Nonlinear Transmission of Financial Shocks: Some New Evidence*

Mario Forni
Università di Modena e Reggio Emilia, CEPR and RECent

Luca Gambetti†
Universitat Autònoma de Barcelona, BSE, Università di Torino, CCA

Nicolò Maffei-Faccioli
Norges Bank

Luca Sala
Università Bocconi, IGIER, Baffi CAREFIN

Abstract

Financial shocks generate a protracted and quantitatively important effect on real economic activity and financial markets only if the shocks are both negative and large. Otherwise, their role is quite modest. Financial shocks have become more important for economic fluctuations after the 2000 and have contributed substantially to deepening the recessions of 2001 and 2008. The evidence is obtained using a new econometric procedure based on a VMA representation that includes a nonlinear function of the financial shock.

JEL classification: C32, E32.

Keywords: SVAR, Financial shocks, Non-linearity, Asymmetry, Financial crisis.

---

*This paper should not be reported as representing the views of Norges Bank. The views expressed are those of the authors and do not necessarily reflect those of the Norges Bank. The authors acknowledge the financial support of the Italian Ministry of Research and University, PRIN 2017, grant J44I20000180001.

†Luca Gambetti acknowledges the financial support from the Spanish Ministry of Science and Innovation, through the Severo Ochoa Programme for Centres of Excellence in R&D (CEX2019-000915-S), the financial support of the Spanish Ministry of Science, Innovation and Universities through grant PGC2018-094364-B-I00, and the Barcelona Graduate School Research Network.

1 Introduction

The 2008 recession has sparked a renewed interest in understanding how financial crises propagate to the real economy. There is a long-standing literature on the interactions and linkages between the two sectors. Early works include Bernanke and Gertler (1989), Carlstrom and Fuerst (1997), Bernanke, Gertler, and Gilchrist (1999) and Kiyotaki and Moore (1997). More recent contributions focusing on the role of the financial sector in economic fluctuations include Christiano, Motto and Rostagno (2003, 2007), Curdia and Woodford (2010), Gertler and Karadi (2011), and Gertler and Kiyotaki (2011). Most of these papers rely on log-linear approximations that imply a linear propagation of economic shocks, including the financial shock.

Nevertheless, a few recent contributions have emphasized the importance of the nonlinear amplification of financial shocks on the real economy, see Mendoza (2010), He and Krishnamurthy (2013), Brunnermeier and Sannikov (2014) and the survey in Brunnermeier, Eisenbach, and Sannikov (2013). In Brunnermeier and Sannikov (2014), there are several nonlinearities at work. First, while small shocks tend to be easily absorbed without important consequences, large shocks may bring the economy far from the steady state, in regions where the amplification mechanisms are much stronger and the consequences for real economic activity much more severe. Second, financial disruptions tend to have larger real effects than financial expansions. Third, when the economy moves away from the steady state, for instance in a downturn, even small shocks can have large effects.[1]

On the empirical side, several studies have investigated and measured the

[1] Several studies have pointed out that business cycle fluctuations tend to be asymmetric since recessionary episodes have larger effects on growth than booms, see for instance Neftci (1984), Sichel (1993) and Morley and Piger (2012).

effects of financial shocks on the real economy. In the seminal contribution by Gilchrist and Zakrajšek (2012) (GZ henceforth), the financial shock is identified in a VAR under a recursive ordering as the shock to the excess bond premium. The shock is found to have significant effects on macroeconomic indicators. Similar findings are obtained in other works using alternative identification schemes, for instance sign restrictions (see Gambetti and Musso, 2017, Furlanetto et al., 2019, Meeks, 2012, Peersman, 2011, and Peersman and Wagner, 2015) and penalty functions (see Brianti, 2021, and Caldara et al., 2016). All in all, the empirical evidence supports the conclusion that financial shocks have significant effects on real activity, although the magnitude of the effects varies across studies. In most of the works based on SVAR models, the variance of macroeconomic variables attributable to financial shock is around 15%-20%, while studies using FAVAR models find higher percentages (Gilchrist et al. , 2009 and Boivin, Giannoni and Stevanovic, 2020).

In the vast majority of empirical works, the effects of the financial shock are estimated using linear models. A notable exception is Barnichon, Matthes and Ziegenbein (2020), which provides evidence supporting the existence of a sign asymmetry in the transmission of financial shocks. More specifically, bad financial shocks have larger effects than good shocks on industrial production. Another contribution is Hubrich and Tetlow (2015), which finds that financial shocks affect the economy differently depending on whether financial stress is high or low. When financial stress is high, the effects are much larger. Apart from these contributions, the literature is quite silent about the empirical support of the nonlinear amplification effects embedded in theoretical models.

This paper contributes empirically to shed new light on the nonlinearity of the transmission mechanisms of financial shocks. We extend the GZ identification to a nonlinear setting where the economy has a Vector Moving Average representation augmented with a nonlinear function of the financial shock.[2] The effects of the financial shock on economic variables are a combination of the coefficients associated with the shock and its nonlinear function. This makes the impulse response functions potentially nonlinear and asymmetric. The identification and estimation procedure consists of two steps within a single model. In the first step, we identify the financial shocks as in GZ. In the second step, we use the shock and its square to estimate the nonlinear transmission mechanisms. In the baseline version of the model, we use the square of the shock and in an extended version we also allow for state-dependent effects.

All in all, our evidence points to the existence of significant nonlinearities in the transmission mechanisms of financial shocks. The main results are as follows:

(i) The effects of big shocks, two standard deviations or more, are highly asymmetric: big financial disruptions generate a sizable and protracted downturn in real economic activity and the stock market, while big financial booms have modest effects. Big negative shocks also turn out to be a major driver of economic slowdowns in the last two decades. We find two such shocks in the sample considered, in early 2000 and in 2008 respectively. In both episodes, the nonlinear term plays a major role and contributes substantially to deepening the recessions that followed.

(ii) One-standard deviation shocks, or smaller, have symmetric effects, i.e. equal in absolute value for positive and negative shocks. Although their effects are significant, from a quantitative point of view their role for economic fluctuations is quite modest.

[2] The approach is similar in spirit to the local projection approach for instance used in Tenreyro and Thwaites (2016) and Alpanda, Granziera and Zubairy (2019) to study the asymmetric and state-dependent effects of monetary policy shocks.

(iii) The effects of financial shocks are further amplified in periods of high uncertainty. No other state variables appear to matter for the transmission of financial shocks.

(iv) Overall, financial shocks have become much more important for economic fluctuations in the last two decades.

Our findings confirm the existence of the type of asymmetries found in Barnichon, Matthes and Ziegenbein (2020), but only for unusually large shocks, since for standard shocks, or smaller, the effects are symmetric. All in all, our results lend empirical support to the nonlinear transmission mechanisms of financial shocks embedded in recent theoretical models such as Brunnermeier and Sannikov (2014) where big negative shocks matter, while small shocks appear to be quickly and easily absorbed without important real consequences.

From a methodological point of view, the approach employed here represents an alternative to that of Barnichon and Matthes (2018), which consists of the estimation, via maximum likelihood, of a nonlinear Vector Moving Average. Both methodologies are flexible enough to handle various types of nonlinearities but have different strengths. The approach of Barnichon and Matthes (2018) can accommodate various types of identification of the economic shock beyond the recursive scheme. Our method does not require any assumption about the probability distribution of economic shocks and is very easy to implement since it only requires OLS estimation. Both, however, are valid alternatives that can be used to study nonlinear transmission mechanisms.

The remainder of the paper is organized as follows. Section 2 discusses the econometric approach. Section 3 presents the results and some robustness checks. Section 4 concludes.

2 Econometric approach

In this section, we present our econometric approach and show how to estimate the nonlinear effects of financial shocks on macroeconomic and financial variables.

2.1 The nonlinear model

We assume that the macroeconomic and financial variables in the n-dimensional vector x_t have the following structural representation

$$x_t = \nu + \beta(L)g(u_{ft}) + B(L)u_t \tag{1}$$

where u_t is a serially independent n-dimensional vector of structural shocks with zero mean and covariance equal to the identity matrix, u_{ft} is the financial shock and is the f-th element of the vector u_t, $g(u_{ft})$ is a contemporaneous nonlinear function of the financial shock, ν is a vector of constants, $B(L) = (I + B_1 B_0^{-1} L + B_2 B_0^{-1} L^2 + ...)B_0$ is a $n \times n$ matrix of impulse response functions and $\beta(L) = \beta_0 + \beta_1 L + \beta_2 L^2 + ...$ is a n-dimensional vector of impulse response functions. One can think of (1) as a generalization of the standard Vector Moving Average (VMA) representation underlying Structural VARs, which, as discussed below, gives origin to nonlinear dynamics for the shock u_{ft}. Model (1) can also be seen as a restricted version of the Volterra representation of y_t. Of course there could be other shocks having nonlinear dynamics but they are not modeled here. Nonetheless, in the simulation section, we assess the performance of our econometric procedure in a more general context where also other shocks have a nonlinear transmission.

We further develop the model and derive an equivalent representation, which is the one we will estimate. Assuming invertibility of the linear term $B(L)u_t$, we obtain the representation

$$D(L)x_t = \mu + D(L)\beta(L)g(u_{ft}) + B_0 u_t \tag{2}$$

where $D(L) = (I + B_1 B_0^{-1} L + B_2 B_0^{-1} L^2 + ...)^{-1} = I - \tilde{D}(L)$, and $\mu = D(1)\nu$. For simplicity we also assume that no lags of $g(u_{ft})$ enter equation (2), i.e. $D(L)\beta(L) = \beta_0$. We will relax this assumption as a robustness check. The model can therefore be rewritten as

$$\begin{aligned} x_t &= \mu + \tilde{D}(L)x_t + \beta_0 g(u_{ft}) + B_0 u_t \\ &= \mu + \tilde{D}(L)x_t + \beta_0 g(u_{ft}) + \alpha_0 u_{ft} + B_{-f0} u_{-ft} \end{aligned} \tag{3}$$

where α_0 is the column of B_0 corresponding to the financial shock, B_{-f0} is the matrix formed by the $n-1$ columns of B_0 excluding α_0 and u_{-ft} is the $(n-1)$-dimensional vector containing the remaining structural shocks other than u_{ft}. Notice that the linear SVAR is nested in our model.

From equations (2) and (3) it is seen that the impulse response functions to u_{ft} and $g(u_{ft})$ are $\alpha(L) = D(L)^{-1}\alpha_0$ and $\beta(L) = D(L)^{-1}\beta_0$ respectively. The total effect is nonlinear and can be found by combining the two terms as

$$\text{IRF}(g(u_{ft}), u_{ft} = u^*) \;\; = \;\; \alpha(L)u^* + \beta(L)g(u^*). \tag{4}$$

So, if nonlinearity is unimportant, i.e. $\beta(L) = 0$, then the impulse response functions will be identical to those of a linear SVAR. On the contrary, if nonlinearity is actually important, then the propagation mechanisms of the financial shock will differ.

Now suppose, $g(u_{ft}) = u_{ft}^2$, which, as discussed below, is our baseline specification. The effect of the financial shock will then be

$$\text{IRF}(u_{ft}^2, u_{ft} = u^*) \quad = \quad \alpha(L)u^* + \beta(L)(u^*)^2. \tag{5}$$

A few remarks are in order. First, in equation (4) the coefficients $\beta(L)$ generate an asymmetry between positive and negative shocks. In particular, when $u^* = 1$, the effect is $\alpha(L) + \beta(L)$. When $u^* = -1$, the effect is $-\alpha(L) + \beta(L)$.

Second, a nonlinearity in terms of magnitude also arises. For $u^* = 1$, the effect is $\alpha(L) + \beta(L)$; for $u^* = 2$, the effect is $\alpha(L)2 + \beta(L)4$. Hence a shock of double magnitude will not have twice the effects.

Notice that other types of nonlinearity can be considered. For instance, one can consider state dependence. Suppose we are interested in understanding whether the financial shock has different effects in different regimes. Let d_t be the state variable of interest, and let $d_t = 1$ if regime one is in place and $d_t = 0$ if regime two is. Defining $g(u_{ft}) = d_t u_{ft}$, the impulse response functions are

$$\text{IRF}(d_t u_{ft}, u_{ft} = u^*) \quad = \quad \alpha(L)u^* + \beta(L)d_t u^*$$

so that $\alpha(L)u^* + \beta(L)u^*$ is the response in regime one and $\alpha(L)u^*$ is the response in regime two.

The framework can also be extended to consider more than one nonlinear function. For instance, one can include both state dependence and the square term

$$\text{IRF}(d_t u_{ft}, u_{ft}^2, u_{ft} = u^*) \quad = \quad \alpha(L)u^* + \beta(L)(u^*)^2 + \gamma(L)d_t u^* \tag{6}$$

where $\gamma(L)$ is another column vector of impulse response functions. In this specification, the nonlinearity depends on the sign and the size of the shock as well as the regime in place. We explore this extension in the empirical section.

2.2 Identification and estimation

The estimation of the effects of the financial shock is based on a two-step procedure. In the first step, an estimate of the financial shock is obtained. In the second step, the nonlinear effects of the shock are obtained using the estimated shock and its nonlinear function as regressors in equation (3).

Step 1

The financial shock is estimated using GZ's strategy.[3] Let us briefly recall their approach. They estimate a VAR with a set of slow-moving variables ordered before the excess bond premium (EBP), and a set of fast-moving variables ordered after. They then impose a Cholesky scheme where the financial shock is the f-th element of the Cholesky representation, where f corresponds to the position of the excess bond premium in the vector x_t.

We adapt their identification strategy to our framework by imposing:

(A) B_0 is lower triangular, as in GZ. This entails that the financial shock has no contemporaneous effect on the slow-moving variables through the linear term, i.e. $\alpha_{0,j} = 0$, $j = 1, ..., f-1$;

(B) the financial shock does not affect the slow-moving variables on impact through the nonlinear term, i.e. $\beta_{0,j} = 0$, $j = 1, ..., f-1$, consistently with assumption (A);

(C) the financial shock does not affect the EBP on impact through the

[3] Note that more recently other identification procedures have been proposed. For instance Caldara et al. (2016) employs a penalty function approach. While acknowledging that other schemes can be successful in identifying financial shocks, we believe the seminal GZ procedure is the best starting point for our analysis.

nonlinear term, i.e. $\beta_{0,f} = 0$.[4]

Under restrictions A, B and C, it is seen that the financial shock is identified and estimated consistently as in GZ, simply by imposing a Cholesky scheme in a standard linear VAR, despite the fact that the standard VAR is misspecified since the true model is the VARX in equation (3). This is shown in the Appendix. The intuition of this result is that restrictions A, B and C imply that the nonlinear term enters only the equations of the fast-moving variables. Therefore the first f shocks, including the financial shock, are simply combinations of the current and past values of the variables as in the linear VAR.

Note that while assumption A is necessary to identify the shock, assumptions B and C could be already satisfied in the data and therefore not imposed ex-ante. Indeed, in the robustness section we estimate the model without imposing restrictions B and C and find almost identical results, suggesting that they hold in the data. This shows that the truly important assumption is A, which in turn implies that identification in the nonlinear model is, *de facto*, the same as in the linear model.

Step 2

We use the estimates of the shock and its nonlinear functions, $\hat{u}_{ft}$ and $g(\hat{u}_{ft})$, as regressors in equation (3). We estimate equation (3) by equation with OLS imposing B and C, obtaining an estimate of α_0, β_0, $\tilde{D}(L)$ and $D(L) = I - \tilde{D}(L)$. An estimate of the impulse response functions is obtained as $\hat{\alpha}(L) = \hat{D}(L)^{-1}\hat{\alpha}_0$ and $\hat{\beta}(L) = \hat{D}(L)^{-1}\hat{\beta}_0$, and the total effects are obtained from equation (4).

It should be noted that, unlike other approaches in the literature, ours does not require any distributional assumptions on the shocks, in addition to serial

[4] Note that the model even under A, B and C is still more general than a linear SVAR under A, i.e. the linear SVAR is nested in our model. The reason is that there is the nonlinear term which enters the equations of the fast-moving variables.

independence and orthogonality of structural shocks.

Note also that, having a narrative measure of the financial shocks, one could skip step 1 and go directly to estimation of the VARX (3), without imposing identification restrictions.

2.3 Variance decomposition and inference

Let us now consider the variance decomposition. Standard formulas are not appropriate in the present context, since $\hat{u}_{ft}$ and $g(\hat{u}_{ft})$ are not orthogonal in general. A simple way to overcome this problem is to compute, for each horizon, the prediction error due to u_{ft}, including the nonlinear term, and divide its sample variance by the sample variance of the total prediction error. Precisely, the h-step ahead prediction error implied by equation (1) is

$$e_{t+h} = \sum_{k=0}^{h-1} \beta_k g(u_{ft}) + \sum_{k=0}^{h-1} B_k u_{t+h-k}$$

and, according to equation (4), the component of the prediction error driven by the financial shock is given by

$$e_{f,t+h} = \sum_{k=0}^{h-1} \alpha_k u_{ft} + \sum_{k=0}^{h-1} \beta_k g(u_{ft}).$$

We use the estimated coefficients and shocks to estimate $e_{f,t+h}$ and e_{t+h} according to the equations above. The variance contribution is then computed as the ratio of the sample variances.

Finally, the confidence bands of the impulse responses are constructed using a bootstrap procedure which accounts for the generated regressors problem in the second step. The approach relies on the following steps: (i) we bootstrap the estimated shock u_{ft} obtained from step 1 and the estimated VARX residuals obtained by estimating (3); (ii) using the bootstrapped shock and residuals and the estimated coefficients of equation (3), we create a new sample; (iii) we re-estimate the financial shock as in step 1 and model (3) as in step 2 with the new sample and compute the impulse response functions. We repeat steps (i)-(iii) 1000 times to construct the confidence bands of the impulse responses.

2.4 Simulations

We use a simulation exercise to assess our econometric procedure. We consider a three-variables VARX(1) where x_{1t} is the slow moving variable, x_{2t} is the excess bond premium, and x_{3t} is the fast moving variable. The financial shock, u_{ft}, is the second shock in u_t. We assume $u_t \sim N(0, I)$ and $g(u_{ft}) = u_{ft}^2$. The shock u_{ft} has zero impact effects on x_{1t} and non-zero effects on x_{2t} and x_{3t}. The coefficients are

$$\tilde{D}_1 = \begin{pmatrix} 0.2 & 0.4 & 0.2 \\ 0.3 & 0.7 & -0.1 \\ 0.3 & -0.2 & 0.6 \end{pmatrix}$$

and the impact effects are

$$B_0 = \begin{pmatrix} 0.6 & 0 & 0 \\ -0.3 & 0.5 & 0 \\ -0.4 & -0.1 & 0.5 \end{pmatrix}$$

Moreover, we assume that $\beta_0 = [0\ 0\ 0.5]'$ so that the restriction $\beta_i = 0$, $i = 1, 2$ holds and $\mu = 0$. The parametrization is of course arbitrary. This however does not affect the results as long as assumptions A, B and C are satisfied. With other parametrizations, the same conclusions about the performance of our procedure are reached. The restriction implies that the nonlinear term does not enter the first and second structural equation. We generate $N = 5000$ datasets of length $T = 518$ (as in the dataset used in the empirical section). For each dataset, we apply our two-step procedure and estimate $\alpha(L)$ and $\beta(L)$. Panel (a) of Figure 1 displays the results. The left column plots $\alpha(L)$, while the right column reports $\beta(L)$. The solid black lines are the mean (across dataset) point estimates, the gray area is the region included in the 16th and 84th percentile of the point estimates distribution and the dash-dotted line is the theoretical response. The mean responses and the theoretical ones exactly overlap, suggesting that our procedure is successful in estimating the effects of the financial shock.

In a second simulation, we use the same value of the parameters but we assume that the nonlinear function does not have any effect on any of the three variables. The simulation is performed as before with the only difference that now $\beta_0 = [0\ 0\ 0]'$. Panel (b) of Figure 1 displays the results. Our procedure is able to detect that the nonlinear term does not have any impact (second column).

In the third simulation, we repeat the first simulation assuming a different distribution for u_{ft}. More specifically we assume a Pearson type IV distribution with zero mean, standard deviation equal to 1, skewness equal to 0.5 and kurtosis equal to 4. Now the distribution of u_{ft} is asymmetric and the correlation between u_{ft} and $g(u_{ft})$ is no longer zero. Panel (c) reports the results of the simulation. As before, the procedure is able to estimate the true effects. The procedure, as discussed earlier, works independently of the type of distribution of the shock.

In the fourth simulation we extend the model assuming that also u_{3t} has nonlinear effects. The model therefore becomes

$$x_t \;\; = \tilde{D}_1 x_{t-1} + \beta_0 u_{ft}^2 + \delta_0 u_{3t}^2 + B_0 u_t \tag{7}$$

where $\delta_0 = [0\ 0\ \ -0.2]'$. Notice that still the conditions for identification are satisfied. Panel (d) reports the results. Again the estimated effects and the true effects overlap, suggesting that the procedure is robust even when other shocks have a nonlinear transmission.

3 Results

In our empirical application we use monthly US data spanning the period 1973:M1 - 2016:M8. In our baseline specification, the vector x_t in equation (3) includes the following variables (in parenthesis, the abbreviation used throughout the paper), in the following order: the log of industrial production (INDPRO), CPI Inflation (CPI), the unemployment rate (UNRATE), the GZ excess bond premium index (EBP), the Chicago Fed's National Financial Conditions Index (NFCI),[5] the S&P Composite Stock Price Index (SP500) and the federal funds rate (FFR).[6] We use five lags in $\tilde{D}(L)$. We set $g(u_{ft}) = u_{ft}^2$ since the quadratic function can accommodate both sign and size asymmetries.

[5]The Chicago Fed's National Financial Conditions Index (NFCI) provides a weekly estimate on U.S. financial conditions in money markets, debt and equity markets and the traditional and shadow banking systems. The index is a weighted average of 105 measures of financial activity. When the NFCI is positive, financial conditions are tighter than average. The methodology used to compute the NFCI is described in Brave and Butters (2012).

[6]Data on industrial production, the CPI deflator, the unemployment rate and the NFCI index is retrieved from the Federal Reserve Economic Data (FRED). The federal funds rate is the effective federal funds rate until 2008:M8. From 2008:M9 - 2016:M8, it corresponds to the shadow rate of Wu and Xia (2016). The S&P Composite Stock Price Index is retrieved from Robert Shiller's webpage.

3.1 The financial shock

Figure 2 plots the financial shock (top panel) estimated in the first step and its square (bottom panel). We observe two major positive spikes in August 2002 and November 2008, of about 4 and 6 standard deviations respectively, which correspond to the stock market downturn of 2002 and the global financial crisis. We also observe two major negative spikes, of about 5 standard deviations, in November 1987 and December 2001, which reflect the rebound from the stock market crash of 1987 and 9/11 respectively. In general, we observe a number of large financial shocks in connection with three key events of turbulence in the history of US financial markets: the stock market crash of 1987, the Dot-com bubble of the early 2000s and the 2008 financial crisis. In contrast, the period between 1990 and 2000 is characterized by low volatility of the shock. The financial shock exhibits high kurtosis (equal to 9), meaning that the distribution of the shock is characterized by substantially fatter tails with respect to a normal.

Overall, the shock seems to effectively capture key developments in US financial markets. In the following subsections, we evaluate the effects of the financial shock and its square on macroeconomic and financial variables.

3.2 Is nonlinearity significant?

Figure 3 reports the results of our two-step procedure using the squared shock. The black solid lines are the point estimates, while the gray areas are the 68% and 90% confidence bands. The first column shows the linear response to the financial shock, while the second column reports the responses of the square term estimated in the second step. The horizontal axis measures time in months from impact to 48 months after innovations have occurred.

The responses to the linear term (first column) are very much in line with previous findings in the literature (see GZ). The financial shock significantly reduces output and inflation and increases the unemployment rate, thus behaving like a typical demand shock. The effects on industrial production and unemployment are persistent and sizable. Additionally, the shock leads to a short-run increase in the NFCI index, highlighting a worsening of overall financial conditions, and to a marked and persistent decline in stock prices. The federal funds rate declines with a few months of delay, suggesting an endogenous monetary easing in response to the adverse economic developments.

Focusing on the responses to the nonlinear function of the financial shock (second column), the square has significant effects on real activity and financial variables. We observe a persistent decline in industrial production and a persistent increase in the unemployment rate. This result implies that the reduction in industrial production and the increase in unemployment following a bad financial shock are significantly amplified by the nonlinear term. Having a closer look at financial variables, the nonlinear effects are particularly relevant for stock prices, both in terms of magnitude and persistence. Overall, the results point towards significant nonlinear effects of financial shocks.

3.3 Nonlinear effects of financial shocks

Here, we focus on the total effects of financial shocks by summing the two components whose separate effects have been discussed above.

The square term introduces two potential asymmetries, one in terms of the sign of the shock, and one in terms of the size of the shock. In the following, we assess their role for the overall response of macroeconomic variables to financial shocks. Figure 4 shows the overall effects of positive (solid black line) and negative (dotted red line) financial shocks for different magnitudes of the shock: one, two and four standard deviations. Negative shocks are normalized to have the same sign as positive shocks, for the sake of comparison.

When considering a shock of one standard deviation, negative and positive shocks have very similar effects. The responses basically overlap for inflation, the excess bond premium, the NFCI index and the federal funds rate, while the effects on industrial production and the unemployment rate appear slightly larger for positive shocks. In this case, stock prices are the only variable that display a significantly different response to positive and negative shocks.

The picture changes substantially when we consider shocks of larger magnitudes: bad financial shocks have markedly larger effects. The asymmetry between positive and negative shocks is clear in the case of a shock of two standard deviations and becomes very large in the case of a four standard deviations shock (third column). To give some figures, a four standard deviations positive shock leads to a decline in industrial production of about 2.8 percent after one year, while a negative shock of the same magnitude increases industrial production by 1.6 percent. For the unemployment rate, a negative shock produces a decrease of 0.3 percentage points, while a positive shock implies a rise of 0.7 percentage points. For stock prices, positive shocks have more than twice the effect of negative shocks.

The results seem to be very much in line with the theoretical predictions in Brunnermeier and Sannikov (2014). Relatively small shocks are easily absorbed by economic agents and the effects, although significant, are modest and symmetric. However, the size of the effects increases more than linearly with the size of the shock. Big financial disruptions have very large effects, while large financial expansions do not.

At this point, a natural question is how relevant financial shocks and their nonlinearities are in explaining fluctuations in the macroeconomic and financial variables we include in the analysis. To address this question we compute for each variable a historical decomposition, which allows us to evaluate the relative importance of the different disturbances at each point in time. Figure 5 reports the results. The black solid line represents, for each variable, the sum of the contributions of each shock in the system, that is the variable in deviation from its deterministic components. The colored areas represent the contributions of the financial shock, its square and the remaining shocks in the system, which we label as "residual" for simplicity. The residual shocks include any macroeconomic shock other than the financial one.

Until the late 90s, fluctuations in industrial production, with the exception of a few years at the end of the 80s, are essentially explained by the residual shocks. The role of financial shocks was very modest in the pre-2000 sample. However, after 2000, financial shocks become a major driver of fluctuations in industrial production. In particular the shock substantially contributed to deepening the economic recessions of 2001 and 2008. Notably, the square of the financial shock plays a marked amplification role in both crises, magnifying the contribution of financial shocks to the decline in industrial production. In the aftermath of the Great Recession, the role of financial shocks essentially doubles due to the nonlinear component. A very similar pattern is present for stock prices and the unemployment rate. In linear SVAR models, the nonlinear component remains uncovered and this can explain why financial shocks have been found in previous contributions to account for a relatively small part of real economic activity fluctuations. When the nonlinear part is taken into account, the effects of financial shocks are substantially magnified, at least in the post-2000 period.

By comparing the time series of the square of financial shocks (see Figure 2) with the historical decomposition, a clear-cut result emerges. When large shocks occur, the nonlinear term significantly amplifies the transmission of financial shocks to industrial production, unemployment and stock prices, confirming that only big bad financial shocks generate quantitatively important effects on the real economy. Notice, however, that the square of the shock does not seem relevant in explaining fluctuations in inflation, the excess bond premium and the federal funds rate. This is not too surprising, given that the relevance of the nonlinear term is small for those variables, see Figure 3. Overall, the results are consistent with the impulse responses presented above, which point towards the size and the sign of the shock as the key sources of nonlinearity of financial shocks.

Table 1, lower panel, reports the variance decomposition computed as explained in Section 2.3. At the two-year horizon, the financial shock, including both the linear and nonlinear term, explains around 32%, 33% and 43% of the variance in industrial production, unemployment rate and stock prices, respectively. These numbers point to a more important role of the financial shock for real economic activity relative to that typically found in linear VARs. Indeed these numbers are substantially larger than those of the upper panel of the table, obtained by considering only the component of the prediction error driven by the linear term.[7] In conclusion, the variance decomposition confirms an important role for the nonlinear term and the relevance of nonlinearities for the transmission of the financial shock.

[7]Note however that the two components of the prediction error, the one driven by the linear term and the one driven by the nonlinear term, are correlated to some extent, so that the variance explained by the two terms is not equal to the sum of the two variances.

3.4 Adding state dependence

Another potential source of nonlinearity might be represented by the state of the economy. When the economy is going through a downturn, the effects of financial shocks are amplified than in periods of expansion. Here we assess this prediction by estimating the model augmented with a term which captures the state of the economy. The model is now

$$x_t = \nu + \beta(L)u_{ft}^2 + \gamma(L)d_t u_t + B(L)u_t \tag{8}$$

where d_t is a dummy variable indicating which of the two possible states is in place. The VARX representation is

$$x_t = \mu + \tilde{D}(L)x_{t-1} + \beta_0 u_{ft}^2 + \gamma_0 d_t u_{ft} + \alpha_0 u_{ft} + B_{-f0}(L)u_{-ft}$$

As before, we assume $\alpha_{i0} = 0$ for $i = 1, 2, 3$, $\beta_{i0} = 0$ for $i = 1, 2, 3, 4$. We also impose $\gamma_{i0} = 0$ for $i = 1, 2, 3, 4$, so to make sure that the interaction term does not appear in the slow moving variables' equations. The impulse responses in this model are as in equation (6).

We consider several state variables: tightness of financial conditions, as reflected by the NFCI being above its average; monetary policy tightening cycle, as reflected by the 10-year government bond being higher than its 5 years moving average as in Alpanda, Granziera and Zubairy (2020); recessions and booms, as reflected by the unemployment rate being higher or lower than its average; high-low macroeconomic uncertainty, as reflected by the Ludvigson, Ma and Ng (2021) measure of uncertainty being higher or lower than its average.

For each of these variables we estimate the model and compute the impulse response functions. The only state-dependence that matters is the one associated to macroeconomic uncertainty, see Figure 6. The first column reports $\alpha(L)$, the second column reports $\beta(L)$, while the third column shows $\gamma(L)$. The shock is assumed to be $u^* = 1$. In this case, the effects of financial shocks are significantly amplified in a regime of high uncertainty for all the variables but inflation and the federal funds rate. The effects of the square term are quantitatively very similar to those obtained in the baseline model and significant. The same nonlinear amplification mechanisms obtained in the baseline model are at work even in this specification, but are enhanced in periods of high macroeconomic uncertainty.

Figures 7 reports $\gamma(L)$ for the the other three state variables. In none of the three cases does state-dependence generate significant effects on any of the variables. On the contrary, the square term remains significant at all times (not reported in the figure but available upon request).

3.5 Robustness

Here we assess the robustness of the results using the absolute value of the shock, instead of the square. The impulse response functions are

$$\text{IRF}(|u_{ft}|, u_{ft} = u^*) \quad = \quad \alpha(L)u^* + \beta(L)|u^*| \tag{9}$$

Figure 8 reports the results. The nonlinear term has significant effect on industrial production, unemployment and stock prices. Again, for these three variables the signs are the same as those of the linear term: the nonlinear term reinforces the effects of an exogenous increase in the excess bond premium. The third column shows the effects of positive (solid black line) and negative (dotted red line) financial shocks using the nonlinear function (9). As before, the negative financial shock is normalized to have the responses of the same sign of a positive shock. In line with the findings of Barnichon, Matthes and Ziegenbein (2020), the evidence points to a significantly more important role of financial disruptions than expansions: the effects of an increase in the excess bond premium are larger than the effects of a decrease, especially for industrial production, unemployment and stock prices. Note that, with the absolute value, the size of the shock is no longer a source of nonlinearity. Thus, in the light of the empirical evidence discussed in the previous subsection, the asymmetry obtained here using the absolute value can be thought of as a sort of average of big shocks with large asymmetric effects and small shocks which are essentially symmetric.

As a second robustness check, we relax the restriction $D(L)\beta(L) = \beta_0$ and include also the lags of the nonlinear term, i.e. $D(L)\beta(L) = \tilde{\beta}(L) = \beta_0 + \tilde{\beta}_1 L + ...$ Figure 9 reports the impulse response functions. The estimated responses are very similar, almost identical, to those obtained in the baseline model.

Finally, we estimate the second step by relaxing the assumption that the nonlinear term has no contemporaneous impact on industrial production, inflation, the unemployment rate and the excess bond premium. Figures 10 and 11 report the results. All in all, the main conclusions regarding both the impulse response analysis and the historical decomposition are essentially unchanged. If anything, the confidence bands of the nonlinear term are slightly wider when no restriction on impact is imposed on the first three variables in the system. The fact that the results are very similar depends on the fact that none of the impact effects of the nonlinear term on the slow-moving variables and the excess bond premium is significantly different from zero. In other words, the assumption $\beta_{i0} = 0$ appears to hold in the data and constitutes a validation of the restrictions discussed in Section 2.2.

4 Concluding remarks

Financial shocks play an important role for the real economy and the financial markets only when they are negative and big (two standard deviations or more). One standard deviation (or smaller) shocks and large positive shocks play a modest role. Two large negative shocks are found in early 2000 and 2008, both contributing significantly to the economic downturn that followed.

This marked nonlinearity is obtained using a new econometric procedure based on the estimation of a VMA representation which includes a nonlinear function of the financial shock, with the financial shock identified along the lines of Gilchrist and Zakrajšek (2012).

Barnichon, Matthes and Ziegenbein (2020) show that bad financial shocks have larger effects than good shocks. Here, complementing their analysis, we show that this sign asymmetry exclusively originates from big shocks, since small shocks (either positive or negative) are largely symmetric. Our findings are in line with the theoretical predictions in Brunnermeier and Sannikov (2014), where both sign and size asymmetries in the amplification of financial shocks emerge.

References

[1] Alpanda, S., E. Granziera and S. Zubairy (2019), "State dependence of monetary policy across business, credit and interest rate cycles", Norges Bank Working Paper 2019/21.

[2] Barnichon, R., and C. Matthes (2018), "Functional approximation of impulse responses", Journal of Monetary Economics 99(C), 41-55.

[3] Barnichon, R., C. Matthes and A. Ziegenbein (2020), "Are the effects of financial market disruptions big or small?", Review of Economics and Statistics, forthcoming.

[4] Bernanke, B, M. Gertler, and S. Gilchrist (1999), "The financial accelerator in a quantitative business cycle framework"., in Handbook of Macroeconomics. Vol. 1, , ed. John B. Taylor and Michael Woodford, Chapter 21, 1341-1393. Elsevier.

[5] Boivin, J., M. P. Giannoni and D. Stevanovic (2020), "Dynamic effects of credit shocks in a data-rich environment", Journal of Business & Economic Statistics, 38(2), 272-284.

[6] Brianti, M. (2021), "Financial shocks, uncertainty shocks, and monetary policy trade-offs", Working Papers 2021-5, University of Alberta, Department of Economics.

[7] Brunnermeier, M. K., T. Eisenbach, and Y. Sannikov (2013), "Macroeconomics with financial frictions: a survey". Advances in Economics and Econometrics, Tenth World Congress of the Econometric Society. New York: Cambridge University Press.

[8] Brunnermeier, M. K., and Y. Sannikov (2014), "A macroeconomic model with a financial sector", American Economic Review, 104(2): 379-421.

[9] Caldara, D., C. Fuentes-Albero, S. Gilchrist, and E. Zakrajsek (2016), "The macroeconomic impact of financial and uncertainty shocks", European Economic Review 88, 185-207.

[10] Carlstrom, C. T., and T.S. Fuerst (1997), "Agency costs, net worth, and business fluctuations: a computable general equilibrium analysis", American Economic Review, 87(5), 893-910.

[11] Christiano, L., R. Motto, and M. Rostagno (2003), "The Great Depression and the Friedman-Schwartz hypothesis", Journal of Money, Credit & Banking, 35(6), 1119-1197.

[12] Christiano, L., R. Motto, and M. Rostagno (2007), "Shocks, structures or monetary policies? The Euro Area and US after 2001", Journal of Economic Dynamics and Control, 32(8), 2476-2506

[13] Curdia, V. and M. Woodford (2010), "Credit spreads and monetary policy", Journal of Money, Credit & Banking, 42(S1): 47-74.

[14] Furlanetto, F., F. Ravazzolo, and S. Sarferaz (2019), "Identification of financial factors in economic fluctuations", Economic Journal, vol. 129(617), pages 311-337.

[15] Gambetti, L., and A. Musso (2017), "Loan supply shocks and the business cycle", Journal of Applied Econometrics, vol. 32(4), pages 764-782.

[16] Gertler, M. and P. Karadi (2011), "A model of unconventional monetary policy", Journal of Monetary Economics, 58(1): 17-34.

[17] Gertler, M. and N. Kiyotaki (2011) “Financial intermediation and credit policy in business cycle analysis”, in Handbook of Monetary Economics, eds. Michael Woodford and Benjamin M. Friedman. North Holland, Elsevier.

[18] Gilchrist, S., V. Yankov, and E. Zakrajšek (2009), “Credit market shocks and economic fluctuations: evidence from corporate bond and stock markets”, Journal of Monetary Economics, 56 (4): 471-93.

[19] Gilchrist, S. and E. Zakrajšek (2012), “Credit spreads and business cycle fluctuations”, American Economic Review, 102(4), 1692-1720.

[20] He, Z., and A. Krishnamurthy (2013), “Intermediary asset pricing”, American Economic Review, 103(2): 732-70.

[21] Hubrich, K., and R. J. Tetlow (2015), "Financial stress and economic dynamics: The transmission of crises," Journal of Monetary Economics, Elsevier, vol. 70(C), pages 100-115.

[22] Kiyotaki, N. and J. Moore (1997), “Credit cycles”, Journal of Political Economy, 105(2): 211-248

[23] Ludvigson, S., Ma, S., and Ng, S. (2019), “Uncertainty and business cycles: exogenous impulse or endogenous response?”, American Economic Journal: Macroeconomics, forthcoming.

[24] Meeks, R., (2012), “Do credit market shocks drive output fluctuations? Evidence from corporate spreads and defaults," Journal of Economic Dynamics and Control, vol. 36(4), 568-584.

[25] Mendoza, E. G. (2010), “Sudden stops, financial crises, and leverage”, American Economic Review, 100(5): 1941-66.

[26] Mertens, K. and M.O. Ravn (2013). “The dynamic effects of personal and corporate income tax changes in the United States”, American Economic Review 103, pp. 1212-1247.

[27] Morley, J., and J. Piger (2012), “The asymmetric business cycle”, Review of Economic and Statistics, 94:208-221.

[28] Neftci, S. (1984), “Are economic time series asymmetric over the business cycle?”, Journal of Political Economy, 92: 307-328.

[29] Peersman, G., (2011), “Bank lending shocks and the Euro area business cycle," Working Papers 11/766, Ghent University.

[30] Peersman, G., and W. Wagner (2015), “Shocks to bank lending, risk-taking, securitization, and their role for U.S. business cycle fluctuations”, CEPR Discussion Papers 10547.

[31] Sichel, D., 1993, “Business cycle asymmetry: a deeper look”, Economic Inquiry, 31, 224-236.

[32] Tenreyro, S., and G. Thwaites (2016), “Pushing on a string: US monetary policy is less powerful in recessions”, American Economic Journal: Macroeconomics, 8(4): 43-74.

Appendix

In this Appendix we show that if the true model is the VARX (3), with B_0 lower triangular and $\beta_{0j} = 0$ for $j = 1, \dots, f$, then the financial shock u_{ft} is identical to the financial shock of a misspecified (linear) VAR model where we impose a Cholesky identification scheme. As a result, we can consistently estimate the financial shock following GZ's strategy as explained in Section 2.2, Step 1.

Consider the misspecified VAR model

$$x_t = \alpha + C(L)x_{t-1} + \eta_t,$$

where η_t is orthogonal to x_{t-k}, $k > 0$. Under the recursive identification scheme the VAR innovations are linked to the (misspecified) structural shocks $\tilde{u}_t$ by the relation $\Gamma\tilde{u}_t = \eta_t$, where Γ is the lower triangular matrix such that $\Gamma\Gamma' = E(\eta_t\eta_t')$.

Now, let A^f be the submatrix of A including only the first f rows of A. Then, by imposing our restrictions $\beta_{0j} = 0$, $j = 1, \dots, f$, the first f equations in (3) can be written as

$$x_t^f = \mu^f + \frac{\tilde{D}(L)^f}{L}x_{t-1} + B_0^f u_t.$$

By comparing the VAR and the VARX (and observing that both η_t and u_t are orthogonal to past x's) it is seen that the first f equations are the same, so that

$$\mu^f = \alpha^f, \qquad \frac{\tilde{D}(L)^f}{L} = C(L)^f, \qquad B_0^f u_t = \eta_t^f = \Gamma^f\tilde{u}_t.$$

Since both B_0 and Γ are lower triangular, all entries of the last $n-f$ columns of both B_0^f and Γ^f are zero, so that, denoting by B_0^{ff} and Γ^{ff} the $f \times f$ matrices obtained by eliminating these columns from B_0^f and Γ^f, respectively, we get $B_0^{ff} u_t^f = \Gamma^{ff} \tilde{u}_t^f$. Since $E(u_t^f u_t^{f'}) = E(\tilde{u}_t^f \tilde{u}_t^{f'}) = I_f$ and both B_0^{ff} and Γ^{ff} are lower triangular, uniqueness of the Cholesky factorization implies $B_0^{ff} = \Gamma^{ff}$ and $u_t^f = \tilde{u}_t^f$.

In conclusion, under our assumptions, the first f equations of the VARX model are identical to the corresponding equations of the VAR model where the nonlinear term is omitted, and the first f Cholesky shocks are the same.[8] If follows that GZ's financial shock is estimated consistently even if the VAR is misspecified. Of course, the remaining $n - f$ Cholesky shocks and all impulse-response functions are not.

The above result is still valid under more general conditions. In particular, the assumption that the lags of the nonlinear term do not appear in equation (3), $D(L)\beta(L) = \beta_0$, can be relaxed and we can allow $g(u_{t-k}^f)$, $k > 0$, to affect the slow-moving variables (but not the EBP).

[8]It is well-known that the shocks obtained with the Cholesky identification scheme are equal, up to a normalization, to the residuals of a recursive VAR. By observing this, it is seen that these shocks can be obtained without resorting to the fast moving variables and the related equations.

	Linear			
	$h = 0$	$h = 12$	$h = 24$	$h = 48$
INDPRO	0.0	15.0	17.7	11.9
CPI	0.0	2.4	2.5	2.6
UNRATE	0.0	14.0	19.2	14.5
EBP	97.6	81.2	78.8	67.7
NFCI	3.6	6.7	6.9	6.9
SP500	8.8	19.0	16.0	13.3
FFR	0.7	2.5	6.2	8.0
	Total			
	$h = 0$	$h = 12$	$h = 24$	$h = 48$
INDPRO	0.0	20.5	32.2	28.4
CPI	0.0	2.5	2.6	3.2
UNRATE	0.0	19.4	33.4	31.2
EBP	97.6	86.0	83.4	75.6
NSCI	6.3	11.9	10.8	9.0
SP500	14.9	38.0	43.4	41.8
FFR	0.7	2.9	8.1	11.0

Table 1: Variance Decomposition.

Figures

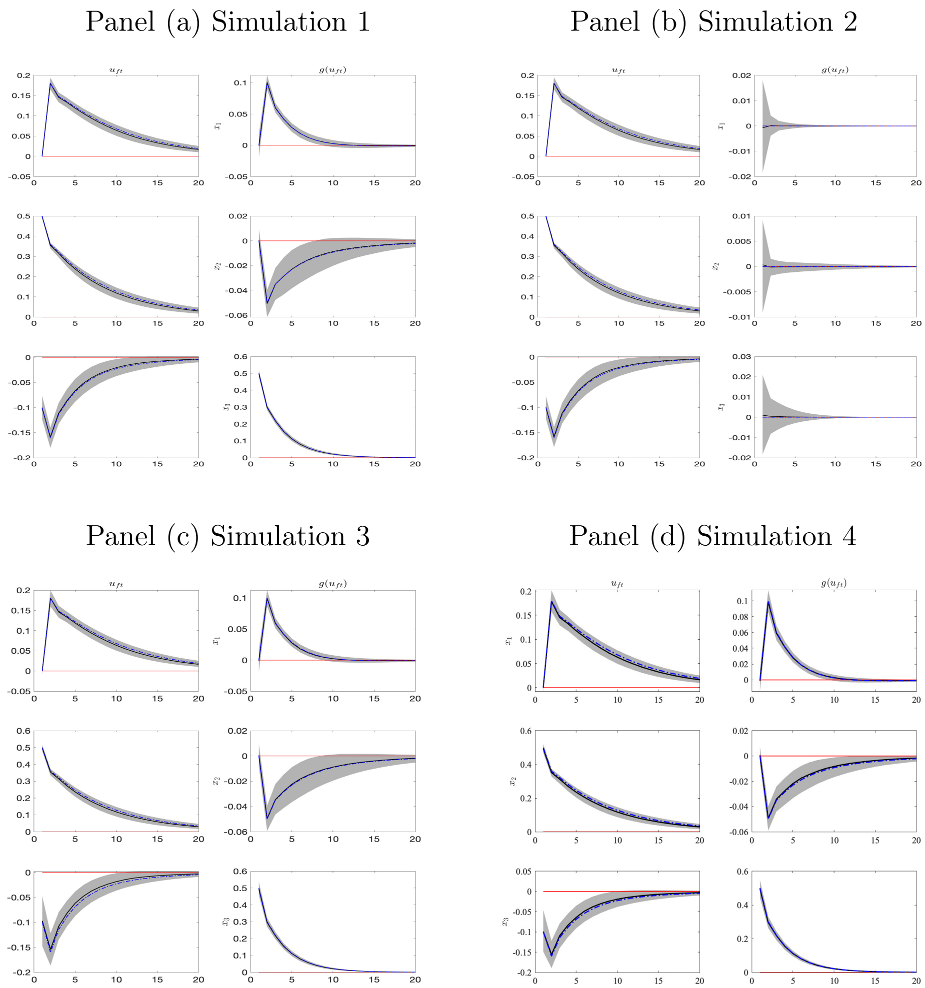

Figure 1: Monte Carlo Simulations. Black solid lines are the average of the point estimates, the gray areas are the interval between the 16th and the 84th percentile and the 10th and 90th percentile respectively of the distribution of the point estimates. The dash-dotted blue lines are the theoretical responses.

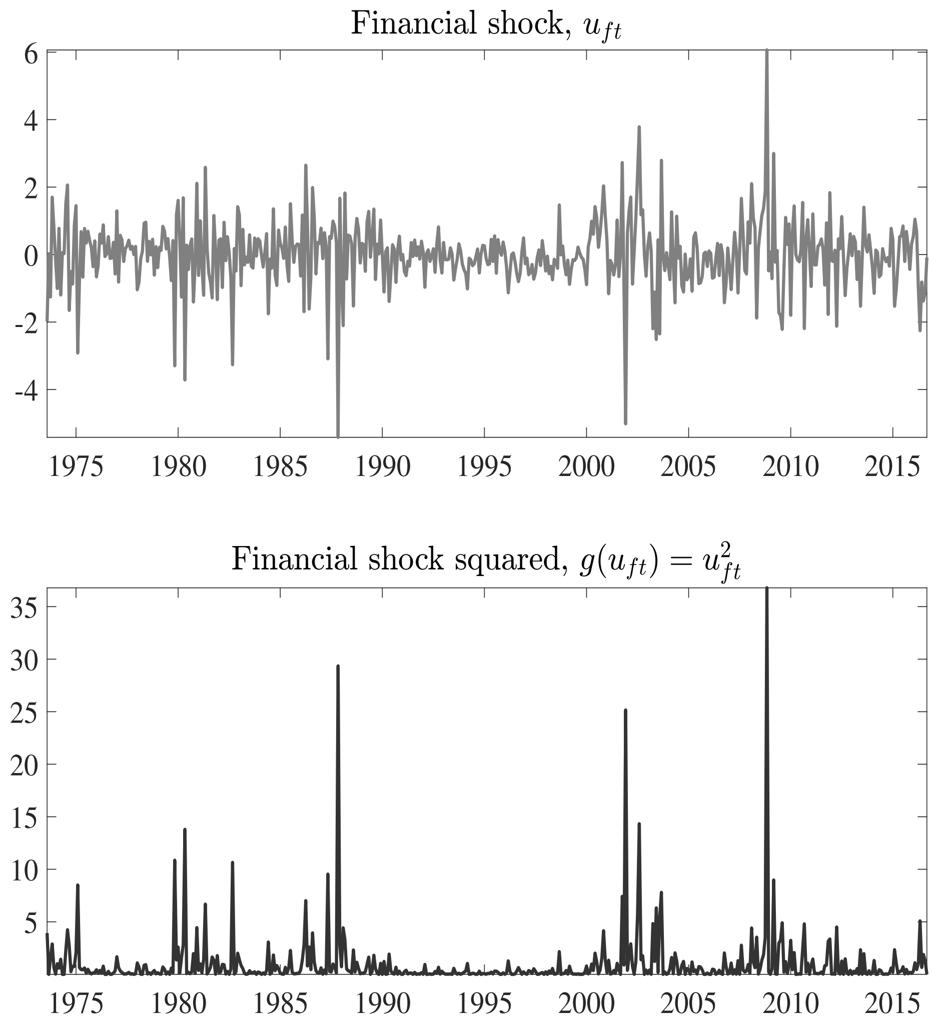

Figure 2: Financial shock (light gray line) and its square (dark gray line) obtained from the first step.

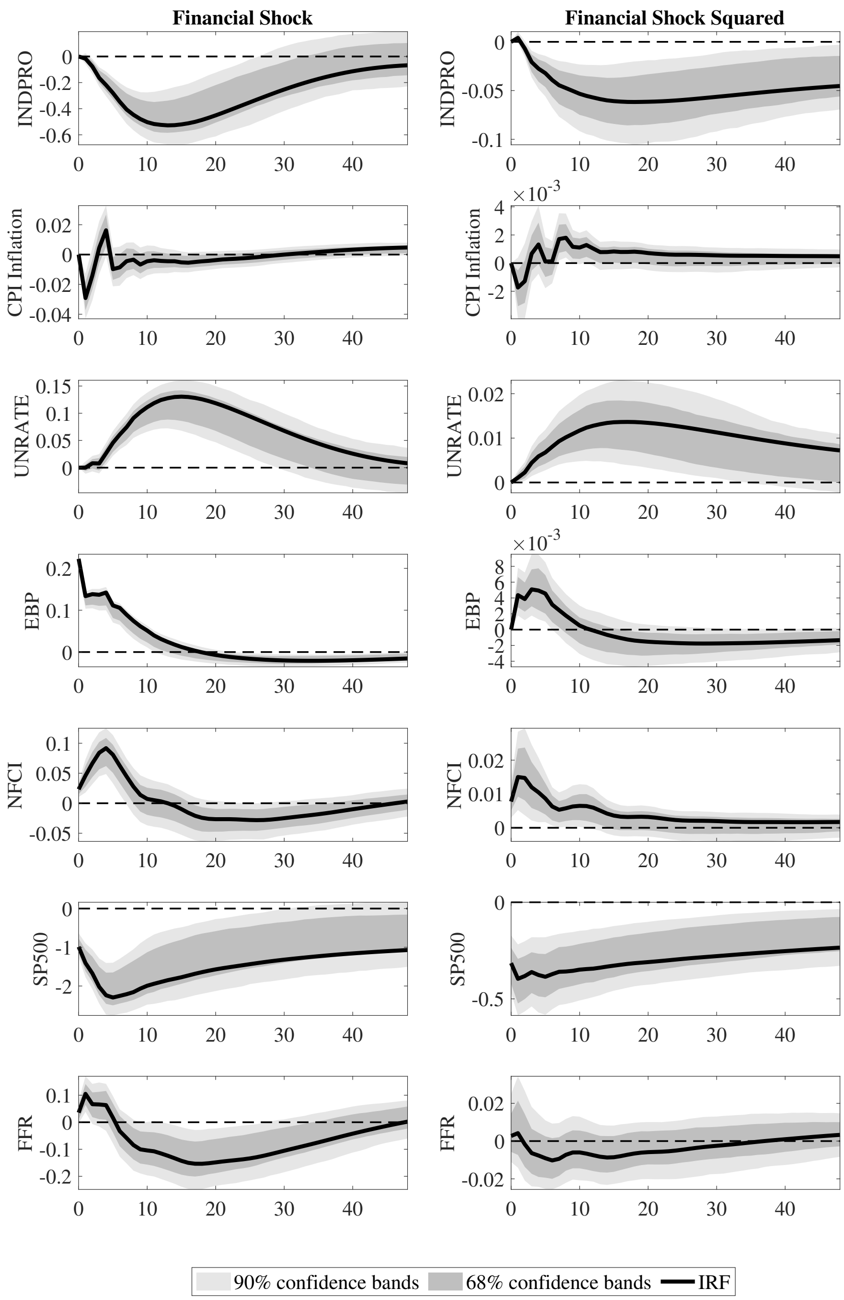

Figure 3: Impulse response functions of the VARX using $g(u_{ft}) = u_{ft}^2$. Black solid lines are the point estimates, the gray areas the 68% and 90% confidence bands of the financial shock.

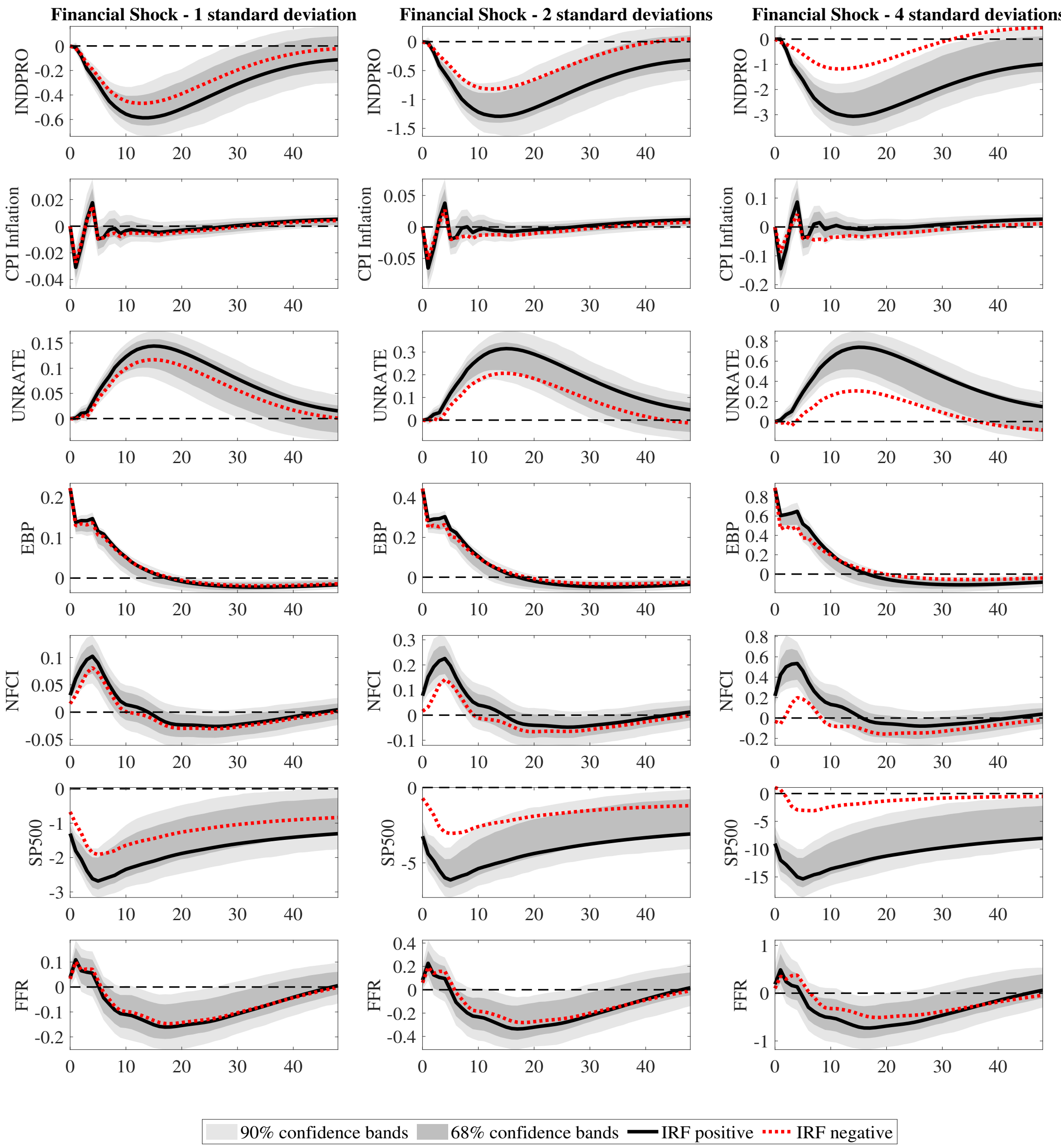

Figure 4: Impulse response functions of the VARX using $g(u_{ft}) = u_{ft}^2$. Black solid lines are the point estimates, the gray areas the 68% and 90% confidence bands of the bad financial shock. Red dotted lines are the point estimates of the good financial shock.

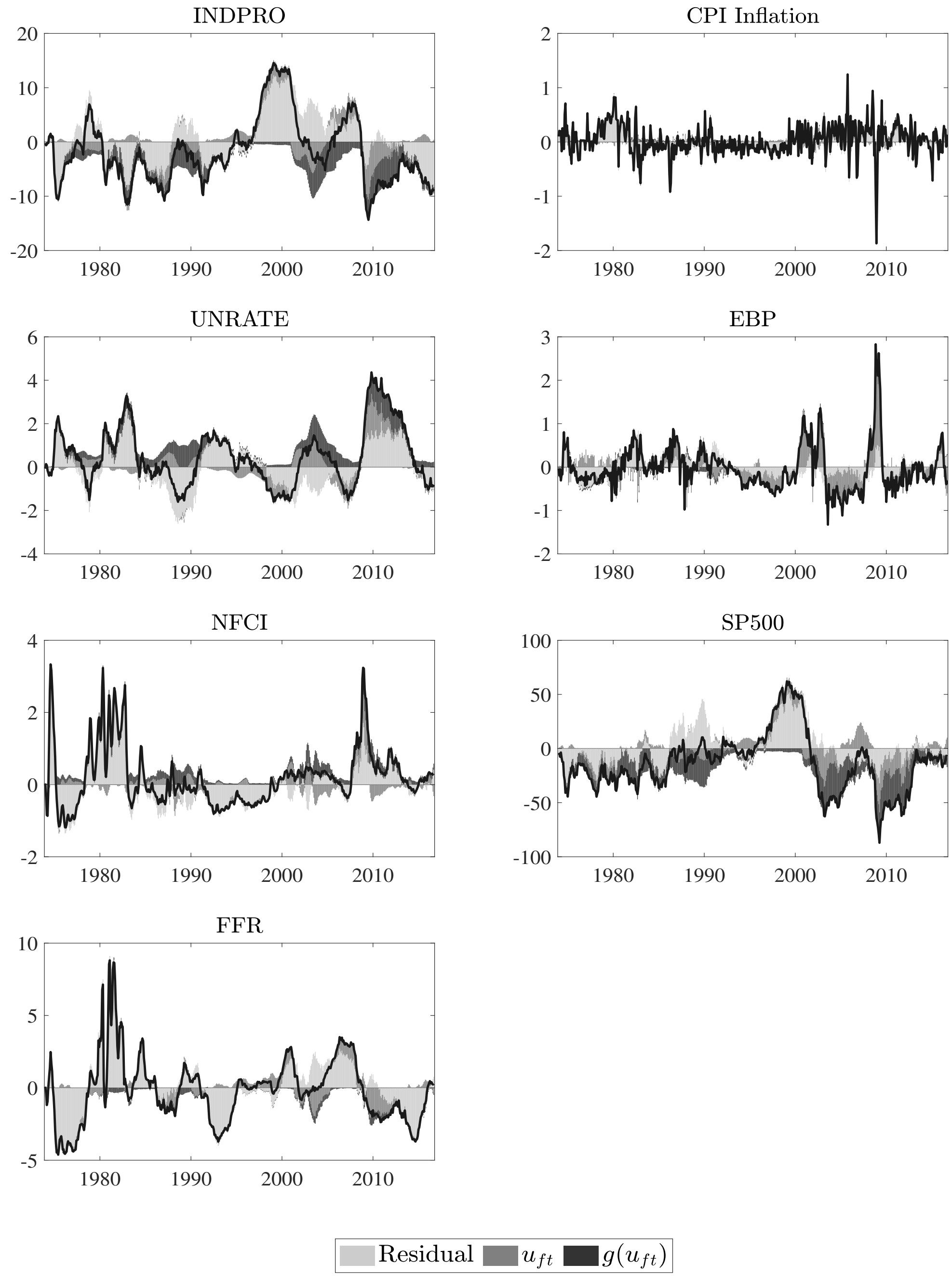

Figure 5: Historical decomposition using $g(u_{ft}) = u_{ft}^2$.

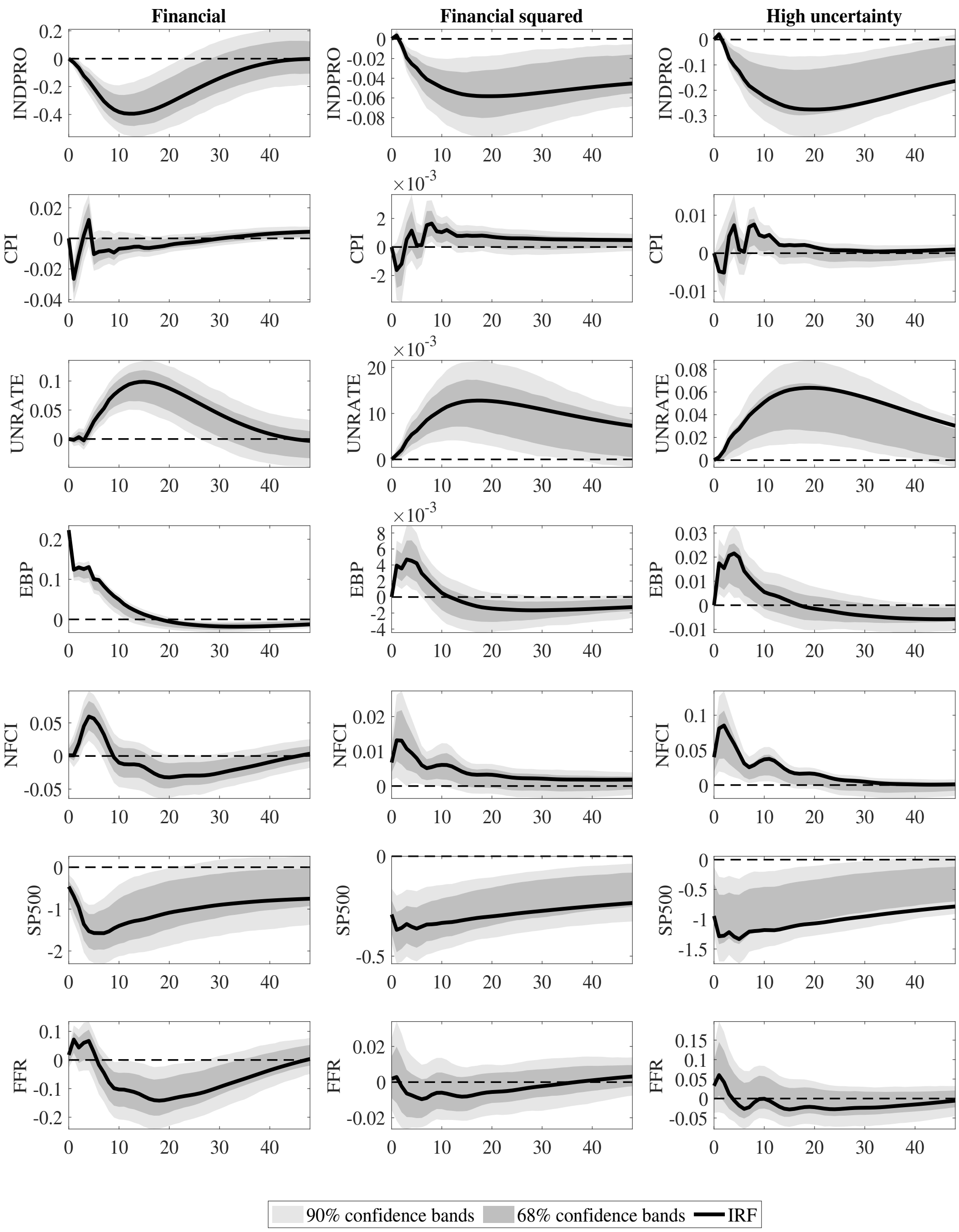

Figure 6: Impulse response functions of the VARX using $g(u_{ft}) = u_{ft}^2$ and $d_t = 1$ if macroeconomic uncertainty is high. Black solid lines are the point estimates, the gray areas the 68% and 90% confidence bands of the financial shock.

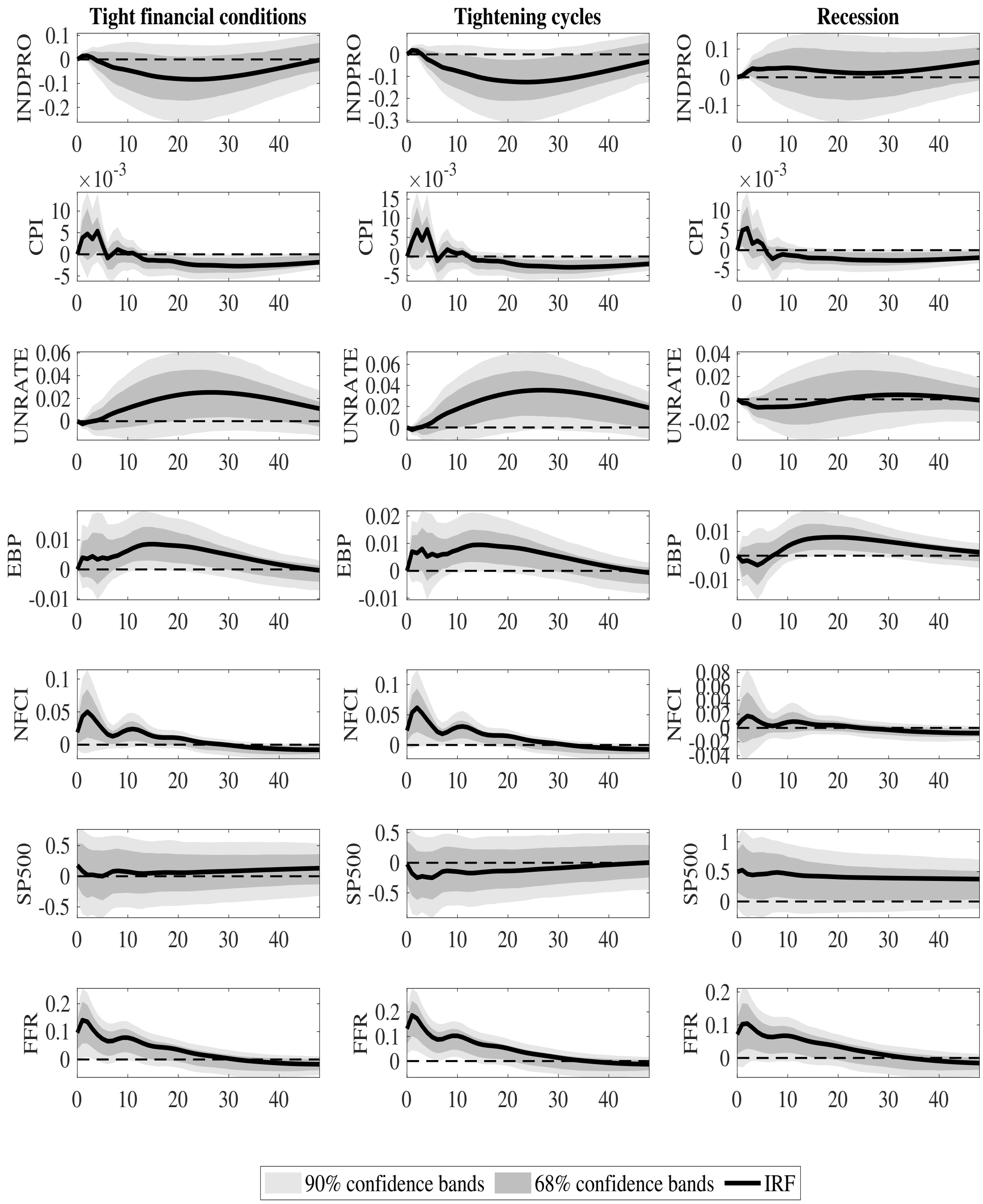

Figure 7: Impulse response functions of the VARX using $g(u_{ft}) = u_{ft}^2$ and $d_t = 1$ if financial conditions are tight (first column), interest rates are high (second column) or the economy is in a recession (third column). Black solid lines are the point estimates, the gray areas the 68% and 90% confidence bands of the financial shock.

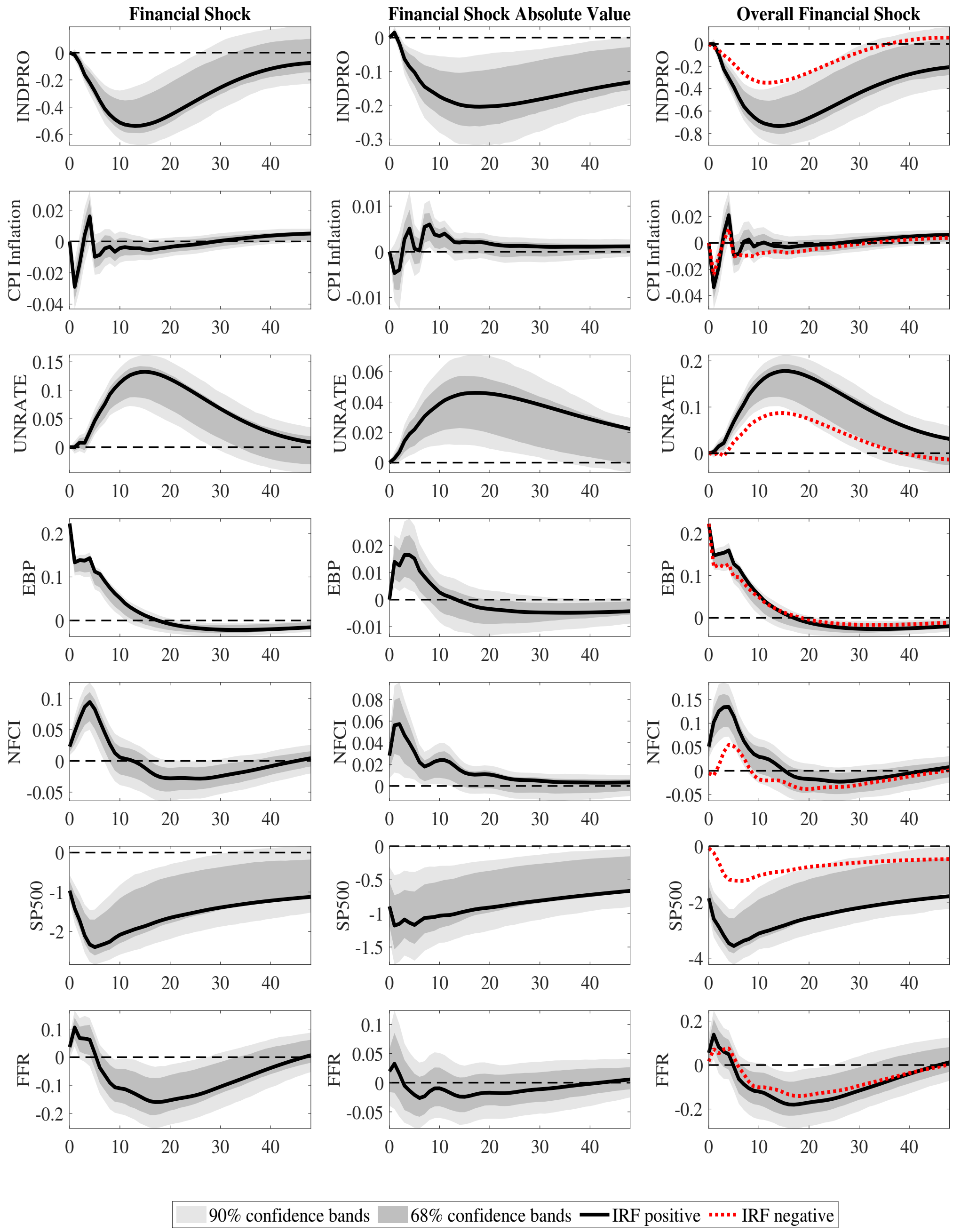

Figure 8: Impulse response functions of the VARX using $g(u_{ft}) = |u_{ft}|$. Black solid lines are the point estimates, the gray areas the 68% and 90% confidence bands of the financial shock.

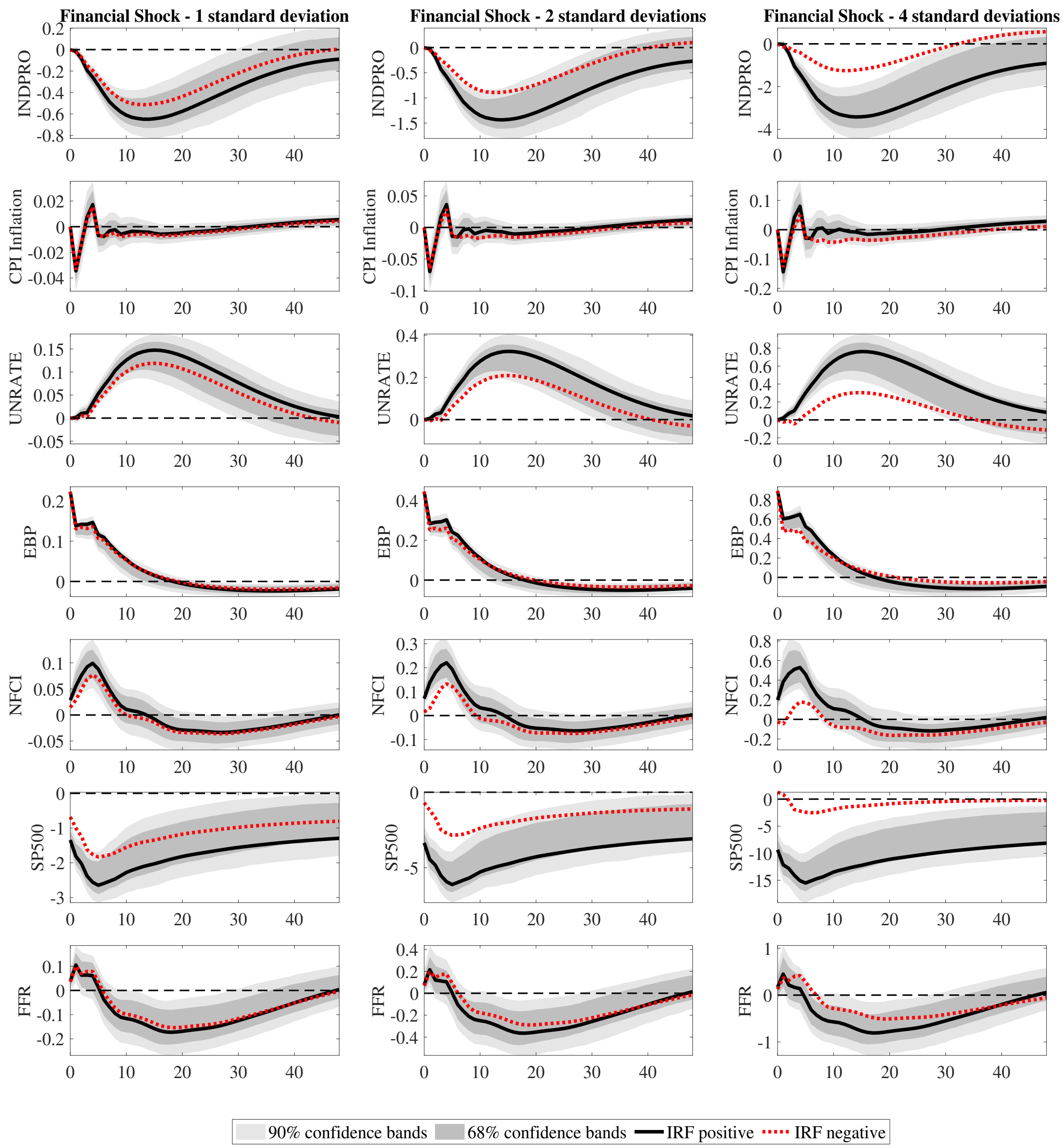

Figure 9: Impulse response functions of the VARX using $g(u_{ft}) = u_{ft}^2$ and 5 lags of the nonlinear term. Black solid lines are the point estimates, the gray areas the 68% and 90% confidence bands of the bad financial shock. Red dotted lines are the point estimates of the good financial shock.

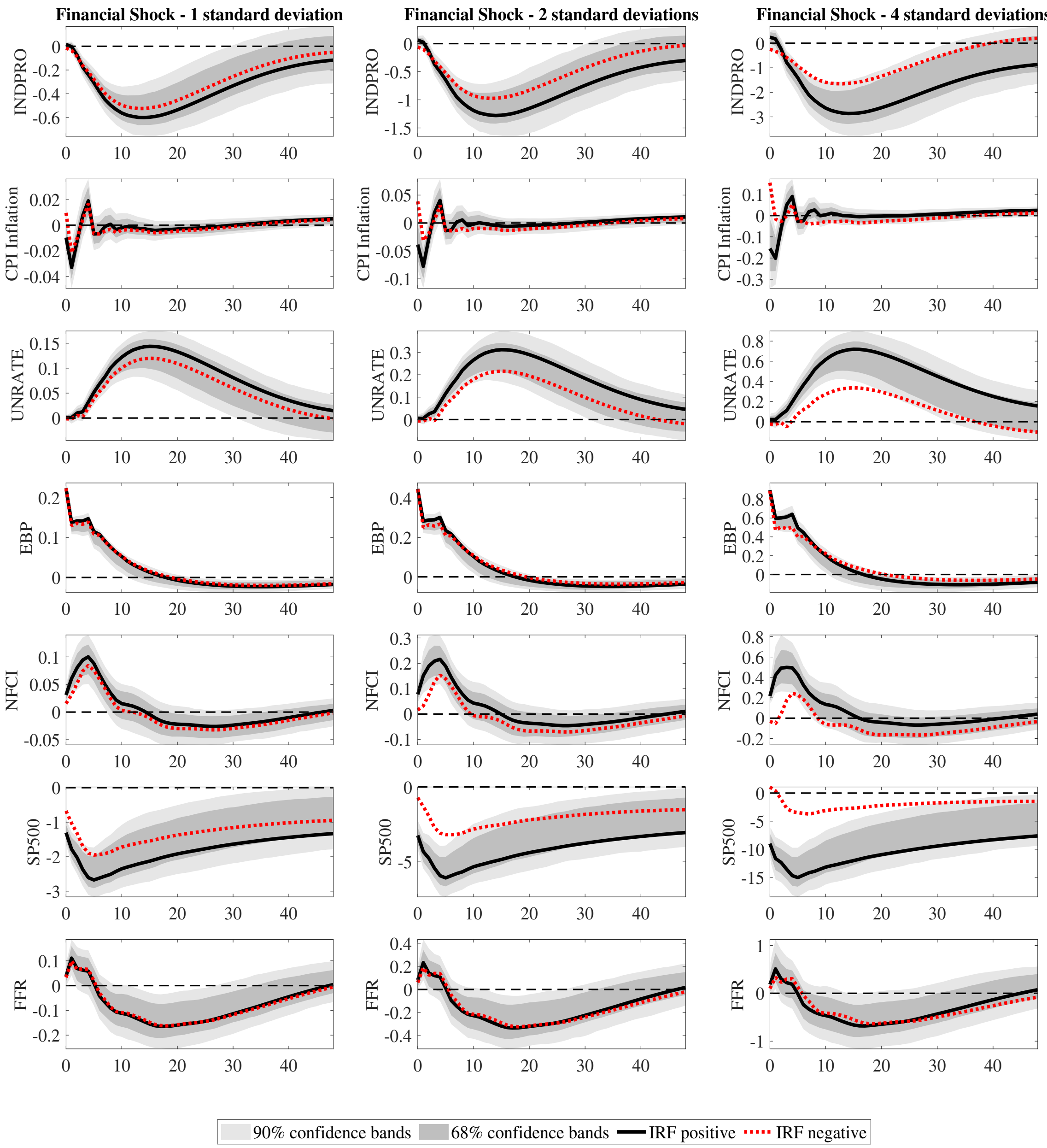

Figure 10: Impulse response functions of the VARX using $g(u_{ft}) = u_{ft}^2$ and leaving the impact response of INDPRO, CPI, UNRATE and EBP to the nonlinear term unrestricted. Black solid lines are the point estimates, the gray areas the 68% and 90% confidence bands of the bad financial shock. Red dotted lines are the point estimates of the good financial shock.

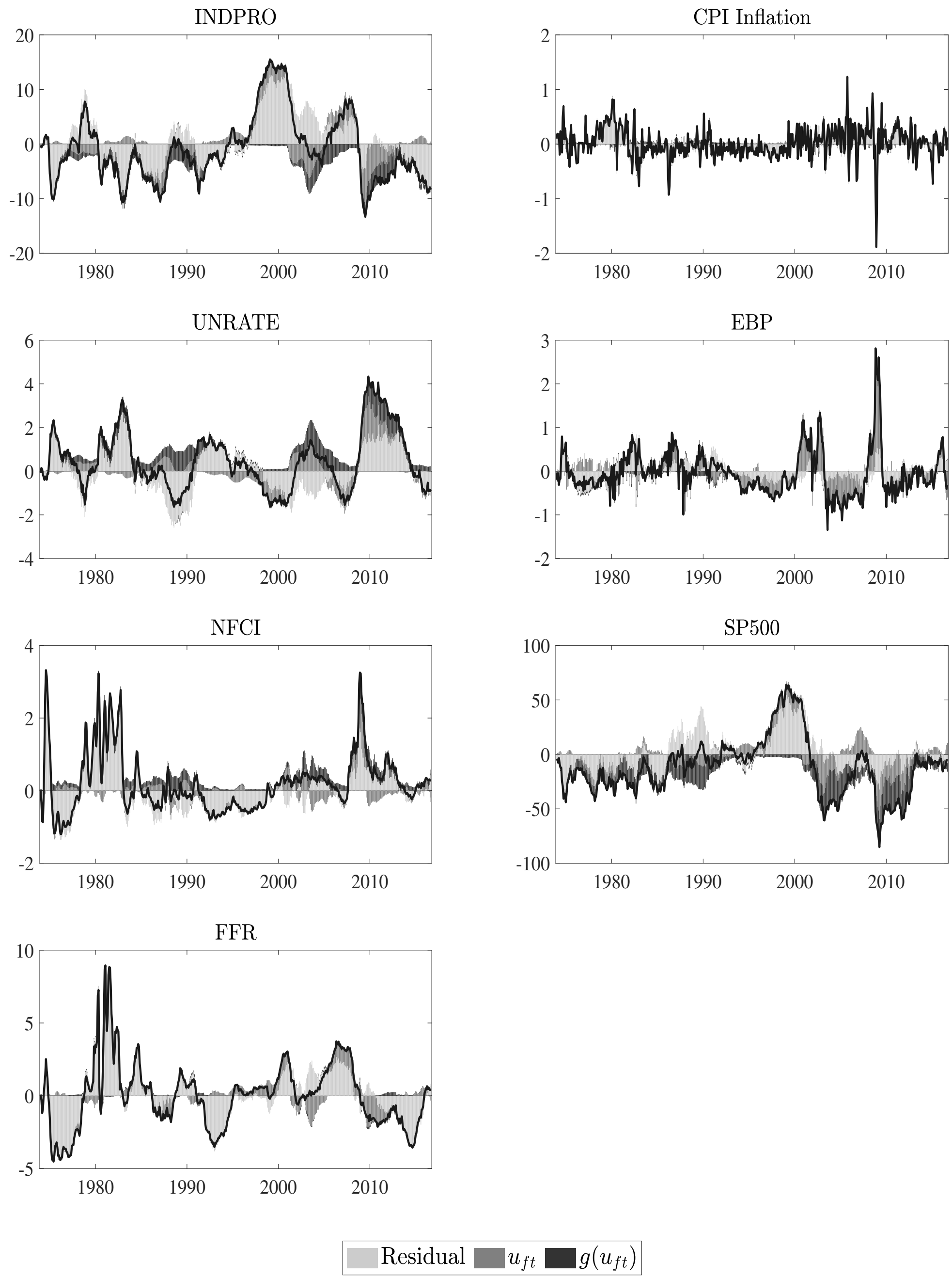

Figure 11: Historical decomposition using $g(u_{ft}) = u_{ft}^2$ and leaving the impact response of INDPRO, CPI and UNRATE to the nonlinear term unrestricted.