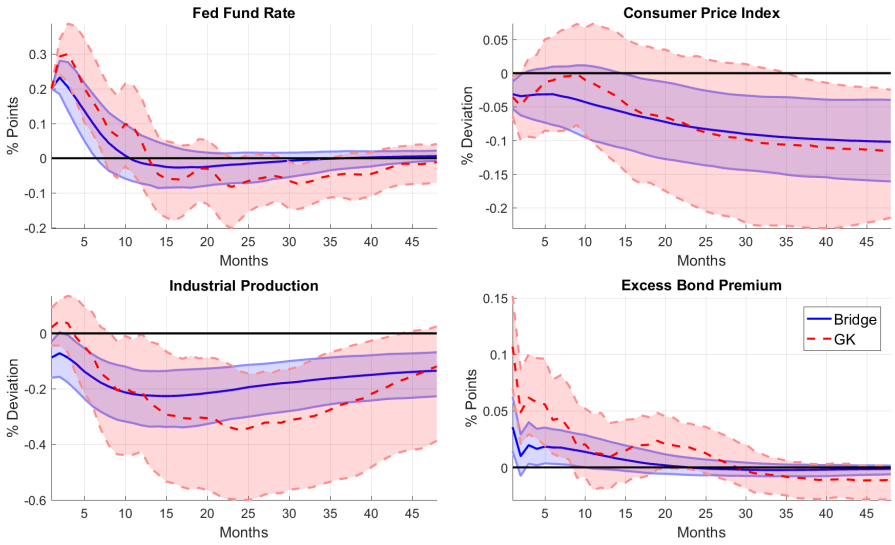

Figure 2.9: IRFs TFFR

IRFs to a monetary policy shock identified by instrumenting the Fed Fund Rate with our series of shocks in the Target Fed Fund rate recovered from our daily VAR. From the first stage, $F-stat=11$. The VAR includes [FFR, CPI, Industrial Production, Excess Bond Premium] and it is estimated in log-levels including the optimal number of lags (2) and a deterministic constant. Shaded areas correspond to 95% bootstrapped confidence bands from 1000 replications.

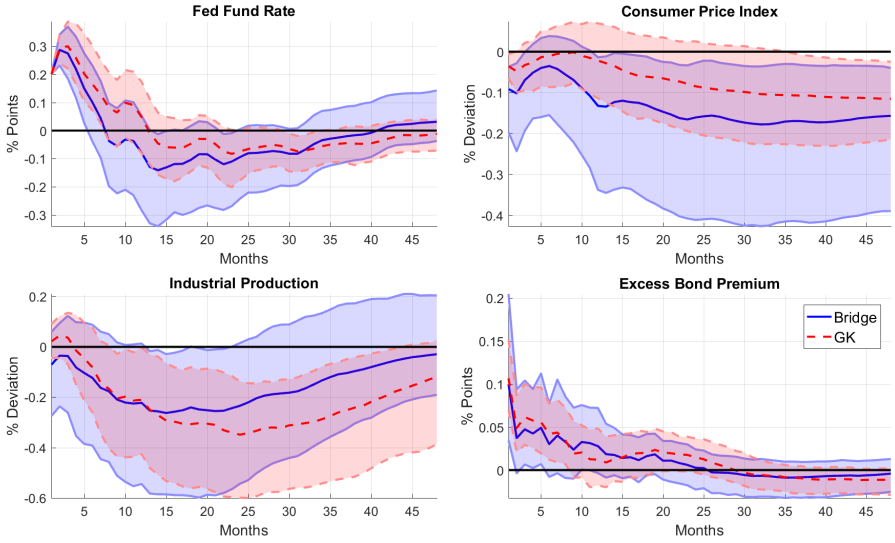

Figure 2.10: IRFs FF4 comparable with Gertler and Karadi (2015)

IRFs to a monetary policy shock identified by instrumenting the Fed Fund Rate with the series of shocks in the Fed Fund Future 3 month ahead recovered from our daily VAR. We assign each FOMC meeting day only to the corresponding month (without imputing it to other months). From the first stage, $F-stat=7.5$. We employ exactly the same specification of Gertler and Karadi (2015): the VAR includes [FFR, CPI, Industrial Production, Excess Bond Premium] and it is estimated in log-levels including 12 lags and a deterministic constant. Shaded areas correspond to 95% bootstrapped confidence bands from 1000 replications.

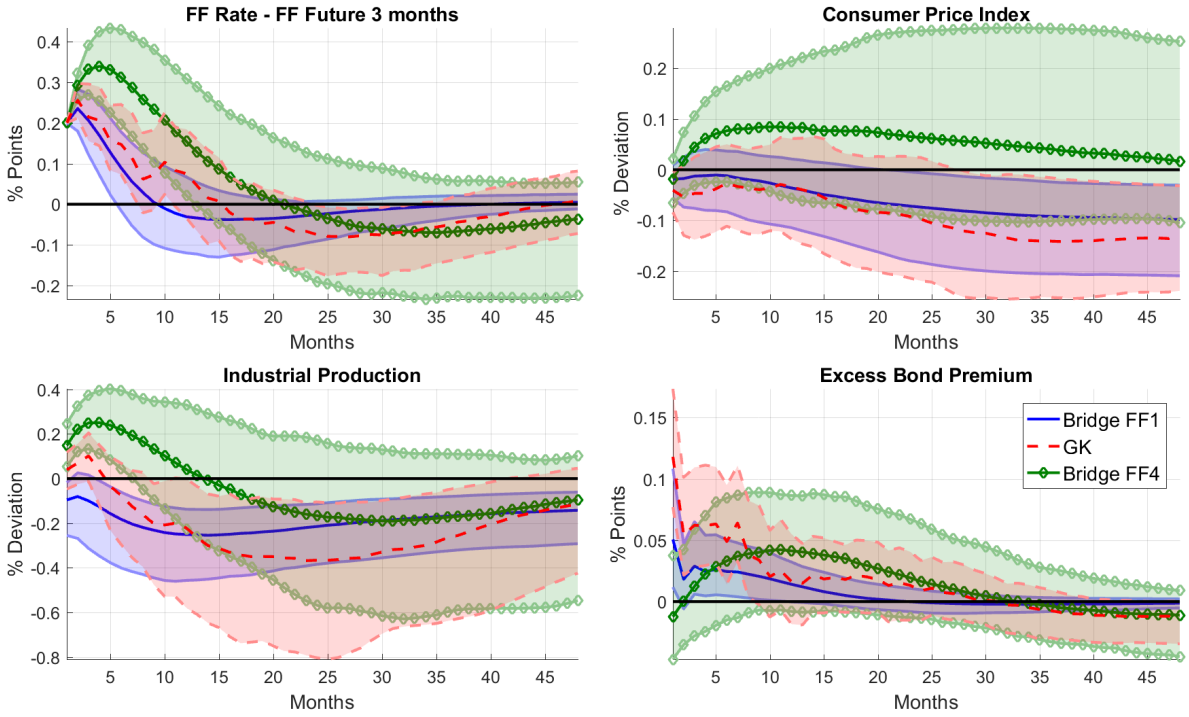

Figure 2.11: IRFs - current and future path

IRFs to a monetary policy shock identified by instrumenting the Fed Fund Rate (Fed Fund Future 3 month ahead) in blue (green) with the series of shocks in the Fed Fund Future 1 (Fed Fund Future 3) month ahead recovered from our daily VAR. From the first stage, for FF1 $F-stat = 16.2$ *and for FF4* $F-stat = 25.6$*. The VAR includes [FFR, CPI, Industrial Production, Excess Bond Premium] and it is estimated in log-levels with the optimal number of lags (2) and includes a deterministic constant. Shaded areas correspond to 95% bootstrapped confidence bands from 1000 replications.*

Chapter 3

The Real Effect of Liquidity Shocks in Sovereign Debt Markets: Evidence from Italy

Joint with Alejandro Vicondoa

3.1 Introduction

The sovereign debt crisis has dramatically affected European countries since 2010. In particular, southern European countries like Greece, Italy, Portugal and Spain (GIPS) have been facing increasing unemployment rates and worsening credit conditions for governments, households and firms. Both the media and economic researchers have focused on the behavior of spreads in yields and credit default swaps (CDS), which are supposed to reflect default risk. However, sovereign bonds are highly demanded for their liquidity properties that have also fluctuated during the crisis.

In this paper, we examine liquidity, understood as the ease in releasing an asset quickly without incurring additional costs (i.e. market liquidity), as a different but complementary dimension of financial tensions. We measure liquidity by using the most popular measure: the *Bid-Ask Spread* (BAS). The BAS measures the distance between the highest bid price and the lowest ask price for an asset. A narrower BAS denotes liquidity because the lower the BAS the easier trading the asset quickly without transaction costs. We also employ an alternative indicator

which takes into account the volumes traded in secondary markets. Government bonds are the most liquid assets in the economy, after money itself. European banks hold large amounts of these assets in their portfolio due to their historical low default risk and liquidity risk. Abrupt changes in the liquidity of sovereign bonds could affect the lending decisions of banks.

To the best of our knowledge, this is the first empirical investigation on the macroeconomic effects of exogenous changes in liquidity in sovereign debt markets, which we call *liquidity shocks*. The Euro crisis constitutes an ideal laboratory for such analysis because indicators of liquidity and default risk display different patterns that can be used for identification. Figure 3.1 shows the evolution of the *Bid-Ask Spread* (BAS), CDS and yield for Italy, which accounts for 26% of European sovereign debt, between 2004 and 2014.[1] While during 2007-2011 the yield and BAS move in opposite directions, between 2011-2012 both of them increase. Moreover, the CDS displays different dynamics with respect to the other variables. Considering the fluctuations in Italian business cycle during this period, we identify the effects and transmission channels of liquidity shocks. We base our analysis on Vector Autoregression models (VAR) and our identification strategy relies both on the standard recursive ordering and on the Proxy-SVAR methodology. The latter uses exogenous changes in liquidity identified in a financial daily VAR as an instrument for structural liquidity shocks.

Liquidity, as we show, has been a major driver for the Italian economy during the sovereign debt crisis. The Forecast Error Variance (FEV) decomposition shows that liquidity shocks explain a relevant share of the volatility of unemployment (15%) and confidence indicators, like consumer confidence, business confidence and stock prices. A BAS shock generates macroeconomic effects that are at least as strong as the effects generated by a raise in yield spreads.[2]

In order to understand the transmission mechanism of liquidity shocks, we turn to survey data. The *Bank Lending Survey* and the *ISTAT Business Confidence Survey* reveal that liquidity shocks affect the lending behavior of banks through their liquidity position and costs related to their capital position. Shocks to

[1] European sovereign debt markets are concentrated with Italy and France accounting for roughly 50% of the total public debt. Source: European Central Bank Statistics. Italy: 26.4%, France 22.7%, and Germany 18.3%. The three variables are expressed as monthly averages.

[2] The joint contribution of BAS and yield spread shocks to the FEV of unemployment is 20% across 2004-2014 (15% + 5% respectively) and raises up to 30% aver 2009-2014 (15% + 15% respectively).

sovereign yield spreads do not generate worse lending conditions through the same channels. Our findings are particularly relevant to improve the understanding of the relationship between real economy and financial markets.

Our empirical results can be interpreted using the theoretical framework developed by Cui and Radde (2015). They build a real DSGE model with search and matching frictions in asset markets, where the financial sector intermediates between buyers and sellers of financial assets. In this framework, an exogenous increase in financial intermediation costs affects the market participation of buyers more than the one of sellers and induces a fall in the liquidity of financial assets. Market liquidity produces relevant implications for the real economy by tightening the financial constraints of firms and reducing their financing possibilities.[3] Cui and Radde (2015) mainly focus on private assets since, in the U.S., sovereign bonds did not experience a fall in liquidity during the crisis. On the contrary, as Figure 3.1 displays, in the European (Italian) case, the liquidity of sovereign bonds has fluctuated significantly.[4] Moreover, their setup can accommodate both market-based and bank-based financial intermediation, with the latter characterizing European economies. Our empirical findings and their theoretical results are consistent in terms of: the observed fall in output, fall in consumption and investment (proxied by business and consumer confidence indicators), turnover (i.e. traded volume relative the outstanding amount of the asset), and asset prices. The only (qualitative) difference consists in their responses being starker than our IRFs because they rely on a model without nominal frictions. In a similar setup to Cui and Radde (2015), Cui (2016) studies monetary and fiscal interactions with market liquidity, and draws conclusions on optimal policies by considering government debt as provider of liquidity services.

Further works have also studied liquidity in theoretical frameworks: Del Negro et al. (2011) and Benigno and Nistico (2014) study the effects of shocks to an exogenous liquidity constraint, which restricts the fraction of an asset which can be used to purchase goods. While Del Negro et al. (2011) impose this constraint on the fraction of equity holdings that a household can resale, Benigno and Nistico

[3]Notice that, contrary to the existing literature, they are able to generate the comovement between asset turnover and asset prices.

[4]Notice that we have also found similar macroeconomic results for the liquidity of corporate bonds and for the spread in liquidity between corporate and sovereign bonds. Nonetheless, in all the specifications, shocks to the liquidity of sovereign bonds induce sizeable macroeconomic effects.

(2014) restrict the fraction of government bonds that can be exchanged for goods. Unlike Cui and Radde (2015), these papers do not endogenize the dynamics of asset liquidity. Both papers conclude that liquidity shocks (i.e a decrease in the release fraction of these assets) produce strong and negative effects on GDP and prices, which in both cases are partially explained by a fall in private consumption. These results differ from our empirical findings since we do not find that liquidity shocks induce a significant effect on CPI inflation. Passadore and Xu (2014) investigate how liquidity risk and credit risk explain sovereign spreads through the optimal behavior of buyers and sellers. In an endowment economy with incomplete markets and search and matching frictions in the sovereign debt markets, they find that the liquidity component can explain up to 50% of sovereign spread during the Argentinian crisis in 2001. Although the model matches the correlations and standard deviations of consumption and net exports, they do not consider the effects on output. Overall, we contribute to this literature by characterizing the empirical effects of liquidity shocks and by identifying its transmission through the banking sector. In light of our empirical findings and of the existing models, we believe that financial intermediation and search frictions are a key feature to be taken into account when studying liquidity.

This paper is also related to the strand of the literature that analyzes the macroeconomic effects of shocks to the spread in yields. Bahaj (2014) and Neri and Ropele (2015) study the macroeconomic effects of yield shocks and find that they explain a relevant fraction of business cycle fluctuations in European countries. However, they do not consider sovereign debt liquidity in their analysis and this omitted dimension could affect their conclusions. Regarding the transmission channels, tensions in sovereign debt markets induce a tightening in credit conditions through an increase in the funding costs of banks (De Marco (2016)) or through the Repo market (Boissel et al. (2014) and Mancini et al. (2014)). In this paper, we show that liquidity shocks have strong macroeconomic effects and identify its transmission through the banking sector. We find that liquidity is at least as relevant as spread in yields to explain fluctuations in economic activity in Italy and Spain and that commercial banks respond to liquidity shocks in a different way than to a yield shock.

The remainder of this paper is organized as follows. Section 3.2 describes the high frequency variables that characterize Italian sovereign debt market. Section 3.3 presents the empirical specification and results using different identification

schemes. Section 3.4 investigates the transmission channels by exploiting survey data. Section 3.5 compares the Italian results to France, Germany and Spain and Section 3.6 concludes.

3.2 Data Description

Sovereign debt markets can be characterized by different indicators: Spread in Yields (Spread), Credit Default Swaps (CDS), and Bid-Ask Spread (BAS). The first one captures the difference in yields that a country has to pay in order to issue sovereign debt with respect to a safe asset, which in this case is the German sovereign bond with the same maturity. CDS is a proxy for credit risk. Finally, the third is a widely-used indicator of sovereign debt liquidity (see for example Pericoli and Taboga (2015) and Pelizzon et al. (2015)).[5] These variables enable us to characterize the sovereign debt markets. For our analysis, we use data from Italy for the period February 2004 until November 2014. The Italian sovereign debt market is one of the most important in Europe, accounting for 26% of the European government debt.[6] Before proceeding to the analysis, we describe briefly the relationship between the three indicators. Table 3.1 displays the daily correlation between these variables, both in levels and growth rates.

CDS is highly correlated (0.91) with the Spread while the BAS displays a relative low correlation with the other two variables. This fact also holds if we consider the variables in daily growth rates instead of in levels. In particular, the daily changes of the BAS are uncorrelated with the other financial variables while CDS and Spread are positively correlated. From this preliminary description, we can see that movements in Spread are more associated with credit risk (proxied by the CDS) than liquidity risk, a similar finding with Pericoli and Taboga (2015).[7] However, these variables maybe correlated with other financial ones like stock prices,

[5]Alternatively, people also look at the volume traded or at a combination of both. Figure A1 in Appendix C.1 displays the evolution of the volume traded together with the BAS. We use the BAS for our empirical analysis and present the results using the *Liquidity Index*, which incorporates both BAS and Turnover, in Appendix C.4.1.

[6]Source: European Central Bank Statistics.

[7]Notice that there is still no consensus in the finance literature. For example, Schwarz (2014) highlights, through a novel measure of liquidity, that liquidity risk explains a large share of the raising yields during the Euro crisis. Beber et al. (2009) show that, during period of market stress, investors chase liquidity and not credit quality.

interest rates or the equity implied volatility from options. Figure 3.2 displays the evolution of these financial variables at daily frequency.[8]

The peaks in the VSTOXX index reflect the two main periods of financial stress: the second part of 2008, associated with the collapse of Lehman Brothers, and between the second half of 2011 and 2012, related to problems in the European Sovereign Debt markets.[9] These periods of stress are reflected in a different way for each financial variable. On the one hand, the Italian stock price index (FTSE MIB) falls with these two events and recover afterwards, without reaching the peak of 2007. The response of the Eonia rate is similar and reflects the interest rate decisions of the ECB and interbank market stress. On the other hand, financial variables associated with sovereign debt markets display different dynamics. The BAS spikes in 2009 and exhibits an abrupt change in volatility after January 14, 2011, when the *Fitch agency* downgraded Greek sovereign debt to junk status.[10] The dynamics of CDS and Spread are similar during 2012, in line with the correlations reported in Table 3.1, but the Spread declines at a lower pace after the spikes than the CDS. During 2014, we observe some spikes in the BAS whereas Spread and CDS decline steadily. The key point for identification is that the six financial variables display different patterns.

Since in this paper we are going to focus on shocks to BAS, we analyze whether fluctuations in this variable are associated with particular European events. This analysis enables to us to understand better the underlying dynamics of this variable and its sources of variation. Figure 3.3 displays the dynamics of the BAS together with some key events related to the European Sovereign debt crisis, which are reported in Table C.1.

First of all, as we mentioned before, the series displays a clear change in volatility after January 14 2011. After that date, many events related to Portugal, Spain, Greece, and Italy are reflected as spikes in this variable. Additionally, other European events coincide with BAS local maxima or local minima. In particular, the BAS reached a minimum, comparable to pre-crisis levels, when Mario Draghi stated the "Whatever it takes to save the Euro". Liquidity in the Italian sovereign

[8] We use the European Volatility Index (VSTOXX) instead of the one based on FTSE MIB index because it is available for the whole period and it is representative also for the Italian economy. Both indexes are highly correlated for the period when they coincide.

[9] In fact, the decline in the implied volatility happens after the famous speech of Mario Draghi, president of the ECB, on July 26 2012.

[10] This fact holds for Spain only a few days later.

debt market reflects important economic news, which is key for identification because many of those events can be considered as exogenous with respect to the Italian economy.

3.3 Empirical Analysis

To analyze the effects of liquidity shocks we rely on different VAR specifications. In Section 3.3.1, we estimate a small scale VAR used to identify the effects of liquidity shocks. Then, we use an enlarged VAR for a better identification of the shocks and to characterize in higher detail the results and the transmission mechanisms (Section 3.3.2). Both specifications rely on the Cholesky decomposition to identify liquidity shocks. Given that imposing zero contemporaneous restrictions on some financial variables can be controversial, in Section 3.3.3 we employ a more agnostic identification strategy, the Proxy-SVAR, which places no restrictions on the timing or sign of the responses. Finally, in Section 3.3.4 we present extensions and additional exercises to further investigate liquidity and assess the robustness of our findings.

3.3.1 Basic Specification

As a first step, we estimate the effect of BAS shocks on Italian business cycles using a small scale VAR. In particular, we specify a VAR that includes the *Unemployment Rate,* as a proxy for economic activity; *Consumer Price Inflation* expressed as an annual rate, to capture price dynamics; *FTSE MIB,* which is the main index of Stock Prices in Italy; *Sovereign Spread*; and *BAS.* While the first two variables are useful to capture the transmission to the real economy, the last three are necessary to identify a liquidity shock. Our sample runs from February 2004 through November 2014. To deal with the different frequencies, we include the financial variables as monthly averages in order to capture all the dynamics during the period.[11] Following Sims et al. (1990), we estimate the model in (log-)levels by OLS, without explicitly modeling the possible cointegration relations among them.[12]

[11] In Appendix we report summary statistic of the main financial variables aggregated at monthly frequency.

[12] Sims et al. (1990) show that if cointegration among the variables exists, the system's dynamics can be consistently estimated in a VAR in levels.

In addition to a constant, we also include a deterministic trend. The lag order is selected following the three information criteria and it is always one.[13]

We identify a liquidity shock using a standard Cholesky decomposition, which is based on recursive ordering. The variables are ordered in the VAR from the most exogenous to the most endogenous, which are allowed to respond contemporaneously to all structural shocks. Thus, we order Unemployment and Inflation, assuming that they cannot react to the shock on the same month. A severe problem arises from the three financial variables that our VAR incorporates. Obviously, they always react to all the available information and so there is no convincing way of ordering them. Considering this issue, we take a more agnostic stance. Within the financial block, we consider all the possible orderings and we report the median and percentiles of the impulse responses and Forecast Error Variance (FEV). In this way, we identify 6 rotations and, for each of those, we compute 100 bootstrap replications. Figure 3.4 displays the Impulse Response Functions (IRFs) to a one standard deviation BAS shock (i.e. a decrease in liquidity). We report the median together with 68% and 90% confidence bands that include both the identification (from the different Cholesky orderings) and statistical uncertainty.

An increase in the BAS induces an increase in Unemployment which lasts 10 months and a slight decrease in CPI inflation. However, the remaining financial variables do not react to the BAS shock. Similar results hold if we estimate the same VAR using the pre-2009 and the crisis sample.[14] Thus, shocks to the BAS have strong effects on economic activity. In order to understand the channels behind this relationship and to see whether results are robust, in the next section we consider a large scale VAR.

3.3.2 Full Specification

We aim at assessing the macroeconomic effects of BAS shocks, with special emphasis on the comparison with other financial shocks. For this purpose, we enlarge the previous VAR system with other variables. This system features six macroeconomic variables (Unemployment, CPI Inflation, Public Debt, ECB Repo, Italian M2, Consumer and Business Confidence) plus five financial indicators (stock prices, Spread, CDS, BAS and VSTOXX). This set of variables is necessary

[13] We check that the residuals are normally distributed and they do not exhibit autocorrelation.
[14] For ease of exposition, we present these results in the Appendix.

to identify financial shocks and assess their transmission to the real economy.[15] Like in Section 3.3.1, we identify the liquidity shock through recursive ordering. In particular, we assume that macroeconomic variables cannot react contemporaneously to the financial shocks and we order them as follows: [UNEMPLOYMENT, CPI, PUBLIC DEBT, M2, CONSUMER CONFIDENCE, BUSINESS CONFIDENCE].

Again, within the financial block, we consider all the possible orderings (120 rotations), compute five bootstrap replications for each of them and report the median and percentiles of the impulse responses and FEV. Different possible orderings across the financial block lead to very similar results, which means that the covariance matrix of the reduce form residuals is close to a diagonal matrix.

Figure 3.5 displays the IRFs to a one standard deviation BAS shock, where 68% and 90% confidence bands include both the identification (from the different Cholesky orderings) and statistical uncertainty. A negative liquidity shock induces an increase in unemployment that reaches its maximum after four months without a significant effect on inflation, comparable to the findings of the VAR presented in Section 3.3.1. The stock of government debt falls with a lag whereas there is no reaction in the Repo rate and M2. Both business and consumer confidence indicators decline in response to the shock and reach their trough four months after the shock. The response of confidence is strong across all the specifications and could reflect a fall both in current and future consumption, which may help to explain the strong response of unemployment (Ludvigson (2004)). Moreover, these dynamics are consistent with the findings of Garcia and Gimeno (2014) for flight-to-liquidity episodes. The FEV contributions of BAS to consumer confidence, business confidence and stock prices are respectively 15%, 9% and 7% one year after the shock. Moving to the financial block, the equity premium, CDS and spread increase and the FTSE declines by 1%, all of them with a lag. Responses of financial variables are in line expected movements: a decrease in the BAS, which could be interpreted as an increase in the uncertainty regarding the value of the underlying asset, reduces prices (i.e. increases the Yield), confidence, and stock prices and increases volatility and CDS.

A key point in our analysis, in light of the outstanding literature on the Euro Crisis, consists of the comparison between BAS (Figure 3.5) and Spread shocks

[15] As in Section 3.3.1, we estimate the VAR in (log) levels by OLS equation by equation. The optimal number of lags is one. Our sample consists of 130 observations which leaves us with enough degrees of freedom for the estimation (15 coefficients in each equation).

(Figure 3.6). The Spread shock induces a similar effect on unemployment slightly less persistent and significant. However, this shock has a negative effect on CPI inflation, which declines by 0.04% points 2 months after the shock. Even if the response of CPI inflation is different with respect to a BAS shock, in Section 3.3.3 we show that, by using the Proxy-SVAR, the IRF of CPI to a BAS shock is also negative.[16] Unlike in the previous case, consumer confidence and business confidence do not display a significant reaction. Regarding the financial block, the responses are similar in magnitude (even if less significant) but less lagged than the case of a BAS shock. An increase in Spread induces a delayed raise in BAS. While the effects on unemployment are similar to the ones reported by Neri and Ropele (2015) using a similar sample, the ones on inflation are the opposite from theirs. This difference may be due to the omission of the liquidity dimension.

For a more comprehensive comparison among financial shocks, in Figure 3.7 we report the FEV decomposition of unemployment (i.e. how much each financial shock explains of unemployment's volatility). BAS shocks explain approximately 15% of unemployment fluctuations at a two year horizon. The second largest shock in relevance is the stock prices, accounting for 7%. The remaining financial shocks do not explain a significant fraction of fluctuations in unemployment. All in all, exogenous fluctuations in financial variables explain around 30% of the total variability of unemployment. From this analysis, we can conclude that liquidity is a major driver of unemployment, out of all the financial variables, for the period under analysis.[17]

3.3.3 Proxy-SVAR

While the results of Section 3.3.2 are robust to the different Cholesky orderings, still, in each rotation, we are constraining (some) financial variables not to react on impact to other financial shocks. In this section, we relax this assumption by applying the so called *Proxy-SVAR* identification developed by Stock and Watson (2012) and Mertens and Ravn (2013). The main idea is to use information external to the VAR system as a proxy for the structural shock of interest, the BAS shock

[16] As we show later on, CPI is the only variable whose dynamics changes across the two methodologies. Notice that this difference comes from the years 2004-2009 as we display in Figure A2. The response of Spread is robust for the sub-sample 2009-2014.

[17] The relative contribution of each financial shock changes if we consider the sub-sample 2009-2014 (Figure A3 in Appendix C.4). In this case, the contribution of spread is similar to the one of BAS, which is quantitatively stable over the full sample.

in our case. In practice, the proxy constitutes an instrument for the reduced form residuals of the VAR and provides partial identification of the structural shocks. The instrument is assumed to be correlated with the structural shock of interest but not with the remaining ones. An advantage of this technique is that, as long as the proxy is a relevant and valid instrument, the identification relies on a much weaker set of assumptions than the recursive identification scheme.[18] In other words, no assumptions are made on the contemporaneous relationship among the variables in the system. Appendix C.3 contains a detailed explanation of this methodology.

In order to obtain a valid instrument for BAS, we propose a new way to identify the proxy for the Proxy-SVAR at high frequency. We label this identification *"Bridge Proxy-SVAR"* because the Proxy-SVAR links two VAR systems that include data at different frequencies. In Gazzani and Vicondoa (2016b), we illustrate analytically the properties of *Bridge Proxy-SVAR* the and test it via Monte Carlo simulations. The procedure consists of the following steps:

1. Construct two VARs systems. The first one is a VAR that incorporates daily financial variables relevant for the analysis, defined as High Frequency VAR (HF-VAR). This VAR features $[BAS, CDS, Yield, FTSE, Eonia, VIX]$. The second one is a VAR, defined as Low Frequency VAR (LF-VAR), that includes variables at monthly frequency. In particular, it is the same system that we define in Section 3.3.2. Again, the financial variables in the LF-VAR are included as monthly averages.

2. Estimate the HF-VAR and identify the structural shock of interest ε_{HF}^{BAS} with the most appropriate identification scheme. Given that economic theory does not support identification via sign restrictions, we apply the recursive ordering Cholesky decomposition. Notice that the biases implied by Cholesky in the HF-VAR are much lighter than in the LF-VAR. Since we observe a structural break in the daily volatility of financial variables in 2009, we estimate a VAR at daily frequency to identify structural innovations in the BAS during the period 2009m1-2014m11 and we use them as an instrument for the structural BAS shocks at monthly frequency.

3. Aggregate ε_{HF}^{BAS} into monthly frequency obtaining $\bar{\varepsilon}_{HF}^{BAS}$.

[18] The proxy is not assumed to be perfectly correlated with the structural shock, but only to be a component of it.

4. Estimate the LF-VAR and apply the Proxy-SVAR identification, where $\bar{\varepsilon}_{HF}^{BAS}$ is employed as a proxy for the for the structural shock of interest in the LF-VAR ε_{LF}^{BAS}. Namely, the reduced form residual u_{LF}^{BAS} is instrumented with $\bar{\varepsilon}_{HF}^{BAS}$. Again, the underlying assumptions concern the relevance, $corr\left(\bar{\varepsilon}_{HF}^{BAS}, \varepsilon_{LF}^{BAS}\right) \neq 0$, and the validity, $corr\left(\bar{\varepsilon}_{HF}^{BAS}, \varepsilon_{LF}^{j}\right) = 0 \ \forall j \neq BAS$, of the instrument.

This proxy explains a significant fraction of BAS reduced form residuals from the monthly VAR. The statistics of the first stage are F-stat $= 29.465$ and $R^2 = 0.30231$, which satisfies the requirements of a strong instrument suggested by Stock and Yogo (2002). This means that a relevant fraction of the reduced form residuals are explained by the daily shocks to the BAS.[19] Figure 7 reports the IRFs to an instrumented shock to the BAS. The BAS shock induces a significant and persistent effect on unemployment, very similar both quantitatively and qualitatively to the ones described in Section 3.3.2. Unlike with the recursive ordering, CPI inflation decreases by 0.02% after the shock. As displayed in Figure A2, this difference is not due to the methodology but to the shorter sample used. The remaining variables in the macroeconomic block display a comparable reaction to the recursive ordering case. In particular, the BAS shock generates a strong response in the indicators of confidence. All the financial variables display a significant lagged response, except for the Equity Premium that reacts on impact.

Even if the Proxy-SVAR relies on a weaker set of assumptions, we include it only as an alternative because this approach just reaches partial identification. This implies that we cannot explicitly compare liquidity and spread shocks. Nonetheless, the results from the Proxy-SVAR confirm the validity of the recursive ordering identification previously applied, that is the standard methodology. Notice that, with the Proxy-SVAR, even without imposing any contemporaneous restriction, financial variables do not display a significant response on impact (apart from the Equity Premium). However, under this methodology, we can still compute the historical contribution of liquidity shocks to unemployment, which help us to assess the relevance of these shocks during the recent crisis. In fact, Figure 3.9 provides the historical interpretation of our results by displaying the

[19] Figure 8 in Appendix C.3 includes a figure with the first stage results.

component of unemployment explained by the BAS. In the upper panel, unemployment is expressed in deviation from the trend whereas, in the lower one, at the business cycle frequency.

The BAS explains the initial increase of unemployment, with respect to its trend, in 2010 and 2013 and also the reduction observed in 2014. Finally, it is also relevant to explain the increase observed during the last stage of 2014. Similar conclusions hold if we look the contribution at business cycle frequencies.

Our findings, which are robust across the two different identification strategies, suggest that liquidity shocks have significant effects on unemployment. These results also hold if we consider industrial production and the *ITA-coin*.[20] A question that may arise naturally is why this peculiar financial variable, not even on the focus of media's attention, has so strong real effects. First, we find that all the measures of confidence decline significantly in response to the decrease in liquidity. This could point to a decrease in aggregate demand that explains the decrease in economic activity (Ludvigson (2004)). Second, in Section 3.4, we show that commercial banks change their lending conditions in response to liquidity shocks.

3.3.4 Alternative VAR Specifications

Shocks to the BAS are a major driver of unemployment for the period under analysis. In this subsection, we consider additional specifications to assess the robustness of our findings. For the ease of exposition, the IRFs of the exercises performed in this section are presented in the Appendix C.4.

3.3.4.1 Indicator of Liquidity

The BAS is one of the most popular indicators of liquidity. However, it captures only the price dimension of liquidity while another relevant feature is the quantity side. A fall in liquidity equally distributed across price and quatities would generate an increase in the BAS and a fall in the quantity traded. In order to explore whether this relationship holds in our analysis, we estimate the Full VAR including the Turnover, volume traded normalized by the stock of the outstanding asset,

[20] Appendix C.4 displays the IRFs using each indicator.

as an additional variable in the system. While responses of macroeconomic variables to a BAS shock remain unchanged, the turnover displays a significant reduction. This result conforms with the theoretical predictions of the model proposed by Cui and Radde (2015).

In order to explicitly take this double dimension of liquidity into account, we compute a liquidity index indicator that is defined as the ratio between the Turnover and the BAS.[21] Thus, when the liquidity index is higher (lower), the asset can be considered more (less) liquid. We estimate the same baseline VAR but replacing the BAS with the Liquidity Index. Both responses of variables in the system and the contribution of liquidity to explain fluctuations in unemployment remain practically unchanged.

3.3.4.2 Measures of Economic Activity

All the results presented so far rely on Unemployment as a proxy for economic activity. Alternatively, we estimate the VAR including Industrial Production and a Coincident Indicator of Economic Activity (*Indicatore Ciclico Coincidente (ITA-coin)*), a monthly indicator of economic activity published by the Bank of Italy.[22] Results are comparable with the ones using Unemployment.

3.3.4.3 Different Samples

Figure 3.2 shows that financial variables display a change in volatility at daily frequency after 2009. Moreover, in the same window there is also a stark fall in interest rates that can constitute another source of structural break. To see whether this fact affects our findings, we estimate our baseline VAR for the sub-sample 2009-2014. The main results remain unchanged. To tackle the possibility that our results are driven only by the Euro crisis, we run the same analysis in 3.3.1 over the sample 2004-2009. Once again, we find very similar results in this short sample.

[21] The correct measure would employ the quantity bid and asked, but unfortunately we cannot access this data. Therefore, we use the actual number of trades (turnover on the secondary market) compiled by MTS.

[22] See `https://www.bancaditalia.it/statistiche/tematiche/indicatori/indicatore-ciclico-coincidente/` for more information about *ITA-coin*.

3.3.4.4 Corporate Liquidity

The finance literature has reported sizable fluctuations of the market liquidity of corporate bonds in the U.S during the financial crisis (see Dick-Nielsen et al. (2012)). Even if Italian firms rely more on banks as a source of finance, we analyze the interrelation between sovereign and corporate liquidity. For this aim, we use the BAS of a representative corporate bond and include it in the VAR instead of the Equity Premium.[23] A couple of interesting facts emerge. First, the effects of sovereign BAS shocks remain unchanged. Second, an exogenous increase in the private BAS generates a significant effect on Unemployment, which is comparable to the one induced by the sovereign BAS. Finally, an exogenous change in the private BAS does not affect significantly the sovereign BAS. These findings suggest that both BAS are relevant to explain economic activity. Finally, we also consider the BAS as a spread between the corporate and sovereign. A shock to this spread induces also sizable effects on economic activity.

3.3.4.5 Market Stress Index

As we show in Figure 3.3, the BAS reflects some relevant European events, which may be regarded as periods of Market Stress. To assess potential omitted variable biases, we replace the Equity Premium with the *Composite Indicator of Systemic Stress* (computed by the ECB) in our VAR. IRFs are comparable with respect to the baseline specification. Thus, these results confirm that our results are not biased by omitting other measures of stress in financial markets.

3.3.4.6 Financial Volatility

Financial variables display a time varying volatility at high frequency which is not reflected at monthly frequency. To control for these changes, we compute the monthly volatility of BAS, CDS and Spread using daily data. We build the first principal component that explains 78% of the variability of these three measures. Then, we include this index in the VAR instead of the Equity Premium. The IRFs and the FEV are unaffected. This suggests that previous findings are not driven by changes in volatility.

[23] We use the BAS of a bond issue by Telecom (*TELECOM ITALIA TITIM 5 3/8 01/19*) which is the longest series available. Moreover, it is highly correlated with the liquidity of the other bonds (e.g. 0,91 with Unicredit - *UCGIM 4 3/8 01/20* and 0,65 with ENI - *ENI INTERNATIONAL FINANCE ENIIM 5 1/27/19*. Source: Bloomberg.

3.4 Transmission Channels

The easiness of trading sovereign bonds is particularly relevant for Italian banks because they hold exceptional amounts of Italian sovereign debt. Gennaioli et al. (2014) show that banks hold large amounts of public bonds due to their liquidity properties. The *European Stress Test* carried out in 2010 provides some insights on the amount of these assets held by the main Italian commercial banks: Banca Popolare, Intesa San Paolo, Monte dei Paschi, UBI Banca and Unicredit. Italian banks' holding of national securities accounts for 74% of their total government bond holdings. This share is even higher if we consider only the trading book: 84%.[24] Moreover, Italian sovereign bonds constitute 6.13% of the total assets owned by those five Italian banks (Gennaioli et al. (2014)). In this Section, we assess whether and how changes in sovereign debt liquidity and spread affect banks' lending decisions using two official surveys. First, we employ the *ISTAT Business Confidence Survey*, which is carried out at monthly frequency. Second, we use the *Bank Lending Survey* from the Bank of Italy, which is available at quarterly frequency. Unlike statistics about total amount of loans that include both demand and supply effects, survey data allows us to disentangle more precisely the transmission channels.

3.4.1 ISTAT Business Confidence Survey

We employ data from the *ISTAT Business Confidence Survey* to assess the effects of liquidity and spread shocks on firms' credit conditions. This survey, which is carried out by ISTAT at a monthly frequency since March 2008, covers a representative sample of 4,000 firms in the manufacturing sector and includes information about firms' assessments and expectations on the Italian economic situation.[25] To assess how changes in sovereign debt liquidity and spread affect the credit market, we focus on questions regarding credit supply and demand and include them

[24]For regulatory purposes, banks divide their activities into two main categories: banking and trading. The trading book was devised to house market-related assets rather than traditional banking activities. Trading book assets are supposed to be highly liquid and easy to trade.

[25]See `http://siqual.istat.it/SIQual/visualizza.do?id=8888945&refresh=true&language=UK` for a detailed description of this survey. There is an analogous survey for the service sector but the sample is shorter. However, results are similar to the ones reported in this section.

as an additional variable in our baseline VAR.[26] Given that the sample is shorter, we estimate the baseline VAR described in section 3.3.2 since August 2009, when all the variables are available, including one variable at the time to avoid loosing degrees of freedom. In particular, we assume that credit decisions cannot react on impact to financial shocks and place these credit variables before the consumer confidence, business confidence and the financial block.[27] Figure 3.10 displays the IRFs to a liquidity deterioration and a positive sovereign spread shock.

Liquidity and sovereign spread shocks have different effects on the credit market. On the one hand, a BAS shock (i.e. a decrease of liquidity) does not change the index on perceived credit conditions but induces worse conditions in terms of interest rate, size of the credit, and costs other than the interest rate. Moreover, the BAS leads to an rise in the number of denied loans by banks with a lag. On the other hand, a spread shock immediately reduces the credit access and increases the number of denied loans by banks and a rise in the interest rate charged by banks. Notably, the reason why credit is not obtained by firms (*credit not obtained*) is not related with firms rejecting the loans offered by the banks (*credit not obtained - too heavy conditions*), but due to banks denying the loan (*credit not obtained - bank denial*). In other words, credit supply is driving the lower access to credit. While the spread shock affects mostly the interest rate and the size of the credit, a liquidity shock also induces higher costs (apart from the interest rate). These higher costs reflect higher commissions, extra-costs and tighter deadlines. For what concerns the timing, we observe a more lagged response to a liquidity shock than to a spread one. This is consistent with the delayed response of financial variables presented in Section 3.3.3.

After analyzing firm's survey responses, in the next subsection we assess whether these results are consistent with bank's replies. Additionally, we investigate the reasons that drive banks behavior.

3.4.2 Bank Lending Survey

We exploit the *Bank Lending Survey (BLS)* on Italian commercial banks to determine the effects of liquidity and spread shocks. This survey, which is carried out by *Banca d'Italia* in collaboration with the *European Central Bank* at quarterly

[26] TheAppendix contains the questions that we consider from the ISTAT Business Confidence Survey.

[27] Results remain unchanged if we place this variable last in the VAR.

frequency since January 2003, contains very detailed information about bank's decisions on different dimensions.[28] Unlike in the previous subsection, we cannot include the replies to the survey in the baseline VAR due to the differences in frequencies. For this reason, we aggregate the monthly BAS and spread shocks identified in section 3.3.2 to quarterly frequency and estimate the following equation:

$$\Delta BLS_t^i = \alpha + \sum_{j=1}^{8} \delta_j \Delta BLS_{t-j}^i + \sum_{j=0}^{12} \beta_j shock_{t-j}^k \tag{3.1}$$

where $\Delta BLS_t^i, shock_t^k$ denote the change in bank's behavior and quarterly BAS and spread shocks, respectively. We follow Romer and Romer (2004) and choose eight lags for the autoregressive part and twelve for the effect of the shock. Then, we compute the IRF to a BAS and spread shock for the main bank decisions available in the Survey (Figure 3.11).[29]

Banks increase their credit standards to firms in response to liquidity and spread shocks with a similar magnitude. However, the reasons for increasing standards differ. On the one hand, in response to negative liquidity shock, banks react due to changes in their liquidity position and costs related to their capital position. On the other hand, banks do not report changes in the relevance of the asset and liquidity position in response to a spread shock. These differences in behavior suggest that banks increase their focus on their own balance sheet in case of a liquidity deterioration in sovereign debt markets. Moreover, banks adjust immediately their standards for mortgage loans while they do not change it for the case of spread shocks. Mortgages are collateralized loans and, in case of no repayment and liquidity problems, banks may not find it easy to release the house and that may explain why they increase their standards. Finally, both shocks are associated with an increase of similar magnitude in the perception of risk about economic activity.

With the evidence presented in Sections 3.3 and 3.4, we conclude that liquidity shocks have relevant real effects on the Italian economy and we document that transmission is through changes in the credit supply. In the next section, we analyze whether liquidity shocks are also relevant for the other three major Eurozone economies: Germany, France, and Spain.

[28] More information about this survey can be found at BLS .

[29] The Data Appendix contains the detailed questions we consider from the Bank and Lending Survey.

3.5 Comparison with other European Countries

In order to assess whether liquidity shocks are also relevant drivers of the business cycle in other European economies, we perform the previous analysis also for Germany, France, and Spain. First, in Table 3.2 we analyze if sovereign BAS are correlated across countries, which would indicate to what extent they are explained by common shocks. We observe that BASs are positively correlated across the biggest four Eurozone economies. While BAS for Germany seems to be less correlated with the rest of the countries, the correlation is stronger between France, Italy and Spain.

Second, we estimate the baseline VAR described in Section 3.3.2 for each country to determine whether the macroeconomic results for Italy also hold for the other countries.[30] A first relevant finding is that the identified BAS shocks are positively correlated across countries: the correlation ranges from 0.3, France-Germany, to 0.21, France-Italy.[31] Both the correlation of the variables in levels and of the shocks indicate that liquidity in sovereign markets is driven by a relevant European component.

We present the macroeconomic relevance of the financial shocks, across the four countries, in Figure 11 through the FEV decomposition of unemployment. There is a clear heterogeneity between the Mediterranean countries and the central European ones. On the one hand, changes in BAS are an important driver of unemployment for Spain and Italy. For both cases, BAS shocks account for 15% of unemployment fluctuations.[32] A special feature of Spain is the relevance of CDS, which might be due to the perceived higher default risk. On the other hand, exogenous fluctuations in stock markets are the most relevant source of unemployment fluctuations for Germany and France. In fact, neither BAS nor sovereign spread seem to be relevant to explain unemployment fluctuations in these countries. Even if financial shocks explain a similar fraction of the total variability of unemployment (around 30%), the relevance of each financial shock differs across countries. Although the sources of this difference are beyond the scope of this

[30] The sample is February 2004-November 2014 for Germany, Italy and Spain. Due to the lack of CDS data before 2005, the sample for France starts in August 2005. All financial variables are expressed as monthly averages.

[31] In particular, the estimated cross-country correlations are statistically significant for all the cases but between France and Spain.

[32] Moreover, the IRF to a BAS shock has similar effects both in terms of magnitude and persistence.

paper, one possible reason could be the lower tensions in sovereign debt markets in France and Germany. Moreover, while Italian and Spanish banks are heavily exposed to their national sovereign debt (around 75% in 2010 according to the European Stress Test), French and German financial institutions hold a more diversified portfolio.

3.6 Conclusions

Economists have been focusing on sovereign debt markets due the European Sovereign Debt Crisis. Contrary to the growing number of theoretical models that analyze changes in liquidity in those markets, the empirical evidence on their real effects is still null. In this paper, we provide novel empirical evidence on the macroeconomic effects of changes in liquidity in secondary sovereign debt markets. We focus on the Italian economy that was hit both by credit risk and liquidity shocks during the recent crisis. We use monthly data from 2004 to 2014 in a VAR analysis and consider two alternative identification strategies: recursive ordering and the Proxy-SVAR, which yield consistent results. The former takes into account all the possible orderings among financial variables. The Proxy-SVAR exploits a daily financial VAR to control for all high-frequency changes in financial markets. Specifically, we use daily BAS structural shocks as proxy for the monthly BAS structural shocks. We find that, contrary to popular perceptions, liquidity is a major financial driver of economic activity. An exogenous raise in this variable generates a strong (15% of the Forecast Error Variance) and persistent (10 months) surge in unemployment. The other variables that are mostly affected are confidence indicators as Stock Prices, and Consumer and Business Sentiment. Banks and firms survey data reveal that liquidity shocks have significant effects on banks standard, in terms of loan's size and through additional costs, particularly due to the asset and liquidity position of Italian banks. Similar macroeconomic effects hold for Spain, whereas liquidity shocks are not a significant driver for France and Germany.

Our results differ from existing models, as Del Negro et al. (2011) and Benigno and Nistico (2014), where liquidity shocks induce a pronounced deflation. Therefore, in particular in the light of our findings related to the banking channel, we believe that models that focus on the asset and liquidity position of financial intermediaries can enhance our understanding of these phenomena. We regard Cui

and Radde (2015) as a first step towards this interesting direction for future research. Frameworks of this kind, which can generate macroeconomic effects consistent with the empirical evidence, can be used to assess whether and how policy makers should react to changes in liquidity (Cui (2016)). They mainly focus on the liquidity of corporate bonds as their reference is the US economy. Instead, by studying European economies we conclude that the liquidity of sovereign bonds is a key financial dimension for the business cycle. Liquidity shocks to these two different assets may involve diverse policy reactions and have different implications.

3.7 Tables

Levels	BAS	Spread	CDS
BAS	1	0.24***	0.36***
Spread	0.24***	1	0.91***
CDS	0.36***	0.91***	1

Growth Rates	BAS	Spread	CDS
BAS	1	-0.03	-0.03
Spread	-0.03	1	0.23
CDS	-0.03	0.23***	1

Table 3.1: Contemporaneous correlation between financial variables

*Contemporaneous daily correlation between Italian financial variables at daily frequency: BAS, Spread, CDS. All the variables correspond to 2 years maturity. Left-panel in levels, right-panel in growth rates. ***, **, * denote 99%, 95% and 90% confidence intervals.*

	Italy	Spain	France	Germany
Italy	1	0.49***	0.56***	0.24***
Spain	0.49***	1	0.69***	0.32***
France	0.56***	0.69***	1	0.42***
Germany	0.24***	0.32***	0.42***	1

Table 3.2: Daily correlation of European BAS

Daily correlations of 2 year sovereign BAS across countries (source: Bloomberg).

3.8 Figures

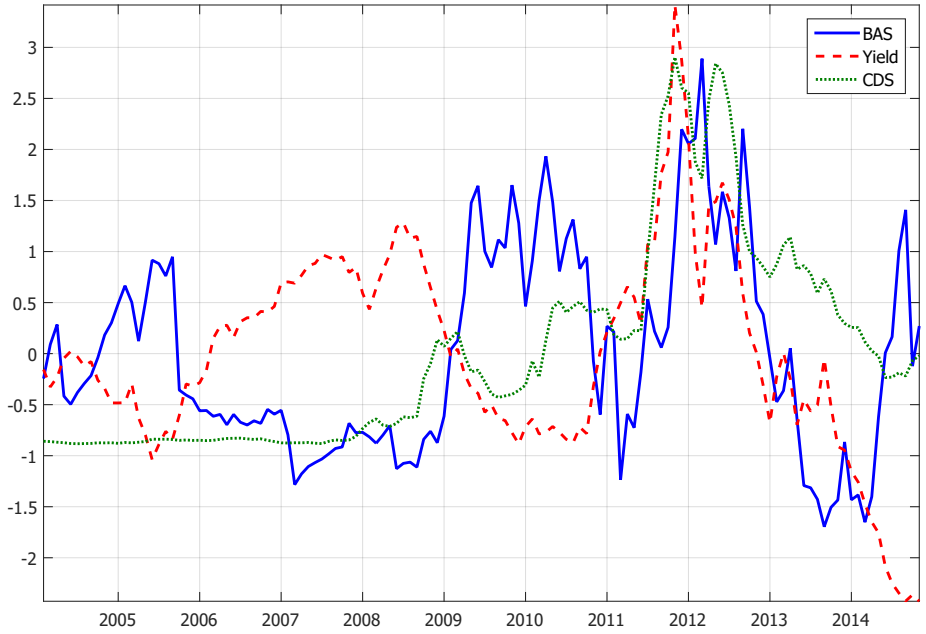

Figure 3.1: Key financial variables

Italian (standardized) BAS, CDS and Yield (monthly average). Each variable corresponds to the first principal components of 2, 5, 10 years bond maturities. Source: Bloomberg (BAS) and Banca d'Italia.

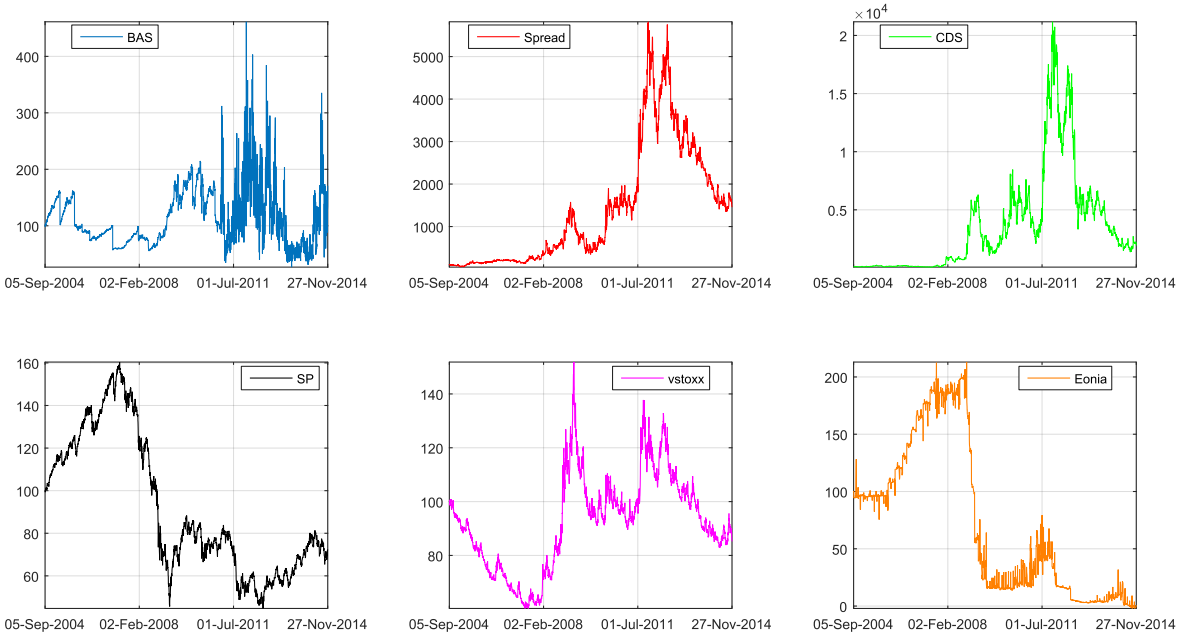

Figure 3.2: Daily dynamics of the main financial variables

Financial variables: BAS Italy, Spread Italy, CDS Italy, FTSE MIB (main Italian Stock Price index), Vstoxx (European Implied Volatiliy Index), Euro Overnight Index Average (Eonia). All variables are expressed in levels for all the business days since September 2004 to November 2014. All variables but the Spread are expressed as an index=100 at the beginning of the sample. Spread is computed as the difference between German and Italian yields and expressed in basis points times 10.

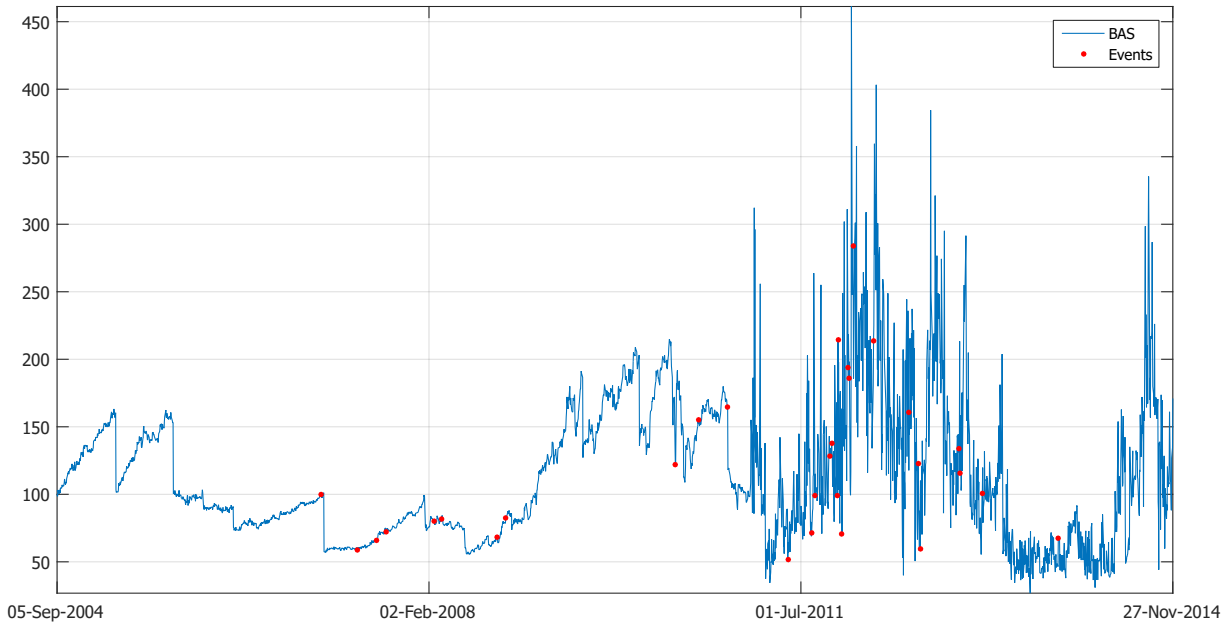

Figure 3.3: Daily BAS and key European events

Daily BAS Italy 2 Years (blue line) and key European events (red dots). Appendix A displays the list of all the events.

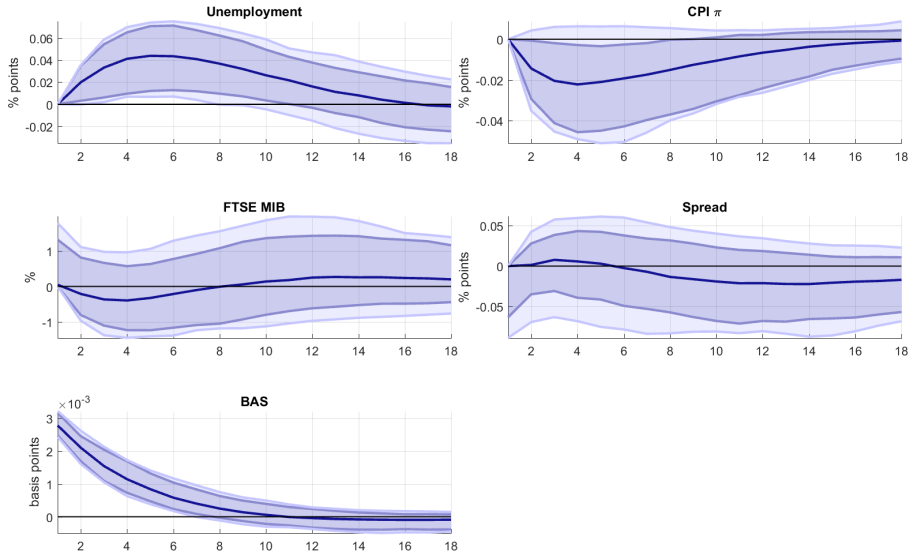

Figure 3.4: IRF to a BAS shock in the small system

IRFs to a 1 std BAS shock (liquidity deterioration) identified through the following ordering [Unemployment, π, FTSE, Spread, BAS]. The median point estimate, 68% and 90% confidence bands are reported in blue and light blue, respectively. 50%, 68% and 90% bands include statistical and identification uncertainty (from all the possible ordering within the financial block).

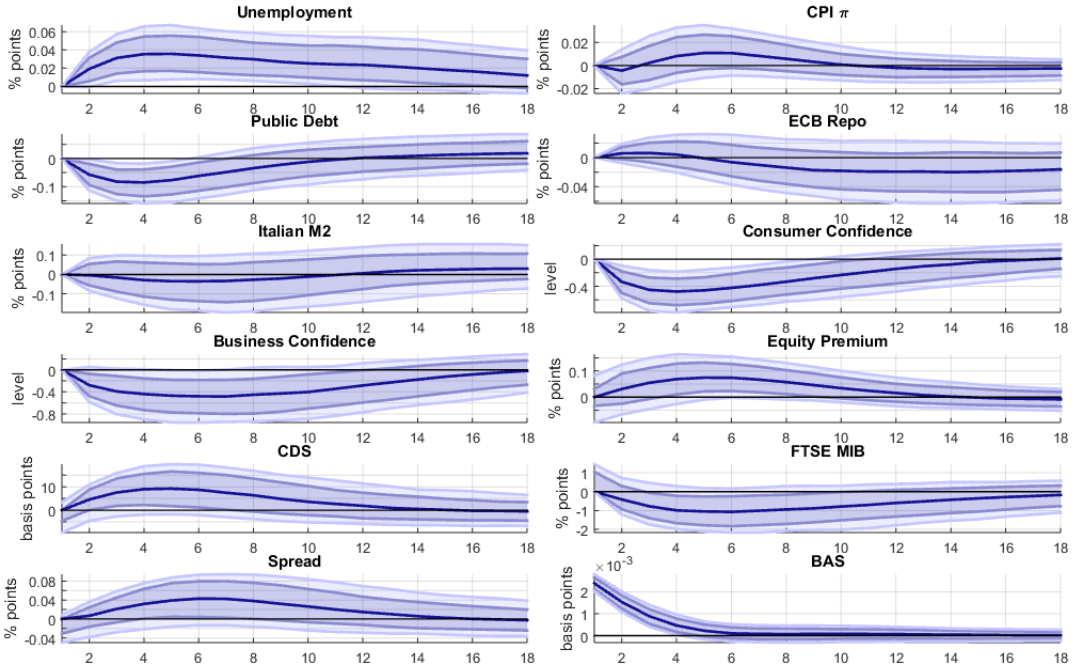

Figure 3.5: IRF to a BAS shock in the large system

IRFs to a 1 std deviation BAS shock (liquidity deterioration) identified through the following ordering [Unemployment, π, Public Debt, R, M2, CC, BC, Financial Block]. The median point estimate, 68% and 90% confidence bands are reported in blue and light blue, respectively. 50%, 68% and 90% bands include statistical and identification uncertainty (from all the possible ordering within the financial block).

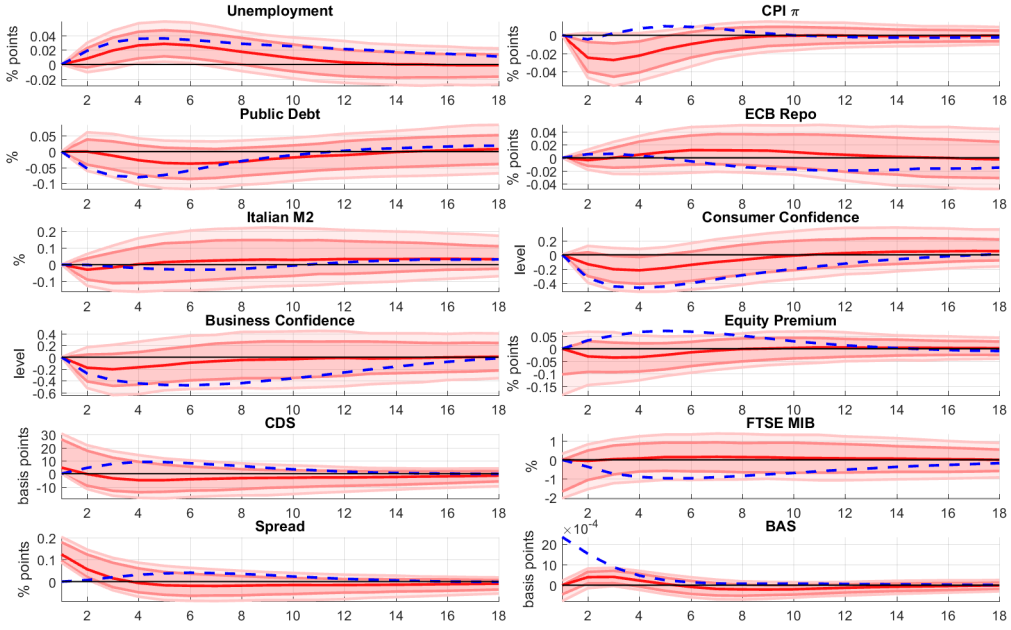

Figure 3.6: IRF to a Spread shock

IRFs to a 1 std deviation Spread shock identified through the following ordering [Unemployment, π, Public Debt, R, M2, CC, BC, Financial Block]. The median point estimate, 68% and 90% confidence bands are reported in red and light red, respectively. 50%, 68% and 90% bands include statistical and identification uncertainty (from all the possible ordering within the financial block). Dotted line denotes the mean response to a 1 std deviation shock to BAS.

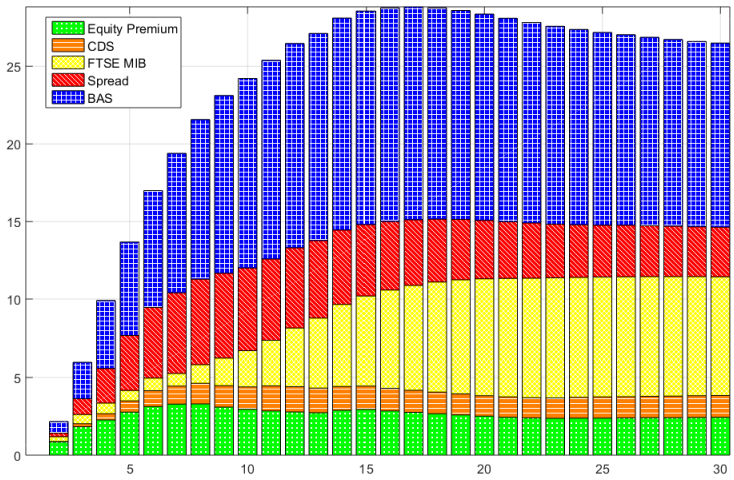

Figure 3.7: FEV of unemployment

FEV of Unemployment in the VAR [Unemployment, π, Public Debt, R, M2, CC, BC, Financial Block]. The bars denote the contribution of each financial shock in explaining the volatility of Unemployment at each horizon (expressed in months).

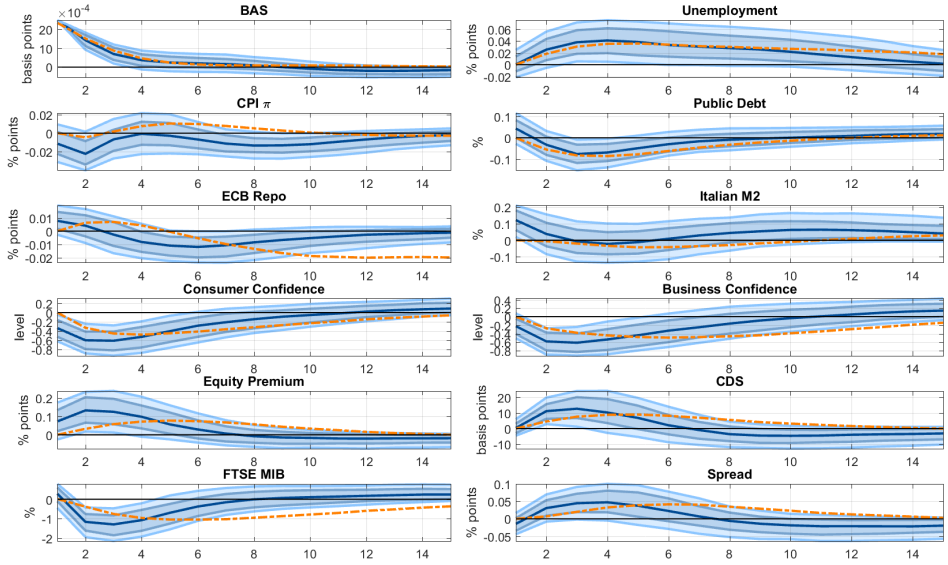

Figure 3.8: IRF to a BAS shock: Bridge Proxy-SVAR

IRFs to a 1 standard deviation BAS shock (liquidity deterioration) in the VAR [Unemployment, π, Public Debt, R, M2, CC, BC, Financial Block]. The shock is identified through the unpredictable variation of the BAS in a daily VAR system. Sample: Jan:2009-Nov:2014. The median point estimate, 68% and 90% confidence bands are reported in blue and light blue, respectively. Confidence bands are computed using wild bootstrap with 1,000 replications. Dotted lines denote the mean responses of each variable to a 1 standard deviation BAS shock identified via recursive ordering.

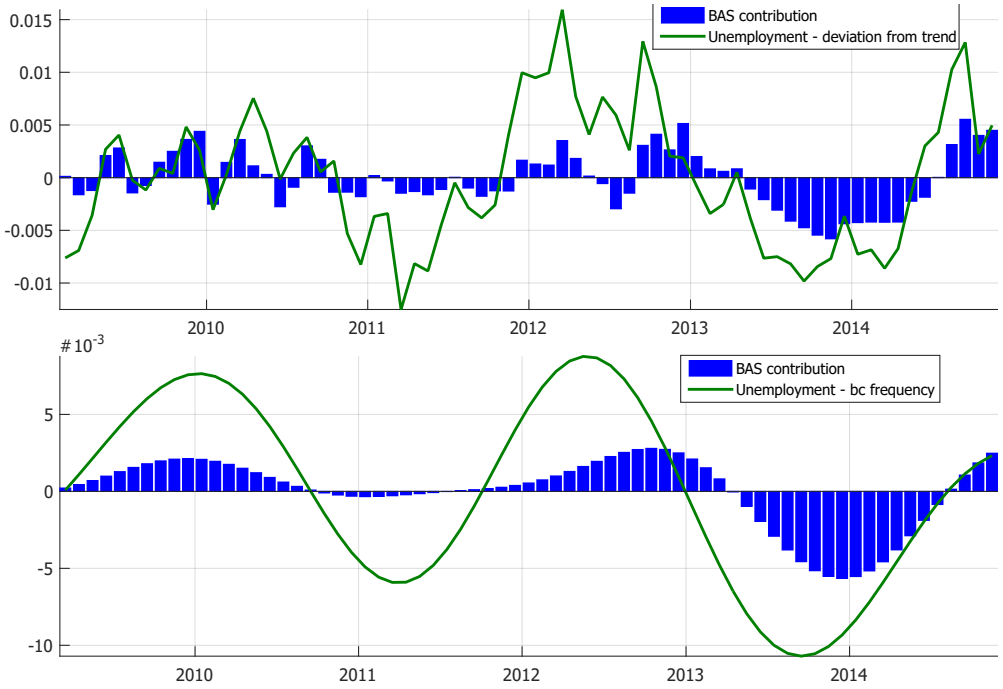

Figure 3.9: Historical contribution of BAS to unemployment: Bridge Proxy-SVAR

Historical contribution of BAS to Unemployment. Identified in the VAR [Unemployment, π, Public Debt, R, M2, CC, BC, Financial Block] through the unpredictable variation of the BAS in a daily VAR system. Upper panel - Unemployment in deviation from trend. Lower panel - Unemployment at the business cycle frequency (18 to 96 months).

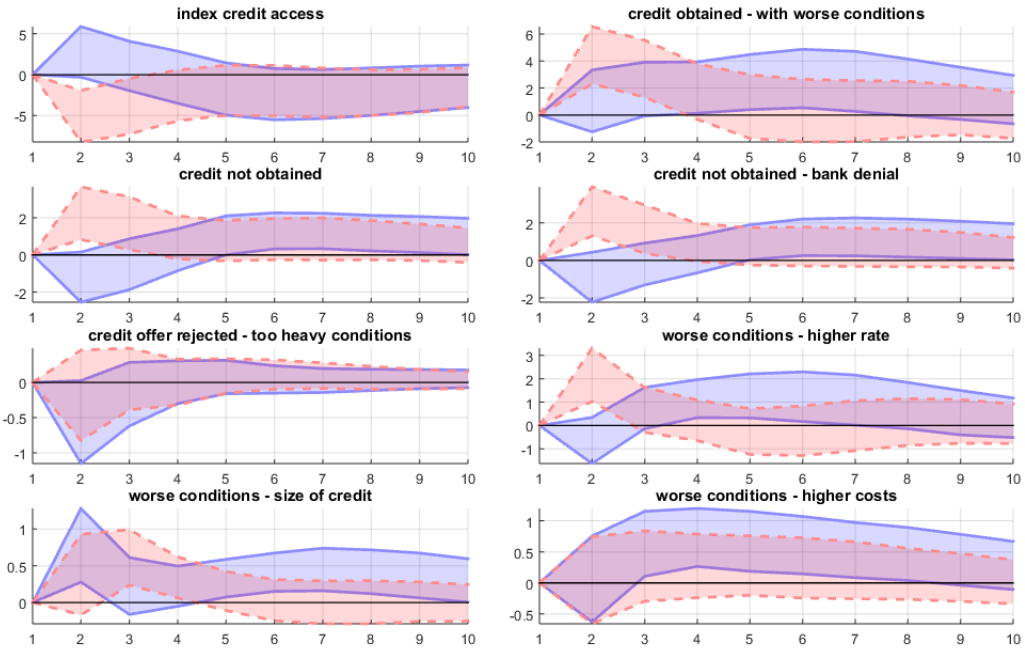

Figure 3.10: Changes in credit market conditions for manufacturing firms

NOTE. *Changes in the credit market for manufacturing firms in response to a one standard positive BAS (blue) and sovereign spread (red) shocks. All figures denote change in the corresponding index reported by ISTAT. Blue and red areas denote the 68% confidence intervals computed using bootstrap and include both identification and statistical uncertainty.*

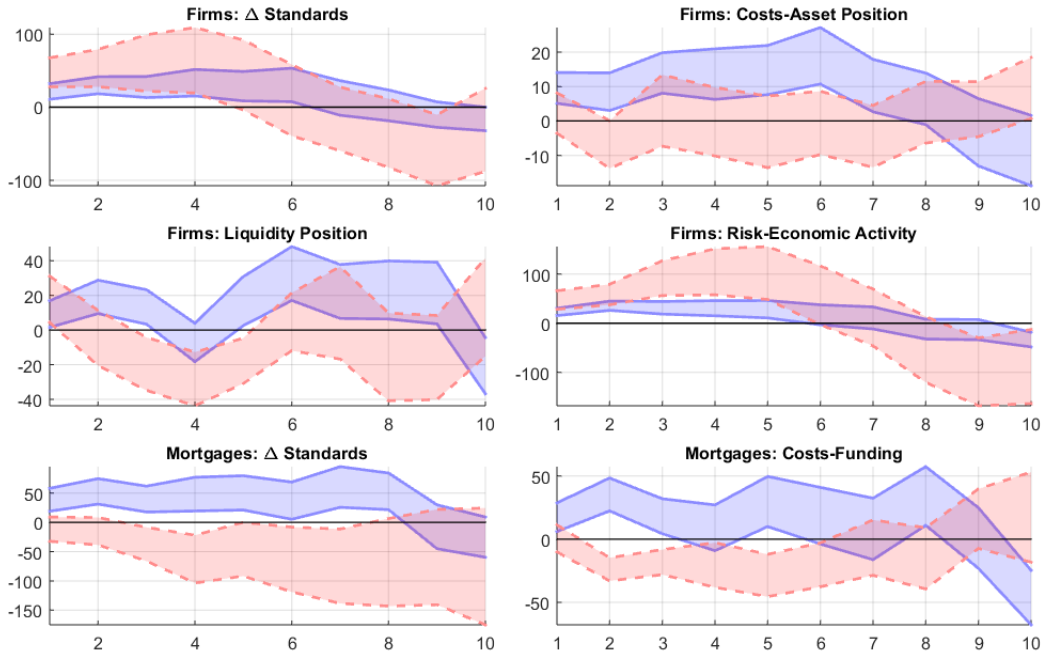

Figure 3.11: Change in banks lending decisions

Change in banks decisions in response to a positive shock in BAS and Spread. All the figures denote the change in the corresponding index as reported in the BLS. Blue and red areas denote the 90% confidence intervals computed using 500 bootstrap replications.

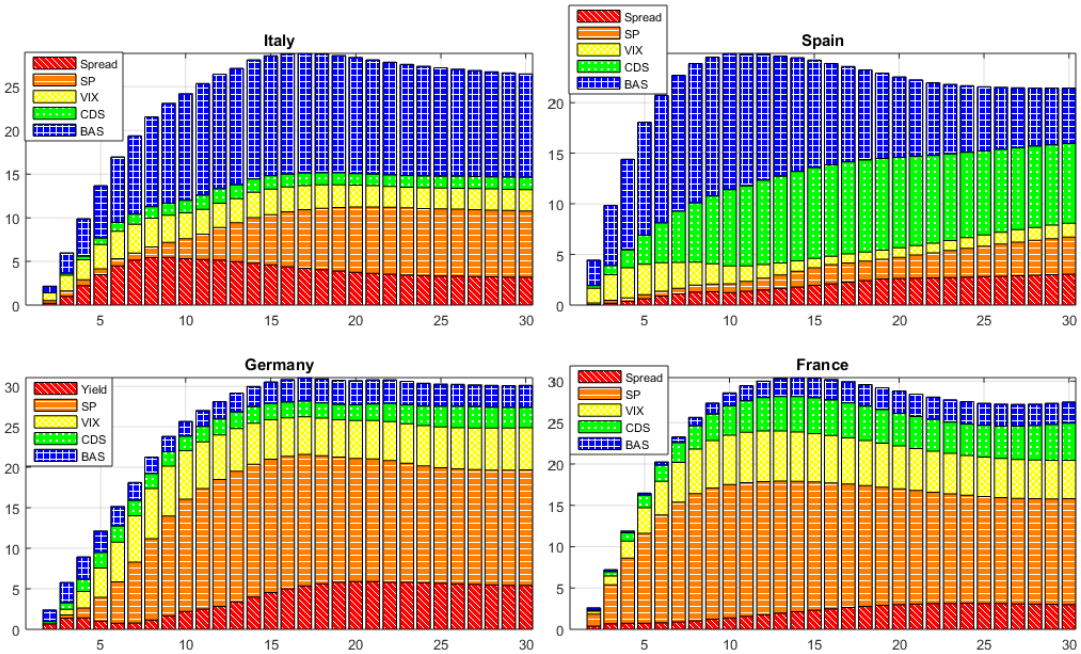

Figure 3.12: FEVD of unemployment for European countries

FEVD *of Unemployment for Italy, France, Germany, and Spain. The FEVD is computed estimating a VAR for each country that includes: [Unemployment, π, Public Debt, R, M2, CC, BC, Financial Block]. BAS shocks are identified from all the possible rotations across the financial variables.*

Bibliography

ADAM, K., P. KUANG, AND A. MARCET (2012): "House Price Booms and the Current Account," *NBER Macroeconomics Annual*, 26, 77–122.

AGNELLO, L. AND L. SCHUKNECHT (2011): "Booms and busts in housing markets: Determinants and implications," *Journal of Housing Economics*, 20, 171 – 190.

AMBROSE, B. W., P. EICHHOLTZ, AND T. LINDENTHAL (2013): "House Prices and Fundamentals: 355 Years of Evidence," *Journal of Money, Credit and Banking*, 45, 477–491.

ANDERSON, B., M. DEISTLER, E. FELSENSTEIN, B. FUNOVITS, AND L. KOELBL (2016a): "The Structure of Multivariate AR and ARMA Systems: Regular and Singular Systems; the Singular and the Mixed Frequency Case," *Journal of Econometrics*, 192, 366–373.

ANDERSON, B., M. DEISTLER, E. FELSENSTEIN, B. FUNOVITS, L. KOELBL, AND M. ZAMANI (2016b): "Multivariate AR Systems and Mixed Frequency Data: G-Identifiability and Estimation," *Econometric Theory*, 32, 793–826.

ANGELETOS, G.-M. AND J. LA'O (2013): "Sentiments," *Econometrica*, 81, 739–779.

ANGELINI, E., M. BANBURA, AND G. RUNSTLER (2010): "Estimating and Forecasting the Euro Area Monthly National Accounts from a Dynamic Factor Model," *OECD Journal: Journal of Business Cycle Measurement and Analysis*, 2010, 1–22.

BAFFIGI, A., R. GOLINELLI, AND G. PARIGI (2004): "Bridge Models to Forecast the Euro Area GDP," *International Journal of Forecasting*, 20, 447–460.

BAHAJ, S. (2014): "Systemic Sovereign Risk: Macroeconomic Implications in the Euro Area," .

BARSKY, R. B. AND E. R. SIMS (2012): "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence," *American Economic Review*, 102, 1343–77.

BEAUDRY, P. AND F. PORTIER (2006): "Stock Prices, News, and Economic Fluctuations," *The American Economic Review*, 96, 1293–1307.

BEBER, A., M. BRANDT, AND K. KAVAJECZ (2009): "Flight-to-Quality or Flight-to-Liquidity? Evidence from the Euro-Area Bond Market," *Review of Financial Studies*, 22, 925–957.

BENIGNO, P. AND S. NISTICO (2014): "Safe Assets, Liquidity and Monetary Policy," .

BLANCHARD, O. J., J.-P. L'HUILLIER, AND G. LORENZONI (2013): "News, Noise, and Fluctuations: An Empirical Exploration," *American Economic Review*, 103, 3045–70.

BLUWSTEIN, K. AND F. CANOVA (2015): "Beggar-thy-Neighbor? The International Effects of ECB Unconventional Monetary Policy Measures," .

BOISSEL, C., F. DERRIEN, E. ORS, AND D. THESMAR (2014): "Sovereign Debt Crisis and Bank Financing: Evidence from the European Repo Market," .

CAINES, C. (2015): "Can Adaptive Learning Explain Boom-Bust Cycles In Asset Prices? An Application to the US Housing Boom," *Working Paper*.

CAMPBELL, J. R., C. L. EVANS, J. D. FISHER, AND A. JUSTINIANO (2012): "Macroeconomi Effects of Federal Reserve Forward Guidance," Federal Reserve Bank of Chicago, Working Paper Series WP-2012-03.

CAMPBELL, J. R., J. D. FISHER, A. JUSTINIANO, AND L. MELOSI (2016): "Forward Guidance and Macroeconomic Outcomes Since the Financial Crisis," in *NBER Macroeconomics Annual 2016*, National Bureau of Economic Research, Inc, NBER Chapters.

CAMPBELL, J. Y. AND R. SHILLER (1988): "Stock Prices, Earnings, and Expected Dividends," *The Journal of Finance*, 43, 661–676.

CAMPBELL, S. D., M. A. DAVIS, J. GALLIN, AND R. F. MARTIN (2009): "What moves housing markets: A variance decomposition of the rent to price ratio," *Journal of Urban Economics*, 66, 90 – 102.

CANOVA, F. AND M. H. SAHNEH (2016): "Are Small-Scale SVARs Useful for Business Cycle Analysis? Revisiting Non-Fundamentalness," *Working Paper*.

CAPASSO, M. AND A. MONETA (2016): "Macroeconomic Responses to an Independent Monetary Policy Shock: a (More) Agnostic Identification Procedure," LEM Working Paper 2016/36.

CARRIERO, A., H. MUMTAZ, K. THEODORIDIS, AND A. THEOPHILOPOULOU (2015): "The Impact of Uncertainty Shocks under Measurement Error: a Proxy SVAR Approach," *Journal of Money, Credit and Banking*, 47, 1223–1238.

CASE, K., R. SHILLER, AND A. K. THOMPSON (2015): "What Have They Been Thinking? Home Buyer Behavior in Hot and Cold Markets – A 2014 Update," Cowles Foundation Discussion Papers 1876R, Cowles Foundation for Research in Economics, Yale University.

CASE, K. E., J. M. QUIGLEY, AND R. J. SHILLER (2005): "Comparing wealth effects: The stock market versus the housing market," *The B.E. Journal of Macroeconomics*, 5, 1–34.

——— (2012): "Wealth Effects Revisited 1975-2012," *Critical Finance Review*, 2, 101–128.

COCHRANE, J. H. (1994): "Shocks," *Carnegie-Rochester Conference Series on Public Policy*, 41, 295–364.

COIBION, O., Y. GORODNICHENKO, L. KUENG, AND J. SILVIA (2012): "Innocent Bystanders? Monetary Policy and Inequality in the U.S." NBER Working Paper No.18170.

CROWE, C., G. DELL'ARICCIA, D. IGAN, AND P. RABANAL (2013): "How to deal with real estate booms: Lessons from country experiences," *Journal of Financial Stability*, 300–319.

CUI, W. (2016): "Monetary-Fiscal Interactions with Endogenous Liquidity Frictions," Unpublished manuscript.

CUI, W. AND S. RADDE (2015): "Search-based Endogenous Illiquidity and the Macroeconomy," Unpublished manuscript.

DE MARCO, F. (2016): "Bank Lending and the European Sovereign Debt Crisis," .

DEL NEGRO, M., G. EGGERTSSON, A. FERRERO, AND N. KIYOTAKI (2011): "The Great Escape? A Quantitative Evaluation of the Fed's Liquidity Facilities," *Federal Reserve Bank of New York Staff Reports*, 520.

DELL'ARICCIA, G., D. IGAN, L. LAEVEN, AND H. TONG (2016): "Credit booms and macrofinancial stability," *Economic Policy*, 31, 299.

DICK-NIELSEN, J., P. FELDHUTTER, AND D. LANDO (2012): "Corporate Bond Liquidity Before and After the Onset of the Subprime Crisis," *Journal of Financial Economics*, 103, 471–492.

DIRON, M. (2008): "Short-Term Forecasts of Euro Area Real GDP Growth. An Assesment of Real-Time Performance Based on Vintage Data," *Journal of Forecasting*, 27, 371–390.

EICHHOLTZ, P., S. STRAETMANS, AND M. THEEBE (2012): "The Amsterdam rent index: The housing market and the economy, 1550-1850," *Journal of Housing Economics*, 21, 269 – 282.

ENGSTED, T., S. J. HVIID, AND T. Q. PEDERSEN (2016): "Explosive bubbles in house prices? Evidence from the OECD countries," *Journal of International Financial Markets, Institutions and Money*, 40, 14 – 25.

ERAKER, B., C. W. CHIU, A. FOERSTER, T. B. KIM, AND H. SEOANE (2015): "Bayesian Mixed Frequency VARs," *Journal of Financial Econometrics*, 13, 698–721.

FORNI, M. AND L. GAMBETTI (2014): "Sufficient information in structural VARs," *Journal of Monetary Economics*, 66, 124 – 136.

FORNI, M., L. GAMBETTI, M. LIPPI, AND L. SALA (2016): "Noise Bubbles," *The Economic Journal*, n/a–n/a.

——— (forth): "Noisy News in Business Cycles," *American Economic Journal: Macroeconomics*.

FORONI, C., E. GHYSELS, AND M. MARCELLINO (2013): "Mixed-Frequency Vector Autoregressive Models," *Advances in Econometrics*, 32, 247–272.

FORONI, C. AND M. MARCELLINO (2016): "Mixed Frequency Structural VARs," Forthcoming, Journal of the Royal Statistical Society-Series A.

GALLIN, J. (2008): "The Long-Run Relationship Between House Prices and Rents," *Real Estate Economics*, 36, 635–658.

GARCIA, J. A. AND R. GIMENO (2014): "Flight-to-Liquidity Flows in the Euro Area Sovereign Debt Crisis," *Banco de Espana Working Paper*, 1429.

GARRIGA, C., R. E. MANUELLI, AND PERALTA-ALVA (2012): "A Model of Price Swings in the Housing Market," *Federal Reserve Bank of St Louis Working Paper 2012-022A*.

GAZZANI, A. AND A. VICONDOA (2016): "The Real Effects of Liquidity Shocks in Sovereign Debt Markets: Evidence from Italy," .

——— (2016b): "Proxy SVAR as a Bridge between Mixed Frequencies," Unpublished manuscript.

GELAIN, P. AND K. J. LANSING (2014): "House prices, expectations, and time-varying fundamentals," *Journal of Empirical Finance*, 29, 3 – 25.

GENNAIOLI, N., A. MARTIN, AND S. ROSSI (2014): "Banks, Government Bonds, and Default: What do the Data Say?" *IMF Working Paper*, 14/120.

GERTLER, M. AND P. KARADI (2015): "Monetary Policy Surprises, Credit Costs, and Economic Activity," *American Economic Journal: Macroeconomics*, 7, 44–76.

GHYSELS, E. (2016): "Macroeconomics and the Reality of Mixed Frequency Data," *Journal of Econometrics*, 193, 294–314.

GIGLIO, S., M. MAGGIORI, AND J. STROEBEL (2016): "No-Bubble Condition: Model-Free Tests in Housing Markets," *Econometrica*, 84, 1047–1091.

GLAESER, E. L., J. GYOURKO, E. MORALES, AND C. G. NATHANSON (2014): "Housing dynamics: An urban approach," *Journal of Urban Economics*, 81, 45 – 56.

GOURIEROUX, C., A. MONFORT, AND J.-P. RENNE (2017): "Statistical Inference for Independent Component Analysis: Application to Structural VAR Models," *Journal of Econometrics*, 196, 111–126.

GRANZIERA, E. AND S. KOZICKI (2015): "House price dynamics: Fundamentals and expectations," *Journal of Economic Dynamics and Control*, 60, 152 – 165.

GUERRIERI, L. AND M. IACOVIELLO (2016): "Collateral Constraints and Macroeconomic Asymmetries," *Working Paper*.

GURKAYNAK, R., B. SACK, AND E. SWANSON (2005): "Do Actions Speak Louder than Words? The Response of Asset Prices to Monetary Policy Actions and Statements," *International Journal of Central Banking*, 1, 55–93.

HARDING, D. AND A. PAGAN (2002): "Dissecting the cycle: a methodological investigation," *Journal of Monetary Economics*, 49, 365 – 381.

IACOVIELLO, M. AND S. NERI (2010): "Housing Market Spillovers: Evidence from an Estimated DSGE Model," *American Economic Journal: Macroeconomics*, 2, 125–64.

JENTSCH, C. AND K. LUNSFORD (2016): "Proxy SVARs: Asymptotic Theory, Bootstrap Inference, and the Effects of Income Tax Changes in the United States," Federal Reserve Bank of Cleveland Working Paper No.16-19.

JORDA, O., S. M., AND T. A. M. (2015): "Betting the house," *Journal of International Economics*, 96, Supplement 1, S2 – S18, 37th Annual NBER International Seminar on Macroeconomics.

——— (forth): "Leveraged Bubbles," *Journal of Monetary Economics*.

LAKDAWALA, A. K. (2016): "Decomposing the Effects of Monetary Policy Using an External Instruments SVAR," Working Paper.

LIPPI, M. AND L. REICHLIN (1994): "VAR analysis, nonfundamental representations, blaschke matrices," *Journal of Econometrics*, 63, 307 – 325.

LORENZONI, G. (2009): "A Theory of Demand Shocks," *The American Economic Review*, 99, 2050–2084.

LUDVIGSON, S. (2004): "Consumer Confidence and Consumer Spending," *Journal of Economic Perspectives*, 18, 29–50.

LUDVIGSON, S., S. MA, AND S. NG (2015): "Uncertainty and Business Cycles: Exogenous Impulse or Endogenous Response?" NBER Working Paper No.21803.

LUNSFORD, K. (2015): "Identifying Structural VARs with a Proxy Variable and a Test for Weak Proxy," Federal Reserve Bank of Cleveland, Working Paper No. 15-28.

LUTKEPOHL, H. AND G. MILUNOVICH (2015): "Testing for Identification in SVAR-GARCH Models: Reconsidering the Impact of Monetary Policy Shocks on Exchange Rates," Discussion Papers of DIW Berlin 1455.

MANCINI, L., A. RANALDO, AND J. WRAMPELMEYER (2014): "The Euro Interbank Repo Market," *Swiss Finance Institute Research Paper*.

MARCELLINO, M. (1999): "Some Consequences of Temporal Aggregation in Empirical Analysis," *Journal of Business and Economic Statistics*, 17, 129–136.

MARIANO, R. AND Y. MURASAWA (2010): "A Coincident Index, Common Factors, and Monthly Real GDP," *Oxford Bulletin of Economics and Statistics*, 72, 27–46.

MERTENS, K. AND M. O. RAVN (2010): "Measuring the Impact of Fiscal Policy in the Face of Anticipation: A Structural VAR Approach*," *The Economic Journal*, 120, 393–413.

——— (2013): "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States," *American Economic Review*, 103, 1212–47.

——— (2014): "A Reconciliation of SVAR and Narrative Estimates of Tax Multipliers," *Journal of Monetary Economics*, 68, S1–S19.

NERI, S. AND T. ROPELE (2015): "The Macroeconomic Effects of the Sovereign Debt Crisis in the Euro Area," *Bank of Italy Working Papers*, 1007.

PASSADORE, J. AND Y. XU (2014): "Illiquidity in Sovereign Debt Markets," .

PELIZZON, L., M. SUBRAHMANYAM, D. TOMIO, AND J. UMO (2015): "Sovereign Credit Risk, Liquidity, and ECB Intervention: Deus Ex Machina?" *SAFE Working Paper No. 95*.

PERICOLI, M. AND M. TABOGA (2015): "Decomposing Euro Area Sovereign Spreads: Credit, Liquidity and Convenience," *Bank of Italy Working Papers*, 1021.

PIAZZESI, M. AND M. SCHNEIDER (2009a): "Inflation and the Price of Real Assets," *Federal Reserve Bank of Minneapolis*, Research Department Staff Report 423.

——— (2016): "Housing and Macroeconomics," *Handbook of Macroeconomics Volume 2*.

POTERBA, J. (1991): "House Price Dynamics: The Role of Tax Policy," *Brookings Papers on Economic Activity*, 22, 143–204.

POTERBA, J. M. (1984): "Tax Subsidies to Owner-Occupied Housing: An Asset-Market Approach," *The Quarterly Journal of Economics*, 99, 729.

RAMEY, V. (2016): "Macroeconomic Shocks and Their Propagation," NBER Working Papers 21978.

RIGOBON, R. (2003): "Identification through Heteroskedasticity," *The Review of Economics and Statistics*, 85, 777–792.

ROMER, C. D. AND D. H. ROMER (2004): "A New Measure of Monetary Shocks: Derivation and Implications," *American Economic Review*, 94, 1055–1084.

SCHOFHEIDE, F. AND D. SONG (2013): "Real Time Forecasting with a Mixed-Frequency VAR," NBER Working Papers 19712.

SCHWARZ, K. (2014): "Mind the Gap: Disentangling Credit and Liquidity in Risk Spreads," .

SIMS, C. A., J. H. STOCK, AND M. W. WATSON (1990): "Inference in Linear Time Series Models with Some Unit Roots," *Econometrica*, 58, 113–44.

STOCK, J., C. SIMS, AND M. WATSON (1990): "Inference in Linear Time Series Models with Some Unit Roots," *Econometrica*, 58, 113–144.

STOCK, J. AND M. WATSON (2012): "Disentangling the Channels of the 2007-2009 Recession," *Brookings Papers on Economic Activity*, Spring, 81–135.

STOCK, J. H. AND M. YOGO (2002): "Testing for Weak Instruments in Linear IV Regression," NBER Technical Working Papers 0284, National Bureau of Economic Research, Inc.

VICONDOA, A. (2016): "Monetary News, U.S. Interest Rate and Business Cycle in Emerging Economies," EUI Working Paper ECO 2016/10.

ZADROZNY, P. (1988): "Analytic Derivatives for Estimation of Discrete-Time," *Econometrica*, 56, 467–472.

ZHAO, B. (2015): "Rational housing bubble," *Economic Theory*, 60, 141–201.

Appendix A

Appendix: News and Noise Bubbles in the Housing Market

A.1 Econometric Framework

I describe briefly the methodology of FGLS in what follows: I first present a simple case, in which the fundamental news shock is anticipated one period ahead, to provide intuitively the mechanism behind the identification and then I describe a more general case. Notice that the actual identification employs (rents at) 40 quarters as the horizon to determine whether a shock to the signal is fundamental or noisy.

If we consider eq.(3) and (4) in a MA representation

$$\begin{pmatrix} \Delta r_t \\ s_t \end{pmatrix} = \begin{pmatrix} L & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} f_t \\ n_t \end{pmatrix} \tag{A.1}$$

it is trivial to see that the associated matrix has determinant 0 for $L = 0$ (comes from the lagged impact of the news shock). Therefore, the MA representation is non-fundamental and non-invertible. In this case, noise and news shock cannot be expressed as a linear combination of present and past reduced form residuals. Thus, a VAR representation in the *structural shocks*, news and noise, does not exist. Intuitively, agents cannot distinguish the two shocks given their information set and the same holds for the econometrician. Adding other variables to the system cannot solve this issue. What the econometrician can recover is the following

fundamental representation:

$$\begin{pmatrix} \Delta r_t \\ s_t \end{pmatrix} = \begin{pmatrix} 1 & L\frac{\sigma_f^2}{\sigma_s^2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_t \\ s_t \end{pmatrix} = \begin{pmatrix} u_t + L\frac{\sigma_f^2}{\sigma_s^2} s_t \\ s_t \end{pmatrix} \tag{A.2}$$

where u_t can be defined as *unanticipated fundamental shock.* The signal extraction problem depends on the relative importance of the news and noise shocks in driving the signal: $\mathbb{E}_{t-1}(\Delta r_t) = \frac{\sigma_f^2}{\sigma_s^2} s_{t-1}$. In other words, u_t is the forecast error of the fundamental:

$$u_t = \Delta r_t - \mathbb{E}_{t-1}(\Delta r_t) = \quad f_{t-1} - \frac{\sigma_f^2}{\sigma_s^2} s_{t-1} = \quad \frac{\sigma_n^2}{\sigma_s^2} f_{t-1} - \frac{\sigma_f^2}{\sigma_s^2} n_{t-1} \tag{A.3}$$

We can express $\begin{pmatrix} u_t \\ s_t \end{pmatrix}$ as combinations of present and past structural shocks $\begin{pmatrix} f_t \\ n_t \end{pmatrix}$:

$$\begin{pmatrix} u_t \\ s_t \end{pmatrix} = \begin{pmatrix} L\frac{\sigma_n^2}{\sigma_s^2} & -L\frac{\sigma_f^2}{\sigma_s^2} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} f_t \\ n_t \end{pmatrix} = \begin{pmatrix} \frac{\sigma_n^2}{\sigma_s^2} f_{t-1} - \frac{\sigma_f^2}{\sigma_s^2} n_{t-1} \\ f_t + n_t \end{pmatrix} \tag{A.4}$$

$\begin{pmatrix} u_t \\ s_t \end{pmatrix}$ can be identified through a standard VAR and, once the news to noise variance ratio is estimated, we can use this information to recover $\begin{pmatrix} f_t \\ n_t \end{pmatrix}$ as follows:

$$\begin{pmatrix} f_t \\ n_t \end{pmatrix} = \begin{pmatrix} L^{-1} & \frac{\sigma_f^2}{\sigma_s^2} \\ -L^{-1} & \frac{\sigma_n^2}{\sigma_s^2} \end{pmatrix} \begin{pmatrix} u_t \\ s_t \end{pmatrix} = \begin{pmatrix} L^{-1}u_t + & \frac{\sigma_f^2}{\sigma_s^2} s_t \\ -L^{-1}u_t + & \frac{\sigma_n^2}{\sigma_s^2} s_t \end{pmatrix} = \begin{pmatrix} u_{t+1} + & \frac{\sigma_f^2}{\sigma_s^2} s_t \\ -u_{t+1} + & \frac{\sigma_n^2}{\sigma_s^2} s_t \end{pmatrix} \tag{A.5}$$

Notice that by inverting L we are employing present and *future* values of the unanticipated fundamental and signal shocks, which, in other words, means we are using future reduced form residuals.[1A]

The news shock can be expressed and thus recovered as the sum of the ex-ante expectation of the fundamental and the realized forecast-error of the fundamental:

$$f_t = u_{t+1} + \frac{\sigma_f^2}{\sigma_s^2} s_t \quad = \quad \Delta r_{t+1} - \frac{\sigma_f^2}{\sigma_s^2} s_t + \frac{\sigma_f^2}{\sigma_s^2} s_t = \Delta r_{t+1} \tag{A.6}$$

The noise shocks is instead the component of the signal that is not reflected in future changes of the fundamental:

$$n_t = -u_{t+1} + \frac{\sigma_n^2}{\sigma_s^2} s_t \quad = - \quad \Delta r_{t+1} + \frac{\sigma_f^2}{\sigma_s^2} s_t + \frac{\sigma_n^2}{\sigma_s^2} s_t = s_t - \Delta r_{t+1} \tag{A.7}$$

Consider a more comprehensive case, using a more general polynomial structure for the bivariate case (it is very easy to extend the scheme to the multivariate case). We define

$$\Delta r_t = c(L) f_t \tag{A.8}$$

and the Blaschke factor

$$b(L) = \prod_{j=1}^{n} \frac{L - k_j}{1 - \overline{k_j} L} \tag{A.9}$$

with k_j $j = 1, 2, ..., n$ are the roots of c(L) smaller than one in modulus with $\overline{k}_j$ the respective complex conjugates. Following **?**, it is not possible to invert $b(L)$ in the past, but it is possible in the future: $b(L)^{-1} = b(L^{-1}) = b(F)$.

$$\begin{pmatrix} \Delta r_t \\ s_t \end{pmatrix} = \begin{pmatrix} a_{11}(L) & a_{12}(L) \\ a_{21}(L) & a_{22}(L) \end{pmatrix} \begin{pmatrix} u_t \\ s_t \end{pmatrix} = \begin{pmatrix} \frac{c(L)}{b(L)} & c(L)\frac{\sigma_f^2}{\sigma_s^2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_t \\ s_t \end{pmatrix} \tag{A.10}$$

[1A] This is quite intuitive: as $\begin{pmatrix} u_t \\ s_t \end{pmatrix}$ are combinations of present and past structural shocks $\begin{pmatrix} f_t \\ n_t \end{pmatrix}$, than $\begin{pmatrix} f_t \\ n_t \end{pmatrix}$ are combinations of present and future structural shocks $\begin{pmatrix} u_t \\ s_t \end{pmatrix}$.

$$\begin{pmatrix} u_t \\ s_t \end{pmatrix} = \begin{pmatrix} b(L)\frac{\sigma_n^2}{\sigma_s^2} & -b(L)\frac{\sigma_f^2}{\sigma_s^2} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} f_t \\ s_t \end{pmatrix} \tag{A.11}$$

We can generalize the system by assuming that, even if the agents' expectations are not perfectly observable, the econometrician has access to a variable informative enough about the signal (z_t):

$$\begin{aligned} \begin{pmatrix} \Delta r_t \\ z_t \end{pmatrix} &= \begin{pmatrix} a_{11}(L) & a_{12}(L) \\ a_{21}(L) & a_{22}(L) \end{pmatrix} \begin{pmatrix} \frac{u_t}{\sigma_u} \\ \frac{s_t}{\sigma_s} \end{pmatrix} = \begin{pmatrix} \frac{c(L)\sigma_u}{b(L)} & \frac{c(L)\sigma_f^2}{\sigma_s} \\ d(L)\sigma_u & f(L)\sigma_s \end{pmatrix} \begin{pmatrix} \frac{u_t}{\sigma_u} \\ \frac{s_t}{\sigma_s} \end{pmatrix} \\ &= \begin{pmatrix} \frac{c(L)\sigma_u}{b(L)} & \frac{c(L)\sigma_f^2}{\sigma_s} \\ d(L)\sigma_u & f(L)\sigma_s \end{pmatrix} \begin{pmatrix} b(L)\frac{\sigma_n}{\sigma_s} & -b(L)\frac{\sigma_f}{\sigma_s} \\ \frac{\sigma_f}{\sigma_s} & \frac{\sigma_n}{\sigma_s} \end{pmatrix} \begin{pmatrix} \frac{f_t}{\sigma_f} \\ \frac{n_t}{\sigma_n} \end{pmatrix} \\ &= \begin{pmatrix} c(L)\sigma_f & 0 \\ d(L)b(L)\frac{\sigma_f\sigma_n^2}{\sigma_s^2} + f(L)\sigma_f & -b(L)d(L)\frac{\sigma_f^2\sigma_n}{\sigma_s^2} + f(L)\sigma_n \end{pmatrix} \begin{pmatrix} \frac{f_t}{\sigma_f} \\ \frac{n_t}{\sigma_n} \end{pmatrix} \end{aligned}$$

The steps above exploit the relationship $\sigma_u = \frac{\sigma_f\sigma_n}{\sigma_s}$. The identification strategy comprises of the following steps:

Step 1: Estimate a standard VAR for $\begin{pmatrix} \Delta r_t \\ z_t \end{pmatrix}$ and obtain the corresponding MA representation

Step 2: $a_{11}(0) = \frac{c(0)\sigma_u}{b(0)} = 0 \Rightarrow c(0) = 0$. This restriction implies that the signal does not affect the fundamental measure contemporaneously. Unanticipated fundamental and signal shocks are identified at this point for the bivariate case.

Step 3: Given the estimate $\hat{a}_{12}(L) = \widehat{\frac{c(L)\sigma_f^2}{\sigma_s}}$ take the roots of $\hat{a}_{12}(L)$ smaller than one in modulus in order to estimate $b(L)$ as shown in (14)

Step 4: $\hat{a}_{11}(1)$ is estimated as $\widehat{\frac{c(1)\sigma_u}{b(1)}}$. Notice that since $b(1) = 1$ and $\sigma_u = \frac{\sigma_f\sigma_n}{\sigma_s}$, the following condition holds for the ratio of variances of news and noise shocks: $\frac{a_{12}(1)}{a_{11}(1)} = \frac{\sigma_f}{\sigma_n}$ estimated as $\frac{\hat{a}_{12}(1)}{\hat{a}_{11}(1)} = \widehat{\frac{\sigma_f}{\sigma_n}}$. [2A]

[2A] In practice, the ratio $\widehat{\frac{\sigma_f}{\sigma_n}}$ is computed as the ratio of the cumulated long-run responses $\frac{CIRF(\Delta r_t to\, s_t)}{CIRF(\Delta r_t to\, u_t)}$. Notice that the theoretical restriction of a null effect of the noise shock

Step 5: Since $\frac{\sigma_f^2}{\sigma_s^2} + \frac{\sigma_n^2}{\sigma_s^2} = 1$, $\widehat{\sigma_f} = sin(arctan(\frac{\widehat{\sigma_f}}{\sigma_n}))$ and $\widehat{\sigma_n} = cos(arctan(\frac{\widehat{\sigma_f}}{\sigma_n}))$ can be directly computed. At this point the variance of the news and noise shock is identified.

Step 6: Finally, using $\begin{pmatrix} f_t \\ n_t \end{pmatrix} = \begin{pmatrix} b(F)\frac{\sigma_n}{\sigma_s} & \frac{\sigma_f^2}{\sigma_s^2} \\ -b(F) & \frac{\sigma_n^2}{\sigma_s^2} \end{pmatrix} \begin{pmatrix} u_t \\ s_t \end{pmatrix}$ one can recover the structural shocks.

on the fundamental should hold at every horizon. In practice, this is imposed on impact and in the long-run (40 quarters), but it is used for testing at the other horizons (noise has no significant effect on the fundamental at each horizon).

A.2 Empirical Results

A.2.1 Risk Free Rate and Risk Premia Shock

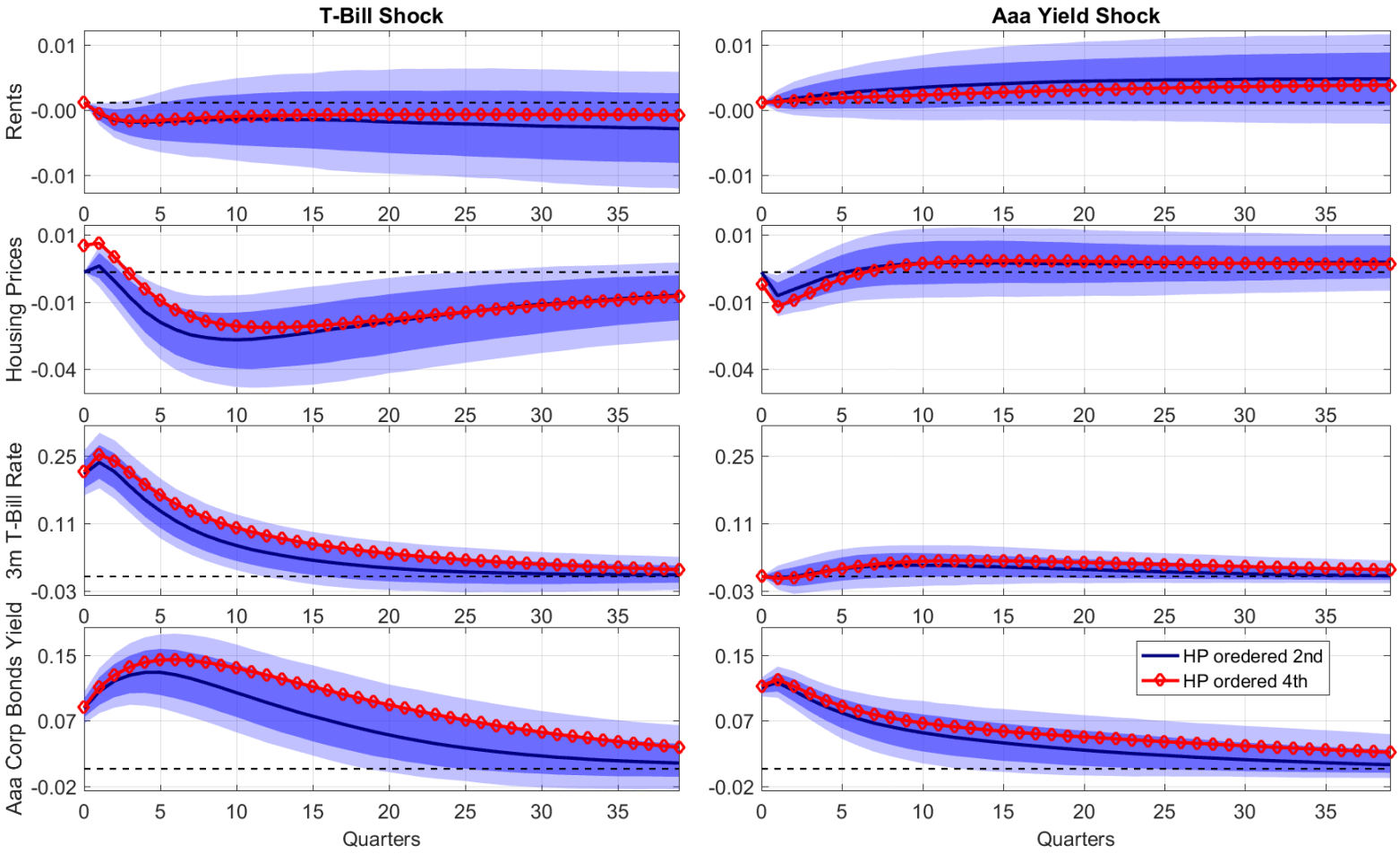

Figure A.1: IRFs to risk free rate and risk premium shock - decomposition

IRFs to a risk free rate shock and to risk premium shock. The responses are reported in terms of the standard deviation of the variables in the system. The solid blue line is the median, the dark and light blue shaded areas represents 68% and 90% confidence bands respectively (2000 bootstrap replications). The shocks are identified through the following ordering: [Rents, Housing Prices, 3 Months Bill Rate, Aaa Moody's Corporate Bond Yield]. The red line reports the median IRFs obtained by the recursive ordering [Rents, 3 Months Bill Rate, Aaa Moody's Corporate Bond Yield, Housing Prices]. Sample: 1963:Q1 - 2016:Q3.

A.2.2 Granger Tests

Variable	Source	Code
Real GDP	FRED	GDPC1
GDP Implicit Price Deflator	FRED	GDPDEF
Private Residential Fixed Investment	FRED	PRFI
Average Sales Price of Houses Sold	FRED	APSUS
Median Price of New Houses Sold	FRED	MSPNHSUS
Consumer Price Index for All Urban Consumers: Rent of primary residence	FRED	CUUR0000SEHA
Personal Consumption Expenditures: Rents Quantity Index	FRED	DTENRA3
Personal Consumption Expenditures: Rents Price Index	FRED	DHUTRC1Q027SBEA
Shiller Housing Price Index	R. J. Shiller	
New Private Housing Units Started	FRED	HSTARTS
New Housing Permits	FRED	HPERMITS
Standard and Poor Composite Index	R. J. Shiller	
Effective Fed Fund Rate	FRED	FFR
Moody's Seasoned Aaa Corporate Bond Yield	FRED	AAA_yield
Civilian Non-institutional Population	FRED	CNP16OV_NBD19480101
US Treasury Yield Adjusted to Constant Maturity 10 years	FRED	GS10
US Treasury Yield Adjusted to Constant Maturity 20 years	FRED	GS20
3 Month Treasury Bill Rate	FRED	TB3MS
Buying Conditions for Housing	Michigan Survey of Consumers	Table 41
Personal Consumption Expenditures: Price Deflator	FRED	DPCERD3Q086SBEA
Households and Nonprofit Organizations: Home Mortgages	FRED	HHMSDODNS
Mortgage Debt Outstanding, All holders	FRED	MDOAH
Real Estate Loans, All Commercial Banks	FRED	REALLN
Price of New One-Family Houses Sold Including Value of Lot	FRED	USHSN1FLF?
Civilian Unemployment Rate	FRED	UNRATE

Table A.1: Variable employed in orthogonality test
Variables employed for the Granger Test of information sufficiency

Shock	Lags	Principal Components					
		(1)	(2)	(3)	(4)	(5)	(6)
Surprise	2	0.17	0.34	0.36	0.12	0.00	0.00
	4	0.30	0.22	0.28	0.11	0.02	0.02
Signal	2	0.61	0.72	0.33	0.04	0.10	0.05
	4	0.36	0.24	0.03	0.03	0.08	0.06
News	2	0.47	0.25	0.09	0.02	0.02	0.01
	4	0.48	0.52	0.02	0.05	0.04	0.03
Noise	2	0.18	0.13	0.19	0.11	0.20	0.06
	4	0.02	0.00	0.01	0.01	0.03	0.00

Table A.2: Orthogonality Test - Decomposition

Results of the test for informational sufficiency in the four variables VAR including [Rents, Housing Prices, 3 Month T-Bill Rate, Aaa Moody's Corporate Yield]. The identified shocks are regressed on the lagged (two and four lags) principal components of the variables listed in Table A.1. The table reports the p-values of the F-test in the regression.

Shock	Lags	Principal Components					
		(1)	(2)	(3)	(4)	(5)	(6)
Surprise	2	0.38	0.68	0.80	0.49	0.49	0.61
	4	0.64	0.72	0.87	0.36	0.18	0.33
Signal	2	0.86	0.96	0.64	0.76	0.84	0.71
	4	0.92	0.93	0.93	0.98	1.00	0.91
News	2	0.57	0.84	0.70	0.82	0.86	0.80
	4	0.92	0.84	0.47	0.65	0.64	0.60
Noise	2	0.81	0.86	0.82	0.91	0.96	0.96
	4	0.87	0.90	0.90	0.94	0.92	0.88

Table A.3: Orthogonality Test - Macro Analysis

Results of the test for informational sufficiency in the six variables VAR including [Rents, Housing Prices, GDP, Residential Investment, Aaa Moody's Corporate Yield, S&P Composite Index]. The identified shocks are regressed on the lagged (two and four lags) principal components of the variables listed in TableA.1. The table reports the p-values of the F-test in the regression.

A.2.3 Expectations from Surveys

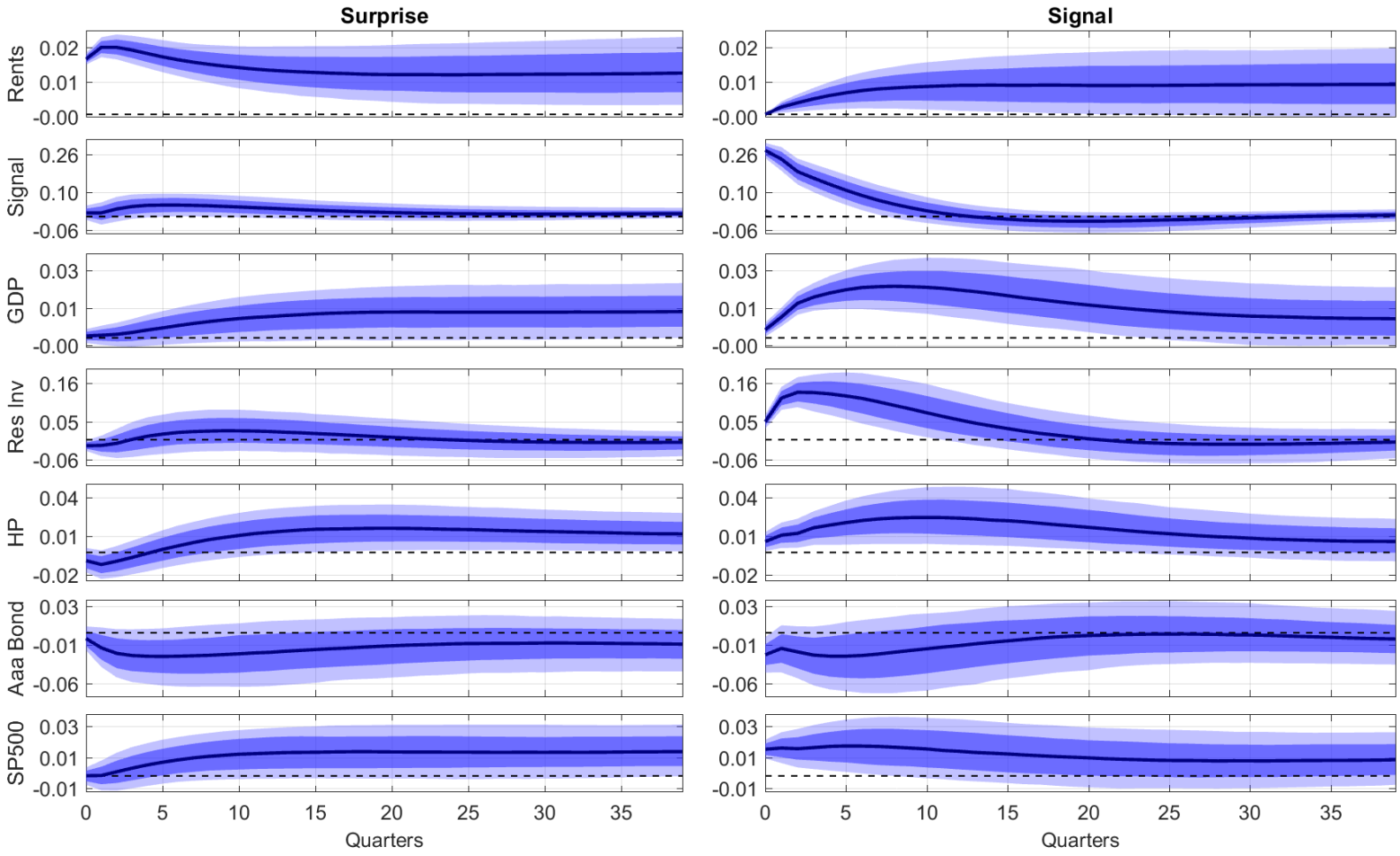

Figure A.2: IRFs Surprise and Signal Shocks - Expectations from surveys

IRFs to a surprise shock to rents and to a signal shock. The responses are reported in terms of the standard deviation of the variables in the system. The solid blue line is the median, the dark and light blue shaded areas represents 68% and 90% confidence bands respectively (2000 bootstrap replications). The shocks are identified through the following ordering: [Rents, Buying Conditions for Housing, GDP, Residential Investment, Housing Prices, Aaa Moody's Corporate Bond Yield, S&P Index]. Sample: 1963:Q1 - 2016:Q3.

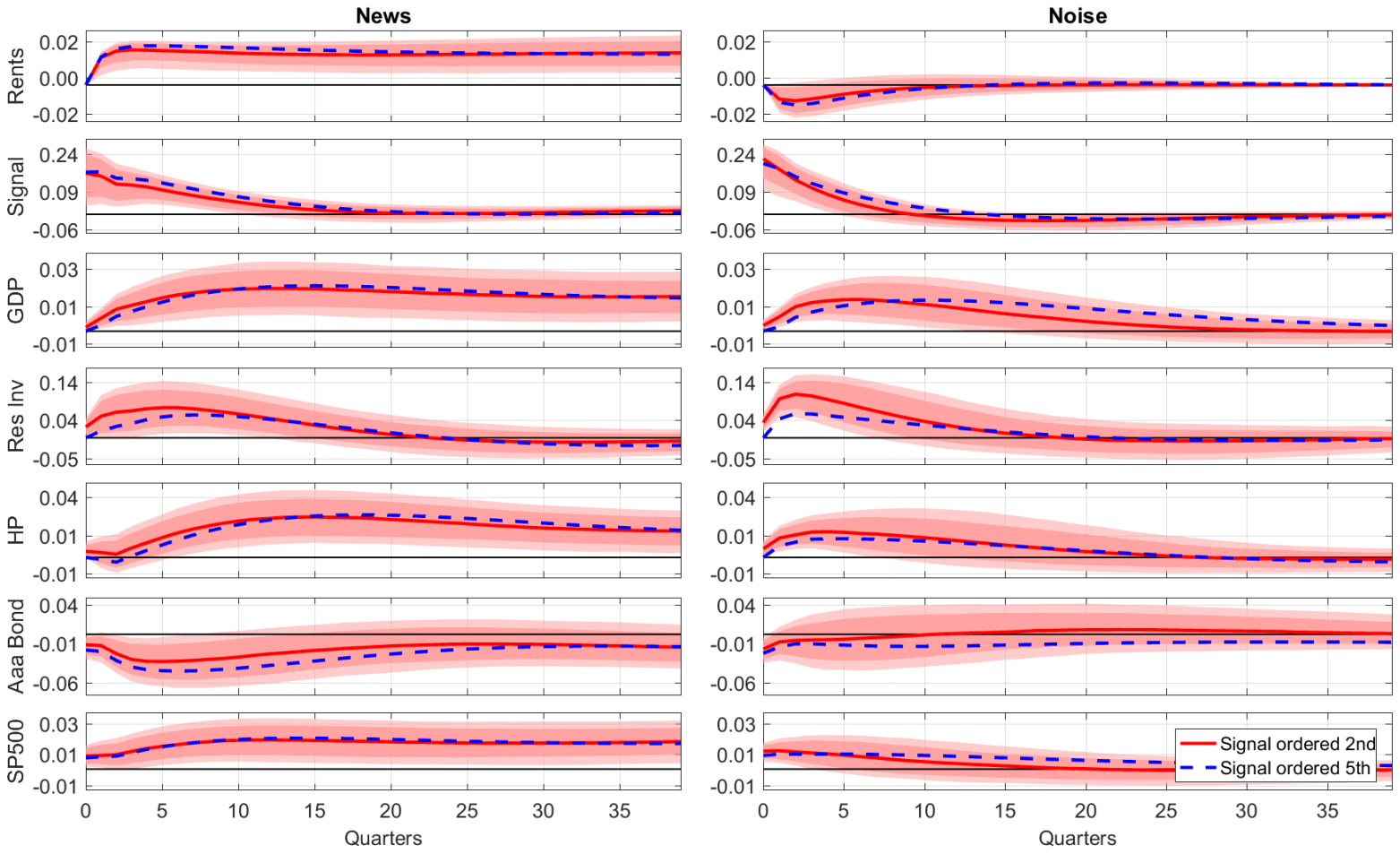

Figure A.3: IRFs News and Noise Shocks - Expectations from surveys

IRFs to news and noise shocks. The responses are reported in terms of the standard deviation of the variables in the system. The solid red line is the median, the dark and light red shaded areas represents 68% and 90% confidence bands respectively (2000 bootstrap replications). The shocks are identified through the following ordering: [Rents, Buying Conditions for Housing, GDP, Residential Investment, Housing Prices, Aaa Moody's Corporate Bond Yield, S&P Index]. The blue dotted line reports the median IRFs obtained by the recursive ordering [Rents, GDP, Residential Investment, Aaa Moody's Corporate Bond Yield, Housing Prices, Buying Conditions for Housing, S&P Index]. Sample: 1963:Q1 - 2016:Q3.

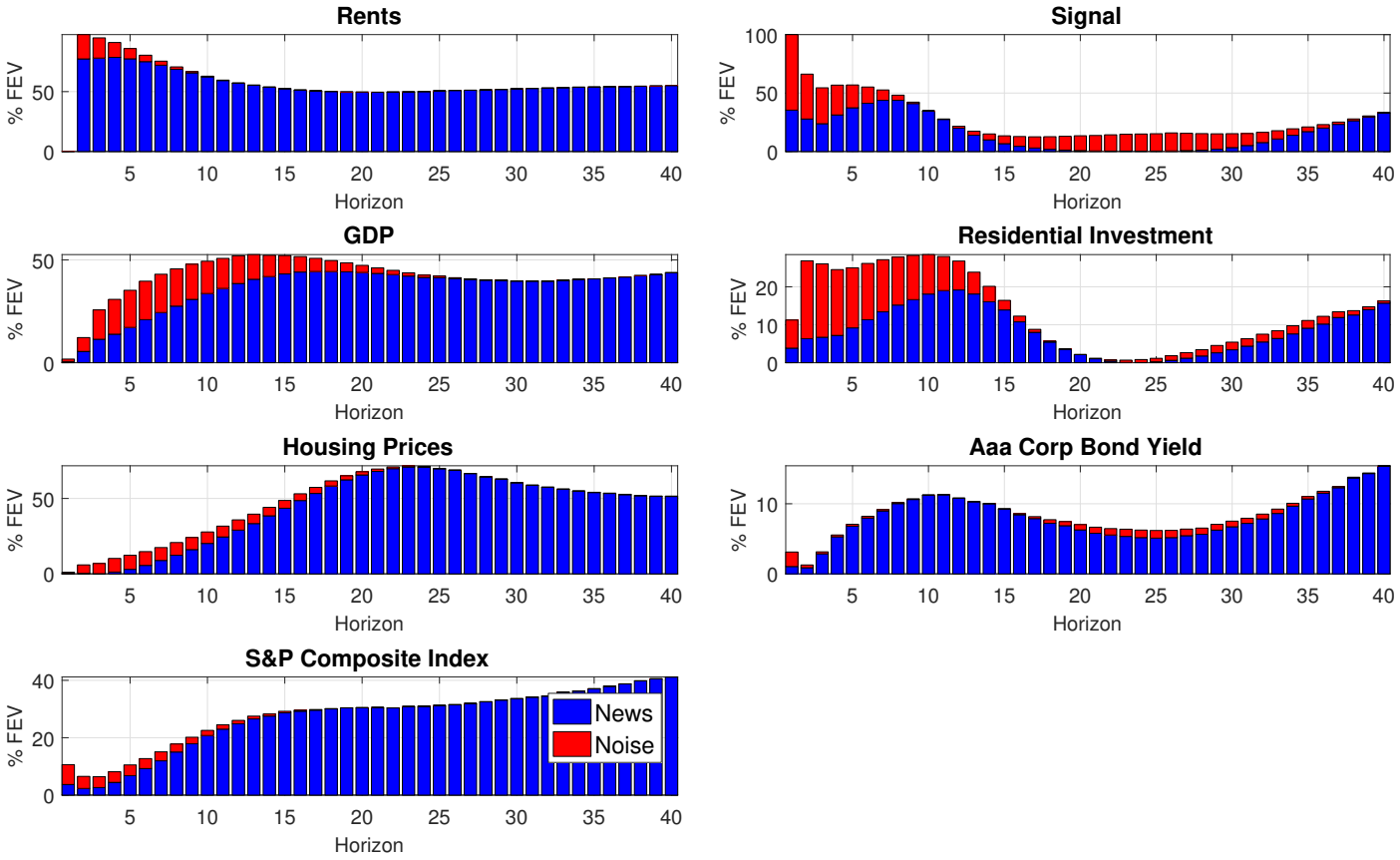

Figure A.4: Forecast Error Variance Decomposition - Expectations from surveys

Forecast error variance decomposition of the variables in the system. The plot display the share of the variance explained by news and noise at each horizon (not cumulatively). The shocks are identified through the following ordering: [Rents, Buying Conditions for Housing, GDP, Residential Investment, Housing Prices, Aaa Moody's Corporate Bond Yield, S&P Index]. Sample: 1963:Q1 - 2016Q3.

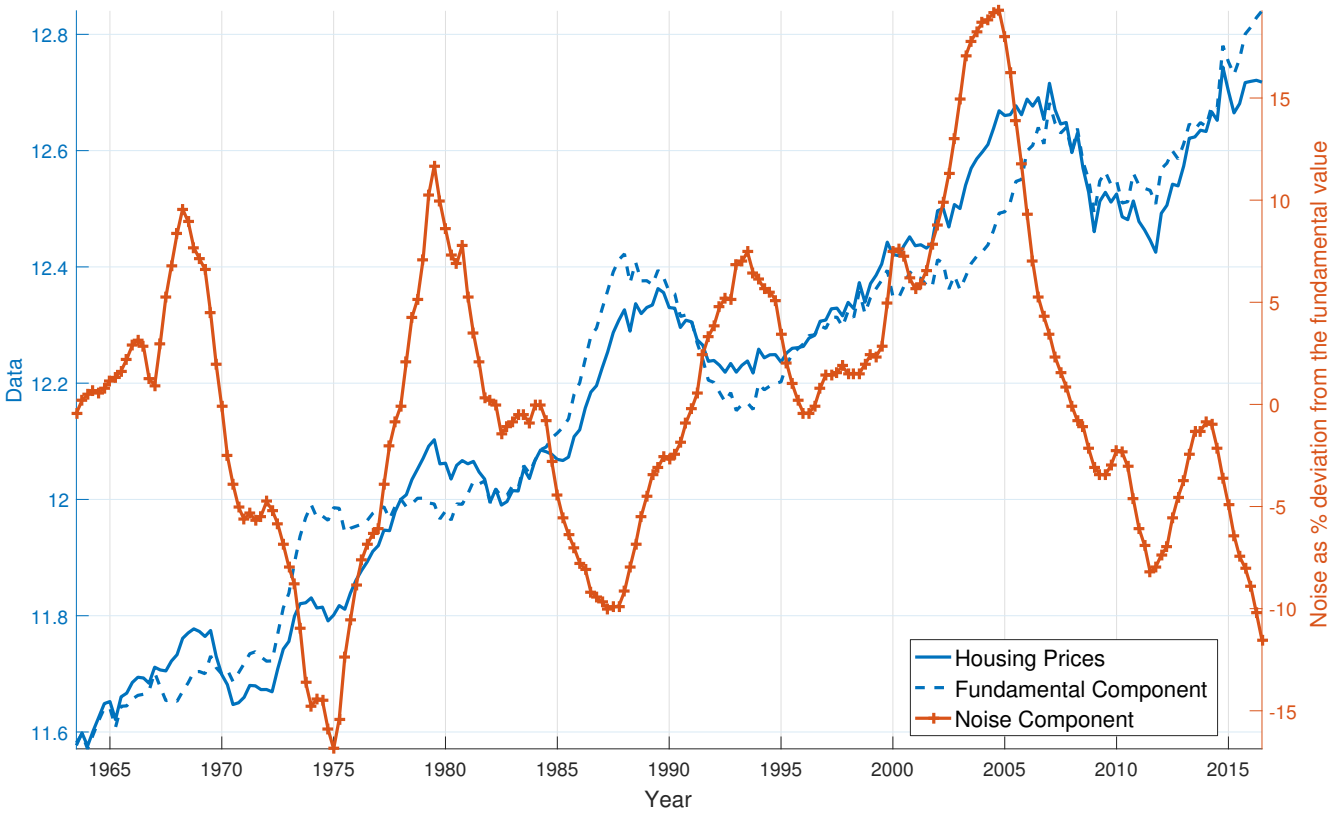

Figure A.5: Historical Decomposition - Expectations from surveys

Historical decomposition of housing prices (dotted blue) into a fundamental component (blue) and noisy component (orange). The shocks are identified through the following ordering: [Rents, Buying Conditions for Housing, GDP, Residential Investment, Housing Prices, Aaa Moody's Corporate Bond Yield, S&P Index]. Sample: 1963:Q1 - 2016Q3.

A.2.4 Expectations from the Stock Market

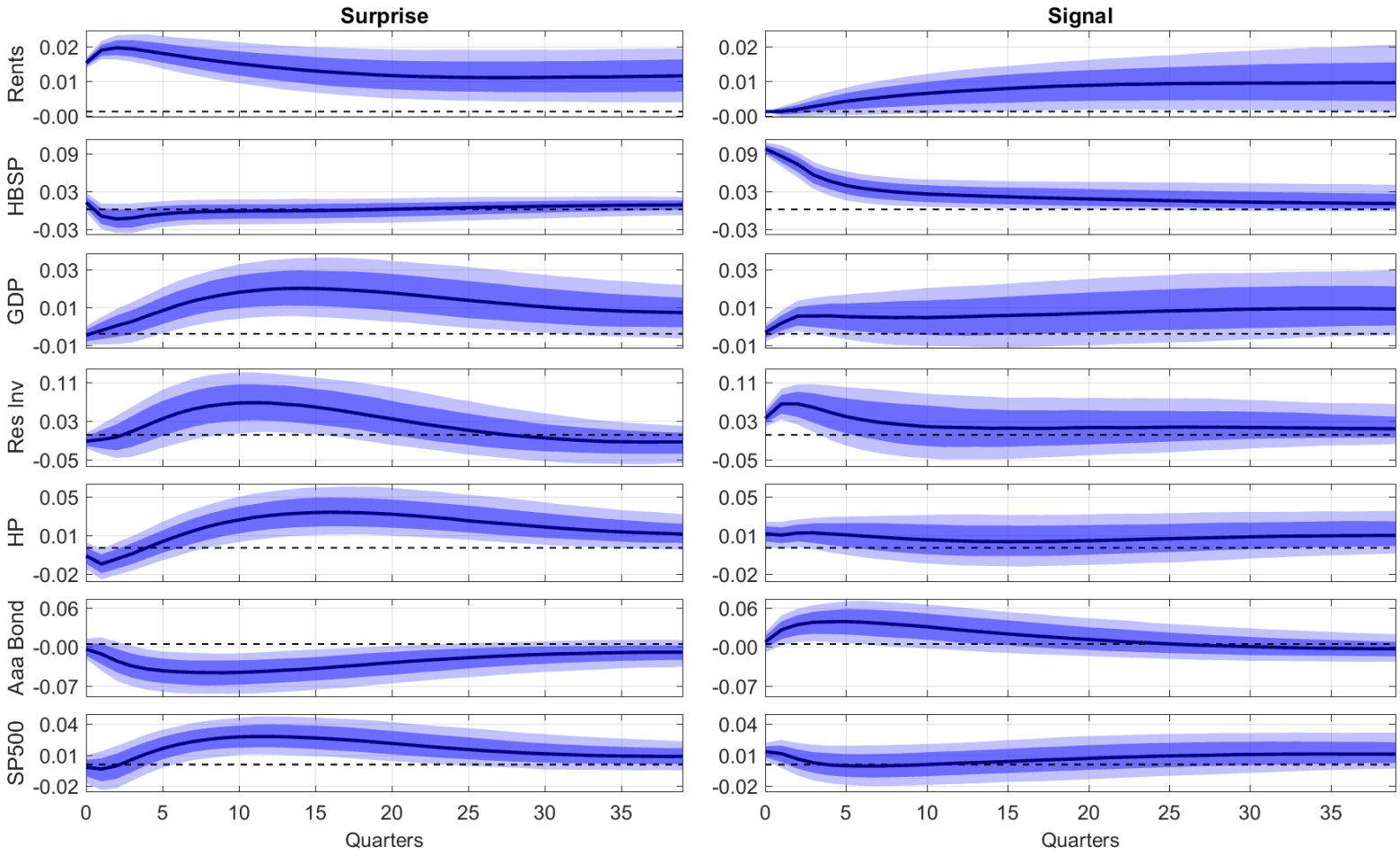

Figure A.6: IRFs to Surprise and Signal Shocks - Expectations from stock prices

IRFs to a surprise shock to rents and to a signal shock. The responses are reported in terms of the standard deviation of the variables in the system. The solid blue line is the median, the dark and light blue shaded areas represents 68% and 90% confidence bands respectively (2000 bootstrap replications). The shocks are identified through the following ordering: [Rents, Home Builders Stock Price Index, GDP, Residential Investment, Housing Prices, Aaa Moody's Corporate Bond Yield, S&P Index]. Sample: 1973:Q1 - 2016:Q3.

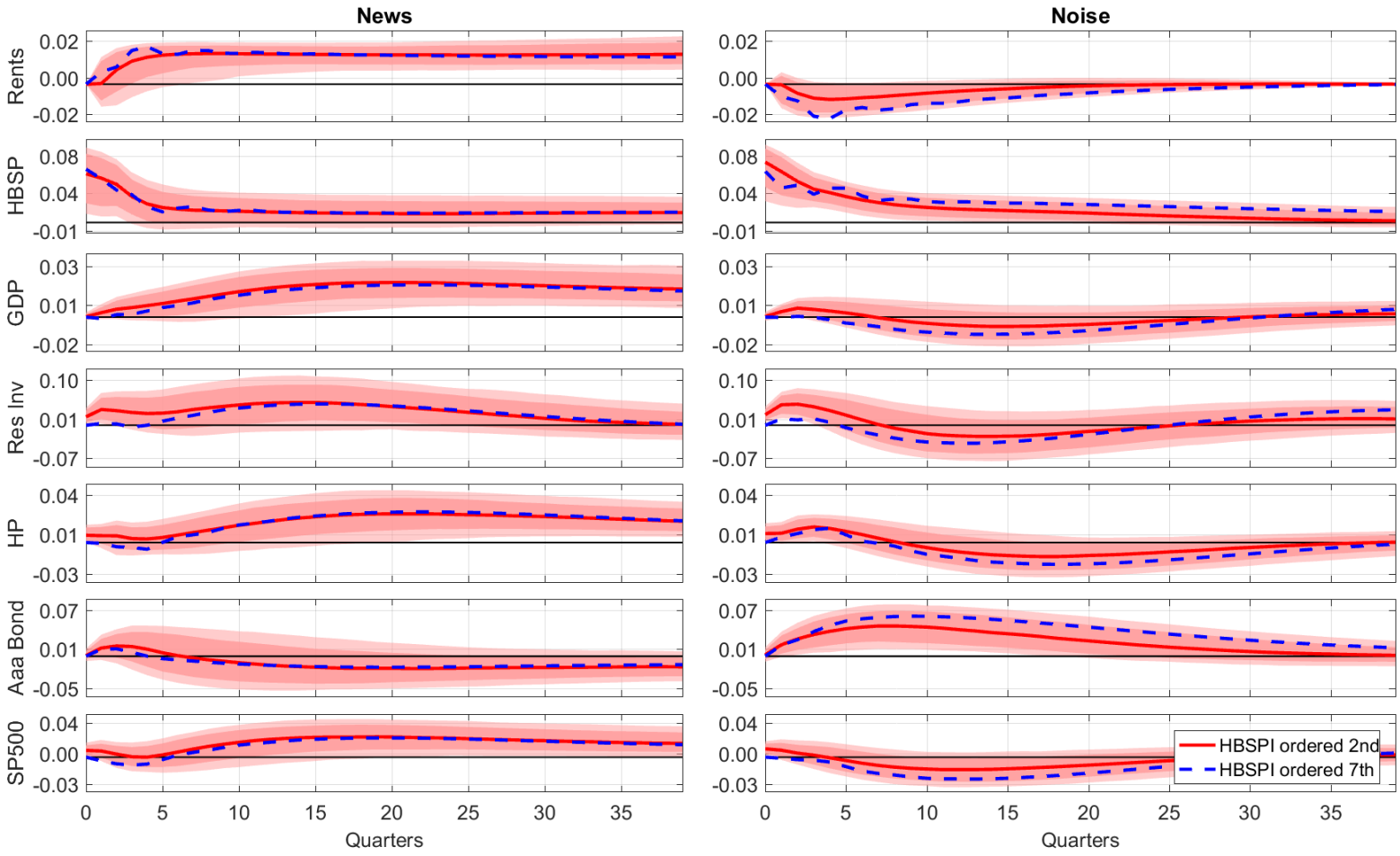

Figure A.7: IRFs to News and Noise Shocks - Expectations from stock prices

IRFs to news and noise shocks. The responses are reported in terms of the standard deviation of the variables in the system. The solid red line is the median, the dark and light red shaded areas represents 68% and 90% confidence bands respectively (2000 bootstrap replications). The shocks are identified through the following ordering: [Rents, Home Builders Stock Price Index, GDP, Residential Investment, Housing Prices, Aaa Moody's Corporate Bond Yield, S&P Index]. The blue dotted line reports the median IRFs obtained by the recursive ordering [Rents, GDP, Residential Investment, Aaa Moody's Corporate Bond Yield, Housing Prices, Home Builders Stock Price Index, S&P Index]. Sample: 1973:Q1 - 2016:Q3.

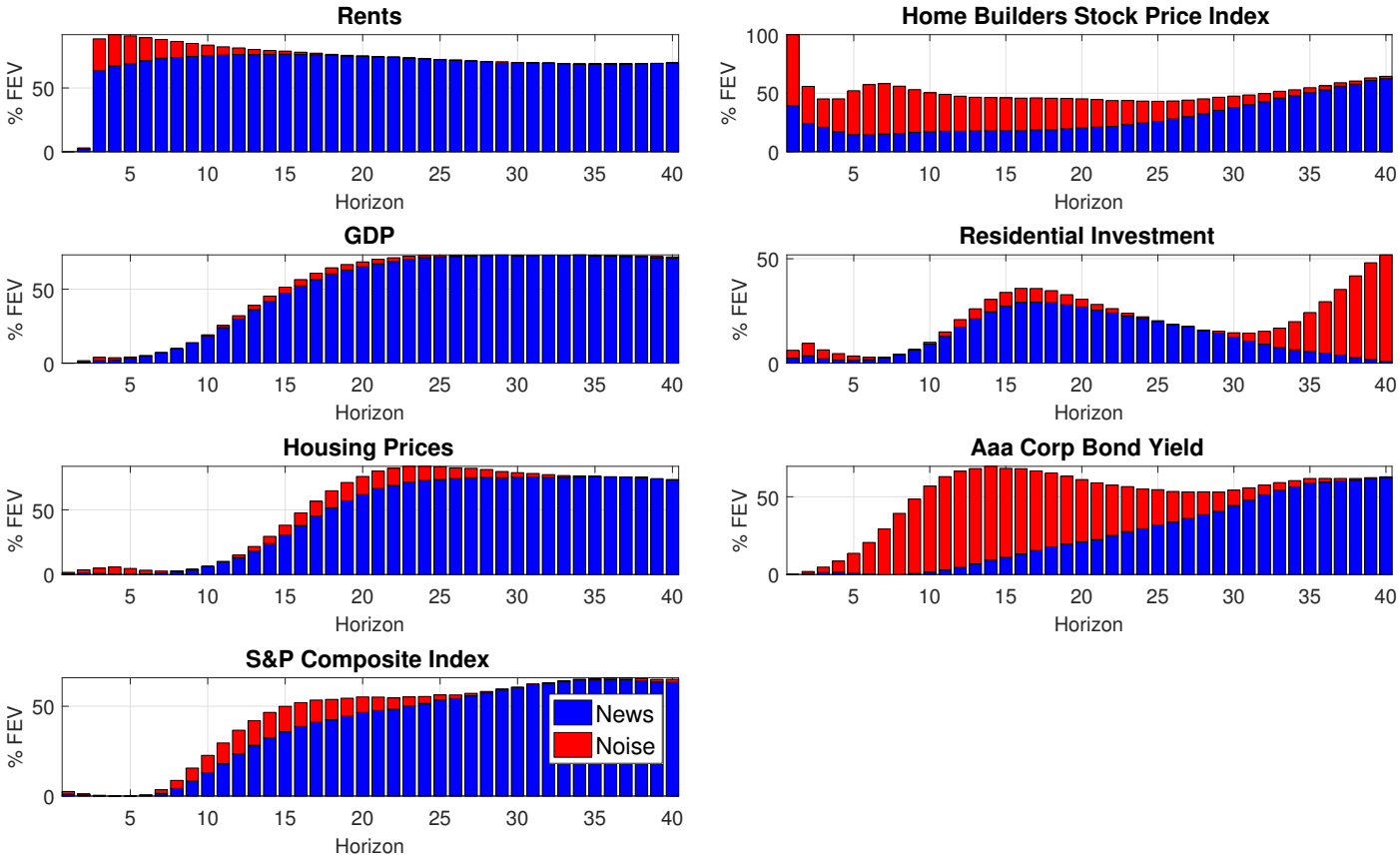

Figure A.8: Forecast Error Variance Decomposition - Expectations from stock prices

Forecast error variance decomposition of the variables in the system. The plot display the share of the variance explained by news and noise at each horizon (not cumulatively). The shocks are identified through the following ordering: [Rents, Home Builders Stock Price Index, GDP, Residential Investment, Housing Prices, Aaa Moody's Corporate Bond Yield, S&P Index]. Sample: 1973:Q1 - 2016Q3.

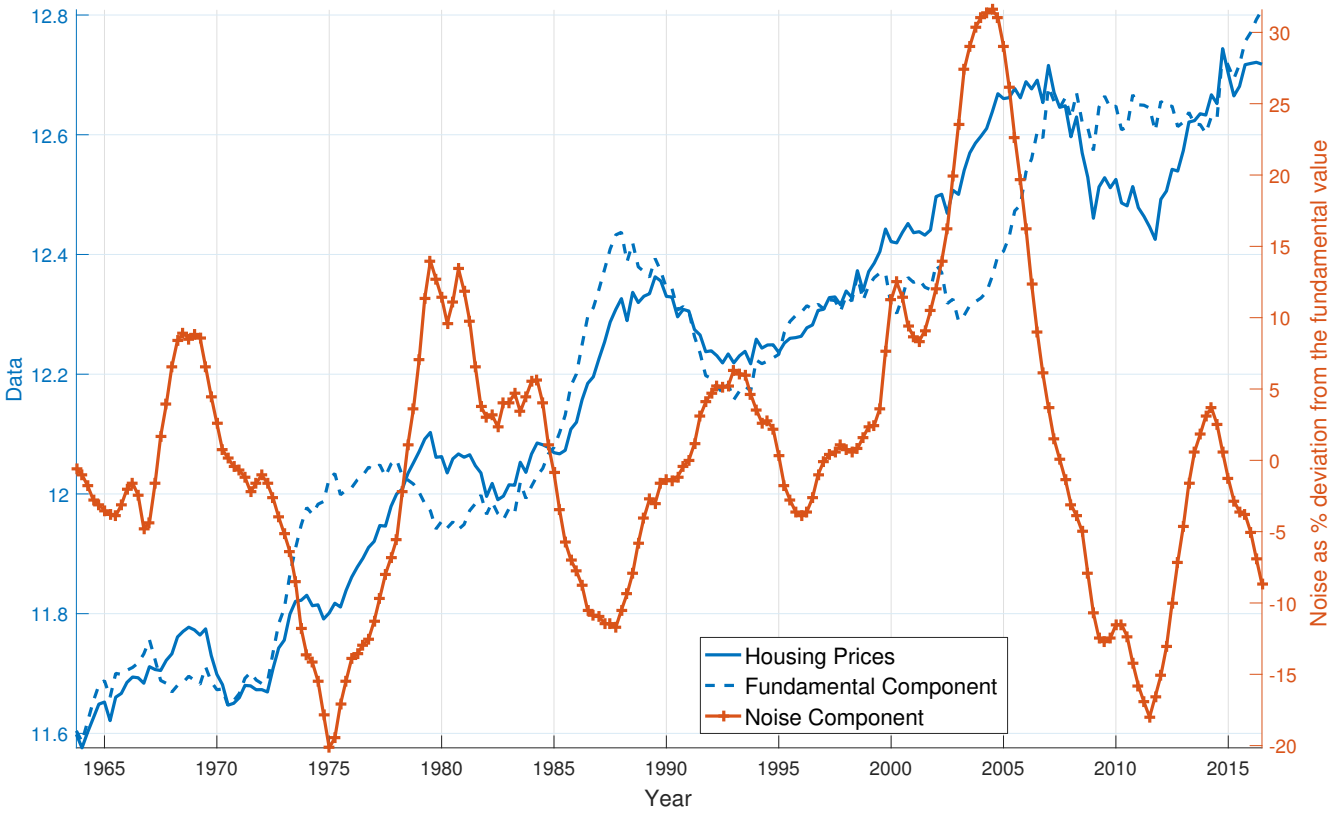

Figure A.9: Historical Decomposition - Expectations from stock prices

Historical decomposition of housing prices (dotted blue) into a fundamental component (blue) and noisy component (orange). The shocks are identified through the following ordering: [Rents, Home Builders Stock Price Index, GDP, Residential Investment, Housing Prices, Aaa Moody's Corporate Bond Yield, S&P Index]. Sample: 1963:Q1 - 2016Q3.

A.3 Robustness Exercises

A.3.1 Long Term Rates

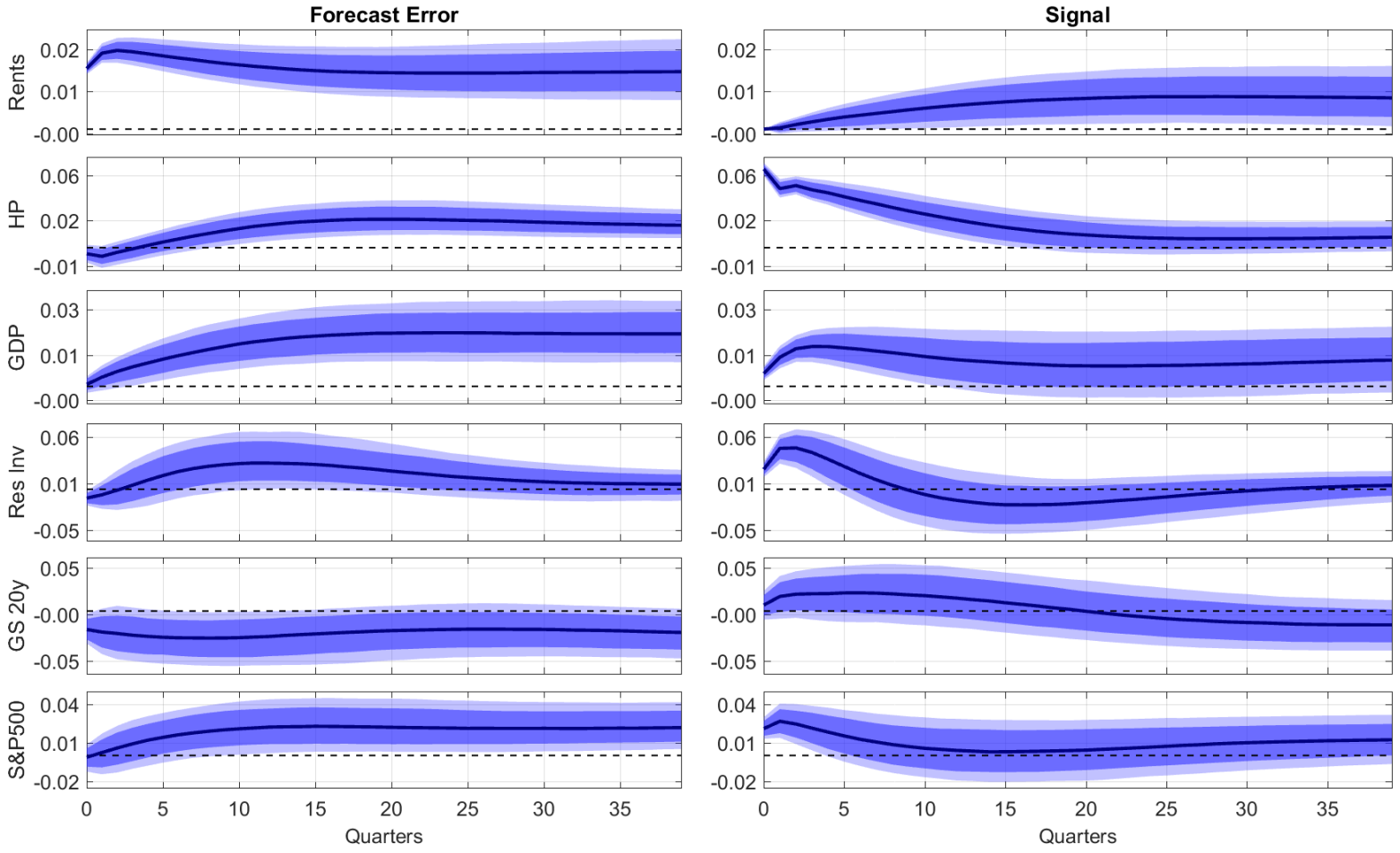

Figure A.10: IRFs to Surprise and Signal Shocks - Long term rates

IRFs to a surprise shock to rents and to a signal shock. The responses are reported in terms of the standard deviation of the variables in the system. The solid blue line is the median, the dark and light blue shaded areas represents 68% and 90% confidence bands respectively (2000 bootstrap replications). The shocks are identified through the following ordering: [Rents, Housing Prices, GDP, Residential Investment, Treasury Yield at Constant 20 year Maturity, S&P Index]. Sample: 1963:Q1 - 2016:Q3.

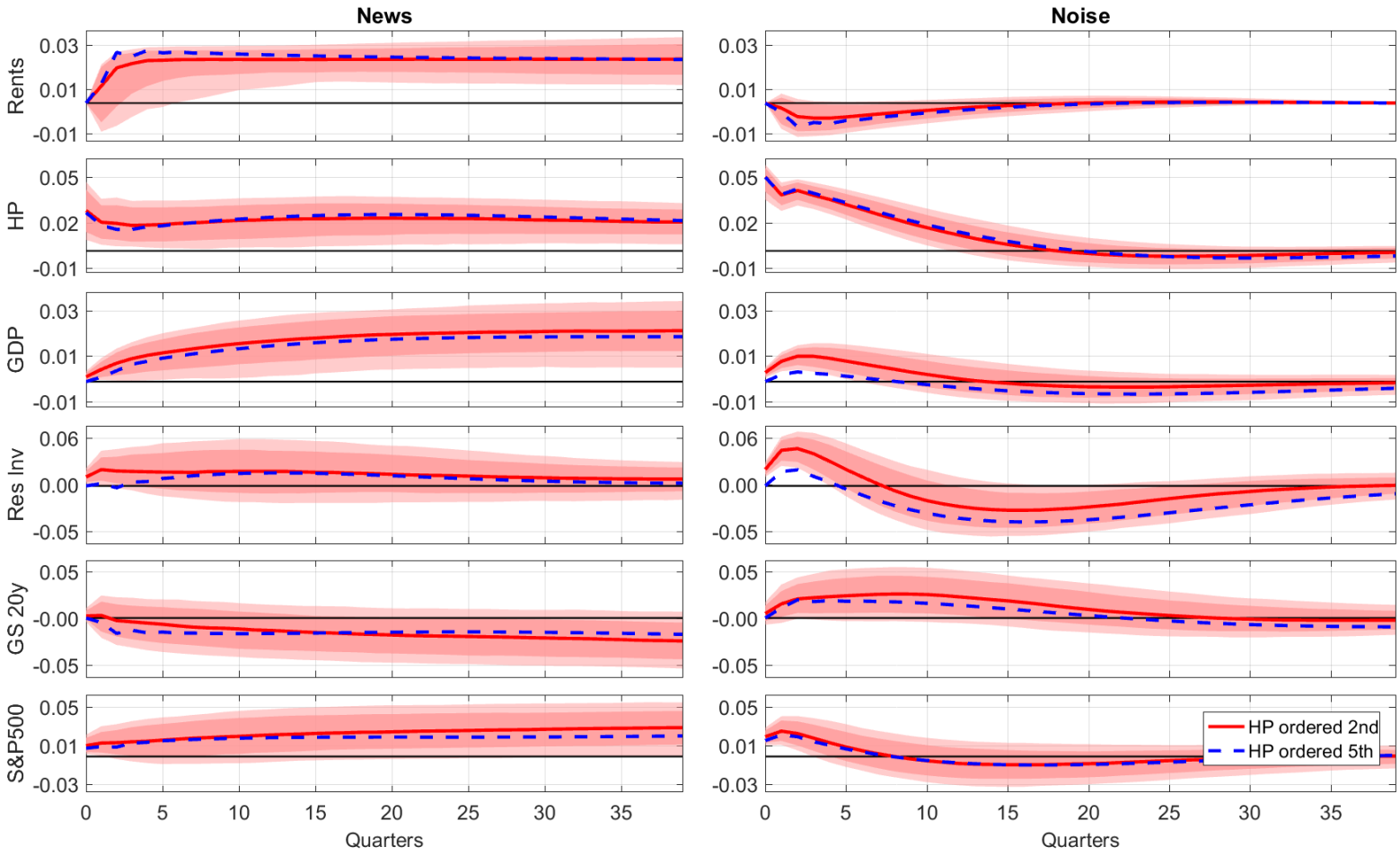

Figure A.11: IRFs to News and Noise Shocks - Expectations from stock prices

IRFs to news and noise shocks. The responses are reported in terms of the standard deviation of the variables in the system. The solid red line is the median, the dark and light red shaded areas represents 68% and 90% confidence bands respectively (2000 bootstrap replications). The shocks are identified through the following ordering: [Rents, Housing Prices, GDP, Residential Investment, Treasury Yield at Constant 20 year Maturity, S&P Index]. The blue dotted line reports the median IRFs obtained by the recursive ordering [Rents, GDP, Residential Investment, Treasury Yield at Constant 20 year Maturity, Housing Prices, S&P Index]. Sample: 1963:Q1 - 2016:Q3.

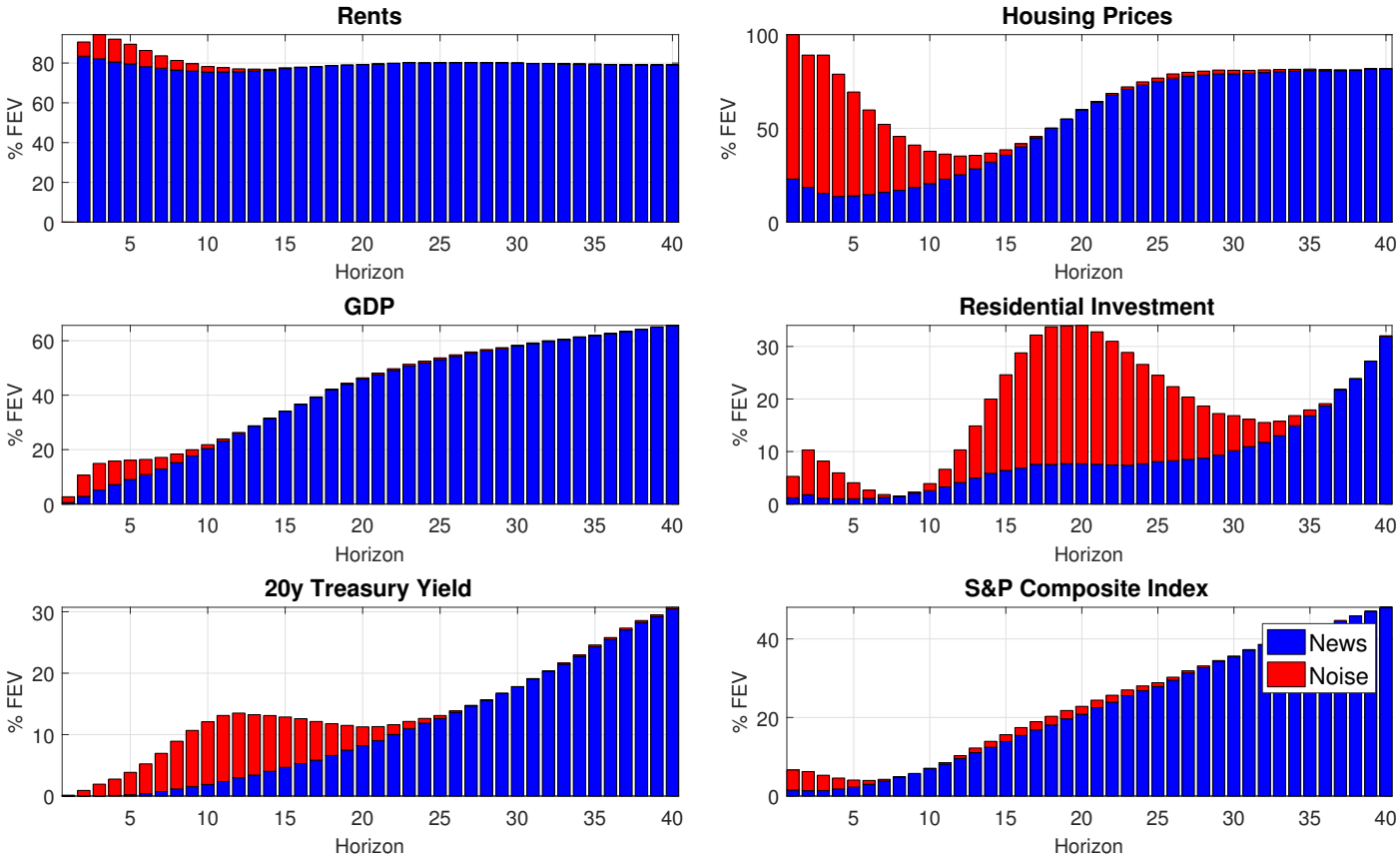

Figure A.12: Forecast Error Variance Decomposition - Long term rates

Forecast error variance decomposition of the variables in the system. The plot display the share of the variance explained by news and noise at each horizon (not cumulatively). The shocks are identified through the following ordering: [Rents, Housing Prices, GDP, Residential Investment, Treasury Yield at Constant 20 year Maturity, S&P Index]. Sample: 1963:Q1 - 2016Q3.

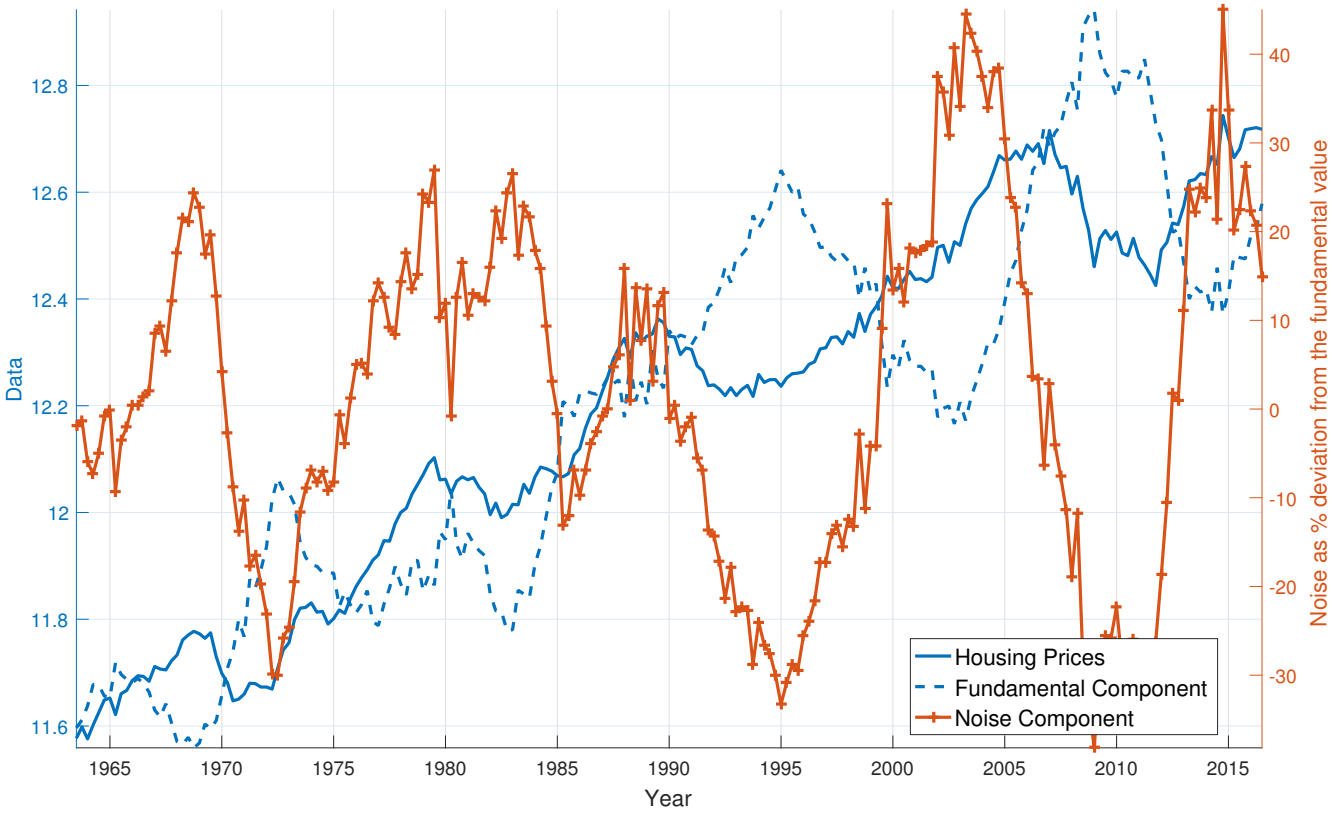

Figure A.13: Historical Decomposition - Long term rates

Historical decomposition of housing prices (dotted blue) into a fundamental component (blue) and noisy component (orange). The shocks are identified through the following ordering: [Rents, Housing Prices, GDP, Residential Investment, Treasury Yield at Constant 20 year Maturity, S&P Index]. Sample: 1963:Q1 - 2016Q3.

A.3.2 Case & Shiller Corelogic Home Price Index

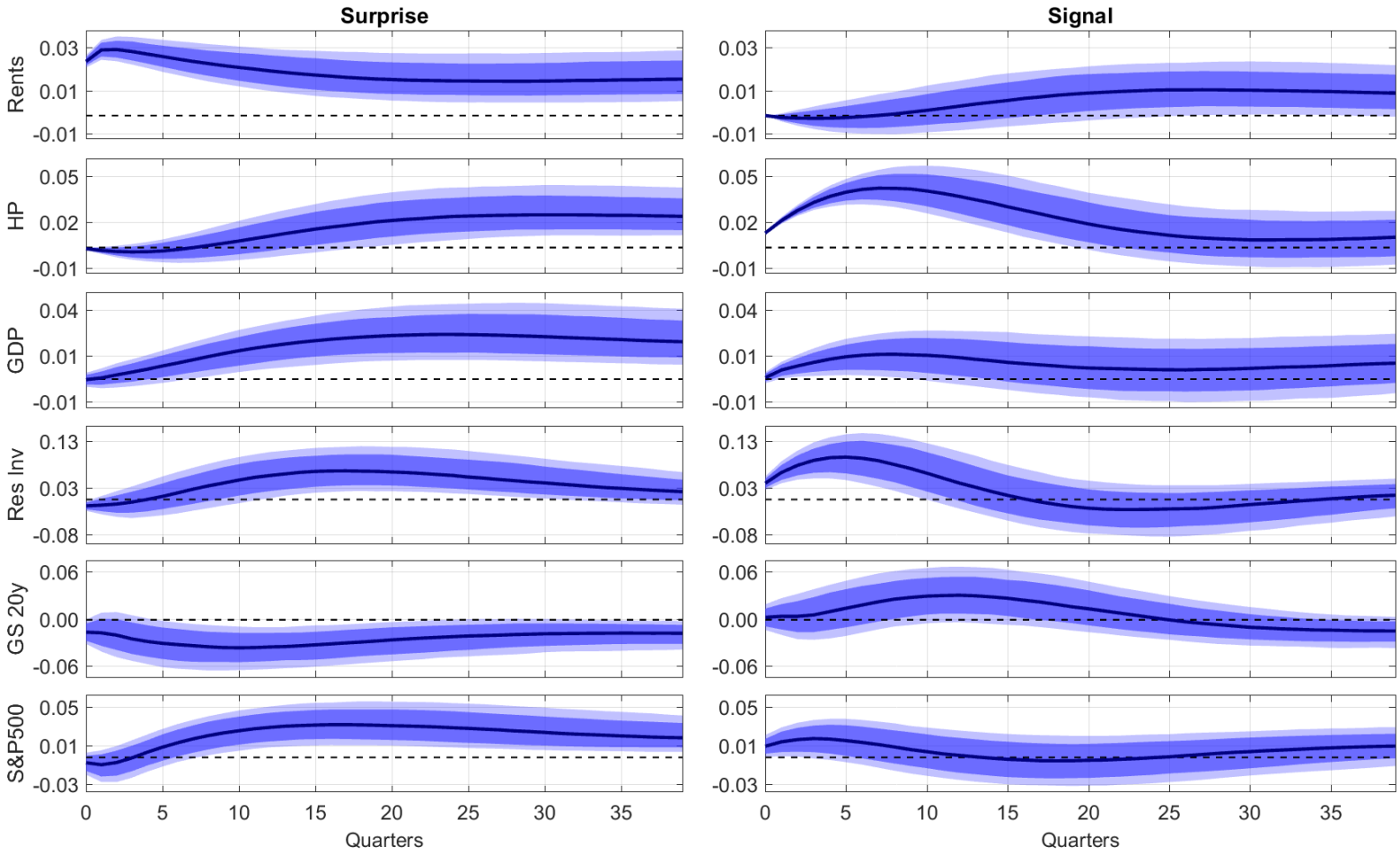

Figure A.14: IRFs to Surprise and Signal Shocks - C&S Corelogic

IRFs to a surprise shock to rents and to a signal shock. The responses are reported in terms of the standard deviation of the variables in the system. The solid blue line is the median, the dark and light blue shaded areas represents 68% and 90% confidence bands respectively (2000 bootstrap replications). The shocks are identified through the following ordering: [Rents, Housing Prices, GDP, Residential Investment, Treasury Yield at Constant 20 year Maturity, S&P Index]. Sample: 1963:Q1 - 2016:Q3.

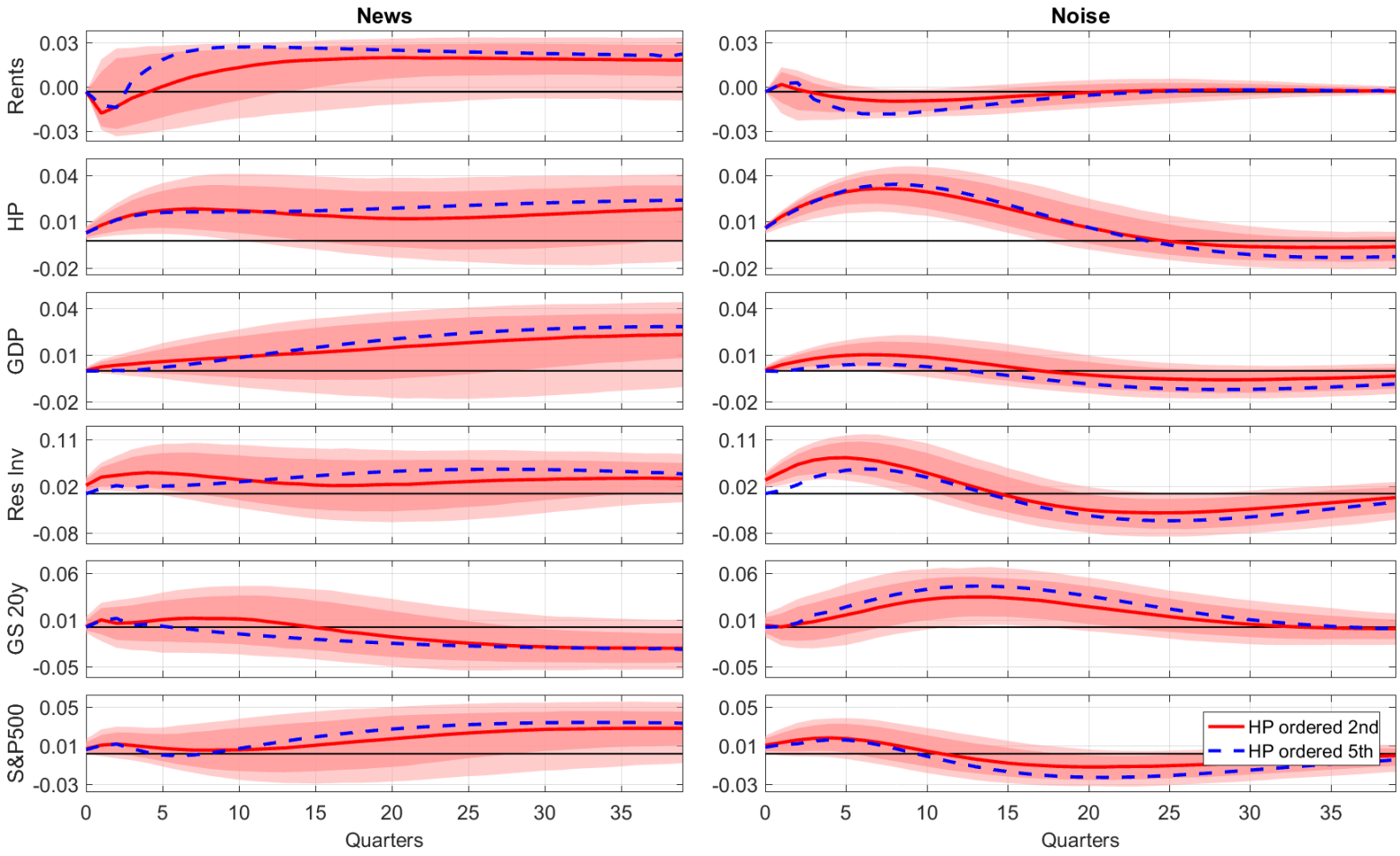

Figure A.15: IRFs to Surprise and Signal Shocks - C&S Corelogic

IRFs to news and noise shocks. The responses are reported in terms of the standard deviation of the variables in the system. The solid red line is the median, the dark and light red shaded areas represents 68% and 90% confidence bands respectively (2000 bootstrap replications). The shocks are identified through the following ordering: [Rents, Housing Prices, GDP, Residential Investment, Treasury Yield at Constant 20 year Maturity, S&P Index]. The blue dotted line reports the median IRFs obtained by the recursive ordering [Rents, GDP, Residential Investment, Treasury Yield at Constant 20 year Maturity, Housing Prices, S&P Index]. Sample: 1963:Q1 - 2016:Q3.

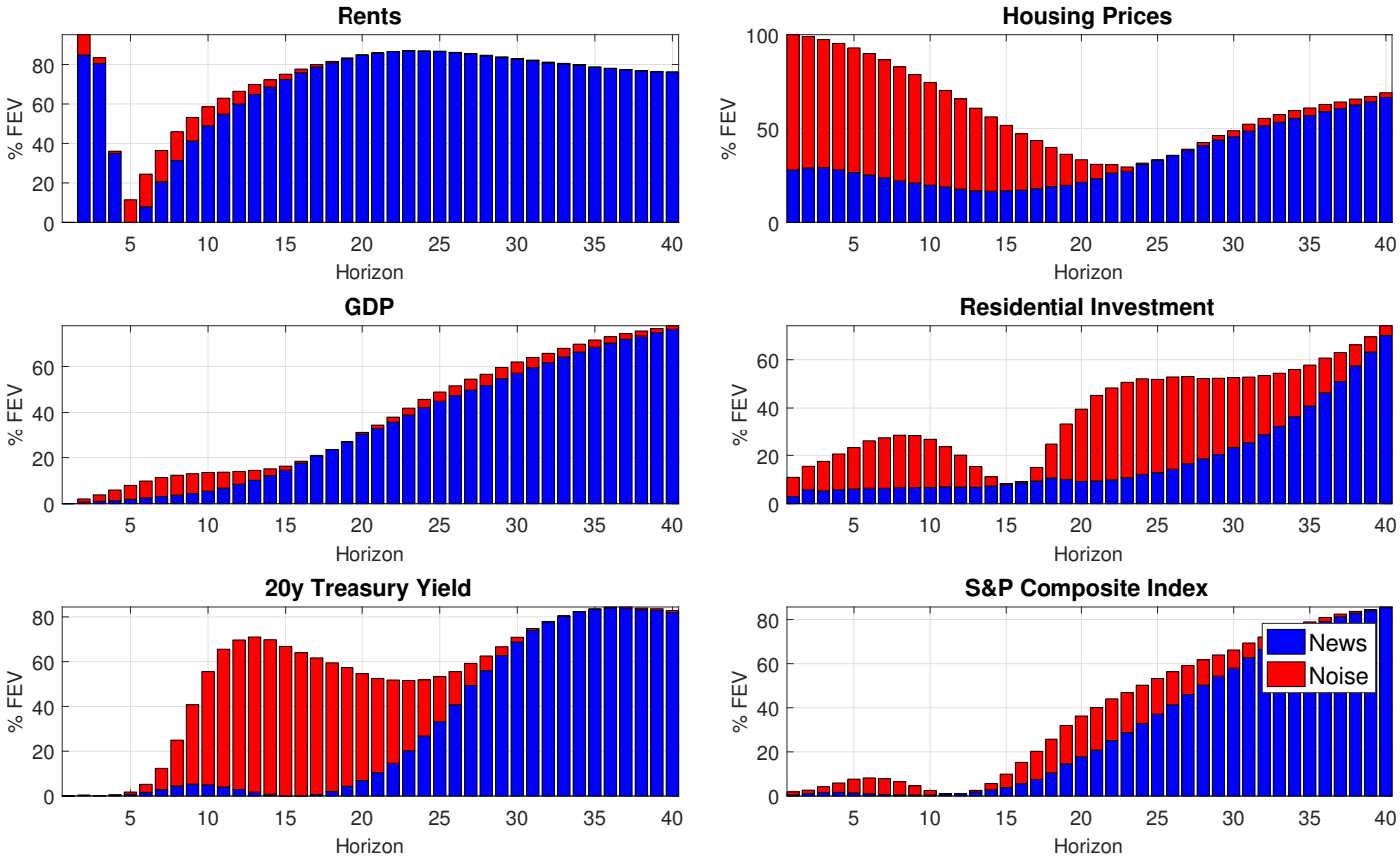

Figure A.16: Forecast Error Variance Decompostion - C&S Corelogic

Variance Decomposition - share of the variance explained by News and Noise at each quarter (not cumulative)

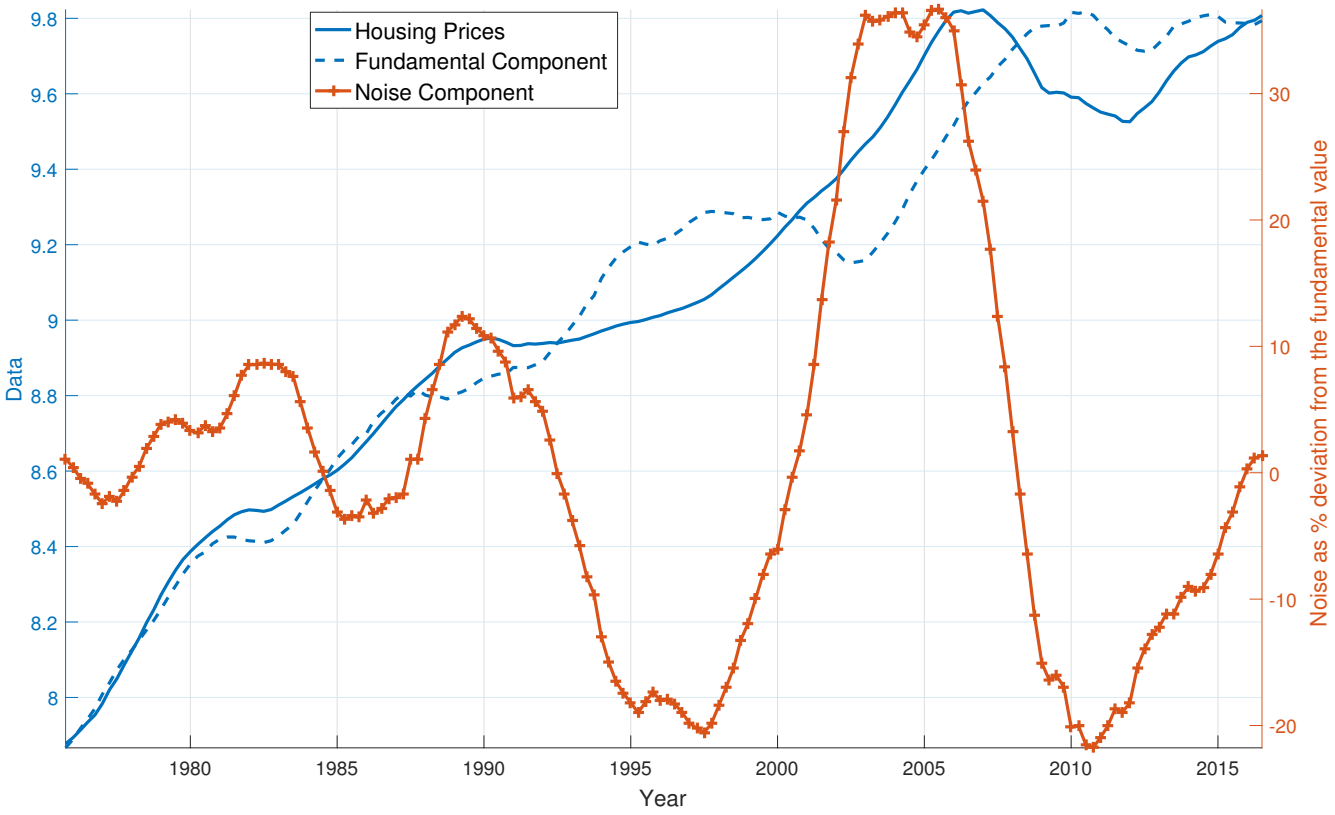

Figure A.17: Historical Decomposition - C&S Corelogic

Historical decomposition of housing prices (dotted blue) into a fundamental component (blue) and noisy component (orange). The shocks are identified through the following ordering: [Rents, Housing Prices, GDP, Residential Investment, Treasury Yield at Constant 20 year Maturity, S&P Index]. Sample: 1963:Q1 - 2016Q3.

A.3.3 Pre-2007 Crash Sample

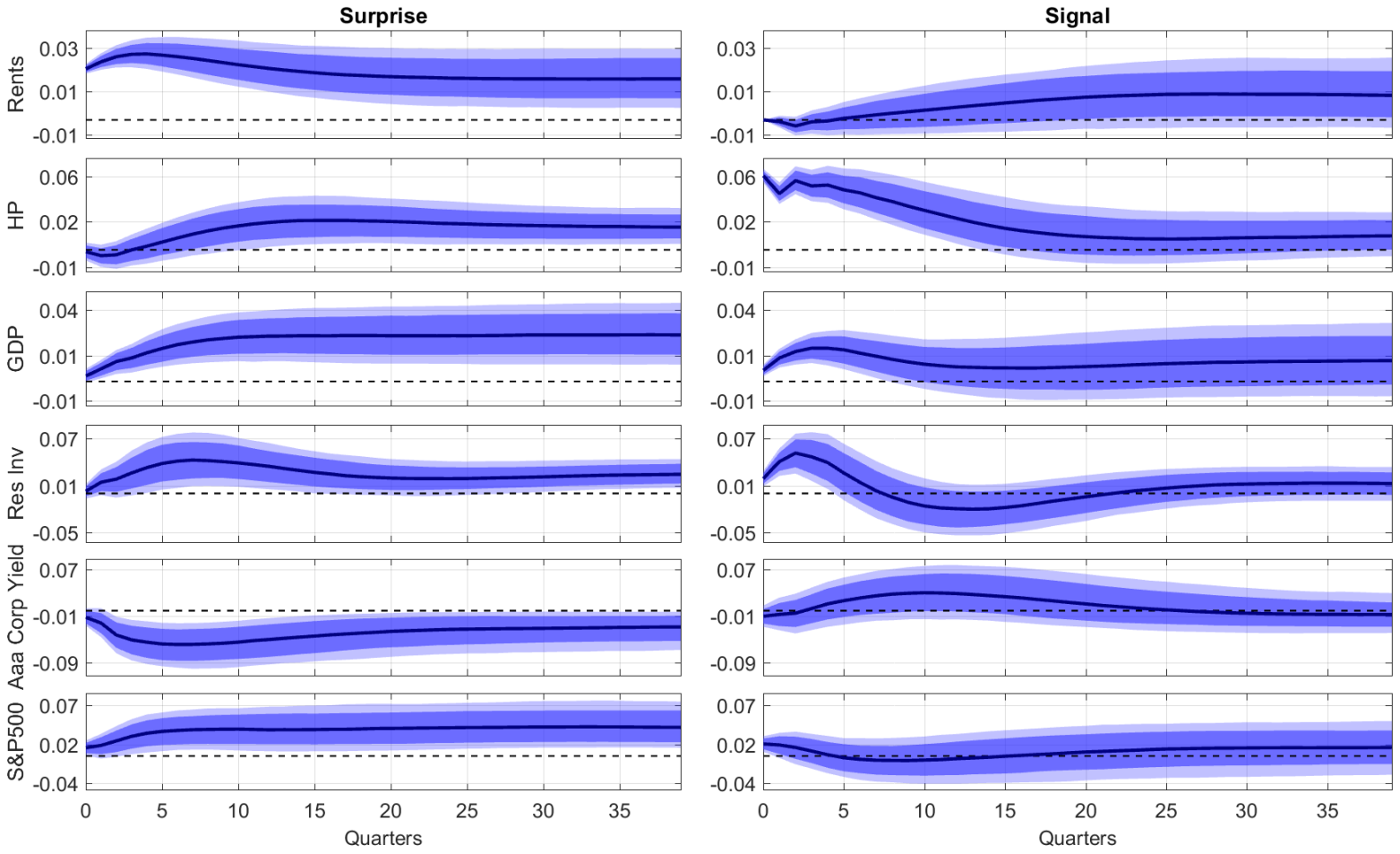

Figure A.18: IRFs to Surprise and Signal Shocks - pre 2007

IRFs to a surprise shock to rents and to a signal shock. The responses are reported in terms of the standard deviation of the variables in the system. The solid blue line is the median, the dark and light blue shaded areas represents 68% and 90% confidence bands respectively (2000 bootstrap replications). The shocks are identified through the following ordering: [Rents, Housing Prices, GDP, Residential Investment, Aaa Moody's Corporate Bond Yield, S&P Index]. Sample: 1963:Q1 - 2006:Q4.

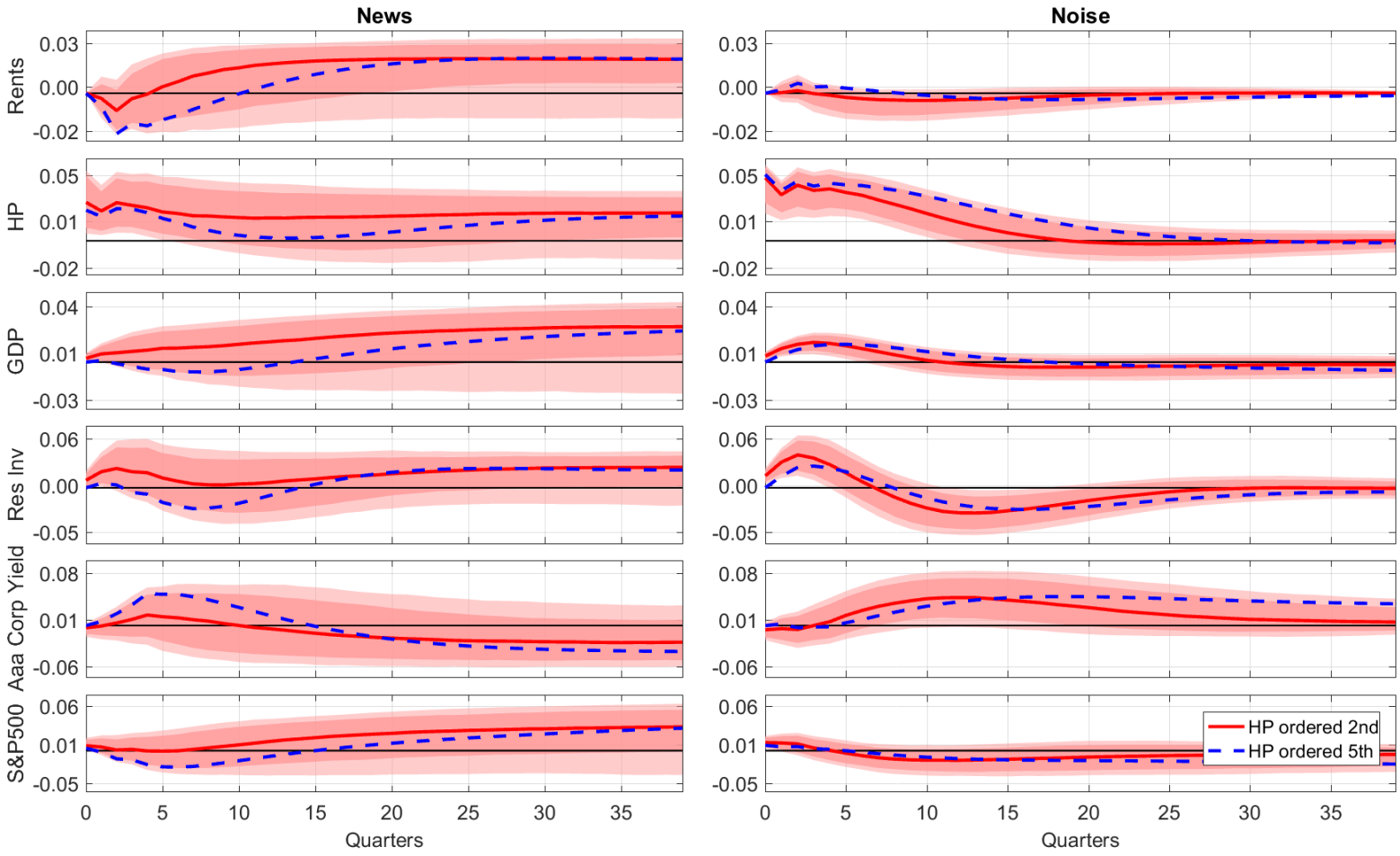

Figure A.19: IRFs to Surprise and Signal Shocks - pre 2007

IRFs to news and noise shocks. The responses are reported in terms of the standard deviation of the variables in the system. The solid red line is the median, the dark and light red shaded areas represents 68% and 90% confidence bands respectively (2000 bootstrap replications). The shocks are identified through the following ordering: [Rents, Housing Prices, GDP, Residential Investment, Aaa Moody's Corporate Bond Yield, S&P Index]. The blue dotted line reports the median IRFs obtained by the recursive ordering [Rents, GDP, Residential Investment, Aaa Moody's Corporate Bond Yield, Housing Prices, S&P Index]. Sample: 1963:Q1 - 2006:Q4.

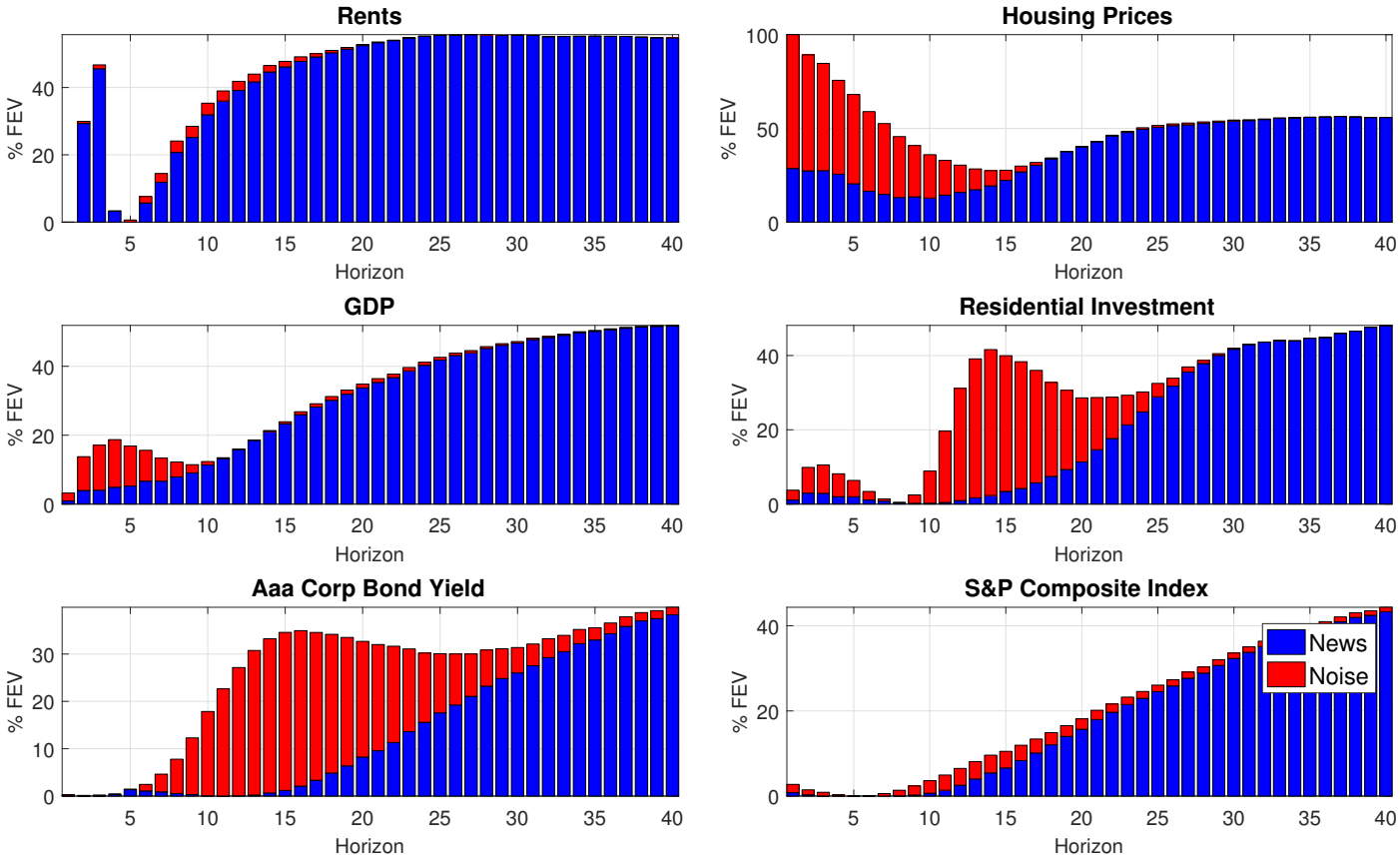

Figure A.20: Forecast Error Variance Decompostion - pre 2007

Forecast error variance decomposition of the variables in the system. The plot display the share of the variance explained by news and noise at each horizon (not cumulatively). The shocks are identified through the following ordering: [Rents, Housing Prices, GDP, Residential Investment, Aaa Moody's Corporate Bond Yield, S&P Index]. Sample: 1963:Q1 - 2006Q4.

Appendix B

Appendix: Bridge Proxy-SVAR

B.1 Conservative Identification - Orthogonalization

Our contribution concerns the way of studying the relationship between HF and LF variables, independently of the particular identification scheme chosen. Nonetheless, we can take an additional step if we restrict the class of DGPs to the subset in which each structural shock is associated with one variable.[1B] Using the representation in eq. (2.3), this assumption means that $B_{11} > B_{21}$; $B_{22} > B_{12}$.[2B] Then, consider a case in which the HF identification employs a VAR, and the researcher does not dispose of other, economic based, identification schemes (first best). In this setting, we can think of a recursive ordering where y is placed last, after all the variables that constitute the information set Ψ, as a second best identification. Such procedure is namely an orthogonalization and it is equivalent to use the residuals from the regression of the variable of interest y on its previous lags p (where p are the lags included in the HF-VAR) and on the contemporaneous values and lags of Ψ:

$$y_t = \sum_{l=1}^{p} \beta^l y_{t-l} + \sum_{l=0}^{p} \alpha^l \Psi_{t-l} + e_t \qquad e_t \sim WN \tag{B.1}$$

If each shock is associated with a variable, regressing the variable of interest y_t on Ψ_t yields the new information introduced in the system uniquely by y_t, that we label ε_t^y.

[1B] This means that each innovation enter the system mainly through a specific variable. For example, we call structural shock an innovation in the variable y which is orthogonal to the innovations in other variables. Notice that this is one of the many interpretations of innovation.

[2B] The assumption is implicit in our notation ε_t^y and ε_t^x, but Section 2.2.2 is actually more general.

Intuitively, the econometrician is likely to face identification trade-offs across different schemes in applied research. The researcher observes the high frequency reduced form residual $\mathring{u}_t^y$ which is a linear combination of the structural shocks:

$$\begin{aligned} \mathring{u}_t^y &= b_{22}\varepsilon_t^y + b_{21}\varepsilon_t^x \\ &= b_{22}\left(\mu_1 \varsigma_t + \mu_2 \phi_t\right) + b_{21}\varepsilon_t^x \end{aligned} \tag{B.2}$$

Suppose that ς_t satisfies the strength requirement of an IV, such that the resulting estimates are statistically reliable: $\mathbb{E}\left[\varsigma_t u_\tau^y\right] = \mu_1 \neq 0$. Given this condition, the econometrician should favor the most conservative HF identifications that, even washing out the component ϕ_t, does no capture in the proxy any other shocks ε_t^x. While the former issue does not yield distorted estimates, this latter event would induce biases by violating the exclusion restriction.

Furthermore, we wish to highlight two advantages of this conservative identification. First, the orthogonalization is robust to misspecifications thanks to the instrumental variable approach embedded into it. The IV approach allows us to employ only an exogenous variation (a component of the true structural shock) and not the whole structural shocks. Second, this identification yields identified shocks orthogonal with respect to the remainder of the current and past information set. Macroeconomic variables are explicitly unobservable at LF and cannot be included in the HF system. However, financial variables respond to the new available information on macroeconomic variables in real-time.

B.1.1 An Illustrative Example

Let us consider how the conservative identification performs with respect to a more relaxed identification. We study a simply bivariate system and compare violations in the exclusion restriction in our instrument $\hat{\varepsilon}_t^y$, i.e. how large is the component of ε_t^x captured in $\hat{\varepsilon}_t^y$. The system is structured as

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_t^x \\ \varepsilon_t^y \end{bmatrix} \tag{B.3}$$

where we normalized $b_{11} = b_{22} = 1$. Recall the assumption $b_{11} > b_{12}$ and $b_{22} > b_{21}$ such that there is a mapping between variables and shocks. We restrict the parameter space to positive values of b_{12} and b_{21} to simplify the analysis. Moreover,

we are only interested in studying the impact matrix B, so we consider a process without persistence:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_t^x \\ \varepsilon_t^y \end{bmatrix} \tag{B.4}$$

Under the relaxed identification scheme, we simply take the reduced form residual of y as structural shock. The component of ε_t^x captured in this measure is b_{21}, i.e. how much ε_t^x impacts on y_t:

$$\hat{\varepsilon}_t^{y_R} = b_{21}\varepsilon_t^x + \varepsilon_t^y \tag{B.5}$$

Under the conservative identification scheme, we regress y_t on x_t and take the residuals:

$$\begin{aligned} y_t = & \quad \Theta x_t + \epsilon_t & \epsilon_t \sim WN \\ b_{21}\varepsilon_t^x + \varepsilon_t^y = & \quad \Theta\left(\varepsilon_t^x + b_{12}\varepsilon_t^y\right) + \epsilon_t & \epsilon_t \sim WN \end{aligned} \tag{B.6}$$

Applying the definition of OLS we obtain:

$$\begin{aligned} \hat{\Theta}_{OLS} &= \mathbb{E}\left[x_t x_t\right]^{-1}\mathbb{E}\left[x_t y_t\right] \\ &= \frac{\mathbb{E}\left[\left(b_{21}\varepsilon_t^x + \varepsilon_t^y\right)\left(\varepsilon_t^x + b_{12}\varepsilon_t^y\right)\right]}{\mathbb{E}\left[\left(\varepsilon_t^x + b_{12}\varepsilon_t^y\right)\left(\varepsilon_t^x + b_{12}\varepsilon_t^y\right)\right]} \\ &= \frac{b_{21} + b_{12}}{1 + b_{12}^2} \end{aligned} \tag{B.7}$$

The residuals are computed as

$$\begin{aligned} y_t - x_t\hat{\Theta}_{OLS} &= b_{21}\varepsilon_t^x + \varepsilon_t^y - \hat{\Theta}_{OLS}\left(\varepsilon_t^x + b_{12}\varepsilon_t^y\right) \\ &= \left(1 - b_{12}\hat{\Theta}_{OLS}\right)\varepsilon_t^y + \left(b_{21} - \hat{\Theta}_{OLS}\right)\varepsilon_t^x \\ &= \left(1 - \frac{b_{12}^2 + b_{21}b_{12}}{1 + b_{12}^2}\right)\varepsilon_t^y + \left(b_{21} - \frac{b_{12} + b_{21}}{1 + b_{12}^2}\right)\varepsilon_t^x \\ \hat{\varepsilon}_t^{y_C} &= \Lambda\varepsilon_t^y + \Gamma\varepsilon_t^x \end{aligned} \tag{B.8}$$

Γ represents a measure of violation in the exclusion restriction. In two extreme cases: $b_{21} = 0 \Rightarrow \parallel \Gamma \parallel = \frac{b_{12}}{1+b_{12}^2}$ and $b_{12} = 0 \Rightarrow \parallel \Gamma \parallel = 0$. The comparison between relaxed and conservative identification reduces to the comparison between Γ and b_{21}. The condition $\Gamma < b_{21}$ is satisfied $\forall\{b_{12}, b_{12}\}$ as ε_t^x enters negatively

in $\hat{\varepsilon}_t^{y_C}$. This is likely to downward bias $\hat{\varepsilon}_t^{y_C}$ and make the first stage in the *Bridge* ineffective. However, let us consider the modulus of Γ for completeness:

$$\begin{aligned} \| \Gamma \| &< b_{21} \Rightarrow -b_{21} < \Gamma < b_{21} \\ b_{21} &> \frac{b_{12}}{(2b_{12}^2 + 1)} \end{aligned} \tag{B.9}$$

A graphical representation of the analytical results is provided below in Fig. B.1. The same results hold in a simulation design (Fig. B.2). The conservative identification is overall better in building an exogenous instrument than a more relaxed identification. The exception comes from low values of b_{21}. However, when b_{21} overcomes a certain threshold than the gains from the conservative over the relaxed identification are exponentially increasing (and the value of b_{12} does not matter anymore). In terms of economic interpretation, the *Bridge* is designed to study the effect of a shock to an HF variable y. b_{21} represents how much y responds to other shocks on impact. We can realistically state that, if y is financial variable, b_{21} takes large values and, in such a way, the conservative identification dominates the relaxed identification.

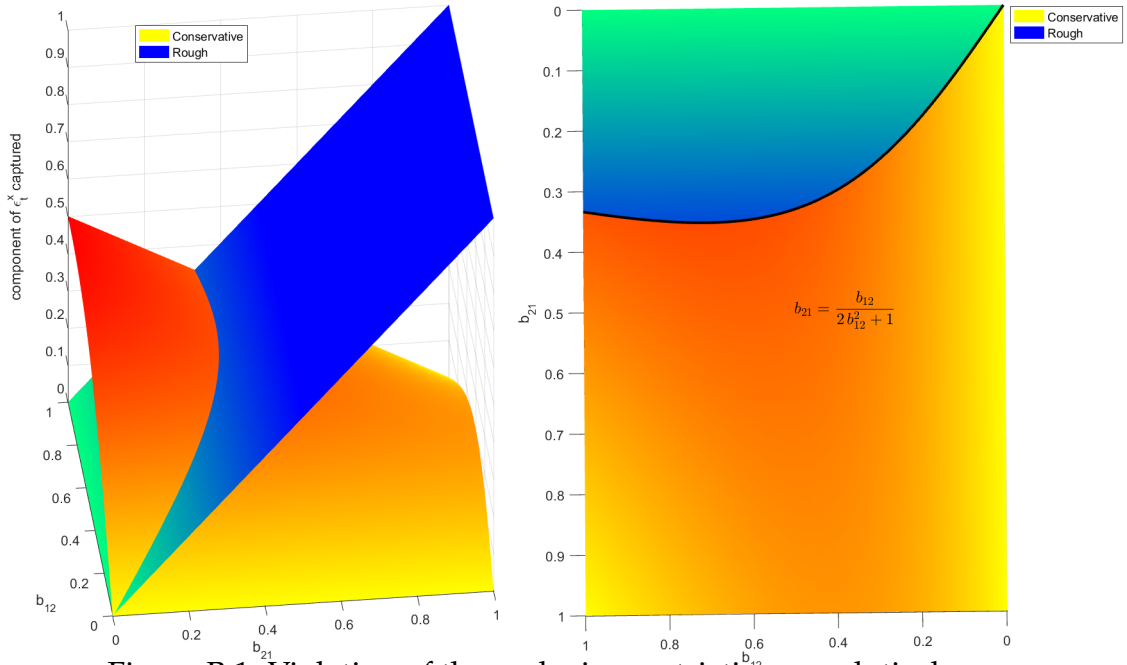

Figure B.1: Violation of the exclusion restriction - analytical case

Comparison of the violation of the exclusion restrictions between our conservative and rough (relax) identifications over the parameter space $\{b_{12}, b_{21}\} = \{0,1\}x\{0.1\}$. The left panel is a 3D plot, while in the right panel the size of the violation of the exclusion restriction have been collapsed. Where colors are cold $b_{21} < \Gamma$, where they are warm $b_{21} > \Gamma$. In black we report the analytical condition where b_{21} crosses Γ.

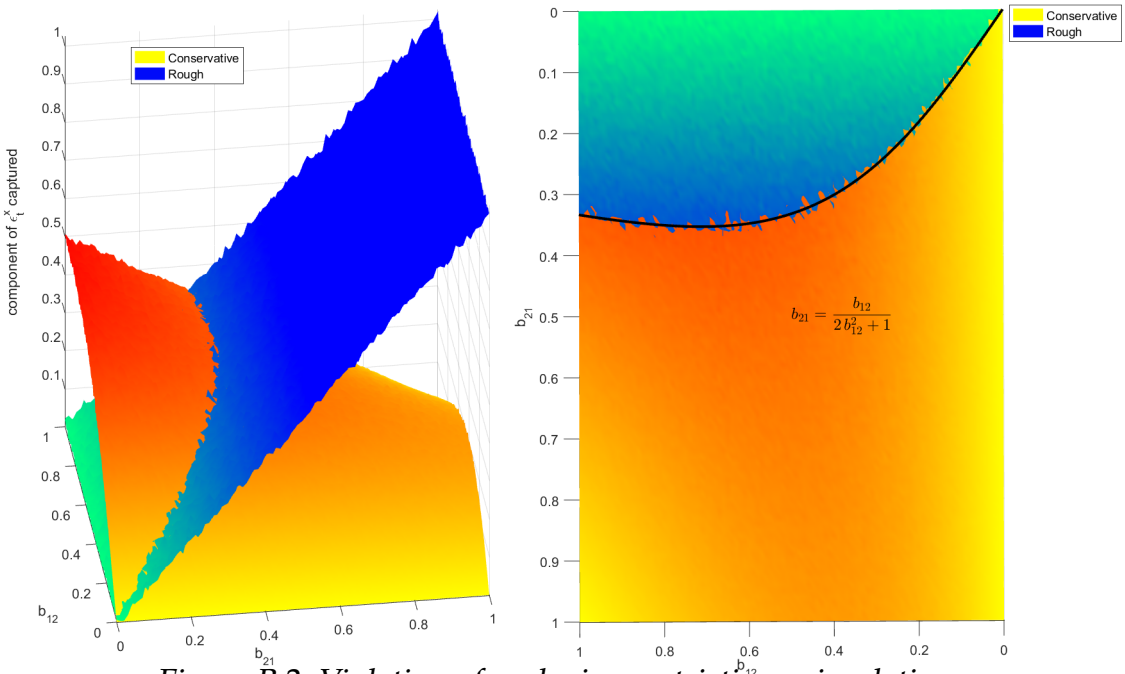

Figure B.2: Violation of exclusion restriction - simulation

Comparison of the violation of the exclusion restrictions between our conservative and rough (relax) identifications over the parameter space $\{b_{12}, b_{21}\} = \{0,1\}x\{0.1\}$. The left panel is a 3D plot, while in the right panel the size of the violation of the exclusion restriction have been collapsed. Where colors are cold $b_{21} < \Gamma$, where they are warm $b_{21} > \Gamma$. In black we report the analytical condition where b_{21} crosses Γ.

B.1.2 Monte Carlo Performances

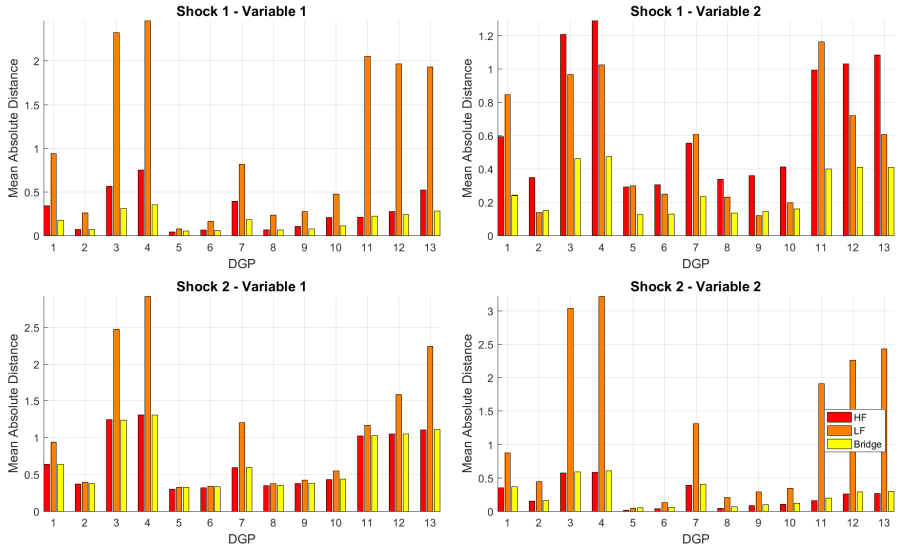

Figure B.3: MAD comparison in the two variable system: mispecification

Mean Absolute Distance (MAD) of IRFs estimated with the HF-VAR, LF-VAR and Bridge Proxy-SVAR in the 13 DGP cases. Time aggregation follows a skip-sampling scheme. Our conservative identification at HF is applied in this case. IRFs are standardize with respect to the true size of the shock.

MAD GAINS OVER LF-VAR		
Identification	Temporal Aggregation Scheme	
	SKIP-SAMPLING	AVERAGING
Full information at HF for *Bridge*: Quarterly-Monthly Frequency Mismatch		
HF-VAR	21.2%	41.4%
Bridge	20%	36.7%
Bridge - conservative identification	21.7%	38.3%
Full information at HF for *Bridge*: Monthly-Daily Frequency Mismatch		
HF-VAR	70%	81.2%
Bridge	65.6%	72.6%
Bridge - conservative identification	65.2%	74.7%

Table B.1: Performance comparison in Monte Carlo simulations - additional cases

Performance comparison across the counter-factual HF-VAR, the LF-VAR and the Bridge Proxy-SVAR. Performances are evaluated in terms of the Mean Absolute Distance (MAD) between the true IRFs and the estimated IRFs in 100 randomly parametrized DGPs. One summary statistic is computed based all the combinations of shocks-variables in the system. The gains are expressed as percentage MAD gains over the LF-VAR. We analyze different cases for a VAR(1) DGP: I) The Bridge employs full information at HF and the impact matrix B is diagonally dominated; II) The Bridge employs full information at HF and no restrictions are imposed on the impact matrix B; III) The Bridge employs only partial information at HF and no restrictions are imposed on the impact matrix B. The system features nine variables and the frequency mismatch is three (quarterly-monthly case). When possible, i.e. under full information, for the Bridge, we report both the results under the same identification of LF/HF-VAR and our conservative identification.

B.2 Skip Sampling Temporal Aggregation

B.2.1 Temporal Aggregation Bias

Following the recursive structure embodied in the impact matrix B, a Cholesky decomposition on the reduced form residuals at HF would yield the impact matrix itself:

$$Chol(BB') = \begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix} = B \tag{B.10}$$

However, when we move to the time aggregation case, even the correct identification scheme yields biases. In fact, we impose the zero restriction on the time aggregated reduced form residuals, whose variance-covariance matrix is given by:

$$\begin{aligned} \Omega &= BB' + ABB'A' \\ &= \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix} \end{aligned} \tag{B.11}$$

$$\begin{aligned}
\omega_{11} &= a_{12}\left[a_{12}b_{22}^2 + b_{21}\left(a_{11}b_{11} + a_{12}b_{21}\right)\right] + b_{11}^2 + a_{11}b_{11}\left(a_{11}b_{11} + a_{12}b_{21}\right) \\
\omega_{12} &= a_{22}\left[a_{12}b_{22}^2 + b_{21}\left(a_{11}b_{11} + a_{12}b_{21}\right)\right] + b_{11}b_{21} + a_{21}b_{11}\left(a_{11}b_{11} + a_{12}b_{21}\right) \\
\omega_{21} &= a_{12}\left[a_{22}b_{22}^2 + b_{21}\left(a_{21}b_{11} + a_{22}b_{21}\right)\right] + b_{11}b_{21} + a_{11}b_{11}\left(a_{21}b_{11} + a_{22}b_{21}\right) \\
\omega_{22} &= a_{22}\left[a_{22}b_{22}^2 + b_{21}\left(a_{21}b_{11} + a_{22}b_{21}\right)\right] + b_{21}^2 + b_{22}^2 + a_{21}b_{11}\left(a_{21}b_{11} + a_{22}b_{21}\right)
\end{aligned}$$

The Cholesky decomposition of Ω yields:

$$Chol(\Omega) = \begin{bmatrix} c_{11} & 0 \\ c_{21} & c_{22} \end{bmatrix} \tag{B.12}$$

$$
\begin{aligned}
c_{11} &= \left(a_{11}^2 b_{11}^2 + 2a_{11}a_{12}b_{11}b_{21} + a_{12}^2 b_{21}^2 + a_{12}^2 b_{22}^2 + b_{11}^2\right)^{\frac{1}{2}} \\
c_{21} &= \left(\frac{b_{11}b_{21} + a_{11}a_{21}b_{11}^2 + a_{12}a_{22} * b_{21}^2 + a_{12}a_{22}b_{22}^2 + a_{11}a_{22}b_{11}b_{21} + a_{12}a_{21}b_{11}b_{21}}{a_{11}^2 b_{11}^2 + 2a_{11}a_{12}b_{11}b_{21} + a_{12}^2 b_{21}^2 + a_{12}^2 b_{22}^2 + b_{11}^2}\right)^{\frac{1}{2}} \\
c_{22} &= \Big[\left(b_{21}^2 + b_{22}^2 + a_{21}^2 b_{11}^2 + a_{22}^2 b_{21}^2 + a_{22}^2 b_{22}^2 + 2a_{21}a_{22}b_{11}b_{21}\right) * \\
&\quad \left(a_{11}^2 b_{11}^2 + 2a_{11}a_{12}b_{11}b_{21} + a_{12}^2 b_{21}^2 + a_{12}^2 b_{22}^2 + b_{11}^2\right) + \\
&\quad - \left(b_{11}b_{21} + a_{11}a_{21}b_{11}^2 + a_{12}a_{22}b_{21}^2 + a_{12}a_{22}b_{22}^2 + a_{11}a_{22}b_{11}b_{21} + a_{12}a_{21}b_{11}b_{21}\right)^2\Big]^{\frac{1}{2}} * \\
&\quad (a_{11}^2 b_{11}^2 + 2a_{11}a_{12}b_{11}b_{21} + a_{12}^2 b_{21}^2 + a_{12}^2 b_{22}^2 + b_{11}^2)^{-1/2}
\end{aligned}
$$

where $\{c_{11}, c_{12,} c_{22}\} \neq \{b_{11}, b_{12,} b_{22}\}$ and the bias depends on the parametrization of of the DGP.

B.2.2 Monte Carlo - Additional Content

The parametrizations of the DGP (eq. 2.28) that we employ in the bivariate Monte Carlo simulations are:

$$
\begin{aligned}
\{\rho, \rho_l, , \rho\} = \; & \{0.5, 0.4, , 0.4\}\,; \{0.5, 0.08, , 0.4\}\,; \{0.9, 0.08, , 0.08\}\,; \{0.9, 0.1, , 0.08\}\,; \\
& \{0.1, 0.1, , 0.1\}\,; \{0.1, 0.4, , 0.4\}\,; \{0.1, 0.08, , 0.08\}\,; \{0.5, 0.1, , 0.1\}\,; \\
& \{0.5, 0.2, , 0.2\}\,; \{0.5, 0.4, , 0.2\}\,; \{0.9, 0.01, , 0.01\}\,; \{0.9, 0.04, , 0.04\}\,; \\
& \{0.9, 0.08, , 0.04\}\,;
\end{aligned}
$$

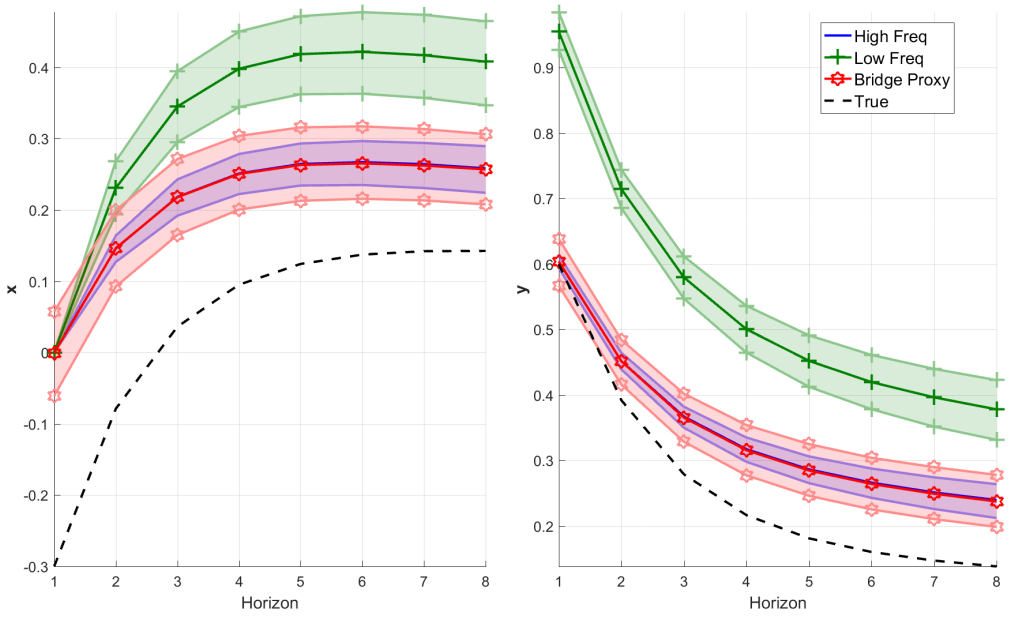

Figure B.4: IRFs2 in the two variable system: misspecification

IRFs to a shock in the second variable (y) in the bivariate system. The true IRF is represented by the dotted black line. The shock is identified through a wrong recursive structure in the HF system (blue), LF system (green) and Bridge Proxy (red). Shaded areas correspond to the 90% confidence bands across 1000 replications. Time aggregation follows a skip-sampling scheme.

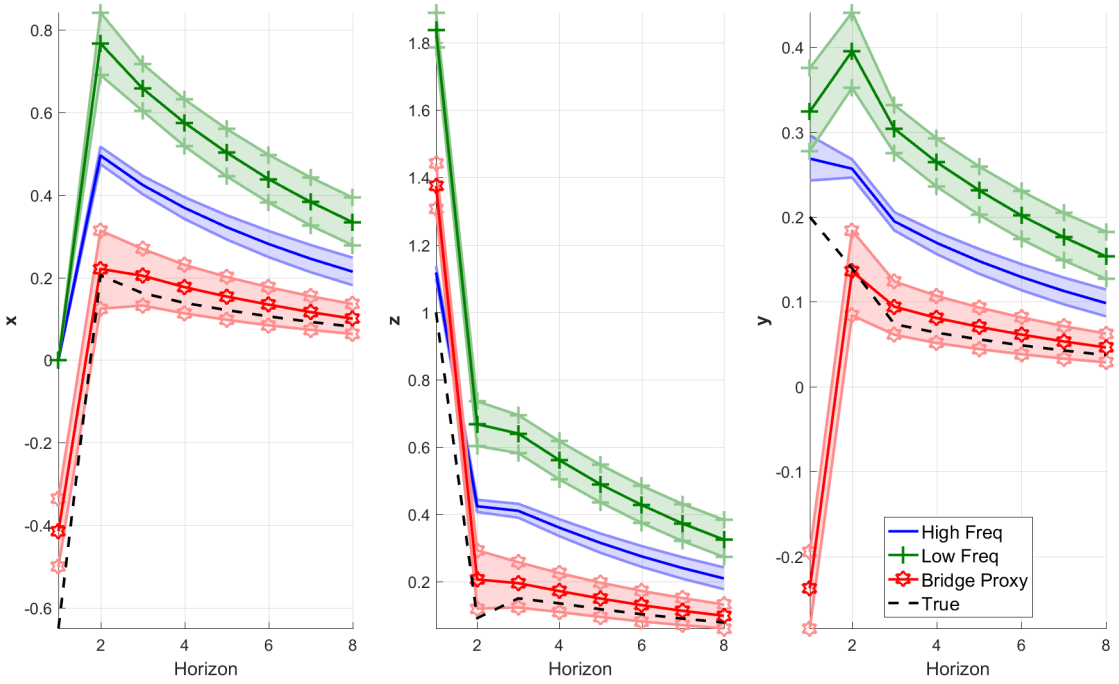

Figure B.5: IRF2 in the practical case

IRFs to a shock in the second variable (z) in the three variable system. Left panel - first variable (x); middle panel - second variable (z); right panel - third variable (y). The shock is identified through a wrong Cholesky in the HF system (blue), LF system (green) and Bridge Proxy (red). Shaded areas correspond to the 90% confidence bands. The black line is the true IRF. Time aggregation follows a skip-sampling scheme.

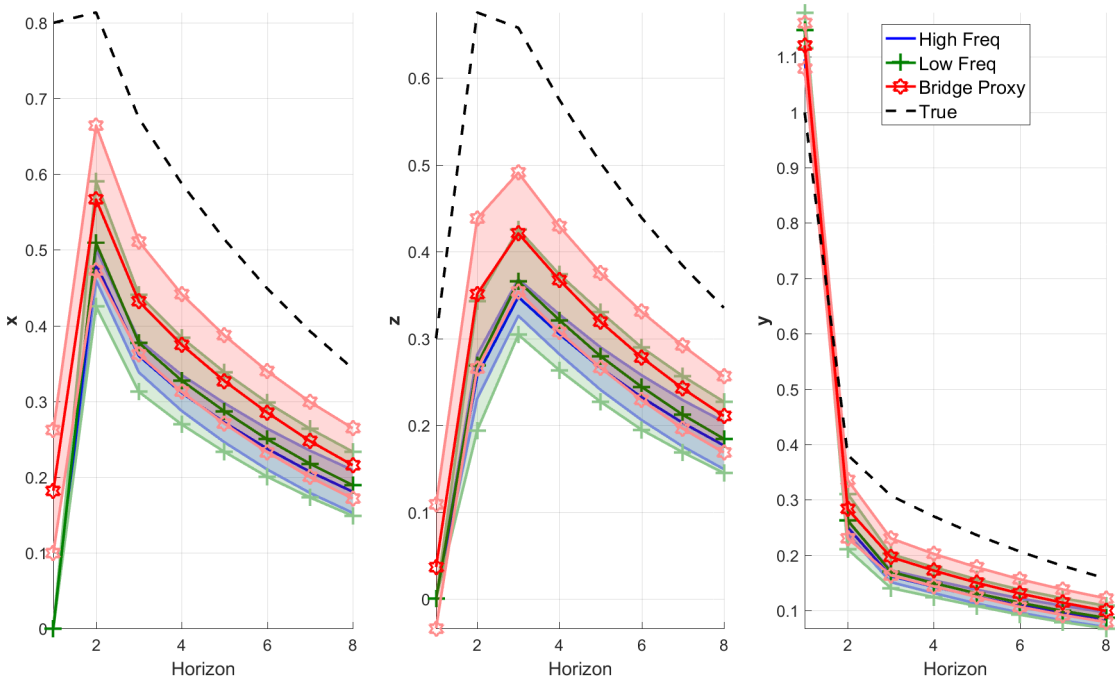

Figure B.6: IRF3 in the practical case

IRFs to a shock in the third variable (y) in the three variable system. Left panel - first variable (x); middle panel - second variable (z); right panel - third variable (y). The shock is identified through wrong a Cholesky in the HF system (blue), LF system (green) and Bridge Proxy (red). Shaded areas correspond to the 90% confidence bands. The black line is the true IRF. Time aggregation follows a skip-sampling scheme.

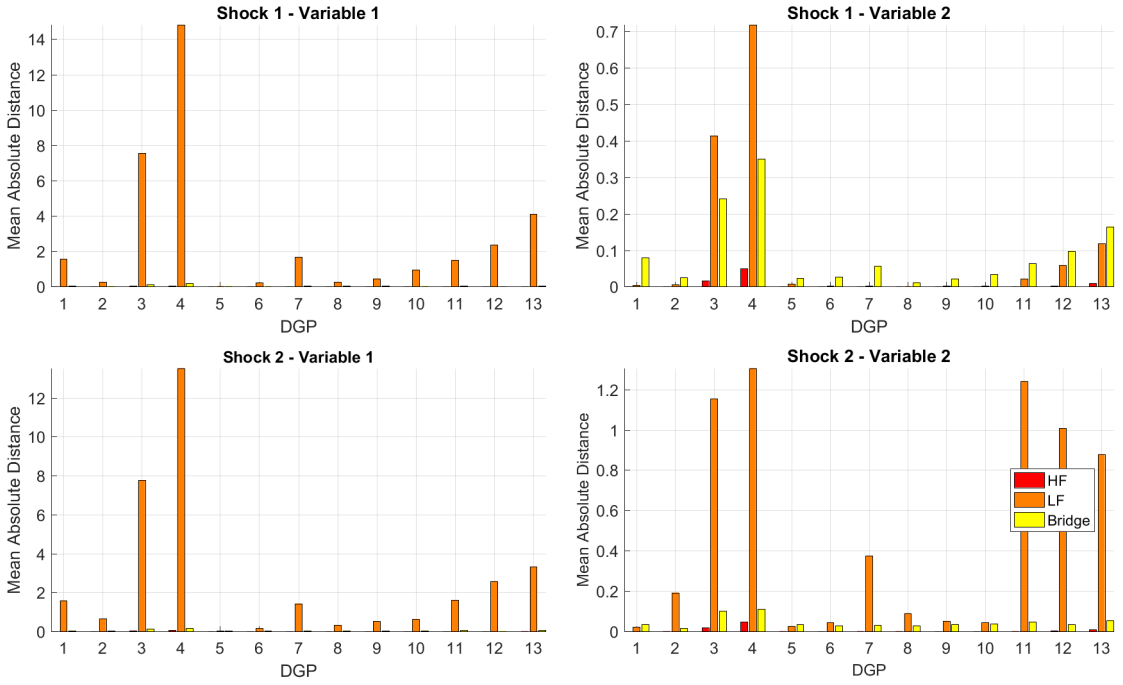

Figure B.7: MAD in the two varriable system: wider frequency mismatch

Mean Absolute Distance (MAD) between the true IRFs and the IRFs estimated by the HF-VAR, LF-VAR and Bridge Proxy-SVAR (through the correct recursive scheme). Results are reported for 13 parametrization of the DGP. The MAD is computed by averaging the MAD over the 1000 replications. Time aggregation follows a skip-sampling scheme.

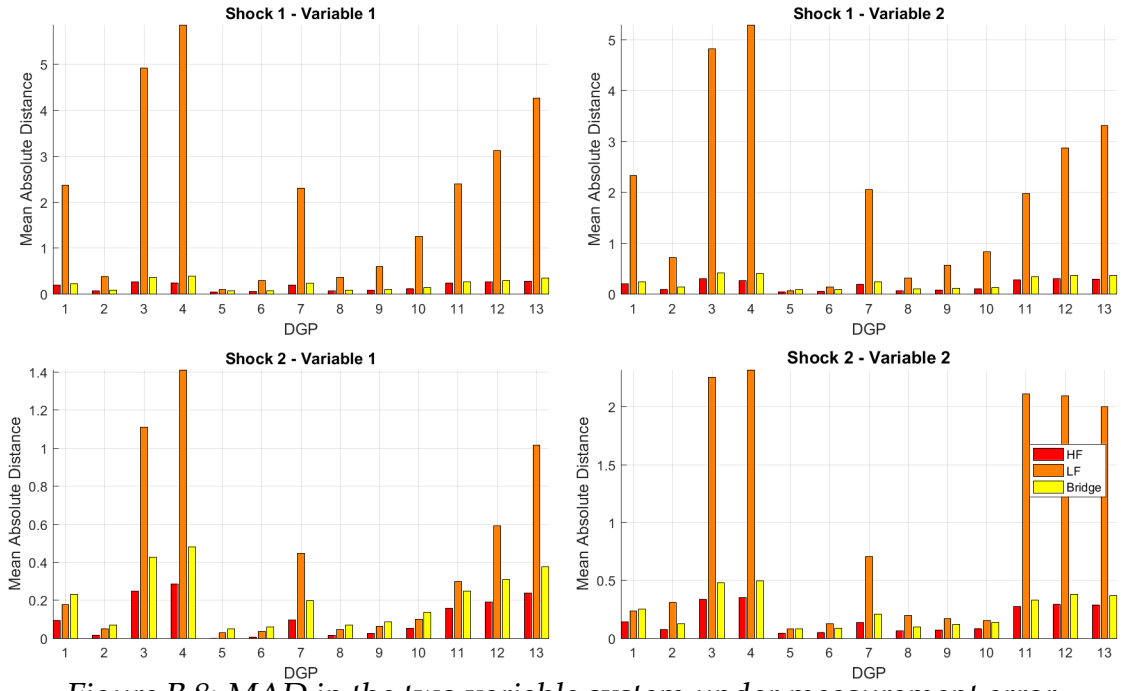

Figure B.8: MAD in the two variable system under measurement error

Mean Absolute Distance (MAD) between the true IRFs and the IRFs estimated by the HF-VAR, LF-VAR and Bridge Proxy-SVAR. Results are reported for 13 parametrization of the DGP. The MAD is computed by averaging the MAD over the 1000 replications. Time aggregation follows a skip-sampling scheme.

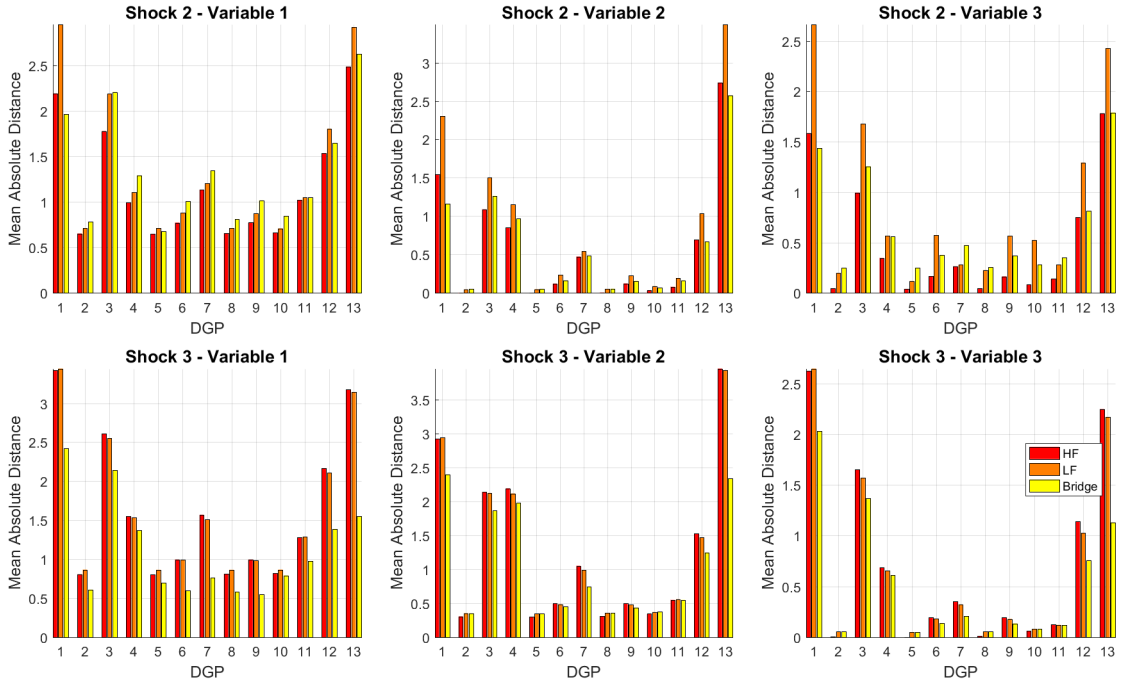

Figure B.9: MAD in the practical case: the wrong high frequency

Mean Absolute Distance (MAD) between the true IRFs and the IRFs estimated by the HF-VAR, LF-VAR and Bridge Proxy-SVAR. Results are reported for 13 parametrization of the DGP. The MAD is computed by averaging the MAD over the 1000 replications. Time aggregation follows a skip-sampling scheme.

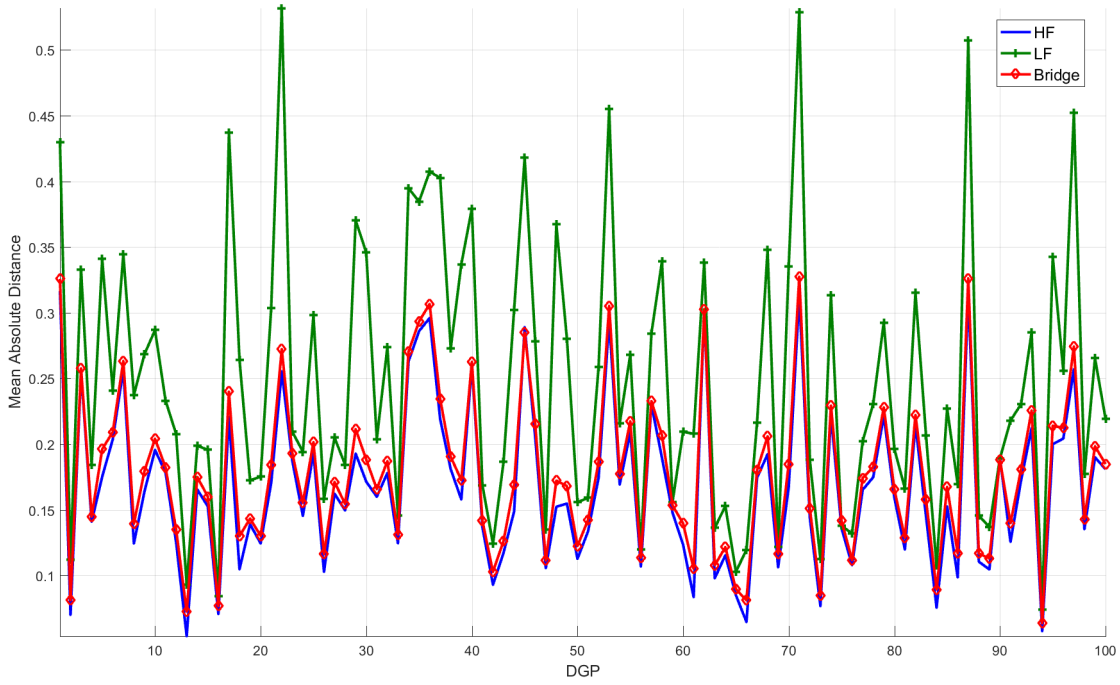

Figure B.10: MAD in each of the 100 large randomly parametrized systems

Mean Absolute Distance performances in the 100 randomly parametrized large systems of the HF-VAR, LF-VAR and Bridge Proxy-SVAR. The summary static is based on the percentage MAD between the true and estimated IRFs in each combination of shocks-variables in the system. Time aggregation follows a skip-sampling scheme.

B.3 Averaging Temporal Aggregation

B.3.1 An Illustrative Example

This section presents the same derivations of Section 2.2.3.1 but when time aggregation follows an averaging scheme. Averaging usually modifies the AR component in the same way as point-in-time sampling but induces higher order MA components.

$$\begin{aligned} Y_t &= AY_{t-1} + B\varepsilon_t \qquad \varepsilon_t \sim \mathcal{N}(0, \mathcal{I}) \\ (I - AL)\, Y_t &= B\varepsilon_t \qquad \varepsilon_t \sim \mathcal{N}(0, \mathcal{I}) \end{aligned} \tag{B.13}$$

To move to the time aggregated representation under averaging, we first apply the filter $w(L) = I + L$ to transform the series as sum (average is just a linear

transformation of it) and then we skip-sample through $D(L) = I + AL$:

$$\begin{aligned} D(L)w(L)\left(I - AL\right)Y_t &= D(L)Bw(L)\varepsilon_t \\ \left(I - A^2L^2\right)\left(I + L\right)Y_t &= \left(I + L\right)\left(I + AL\right)B\varepsilon_t \\ Y_t + Y_{t-1} &= A^2\left(Y_{t-2} + Y_{t-3}\right) + B\left(\varepsilon_t + \varepsilon_{t-1}\right) + AB\left(\varepsilon_{t-1} + \varepsilon_{t-2}\right) \\ Y_\tau &= CY_{\tau-1} + v_\tau \qquad v_\tau \sim \left(0, BB' + \left(\mathcal{I} + A\right)BB'\left(\mathcal{I} + A\right)' + ABB'A'\right) \\ Y_\tau &= CY_{\tau-1} + B\xi_t + AB\xi_{t-1} \qquad \xi_t \sim \left(0, \mathcal{I}\right),\ corr\left(\xi_t, \xi_{t-1}\right) = AB'B \end{aligned} \tag{B.14}$$

where $C = A^2$. Let us consider a bivariate system in extended notation:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix}\begin{bmatrix} \varepsilon_t^x \\ \varepsilon_t^y \end{bmatrix} \tag{B.15}$$

which is observed in time aggregation as

$$\begin{bmatrix} x_\tau \\ y_\tau \end{bmatrix} = \begin{bmatrix} a_{11}^2 + a_{12}a_{21} & a_{11}a_{12} + a_{12}a_{22} \\ a_{11}a_{21} + a_{21}a_{22} & a_{12}a_{21} + a_{22}^2 \end{bmatrix}\begin{bmatrix} x_{\tau-1} \\ y_{\tau-1} \end{bmatrix} + \begin{bmatrix} v_\tau^x \\ v_\tau^y \end{bmatrix} \tag{B.16}$$

where

$$\begin{bmatrix} v_\tau^x \\ v_\tau^y \end{bmatrix} = \begin{bmatrix} b_{11}\left(\varepsilon_t^x + \varepsilon_{t-1}^x\right) + \left(a_{11}b_{11} + a_{12}b_{21}\right)\left(\varepsilon_{t-1}^x + \varepsilon_{t-2}^x\right) + a_{12}b_{22}\left(\varepsilon_{t-1}^y + \varepsilon_{t-2}^y\right) \\ b_{21}\left(\varepsilon_t^x + \varepsilon_{t-1}^x\right) + b_{22}\left(\varepsilon_t^y + \varepsilon_{t-1}^y\right) + \left(a_{21}b_{11} + a_{22}b_{21}\right)\left(\varepsilon_{t-1}^x + \varepsilon_{t-2}^x\right) + a_{22}b_{22}\left(\varepsilon_{t-1}^y + \varepsilon_{t-2}^y\right) \end{bmatrix}$$

In this case, we employ as a proxy the first HF shock in the LF period to recover the true impact matrix. Namely, $z_\tau = \varepsilon_{t-1}^y$. The first stage in our IV procedure reads:

$$\begin{aligned} \hat{\beta}_{1s} &= \mathbb{E}\left[z_\tau' z_\tau\right]^{-1}\mathbb{E}\left[z_\tau' v_\tau^y\right] \\ &= \frac{\mathbb{E}\left[\varepsilon_{t-1}^y\left\{b_{21}\left(\varepsilon_t^x + \varepsilon_{t-1}^x\right) + b_{22}\left(\varepsilon_t^y + \varepsilon_{t-1}^y\right)\right\}\right]}{\mathbb{E}\left[\left(\varepsilon_{t-1}^y\right)\left(\varepsilon_{t-1}^y\right)\right]} \\ &\quad + \frac{\mathbb{E}\left[\varepsilon_{t-1}^y\left\{\left(a_{21}b_{11} + a_{22}b_{21}\right)\left(\varepsilon_{t-1}^x + \varepsilon_{t-2}^x\right) + a_{22}b_{22}\left(\varepsilon_{t-1}^y + \varepsilon_{t-2}^y\right)\right\}\right]}{\mathbb{E}\left[\left(\varepsilon_{t-1}^y\right)\left(\varepsilon_{t-1}^y\right)\right]} \\ &= \left(b_{22} + a_{22}b_{22}\right) \\ &= b_{22}\left(1 + a_{22}\right) \end{aligned} \tag{B.17}$$

and the fitted values are

$$\hat{\beta}_{1s} z_\tau = b_{22} (1 + a_{22}) \varepsilon^y_{t-1}$$

The second stage regression reads

$$\xi^x_\tau = \beta_{2s} \left(\hat{\beta}_{1s} z_\tau \right) + \varphi_\tau \qquad \varphi_\tau \sim WN$$

$$\begin{aligned}
\hat{\beta}_{2s} &= \mathbb{E}\left[\left(\hat{\beta}_{1s} z_\tau \right) \hat{\beta}_{1s} z_\tau \right]^{-1} \mathbb{E}\left[\hat{\beta}_{1s} z_\tau v^x_\tau \right] \\
&= \left(\hat{\beta}_{1s} \right)^{-1} \mathbb{E}\left[z_\tau z_\tau \right]^{-1} \mathbb{E}\left[z_\tau \xi^x_\tau \right] \\
&= \left(\hat{\beta}_{1s} \right)^{-1} \mathbb{E}\left[\varepsilon^y_{t-1} \left\{ b_{11} \left(\varepsilon^x_t + \varepsilon^x_{t-1} \right) + \left(a_{11} b_{11} + a_{12} b_{21} \right) \left(\varepsilon^x_{t-1} + \varepsilon^x_{t-2} \right) + a_{12} b_{22} \left(\varepsilon^y_{t-1} + \varepsilon^y_{t-2} \right) \right\} \right] \\
&= \frac{a_{12} b_{22}}{b_{22} (1 + a_{22})} \\
&= \frac{a_{12}}{1 + a_{22}}
\end{aligned} \tag{B.18}$$

We obtain an equivalent result if we apply straight the definition of IV estimator:

$$\begin{aligned}
\hat{\beta}_{Proxy} &= \mathbb{E}\left[z_\tau v^y_\tau \right]^{-1} \mathbb{E}\left[z_\tau v^x_\tau \right] \\
&= \frac{\mathbb{E}\left[\varepsilon^y_{t-1} \left\{ b_{11} \left(\varepsilon^x_t + \varepsilon^x_{t-1} \right) + \left(a_{11} b_{11} + a_{12} b_{21} \right) \left(\varepsilon^x_{t-1} + \varepsilon^x_{t-2} \right) + a_{12} b_{22} \left(\varepsilon^y_{t-1} + \varepsilon^y_{t-2} \right) \right\} \right]}{\mathbb{E}\left[\varepsilon^y_{t-1} \left\{ b_{21} \left(\varepsilon^x_t + \varepsilon^x_{t-1} \right) + b_{22} \left(\varepsilon^y_t + \varepsilon^y_{t-1} \right) + \left(a_{21} b_{11} + a_{22} b_{21} \right) \left(\varepsilon^x_{t-1} + \varepsilon^x_{t-2} \right) + a_{22} b_{22} \left(\varepsilon^y_{t-1} + \varepsilon^y_{t-2} \right) \right\} \right]} \\
&= \left[b_{22} (1 + a_{22}) \right]^{-1} a_{12} b_{22} \\
&= \frac{a_{12}}{1 + a_{22}}
\end{aligned} \tag{B.19}$$

It is important to highlight that, even if we are able to recover the true $IRFs$ on impact, the estimated autoregressive matrix of the LF-VAR is biased due to the VARMA structure of the temporally aggregated process.[3B] VARMA models are not used in empirical application due the high parametrization and severe problems in defining an economic interpretable structure (SVARMA). Therefore, we do not tackle this issue as the improvement in identification over a LF-VAR is the

[3B] The bias in the estimated A matrix induces a bias also in the estimated reduced form residuals. However, the IRFs on impact (B) would be biased only if the bias in the A matrix were correlated with the structural shocks. In a simple AR(1) process, the bias is a constant and so does not interfere with the estimates of the B matrix. Moreover, our simulations of more complex processes indicate that the *Bridge* always recover the impact response.

best we can reach through our methodology. This steams from the fact that we derive identifying restrictions at HF but we still rely on the LF-VAR representation for the transmission of the shocks. On the contrary, the state space MF-VAR improves the estimates of the A matrix by shifting the representation of the LF variables at HF.

B.3.2 Comparison *Bridge* - Mixed Frequency VAR

If financial processes are part of the analysis, the shortcoming of the MF-VAR consists of the inability to use daily data.[4B] To the best of our knowledge, the MF-VAR can exploit at most weekly data. Therefore, there is a trade-off between the identification of the impact matrix B, favorable to the *Bridge*, and the estimates of the autoregressive matrix A, favorable to the MF-VAR.[5B] Finally, notice that sample size is quite relevant in this trade-off: the biases in the estimate A matrix are decreasing in the sample size as the VARMA process is well approximate by a VAR in large samples but not in short samples.[6B]

We design two Monte Carlo experiments to compare the performances of the *Bridge* versus the MF-VAR. On the one hand, we quantitatively illustrate this trade-off. On the other hand, and more importantly, our goal is to study the dependence of the relative performances of the two methodologies on the parametrization of the DGP. Our intuition suggests that when the variables in the system are very responsive to other shocks on impact, i.e. the simultaneity problem is very severe, improving the estimation of the impact matrix is crucial.

We consider both a full information and partial information setup. In the full information case, the *Bridge* employs all variables in both stages, whereas the MF-VAR is actually the counter-factual HF-VAR. In the partial information case, we

[4B] For example, in a quarterly-weekly ($m = 12$) Monte Carlo simulation Foroni and Marcellino (2016) report:

1. "For computational reasons (the number of missing values is high and therefore the computational time increases substantially), we fix the number of replications to $R = 500$."
2. "Due to the higher number of missing values when $m = 12$, we increase the size to 300 quarterly observations to obtain more stable results when running the Kalman filter."

[5B] If the true process occurs at daily frequency while the MF-VAR employs weekly data, the estimates of the A matrix will still be biased, even if less than the monthly estimates.

[6B] Notice that, on the one hand, the strength of the instrument and the precision of the estimates is increasing with the sample size for the *Bridge*. On the other hand, the computational burden of the MF-VAR increases with the length of the sample.

run the practical case presented in Section 2.3.3. The first variable in the system is effectively unobservable at HF, so the *Bridge* employs only two variables in recovering the shocks at HF (first stage). The MF-VAR estimates in a state space representation the missing observations of the LF variable.

Full Information We employ a nine variable system to quantitatively evaluate the *A-B* trade-off, but we study also a two variable system to illustrate how this trade-off depends on simultaneity. The true frequency of the process is daily but macro variables are available only at the monthly frequency. We compare the best performances of a MF-VAR (HF-VAR) on weekly data with the best performances of the *Bridge* (full information) using daily as HF data and monthly as LF data. Once again, we run a 100 random parametrization experiment in a three variable system as we want to analyze the trade-off between *Bridge* (advantage in identifying the impact matrix) versus MF-VAR (advantage in estimating the autoregressive matrix). We do not constrain the generated parameter in anyway other than maintaining a mapping variables-shocks. Overall, we obtain the results displayed in Table B.2.

More importantly, for the bivariate case we build an index of relative performances for the cross impacts of the shocks and regress it on the parameters of the B matrix. Our index capture the percentage difference in the MAD between the MF-VAR and *Bridge*. Table B.3 confirms our priors: when the off-diagonal elements in the B matrix are large, the (daily-monthly) *Bridge* is preferred to the (weekly) MF-VAR.

Identification	MAD GAINS OVER LF-VAR (MONTHLY)
Bivariate system	
MF-VAR (HF-VAR weekly)	70.6%
Bridge (full-information daily)	78.7%
9 variable randomized system	
MF-VAR (HF-VAR weekly)	67.4%
Bridge (full-information daily)	66.2%

Table B.2: Performance comparison in Monte Carlo simulations - Bridge and MF-VAR

Performance comparison across the MF-VAR (weekly HF-VAR), the LF-VAR (monthly) and the (full information) Bridge Proxy-SVAR (daily-monthly). Performances are evaluated in terms of the Mean Absolute Distance (MAD) between the true IRFs and the estimated IRFs. The gains are expressed as percentage MAD gains over the LF-VAR. We report the results for I) the bivariate case used to evaluate the dependence of the performances on the structure of the DGP; II) a 9 variable randomly parametrized system in 100 randomly parametrized DGPs.

	(1) OLS	(2) OLS	(3) Probit	(4) Probit
VARIABLES	% Δ MAD MF-B Variable 2 - Shock 1	% Δ MAD MF-B Var 1 - Shock 2	% Δ MAD MF-B Var 2 - Shock 1	% Δ MAD MF-B Var 1 - Shock 2
$\lvert b_{12}\rvert$	0.40***	-0.25	1.07***	0.93*
	(0.13)	(0.18)	(0.27)	(0.49)
$\lvert b_{21}\rvert$	-0.05	0.47***	-0.31	0.51
	(0.12)	(0.17)	(0.23)	(0.38)
Constant	-0.19	1.10***	-0.73***	0.39
	(0.15)	(0.21)	(0.28)	(0.38)
Observations	100	100	100	100
R-squared	0.09	0.09	0.14 (pseudo)	0.11 (pseudo)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B.3: MAD comparison as function of DGP: full information

Relationship between relative performances of the (daily-monthly) Bridge over the (weekly) MF-VAR and the structure of the impact matrix. In particular, we study the relationship between the estimated cross IRFs with the absolute values of the off-diagonal elements in the B matrix: b_{12} and b_{21}. The higher the degree of simultaneity, the wider the gains from using daily data (Bridge) over weekly data (MF-VAR).

Partial Information We turn next to a three variable system where one variable is actually unobserved a HF and compare how the MF-VAR and *Bridge* cope with this lack of information. The LF variable is observable only once each 24 periods as average.[7B] The MF-VAR aggregates the HF over 8 periods and jointly estimate the relationship with the LF variable. Basically, the MF-VAR reverse to the monthly-quarterly case. Finally, the *Bridge* recovers shocks at the true frequency by using a bivariate system with the two variables available at HF. In terms of MAD percentage gains over the LF-VAR, the MF-VAR improves by 46.7%, while the *Bridge* by 70.5%.

However, more than providing a quantitative comparison across the two methodologies, we are interested in analyzing the cases that suit one or another procedure. As in the previous case, we regress the relative performances of the *Bridge* versus the MF-VAR on the parametrization of the B matrix. In particular, we focus on the simultaneity between the variables observable at the highest frequency. We analyze how this simultaneity affects the bias in the estimated responses of the low frequency variable to the high frequency shocks. Namely, we regress the bias in the IRF of variable x to shocks in z and y on b_{23} and b_{32}.[8B] The results presented Table highlight that the gains from using the *Bridge* increasing in the simultaneity across the high frequency variable. This finding suggests that the Bridge is particularly suitable to study macro-financial linkages where high frequency variables contemporaneous co-move significantly.

[7B] This number may be interpreted as the working days within one month.

[8B] We include b_{22} and b_{33} to take into account the size of the shock.

VARIABLES	(1) % Δ MAD MF-Bridge Variable 1 - Shock 2	(2) Δ MAD MF-Bridge Variable 1 - shock 3
$\lvert b_{32} \rvert$	2.49**	2.13*
	(1.15)	(1.27)
$\lvert b_{23} \rvert$	-0.63	2.22*
	(1.18)	(1.27)
$\lvert b_{33} \rvert$	0.71	-0.99
	(1.01)	(1.08)
$\lvert b_{22} \rvert$	-0.86	0.034
	(0.81)	(0.91)
Constant	1.49	0.39
	(1.84)	(1.99)
Observations	99	96
R-squared	0.058	0.074

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B.4: MAD comparison as function of DGP: partial information

Relationship between relative performances of the (daily-monthly) Bridge over the (weekly) MF-VAR and the structure of the impact matrix. In particular, we study the relationship between the estimated cross IRFs with the absolute values of the off-diagonal elements in the B matrix: b_{23} and b_{32}. These two parameters represent the degree of simultaneity between variable 2 (z) and variable 3 (y). The higher the degree of simultaneity, the wider the gains from using daily data (Bridge) over weekly data (MF-VAR).

B.3.3 Monte Carlo Simulations - Averaging Case

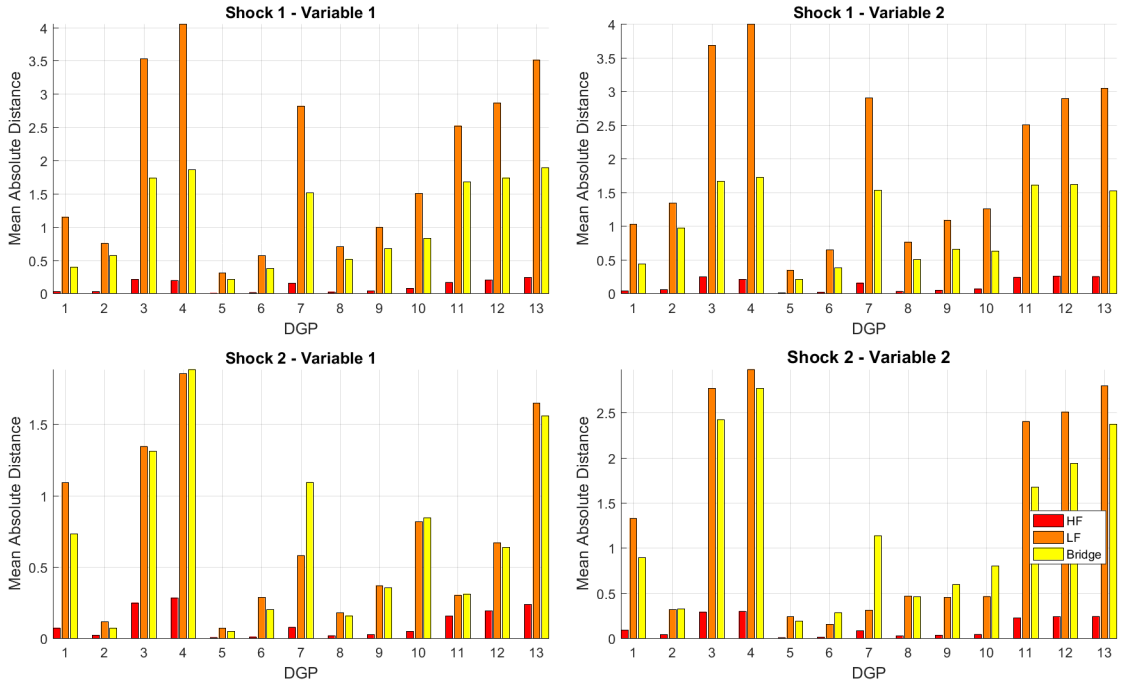

Figure B.11: MAD in the two variable system - averaging

Mean Absolute Distance (MAD) between the true IRFs and the IRFs estimated by the HF-VAR, LF-VAR and Bridge Proxy-SVAR (through the correct recursive scheme). Results are reported for 13 parametrization of the DGP. The MAD is computed by averaging the MAD over the 1000 replications. Time aggregation follows an averaging scheme.

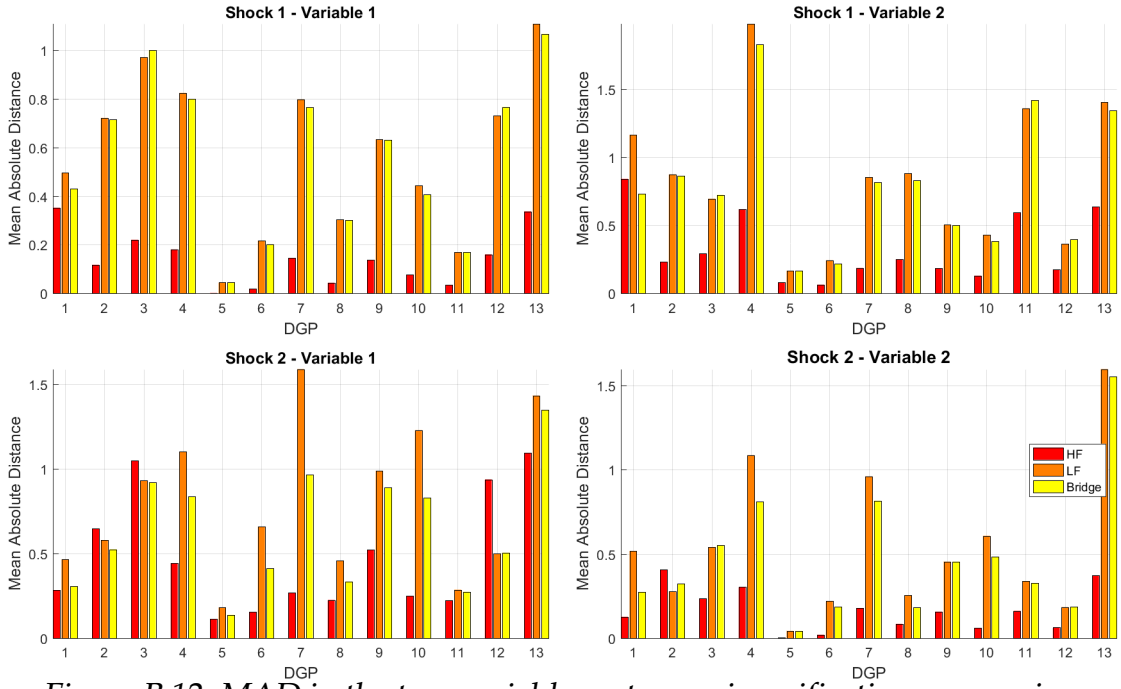

Figure B.12: MAD in the two variable system: mispecification - averaging

Mean Absolute Distance (MAD) between the true IRFs and the IRFs estimated by the HF-VAR, LF-VAR and Bridge Proxy-SVAR. Results are reported for 13 parametrization of the DGP. The MAD is computed by averaging the MAD over the 1000 replications. Time aggregation follows a skip-sampling scheme.

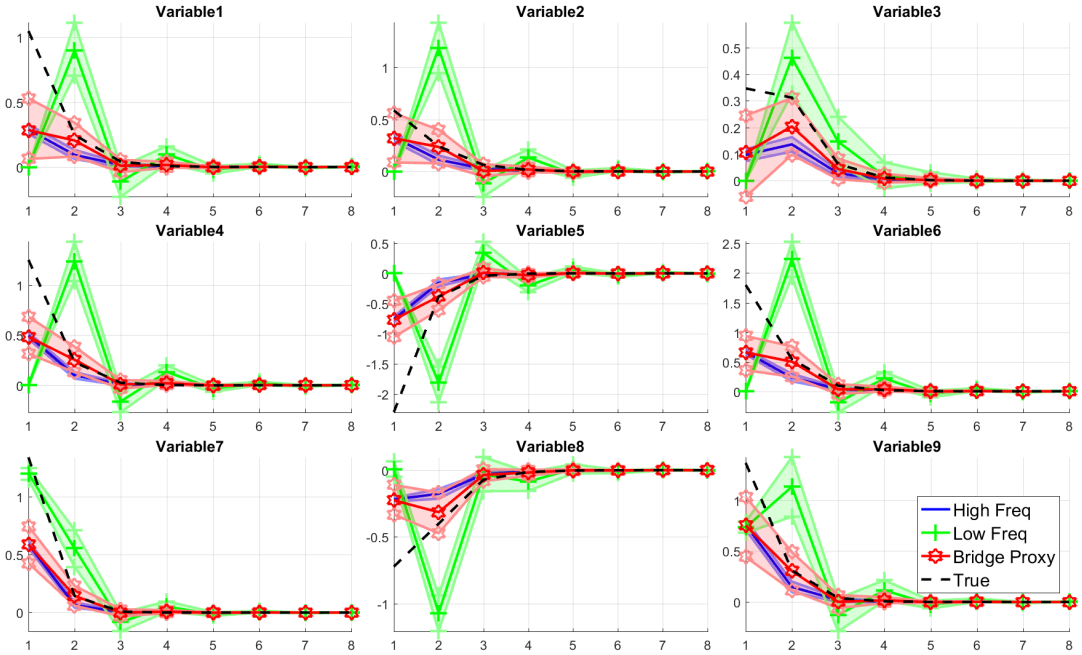

Figure B.13: IRFs from large randomized Monte Carlo experiment - averaging

Example of the IRFs of the system to a shock in the first variable in the system, estimated by the HF-VAR, LF-VAR and Bridge Proxy-SVAR in one of the 100 randomly parametrized DGPs. Shaded areas correspond to the 90% confidence bands across 1000 replications. The true IRF is represented by the dotted black line. Time aggregation follows an averaging scheme.

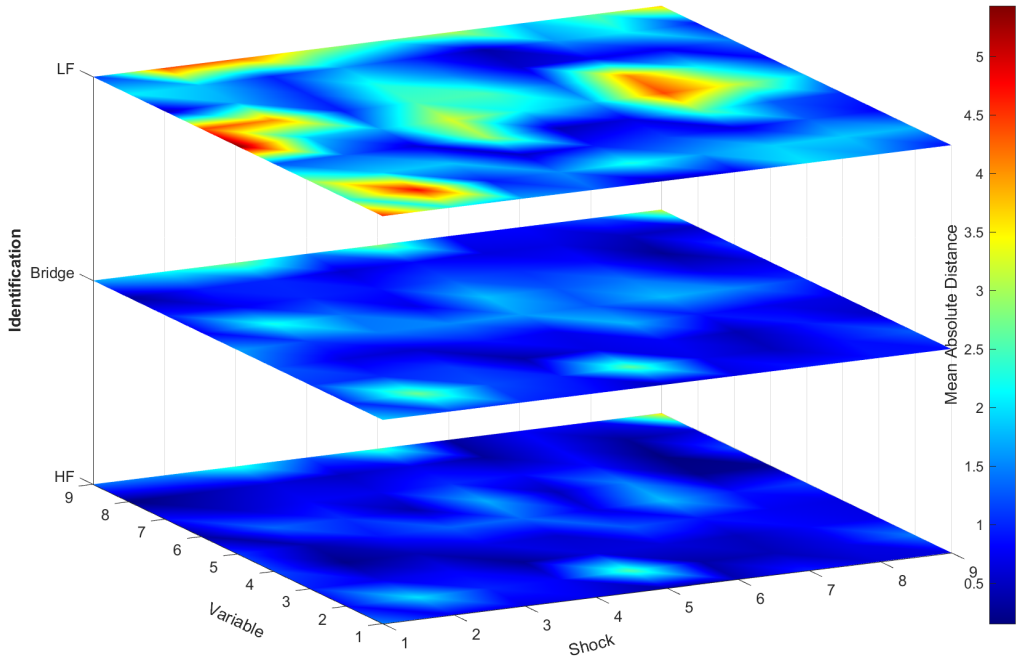

Figure B.14: MAD heatmap from large randomized Monte Carlo experiment - averaging

Mean Absolute Distance (MAD) between the true IRFs and the IRFs estimated by the HF-VAR, LF-VAR and Bridge Proxy-SVAR in one of the 100 randomly parametrized DGPs. Results are reported for each combination of shocks-variables in the system (81). The MAD is computed by averaging the MAD over the 1000 replications. Time aggregation follows a skip-sampling scheme.

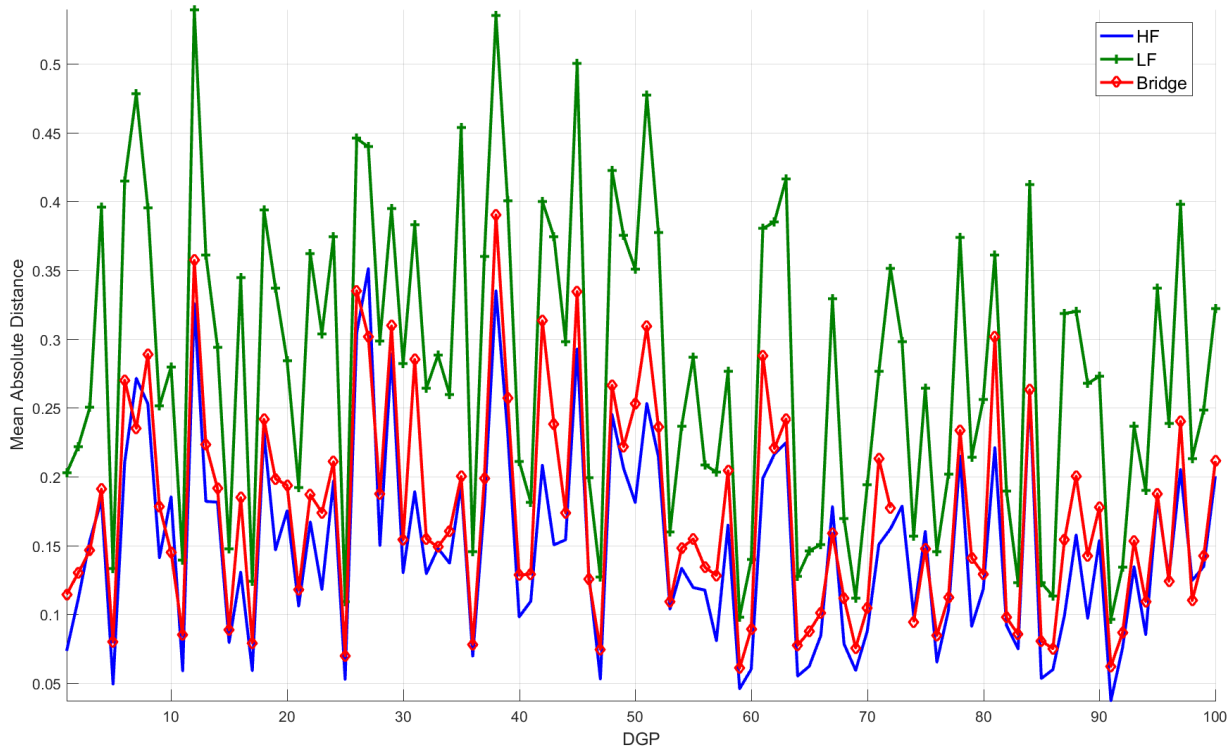

Figure B.15: MAD in each of the 100 large randomly parametrized systems

Mean Absolute Distance performances in the 100 randomly parametrized large systems of the HF-VAR, LF-VAR and Bridge Proxy-SVAR. The summary static is based on the percentage MAD between the true and estimated IRFs in each combination of shocks-variables in the system. Time aggregation follows a skip-sampling scheme.

B.4 Empirical Application

Name	Datastream Code
Fed Funds Future 3 months ahead	CFFCS30
S&P 500	S&PCOMP
Oil Price Index	OILBREN
Oil Price Future 3 months ahead	NCLCS30
BBA Corporate Spread	LHIGBAA
Dollar-Euro Exchange Rate	USEURSP
Dollar-Sterlin Exchange Rate	USDOLLR
Commodity Price Index	CRBSPOT
Gold Price Index	GOLDHAR
Oil Future 3 months ahead	NCLCS30
Eurodollar Future 3 months ahead	NCLCS30
Cleveland Financial Stress Index	USCVFSI
CBOE VXO - Stock Volatility Index	CBOEVXO
Bid Cover Ratio in Trasuries Auctions (26 weeks)	USBCR26
Bank of America Merril Lynch Asset Backed Security Index	MLR0A2L
US Federal Funds Target Rate	USFDTRG
US Treasury Term Premia 1 years	USTTP1Y
US Treasury Term Premia 5 years	USTTP5Y
US Treasury Term Premia 10 years	USTTY10
Conventional Fixed Mortgage Rate	FRCMORT

Table B.5: Data description

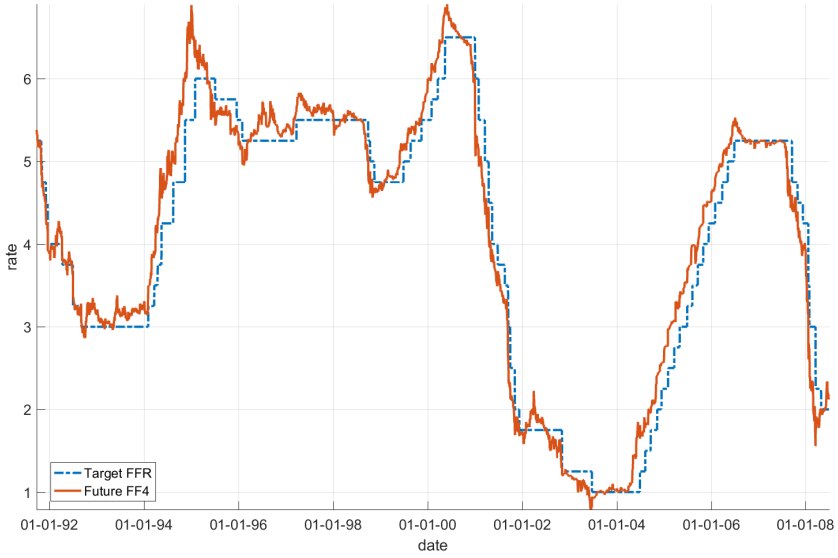

Figure B.16: Comparison TFFR and FF4
Comparison Target Fed Fund Rate - Fed Fund Rate Future 3 month ahead

B.4.1 Shocks identified from the Daily VAR

B.4.1.1 Baseline Identification

Table B.6-B.7 point out that, even without imposing any particular role for the FOMC meeting days, our conservative identification highlights a special role for these days. In fact, both mean and standard deviation of the shocks on FOMC meeting days are twice as sizable as the same statistics computed over the whole sample. Not surprisingly, this difference is more relevant for future contracts at shorter horizons. More formally, we also regress the size of the shocks over a dummy that reflect the FOMC meeting days, finding the same pattern (Table B.8).

Finally, we provide anecdotal evidence on the identified shocks. Specifically, the daily framework allows us to track the events that occurred on the days in which we register the most sizable shocks. Description and references are included in Table B.9.

Variable	Mean	Std. Dev.	Min.	Max.
|TFFR*|	0.444	0.838	0	15.136
|fut4* |	0.6	0.747	0	10.156
|fut1|	0.53	0.79	0	15.973
|fut4|	0.598	0.739	0	10.151
|fut7|	0.614	0.726	0	8.268
|fut18|	0.559	0.769	0	15.361
Observations		4352		

Table B.6: Descriptive statistics of monetary policy shocks - comparison across maturities

*Shocks in the whole sample - * refers to section 2.4.1; others show the robustness to using different future contracts (over a slightly shorter sample).*

Variable	Mean	Std. Dev.	Min.	Max.
|TFFR*|	2.832	3.422	0.015	15.136
|fut4* |	1.139	1.303	0.01	7.184
|fut1|	1.092	1.346	0.002	9.587
|fut4|	0.856	0.969	0.008	6.524
|fut7|	0.813	0.930	0.001	7.104
|fut18|	0.765	0.841	0.011	5.966
Observations		148		

Table B.7: Descriptive statistics of monetary policy shocks on FOMC meeting dates - comparison across maturities

*Shocks in the FOMC dates - * refers to section 2.4.1; others show the robustness to using different future contracts (over a slightly shorter sample).*

VARIABLES	(1) | TFFR* |	(2) | fut4* |	(3) | fut1 |	(4) | fut4 |	(5) | fut7 |	(6) | fut18 |
FOMC	2.47***	0.56***	0.58***	0.27***	0.21***	0.21***
	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)	(0.06)
Constant	0.36***	0.58***	0.51***	0.59***	0.61***	0.55***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	4,352	4,352	4,352	4,352	4,352	4,352
R-squared	0.29	0.02	0.02	0.00	0.00	0.00

Standard errors in parentheses
*** p<0.01

Table B.8: Regression of monetary policy shocks on FOMC meeting dates dummy - comparison across maturities

*Daily shocks regressed on FOMC days dummy - * refers to Section 2.4.1; others show the robustness to using different future contracts (over a slightly shorter sample).*

Bridge TFFR	**Event 1**	**Event 2**	**Event 3**
Dates	18 March 2008	22 January 2008	15 November 1994
Description	FOMC meeting	FOMC meeting	FOMC meeting
Reference	Event 1a; Event 1b	Event 2a; Event 2b	Event 3
Shock	$-15\,std$	$-13.1\,std$	$15.9\,std$
Bridge FF4	**Event 1**	**Event 2**	**Event 3**
Dates	02 January 2001	22 January 2008	02 January 1995
Description	Anticipation FOMC 03 Jan 2001	FOMC meeting	\$50 billion bailout Mexican *tequila crisis*
Reference	Event 1	Event 2a Event 2b	Event 3
Shock	$-7.9\,std$	$-7.5\,std$	$10.5\,std$

Table B.9: Largest monetary policy shocks

Main shocks (reported in standard deviation units) identified in our daily VAR and corresponding events - section 2.4.1

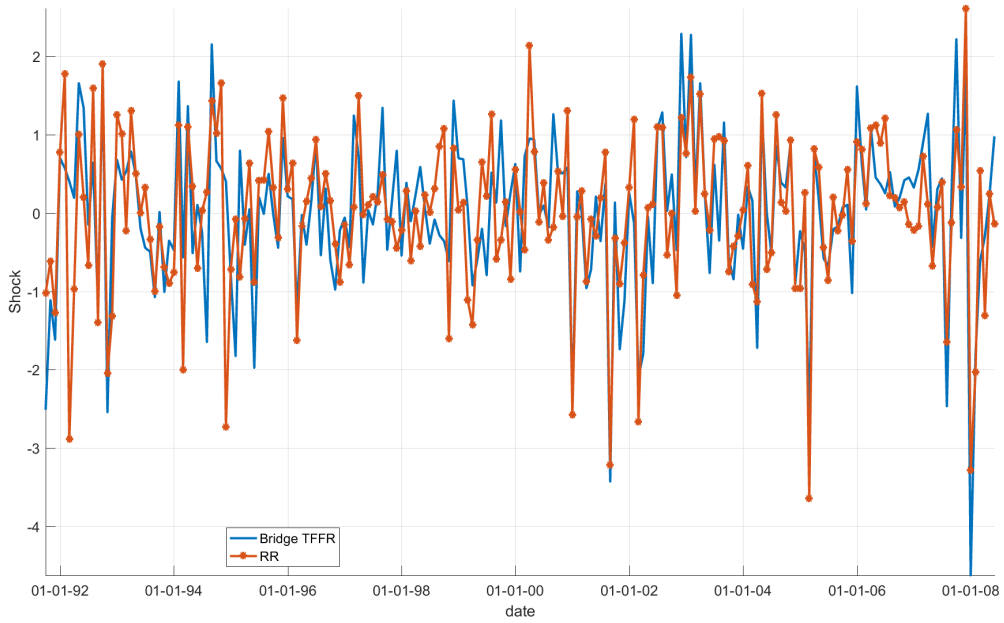

Figure B.17: Comparison of TFFR shocks with Romer and Romer shocks

Comparison of monetary policy shocks from different identifications. Bridge TFFR (blue) refers to the series of shocks identified using our daily VAR. RR refers to the series of shocks build as Romer and Romer (2004), extended by Coibion et al. (2012).

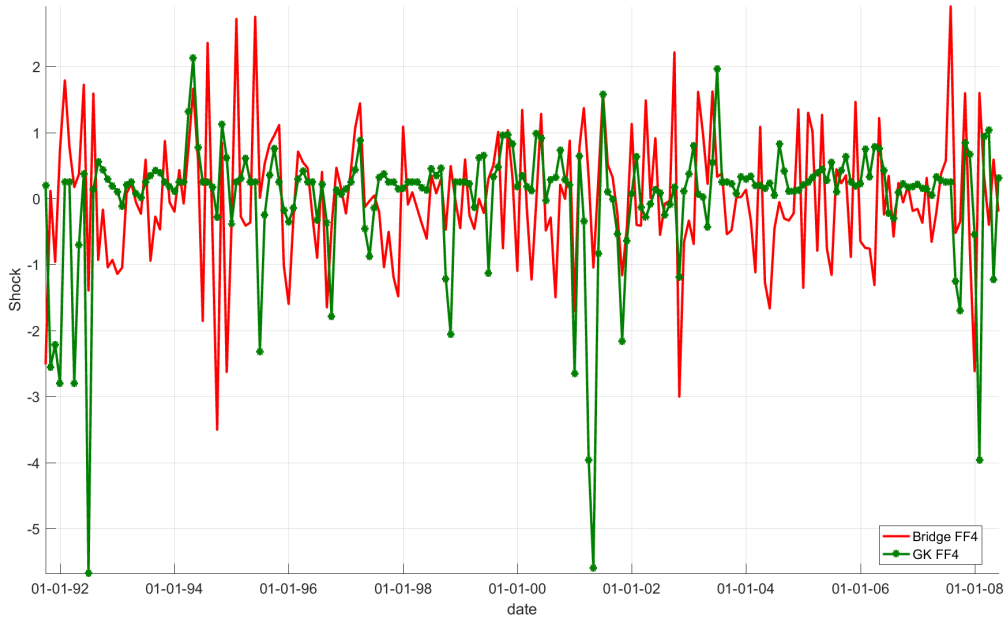

Figure B.18: Comparison of FF4 shocks with Gerter and Kararadi shocks

Comparison of monetary policy shocks from different identifications. Bridge FF4 (red) refers to the series of shocks identified using our daily VAR. GKFF4 refers to the series of shocks employed by Gertler and Karadi (2015).

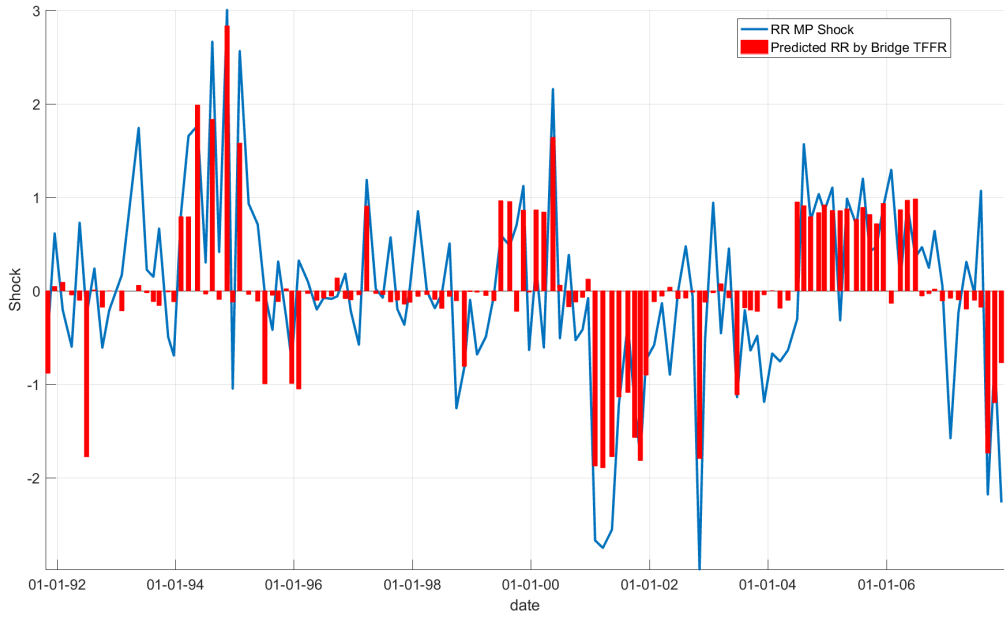

Figure B.19: Explanatory power of TFFR shocks for Romer and Romer shocks
Romer and Romer (2004) shocks, extended by Coibion et al. (2012), fitted by your TFFR series of shocks estimated in a daily VAR.

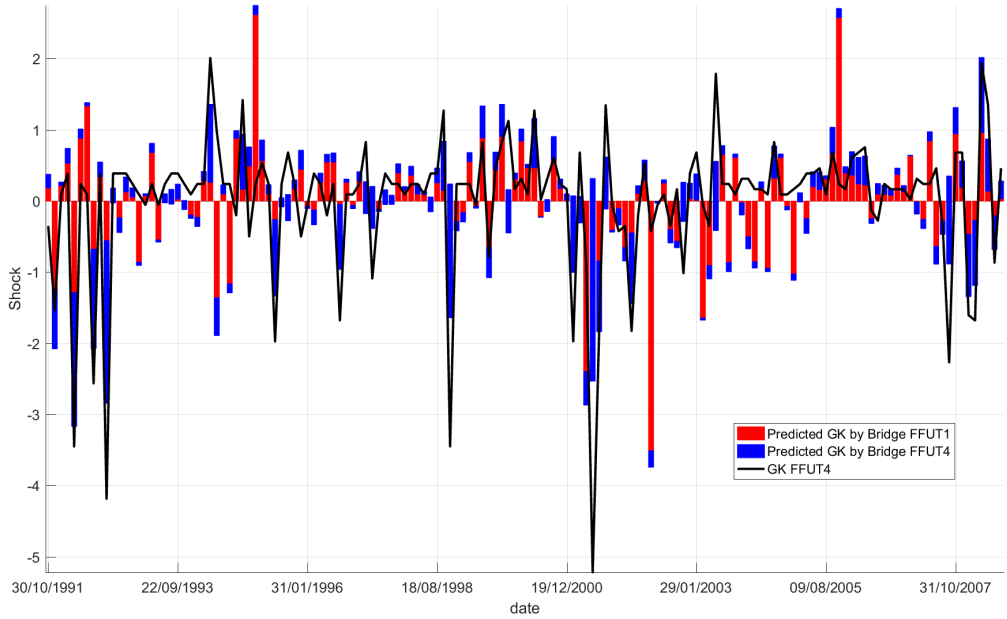

Figure B.20: Explanatory power of TFFR and FF4 shocks for Romer and Romer shocks
Gertler and Karadi (2015) FF4 shocks fitted by your TFFR and FF4 shocks estimated in a daily VAR.

B.4.1.2 Alternative Identifications

Our two alternative identification strategies yield series of daily monetary policy shocks that are very correlated with our baseline series. Moreover, they generate

very similar macroeconomic effects. In Tables B.10-B.11 we report the correlations among the shocks identified with all the strategies that we have employed.

Identification Via Heteroskedasticity

In short, the identification proposed by Rigobon (2003) exploits the change in the volatility of the structural shocks across (at least) two regimes. Consistently with our finding reported in Table B.6-B.7, we assume that the variance of the monetary policy shocks changes across FOMC meeting days and non-FOMC meeting days. We estimate a bivariate VAR including FF4 and SP&500 and exploit the change in the variance of the shocks in FF4 across the two regimes for identification. In this way, we obtain a series of shocks that correlates 0.9998 with the shocks identified by ordering the TFFR last in our large scale VAR. The same result hold in three and four variable daily VARs, which additionally include the commodity price index and commodity price index plus the Cleveland Financial Stress index. Finally, notice that event-based identification is equivalent to the identification via heteroskedasticity where the change in the volatility across the two regimes is assumed to be infinite.

Identification Via Independent Component Analysis

Detailed reference on the application of Independent Component Analysis (ICA) to VARs can be found in Capasso and Moneta (2016) and Gourieroux et al. (2017). Intuitively, ICA can be seen as a generalization of principal component analysis (PCA). While PCA looks for uncorrelated latent components, ICA minimizes the statistical independence among such components. Obviously, if the data is normally distributed, the two concept are equivalent. However, when departing from gaussianity, ICA can solve the identification problem in VARs. While the reduced form residuals can be decomposed in uncorrelated structural shocks in infinite ways, ICA searches for the (unique) combination of the most statistically independent components.

Both visual inspection and the Kolmorogov-Smirnov reject the normality of the 18 reduced form residuals in our daily VAR. We do not assume any particular distribution of the reduce form residuals but we estimate semi-parametrically the independent components.[9B] We consider as monetary policy shock the structural

[9B] We employ the algorithm *Icasso* v1.22 and *FastICA* v2.5.

shock that contributes the most to the variance of the FF4 on impact. The resulting series of structural shocks correlates 0.89 with the shocks in the TFFR and 0.9 with the shocks in the FF4 identified with our baseline recursive ordering.

	Target FFR - Last	FF4 - Last	FF4 - Heteroschedasticity	FF4 - ICA
Target FFR - Last	1	*	*	*
FF4 - Last	0	1	*	*
FF4 - Heteroskedasticity	1*	0	1	*
FF4 - ICA	0	0.92*	0	1

Table B.10: Correlation among monetary policy shocks across different identifications - daily frequency

Correlations among monetary policy shocks recovered at the daily frequency through different identification strategies: 1) Target FFR ordered last in recursive identification; 2) Fed Future (3 months ahead) ordered last in recursive identification; 3) Fed Future (3 months ahead) exploiting the change volatility in FOMC meeting days and other days (heteroskedasticity); 4) Fed Future (3 months ahead) exploiting the non-normality of the reduced form residuals (Independent Component Analysis - ICA) . All coefficients different from 0 are statistically significant at the 1% level.

	Target FFR - Last	FF4 - Last	FF4 - Heteroschedasticity	FF4 - ICA
Target FFR - Last	1	*	*	*
FF4 - Last	0.1	1	*	*
FF4 - Heteroskedasticity	1 *	0.11	1	*
FF4 - ICA	0.1	0.93*	0.11	1

Table B.11: Correlation among monetary policy shocks across different identifications - monthly frequency

*Correlations among monetary policy shocks recovered at the daily frequency through different identification strategies and aggregated at the monthly frequency: 1) Target FFR ordered last in recursive identification; 2) Fed Future (3 months ahead) ordered last in recursive identification; 3) Fed Future (3 months ahead) exploiting the change volatility in FOMC meeting days and other days (heteroskedasticity); 4) Fed Future (3 months ahead) exploiting the non-normality of the reduced form residuals (Independent Component Analysis - ICA) . * denotes statistical significance at the 1% level.*

B.4.1.3 Impulse Response Functions

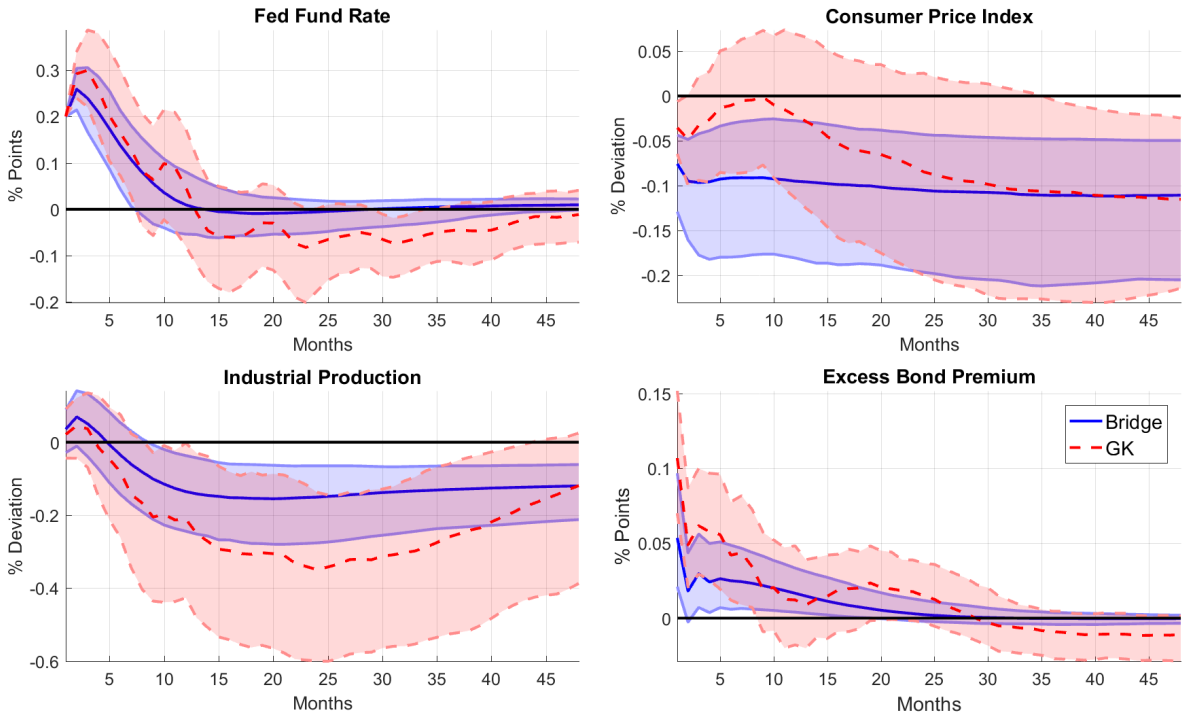

Figure B.21: IRFs FF4

IRFs to a monetary policy shock identified using Bridge Future using all the available days (FOMC and non-FOMC). From the first stage, $F - stat = 7.7$. The VAR is estimated in log-levels with the optimal number of lags (2) and includes a deterministic constant. Shaded areas correspond to 95% bootstrapped confidence bands.

IRFs in the Medium System of Gertler and Karadi

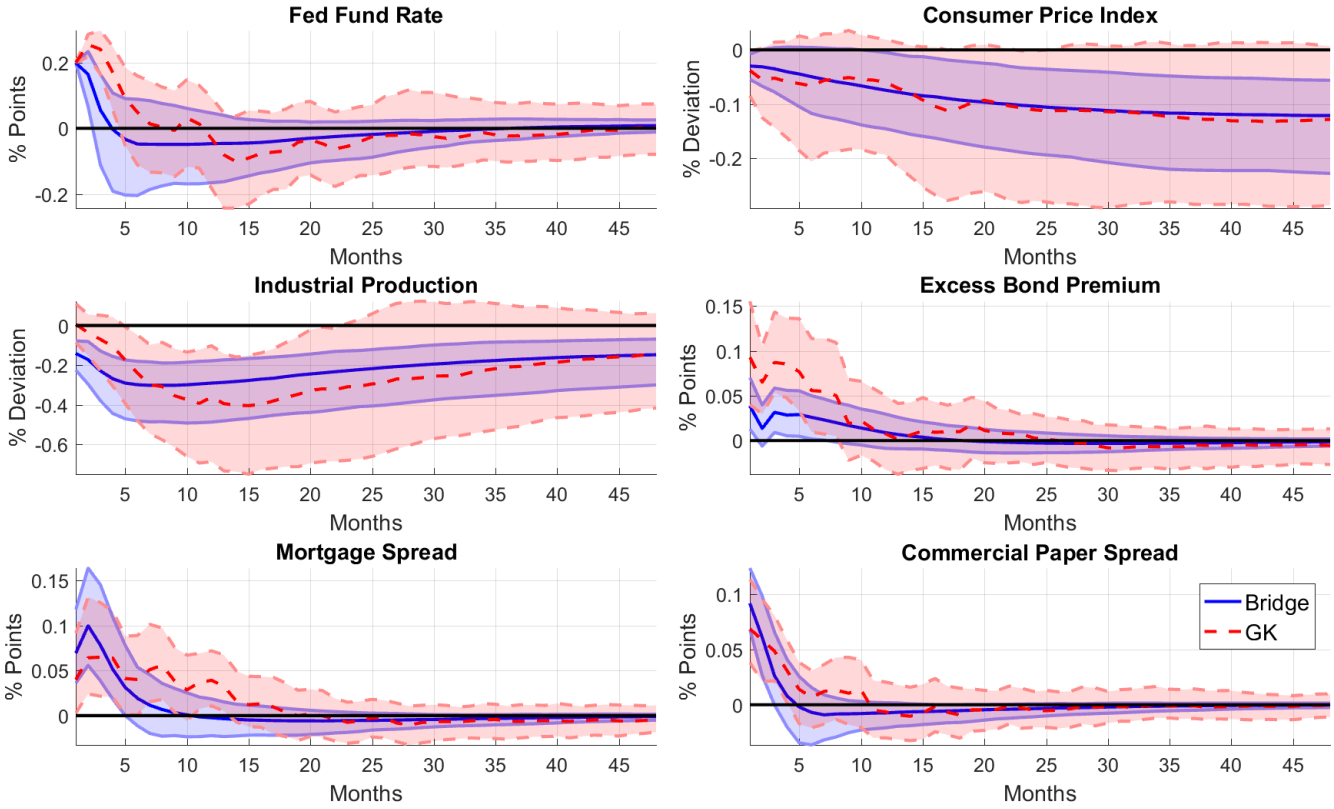

Figure B.22: IRFs TFFR - medium system

IRFs to a monetary policy shock identified using Bridge Target. From the first stage, $F-stat = 10.2$. The VAR includes [FFR, CPI, Industrial Production, Excess Bond Premium, Mortgage Spread, Commercial Paper Spread] and it is estimated in log-levels with the optimal number of lags (2) and includes a deterministic constant.Shaded areas correspond to 95% bootstrapped confidence bands from 1000 replications.

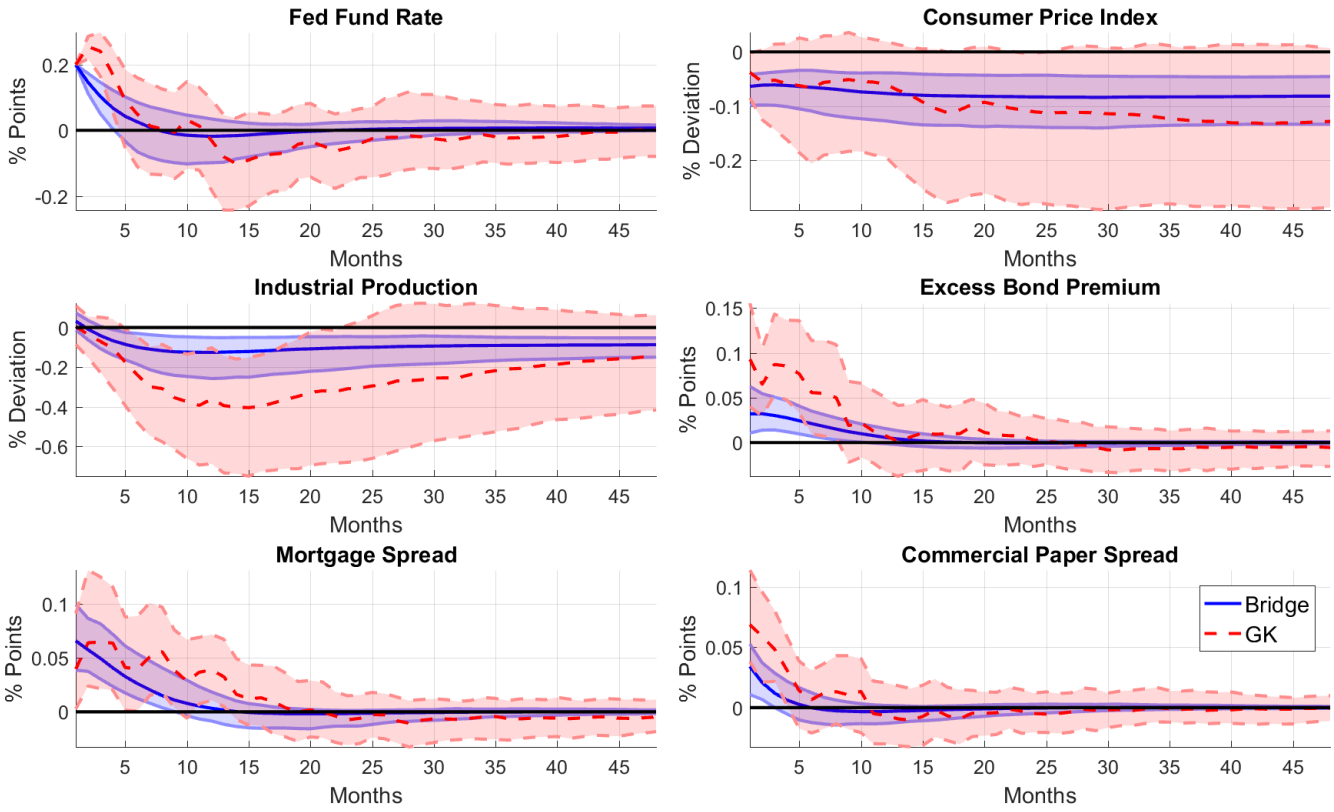

Figure B.23: IRFs FF4 - medium system

IRFs to a monetary policy shock identified using Bridge Future (FOMC and non-FOMC). From the first stage, $F-stat = 7.44$. The VAR includes [FFR, CPI, Industrial Production, Excess Bond Premium, Mortgage Spread, Commercial Paper Spread] and it is estimated in log-levels with the optimal number of lags (2) and includes a deterministic constant. Shaded areas correspond to 95% bootstrapped confidence bands.

Appendix C

Appendix: Liquidity Shocks

	Italy	Spain
Unemployment	ISTAT	Ministry of Economy
Industrial Production	ISTAT	INE
CPI Inflation	ISTAT	INE
Central Government Debt	Bank of Italy	Ministry of Economy
ECB Repo	ECB	ECB
M2	Bank of Italy	Banco de España
Consumer Confidence	ISTAT	Ministry of Economy
Business Confidence	ISTAT	Ministry of Industry
Volatility Index	ASR-Absolute Strategy	VSTOXX
CDS	Thomson Reuters CDS	Thomson Reuters CDS
Bid-Ask Spread	Bloomberg	Bloomberg
Yield Spread	ECB	ECB
Stock Prices	FTSE MIB	IBEX 35

	France	Germany
Unemployment	INSEE	OECD
Industrial Production	INSEE	Federal Statistical Office
CPI Inflation	Thomson Reuters	Thomson Reuters
Central Government Debt	Banque de France	Deutsche Bundesbank
ECB Repo	ECB	ECB
M2	Banque de France	Deutsche Bundesbank
Consumer Confidence	DG ECFIN	DG ECFIN
Business Confidence	DG ECFIN	DG ECFIN
Volatility Index	Euronext Paris	Deutsche Boerse
CDS	Thomson Reuters CDS	Thomson Reuters CDS
Bid-Ask Spread	Bloomberg	Bloomberg
Yield Spread	ECB	ECB
Stock Prices	CAC 40	MDAX Frankfurt

Table C.1: Data Sources

All the variables are seasonally adjusted originally or by using the X-13ARIMA procedure. We deflate nominal variables by the corresponding CPI price index in order to estimate the VAR with real variables.

In Section 3.4.2, we refer to the following questions from the Bank and Lending Survey:

1. *Firm* Δ *Standards:* Changes in bank's credit standards for approving loans or credit lines to enterprises, Overall (all firms and types of loans), Past three months.

2. *Firm: Costs-Asset Position:* Changes in the contribution of cost of funds and balance sheet constraints (costs related to bank's capital position) affecting credit standards for approving loans or credit lines to enterprises.

3. *Firm: Liquidity Position:* Changes in the contribution of cost of funds and balance sheet constraints (bank's liquidity position) affecting credit standards for approving loans or credit lines to enterprises.

4. *Firm: Risk-Economic Activity:* Changes in the contribution of perception of risk about general economic situation and outlook affecting credit standards for approving loans or credit lines to enterprises.

5. *Mortgages:* Δ *Standards:* Changes in credit standards for approving loans to households, loans for house purchase in the last three months.

6. *Mortgages: Costs-Funding:* Changes in the contribution of the following factors affecting credit standards for approving loans to households for house purchase, cost of funds and balance sheet constraints.

Concerning the ISTAT survey, the questionnaire can be found at ISTAT questionnaire (only in Italian). We refer to the following questions/answers:

43 Today, in our opinion, are the credit conditions more or less favorable compared to three months ago? (Possible answers: *More; Constant; Less*)

45 Have you obtained the loan you requested to the bank or financial institution? (Possible answers: *Yes, at the same conditions; Yes, at worse conditions; No; Only asking information*)

46 In case answer to 43 was *No* - Has the bank reject your request or you have not accepted their offer due to the conditions they were setting? (Possible answers: *The bank has not offered a loan; We have not accepted the loan due to not favorable conditions*)

47 In case answer to 45 was *Yes, at worse conditions* - Why the conditions have become worse? (Possible answers: *Higher rate; More personal collateral requested; More real collateral requested; Limits on the amount of the loan; Additional costs*)

C.1 High Frequency Variables

Date	Events
2/7/07	HSBC issue with subprimes
6/7/07	Bearn Sterns first bad news
8/9/07	BNP Paribas
9/13/07	Northern Rock
2/18/08	Northern Rock Nationalized
3/14/08	Bearn Sterns bought by JP Morgan
9/15/08	Lehman
10/16/08	Greek Deficit Surprise
5/7/10	EFSF
7/23/10	Stress Test
10/28/10	ESM
5/17/11	Portugal asks help
8/5/11	Letter to Mr. Berlusconi from ECB
8/16/11	ECB buys after Ita take measures
10/4/11	Downgrade ITA-SPAIN
10/11/11	CDS-ban announced
10/31/11	Draghi takes over
11/1/11	CDS-ban in place
11/14/11	Mr. Monti takes over
12/5/11	Mr. Monti package
12/8/11	LTRO announced
12/21/11	1st LRTO
2/28/12	LTRO announced
6/26/12	Cyprus requests aid
7/26/12	Mr. Draghi whatever it takes
8/2/12	OMT announced
12/10/12	Monti resigns
12/13/12	SSM announced
11/7/13	ECB cuts Rate

Table C.2: List of European and Italian events

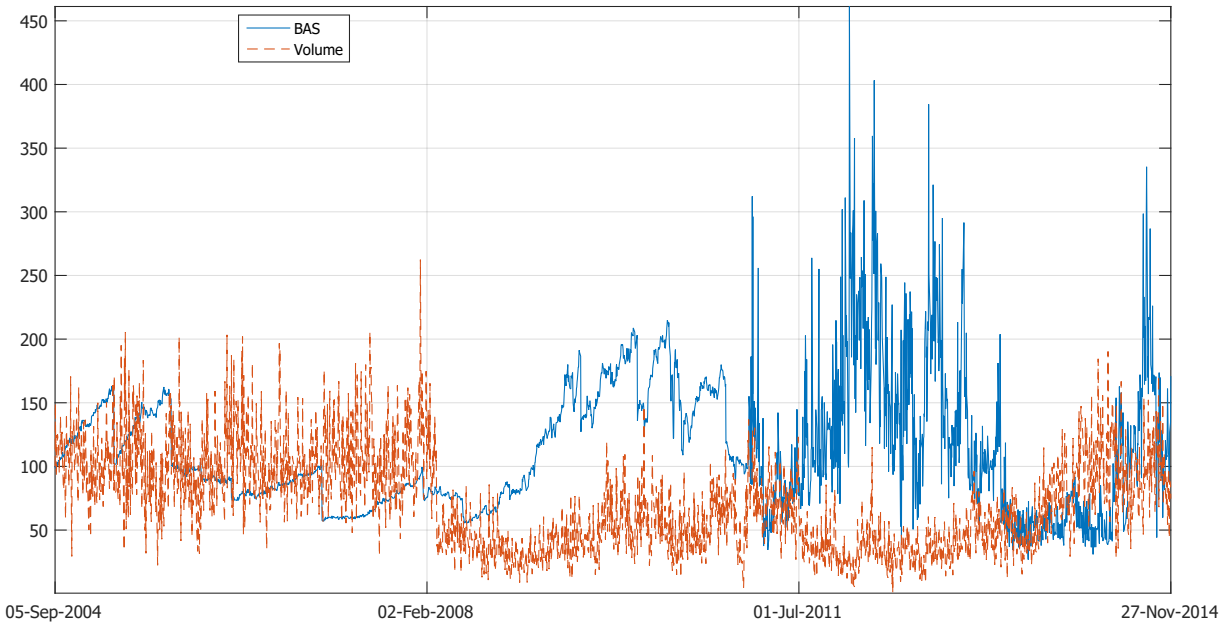

Figure C.1: Italian BAS and Turnover on the MTS platform

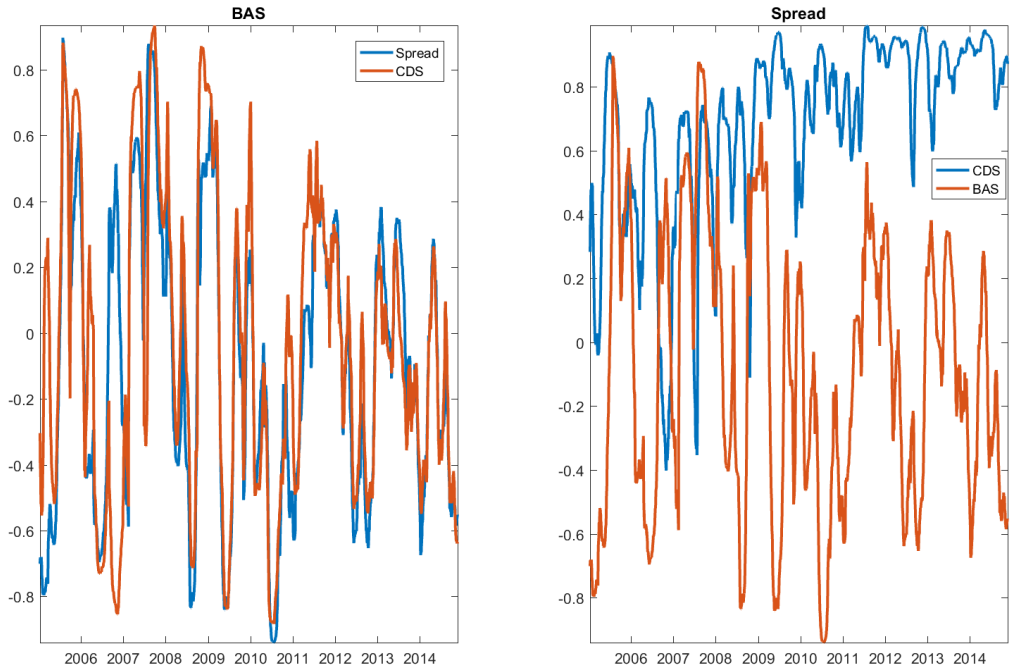

Figure C.2: Dyanmic correlations among Spread, CDS and BAS over 2004-2014. Correlations are computed over a 90 days rolling window

C.2 Financial Variables at Monthly Frequency

Table A3 summarizes statistics of the financial variables used in the empirical analysis at monthly frequency:

	Full Sample			2009-2014		
	BAS	Yield	CDS	BAS	Yield	CDS
Mean	0.017	4.318	98.278	0.020	4.41	169.58
Max	0.037	7.057	546.159	0.037	7.057	546.159
Min	0.007	1.990	2.343	0.007	1.990	36.352
St. Dev.	0.007	0.809	124.411	0.007	1.008	128.619
Auto Corr.	0.836	0.956	0.964	0.782	0.957	0.940

Table C.3: Descriptive statistics of sovereign debt financial variables at monthly frequency. Sources: Bloomberg, Datastream and Bank of Italy. Maturities: BAS and CDS 2 years; Yield 10 years.

There is no significant change in volatility and standard deviation in the period of the sovereign debt crisis at monthly frequency.

C.3 Proxy-SVAR

We describe the the Proxy SVAR methodology that we use to identify the effects of BAS shocks and the first stage results (i.e. the linear projection of the reduced form residuals on the exogenous variations of BAS identified at daily frequency).

C.3.1 Theoretical Reference

Consider the following VAR:

$$Y_t = AY_{t-1} + u_t \tag{C.1}$$

with Y_t a vector of endogenous variables and u_t is a vector of reduced form residuals with variance-covariance matrix Σ_u. The objective is to recover the structural form of the VAR, characterized by the vector of structural shocks $\varepsilon_t = B^{-1}u_t$:

$$Y_t = AY_{t-1} + B\varepsilon_t \tag{C.2}$$

We can rewrite the VAR system as partitioned (or bivariate for a matter of interpretation):

$$\begin{bmatrix} Bas_t \\ X_t \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} Bas_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{bas} \\ \varepsilon_t^{X} \end{bmatrix} \tag{C.3}$$

The Proxy-SVAR is an identification strategy that (potentially) partially identifies the unknown B matrix. Namely, we aim at identifying only the block $\begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}$, which would allows us to compute the IRFs of the system to a structural innovation in the BAS. In order to reach the identification, we exploit information from outside the VAR system. We use the variable z_t as a proxy for the true structural shock ε_t^{bas}. z_t is assumed to be a proxy for (a component of) the true ε_t^{bas} with the following (instrumental variable) properties:

$$\begin{aligned} \mathbb{E}\left[\varepsilon_t^{bas} z_t\right] &\neq 0 \\ \mathbb{E}\left[\varepsilon_t^{X} z_t\right] &= 0 \end{aligned}$$

In fact, under those assumptions, we can obtain consistent estimates of $\begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}$ by taking an instrumental variable approach:

First Stage: regress $u_t^{bas} = \beta z_t + \xi_t$ obtaining $\hat{u}_t^{bas}$

Second Stage: $u_t^X = \frac{B_{21}}{B_{11}} \hat{u}_t^{bas} + \zeta_t$

Given that the BAS reacts one to one to its own structural shock (on impact), we can normalize $\frac{B_{21}}{B_{11}} = B_{21}$. The IRFs to a BAS shock can be then computed across different horizons as:

$$\mathcal{IRF}_0^X = B_{21}$$

$$\mathcal{IRF}_n^X = A^{n-1} \mathcal{IRF}_{n-1}^X \quad \forall n > 0$$

C.3.2 First Stage

Figure A4 displays the RF residuals predicted by the proxy, compared to the original RF innovation series.

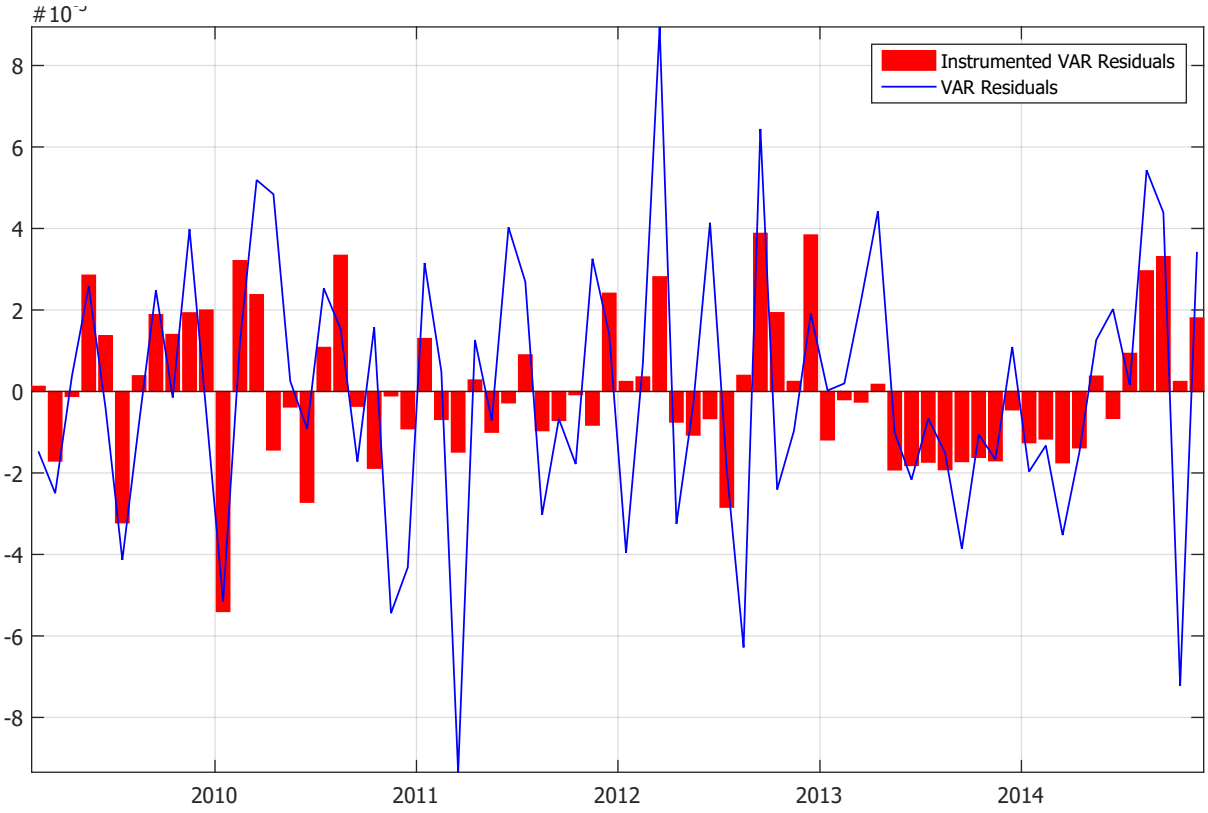

Figure C.3: First stage result of the Bridge Proxy-SVAR identification
The blue line represents the RF residuals of the BAS from the VAR featuring [Unemployment, π, Public Debt, R, M2, CC, BC, Financial Block]; the red bar is the RF residuals predicted by the Proxy (BAS shocks identified in a daily VAR including $[BAS, CDS, Yield, FTSE, Eonia, VIX]$)

C.4 Alternative VAR Specifications

We present the results from alternative VAR specifications described in Section 3.3.4. To keep the appendix short, we only report results using some particular identification schemes (Basic, Full or Proxy SVAR). Results are robust using the other identification schemes and are available from the authors upon request.

C.4.1 Indicator of Liquidity

The following figures report the IRFs to a BAS shock of the Full VAR and Proxy-SVAR specifications including the Turnover instead of the Equity Premium, respectively. Moreover, we also display the IRFs and the FEVD of Unemployment from the Full VAR including the Liquidity Index instead of the BAS. An increase (decrease) in the Liquidity Index is analogous to a decrease (increase) in the BAS.

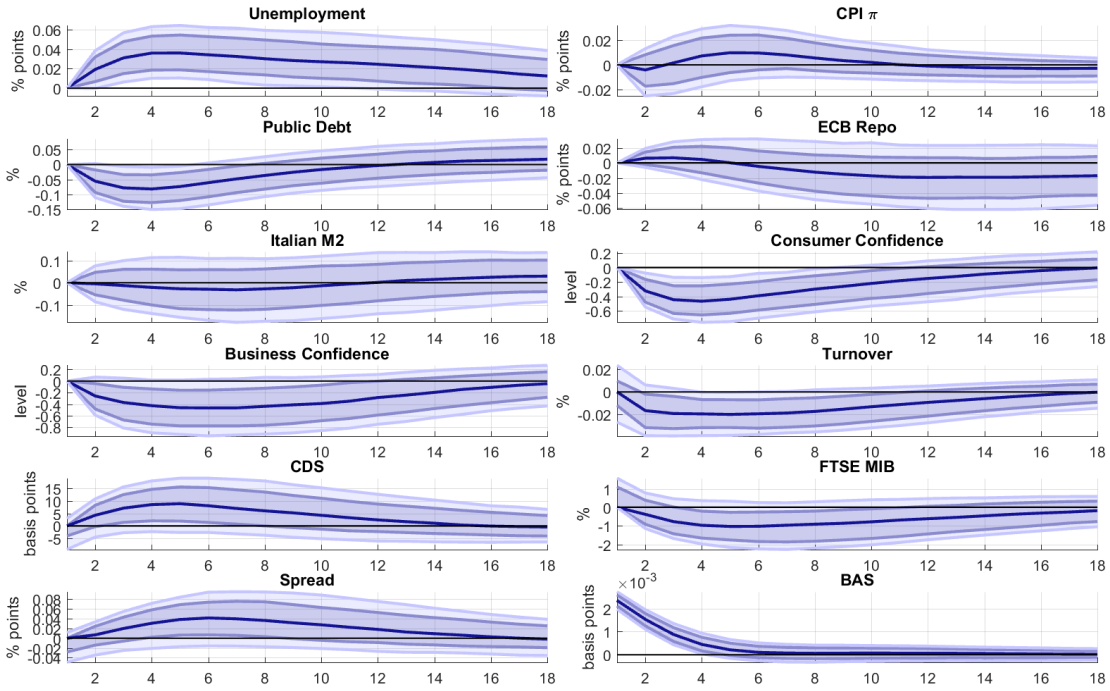

Figure C.4: IRFs to a BAS Shock - Choleski identification

IRFs to a 1 std BAS shock identified through the following ordering [Unemployment, π, Public Debt, R, M2, CC, BC, Financial Block]. The turnover of Italian sovereign bonds is included in place of the equity premium. The median point estimate, 68% and 90% confidence bands are reported in cyan, blue, and light blue, respectively. 50%, 68% and 90% bands include statistical and identification uncertainty (from all the possible ordering within the financial block)

.

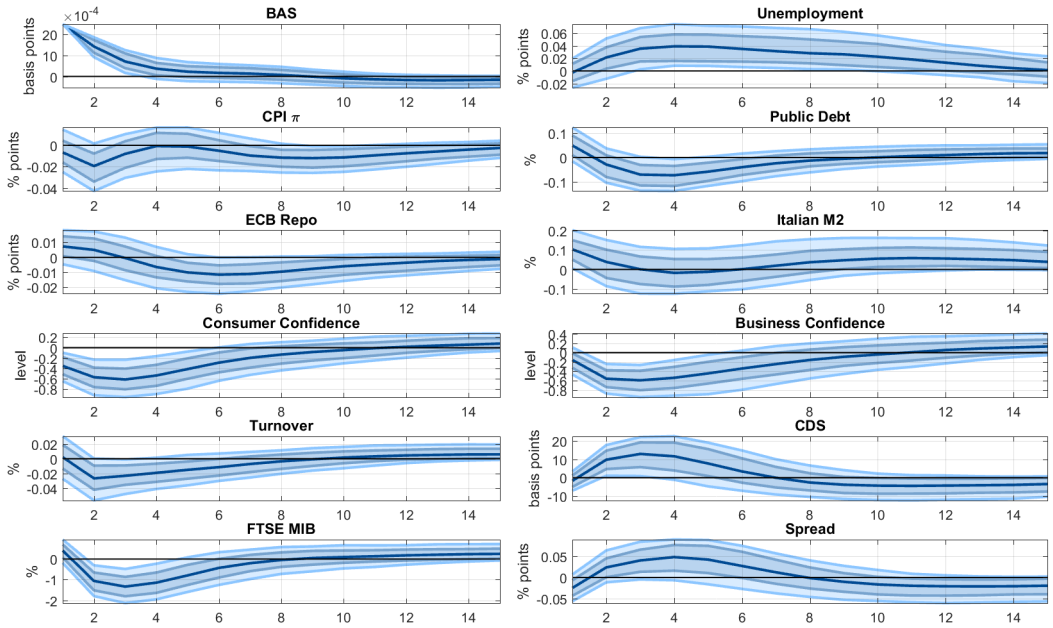

Figure C.5: IRFs to a BAS Shock - Bridge Proxy-SVAR identification

IRFs to a 1 standard deviation BAS shock in the VAR [IP, π, Public Debt, R, M2, CC, BC, Financial Block]. The turnover of Italian sovereign bonds is included in place of the equity premium. The shock is identified through the unpredictable variation of the BAS in a daily VAR system. Sample: Jan:2009-Nov:2014. The median point estimate, 68% and 90% confidence bands are reported in blue and light blue, respectively. Confidence bands are computed using wild

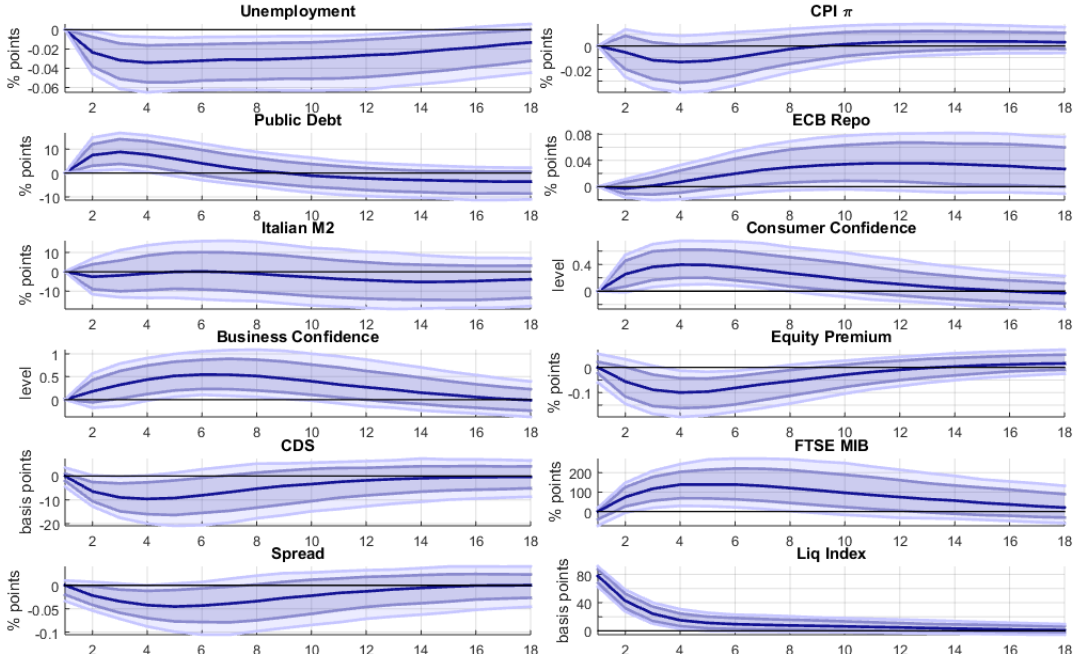

Figure C.6: IRFs to a Liquidity Index shock - Choleski identification

IRFs to a 1 std Liquidity Index shock (liquidity improvement) identified through the following ordering [Unemployment, π, Public Debt, R, M2, CC, BC, Financial Block]. The median point estimate, 68% and 90% confidence bands are reported in cyan, blue, and light blue, respectively. 50%, 68% and 90% bands include statistical and identification uncertainty (from all the possible ordering within the financial block).

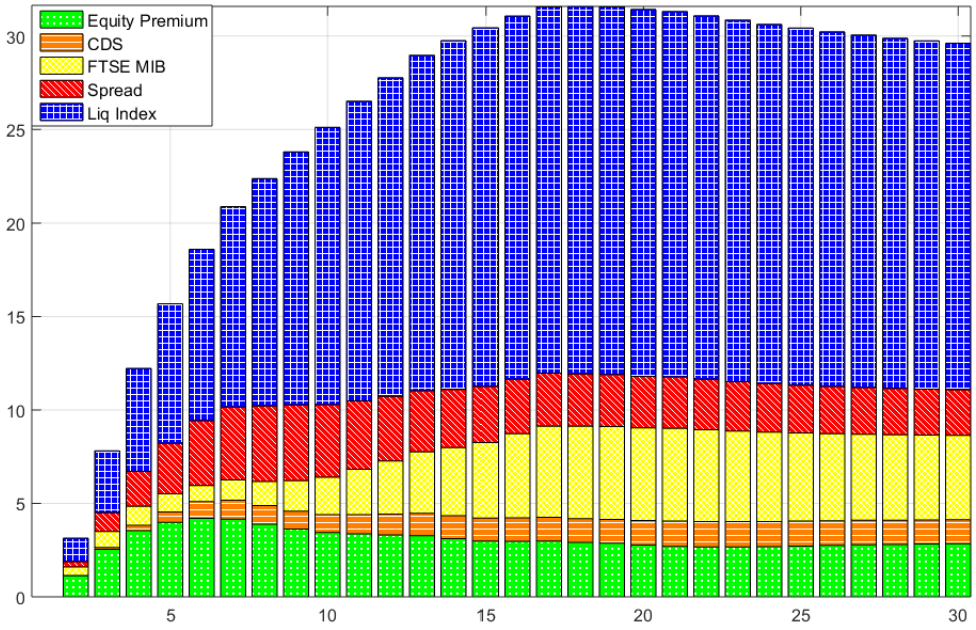

Figure C.7: FEVD of unemployment - Choleski identification
FEVD of unemployment including the Liquidity Index identified through the following ordering [Unemployment, π, Public Debt, R, M2, CC, BC, Financial Block].

Liquidity accounts for around 20% of Unemployment fluctuations in the period under analysis, in line with results presented in Section 3.3.2.

C.4.2 Measures of Economic Activity

In this case, we use alternative measures of economic activity and present the corresponding IRFs. We include results both with our small VAR system and with the Proxy-SVAR. We employ Industrial Production and the ITA-Coin.

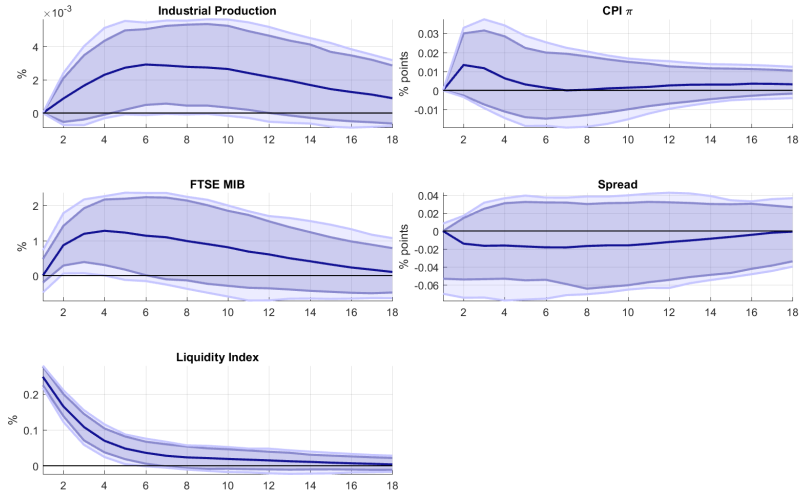

Table C.4: IRFs to a Liquidity Index shock - Choleski identification and industrial production

IRFs to a 1 std Liquidity Index shock (liquidity improvement) identified through the following ordering [Industrial Production, π, FTSE, Spread, BAS]. The median point estimate, 68% and 90% confidence bands are reported in cyan, blue, and light blue, respectively. 50%, 68% and 90% bands include statistical and identification uncertainty (from all the possible ordering within the financial block).

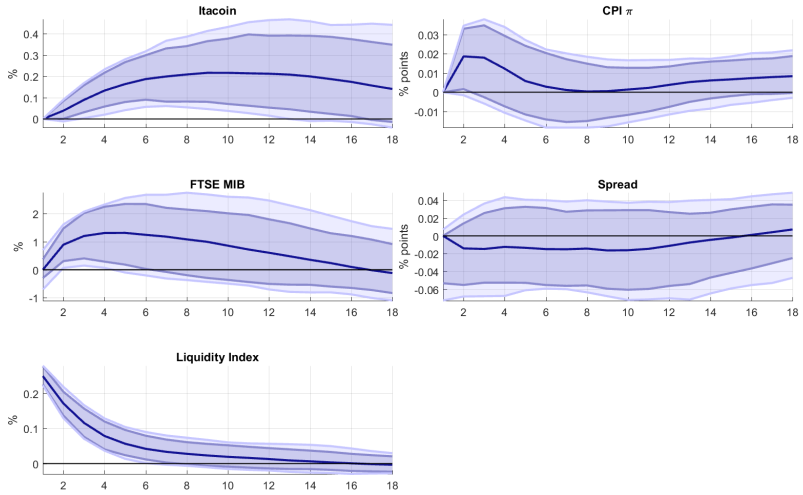

Figure C.8: IRFs to a Liquidity Index shock - Choleski identification; industrial production

IRFs to a 1 std Liquidity Index shock (liquidity improvement) identified through the following ordering [Itacoin, π, FTSE, Spread, BAS]. The median point estimate, 68% and 90% confidence bands are reported in cyan, blue, and light blue, respectively. 50%, 68% and 90% bands include statistical and identification uncertainty (from all the possible ordering within the financial block).

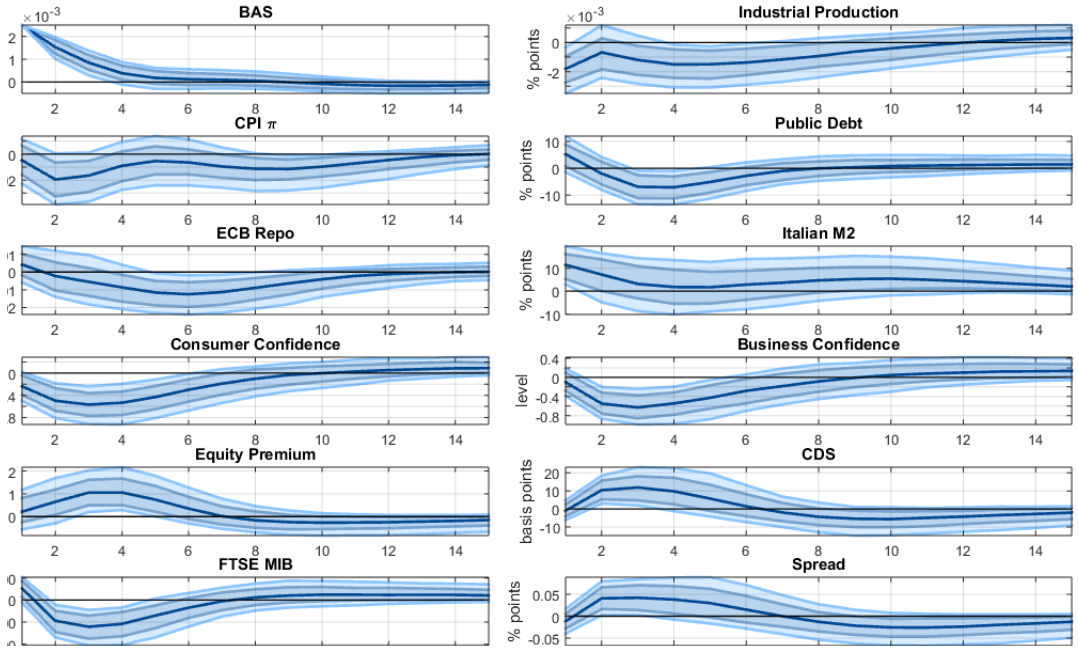

Figure C.9: IRFs to a BAS shock - Bridge Proxy-SVAR identification; industrial production

IRFs to a 1 standard deviation BAS shock (liquidity deterioration) in the VAR [IP, π, Public Debt, R, M2, CC, BC, Financial Block]. The shock is identified through the unpredictable variation of the BAS in a daily VAR system. Sample: Feb:2004-Nov:2014. The median point estimate, 68% and 90% confidence bands are reported in blue and light blue, respectively. Confidence bands are computed using wild bootstrap with 1,000 replications.

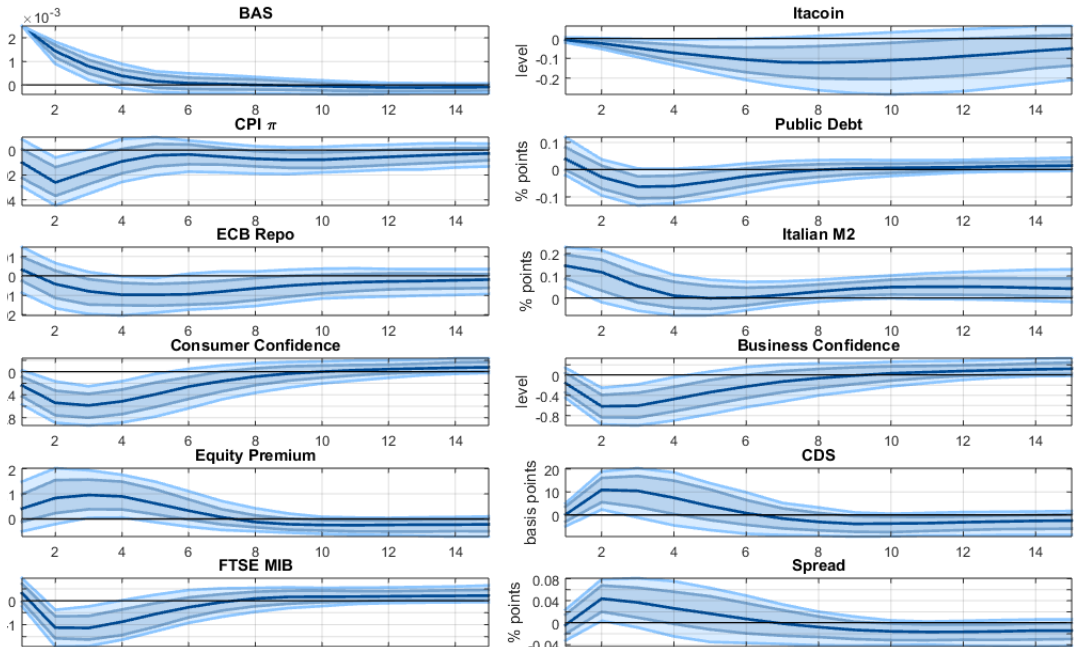

Figure C.10: IRFs to a BAS shock - Bridge Proxy-SVAR identification; Itacoin

IRFs to a 1 standard deviation BAS shock (liquidity deterioration) in the VAR [IP, π, Public Debt, R, M2, CC, BC, Financial Block]. The shock is identified through the unpredictable variation of the BAS in a daily VAR system. Sample: Feb:2009-Nov:2014. The median point estimate, 68% and 90% confidence bands are reported in blue and light blue, respectively. Confidence bands are computed using wild bootstrap with 1,000 replications.

C.4.3 Alternative Samples

We study the dependence of our findings on the sample used. We display the IRFs to a BAS shock and FEV of Unemployment using the sample January 2009-November 2014 and on the pre-crisis sample (February 2004-December 2008). The main conclusions remain unchanged.

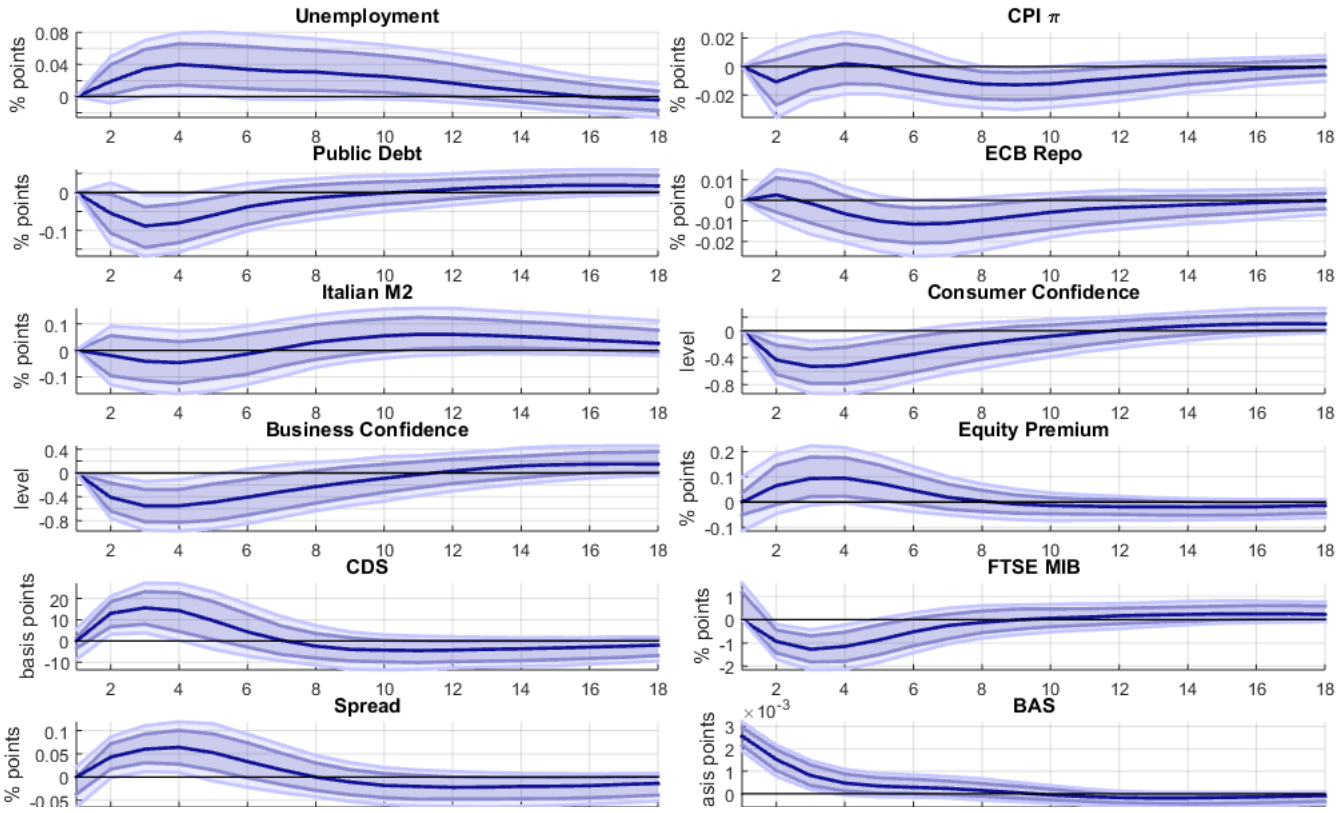

Figure C.11: IRFs to a BAS shock - Choleski; sample 2009-2014

IRFs to a 1 std BAS shock identified through the following ordering [Unemployment, π, Public Debt, R, M2, CC, BC, Financial Block]. The median point estimate, 68% and 90% confidence bands are reported in cyan, blue, and light blue, respectively. 50%, 68% and 90% bands include statistical and identification uncertainty (from all the possible ordering within the financial block).

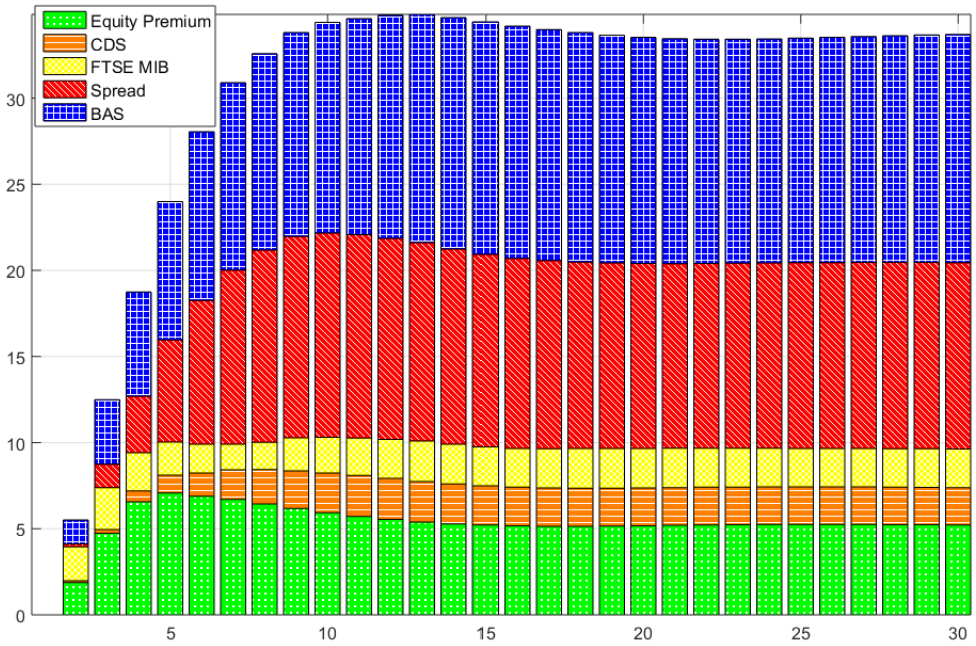

Figure C.12: IRFs to a BAS shock - Choleski; sample 2009-2014
FEVD of unemployment including the Liquidity Index identified through the following ordering [Unemployment, π, Public Debt, R, M2, CC, BC, Financial Block].

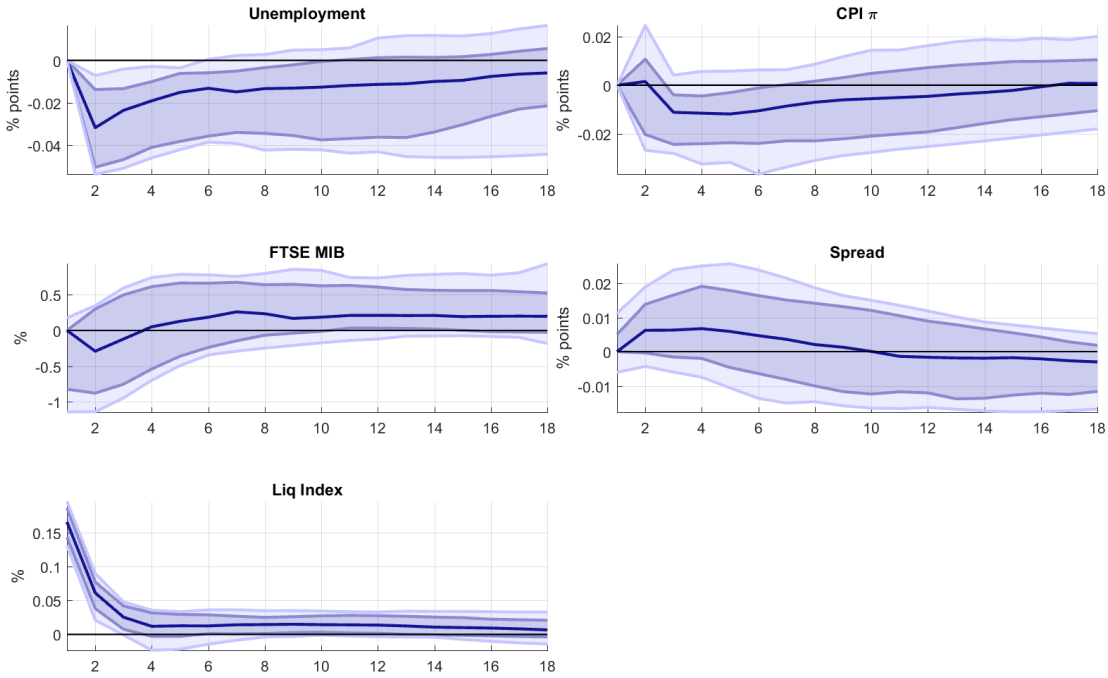

Figure C.13: IRFs to a BAS shock - Choleski; sample 2004-2008

IRFs to a 1 std Liquidity Index shock (liquidity improvement) identified through the following ordering [Unemployment, π, FTSE, Spread, BAS]. The median point estimate, 68% and 90% confidence bands are reported in cyan, blue, and light blue, respectively. 50%, 68% and 90% bands include statistical and identification uncertainty (from all the possible ordering within the financial block).

C.4.4 Corporate Liquidity

In this section, we consider the relationship between the Corporate and Sovereign liquidity. Figure A16 displays the evolution of the Corporate BAS together with sovereign variables aggregated at monthly frequency. Figure A17 displays the IRF to a shock to corporate BAS and compares it to the one to a sovereign BAS. Finally, Figure A18 shows the IRFs using as a variable the spread between Corporate and Sovereign BAS instead of the BAS.

Levels	BAS-S	Spread	CDS	BAS-C
BAS-S	1	-0.08	0.39*	0.31*
Spread	-0.08	1	0.35	0.5*
CDS	0.39*	0.35	1	0.9*
BAS-C	0.31*	0.5*	0.9*	1

Table C.5: Sovereign and Corporate Liquidity
Correlation over the 2004-2014 among Sovereign and Corporate BAS, Spread and CDS (as monthly averages).

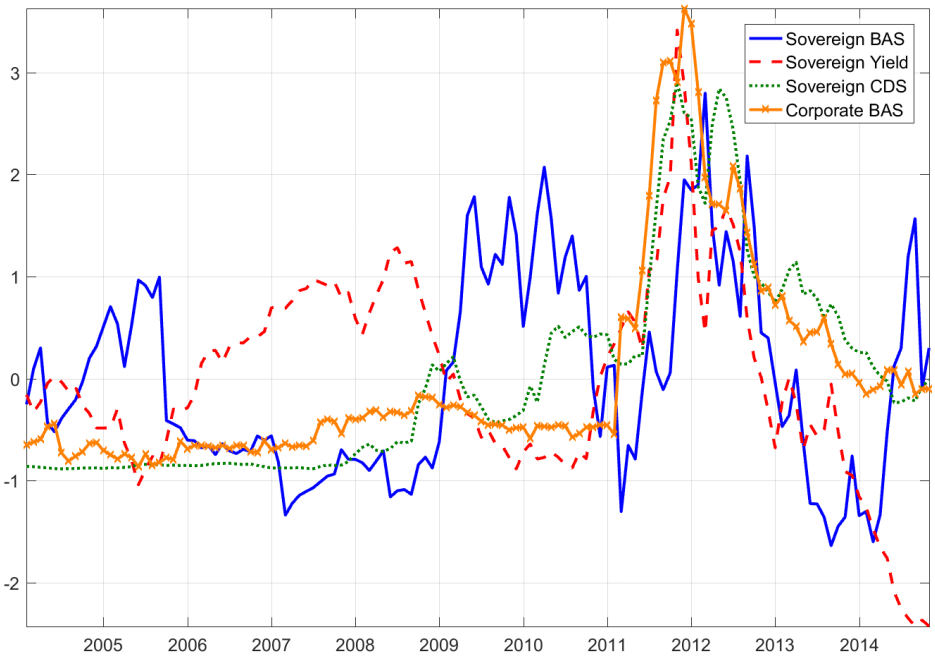

Figure C.14: Comparison among Sovereign and Corporate BAS, Spread and CDS (as monthly averages). Source of Corporate BAS: Bloomberg.

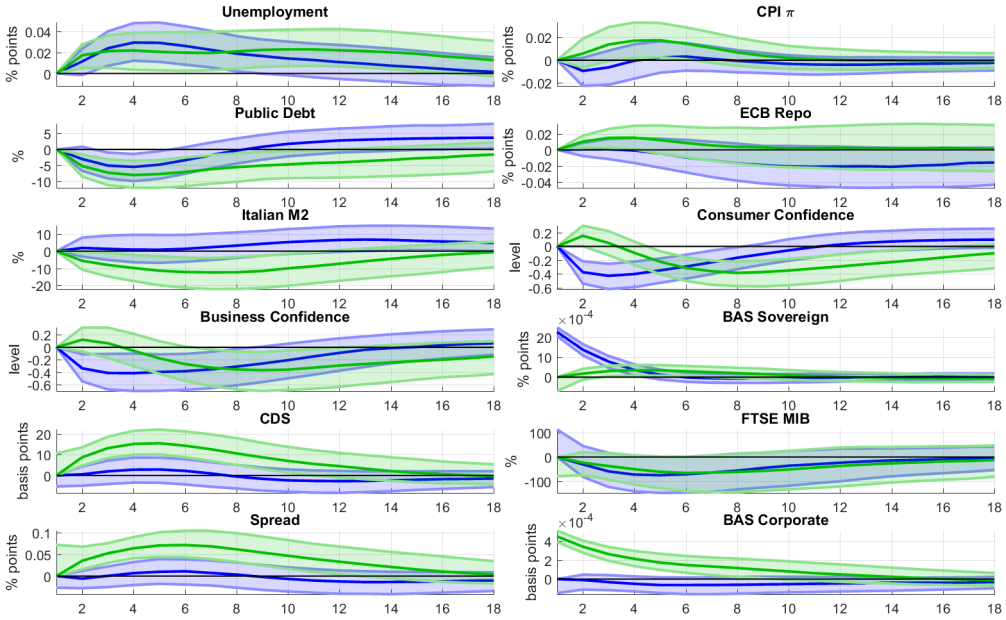

Figure C.15: IRFs to a BAS shock- Choleski identification; sovereign and corporate liquidity

IRFs to a 1 std Corporate BAS shock (compared to a sovereign BAS shock in blue) identified through the following ordering [Unemployment, π, Public Debt, R, M2, CC, BC, Financial Block]. The median point estimate, 68% and 90% confidence bands are reported in cyan, blue, and light blue, respectively. 50%, 68% and 90% bands include statistical and identification uncertainty (from all the possible ordering within the financial block).

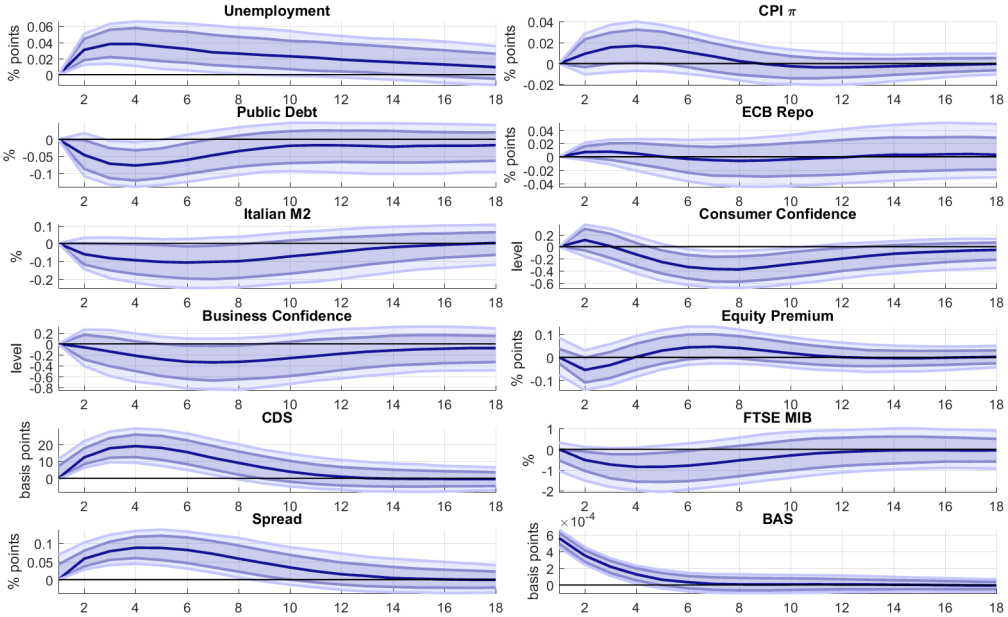

Figure C.16: IRFs to a BAS shock- Choleski identification; corporate bond liquidity

IRFs to a 1 std (Corporate-Sovereign) BAS shock identified through the following ordering [Unemployment, π, Public Debt, R, M2, CC, BC, Financial Block]. The median point estimate, 68% and 90% confidence bands are reported in cyan, blue, and light blue, respectively. 50%, 68% and 90% bands include statistical and identification uncertainty (from all the possible ordering within the financial block).

C.4.5 Market Stress Index

Figure A19 displays the IRFs to a BAS shock of the enlarged VAR that includes the *Composite Indicator of Systemic Stress*, computed by the ECB.

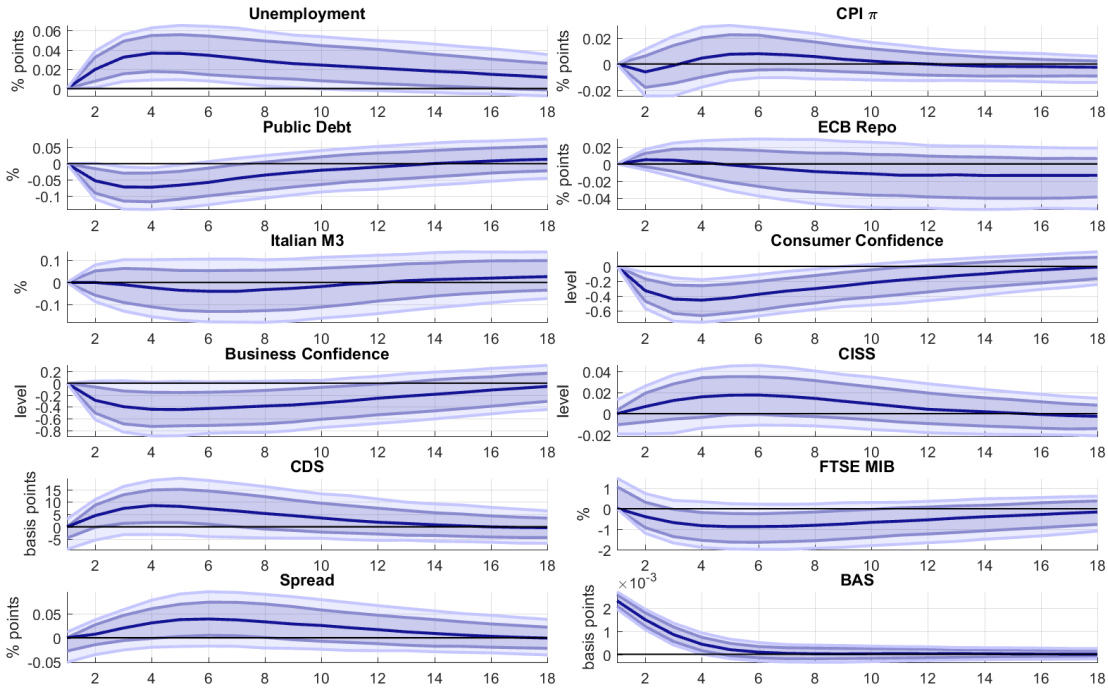

Figure C.17: IRFs to a BAS shock- Choleski identification; CISS

IRFs to a 1 std BAS shock identified through the following ordering [Unemployment, π, Public Debt, R, M2, CC, BC, Financial Block]. The CISS Index is included in place of the the equity premium. The median point estimate, 68% and 90% confidence bands are reported in cyan, blue, and light blue, respectively. 50%, 68% and 90% bands include statistical and identification uncertainty (from all the possible ordering within the financial block).

C.4.6 Financial Volatility

We report the IRFs to a BAS shock of the enlarged VAR that includes an indicator that account for volatility in sovereign debt markets. This indicator is defined as the first principal component of the realized monthly volatility of sovereign BAS, Spread and CDS, computed using daily data.

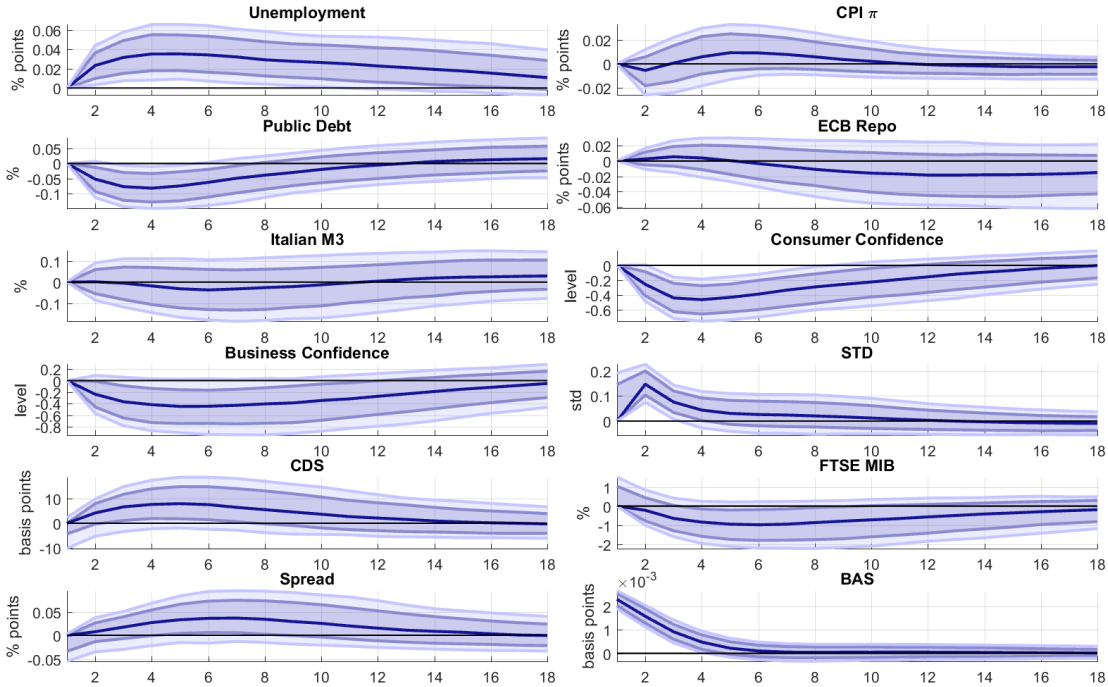

Figure C.18: IRFs to a BAS shock- Choleski identification; financial volatility

IRFs to a 1 std BAS shock identified through the following ordering [Unemployment, π, Public Debt, R, M2, CC, BC, Financial Block]. A principal component that summarizes the volatility of financial variables is included in place of the equity premium. The median point estimate, 68% and 90% confidence bands are reported in cyan, blue, and light blue, respectively. 50%, 68% and 90% bands include statistical and identification uncertainty (from all the possible ordering within the financial block).