House Prices and Aggregate Markups: A VAR Approach

Pablo Lara Hinojos*– CUNY Graduate Center

February 2021

Abstract

Based on the empirical results of Stroebel and Vavra (2019), who find that a local increase in house prices translates into higher local retail markups and prices, I investigate whether this relationship holds at the aggregate level. I first construct a quarterly aggregate markup series and find that indeed house prices and aggregate markups have a positive relationship. This result emerges from a Vector Autoregression (VAR) system that includes the nominal interest rate, the inflation rate, aggregate house prices, aggregate markups and real output. I use the Impulse Response Functions (IRFs) to understand the joint behavior of the variables in the system. I find that this positive relationship between house prices and aggregate markups is robust to different specifications; different time periods; different methods of estimating the responses, such as local projections and generalized IRFs; and a Vector Error Correction Model (VECM) estimation.

1 Introduction

Stroebel and Vavra (2019) find a positive effect of local house prices on the final price that retailers in that same location charge. That is, as house prices increase so do local retail prices. Importantly,

*I am greatly indebted to Rodolfo Oviedo, who acted as my mentor during my stay at Banco de México. His comments and suggestions helped to greatly improve the paper. I would also like to acknowledge the help of Meng-Ting Chen, who provided me with an example for the code used to estimate the necessary parameters for the computation of markups. Additionally, I would like to thank Sangeeta Pratap and Nicolás Amoroso Plaza for very useful comments and suggestions; as well as the seminar participants for useful comments.

this only happens in zip codes with relatively high homeownership rates. The authors analyze various possible explanations for this phenomenon such as: an increase in local rents; an increase in local labor costs; income gentrification[1]; among others. They conclude that this phenomenon is explained by an increase in the markups over marginal cost that retailers set.

In the context of monopolistic competition, the optimal pricing strategy sets the final price equal to the marginal cost times a markup. Furthermore, if the final good is a Dixit-Stiglitz aggregate of the different varieties over which producers compete monopolistically, the optimal markup is a decreasing function of demand elasticity[2]. As house prices increase, the value of housing wealth for homeowners also increases, which does not happen for renters. This could be causing a *wealth effect* or a *relaxation of collateral constraints* that would reduce the price sensitivity of homeowners, leading retailers to optimally increase their markups in response.

These results may have important implications for the aggregate economy and for business cycle fluctuations. In particular, while most medium-scale New Keynesian models such as Christiano, Motto, and Rostagno (2014) or Justiniano, Primiceri, and Tambalotti (2010, 2011) do include time-varying markups, this variability is exogenous. This results in countercyclical sticky-price markups in these models, which might not be appropriate, given that the data is mixed in terms of the cyclicality of markups[3]. Furthermore, assuming exogenous or constant markups may be subject to the Lucas critique. The findings of Stroebel and Vavra (2019) imply that after an expansionary monetary policy, markups should increase, thereby accelerating inflationary dynamics. Current models do not take this possible channel into account. Distinguishing between two sources of inflation may allow to adapt the optimal policy response accordingly. Finally, as Iacoviello (2011) describes, housing wealth represents more than 50% of the total wealth for the typical household in the U.S., so any change in its value has sizable effects on the macroeconomy. The effect that housing wealth may have on aggregate markups represents an additional channel through which it

[1]Meaning increasing average incomes per location.

[2]Specifically, if the final good Y_t is a Dixit-Stiglitz aggregate of the continuous varieties $y_t(z)$ indexed by $z \in [0,1]$, then the production of the final good will be given by the expression: $Y_t = \left(\int_0^1 y_t(z) dz^{\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}}$, where $\varepsilon > 1$ is demand elasticity. This in turn implies that the optimal markup will be equal to: $\mathcal{M} = \frac{\varepsilon}{\varepsilon-1}$. As $\varepsilon \to 1$, the optimal markup tends to infinity.

[3]As Nekarda and Ramey (2013) discusses, the empirical relationship between markups and output is not clear and depends on the way markups are measured. When they are measured as the inverse of the labor share, they are mildly procyclical However, when considering more general production functions, the cyclical properties depends on the empirical implementation. They also argue that all New Keynesian models rely on the countercyclicality of markups, which is at odds with the data.

influences aggregate dynamics.

While the possible implications of the findings of Stroebel and Vavra (2019) are important, their results emerge from using disaggregated data, and it is not evident that they should be reflected at the aggregate level. In particular, their analysis relies on the regional differences between zip codes, concretely, on the homeownership rates of each region. At the aggregate level, the forces acting on a specific region may be offset by those occurring in a different one. Additionally, their findings are valid only for retail markups, whose value added represented on average only 5.7% of yearly GDP in the U.S. during the 2005-2020 period[4]. Therefore, a first step before attempting to understand all the implications from the findings of Stroebel and Vavra (2019) is to verify whether or not this relationship holds when using aggregate data.

In this paper, I set out to find whether or not the positive effect of house prices on markups holds at the aggregate level. To do so, I first build a quarterly series of aggregate markups using the methodology of De Loecker, Eeckhout, and Unger (2020) and data from Compustat[5]. With this series, I am able to estimate a VAR system that includes many of the usual variables in this type of analysis. Specifically, I estimate a non-structural VAR that consists of the nominal interest rate, the inflation rate, real aggregate house prices, aggregate markups, and real output. I then use the impulse response functions (IRFs) from this VAR to measure the relationship between the variables. In particular, I am interested in the response of aggregate markups after a positive shock to housing prices.

The IRFs are estimated with various degrees of significance, with some of them being statistically indistinguishable from zero. However, from those IRFs that are significantly different from zero I extract the following three stylized facts:

i) There is a positive effect from house prices into markups and inflation, which is the aggregate counterpart of the results in Stroebel and Vavra (2019).

ii) After a positive markup shock, output decreases and house prices increase slightly.

iii) After an increase in real output, there is an increase in house prices and markups.

These results are robust to using the aggregate markups for the retail sector, as defined by the

[4]Source: U.S. Bureau of Economic Analysis.

[5]De Loecker et al. (2020) estimate yearly markups.

2-digit NAICS codes 44 and 45; to using the CPI inflation rate, instead of GDP Deflator inflation; to separating the sample period into two sub-periods (i.e. 1976Q1 to 2007Q4, and 1991Q2 to 2019Q4); to using the generalized IRFs as developed in Pesaran and Shin (1998); to using local projections as developed by Jordà (2005); and to using a VECM to account for possible nonstationary behavior and cointegrating relationships. All of these robustness exercises are shown in Appendices C to H.

Despite the apparent robustness of the results, they should be taken with some caution. In particular, it is not possible to disentangle an increase in markups that occurs due to an increase in the market power of firms[6] from one that is caused as an optimal response to a decrease in the price sensitivity of households. Furthermore, rationalizing these results requires a theoretical framework, that would also allow to carry out counterfactuals and policy experiments. This constitutes the future research avenues for this paper.

2 Computing Aggregate Markups

The estimation of quarterly markups is done in two steps. First, I use the yearly data from Compustat[7] to estimate the output elasticity of the variable input (i.e. the parameter θ_{it}^V) [8]. This means that while the computation of markups is done in a quarterly fashion, this parameter is estimated in a yearly basis. This is because θ_{it}^V is a parameter related to the technology of production of firms, and the assumption is that it varies across time. Therefore, it seems more sensible to assume that technology changes occur yearly, rather than quarterly. Secondly, I collect the estimates of θ_{it}^V for each year-industry from the previous step, and combine them with quarterly data to obtain a measure of quarterly markups per firm.

Specifically, the computation of each firm's markups follows the methodology of De Loecker et al. (2020). Following the firm's profit maximization problem, the optimal markup for firm i in period t is a function that depends on the total sales, the expenditure in the variable input, and the output elasticity of the variable input. This is given by the following expression:

$$\mathcal{M}_{it} = \theta_{it}^V \frac{P_{it} Q_{it}}{P_{it}^v V_{it}} \tag{1}$$

[6]For instance due to the foreclosure of firms in a particular sector, or due to important M&A operations.

[7]For details about the data used, see Appendix A.

[8]The estimation of this parameter is central to this methodology, and as such requires special attention. The general method to estimate it is explained in the next subsection.

Where θ_{it}^{V} is the output elasticity of the variable input; $P_{it}Q_{it}$ are firm's sales; and $P_{it}^{V}V_{it}$ is the total expenditure in the variable input.

Once each firm's markup is computed, the aggregate markup per period can be constructed as a weighted average of each firm's individual markup. Specifically, the aggregate markup in each period is given by:

$$\mathcal{M}_t = \sum_i m_{it}\mathcal{M}_{it} \tag{2}$$

Where m_{it} is the weight given to each individual firm. As it will become more apparent, these weights can vary between the share of total sales per period of each firm (Sales), the share of total costs per period (TC), or the share of the total expenditure in the variable input (COGS).

2.1 Estimating the Output Elasticity of the Variable Input

For the estimation of the output elasticity of the variable input, I use yearly data from Compustat and follow the general methodology of De Loecker et al. (2020). However, I follow Chen (2020) in order to include an additional restriction on the returns to scale of the production technology of firms, which is in line with the results of (Flynn, Gandhi, & Traina, 2019), who find that including this restriction improves the estimation of markups and reduces their skewness. This also allows me to implement the estimation method of Wooldridge (2009), which estimates the parameter using a GMM approach, instead of relying on the two-stage approach described on De Loecker et al. (2020), which is in turn based on Ackerberg, Benkard, Berry, and Pakes (2007). The purpose of this method is to increase the precision of the estimates[9].

The estimation is done for each industry-year. Where industry is defined at the 2-digit NAICS level. This definition implies a total of 23 different industries. The data is first deflated by the GDP Deflator and all negative observations are removed. Then the ratio of total sales to COGS is computed, and the data is trimmed to drop the top and bottom 1% of observations. Once the estimates for each year-industry are obtained, I save them in order to be used in the computation of quarterly markups.

[9]This is explained with more detail on Appendix B.

2.2 Estimating Quarterly Markups

The data is transformed in a similar manner to the previous case. That is, I first drop all the negative observations. I then convert the data expressed in CAN into USD by using the quarterly USD-CAN exchange rate. I proceed to transform the nominal quantities into real ones by deflating them with the appropriate GDP Deflator. I then calculate the ratio of total sales to total expenditures in the variable input (i.e. $\frac{P_{it}Q_{it}}{P_{it}^V V_{it}}$). This allows me to trim the data by dropping the top and lowest 1% individual firms (for this step I use the yearly sum of both the numerator and the denominator quantity variables).

Once the data is trimmed, I load the resulting estimates for the output elasticity of the variable input. I merge them by year-industry[10] and use the expression given by (1) in order to obtain a measure of each individual firm's markup in each period.

I then construct three aggregate series using the three possible weights: share of total cost per period (TC), share of total COGS per period (COGS) share of total sales per period (Sales). For the computation of total expenditures, I follow De Loecker et al. (2020) and define it as the sum of expenditures in the variable input (COGS), general expenditures (XSG&A), and capital expenditures (i.e. $TC_t = COGS_t + XSGA_t + R_t K_t$). The latter is computed as the user cost of capital multiplied by the total capital stock (as measured by the variable PPEGT). In turn, the user cost of capital R_t is defined as $R_t = I_t - \Pi_t$; where I_t is the nominal interest rate (as measured by the federal funds rate in quarterly terms), Π_t is the quarterly inflation rate as measured by the GDP Deflator, adjusted for the change in the price of investment goods relative to consumption goods (PIRIC)[11].

2.3 Seasonal Adjustment

Once the three series have been obtained, it becomes evident that they exhibit strong seasonal patterns. Therefore, it is necessary to seasonally adjust the data. To this end, I use the X-13 method recommended by the Census Bureau. With this transformation, the data is ready to be

[10] That is, this parameter does not change within a given calendar year, during which the four observations are identical. They also do not vary across firms within the same industry.

[11] In De Loecker et al. (2020) the authors use the equation: $R_t = I_t - \Pi_t + \Delta$ for the user cost of capital (COC). Where Δ is considered the risk premium, which is set to a constant 12% in yearly terms. Since this is a constant, omitting from the computation of the user COC does not alter the results significantly.

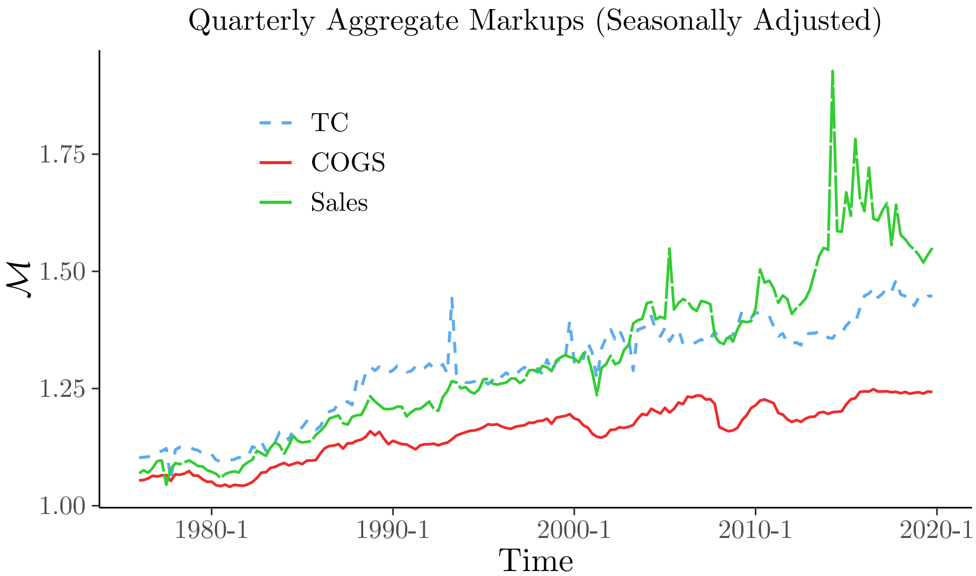
used.

The results of this exercise are shown in Figure 1. As can be seen, all three series exhibit a similar behavior to the one documented by De Loecker et al. (2020). That is, before the 1980s, there was a modest decline in aggregate markups. However, beginning on the early 1980s, there was a reversal of this trend and a steady increase in the level of aggregate markups ensued. All three series exhibit this same behavior, albeit in different magnitudes.

The sales-weighed aggregate markups exhibit a more erratic trend. Particularly from 2000 onward. This is in line with the findings of De Loecker et al. (2020), which show that sales-weighed markups exhibit more variation than total cost-weighed markups. This volatility seems to be exacerbated by the seasonal adjustment of the data.

In the analysis that follows, I choose the COGS-weighted series for various reasons. First, as argued in Edmond, Midrigan, and Xu (2018), firms with market power charge higher prices, decreasing the quantity demanded for their products. A lower demand implies that the quantity sold and the inputs used to produce are also lower, while revenue might remain relatively high for those firms with higher levels of market power. This results in higher revenue weights being allocated to larger firms, relative to the input weights. This means that choosing sales-weighted markups might be overstating the importance of larger firms and biasing the results upwards. This leaves the series of TC and COGS as possible options; however, given that the TC-weighted series seems to have a more volatile behavior, the prudent choice is to select the COGS-weighted series. As Figure 1 indicates, the upward trend of COGS-weighted markups is relatively modest and it seems to exhibit less cyclical volatility than the other series. This implies that any results emerging from this analysis constitute a lower bound on the aggregate relationship between house prices and markups.

Figure 1: Quarterly Aggregate Markups (various weights)



The figure shows the three series for aggregate markups across time. The frequency is quarterly and the sample period goes from 1976Q1 to 2019Q4. The data has been seasonally adjusted using the X-13 method developed by the Census Bureau. Each series results from giving a different weight, m_{it}, to each individual firm's markup, where each individual weight can be either the firm's share of total cost per period (TC, dashed line in blue), the firm's share of total expenditures in the variable input per period (COGS, solid line in red), or the firm's share of total sales per period (Sales, broken line in green). This means $m_{it} \in \{TC, COGS, Sales\}$, according to the notation of Eq. (2).

3 VAR

In order to study the aggregate behavior of the variables, I estimate a non-structural VAR with the following variables: the (nominal) interest rate, the inflation rate, aggregate (real) house prices[12], aggregate markups (COGS-weighted), and real output. I will then use the impulse response functions (IRFs) from this VAR system to measure the joint relationship of the variables. However, some of the series need to be transformed before estimating the VAR. In particular, I will use the log-difference of the GDP Deflator as the measure for inflation; and the cycle component of house prices, markups and real output (expressed as percentage deviation from their respective trends). The identification of the structural shocks relies on the Cholesky decomposition of the variance-covariance matrix of residuals.

The choice of this setting is not arbitrary. It is done following the empirical approach in Iacoviello (2005), who uses a VAR approach both to document stylized facts drawn from the data; as well as

[12] The housing price index is deflated by the GDP Deflator, effectively expressing the evolution of house prices relative to the general price level.

to discipline a DSGE model.

3.1 Transforming and Filtering the Data

I keep the beginning-of-period effective federal funds rate as the measure for the interest rate in each quarter. For the inflation rate, I use the log-difference of the GDP Deflator with respect to the previous quarter.

For house prices, I first deflate the house price index by the GDP Deflator. I then take the natural logarithm of the house price index, the aggregate markups (COGS-weighted), and real output. I filter the three series using a Baxter-King method [13] by choosing upper and lower bands of 32 and 2, respectively. Filtering the series causes the loss of 12 observations at the beginning and at the end of the series. In order to avoid losing these observations, I *AR-pad the series* using forecasts and backcasts of and ARIMA(4,1,0) model, as explained in Steiger, Stock, and Watson (2002)[14]. Furthermore, using the natural logarithm of the series guarantees that the cycle is expressed in terms of percentage deviations from the trend.

Figure 2 shows the three original series with the estimated trend for reference. The figure also shows the NBER recession periods shaded in gray. The graphs show that the three variables exhibit an overall upward trend during the available period [15]. Additionally, it is possible to observe that both aggregate markups and house prices seem to increase above their respective trends before a recession, and fall below it after one. The cycle component for these three series is shown in Figure 3.

3.2 Endogenous Variables

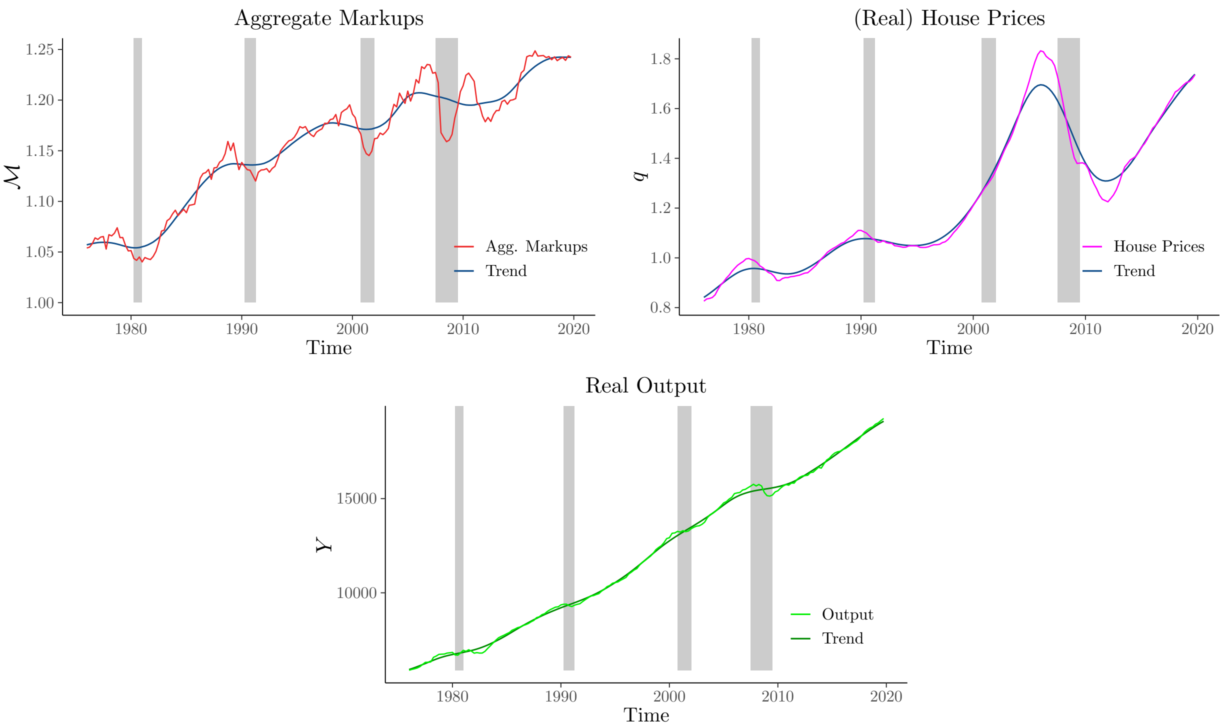
The endogenous variables of the model are: the nominal interest rate R (as measured by the effective federal funds rate); the inflation rate π (as measured by the quarter-on-quarter log difference of the GDP Deflator); real aggregate house prices q (as measured by the cyclical component of the natural logarithm of the aggregate house price index, deflated by the GDPD); aggregate markups $\mathcal{M}$ (as

[13]See Baxter and King (1999).

[14]In summary, this method requires the estimation of an ARIMA(4,1,0) to perform forecasts and backcasts to add 12 observations at the start and at the end of the series.

[15]Each series is available for different ranges and sample periods. Importantly, while at the moment of collecting the data for real GDP was available until the second quarter of 2020, I only use the data from the first available period (1947Q1) until 2019Q4 to avoid biasing the trend downwards with the observations from 2020 which include the effects of the COVID-2019 pandemic.

Figure 2: Aggregate Markups, Real House Prices and Real Output – Series and Trends



The three graphs in the above figure show the actual series and their trend for each variable. The NBER recession periods are shaded in gray. The trend component is obtained by applying a Baxter-King filter with upper and lower limits of 32 and 2, respectively. The VAR uses the cyclical component from the natural logarithm of each series. The time period shown goes from 1976Q1 to 2019Q4. Output is measured in billions of chained 2012 dollars. Importantly, I do not use the full range of data available for real GDP. This is due to the fact that the two first quarters of 2020 were atypical due to the COVID-19 pandemic, and they pull the trend down, implying a significant positive cycle for GDP.

measured by the cyclical component of the natural logarithm of the COGS-weighted aggregate markups); and real output Y (as measured by the cyclical component of the natural logarithm of real output). The units of all the variables are expressed in percentage points.

Figure 3 plots the evolution in time of the five endogenous variables that are included in the VAR.

3.3 Exogenous Variables

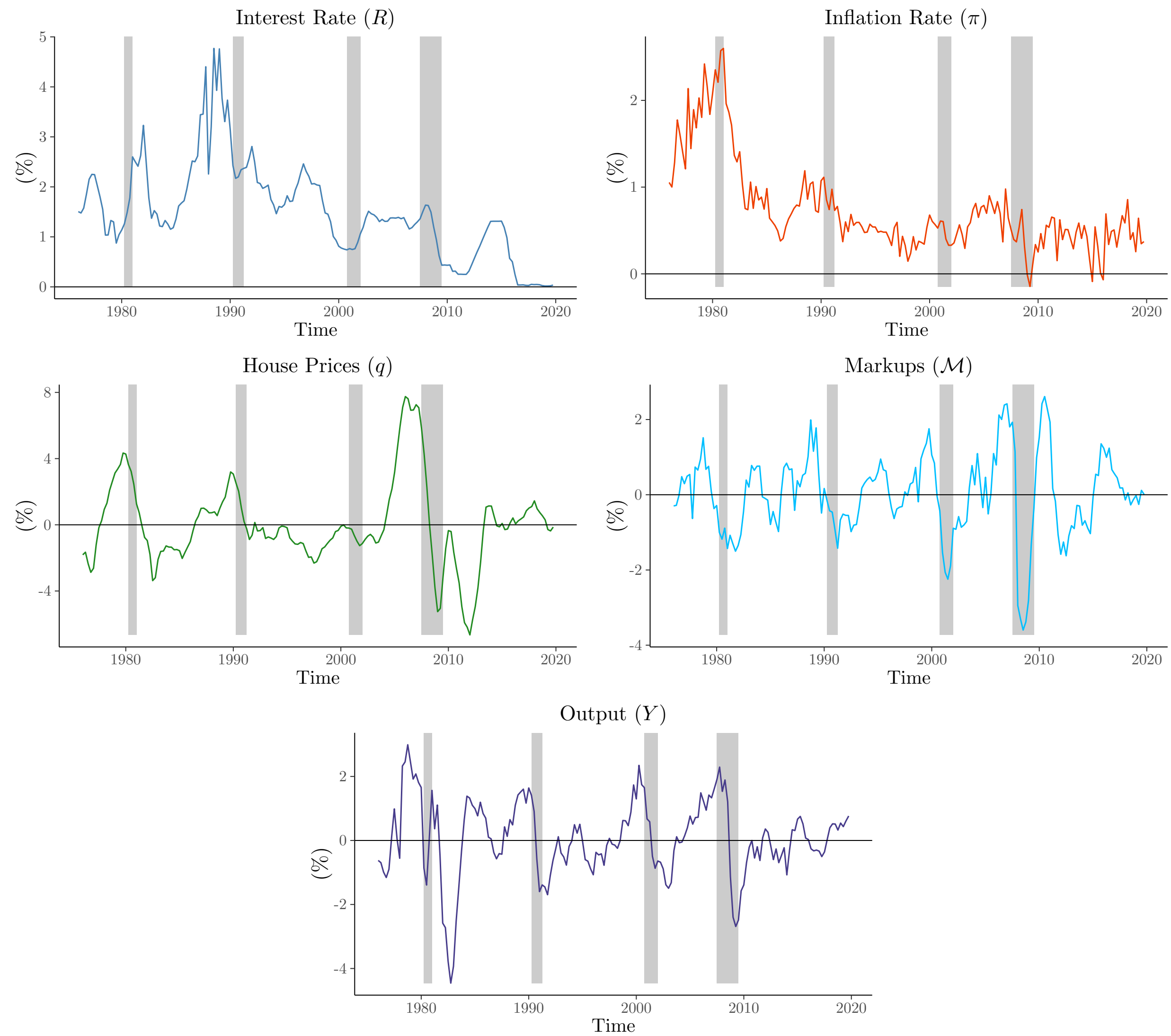
The model is complemented with a set of exogenous variables in order to control for additional factors that may bias the results. In particular, it includes a constant and a linear time trend, to account for the possible non-stationary behavior of the interest rate and the inflation rate.

I also include the natural logarithm of a commodity price index, with a lag, to reduce possible *price puzzle* concerns when plotting the reaction of inflation to the interest rate[16]. I use the World Bank's Commodity Price Index (The Pink Sheet), which is a monthly index for different categories. Among these, the three main categories are energy, non-energy, and precious metals. Within non-energy, there are various sub-classifications such as agriculture, fertilizers, and metals and minerals. I focus only on the energy classification for various reasons. First, according to Sims (1992), the price of oil is one of the main drivers of the price puzzle. Secondly, the behavior of all the rest of the indices is not starkly different from that of the Energy index. Finally, the U.S. economy does not depend on the export or the import of a particular agricultural commodity or metal, so while the general pass-through of commodities might be significant, there is not one single commodity other than oil that requires special attention.

Additionally, I include a dummy to control for the changes in the monetary policy stance that occurred after the tenure of Paul Volcker began. That is, these dummies activate beginning on 1980Q1 to account for the fact that the conduction of monetary policy changed during and after the Volcker period.

[16] A price puzzle refers to the fact that IRFs from VARs often exhibit a positive response from inflation to a monetary shock, when using the interest rate as the monetary variable (as opposed to a monetary aggregate). This result is counter-intuitive, and Sims (1992) argues that this is because the information set available to policy makers at time t is richer than the variables included in the VAR. Among these variables that are available for policy makers is the price of commodities, and in particular oil, which could have an impact on prices. Therefore, what the VAR is capturing is the response of policy makers to the expected rise of inflation. This problem is often solved by including a commodity price index or an oil price index.

Figure 3: Endogenous Variables



The graphs show the five endogenous variables that make up the VAR system. That is the interest rate (R), the inflation rate (π), aggregate house prices (q), aggregate markups ($\mathcal{M}$), and output (Y). They are all expressed in percentage points. The interest rate measures the beginning-of-period monthly effective federal funds rate; the inflation rate is computed as the quarter-on-quarter log difference of the GDP Deflator; and house prices, markups, and output are expressed in terms of their deviation from their respective trends (in percentage). The NBER recession periods are shaded in gray. The time period goes from 1976Q1 to 2019Q4.

3.4 Optimal Lag Selection

All the usual selection criteria point out to an optimal lag structure equal to two. That is, the Akaike Informetion Criterion (AIC), the Hannan-Quinn (HQ), the Schwarz Criterion (SC/BIC), and the Final Prediction Error criterion (FPE) all indicate that the minimum likelihood is achieved when the lag order is equal to $p = 2$. I carry out this test on a maximum of $p = 10$ lags.

3.5 VAR and Cholesky Ordering

Specifically, the $VAR(2)$ that I estimate is given by the following equation:

$$\boldsymbol{y_t} = \boldsymbol{A_1 y_{t-1}} + \boldsymbol{A_2 y_{t-2}} + \boldsymbol{B x_t} + \boldsymbol{\varepsilon_t} \tag{3}$$

Where $\boldsymbol{y_t} = \left(R_t, \pi_t, q_t, \mathcal{M}_t, Y_t\right)'$ is the vector of endogenous variables, which are the interest rate, the inflation rate, the aggregate house prices, aggregate markups, and real output, respectively. Furthermore, since the identification of the structural shock depends on a Cholesky decomposition of the variance-covariance matrix of the residuals, the above ordering of the endogenous variables implies that they are ranked from less to more endogeneity, if the Cholesky decomposition is lower triangular. The ordering is not essential to the main results of this paper.

Alternatively, $\boldsymbol{x_t}$ is a vector of the exogenous variables described above. That is it includes a constant, a linear trend, the lag of the natural logarithm of a commodity price index, and dummies for the change of monetary policy stance in the U.S. The inclusion of the linear trend and the monetary policy dummies responds to the fact that the interest rate and the quarterly inflation rate may be non-stationary.

In this specification, $\boldsymbol{A_1}$ and $\boldsymbol{A_2}$ are 5×5 matrices; $\boldsymbol{B}$ is a 5×4 matrix; and $\boldsymbol{\varepsilon_t} \sim (0, \boldsymbol{\Sigma_\varepsilon})$, with $\boldsymbol{\Sigma_\varepsilon}$ being a 5×5 matrix. By construction, the diagonal elements of $\boldsymbol{A_1}$ and $\boldsymbol{A_2}$ contain the AR components of the endogenous variables with respect to themselves; and the off-diagonal elements contain the influence of the other variables in the system. Finally, the Cholesky decomposition of the estimate for $\boldsymbol{\Sigma_\varepsilon}$ is what identifies the structural shocks and gives the structure to impulse response functions (IRFs) from the next section.

4 Impulse Response Functions

The estimated VAR allows me to obtain the response of the system to an orthogonal shock in each variable. This helps to identify the joint behavior of variables. In particular, given the results of Stroebel and Vavra (2019), I am interested in the response of markups after a shock to house prices. However, the VAR allows me to understand the joint dynamics of the whole system.

Each IRF collects the response of each variable after a one-standard deviation shock. The standard deviation of each variable is measured by the square root of diagonal entries of the variance-covariance matrix of the residuals. By construction, the Cholesky decomposition guarantees that each one of these shocks can be interpreted as an orthogonal structural shock. The confidence intervals for each IRF are computed by running a bootsrap of the IRFs with 1,000 runs.

4.1 Stylized Facts

The responses of the variables to each shock are computed with various degrees of precision. That is, in some cases the upper and lower confidence bounds suggest that the response cannot be distinguished from zero with statistical significance. However, in those cases in which the responses are significantly estimated, I document three main stylized facts:

1. After a shock to housing prices (q), inflation (π) and markups ($\mathcal{M}$) increase significantly.

2. After a shock to markups ($\mathcal{M}$), output (Y) decreases significantly and house prices (q) rise weakly.

3. After a shock to output (Y), markups ($\mathcal{M}$) increase significantly and house prices (q) do so weakly.

These three facts follow from the responses of the system after a shock to house prices, a shock to aggregate markups, and a shock to output, respectively. For completeness sake I present all of the IRFs from this exercise.

Furthermore, the three facts hold under different specifications. For example, when using retail markups instead of aggregate markups[17]; or when using the Consumer Price Index (CPI) for the

[17] Retail markups are calculated by aggregating the observed markups for firms with 2-digit NAICS equal to 44 or 45. See Appendix C for more details.

measure of inflation [18].

4.2 House Prices Shock

Fact 1 is obtained from the responses to a housing shock. Figure 4 shows the response from all the variables in the VAR after such a shock. As can be seen, only the responses of inflation (π) and markups ($\mathcal{M}$) are significantly different from zero. It is possible to observe that the shock to q behaves in a hump-shaped manner, and so do the responses from π and $\mathcal{M}$. This fact would be consistent with the findings of Stroebel and Vavra (2019). Furthermore, such a response could be explained by the same mechanism discussed in their paper. Namely, after a positive house prices shock, there is both a *wealth effect* and a *relaxation of the collateral constraints* for households. This in turn, may decrease price sensitivity, which leads producers to optimally increase their markups in response. The increase in markups then leads to an increase in inflation. The response of markups is positive and statistically significant as house prices increase, and markups fall below trend non-significantly after roughly 7 quarters.

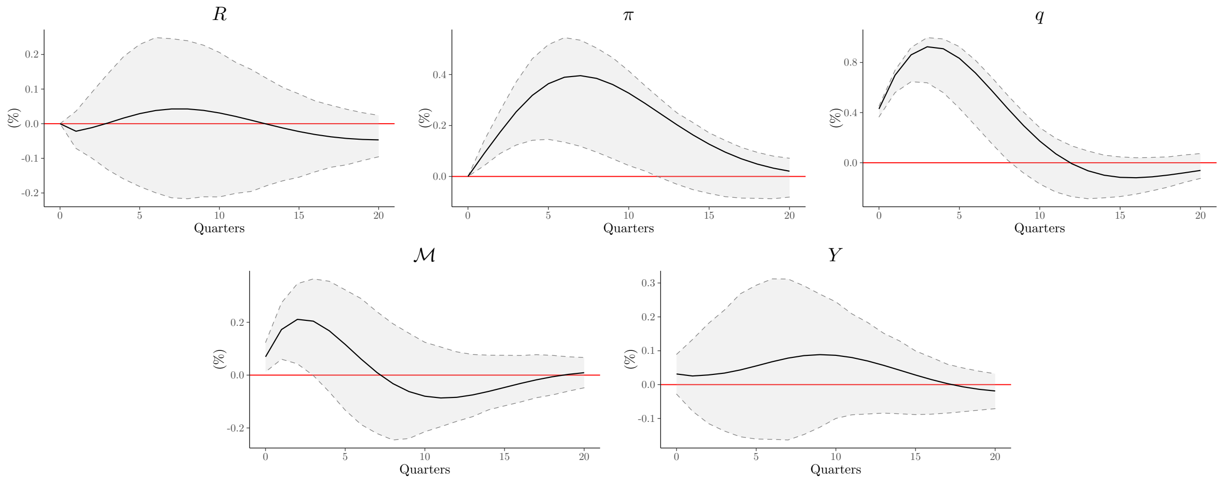
Regarding the responses of the interest rate (R) and real output (Y), neither is estimated significantly different from zero. Both decrease slightly and increase after 6 or 7 quarters. This could be due to the increase in economic activity, which also raises interest rates. However, since these responses are not significant, they are difficult to interpret.

4.3 Markups Shock

Fact 2 emerges from the responses of the system after an orthogonal shock to markups . Figure 5 shows the responses of the VAR to this shock. It can be seen that output (Y) rises non-significantly above its trend upon impact; but after 4 to 5 periods, it significantly falls below trend. In terms of house prices (q) they rise above trend with slight significance, but then the estimated response cannot be significantly distinguished from zero. Such behavior could be explained by looking at the cyclical behavior of the variables in Figure 2, where it can be seen that aggregate markups tend to increase above their trend before a recession, which means they act as a lead indicator of economic activity. House prices also rise above their trend before a recession, so this could explain the co-movement between these two variables and their response to a shock to markups.

[18] See Appendix D for more details.

Figure 4: Shock to House Prices.



The figure shows the responses of the five variables in the VAR after a one-standard-deviation orthogonal shock to house prices (q). The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

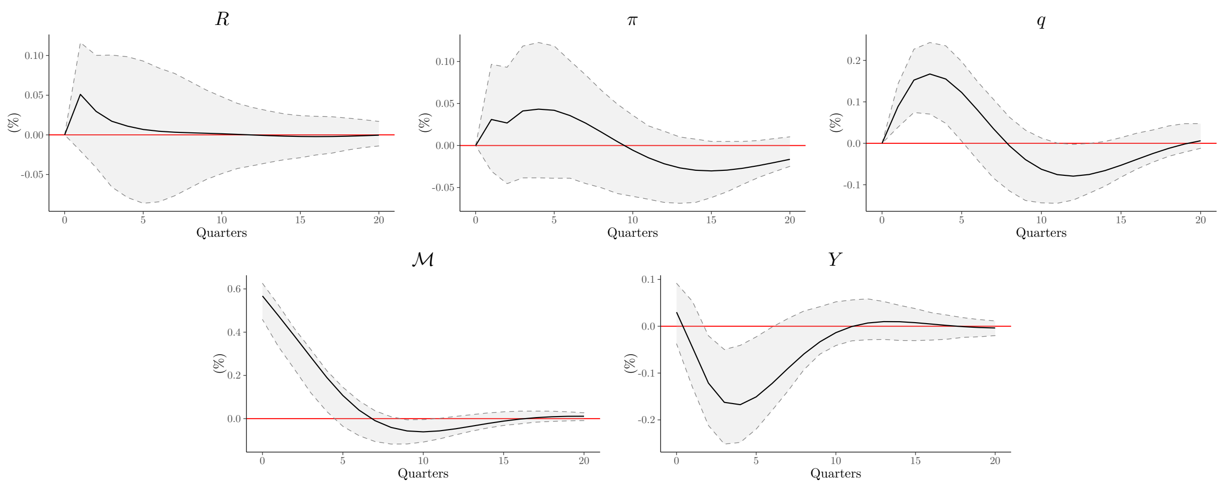
Finally, the responses of the interest rate (R) and inflation (π) cannot be significantly distinguished from zero. The fact that interest rates seem to increase slightly and then fall seem to suggest that the idea of markups acting as a lead indicator would be correct.

4.4 Output Shock

Fact 3 follows from the responses of the VAR to a shock to real output. These responses plotted in Figure 6. After a positive shock to output (Y), both house prices (q) and markups ($\mathcal{M}$) increase. There are a few possible explanations consistent with this behavior. If output increases due to an increase in Total Factor Productivity (TFP), and the supply of housing does not increase proportionally[19], then consumption goods become relatively more abundant than housing, which causes the relative price of housing (q) to increase. This increase in house prices then causes the wealth effect and the relaxation of collateral constraints described before, which decreases the price sensibility of households and drives firms to increase their markups, even as prices should be falling. On the other hand, a positive shock to output can be interpreted as a positive income shock which increases both the demand for housing, and house prices as a response, as well as the markups

[19]This can occur due to there being a fixed supply of housing in the short run; or if the productivity of the technology to produce housing does not react, or does so in a slower fashion, to the shocks to the goods-producing technology. This latter case is explored in Iacoviello and Neri (2010).

Figure 5: Shock to Markups.



The figure shows the responses of the five variables in the VAR after a one-standard-deviation orthogonal shock to aggregate markups ($\mathcal{M}$). The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

that firms can charge. On the other hand, there is a rising literature that shows that larger, more productive firms are able to establish higher markups[20], if the TFP shock affects these firms more intensely, their market power would increase and so would aggregate markups.

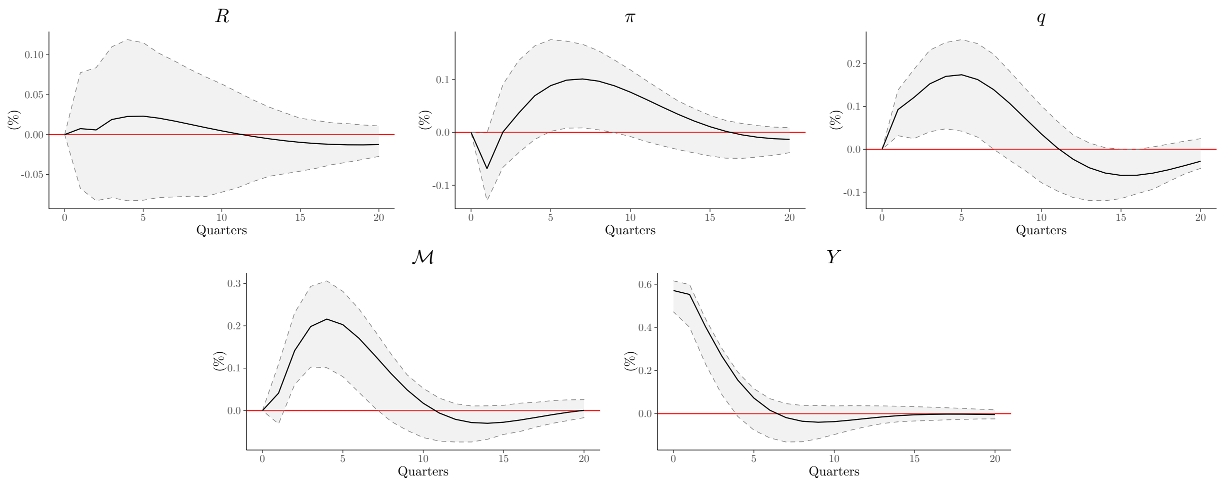
The responses of the interest rate (R) and inflation (π) are not estimated to be significantly different from zero. However, the responses are consistent with the aforementioned stories. In particular, upon a positive TFP shock, prices should decrease, and as house prices and markups increase, so should the price level. Alternatively, most New Keynesian models imply a fall in interest rates after a positive TFP shock, which indeed occurs in the estimated response of the interest rate, albeit non-significantly.

4.5 Interest Rate Shock

The responses of the variables after a shock to interest rate are not significantly estimated different from zero. There is a slight significant fall in markups after an increase in interest rates, but this fall becomes non-significant after one period. The rest of the variables does not react significantly to the interest rates. One possible explanation for this is that the sample period covers the post-2008 period, which has been characterized by very low interest rates, slow recovery in the U.S. and other

[20]See Edmond et al. (2018), for instance.

Figure 6: Shock to real output.



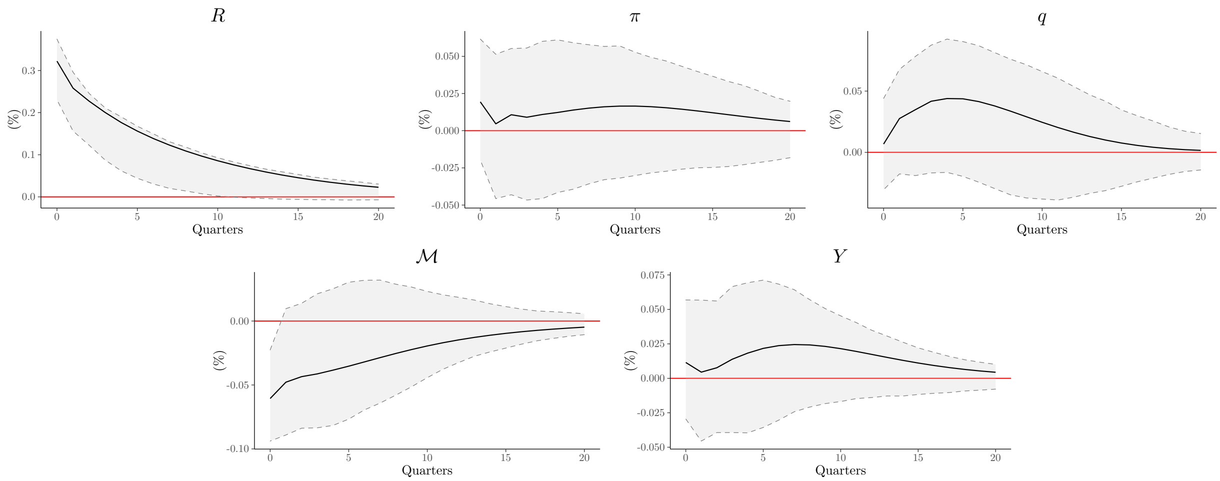
The figure shows the responses of the five variables in the VAR after a one-standard-deviation orthogonal shock to real output (Y). The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

advanced economies, and unconventional monetary policy tools. Figure 7 collects the responses after a shock to R.

4.6 Inflation Shock

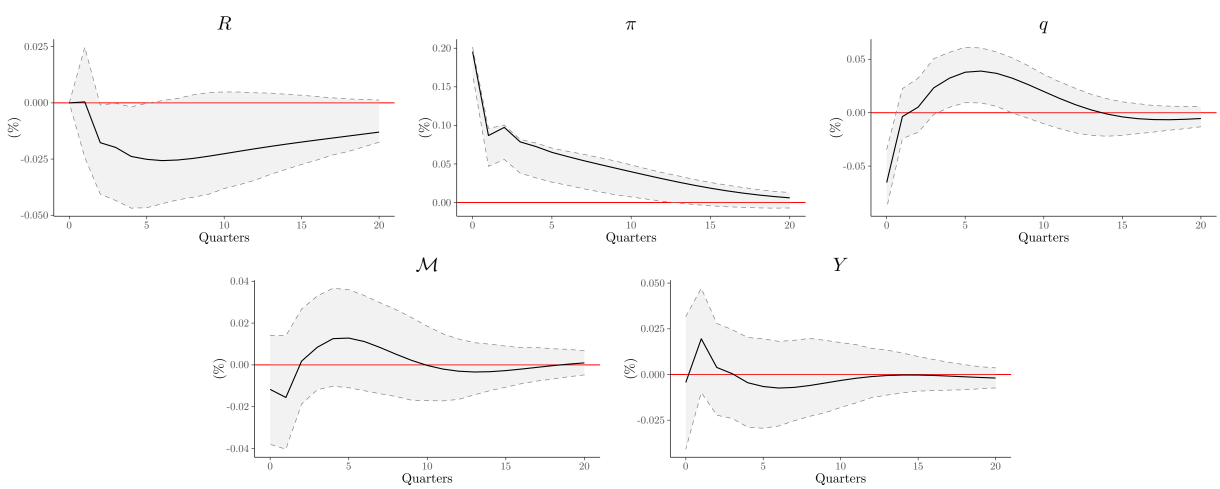
Just as it happened with the responses after a shock in the interest rate (R), the responses of the VAR system after a positive inflation (π) shock are not precisely estimated. After such a shock, house prices decrease significantly and then increase significantly. This initial decrease in house prices after an inflation shock is also documented by Iacoviello (2005). If the Central Bank carries out restrictive monetary policy after the increase in inflation, which is not reflected in an increase in the interest rate, this could explain the observed phenomenon. After four quarters, house prices rise above their trend and so do markups, albeit the significance is not long-lived. Additionally, the responses of the interest rate and output are not significantly different from zero. These responses are shown in Figure 8.

Figure 7: Shock to interest rate.



The figure shows the responses of the five variables in the VAR after a one-standard-deviation orthogonal shock to the nominal interest rate (R). The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 8: Shock to inflation.



The figure shows the responses of the five variables in the VAR after a one-standard-deviation orthogonal shock to the inflation rate (π). The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

5 Concluding Remarks

Following on the results of Stroebel and Vavra (2019), I investigate whether the effect that house prices have on local retail markups (and prices) holds at the aggregate level. To do so, I first construct a quarterly measure of aggregate markups using the methodology of De Loecker et al. (2020) and quarterly data from Compustat. I then use this measure of aggregate markups alongside with other aggregate variables such as the nominal interest rate, the inflation rate, a measure of aggregate house prices, and real output to estimate a VAR. The VAR allows me to compute the Impulse Response Functions (IRFs) after a shock to each one of the variables in order to understand their joint dynamics.

With this methodology, I am able to document three stylized facts. First, after a positive shock to house prices, there is an increase in aggregate markups and in real output, which is evidence in favor of the results of Stroebel and Vavra (2019) holding at the aggregate level. Secondly, after an increase of aggregate markups, real output decreases, while house prices increase. Thirdly, after a positive shock to output house prices and aggregate markups increase. The results hold even when focusing only on the markups of the retail sector, defined at the 2-digit NAICS codes 44 and 45; or when using CPI inflation instead of GDPD inflation.

However, these results should be taken with some caution. In particular, it is not possible to disentangle changes in markups that occur as an optimal response to shifts in demand, from those that happen purely due to supply factors. That is, it is not possible to distinguish between an increase in markups that occurs as the optimal response from a fall in the price sensitivity of consumer; from an increase that occurs solely to an increase in market power.

One way to be able to rationalize these results is to integrate them into a Dynamic Stochastic General Equilibrium (DSGE) model in order to be able to carry out counterfactuals and policy experiments. This constitutes the next avenues for my research. I intend to extend the New Keynesian model in Iacoviello (2005) for this purpose, by allowing markups to respond to house prices. I will also use the IRFs from this paper to discipline the model.

References

Ackerberg, D., Benkard, C. L., Berry, S., & Pakes, A. (2007). Econometric tools for analyzing market outcomes. *Handbook of Econometrics*, *6*, 4171–4276.

Baxter, M., & King, R. G. (1999). Measuring business cycles: approximate band-pass filters for economic time series. *Review of Economics and Statistics*, *81*(4), 575–593.

Bernanke, B. S., Gertler, M., & Gilchrist, S. (1999). The financial accelerator in a quantitative business cycle framework. *Handbook of macroeconomics*, *1*, 1341–1393.

Chen, M.-T. (2020). *Markups, labor share, and wage dispersion.* Mimeo, Department of Economics, CUNY - Graduate Center.

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

De Loecker, J., & Eeckhout, J. (2017). *The rise of market power* (Tech. Rep.). mimeo.

De Loecker, J., Eeckhout, J., & Unger, G. (2020). The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics*, *135*(2), 561–644.

Edmond, C., Midrigan, V., & Xu, D. Y. (2018). *How costly are markups?* (Tech. Rep.). National Bureau of Economic Research.

Flynn, Z., Gandhi, A., & Traina, J. (2019). Measuring markups with production data. *Available at SSRN 3358472*.

Hall, R. E. (2018). *New evidence on the markup of prices over marginal costs and the role of mega-firms in the us economy* (Tech. Rep.). National Bureau of Economic Research.

Iacoviello, M. (2005). House prices, borrowing constraints, and monetary policy in the business cycle. *American economic review*, *95*(3), 739–764.

Iacoviello, M. (2011). Housing wealth and consumption. *FRB International Finance Discussion Paper*(1027).

Iacoviello, M., & Neri, S. (2010). Housing market spillovers: evidence from an estimated dsge model. *American Economic Journal: Macroeconomics*, *2*(2), 125–64.

Ireland, P. N. (2004). A method for taking models to the data. *Journal of Economic Dynamics and Control*, *28*(6), 1205–1226.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica: Journal of the Econometric Society*, 1551–1580.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review*, *95*(1), 161–182.

Justiniano, A., Primiceri, G. E., & Tambalotti, A. (2010). Investment shocks and business cycles. *Journal of Monetary Economics*, *57*(2), 132–145.

Justiniano, A., Primiceri, G. E., & Tambalotti, A. (2011). Investment shocks and the relative price of investment. *Review of Economic Dynamics*, *14*(1), 102–121.

Kiyotaki, N., & Moore, J. (1997). Credit cycles. *Journal of Political Economy*, *105*(2), 211–248.

Liu, Z., Wang, P., & Zha, T. (2019). *A theory of housing demand shocks* (Tech. Rep.). National Bureau of Economic Research.

Mian, A., Rao, K., & Sufi, A. (2013). Household balance sheets, consumption, and the economic slump. *The Quarterly Journal of Economics*, *128*(4), 1687–1726.

Nekarda, C. J., & Ramey, V. A. (2013). *The cyclical behavior of the price-cost markup* (Tech. Rep.). National Bureau of Economic Research.

Olley, S., & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, *64*.

Pesaran, H. H., & Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters*, *58*(1), 17–29.

Piazzesi, M., & Schneider, M. (2016). Housing and macroeconomics. In *Handbook of macroeconomics* (Vol. 2, pp. 1547–1640). Elsevier.

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

Sims, C. A. (1992). Interpreting the macroeconomic time series facts: The effects of monetary policy. *European Economic Review*, *36*(5), 975–1000.

Steiger, D., Stock, J. H., & Watson, M. W. (2002). Prices, wages, and the u.s. nairu in the 1990s. In A. B. Krueger & R. Solow (Eds.), *The roaring nineties: Can full employment be sustained?* (p. 1-8). New York: Russell Sage Foundation.

Stroebel, J., & Vavra, J. (2019). House prices, local demand, and retail prices. *Journal of Political Economy*, *127*(3), 1391–1436.

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, *104*(3), 112–114.

A Data

The data can be separated in two main groups. First, the data relating to firms' financial statements, which comes from Compustat and is not publicly available, is used to estimate firm-level markups and aggregate markups per period of time. Second, the aggregate data available publicly, which is later used in the VAR alongside the measure of aggregate markups obtained with Compustat data. I will therefore describe both groups separately.

A.1 Firm-Level Data

Compustat is available both in yearly and in quarterly frequencies. I use both periodicities for two reasons. I first use yearly data to estimate the output-elasticity of the variable input (θ_{it}^V), following the methodology of De Loecker et al. (2020). I compute estimates of this parameter for each year-industry, the latter defined at the two-digit level according to NAICS. I then use these estimates to compute quarterly markups per firm by using the quarterly data to compute the appropriate ratios. The implicit assumptions are that the output elasticity of the variable input changes yearly, as opposed to quarterly; and that each period, this parameter is common across firms in the same industry[21].

To use the estimation strategy in De Loecker et al. (2020), I follow their choice of variables [22]. Specifically, I use the following variables from the yearly version of Compustat:

These variables will allow me to estimate the output elasticity of the variable input θ_{it}^V by year-industry. The sample period goes from 1950 to 2019.

Alternatively, in order to compute the markups, I need to observe both the total sales for the firm as well as the amount spent on the variable input (COGS), and the amount of capital the firm

[21] Both assumptions are also required in De Loecker et al. (2020). The first one is true, by construction, whereas the second one is necessary to carry out the estimation of the output elasticity of the variable input.

[22] I benefited greatly from the author's own codes, available on Jan De Loecker's personal website: https://sites.google.com/site/deloeckerjan/data-and-code

Table 1: Summary of variables from Compustat (Yearly)

	Acronym in Compustat	Variable	Mean	Median	N
Cost of Goods Sold	COGS	V	1,038,311	40,272.50	443,592
Sales	Sale	PQ	1,516,136	66,782	449,321
Selling, General & Administrative Exp.	XSG&A	X	256,575.4	14,222	358,331
Capital Stock	PPEGT	K	1,360,577	26,656	409,758

has, in order to calculate the total expenditures in capital and equipment by period.

Table 2: Summary of variables from Compustat (Quarterly)

	Acronym in Compustat	Variable	Mean	Median	N
Cost of Goods Sold	`cogsq`	V	318,593.8	12,690	1,443,408
Sales	`saleq`	PQ	441,879	21,750	1,533,902
Selling, General & Adm. Exp.	`xsga`	X	78,048.88	5,075	1,129,504
Capital Stock	`ppegtq`	K	1,570,087	35,820	907,193

The time period goes from January 1961 to December 2019. However, the data quality before 1976 is very poor, so I have to restrict my sample period from 1976Q1 to 2019Q4.

A.2 Aggregate Data

I use various aggregate data sources. In particular, I am interested in series for the interest rate (both for the VAR, and to calculate the user cost of capital); the inflation rate (similarly, for the VAR and for the user cost of capital); the relative price of investment goods with respect to consumption goods; the USD-CAN exchange rate; real output; a housing price index; and a commodity price index. I will describe each one separately. While the data comes from different sources, most of them are downloaded directly from the Federal Reserve of St. Louis Economic Data (FRED) website. There are some exceptions, but they are properly indicated.

Interest Rate

In terms of interest rates, I use the monthly effective federal funds rate. I use the beginning-of-period observation for each quarter. The data comes from the FRED (ID: FEDFUNDS), and I use first available time period (July, 1954) until the second trimester of 2020 (June, 2020). The data is used both to calculate the user cost of capital following De Loecker et al. (2020), and as input in the VAR calculations.

Inflation Rate

In order to obtain a series for the inflation rate, I use the quarterly GDP Deflator calculated by the Bureau of Labor Statistics (BLS). Additionally, I also use this series to deflate the data obtained from Compustat in order to use real quantities. The data is downloaded from FRED (ID: GDPDEF). The sample period goes from January 1947 to June 2020.

Relative Price of Investment

The relative price of investment goods with respect to consumption goods is used to obtain a measure of the user cost of capital, following De Loecker et al. (2020). The data is available at quarterly frequency. I download them from FRED (ID: PIRIC) and use the full available sample period, which goes from 1947Q1 to 2018Q4.

USD-CAN Exchange rate

Since Compustat includes firms in Canada[23], in order to keep these data, I transform their figures into USD dollars. I download the monthly USD-CAN exchange rate from FRED (ID: EXCAN) for the available sample period (January 1971 to June 2020).

Real Output

I use the Real Gross Domestic Product, expressed in Chained 2012 dollars. The data is available on a quarterly basis. I obtain the series from FRED (ID: GDPC1) and use the available sample period, which goes from 1947Q1 to 2019Q4. Although the observations for the first two quarters of 2020 was available at the time of the collection of the data, I omit its use due to the bias in the trend component that occurs as a result of the drop in output during the COVID-19 pandemic. The results do not change significantly if I use the full available period.

Housing Price Index

For the housing price index, I use Freddie Mac's monthly House Price Index for the U.S. The data is seasonally adjusted and published on the website of Freddie Mac. The available sample period

[23] In the sample drawn from Compustat, 177,672 (9.89%) observations use CAN as their listed currency (`curcd` in Compustat).

goes from January, 1975 to June, 2020.

Commodity Price Index

The Commodity Price Index is used in the estimation of the VAR to reduce concerns of price puzzles, as discussed originally in Sims (1992). I use the World Bank's Commodity Price Index (The Pink Sheet). It collects monthly indices of commodities in different categories. I focus on the Energy category, but this choice does not affect the results, as the rest of the indices behave in a similar fashion. The indices use 2010 as the base year, and the available sample period goes from January 1960 until present.

B Estimating Output Elasticity

The following exposition follows De Loecker et al. (2020) and Chen (2020) closely. The estimation of the output elasticity of the variable input relies on running for each industry s, and each firm i, the the following regression:

$$y_{it} = \theta_t^v v_{it} + \theta_t^k k_{it} + \omega_{it} + \epsilon_{it}$$

Where y_{it} is the log of a measure of firm's output; v_{it} represents the log of deflated expenditure in variable inputs; k_{it} the log of the firm capital; ω_{it} is a productivity shock (unobserved); and ϵ_{it} captures measurement error in output. In this case, the parameter of interest is θ_{it}^v, which measures the output elasticity of the variable input. The estimation of this parameter is complicated by the fact that ω_{it} is unobserved, which means OLS estimates will be biased.

The solution is to use the insight in Olley and Pakes (1996): unobserved productivity ω_{it} can be expressed as an unknown function of the firm's state variables and other observables. This means taking the input (or investment) demand and invert the function to obtain productivity, ω_{it}. That is:

$$\omega_{it} = h_t(d_{it}, k_{it}, z_{it})$$

Where d_{it} is the control variable; z_{it} captures output and input market factors that generate variation in factor demand (for input d) across firms, conditional on the level of productivity and capital. In this estimation, I follow Ackerberg et al. (2007) and choose the variable input v_{it} as the control variable. Additionally, and following Chen (2020), I use the 4-digit market share to control for z_{it}.

The method then uses a two-stage approach. In the first stage, the measurement error (ϵ_{it}) and the unanticipated shocks to output (ω_{it}) are purged using a non-parametric projection of output on the inputs and the control variable. In the case of $d_{it} = v_{it}$ this means:

$$y_{it} = \phi_t(v_{it}, k_{it}, z_{it}) + \epsilon_{it}$$

The second stage obtains the output elasticity by constructing moments on the productivity shock by considering a process:

$$\omega_{it} = g(\omega_{it}) + \xi_{it}$$

.

The moment condition of the industry-year specific output elasticity is given by:

$$\mathbb{E}\left(\xi_{it}(\theta_t)\begin{bmatrix} v_{it-1} \\ k_{it} \end{bmatrix}\right) = 0$$

Where $\xi_{it}(\theta_t)$ is obtained by projecting productivity $\omega_{it}(\theta_t)$ on its lag $\omega_{it-1}(\theta_t)$. In this context, θ_t is the vector of the two parameters $\begin{pmatrix} \theta_t^V \\ \theta_t^K \end{pmatrix}$. Productivity is then obtained from $\phi_{it} - \theta_t^V v_{it} - \theta_t^K k_{it}$ using the estimate ϕ_{it} from the first-stage regression. This method identifies the output elasticity of variable inputs assuming that the variable input responds to contemporaneous productivity shocks, but that the lagged values do not. It also assumes that the lag of variable input is correlated with the current observation of variable input use only through serially correlated input and output market conditions captured by z_{it}.

As I have explained, following Chen (2020) (who in turns follows Flynn et al. (2019)), I add a constant-returns-to-scale (CRS) condition, which means $\theta_t^V + \theta_t^K = 1$. Additionally, and using the insights of Wooldridge (2009), the estimation is done in a GMM fashion. This means first applying the two-stage approach of Ackerberg et al. (2007) to obtain initial values for the GMM procedure of Wooldridge (2009). This GMM methodology allows for more precise estimates of the parameter (i.e. the standard error is smaller).

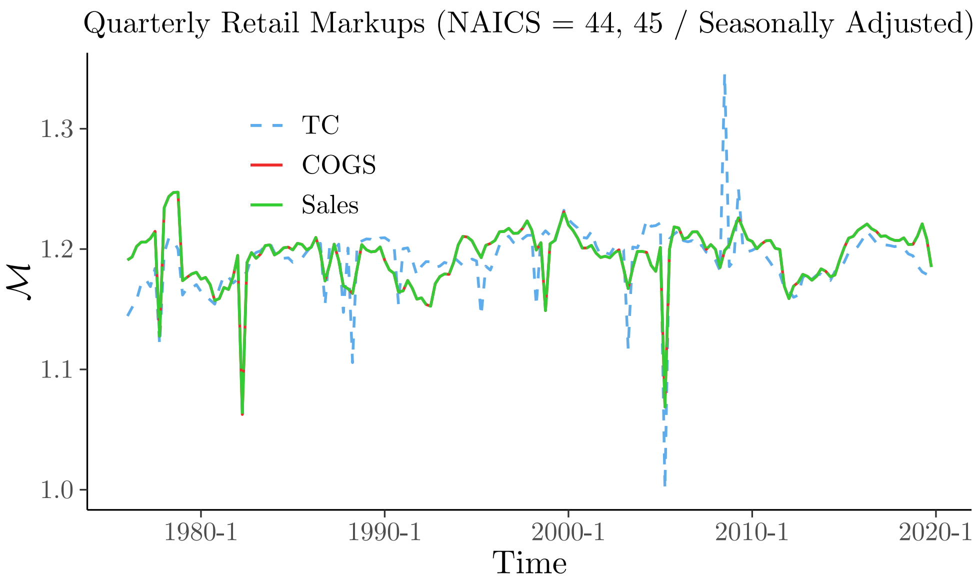
C Retail Markups

The findings of Stroebel and Vavra (2019) rely on the effect of local house prices on local retail markups. In this paper I concentrate in aggregate markups, meaning the average markup per firm across all 23 industries, defined at the 2-digit NAICS code. As a robustness check, I repeat the analysis carried out in this paper, but focusing on the retail sector, as defined by those firms with NAICS classification equal to 44 or 45. This is still a broader definition than the retailers used in Stroebel and Vavra (2019), who focus mainly in food and groceries, but it is closer in spirit. Focusing only on these two industries reduces the individual firms in the sample from 16,051 to 1,178.

I begin the analysis by constructing a quarterly series for retail markups. This is shown in Figure 9, where it is possible to see that, unlike the case for aggregate markups, retail markups do not seem to exhibit an increasing trend. It is also possible to observe that in this case the COGS-weighted and the Sales-weighted series are essentially identical.

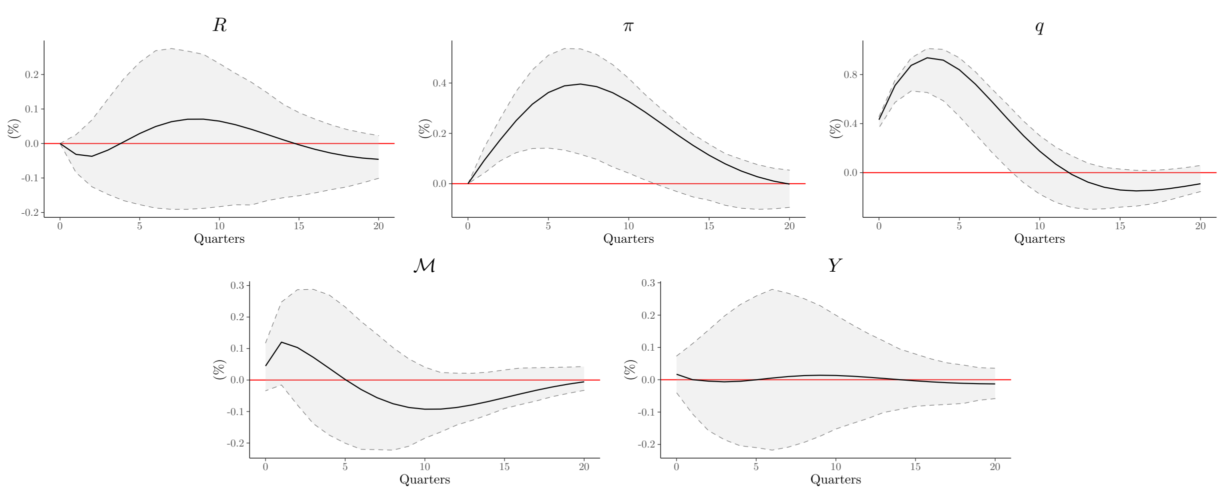
I filter the quarterly series for COGS-weighted retail markups using a Baxter-King filter with upper and lower bands of 32 and 2, respectively. I then estimate the same VAR system described in Section 3 and compute the IRFs of the system. The IRFs indicate that the three stylized facts from Section 4 still hold true overall. However, as Figure 10 shows, the response of markups is estimated with lower precision, although it exhibits the same general shape. These IRFs are shown in Figures 10 to 14.

Figure 9: Quarterly Retail Markups (various weights)



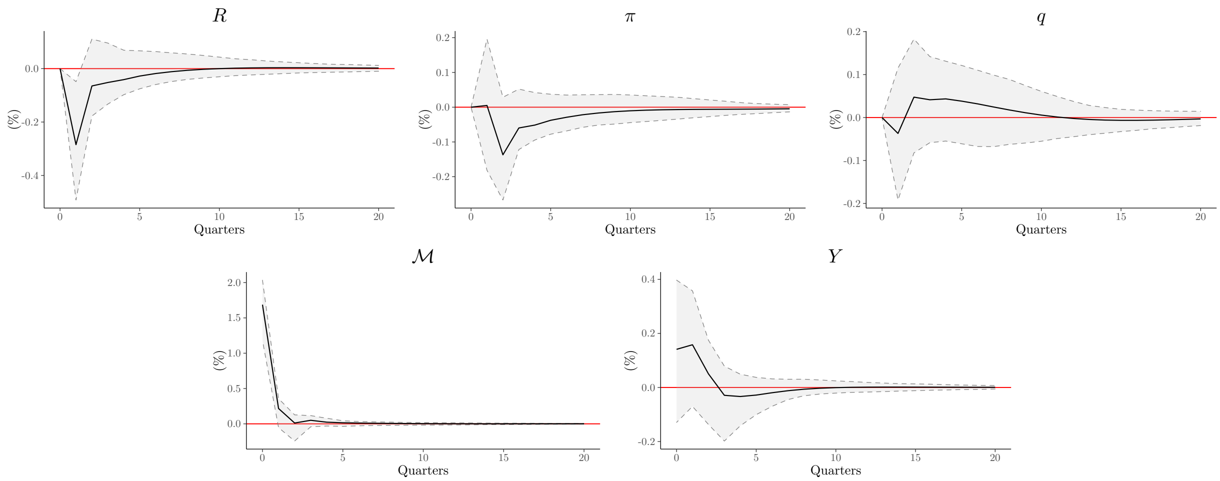
The figure shows the three series for aggregate retail markups across time. That is, the figure plots the estimated aggregate markups for all firms in the retail industries defined as having the 2-digit NAICS codes 44 and 45. The frequency is quarterly and the sample period goes from 1976Q1 to 2019Q4. The data has been seasonally adjusted using the X-13 method developed by the Census Bureau. Each series results from giving a different weight, m_{it}, to each individual firm's markup, where each individual weight can be either the firm's share of total cost per period (TC, dashed line in blue), the firm's share of total expenditures in the variable input per period (COGS, solid line in red), or the firm's share of total sales per period (Sales, broken line in green). This means $m_{it} \in \{TC, COGS, Sales\}$, according to the notation of Eq. (2).

Figure 10: Impulse to house prices.



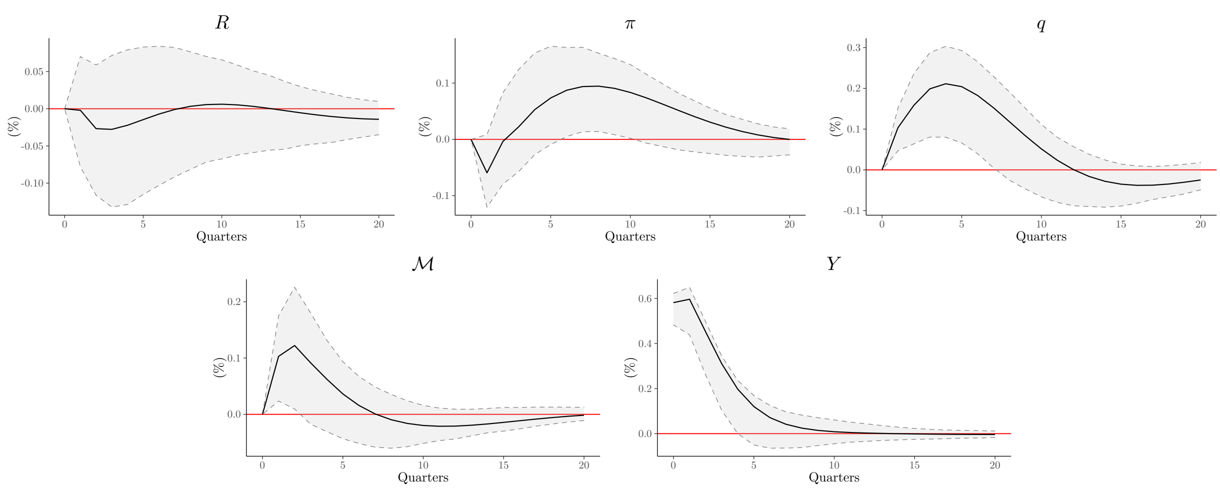
The figure shows the responses of the five variables in the VAR after a shock to house prices (q). Markups are measured as the COGS-weighted aggregate retail markup. The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 11: Impulse to markups.



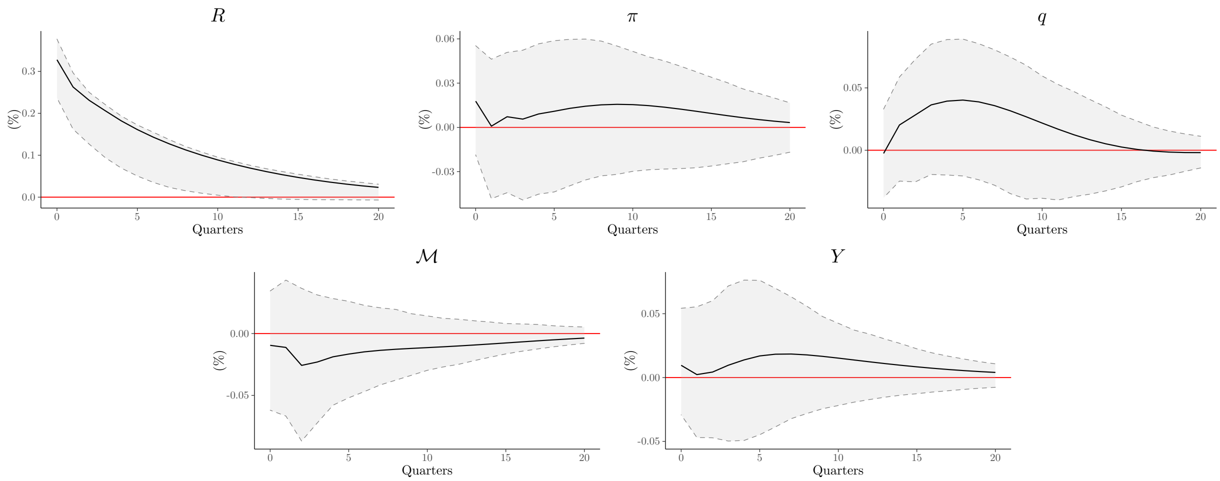
The figure shows the responses of the five variables in the VAR after a shock to aggregate markups ($\mathcal{M}$). The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 12: Impulse to real output.



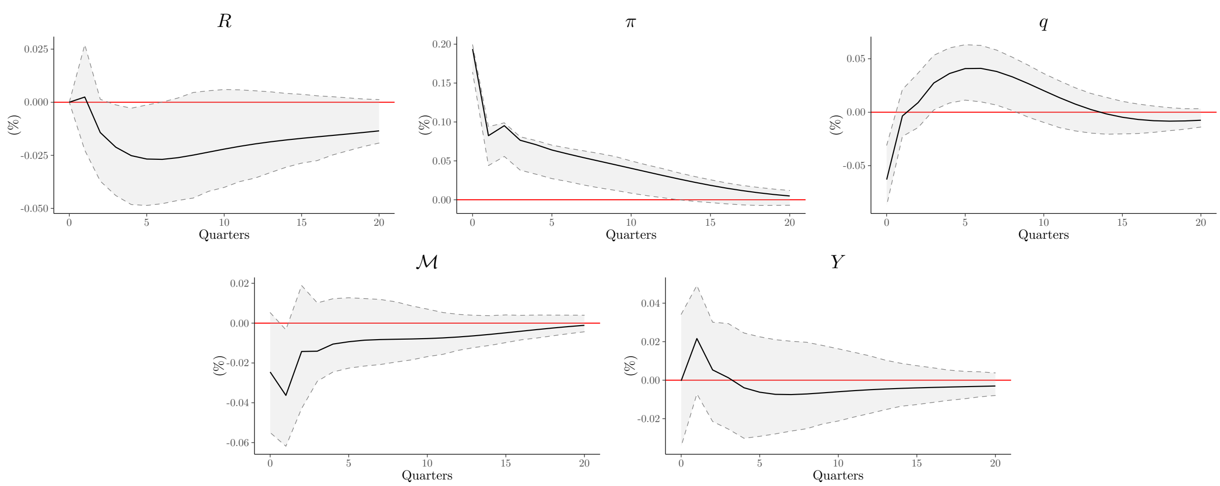
The figure shows the responses of the five variables in the VAR after a shock to real output (Y). The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 13: Impulse to the interest rate.



The figure shows the responses of the five variables in the VAR after a shock to the nominal interest rate (R). The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 14: Impulse to the inflation rate.



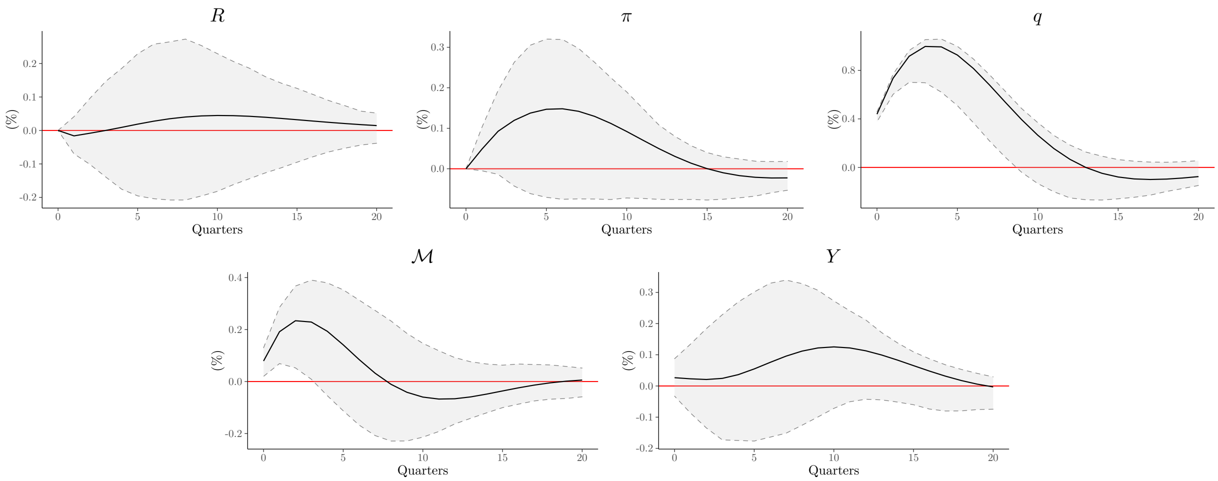
The figure shows the responses of the five variables in the VAR after a shock to the inflation rate (π). The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

D CPI Inflation

Another possible factor driving the results is the fact that Stroebel and Vavra (2019) focus on the final price of retail goods to consumers. It could be argued that the relevant price measure in that case would be the Consumer Price Index (CPI) as opposed to the GDP Deflator, which takes into account all the final goods and services produced in the economy. In that sense, I repeat the analysis carried out in Section 3 by using the quarterly CPI inflation rate, instead of the GDPD inflation rate. I then compute the IRFs for this modified VAR system.

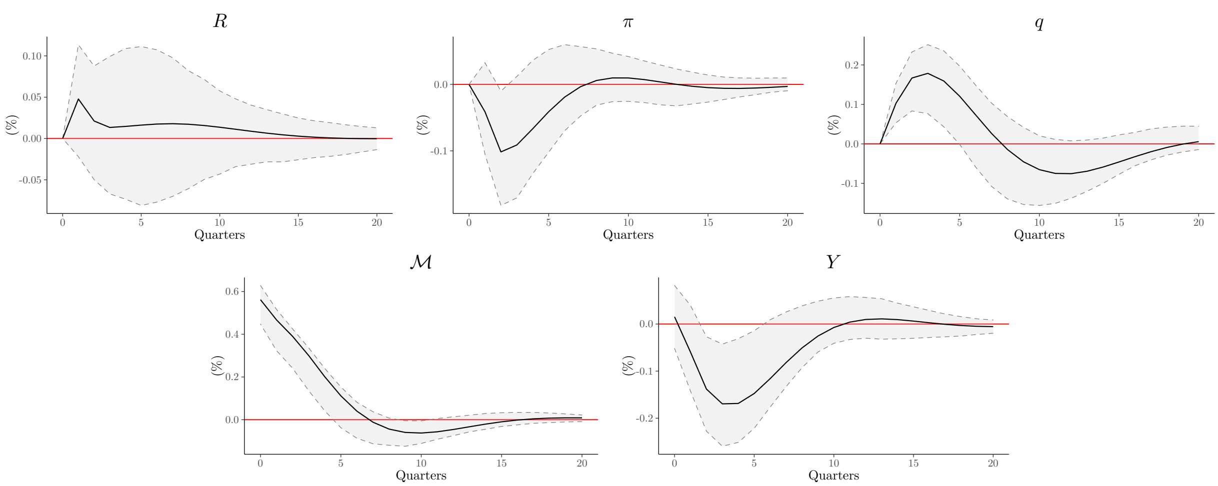
These IRFs are collected in Figures 15 to 19. As can be seen, most of the stylized facts mentioned in Section 4 still hold true. Moreover, some of the responses of the variables seem to be actually amplified.

Figure 15: Impulse to house prices.



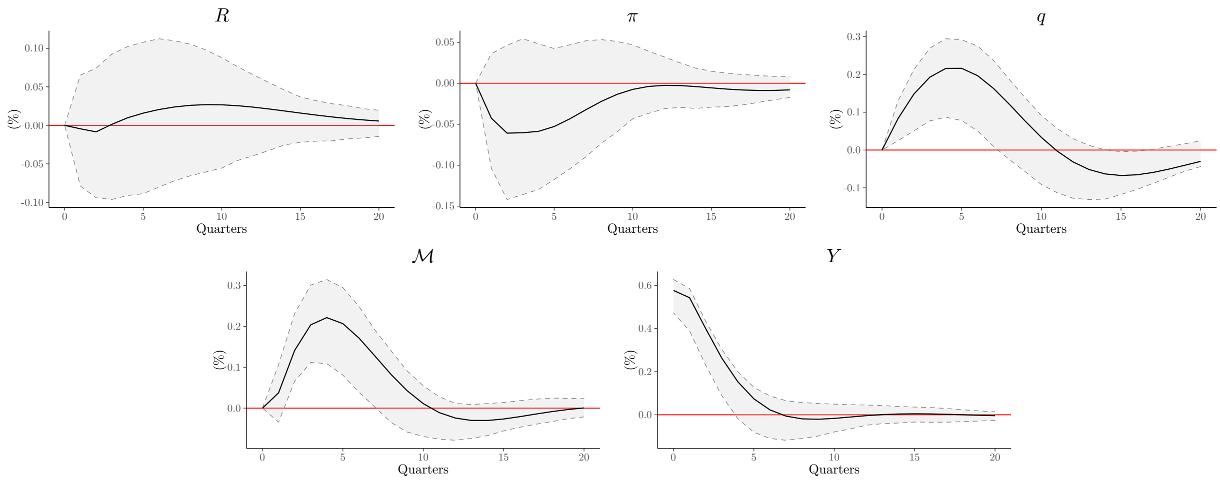
The figure shows the responses of the five variables in the VAR after a shock to house prices (q). The inflation rate is measured as the quarterly log-difference on the Consumer Price Index (CPI). The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 16: Impulse to markups.



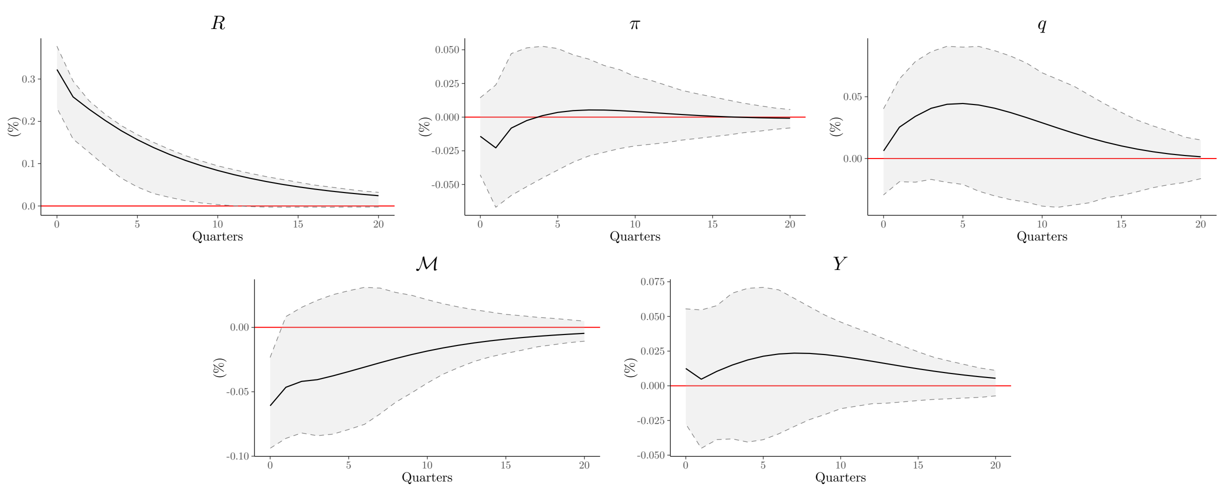
The figure shows the responses of the five variables in the VAR after a shock to aggregate markups ($\mathcal{M}$). The inflation rate is measured as the quarterly log-difference on the Consumer Price Index (CPI). The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 17: Impulse to real output.



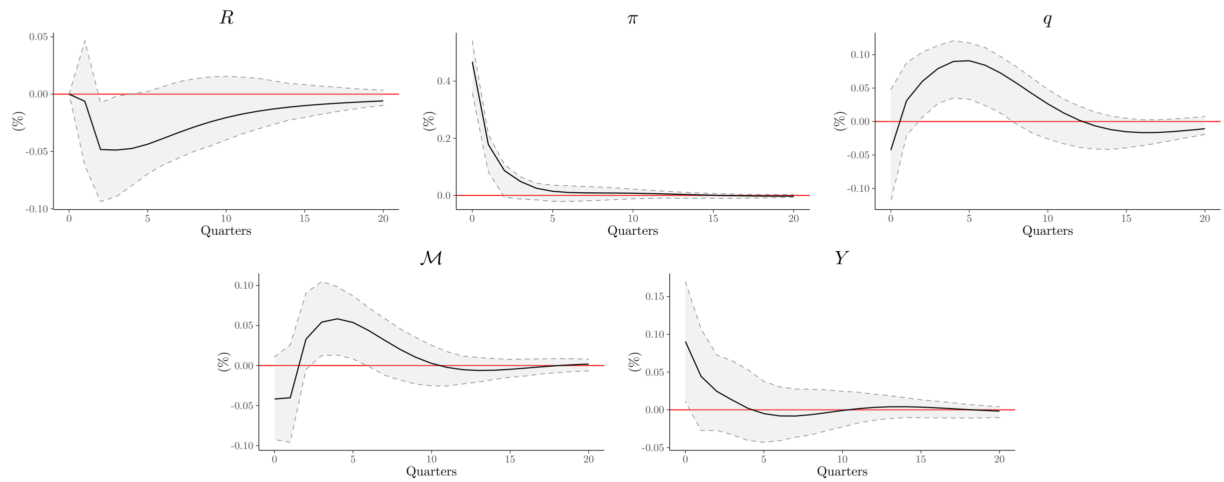
The figure shows the responses of the five variables in the VAR after a shock to real output (Y). The inflation rate is measured as the quarterly log-difference on the Consumer Price Index (CPI). The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 18: Impulse to the interest rate.



The figure shows the responses of the five variables in the VAR after a shock to the nominal interest rate (R). The inflation rate is measured as the quarterly log-difference on the Consumer Price Index (CPI). The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 19: Impulse to the inflation rate (CPI).



The figure shows the responses of the five variables in the VAR after a shock to the inflation rate (π). The inflation rate is measured as the quarterly log-difference on the Consumer Price Index (CPI). The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

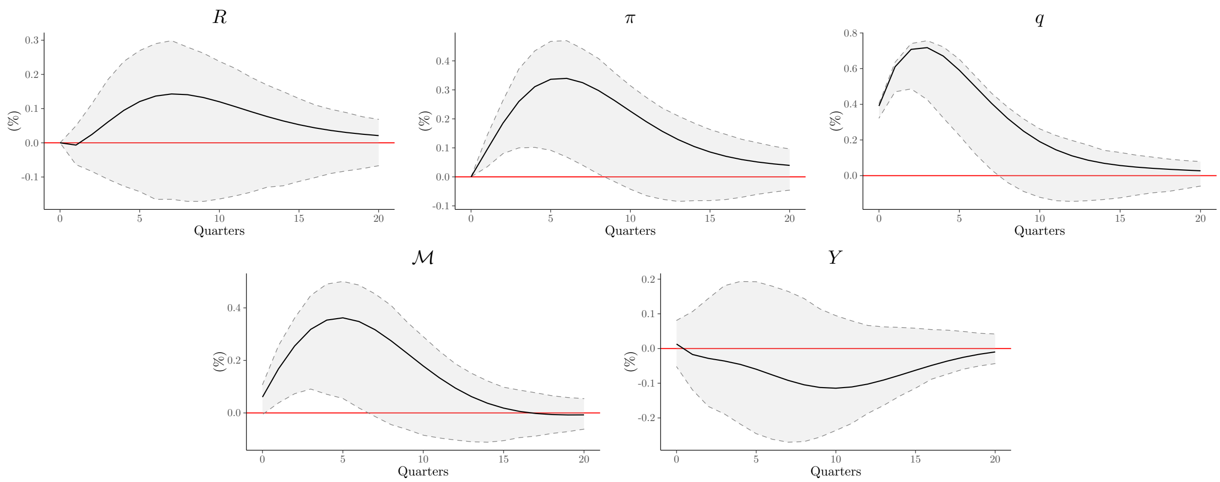
E Different Sample Periods

To investigate whether the results may be driven by focusing on particular time periods, I repeat the analysis over two different sample periods. The first one excludes the 2008 financial crisis an onward, which means the sample period considered goes from 1976Q1 to 2007Q4. For the second one, I focus only on the period that includes the last two significant recessions as identified by the NBER business cycle dates. That is, I consider the period going from 1991Q2 to 2019Q4.

1976Q1 – 2007Q4

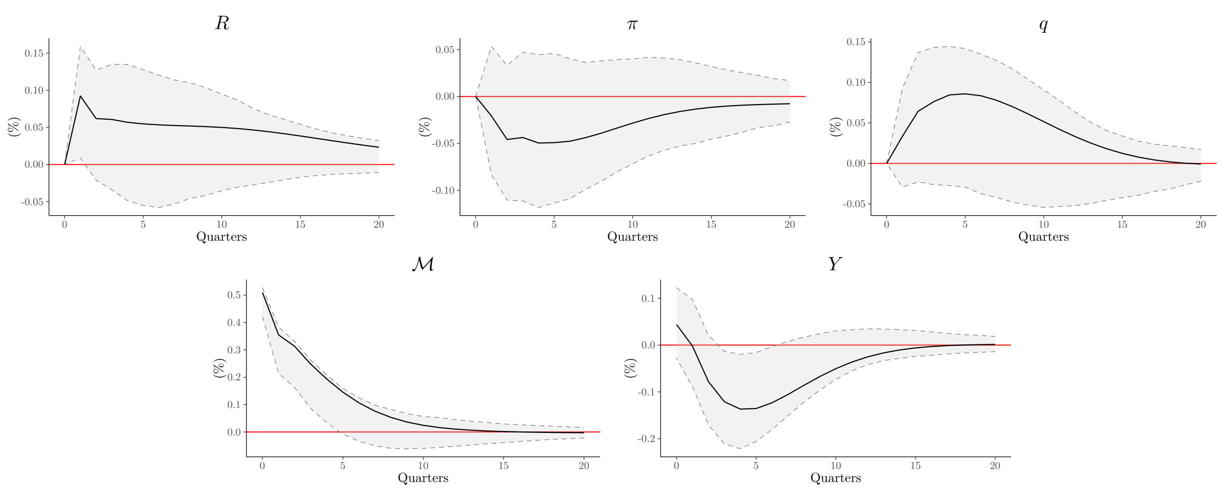
In the case of the sample period going from 1976Q1 to 2007Q4, the three stylized facts still hold in general. In particular, focusing on this period increases the persistence and the significance of the response of markups after a housing shock. However, it decreases the significance of both house prices and output after a shock to markups. In terms of the response of the system after an output shock, the results remain unchanged. Furthermore, the reaction of the system after an interest rate or an inflation shock do not vary significantly.

Figure 20: Impulse to house prices (1976Q1 - 2007Q4).



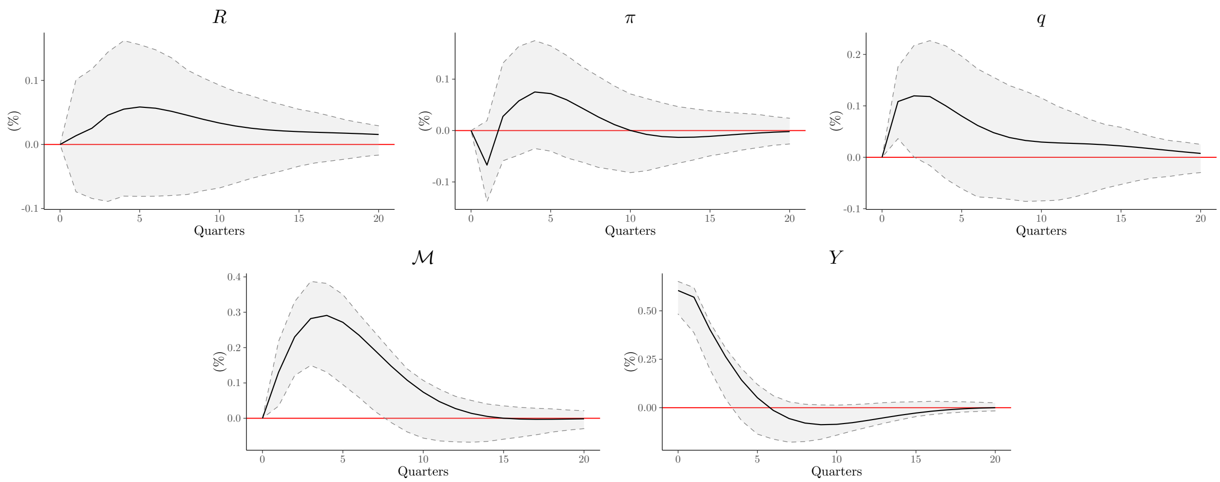
The figure shows the responses of the five variables in the VAR after a shock to house prices (q). The sample period runs from 1976Q1 to 2007Q4. The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 21: Impulse to markups (1976Q1 - 2007Q4).



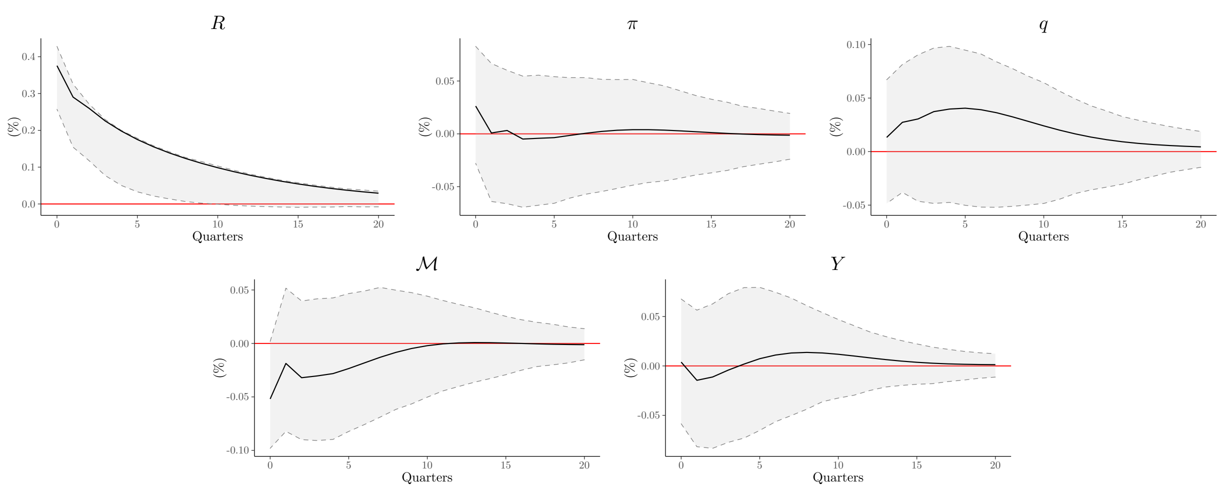
The figure shows the responses of the five variables in the VAR after a shock to aggregate markups ($\mathcal{M}$). The sample period runs from 1976Q1 to 2007Q4. The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 22: Impulse to real output (1976Q1 - 2007Q4).



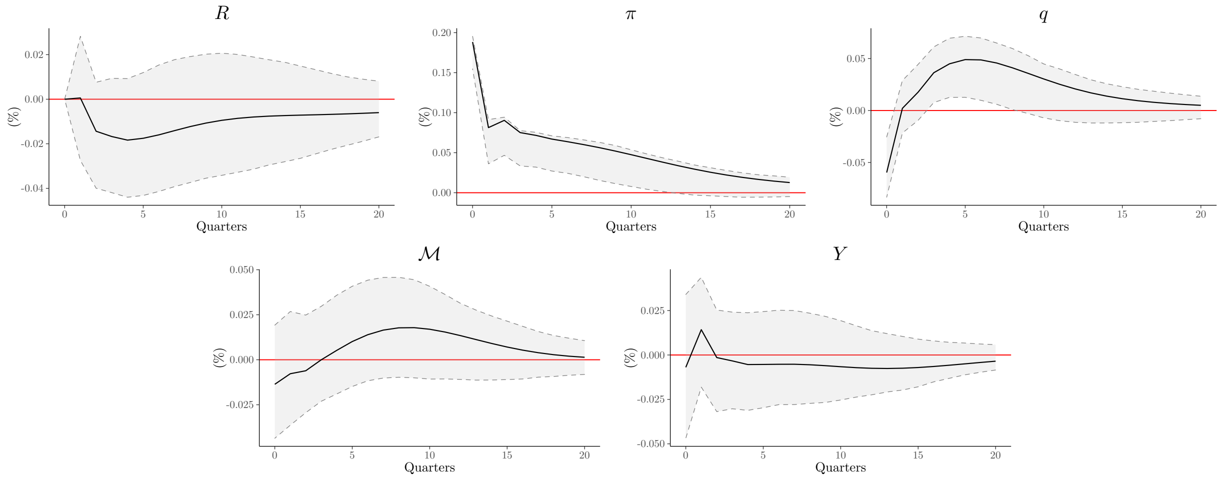
The figure shows the responses of the five variables in the VAR after a shock to real output (Y). The sample period runs from 1976Q1 to 2007Q4. The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 23: Impulse to the interest rate (1976Q1 - 2007Q4).



The figure shows the responses of the five variables in the VAR after a shock to the nominal interest rate (R). The sample period runs from 1976Q1 to 2007Q4. The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 24: Impulse to the inflation rate (1976Q1 - 2007Q4).



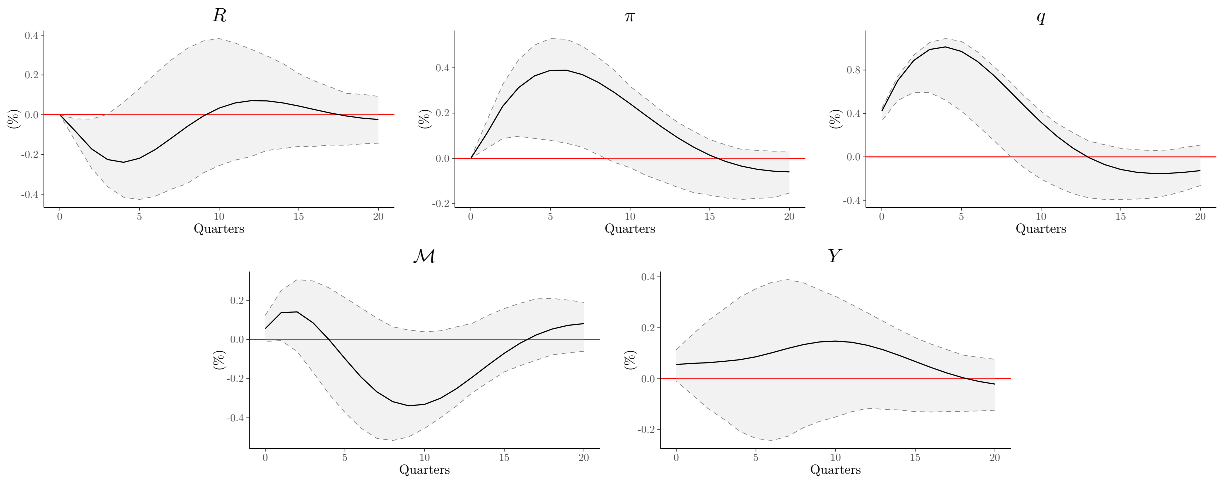
The figure shows the responses of the five variables in the VAR after a shock to the inflation rate (π). The sample period runs from 1976Q1 to 2007Q4. The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

1991Q2 – 2019Q4

For the period going from 1991Q2 to 2019Q4the results do change. In particular, after a shock to house prices inflation does rise significantly, but markups do not do so anymore. While they initially increase, they then decrease below trend, albeit in a not statistically significant manner. After a markup shock, output follows a similar trajectory as in the baseline analysis, however the significance is reduced. Furthermore, now the interest rate significantly falls below trend. The response of the system after an output shock remains unchanged. Alternatively, the system also shows a different response after an interest rate and an inflation shock. After an interest rate shock, house prices increase significantly. After a inflation shock, house prices and output increase and interest rates decrease significantly, albeit for a brief period.

This could be explained by the behavior of the variables in the VAR during the housing boom period going from the mid-90s to 2008, and the recovery dynamics after the crash of 2008. The first period was characterized by rising house prices and changes in interest rates, which could explain the positive relationship between house prices and interest rates. The second period was characterized by decreasing interest rates, and rising house prices as the economy recovered. This could help explain the positive relationship between inflation, house prices and output; and the

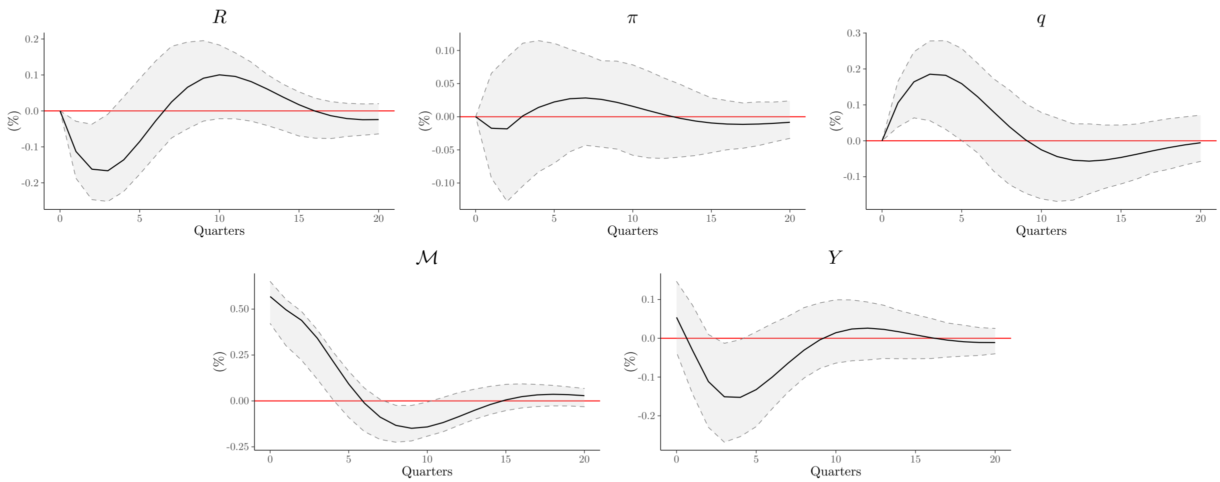
Figure 25: Impulse to house prices (1991Q1 - 2019Q4).



The figure shows the responses of the five variables in the VAR after a shock to house prices (q). The sample period runs from 1991Q2 to 2019Q4. The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

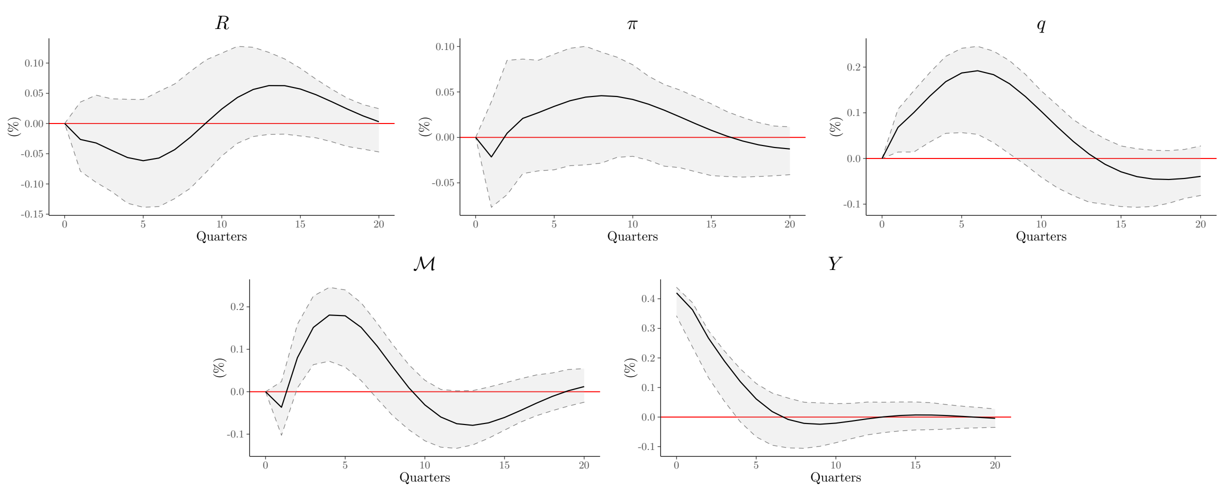
negative relationship between inflation and interest rates.

Figure 26: Impulse to markups (1991Q2 - 2019Q4).



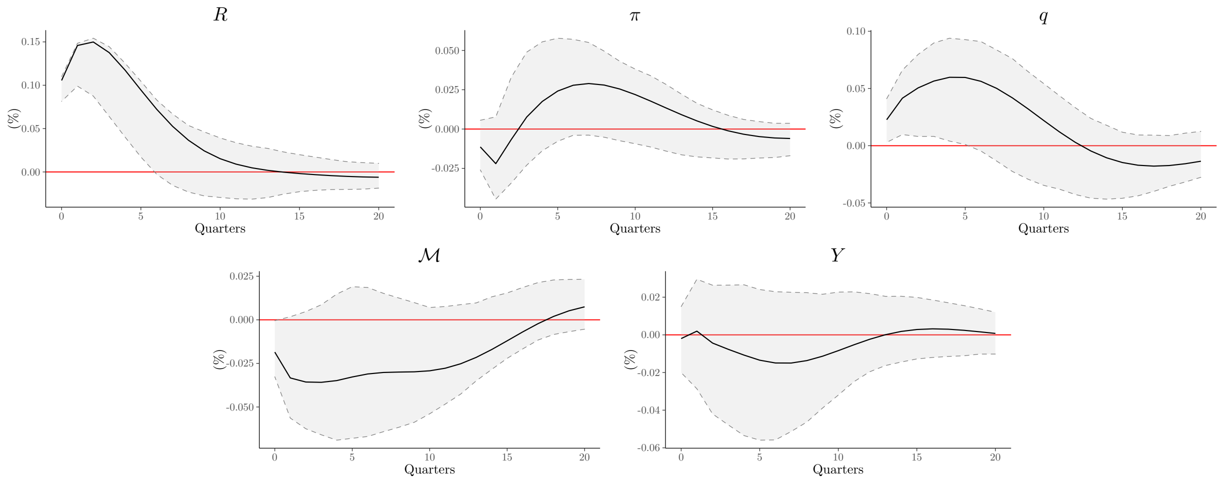
The figure shows the responses of the five variables in the VAR after a shock to aggregate markups ($\mathcal{M}$). The sample period runs from 1991Q2 to 2019Q4. The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 27: Impulse to real output (1991Q2 - 2019Q4).



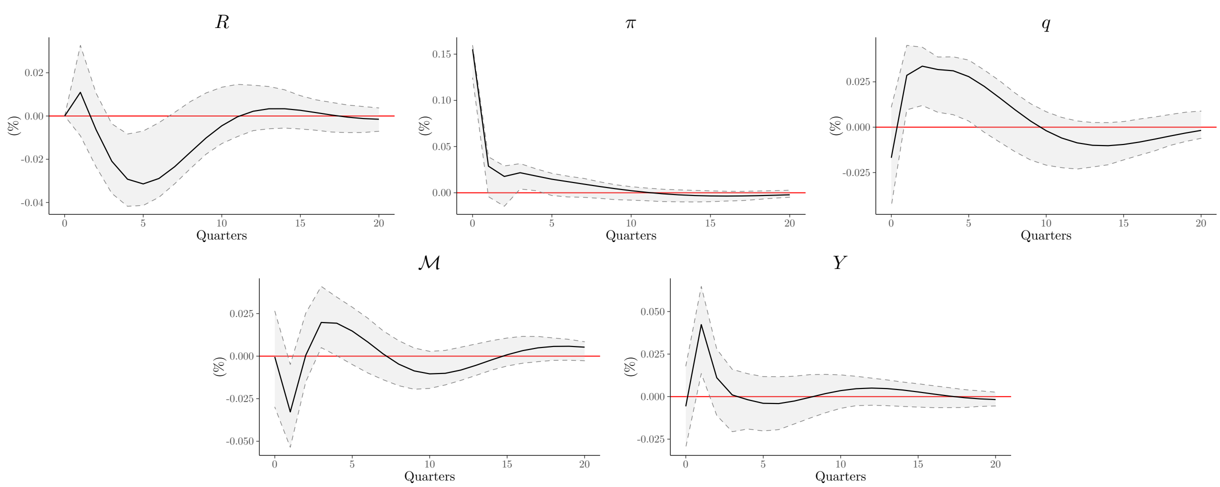
The figure shows the responses of the five variables in the VAR after a shock to real output (Y). The sample period runs from 1991Q2 to 2019Q4. The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 28: Impulse to the interest rate (1991Q2 - 2019Q4).



The figure shows the responses of the five variables in the VAR after a shock to the nominal interest rate (R). The sample period runs from 1991Q2 to 2019Q4. The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 29: Impulse to the inflation rate (1991Q2 - 2019Q4).



The figure shows the responses of the five variables in the VAR after a shock to the inflation rate (π). The sample period runs from 1991Q2 to 2019Q4. The red reference line indicates a level of zero. The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

F Generalized Impulse Response Functions

An additional robustness check is to look at Generalized Impulse Response Functions (GIRFs), which do not depend on the ordering of the variables in the VAR. As developed in Pesaran and Shin (1998), consider the *augmented* vector autoregressive model ($VAR(p)$):

$$\boldsymbol{X_t} = \sum_{i=1}^{p} \boldsymbol{\Phi_i X_{t-i}} + \boldsymbol{\Psi w_t} + \boldsymbol{\varepsilon_t}, \quad t = 1, 2, \ldots, T$$

with $\boldsymbol{X_t} = (x_{1t}, x_{2t}, \ldots, x_{mt})'$ is the $m \times 1$ vector of endogenous variables; $\boldsymbol{w_t}$ is a $q \times 1$ vector of deterministic (exogenous) variables; $\boldsymbol{\Phi_i}$, for $i = 1, \ldots, p$, and $\boldsymbol{\Psi}$ are $m \times m$ and $m \times q$ coefficient matrices.

The following assumptions guarantee that the VAR can be expressed in $MA(\infty)$ representation:

Assumption 2.1: $E(\boldsymbol{\varepsilon_t}) = \boldsymbol{0}$, $E(\boldsymbol{\varepsilon_t \varepsilon_t'}) = \boldsymbol{\Sigma}$ for all t, where $\boldsymbol{\Sigma} = \{\sigma_{ij}, i, j = 1, 2, \ldots, m\}$ is an $m \times m$ positive definite matrix, $E(\boldsymbol{\varepsilon_t \varepsilon_{t'}'}) = \boldsymbol{0}$ for all $t \neq t'$ and $E(\boldsymbol{\varepsilon_t} \mid \boldsymbol{w_t}) = \boldsymbol{0}$.

Assumption 2.2: All the roots of $\mid \boldsymbol{I_m} - \sigma_{i=1}^{p} \boldsymbol{\Phi_i} z^i \mid = \boldsymbol{0}$ fall outside the unit circle.

Assumption 2.3: $\boldsymbol{X_{t-1}}, \boldsymbol{X_{t-2}}, \ldots, \boldsymbol{X_{t-p}}, \boldsymbol{w_t}$, $t = 1, 2, \ldots, T$, are not perfectly collinear.

The $MA(\infty)$ representation is then:

$$\boldsymbol{X_t} = \sum_{i=0}^{\infty} \boldsymbol{A_i \varepsilon_{t-i}} + \sum_{i=0}^{\infty} \boldsymbol{G_i w_{t-i}}, \quad t = 1, 2, \ldots, T$$

where the $m \times m$ coefficient matrices $\boldsymbol{A_i}$ are obtained recursively as follows:

$$\boldsymbol{A_i} = \boldsymbol{\Phi_1 A_{i-1}} + \boldsymbol{\Phi_2 A_{i-2}} + \cdots + \boldsymbol{\Phi_p A_{i-p}},\ i = 1, 2, \ldots,$$

With $\boldsymbol{A_0} = \boldsymbol{I_m}$, $\boldsymbol{A_i} = \boldsymbol{0}$ for $i < 0$ and $\boldsymbol{G_i} = \boldsymbol{A_{i-1} \Psi}$.

Now, assuming an $m \times 1$ hypothetical vector of shocks $\boldsymbol{\delta}$, with $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_m)'$, the IRF may be described as a conceptual problem of comparing the base-line profile of the economy at time $t + n$ against the profile of the economy hit by shock $\boldsymbol{\delta}$ at time t. Letting $\boldsymbol{\Omega_{t-1}}$ be the non-decreasing information set that reflects the known history of the economy up to period $t - 1$, the generalized impulse response function of $\boldsymbol{X_t}$ at horizon n, can be defined as:

$$\boldsymbol{GI_x}(n, \boldsymbol{\delta}, \boldsymbol{\Omega_{t-1}}) = E(\boldsymbol{X_{t+n}} \mid \boldsymbol{\varepsilon_t} = \boldsymbol{\delta}, \boldsymbol{\Omega_{t-1}}) - E(\boldsymbol{X_{t+n}} \mid \boldsymbol{\Omega_{t-1}})$$

Given the previous equations and assumptions, this collapses to: $\boldsymbol{GI_x}(n, \boldsymbol{\delta}, \boldsymbol{\Omega_{t-1}}) = \boldsymbol{A_n}\boldsymbol{\delta}$, which is independent of $\boldsymbol{\Omega_{t-1}}$, but depends on the shocks defined by $\boldsymbol{\delta}$.

The traditional approach of (Sims, 1980) is to choose $\boldsymbol{\delta}$ according to a Cholesky decomposition of $\boldsymbol{\Sigma}$:

$$\boldsymbol{PP'} = \boldsymbol{\Sigma}$$

Where $\boldsymbol{P}$ is an $m \times m$ lower triangular matrix. This means the above $MA(\infty)$ representation can be expressed as:

$$\boldsymbol{X_t} = \sum_{i=0}^{\infty}(\boldsymbol{A_i P})(\boldsymbol{P^{-1}}\boldsymbol{\varepsilon_{t-i}}) + \sum_{i=0}^{\infty}\boldsymbol{G_i w_{t-i}} = \boldsymbol{X_t} = \sum_{i=0}^{\infty}(\boldsymbol{A_i P})(\boldsymbol{\xi_{t-i}}) + \sum_{i=0}^{\infty}\boldsymbol{G_i w_{t-i}}, \quad t = 1, 2, \ldots, T$$

with $\boldsymbol{\xi_t} = \boldsymbol{P^{-1}}\boldsymbol{\varepsilon_t}$, are the orthogonalized disturbances. That is, $E(\boldsymbol{\xi_t \xi t'}) = \boldsymbol{I_m}$. So the vector of orthogonalized IRFs from a unit shock to the jth equation on $\boldsymbol{X_{t+n}}$ is given by:

$$\psi_j^o(n) = \boldsymbol{A_n P e_j},\ n = 0, 1, 2, \ldots$$

And where $\boldsymbol{e_j}$ is an $m \times 1$ vector with ones in the jth element, and zeroes elsewhere.

Given that the Cholesky decomposition gives rise to a lower triangular matrix $\boldsymbol{P}$, the ordering of the variables in the VAR determines the contemporaneous response of some of them vis-à-vis each shock.

The alternative approach of Generalized Impulse Response Functions (GIRFs) circumvents this issue by shocking a single element of $\boldsymbol{\varepsilon_t}$ and integrate out the effects of other shocks using an assumed or the historically observed distribution of the disturbances. In this case, the generalized impulse response function of $\boldsymbol{X_t}$ at horizon n, can be defined as:

$$\boldsymbol{GI_x}(n, \delta_j, \boldsymbol{\Omega_{t-1}}) = E(\boldsymbol{X_{t+n}} \mid \varepsilon_{jt} = \delta_j, \boldsymbol{\Omega_{t-1}}) - E(\boldsymbol{X_{t+n}} \mid \boldsymbol{\Omega_{t-1}})$$

Assuming $\boldsymbol{\varepsilon_t}$ has a multivariate normal distribution:

$$E(\boldsymbol{\varepsilon_t} \mid \varepsilon_{jt} = \delta_j) = (\sigma_{1j}, \sigma_{2j}, \ldots, \sigma_{mj})'\sigma_{jj}^{-1}\delta_j = \boldsymbol{\Sigma e_j}\sigma_{jj}^{-1}\delta_j$$

Hence, the $m \times 1$ vector of (unscaled) generalized impulse response of the effect of a shock in the jth equation at time t of $\boldsymbol{X_{t+n}}$ is given by:

$$\left(\frac{\boldsymbol{A_n \Sigma e_j}}{\sqrt{\sigma_{jj}}}\right)\left(\frac{\delta_j}{\sqrt{\sigma_{jj}}}, \, n = 0, 1, 2, \ldots\right)$$

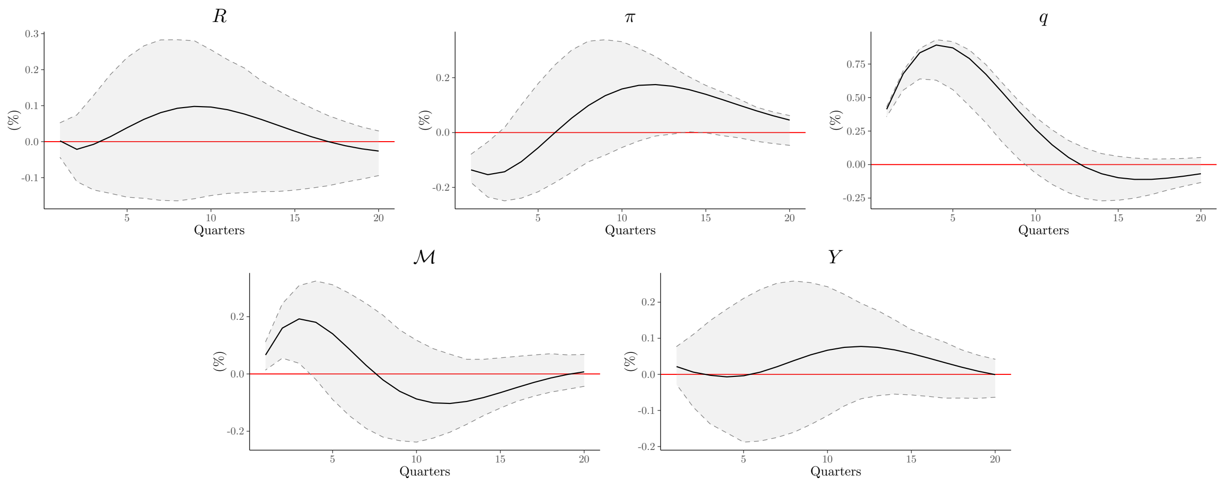
Setting $\delta_j = \sqrt{\sigma_{jj}}$ achieves the scaled generalized impulse response function as:

$$\psi_j^g(n) = \sigma_{jj}^{-\frac{1}{2}}\boldsymbol{A_n \Sigma e_j}, \, n = 0, 1, 2, \ldots$$

The above equation measures the effect one standard error shock to the jth equation at time t on the expected values of $\boldsymbol{X}$ at time $t + n$.

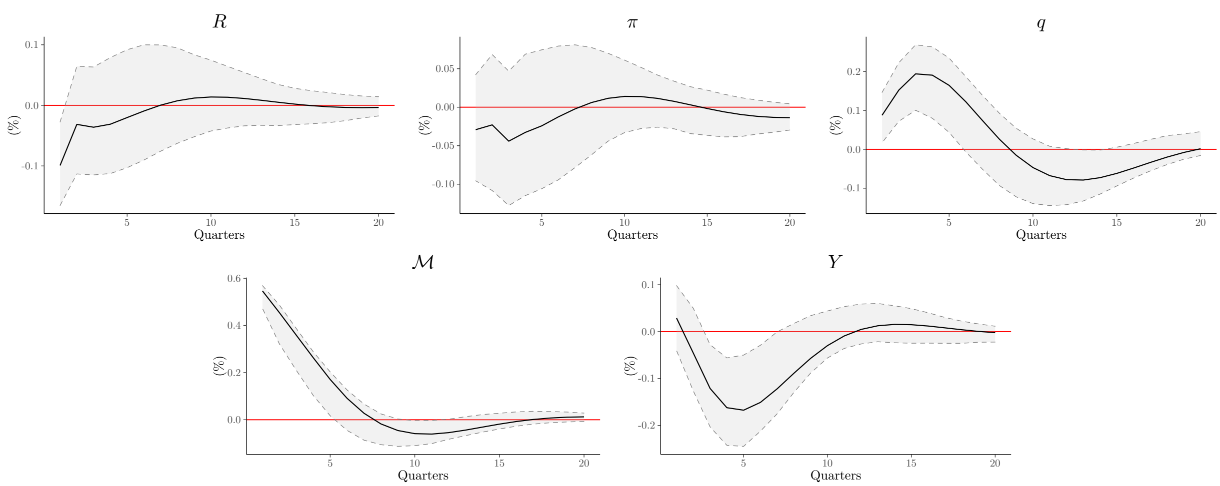
Since the GIRFs do not involve a Cholesky decomposition of the variance-covariance matrix, the responses remain invariant to the ordering of the variables in the VAR. In that sense, they provide a robustness check of the results presented in this paper with respect to the ordering of the VAR. The downside of this approach, is that the shocks are not orthogonal, and cannot be interpreted as structural shocks. As such, the IRFs do not have a causal interpretation. However, this analysis shows that the positive relationship between house prices and markups is robust to the ordering in the VAR. It also shows that After a positive shock to markups, house prices increase and output does so as well, albeit non-significantly. In later periods, output decreases significantly, while house prices return to zero. Finally, after a shock to output, both house prices and markups increase significantly. So, the stylized facts from the paper hold when using GIRFs to study the joint behavior of the variables in the model.

Figure 30: Generalized IRF – Shock to house prices.



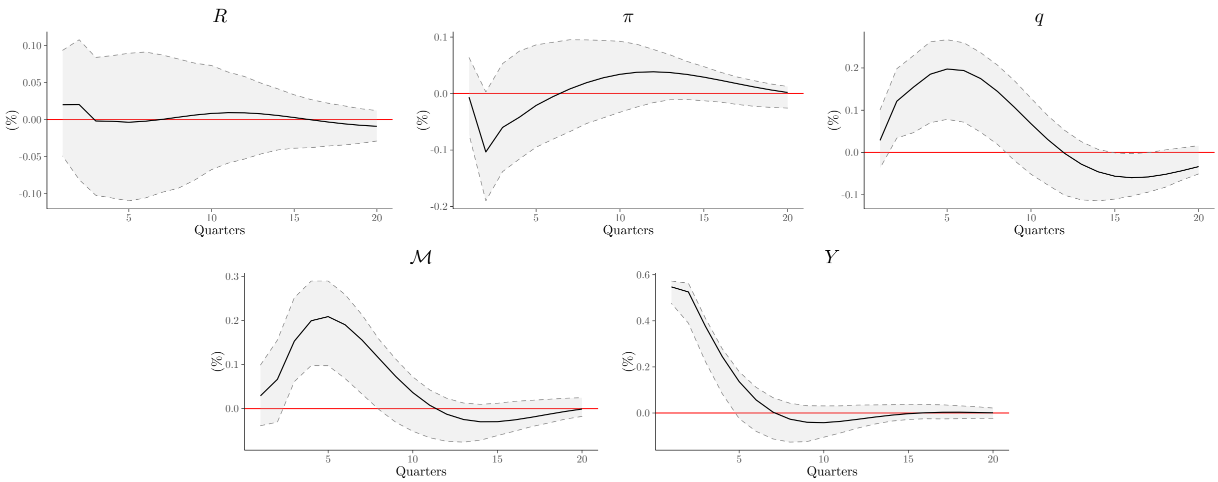
The figure shows the generalized responses of the five variables in the VAR after a shock to house prices (q). The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 31: Generalized IRF – Shock to markups.



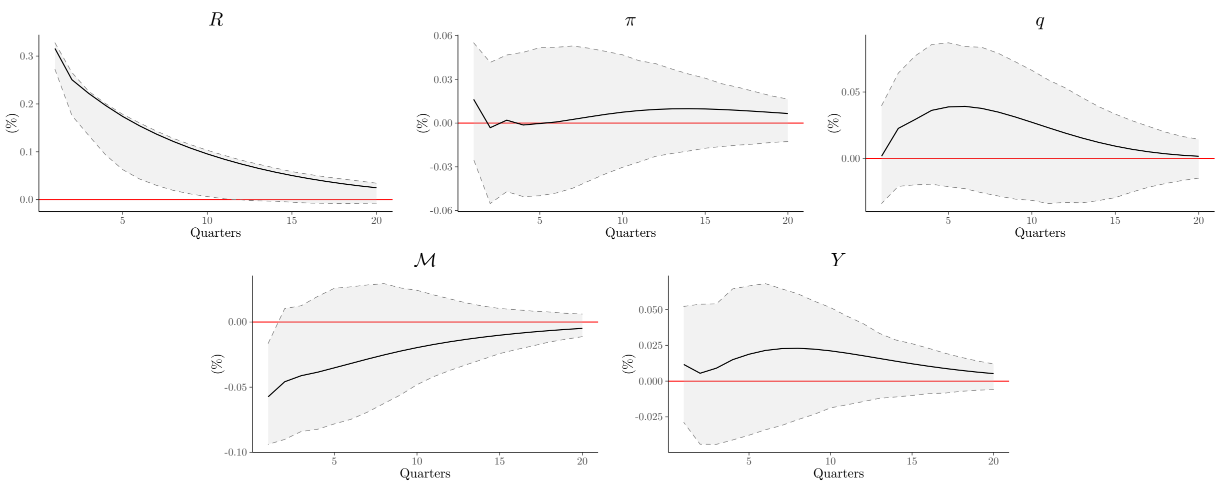
The figure shows the generalized responses of the five variables in the VAR after a shock to markups ($\mathcal{M}$). The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 32: Generalized IRF – Shock to output.



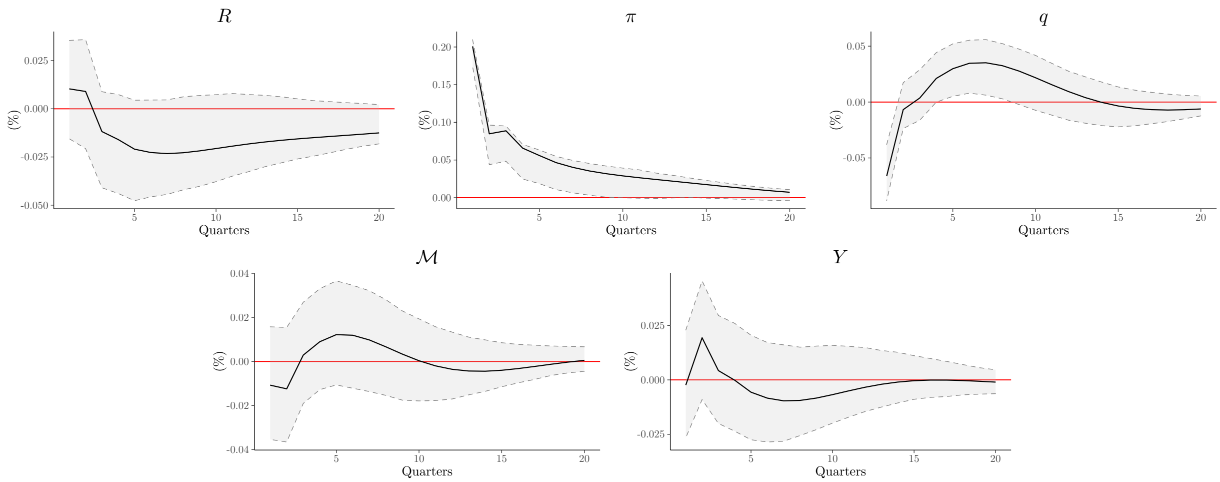
The figure shows the generalized responses of the five variables in the VAR after a shock to real output (Y). The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 33: Generalized IRF – Shock to the interest rate.



The figure shows the generalized responses of the five variables in the VAR after a shock to the nominal interest rate (R). The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

Figure 34: Generalized IRF – Shock to the inflation rate.



The figure shows the generalized responses of the five variables in the VAR after a shock to the inflation rate (π). The graphs show the estimated responses, in the solid dark line; and the 90% confidence regions, shaded in gray. A nonzero response is considered statistically significant when the shaded region lies clearly above or below the reference region. The confidence intervals are obtained by running a bootstrapping exercise with 1,000 draws.

G Local Projections

As developed in Jordà (2005), the usual approach with VARs and IRFs is to extrapolate into increasingly distant horizons from a given model. The ideal of local projections (LPs) is to estimate a projection at each period of interest. There are some advantages of local projections, mainly being more robust to misspecification; they are easy to estimate with simple regression techniques; their joint or point-wise analytic inference is simple; and they can accommodate experimentation with nonlinear and flexible specifications. For these reasons, measuring the responses of the variables in the VAR using local projections constitutes an additional robustness check to the responses found in the main text.

Using the notation of Appendix F, we have that the general impulse response function is given by:

$$\boldsymbol{GI_x}(n, \boldsymbol{\delta}, \boldsymbol{\Omega_{t-1}}) = E(\boldsymbol{X_{t+n}} \mid \boldsymbol{\varepsilon_t} = \boldsymbol{\delta}, \boldsymbol{\Omega_{t-1}}) - E(\boldsymbol{X_{t+n}} \mid \boldsymbol{\Omega_{t-1}})$$

Where $\boldsymbol{\delta}$, with $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_m)'$ is an $m \times 1$ hypothetical vector of shocks , the IRF may be described as a conceptual problem of comparing the base-line profile of the economy at time $t+n$ against the profile of the economy hit by shock $\boldsymbol{\delta}$ at time t.; $\boldsymbol{\Omega_{t-1}}$ is the non-decreasing information set that reflects the known history of the economy up to period $t-1$.

The usual way of estimating the IRFs by recursively iterating on the estimated model obtains the best, mean-squared predictions under the assumption that the data generating process (DGP) is indeed correct. However, the LP approach does a multi-step prediction by using direct forecasting models, which are reestimated for each forecast horizon. Concretely, $\boldsymbol{X_{t+n}}$ is projected onto the linear space generated by $(\boldsymbol{X_{t-1}}, \boldsymbol{X_{t-2}}, \ldots, \boldsymbol{X_{t-p}})'$

$$\boldsymbol{X_{t+n}} = \boldsymbol{\alpha^n} + \boldsymbol{B_1^{n+1}} \boldsymbol{X_{t-1}} + \boldsymbol{B_2^{n+1}} \boldsymbol{X_{t-2}} + \cdots + \boldsymbol{B_p^{n+1}} \boldsymbol{X_{t-p}} + \boldsymbol{u_{t+n}^n},\ n = 0, 1, 2 \ldots, h$$

Where $\boldsymbol{\alpha^n}$ is an $m \times 1$ vector of constants, and $\boldsymbol{B_i^{n+1}}$ are matrices of coefficients for each lag i and horizon $n+1$. The collection of h regressions described above is denoted *local projections*, in reference to nonparametric considerations.

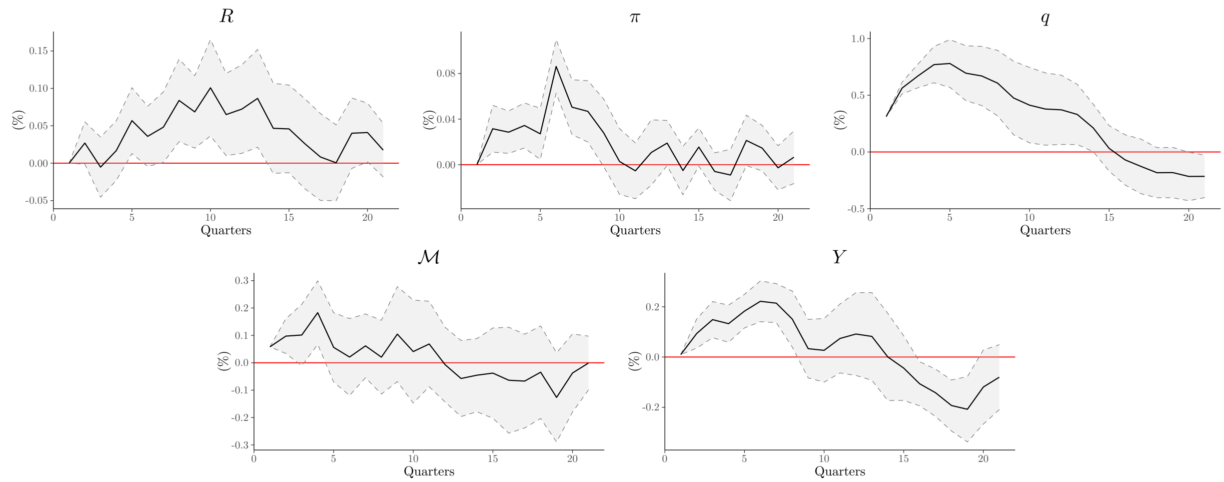
In this context, the general impulse response function from the local-linear projections collapses to:

$$\widehat{\boldsymbol{GI_x}}(n, \boldsymbol{\delta}, \boldsymbol{\Omega_{t-1}}) = \boldsymbol{\hat{B}_1^n} \boldsymbol{d_i},\ n = 0, 1, 2, \ldots, h$$

With the obvious normalization $\boldsymbol{\hat{B}_1^0} = \boldsymbol{I_m}$ and $\boldsymbol{d}_i$ is the "structural shock" to the ith element in $\boldsymbol{X_t}$. As explained more in-depth in Jordà (2005), the consistency of the estimates $\boldsymbol{\hat{B}_1^n}$ follows from the fact that the residuals $\boldsymbol{u_{t+n}^n}$ are a moving average of the forecast errors from time t to $t+s$, and therefore uncorrelated with the regressors, which are dated $t-1$ to $t-p$. It is important to note that the maximum lag p need not be common to each horizon n, and that consistency does not require joint estimation. Instead, the impulse response for the jth variable in $\boldsymbol{X_t}$ can be estimated by a univariate regression of X_{jt} onto $(\boldsymbol{X_{t-1}}, \boldsymbol{X_{t-2}}, \ldots, \boldsymbol{X_{t-p}})'$. The error bands come from a Newey-West estimator of the variance-covariance structure of $\boldsymbol{u_{t+n}^n}$.

The results from the LPs show that some of the stylized facts found in the main text hold. In particular, after a housing price shock, both markups and inflation increase. An additional advantage from using LPs is that the estimated response for the variables after a shock to the interest rate or the inflation rate are more intuitive. For instance, LPs indicate that after an interest rate shock, prices fall initially and rise subsequently (i.e. there is no price puzzle). Also,

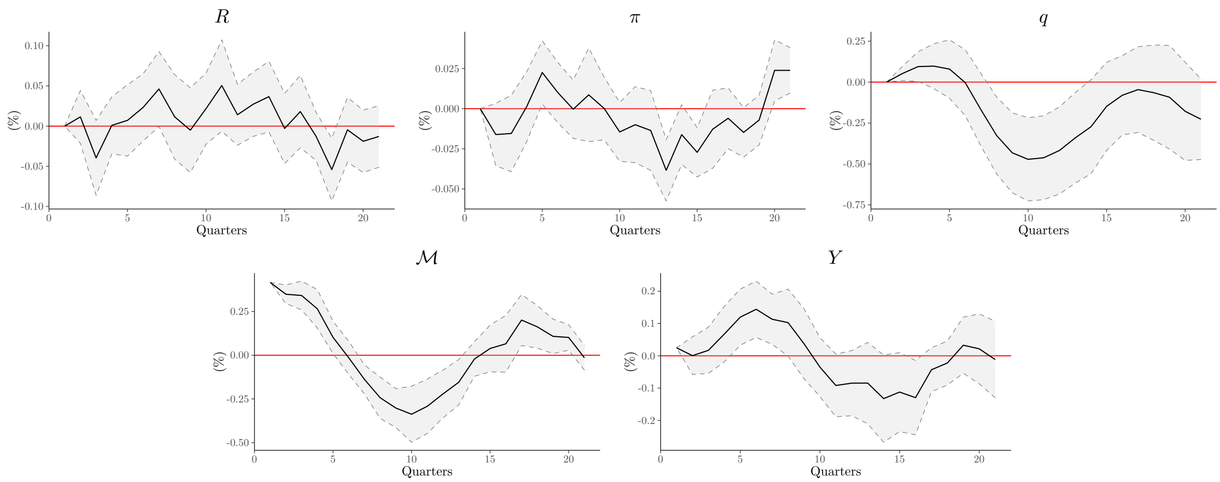
Figure 35: Local Projection – Shock to house prices.



The figure shows the local projections (Jordà, 2005) after a shock to house prices q. The chosen lag-length is equal to 10, to account for possible persistence in some of the variables. The shaded area represents the 90% error region. The bands are estimated using a Newey-West estimator of the residuals $\hat{\boldsymbol{u}}^{n}_{t+n}$.

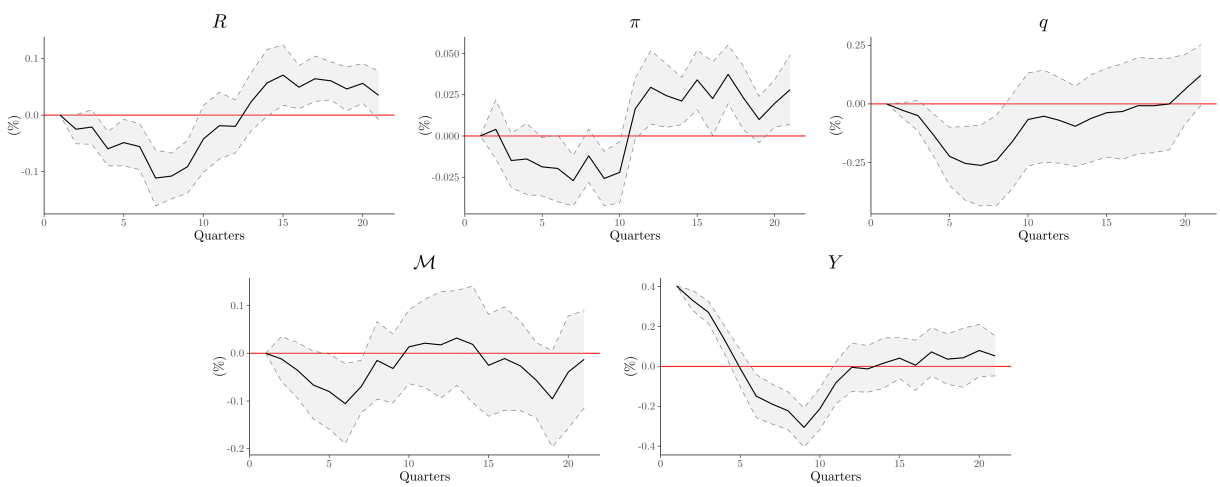
house prices, markups and output fall, which is consistent with the theoretical behavior of these variables after a monetary shock. Additionally, after an inflation shock, the interest rate increases, which cause house prices, markups, and output to decrease eventually. This behavior is also more consistent with conventional results in the theoretical literature. The results of this approach are collected in Figures 35 - 39.

Figure 36: Local Projection – Shock to markups.



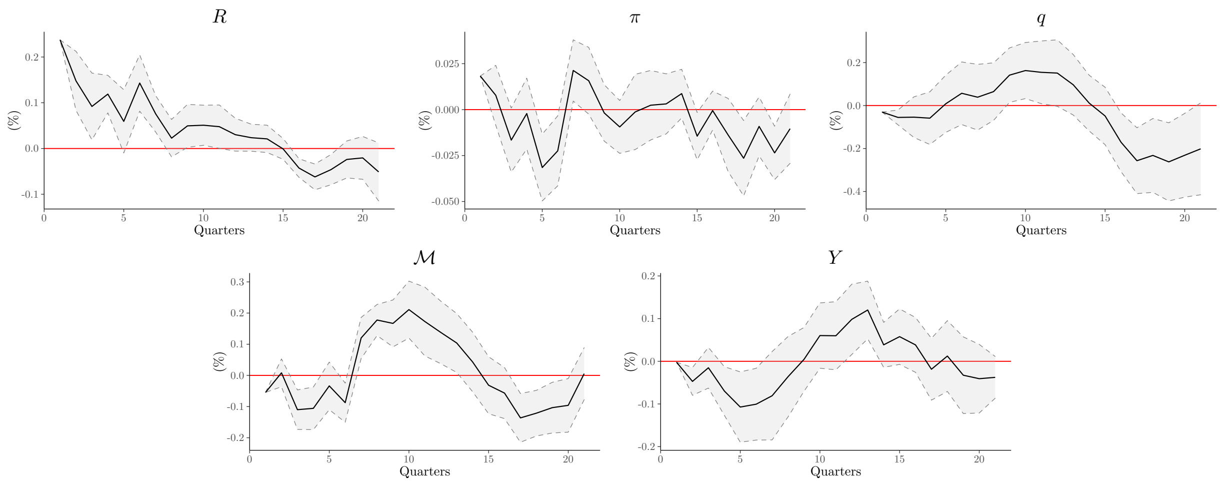
The figure shows the local projections (Jordà, 2005) after a shock to markups $\mathcal{M}$. The chosen lag-length is equal to 10, to account for possible persistence in some of the variables. The shaded area represents the 90% error region. The bands are estimated using a Newey-West estimator of the residuals $\hat{\boldsymbol{u}}^{\boldsymbol{n}}_{\boldsymbol{t+n}}$.

Figure 37: Local Projection – Shock to output.



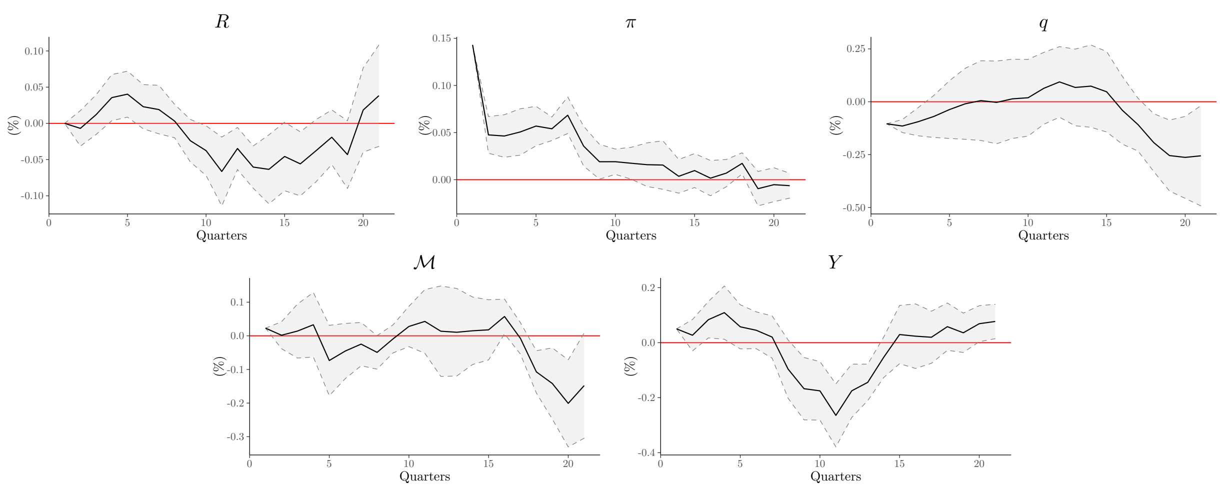
The figure shows the local projections (Jordà, 2005) after a shock to output Y. The chosen lag-length is equal to 10, to account for possible persistence in some of the variables. The shaded area represents the 90% error region. The bands are estimated using a Newey-West estimator of the residuals $\hat{\boldsymbol{u}}^{\boldsymbol{n}}_{\boldsymbol{t+n}}$.

Figure 38: Local Projection – Shock to the interest rate.



The figure shows the local projections (Jordà, 2005) after a shock to the interest rate R. The chosen lag-length is equal to 10, to account for possible persistence in some of the variables. The shaded area represents the 90% error region. The bands are estimated using a Newey-West estimator of the residuals $\hat{\boldsymbol{u}}^{\boldsymbol{n}}_{\boldsymbol{t+n}}$.

Figure 39: Local Projection – Shock to the inflation rate.



The figure shows the local projections (Jordà, 2005) after a shock to house prices π. The chosen lag-length is equal to 10, to account for possible persistence in some of the variables. The shaded area represents the 90% error region. The bands are estimated using a Newey-West estimator of the residuals $\hat{\boldsymbol{u}}^{\boldsymbol{n}}_{\boldsymbol{t+n}}$.

H VECM

A situation in which some of the variables of the VAR may be $I(1)$ may give rise to a situation of a spurious regression. Given that two of the variables (i.e. the inflation rate and the interest rate) in the VAR seem to be non-stationary, one way to correct for this situation would be to find a cointegration relationship between the variables in the model that makes the system stationary and then estimate a Vector Error Correction Model (VECM). That is, find a cointegrating matrix that makes a linear combination of the matrices in the system stationary. This is done following the methodology of Johansen (1991).

Simply put, consider the $VAR(p)$ system:

$$\boldsymbol{X_t} = \boldsymbol{\mu} + \boldsymbol{\Phi D_t} + \boldsymbol{\Pi_p X_{t-p}} + \ldots + \boldsymbol{\Pi_1 X_{t-1}} + \boldsymbol{\varepsilon_t},\ t = 1, 2, \ldots, T$$

Where $\boldsymbol{X_t}$ is the $m \times 1$ vector of endogenous variables, $\boldsymbol{D_t}$ is an $n \times 1$ vector of exogenous or deterministic variables (which can include a deterministic trend), $\boldsymbol{\Pi_i}$ for $i = 1, \ldots, p$ are the $m \times m$ matrices associated with the lags of the endogenous variables, $\boldsymbol{\Phi}$ is the $m \times n$ matrix associated with the exogenous variables, $\boldsymbol{\mu}$ is an $m \times 1$ vector of means. By taking subtracting $\boldsymbol{X_{t-1}}$ from both sides of the equation and adding and subtracting the appropriate terms, the VAR system above can be transformed into its VECM representation, which can be either longrun or transitory.

The longrun VECM is given by:

$$\boldsymbol{\Delta X_t} = \boldsymbol{\mu} + \boldsymbol{\Phi D_t} + \boldsymbol{\Pi X_{t-p}} + \boldsymbol{\Gamma_{p-1} \Delta X_{t-p+1}} + \cdots + \boldsymbol{\Gamma_1 \Delta X_{t-1}} + \boldsymbol{\varepsilon_t},\ t = 1, 2, \ldots, T$$

Where $\boldsymbol{\Delta X_t} := \boldsymbol{X_t} - \boldsymbol{X_{t-1}}$ is the first-difference operator; and $\boldsymbol{\Gamma_i} = \boldsymbol{\Pi_1} + \ldots + \boldsymbol{\Pi_i} - \boldsymbol{I_m}$ for $i = 1, \ldots, p-1$

The transitory VECM is given by:

$$\boldsymbol{\Delta X_t} = \boldsymbol{\mu} + \boldsymbol{\Phi D_t} - \boldsymbol{\Gamma_{p-1} \Delta X_{t-p+1}} - \cdots - \boldsymbol{\Gamma_1 \Delta X_{t-1}} + \boldsymbol{\Pi \Delta X_{t-1}} + \boldsymbol{\varepsilon_t},\ t = 1, 2, \ldots, T$$

With $\boldsymbol{\Gamma_i} = (\boldsymbol{\Pi_{i+1}} + \boldsymbol{\Pi_p})$, $i = 1, \ldots, p-1$

In both cases $\boldsymbol{\Pi} = \boldsymbol{\Pi_1} + \cdots + \boldsymbol{\Pi_p} - \boldsymbol{I_m}$. Both representations give rise to the same inference.

Consider the transitory VECM. If $\boldsymbol{X_t}$ is $I(1)$ then $\boldsymbol{\Delta X_t}$ is $I(0)$ (by assumption, so is $\boldsymbol{\varepsilon_t}$). This means the only term left is $\boldsymbol{\Pi X_{t-1}}$ and it must be also $I(0)$. Put differently, pre-multiplying $\boldsymbol{X_t}$ by $\boldsymbol{\Pi}$ produces linear combinations that are $I(0)$. These linear combinations are called the cointegrating relationships among the elements of $\boldsymbol{X_t}$.

Now, if $\boldsymbol{\Pi}$ is expressed as the product of two $m \times r$ matrices, α and β as follows: $\boldsymbol{\Pi} = \alpha\beta'$, then the element $\boldsymbol{\Pi X_{t-1}}$ can be expressed as:

$$\boldsymbol{\Pi X_{t-1}} = (\alpha\beta')\boldsymbol{X_{t-1}} = \alpha(\beta'\boldsymbol{X_{t-1}})$$

Therefore, $\beta'\boldsymbol{X_{t-1}}$ is an $r \times 1$ vector containing the error correction terms and the r columns of β contain the cointegrating vector. The coefficients of α determine the size of the effects of the r error correction terms in the m equations of the VECM. Therefore, testing for the number of cointegrating relationships r constitutes the first step in estimating a VECM.

The Johansen (1991) relies on the following observation: The number of cointegrating relationships r is a non-negative integer less than or equal to m, the dimension of $\boldsymbol{X_t}$. Additionally, since $\boldsymbol{\Pi} = \alpha\beta'$, then r is also the rank of $\boldsymbol{\Pi}$. The procedure consists on *sequentially* testing for the rank of $\boldsymbol{\Pi}$. This can be done either by doing trace tests or eigenvalue tests.

In the case of the trace test, the null hypothesis is $H_0 : r \leq r*$ where $r*$ is chosen sequentially from $\{0, 1, \ldots, m-1\}$. If $H_0 : r \leq 0$ is not rejected, then this implies that the rank of $\boldsymbol{\Pi}$ is equal to zero. That is, there are no cointegrating relationships and the VECM is a VAR in first-differences. If $H_0 : r \leq 0$ is rejected, then one tests $H_0 : r \leq 1$. This process continues until the hypothesis $H_0 : r \leq r*$ is not rejected, or until the hypothesis $H_0 : r \leq m-1$ is not rejected. The resulting number of cointegrating relationships is set to $r = r*$. If $H_0 : r \leq m-1$ is not rejected, this suggests $\boldsymbol{\Pi}$ is of full rank, which implies $\boldsymbol{\Pi X_{t-1}}$ spans the same vector space as the variables in $\boldsymbol{X_{t-1}}$, which in turn implies the left hand side of the equation is $I(0)$ but the right hand side is $I(1)$, which is a contradiction, given the underlying assumptions.

Carrying out the Johansen test on the VECM form of the VAR presented in the main text yields the following result:

Table 3: Johansen Test

		Critical Values		
H_0	Test statistic	10%	5%	1%
$r \leq 4$	4.07	6.50	8.18	11.65
$r \leq 3$	19.82	15.66	17.95	23.52
$r \leq 2$	40.44	28.71	31.52	37.22
$r \leq 1$	78.16	45.23	48.28	55.43
$r = 0$	120.04	66.49	79.60	78.87

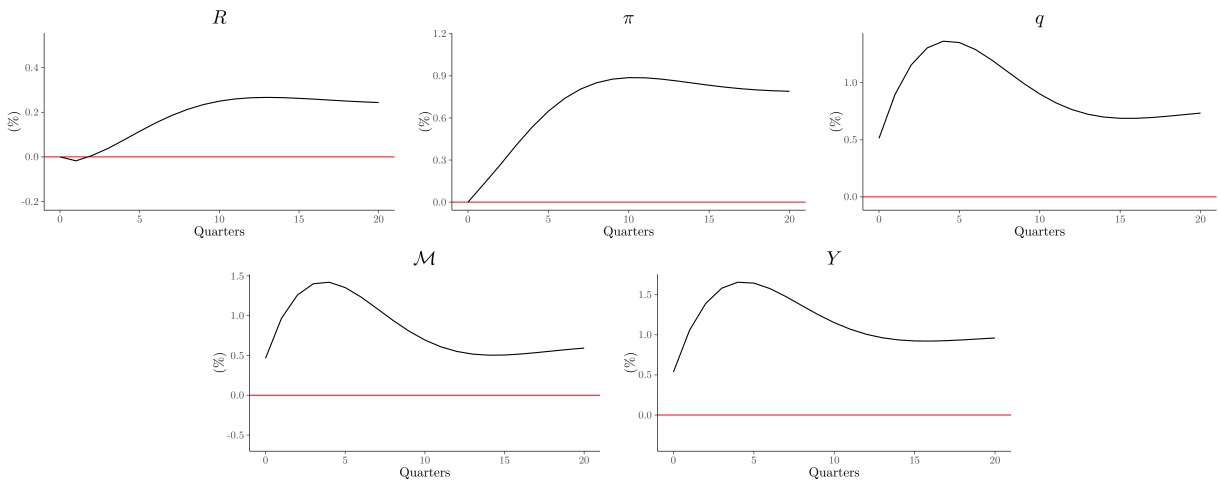
As Table 3 shows, the hypotheses until $H_0 : r \leq 2$ can be comfortably rejected at the 1% level. The hypothesis $H_0 : r \leq 3$ can be rejected only at the 5% level. This suggests that there are three cointegrating relationships in the VAR.

The estimated matrix containing the cointegrating relationships β is given by:

$$\beta = \begin{pmatrix} 1 & 1.04 \times 10^{-17} & -2.77 \times 10^{-17} \\ 0 & 1 & 0 \\ -5.55 \times 10^{-17} & 2.77 \times 10^{-17} & 1 \\ -0.637 & 0.193 & -5.72 \\ -0.18 & 0.148 & -1.28 \end{pmatrix}$$

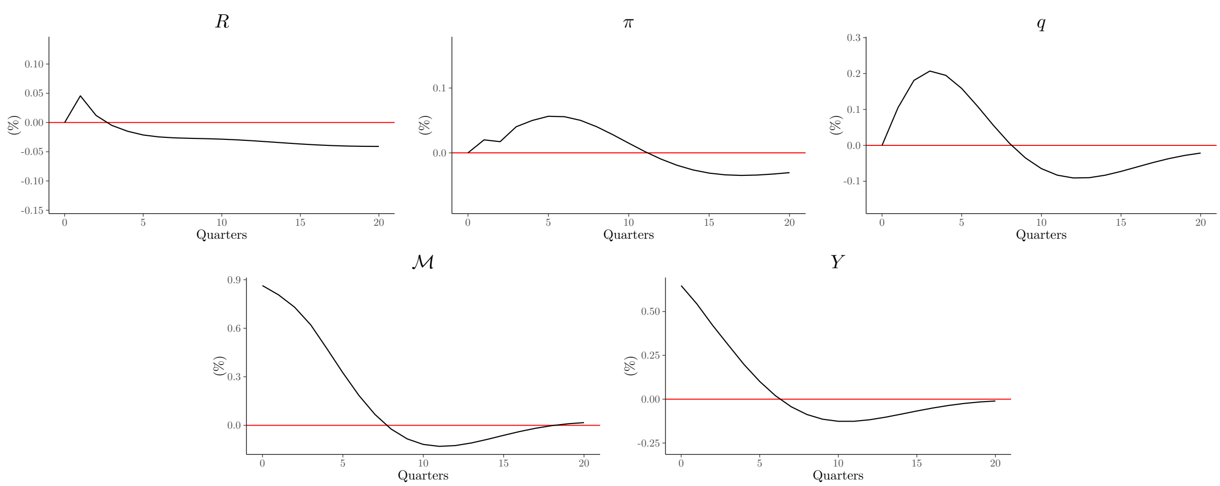
As Figures 40 - 44 show, the positive relationship between house prices and aggregate markups remains under the IRFs. The figures show the IRFs after each structural shock in the VECM. They do not include the error bands, since the usual estimation of the error bands through bootstrapping procedures yields inconsistent results. However, the figures allow to understand the general qualitative relationship between the variables in the VECM.

Figure 40: VECM – Shock to house prices.



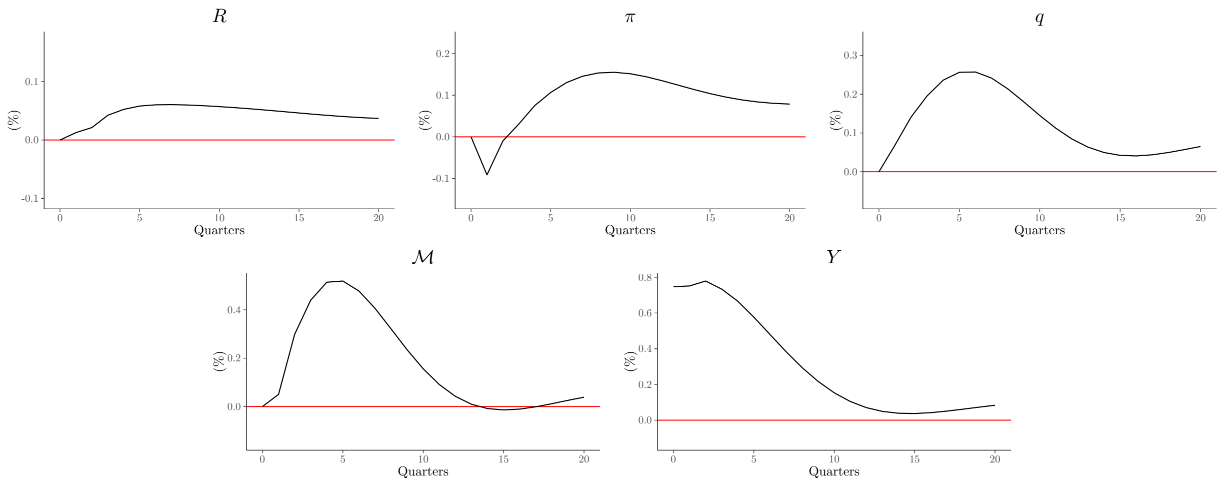
The figure shows the impulse response functions of the VECM after a shock to house prices q. The VECM contains $r = 3$ cointegrating relationships, deterministic regressors (linear trend and a constant) and exogenous variables (the monetary policy dummies and the world commodity index). The figures do not include the error bands because their estimation through bootstrapping yields inconsistent results.

Figure 41: VECM – Shock to markups.



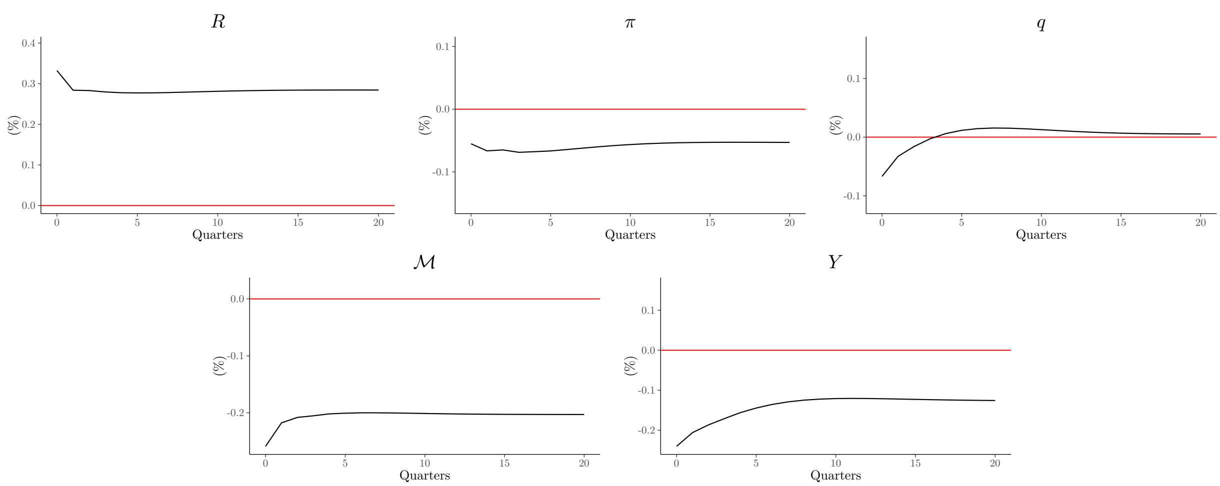
The figure shows the impulse response functions of the VECM after a shock to aggregate markups $\mathcal{M}$. The VECM contains $r = 3$ cointegrating relationships, deterministic regressors (linear trend and a constant) and exogenous variables (the monetary policy dummies and the world commodity index). The figures do not include the error bands because their estimation through bootstrapping yields inconsistent results.

Figure 42: VECM – Shock to output.



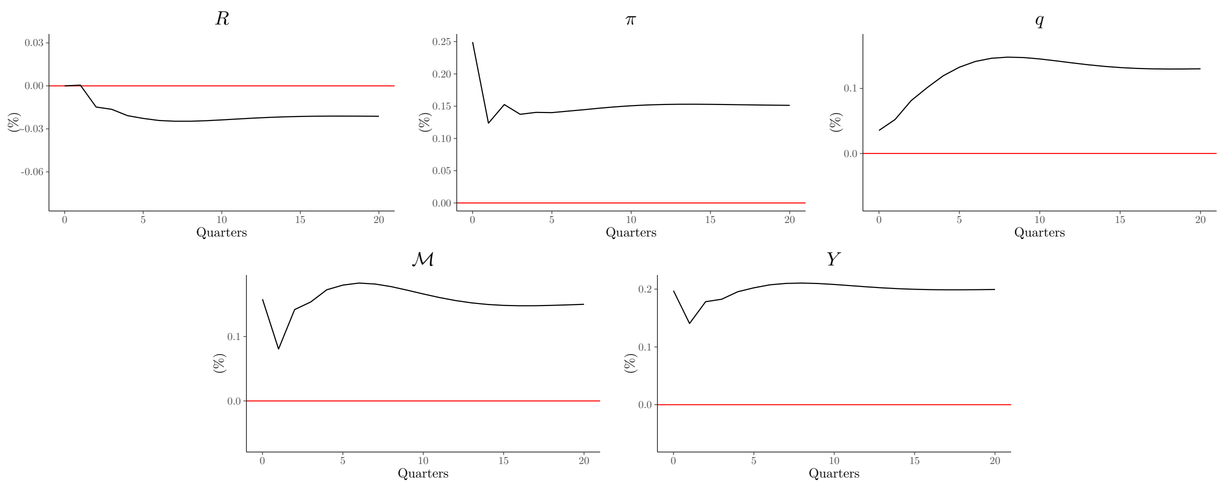
The figure shows the impulse response functions of the VECM after a shock to output Y. The VECM contains $r = 3$ cointegrating relationships, deterministic regressors (linear trend and a constant) and exogenous variables (the monetary policy dummies and the world commodity index). The figures do not include the error bands because their estimation through bootstrapping yields inconsistent results.

Figure 43: VECM – Shock to the interest rate.



The figure shows the impulse response functions of the VECM after a shock to the interest rate R. The VECM contains $r = 3$ cointegrating relationships, deterministic regressors (linear trend and a constant) and exogenous variables (the monetary policy dummies and the world commodity index). The figures do not include the error bands because their estimation through bootstrapping yields inconsistent results.

Figure 44: VECM – Shock to the inflation rate.



The figure shows the impulse response functions of the VECM after a shock to the inflation rate π. The VECM contains $r = 3$ cointegrating relationships, deterministic regressors (linear trend and a constant) and exogenous variables (the monetary policy dummies and the world commodity index). The figures do not include the error bands because their estimation through bootstrapping yields inconsistent results.