DEUTSCHE BUNDESBANK
EUROSYSTEM

Discussion Paper

Deutsche Bundesbank
No 61/2020

Monetary policy, firm exit and productivity

Benny Hartwig
(Goethe University Frankfurt)

Philipp Lieberknecht
(Deutsche Bundesbank)

Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-0

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Internet http://www.bundesbank.de

ISBN 978–3–95729–794–5 (Printversion)
ISBN 978–3–95729–795–2 (Internetversion)

Non-technical summary

Research Question

Firms' decisions to enter or leave markets shape business cycles and are affected by monetary policy. Whereas the previous literature has focused exclusively on the effect of monetary conditions on entry decisions of identical firms, this paper analyzes the importance of firm exit and heterogeneity in productivity for the transmission of monetary policy.

Contribution

This paper provides empirical evidence about the influence of monetary policy shocks on firm dynamics and productivity, using various empirical models and identification procedures. The empirical results are rationalized in a theoretical model with heterogeneous firms and endogenous entry and exit.

Results

Following an exogenous decrease of nominal interest rates, firm exit rates decrease, whereas entry rates increase. After around two years, firm exit overshoots its long-run level. From the viewpoint of the model, expansionary monetary policy shocks increase firm profits by stimulating aggregate demand. This allows low-productivity incumbent firms to remain active and new firms with low productivity to enter the market. In the model, the average productivity of surviving firms is thus lower following expansionary monetary policy shocks.

Nichttechnische Zusammenfassung

Fragestellung

Markteintritte und -austritte von Firmen sind für Konjunkturzyklen von substantieller Bedeutung. Diese Entscheidungen werden zudem von der Geldpolitik beeinflusst. Die existierende Literatur konzentriert sich bislang allerdings vornehmlich auf den Einfluss monetärer Konditionen auf die Eintrittsentscheidungen von identischen Firmen. Dahingegen untersucht dieses Forschungspapier die Wichtigkeit von Marktaustritten und unterschiedlicher Firmenproduktivität für die Transmission von geldpolitischen Impulsen.

Beitrag

Dieses Forschungspapier präsentiert empirische Evidenz zum Einfluss geldpolitischer Schocks auf Firmendynamiken und Produktivität, unter Verwendung verschiedener empirischer Modelle und Identifikationsmethoden. Die empirischen Resultate werden in einem theoretischen Modell mit heterogenen Firmen sowie endogenen Markteintritten und -austritten rationalisiert.

Ergebnisse

Nach einem exogenen Sinken des nominalen Zinses fallen gemäß der empirischen Analyse die Marktaustritte, während Eintritte steigen. Nach ungefähr zwei Jahren lässt sich ein Überschießen von Marktaustritten relativ zur Basislinie beobachten. Aus Sicht des theoretischen Modells stimulieren expansive geldpolitische Schocks die aggregierte Nachfrage und damit auch Firmenprofite. Daher können sich bereits aktive Firmen mit niedriger Produktivität am Markt halten, zudem können neue Firmen mit niedriger Produktivität in den Markt eintreten. Im Modell sinkt daher die durchschnittliche Produktivität der verbleibenden Firmen nach expansiven geldpolitischen Schocks.

DEUTSCHE BUNDESBANK DISCUSSION PAPER NO 61/2020

Monetary Policy, Firm Exit and Productivity*

Benny Hartwig
Goethe University Frankfurt

Philipp Lieberknecht
Deutsche Bundesbank

October 23, 2020

Abstract

We analyze the influence of monetary policy on firms' extensive margin and productivity. Our empirical evidence for the U.S. based on a macro-financial SVAR suggests that expansionary monetary policy shocks stimulate corporate profits, reduce firm exit and increase firm entry. In the medium run, exit overshoots the baseline. We rationalize these findings in a general equilibrium model featuring endogenous entry and exit. In the model, expansionary monetary policy shocks increase firm profits by stimulating aggregate demand and thereby allow less productive firms to remain in the market. As the monetary stimulus fades, these less productive firms become unprofitable such that exit overshoots. This exit channel of monetary policy implies a flatter aggregate supply curve and therefore amplifies output responses, but dampens inflationary effects.

Keywords: Firm exit, firm entry, extensive margin, corporate profits, monetary policy.

JEL classification: E24, E32, E52, E58, L11

*Corresponding author: philipp.lieberknecht@bundesbank.de. Contact address: Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main, Germany. We would like to thank Celine Poilly for an insightful referee report, Refet Gürkaynak for kindly sharing his dataset on FOMC announcement surprises and helpful comments, Aydan Dogan and Federico Ravenna for excellent discussions, as well as Michael Binder, Ambrogio Cesa-Bianchi, Ulrich Grosch, Sören Karau, Joachim Keller and participants of the Bundesbank's research centre seminar, the 2020 VfS Jahrestagung, the Norges Bank 23rd Central Bank Macroeconomic Modelling Workshop and the 2020 ESCB Research Cluster on Monetary Economics Workshop for helpful comments and discussions. The views expressed in this paper are those of the authors and do not necessarily coincide with the views of the Deutsche Bundesbank, the European Central Bank or the Eurosystem.

1 Introduction

It is well known that firms' extensive margin shapes business cycles (Ghironi and Melitz, 2005; Bilbiie et al., 2012). Firms decide to enter and leave markets over the cycle, as their (expected) profitability fluctuates alongside demand and production costs. As such, the extensive margin of firm entry and exit is crucial to understanding the monetary transmission mechanism. However, the previous literature has focused exclusively on the effect of monetary conditions on entry decisions of homogeneous firms (Bergin and Corsetti, 2008; Lewis and Poilly, 2012; Bilbiie et al., 2014).[1] Consequently, these studies are solely focused on one side of the extensive margin, and inherently silent about individual firm characteristics. The aim of this paper is to fill this gap by analyzing the importance of firm exit and heterogeneity for the transmission of monetary policy. While firms are heterogeneous along various dimensions, we focus on differences in idiosyncratic productivity.

Our first contribution is to provide empirical evidence about the influence of monetary policy shocks on profits, firm dynamics and aggregate productivity. To measure entry and exit, we use quarterly U.S. data from 1993:Q2–2017:Q4 on establishment births and deaths. We include these series alongside a variety of aggregate productivity measures, an aggregate corporate profit series and real wages in a structural VAR (SVAR) that we estimate using Bayesian methods. Our identification procedure relies on high-frequency FOMC announcement surprises and a modified sign-restriction approach similar to Jarociński and Karadi (2020). In particular, we identify monetary policy shocks as moving yields and stocks at high- and low-frequency in opposite directions.

We find that expansionary monetary policy shocks stimulate corporate profits and affect both sides of firms' extensive margin. Following an exogenous decrease in nominal interest rates, firm exit rates fall, whereas entry rates rise. After around two years, firm exit overshoots its long-run level and gradually reverts afterwards. Total factor productivity (TFP) increases persistently following more favorable monetary conditions, while the response of utilization-adjusted TFP is largely insignificant.

As a second contribution, we rationalize these findings in a New Keynesian DSGE model with heterogeneous firms and endogenous entry and exit. Firms are ex-ante identical and enter the market if the expected firm value exceeds entry costs. Upon entrance in the market, firms draw an idiosyncratic productivity level, which remains constant throughout their entire life-cycle. As the market is monopolistically competitive, firms with a range of productivity levels are active at any given point in time. If a firm's productivity level is below an endogenously determined cut-off, it leaves the market as its profit is smaller than zero. In turn, only firms with sufficiently high productivity remain in the market each period.

In the model, expansionary monetary policy shocks increase firm profits by stimulating aggregate demand. This allows low-productivity incumbent firms to remain active and new firms with low productivity to enter the market. As a result, firm exit rates fall and firm entry rises. In line with the empirical evidence, exit rates overshoot in the medium-run as the monetary stimulus fades. The average productivity of surviving firms is *lower* following expansionary monetary policy shocks. A model-based measure of (utilization-adjusted) productivity thus *declines*. Against this backdrop, we interpret the empirically insignificant response of utilization-adjusted productivity as reflecting these firm dynamics

[1] In these models, the equilibrium is symmetric such that all firms take the same decisions.

at the extensive margin as well as various other counteracting channels documented in the previous literature (discussed in more detail below).

We further demonstrate that the model's ability to replicate the empirical findings hinges crucially upon nominal rigidities in wage setting. Wage rigidities dampen the rise in production costs following expansionary monetary policy shocks. As a consequence, the effect of higher aggregate demand on firm profits dominates. Firms become more profitable, which lowers the cut-off level for productivity required for profitability. A model that includes only price stickiness and/or a working capital channel implies that profits decrease (ceteris paribus) and thus yields counterfactual results for firm exit.

Endogenous firm exit activates a new transmission channel for monetary policy, which we label the "exit channel". As expansionary monetary policy allows unproductive firms to remain in the market, the number of active firms increases substantially. This raises overall production capacity, causing aggregate supply to rise. In turn, the exit channel amplifies the effect of monetary policy shocks on output relative to a model in which exit occurs only exogenously. At the same time, the excess production capacity exerts downward pressure on product prices via a demand-based variety effect, thereby reducing the inflationary effect.

The survival of unproductive firms after expansionary monetary policy shocks may be labeled as "zombification" (Hoshi and Kashyap, 2004) or "sclerosis" (Caballero and Hammour, 2005). In an efficient allocation, such firms would not remain in the market. This particular effect on firm dynamics arises not only with monetary policy shocks, but whenever the central bank attempts to stabilize the economy over the business cycle. While we leave in-depth normative investigations to future research, this suggests that central banks might face a trade-off between macroeconomic stabilization and the "cleansing effect" of less productive firms in recessions à la Caballero and Hammour (1994).

This paper contributes to three strands of the literature. The first strand is the literature on firm dynamics over the business cycle in general equilibrium models, building on endogenous firm entry and exit along the lines of Hopenhayn (1992), Melitz (2003) and Ghironi and Melitz (2005).[2] Assuming exogenous exit, such models have been used to study the influence of endogenous entry on business cycles (Jaimovich and Floetotto, 2008; Bilbiie et al., 2012), the transmission of monetary policy (Lewis, 2009; Lewis and Poilly, 2012) and its optimal design (Bergin and Corsetti, 2008; Lewis, 2013; Bilbiie et al., 2014; Cacciatore et al., 2016).[3] In contrast to these papers, we investigate the transmission of monetary policy in a framework featuring both endogenous entry *and* exit.[4] This allows us to analyze the effect of monetary policy on both sides of the extensive margin and (endogenously determined) average productivity. A similar approach is taken in two recent studies by Colciago and Silvestrini (2020) and Hamano and Zanetti (2020), which are the closest papers to ours, focusing on market concentration and optimal monetary policy, respectively. In comparison, we provide state-of-the-art macroeconometric evidence on

[2] Hopenhayn (1992) considers perfect competition, whereas the latter two papers introduce monopolistic competition and focus on international trade dynamics. Firms decide endogenously whether to exit export markets, whereas exit from domestic markets is exogenous.

[3] In a different framework based on Schumpeterian creative destruction and menu costs, Oikawa and Ueda (2018) study the effect of steady state nominal growth and inflation on real growth and welfare. Similar to the above-mentioned papers, the exit rate is exogenous in their framework.

[4] In the context of endogenous entry and exit, the propagation of aggregate TFP shocks is studied by Clementi and Palazzo (2016), Hamano and Zanetti (2017, 2018), and Rossi (2019).

the effect of monetary policy shocks on firm exit and measures of productivity. We also investigate the importance of wage and price rigidities in replicating the empirical results and obtain insights into how the Phillips curve is affected by the exit channel.

The second related strand is the literature investigating zombification. Hoshi and Kashyap (2004) argue that zombification was a key driver of the Japanese "lost decade" after the stock market crash in the early 1990s. Peek and Rosengren (2005) show that banks extended credit to zombie firms to avoid painful deleveraging. Caballero et al. (2008) and Kwon et al. (2015) provide evidence that this "zombie lending" led to resource misallocation to unproductive firms and negative effects on productivity growth. However, comparably little attention has so far been devoted to the role of monetary conditions. A notable exception is Acharya et al. (2019), who find that the ECB's announcement of Outright Monetary Transactions (OMT) induced more zombie lending in the euro area.[5] Our analysis complements this largely microeconometric literature by providing macroeconometric evidence on the response of firms' extensive margin to monetary easings, alongside a dynamic structural general equilibrium model rationalizing these findings. In particular, we show that zombification may arise even without credit misallocation and deliberate decisions by banks, but naturally occurs over the business cycle as monetary conditions affect demand and costs.[6]

Third, this paper is related to the literature on monetary policy and aggregate productivity, starting with Evans (1992). In line with this literature, we find empirical evidence which favors an increase of aggregate TFP following expansionary monetary policy shocks. The explanations put forward encompass various mechanisms such as variable capital utilization (Christiano et al., 2005), incentives for R&D (Moran and Queralto, 2018; Garga and Singh, 2019), financial frictions[7] (Midrigan and Xu, 2014; Moll, 2014) and heterogeneous price pass-throughs (Meier and Reinelt, 2020). In contrast, we find no empirical evidence for a significant effect of monetary policy on utilization-adjusted productivity. Through the lens of our theoretical framework, this suggests that the *decline* in average firm-level productivity via the exit channel counteracts other channels which imply a positive effect of more favorable monetary conditions on productivity. The associated inefficiency in the allocation of production factors also links our study to the rich literature on misallocation; see, for example Restuccia and Rogerson (2008), Hsieh and Klenow (2009), Gopinath et al. (2017) and Baqaee and Farhi (2020).

The rest of the paper is structured as follows. Section 2 presents the empirical analysis. In Section 3, we outline the theoretical framework. We show in Section 4 that the model replicates the empirical findings and discuss the associated economic intuition and policy implications. Section 5 concludes and provides an outlook on future research.

[5] This is in line with further microeconometric evidence suggesting that zombie lending became generally more pervasive in the past years in the euro area, in particular in periphery countries and from weakly capitalized banks (Schivardi et al., 2017; Storz et al., 2017; Adalet McGowan et al., 2018; Blattner et al., 2019; Andrews and Petroulakis, 2019; Antoni et al., 2019; Bittner et al., 2020). Tracey (2019) develops a quantitative model with endogenous zombie lending and firm defaults, arguing that zombie lending contributed to low output growth across the euro area in recent years. Acharya et al. (2020) connect zombie lending to low inflation rates in the euro area.

[6] Within the microeconometric literature, zombie firms are identified based on weak financial characteristics or extraordinarily favorable credit conditions. In contrast, we prefer to label firms as zombies based on their low idiosyncratic productivity, which is what matters for macroeconomic effects.

[7] Standard macroeconomic models suggest that expansionary monetary policy alleviates financial frictions, see for example the canonical financial accelerator framework by Bernanke et al. (1999).

2 Empirical Analysis

In this section, we present our empirical analysis aimed at measuring the effects of monetary policy shocks on firms' extensive margin and aggregate productivity.

2.1 Data

Our overall sample covers the period from 1993:Q2 to 2017:Q4.[8] More recent observations are excluded due to data availability limitations with respect to the identification of monetary policy shocks (described in more detail below).

To capture firm entry and exit, we use U.S. data on establishment births and deaths from the Bureau of Labor Statistics (BLS).[9] These are available on a quarterly basis starting in 1993:Q2, thereby determining the start of our sample. As a measure of aggregate productivity, we use the series on aggregate total factor productivity (TFP) by Fernald (2014a). This so-called Solow residual is the most commonly used measure to assess productivity in the economy. We also report results for utilization-adjusted TFP (henceforth TFPu) – a cleaner measure of pure technological change – and labor productivity in the business sector (LP), which reflects the efficiency of labor.[10] Moreover, we use corporate profits after taxes with inventory valuation and capital consumption adjustment (IVA and CCAdj) from the Bureau of Economic Analysis (BEA) as an indicator of firms' aggregate profitability, which we deflate by the GDP implicit price deflator. We also include wages and salaries per employment of the overall economy (deflated by the GDP implicit price deflator) as a measure of real wages.

To identify the effects of a monetary policy shock, we use high-frequency data on financial surprises following monetary policy announcements. Specifically, we use yield and asset price changes identified in a tight window around 209 FOMC announcements during the period 1993:Q2 – 2017:Q4.[11] Refet Gürkaynak kindly provided us an updated version of the Gürkaynak et al. (2005) dataset. We follow Jarociński and Karadi (2020) and use changes in the 3-month federal funds future and changes in the S&P 500 index as our measure of interest rate and stock price surprises, respectively. Changes in the 3-month federal funds future reflect both surprises about actual rate-setting and near-term forward guidance and therefore constitute a broad measure of the overall monetary policy stance.

[8] In Appendix A.1, we provide time-series charts and summary descriptive statistics. We also investigate average conditional responses of the variables around recessions.

[9] An establishment is a single physical location, whereas a firm is an establishment or a combination of establishments. Rossi (2019) uses the establishment series as a proxy for firm entry and exit decisions to investigate the dynamic effects of aggregate technology shocks on firms' extensive margin.

[10] The aggregate productivity series are based on growth accounting techniques proposed by Basu et al. (2006). TFP growth is output growth not explained by (observed) input growth $\Delta TFP = \Delta Y - \alpha\Delta K - (1-\alpha)\Delta L$ where ΔY is real output growth, ΔK is capital growth, ΔL is labor growth and α is the capital share on output. Utilization-adjusted TFP growth is TFP not explained by capital and labor utilization adjustment growth $\Delta TFPu = \Delta TFP - \Delta Util$. Labor productivity growth is defined as growth in output per hour $\Delta LP = \Delta Y - \Delta H$, where ΔH is hours worked in business sector.

[11] These changes are measured in a window starting 10 minutes before and ending 20 minutes after the announcement. Gürkaynak et al. (2005) show that these intra-day changes are solely driven by FOMC announcements and are not due to confounding factors like macroeconomic releases on that day. Those major news may dilute measured surprises at a wider intra-day window and at a daily frequency.

We use the one-year government constant maturity bond yield as our indicator of monetary policy, following Gertler and Karadi (2015) and Jarociński and Karadi (2020). This measure captures the effects of forward guidance and hence moved sufficiently during the zero lower bound period, which constitutes a distinct advantage compared to the effective federal funds rate. As argued by Gürkaynak et al. (2005), forward guidance became an important tool for U.S. monetary policy since the FOMC started issuing press releases in February 1994, which almost coincides with the start of our sample.

Our block of macroeconomic variables consists of real GDP and the GDP implicit price deflator. Regarding financial variables, we include the S&P500 stock price index (deflated by the GDP price deflator) as well as the excess bond premium (EBP) of Gilchrist and Zakrajšek (2012). The EBP is included as a measure of financial frictions in the economy. Including a measure of financial frictions – proxied either by the EBP or the BAA-corporate bond credit spread – is crucial to identify the transmission channel of monetary policy (Gertler and Karadi, 2015) and the systemic rule of monetary policy in the economy (Caldara and Herbst, 2019).

2.2 Model

Our baseline empirical model is a VAR with zero coefficient restrictions for the high-frequency surprises along the lines of Jarociński and Karadi (2020). Let y_t be a vector of macroeconomic and financial variables that includes the interest rate, real GDP, the GDP implicit price deflator, the corporate spread and the stock price. We add the specific variables of interest, namely firm entry, firm exit, aggregate TFP, aggregate utilization-adjusted TFPu, labor productivity, aggregate real corporate profits and real wages on a one-by-one basis to adopt a parsimonious estimation approach. Let m_t be a vector of surprises in financial variables in quarter t, constructed as the sum of intra-day surprises occurring in quarter t on the days with FOMC announcements. m_t is zero in the quarters with no FOMC announcements.

The baseline VAR models the joint dynamics of m_t and y_t under the restriction that m_t has a zero mean and does not depend on the lags of either m_t or y_t,

$$\begin{pmatrix} m_t \\ y_t \end{pmatrix} = \begin{pmatrix} 0 \\ c_Y \end{pmatrix} + \sum_{p=1}^{4} \begin{pmatrix} 0 & 0 \\ A_{p,YM} & A_{p,YY} \end{pmatrix} \begin{pmatrix} m_{t-p} \\ y_{t-p} \end{pmatrix} + \begin{pmatrix} u_t^m \\ u_t^y \end{pmatrix}, \begin{pmatrix} u_t^m \\ u_t^y \end{pmatrix} \sim \mathcal{N}(0, \Sigma), \tag{1}$$

where $\mathcal{N}$ denotes the normal distribution. Note that the above zero restrictions are plausible as long as the financial market surprises are unpredictable.

We estimate this VAR using Bayesian methods in log-levels for all variables except interest rates, spreads and high-frequency surprises, set the maximum lag length to four and use a flat prior for our benchmark results. We use a flat prior rather than a moderately tight Minnesota prior in order to make the VAR estimates more directly comparable to estimates coming from local projections, which we present in a robustness analysis.[12]

[12] Our results are robust to using a moderately tight Minnesota prior and estimating the VAR at once using the Minnesota prior. Estimates under the flat prior of the full VAR are qualitatively similar but are associated with larger credible intervals.

2.3 Identification

A recent literature notes that high-frequency surprises may not only reflect monetary policy shocks. Miranda-Agrippino and Ricco (2018) provide evidence that these surprises may be contaminated by information about the state of the economy because of information asymmetries between private and public agents in the economy. In addition, Jarociński and Karadi (2020) document that a sizable fraction of interest rate surprises is accompanied by a positive co-movement in asset price changes, which is at odds with standard macroeconomic theory. This so-called information channel of monetary policy announcements, alternatively known as the Fed information effect (Romer and Romer, 2000; Nakamura and Steinsson, 2018), or signaling channel of monetary policy (Melosi, 2017), is therefore an important channel to control for when identifying the effects of a monetary policy shock based on high-frequency surprises. Specifically, this channel may compensate the effects of a monetary policy shock and bias the estimated impulse response function towards zero or into the opposite territory, leading to various price and output puzzles.[13] For this reason, special filtering is necessary to extract the relevant information from interest rate surprises in order to identify the effects of a monetary policy shock.

There are two prominent approaches to extracting monetary policy shocks from high-frequency surprises. Jarociński and Karadi (2020) and Cieslak and Schrimpf (2019) propose imposing sign-restrictions on the co-movement between interest rate and asset price surprises to disentangle a monetary policy shock from a central bank announcement shock. In particular, they identify a monetary policy shock as a negative co-movement shock and use orthogonality restrictions to set this shock apart from other shocks. An alternative approach is pursued by Miranda-Agrippino and Ricco (2018), who propose purging interest rate surprises from information rigidities by making them unforecastable and control for the central bank's information set, as summarized by Greenbook forecasts.

For our application, we adopt a sign restriction approach in the spirit of Jarociński and Karadi (2020). This enables us to extract the relevant information from high-frequency data on interest rate and asset price changes to identify a monetary policy shock within a multivariate time series model. Furthermore, this methodology offers us the flexibility to impose additional restrictions that we deem necessary in our specific setup to improve the identification (see Appendix A.3 for more details).[14,15]

[13] In Appendix A.2, Figure A3 shows that an expansionary monetary policy shock identified as interest rate surprise ordered first in the VAR and excluding stock market surprises (blue dashed) initially lowers real GDP and stock market valuation. These puzzling responses can be explained by central bank information shocks (black dashed-dotted), which lead to a substantial decline in these variables.

[14] In comparison to to Jarociński and Karadi (2020), we work with quarterly data and a sample that excludes 1990:M2 – 1993:M3 due to the availability of the establishment series. For our sample, there are no monthly proxies available to interpolate the data on firms' extensive margin and productivity.

[15] In Appendix A.3, we document that the interest rate response exhibits implausible dynamics when identifying monetary policy shocks in a sample starting in 1993:M4 by sign restrictions on high-frequency variables only. In particular, the monetary impulse is small and disproportionally compensated in the medium term. By comparing estimates across different data frequencies and different sample sizes, we conclude that the differences are primarily due to the exclusion of the data from 1990:M2-1993:M3 in our setup (because the entry/exit series are not available earlier). This period coincides with the U.S. savings and loan crisis, which featured large and surprising interest cuts by the Fed and associated positive stock surprises. The exclusion of these particularly informative surprises from our data set requires additional sign restrictions to ensure proper identification. In contrast, the effects of the frequency conversion to quarterly data are limited.

Our identification procedure specifies monetary policy shocks as negative co-movement shocks that move both high-frequency surprises and their low-frequency counterparts (the interest rate and the stock price) in opposite directions. While not the focus of our analysis, we also identify central bank information shocks in the same spirit to ensure that these two shocks are orthogonal to each other. Table 1 summarizes our set of identification restrictions. We furthermore enlarge the rotation space of orthonormal matrices to include the interest rate and the stock price in order to increase the set of structural models that potentially exhibit a strong link between high-frequency surprises and their low-frequency counterparts.[16,17]

Table 1: Identification restrictions

Variable	Shock		
	Monetary policy (negative co-movement)	CB information (positive co-movement)	Other
m_t, high frequency			
Interest rate surprise	+	+	0
Stock price surprise	−	+	0
y_t, low frequency			
Interest rate	+	+	0
Stock price index	−	+	0
Other	•	•	•

Note: Sign restrictions are imposed on the contemporaneous impulse response of variables to shocks. +, −, and • denote the respective sign restrictions and unrestricted response.

Our identification strategy is more stringent than the approach of Jarociński and Karadi (2020). We postulate an a priori strong link between high-frequency surprises and their low-frequency counterparts, while they remain agnostic about the effects of sign-identified high-frequency surprises on all low-frequency variables. However, we regard the additional restrictions on low-frequency variables as not being particularly restrictive for two reasons. First, standard economic theory predicts that a monetary policy shock has persistent effects on the interest rate and the stock market that are likely to prevail for several months. Similarly, the new theory developed around the effects of central bank information shocks shows that it acts as a demand shock which also triggers persistent effects in these variables. Second, our approach also remains agnostic about the potential effects on all other variables except the response of the interest rate and the stock market. Thus, our identification strategy also leaves the response to output, prices and all other variables of interest unrestricted.

[16] For our sample, sign restrictions on low-frequency variables alone are not sufficient to identify a plausible monetary transmission channel.

[17] The estimated impulse response functions of the two residual shocks are unrelated to the identified monetary policy and the central bank information shock.

2.4 Results

Figure 1 shows our baseline empirical results for an expansionary monetary policy shock. On impact, the yield on the one-year government bond decreases by roughly 10 basis points in response to the sign-identified monetary policy shock. The interest rate response is very short-lived and is soon compensated by counteracting measures of US monetary policy. Real economic activity and the aggregate price level increase in response to the expansionary monetary impulse with a delay of a couple of quarters, in line with standard economy theory. Stock prices expand over a prolonged period, while financial frictions, as measured by the excess bond premium, decline on impact, in line with the credit channel view of monetary policy. These impulse responses of key variables are very similar, albeit somewhat smoother, to the monthly estimates reported in Gertler and Karadi (2015), Miranda-Agrippino and Ricco (2018), Caldara and Herbst (2019) and Jarociński and Karadi (2020). Wages and salaries respond comparably sluggishly and rise only slowly over the medium-run, likely reflecting nominal rigidities. Furthermore, corporate profits increase persistently following an expansionary monetary policy shock, in line with Lewis and Poilly (2012).

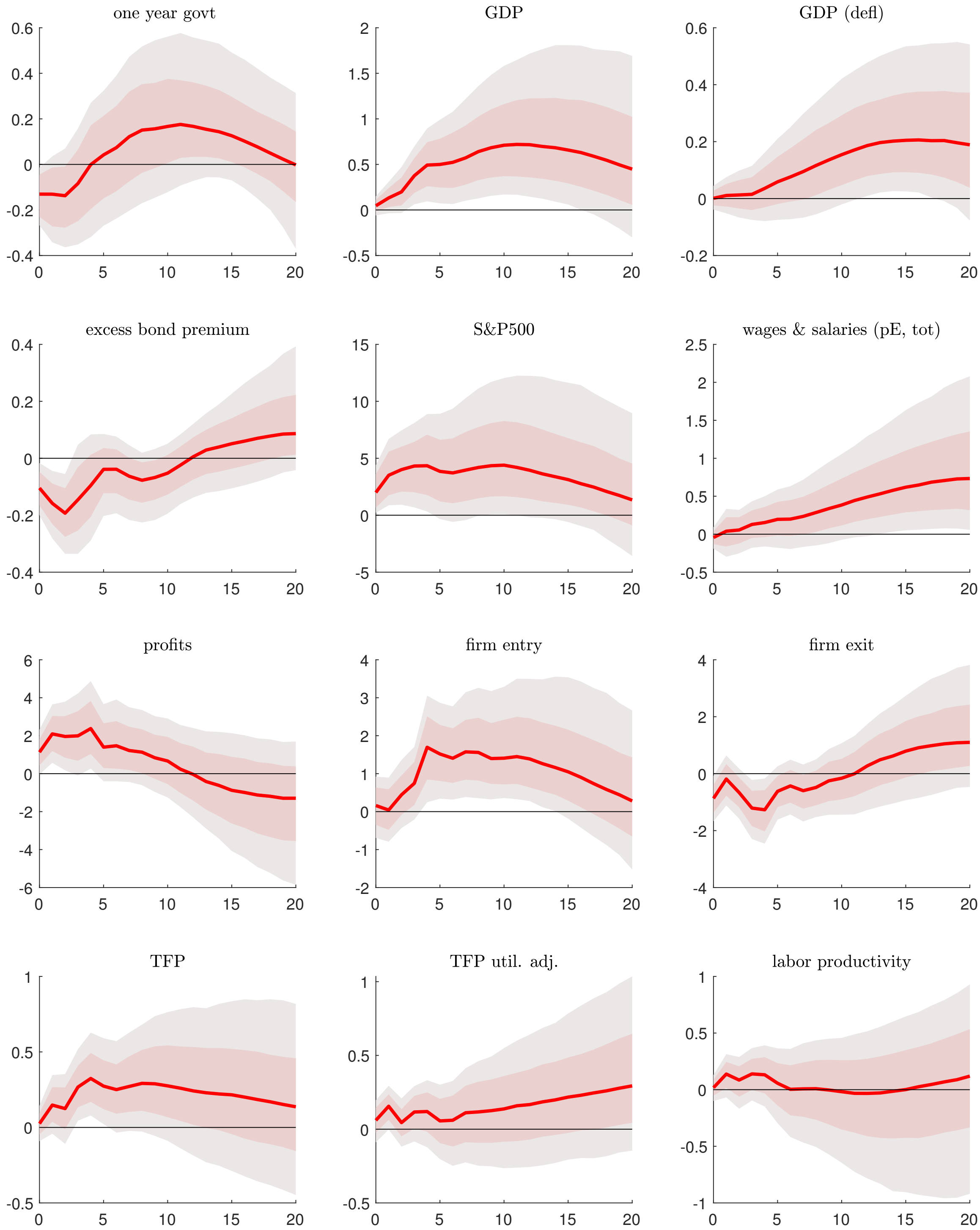
Turning to firm dynamics, our results show that a monetary policy shock affects both sides of firms' extensive margin. Firm entry is procyclical and increases following an expansionary monetary policy shock. The increase in firm entry is persistent, lasting around 3-4 years. At the same time, firm exit is countercyclical and decreases following more favorable monetary conditions. Net business formation (which may be defined as entry minus exit) peaks around 4-6 quarters after the monetary policy shock. After around two years, firm exit overshoots its long-run level and gradually reverts afterwards, while the expansion of economic activity fades.

Our results on firm entry are consistent with previous findings by Lewis (2009), Lewis and Poilly (2012) and Bergin et al. (2018), while Hamano and Zanetti (2020) document similar results for entry and exit. To proxy firms' extensive margin, all of these studies rely on measures from the BEA's Survey of Current Business and/or data compiled by the Dun & Bradstreet Corporation. However, all of these series were discontinued pre-2000. As such, our analysis provides new empirical evidence that an expansionary monetary policy shock induces a rise in firm entry in more recent times, shows that both sides of firms' extensive margin are affected and documents the notable overshooting of firm exit.[18,19]

[18] Note that these studies use different identification procedures: Lewis (2009) uses sign restrictions on conventional variables, Lewis and Poilly (2012), Bergin et al. (2018) and Hamano and Zanetti (2020) use short-run restrictions on the co-movement with activity and prices. As documented by Gürkaynak et al. (2005), forward guidance became an important short-term monetary policy instrument within our more recent sample. Therefore, pure monetary policy shocks cannot be disentangled from central bank information shocks via short-run restrictions on output and prices as they do not fully capture the central bank information set. In Appendix A.2, Figure A4 shows that a monetary policy shock identified using short-run restriction on output and prices (blue dashed) exhibits a price puzzle, which is due to the confounding effect of the central bank information shock (black dashed-dotted).

[19] It is furthermore noteworthy that the conditional countercyclical response of firm exit following monetary policy shocks is in line with firm exit being unconditionally countercyclical in our sample, see Table A1 in Appendix A.1. Campbell (1998) and Jaimovich and Floetotto (2008) also document unconditionally countercyclical firm exit as measured by business failures for earlier time periods at annual frequency.

Figure 1: VAR with FOMC announcement surprises and sign-restrictions



Note: Impulse response functions to a monetary policy shock identified by sign-restrictions on high-frequency and low-frequency variables. The thick line is the median estimate and the red and gray area depict the 68% and 90% credible intervals.

Interestingly, our estimated dynamic effects on firm exit share similarities with earlier findings for productivity shocks. Rossi (2019) documents that firm exit is initially countercyclical following an aggregate technology shock (measured by utilization-adjusted TFP), but overshoots its long-run level after approximately two years. She interprets this result as a signal that too many new entrants with low productivity enter the market during the boom. As the economy reverts to the initial equilibrium, such low-productivity firms exit the market with some delay after the initial shock. We interpret this as suggesting that our empirical results also reflect heterogeneity in firms' productivity. In our theoretical analysis below, we show that this hypothesis is perfectly consistent with firm dynamics in a heterogeneous firm general equilibrium model in the spirit of Hopenhayn (1992).

Turning to the responses of aggregate productivity, TFP is procyclical and rises following more favorable monetary conditions, whereas the responses of utilization-adjusted TFP and labor productivity are largely insignificant. Against the backdrop of the aforementioned interpretation of the overshooting property of firm exits, this result seems initially quite surprising. Anticipating our theoretical results, if particularly low-productivity firms enter the market, one would expect an initial (and perhaps persistent) decrease in (utilization-adjusted) aggregate productivity. However, it is important to note that the overall response of aggregate productivity to changes in monetary conditions reflects numerous channels, as discussed in the introduction. Whereas the vast majority of these imply an increase of productivity following expansionary monetary policy shocks, firm dynamics at the extensive margin imply the opposite. One might hence interpret our empirical findings regarding productivity as suggesting that the various counteracting channels are offsetting each other, at least over the short- and medium-run.

These results stand in contrast to those of Christiano et al. (2005), Moran and Queralto (2018) and Meier and Reinelt (2020). Christiano et al. (2005) document a significant rise in labor productivity following an expansionary monetary policy shock, while Moran and Queralto (2018) find a highly persistent increase in aggregate TFP (for about 15 years). Moreover, Meier and Reinelt (2020) document persistent effects of monetary policy shocks on all three productivity measures. In contrast, our results do not suggest that labor productivity and utilization-adjusted TFP rise significantly, while the increase in aggregate TFP is considerably less persistent. In comparing these results, the former two studies refer to earlier sample periods, which however overlap with the period when forward guidance became an important monetary policy instrument. They identify monetary policy shocks via short-run restrictions on GDP and prices and also document that their impulse response functions exhibit a price puzzle. We argue that their estimates should be interpreted with caution as their identification strategy does not control for the importance of central bank information shocks.

While Meier and Reinelt (2020) use a more recent sample, their approach differs from ours in three respects. First, they use intraday surprises around scheduled FOMC meetings only. Second, they aggregate daily surprises to quarterly frequency using the temporal aggregator of Ottonello and Winberry (2018) that aims to account for the timing of surprises in a given quarter. Third, they estimate the effects via local projections with one lag of the shock and the endogenous series. However, we find that adding macroeconomic and financial control variables in a similarly specified local projection yields insignificant estimates of utilization-adjusted TFP and labor productivity, while aggregate TFP still increases (see the next section for further details on the local projections).

2.5 Robustness

We now present a series of robustness checks regarding our baseline empirical specification.

Sign restrictions on low frequency variables: To explore whether our main empirical results are driven by the additional sign restriction on the low-frequency variables, we explore the sensitivity of our results by using the identification strategy of Jarociński and Karadi (2020), that is, imposing sign restrictions on the co-movement of high-frequency variables only. Figure A7 shows our baseline estimates (red) and those obtained when sign restrictions are only imposed on the co-movement of high-frequency surprises (blue dashed). Within our empirical setup, this identification strategy yields a rather implausible estimate for the impulse response dynamics of the interest rate. In particular, the initial impulse is small and disproportionally compensated in the medium term (see also Footnote 15). However, our main results on firms' extensive margin and the response of other key variables are not affected by this alternative identification strategy. The effects on firm dynamics and profits are even more significant, while the response of aggregate TFP is less significant but still displays procyclicality.

Estimation at monthly frequency: Our empirical specification uses quarterly data, while Jarociński and Karadi (2020) use monthly data. We hence investigate whether our results are sensitive to the frequency conversion, that is, going from monthly to quarterly frequency. For this robustness exercise, we use a monthly data set in which some variables have been interpolated by monthly proxies (if available) or by cubic splines. Specifically, we include the core variables from the Jarociński and Karadi (2020) data set, i.e. surprises, interest rate, activity, prices, excess bond premium and the stock price, in our data set and interpolate profits, wages, measures of firm dynamics and productivity by cubic splines.[20] Figure A8 shows our estimated effects of monetary policy shocks identified by using the baseline sign restrictions (red) and the sign restrictions on high-frequency variables only (blue dashed) for the data set with monthly interpolated data. Overall, our results are robust to the alternative data frequency. At the monthly frequency, the identification strategy of Jarociński and Karadi (2020) still yields implausible dynamics for the interest rate response.[21] Apart from that, the estimated dynamic effects on profits, firms' extensive margin and productivity measures are very similar to those at the quarterly frequency (compare Figure A7).

Are surprises unpredictable? Miranda-Agrippino and Ricco (2018) provide evidence that the 3-month federal funds future surprises are serially correlated and predictable by a set of macro-financial factors for the sample 1990 – 2009 at monthly frequency. We thus explore the sensitivity of our results by abandoning the zero restrictions of the VAR in equation (1) and estimate a fully parameterized VAR. Figure A9 shows our baseline estimates (red) and those of a VAR with unrestricted coefficients (blue dashed). Overall, our main results are not affected by relaxing the restrictions. However, there are some

[20] Real GDP and GDP deflator are each interpolated by industrial production and consumer prices.

[21] In Appendix A.3, we present evidence that the exclusion of the data from 1990:M2 – 1993:M3 may be responsible for the weak monetary impulse identified by sign restriction on high frequency variables only.

subtle differences. The estimated responses of firm dynamics are less pronounced in the medium-term, while their overall dynamic pattern is similar to the baseline estimates. Regarding productivity measures, the increase in utilization-adjusted TFP is now significant in the medium-term.

Surprises from scheduled FOMC announcements only: Decision from the FOMC at unscheduled meetings and conference calls occur often when economic conditions deteriorate abruptly. Consequently, these announcements may be largely unanticipated by market participants and may occur in reaction to other contemporaneous shocks (Nakamura and Steinsson, 2018). Thus, there may be a concern with respect to the proper identification of pure monetary policy shocks as these surprises may constitute an endogenous response of monetary policy. Figure A10 shows our baseline estimates (red) and the estimates when using surprises from scheduled FOMC meetings only (blue dashed). The chart shows that excluding the unscheduled decisions particularly affects the estimates of the interest rate response to a monetary policy shock. The qualitative pattern of all other variables is roughly unchanged. Therefore, we conclude that surprises from unscheduled meetings and conference calls are especially important to disentangle pure monetary policy shocks from central bank information and to identify the monetary transmission channel.

VAR with poor man's proxy: Jarociński and Karadi (2020) also propose a simpler identification of monetary policy shocks based on sign restrictions on the co-movement of the surprises in a given month.[22] We follow their approach and construct the poor man's proxy of a pure monetary policy shock at quarterly frequency. In particular, we impose sign restrictions on the sum of daily surprises in a quarter, i.e., the interest rate and stock price surprise have opposite signs, and consider this as the proxy of a pure monetary policy shock (the proxy is otherwise zero). The implicit assumption for this proxy is that each quarter features a pure monetary policy shock or a central bank information shock. In other words, we use the dominant shock in a quarter as a proxy of a pure monetary policy shock or central bank information shock.[23] We plug the poor man's proxy of monetary policy shock into the VAR with zero restrictions and identify the dynamic effects of a monetary policy shock by ordering the shock first, while imposing short-run restrictions. Figure A11 shows the results. The response of firm entry is somewhat weaker in the medium-term, while the estimated effects on firm exit almost coincides with our baseline for the first three years. The impulse responses of all productivity measures are all slightly shifted downwards. Overall, the qualitative patterns of the impulse responses for all variables remain roughly unchanged.

[22] From Figure 4 in Jarociński and Karadi (2020), it is evident that sign-identified monetary policy and central bank information shocks occur simultaneously in a month. They may move either in the same or in opposite directions.

[23] We obtain similar results when we impose a weaker version of the poor man's sign restriction that allows monetary policy shocks and central bank information to occur simultaneously in a quarter. In particular, this procedure involves imposing ad-hoc sign restrictions on (1) daily and (2) monthly surprises to identify pure monetary policy and central bank information shocks. Subsequently, the shocks are converted from the high-frequency to the low frequency by summing the identified shock series.

Poor man's proxy – VAR and local projections: A common concern with VAR estimates is that if the selected lag order is too small, the estimated impulse response may be biased for more distant lags. The local projection method of Jordà (2005) is a flexible method that overcomes this issue by imposing weaker dynamic restrictions on the data. We hence estimate the impulse responses to a monetary policy shock in a local projection model given by

$$y_{t+h} = \alpha_h + x_t\beta_h + \sum_{i=1}^{2} y_{t-i}\theta_{i,h} + \sum_{i=1}^{2} w_{i,t}\gamma_{i,h} + u_{t+h}$$

where y_t is the dependent variable, x_t is the poor man's proxy of a pure monetary policy shock and $w_{i,t}$ is a vector of additional controls. The coefficient β_h represents the response of y_t at time $t+h$ to a shock x_t at time t. To control for autoregressive dynamics of the series, two lags of the dependent variable are included. In addition, the lagged poor man's proxy, the interest rate, the log of real GDP, the log of GDP implicit price deflator, the log of the S&P500 index and the excess bond premium enter as controls with two lags. We include these additional controls to ensure that the information content in local projections and VAR with zero restrictions and the poor man's proxy are comparable.[24] We also explore sensitivity when we exclude these control variables. To control for serial correlation and heteroskedasticity of the error term u_{t+h}, the asymptotic variance is estimated using the Newey-West estimator.

Figure A12 shows the results of this alternative empirical strategy. The graph shows the estimates of the VAR with the poor man's proxy as the baseline (red), the estimates of the local projection with controls (blue dashed), and the estimates of the local projection without controls (black dashed-dotted). Overall, the estimated impulse responses of the local projection with controls are qualitatively similar, though somewhat more erratic, to our VAR estimates with a sign-identified and a poor man's identified monetary policy shock.[25] This confirms our baseline results. In contrast, the estimates without controls exhibit substantial output and price puzzles and may hence be regarded as uninformative with respect to our main variables of interest. We interpret these findings as suggesting that it is important to include macroeconomic and financial controls when estimating the dynamic effects of a monetary policy shock via local projections.

Alternative measures and sample splits: Lastly, we explore robustness with respect to the choice of variables and sample splits. Figure A13 and Figure A14 show that our baseline results are robust when using alternative measures for monetary policy, economic activity, inflation and the corporate spread.[26] Our main results are also not affected if we consider a sample up to or excluding the Great Recession, as shown in Figure A15 and Figure A16.

[24] Plagborg-Møller and Wolf (2019) show that local projections and VARs estimate the same impulse responses in a recursive VAR when the lag structure is unrestricted. In our case, this implies that for $h \leq 2$, the estimates are equivalent.

[25] The erratic pattern of local projection estimates of impulse response functions as compared to a VAR is due to a loss in efficiency in the estimation and less dynamic restrictions, see Barnichon and Brownlees (2019).

[26] The results are also robust to using alternative measures of real wages such as wages and salaries per hour or compensation per employment or per hour.

Based on our empirical analysis, we conclude that expansionary monetary policy shocks affect both sides of firms' extensive margin: Firm exit rates fall initially and overshoot in the medium run, whereas firm entry rates rise. These empirical results are robust to (a) variations in the sign restrictions, the data frequency and the VAR specification (b) the identification procedure of monetary policy shocks, (c) alternative measures of real activity, prices, spreads and monetary policy indicators and (d) sample splits accounting for the Great Recession.

3 A Model with Endogenous Firm Entry and Exit

To rationalize our empirical findings, we use a structural general equilibrium framework. Our model combines endogenous firm entry and exit à la Hopenhayn (1992), Melitz (2003) and Ghironi and Melitz (2005) with nominal rigidities and a working capital channel in the spirit of Ravenna and Walsh (2006). The economy is populated by a continuum of households that consume a variety of differentiated goods. These goods are produced by firms, who enter and exit the market according to their (expected) profits. Entry is subject to fixed entry costs. Upon entry, each firm draws an idiosyncratic productivity level for the remaining time of operation. Production requires paying fixed operational costs in advance, which firms cover by obtaining loans from financial intermediaries. Firms exit the market if their profit is non-positive. The goods market is monopolistically competitive and firms' price setting is associated with adjustment costs.

3.1 Firms

There is a continuum of firms, each producing a different good $\omega \in \Omega$ using labor as the only production factor.[27] Firm-specific productivity is given by aggregate productivity A_t and idiosyncratic productivity z, with the latter remaining constant over the entire life-cycle of the firm. The production function of a given firm can hence be written as

$$y_t^C(z) = A_t z l_t^C(z) \tag{2}$$

where $y_t^C(z)$ denotes consumption output produced by a firm with individual productivity z and $l_t^C(z)$ is the corresponding amount of labor hired. Aggregate productivity evolves according to an exogenous AR(1) process in logs:

$$ln(A_t) = \rho_A \, ln(A_{t-1}) + \varepsilon_t^A \tag{3}$$

Labor demand and price setting: Production in each period is subject to fixed operational costs f of effective labor units at the beginning of the period. At this stage, firms do not have available funds. In order to hire workers to prepare production, firms obtain loans from financial intermediaries at the nominal gross interest rate R_t. This

[27] We deliberately abstract from physical capital as a production factor to keep the model as simple as possible. As discussed by Bilbiie et al. (2012), a model with labor as the only production factor features correlations with respect to entry, profits and markups that are similar to an extended model. We note, however, that studying the interaction of idiosyncratic productivity and the accumulation of firm-specific physical capital potentially conveys interesting insights.

reflects a working capital channel in the spirit of Ravenna and Walsh (2006).[28] Total costs of production TC_t in real terms are given by

$$TC_t(z) = w_t \left(l_t^C(z) + f \frac{R_t^{\vartheta}}{A_t} \right) \tag{4}$$

where w_t is the real wage. The indicator function ϑ takes the value 1 in the presence of the working capital channel. Setting $\vartheta = 0$ eliminates this channel. Minimizing total costs subject to the production function (with Lagrange multiplier mc_t), taking aggregate variables as given yields

$$mc_t(z) = \frac{w_t}{A_t z} \tag{5}$$

which shows that firms' marginal cost (the shadow value of an extra unit of output) differs across firms, depending on idiosyncratic productivity. As outlined below, household demand for a specific good is given by

$$y_t^C(z) = \left(\frac{p_t(z)}{P_t} \right)^{-\theta} Y_t^C \tag{6}$$

where $p_t(z)$ is the nominal individual price, P_t is the aggregate price index, Y_t^C is overall consumption demand and θ is the (constant) elasticity of substitution between goods. Operating on a monopolistically competitive market, each firm maximizes profits by choosing its price subject to the individual demand schedule, taking agggregate variables as given. Firms face quadratic price adjustment costs following Rotemberg (1982). The costs of adjusting prices in real terms pac_t are

$$pac_t(z) = \frac{\tau}{2} \left(\frac{p_t(z)}{p_{t-1}(z)} - 1 \right)^2 \rho_t(z) y_t^C(z) \tag{7}$$

where

$$\rho_t(z) = \frac{p_t(z)}{P_t} \tag{8}$$

denotes the real price of firm z. The real profit of a given firm is

$$d_t(z) = \rho_t(z) y_t^C(z) - w_t l_t^C(z) - \frac{\tau}{2} \left(\frac{p_t(z)}{p_{t-1}(z)} - 1 \right)^2 \rho_t(z) y_t^C(z) - f \frac{w_t R_t^{\vartheta}}{A_t} \tag{9}$$

where the first term captures revenues, and the remaining terms are costs. The optimal real price that maximizes firms' real profits satisfies

$$\rho_t(z) = \mu_t(z) mc_t(z) \tag{10}$$

[28] We also analyzed a model variant in which the working capital channel refers to total labor input. In this case, the expression for marginal costs reads $mc_t(z) = \frac{w_t R_t}{A_t z}$. The results are qualitatively and quantitatively very similar to the baseline version considered here.

where the markup over marginal costs is given by

$$\mu_t(z) = \frac{\theta}{(\theta-1)\left[1-\frac{\tau}{2}\left(\frac{p_t(z)}{p_{t-1}(z)}-1\right)^2\right]+\tau\Upsilon_t} \tag{11}$$

where

$$\Upsilon_t(z) = \frac{p_t(z)}{p_{t-1}(z)}\left(\frac{p_t(z)}{p_{t-1}(z)}-1\right) - E_t\left[\Lambda_t \frac{y^C_{t+1}(z)}{y^C_t(z)}\frac{P_t}{P_{t+1}}\left(\frac{p_{t+1}(z)}{p_t(z)}-1\right)\left(\frac{p_{t+1}(z)}{p_t(z)}\right)^2\right] \tag{12}$$

and Λ_t denotes the representative household's stochastic discount factor. Optimal prices are thus heterogeneous across firms, as both marginal costs and optimal markups are different. Note that the markup would be identical for all firms in the absence of nominal rigidities for $\tau = 0$ and equivalent to the familiar expression $\theta/(\theta-1)$. Using the optimal pricing condition, real profits of an individual firm can then be written as:

$$d_t(z) = \left(1-\frac{1}{\mu_t(z)}-\frac{\tau}{2}\left(\frac{p_t(z)}{p_{t-1}(z)}-1\right)^2\right)\rho_t(z)^{1-\theta}Y^C_t - f\frac{w_t R^\vartheta_t}{A_t} \tag{13}$$

Entry and exit: Each period, firms enter and exit the market depending on their (expected) profitability. There is an unbounded mass of ex-ante homogeneous prospective entrants. When entering the market, each firm draws an idiosyncratic productivity level z from a distribution $G(z)$ with support on $[z_m, \infty)$ and start to produce in the next period after some time to build. Entry is subject to sunk entry costs f_E in terms of effective labor units. In terms of consumption goods, the entry costs are hence given by $f_E\frac{w_t}{A_t}$. Potential entrants are forward-looking and decide to enter if the expected post entry value of operation v_t is sufficiently high relative to entry costs. In equilibrium, firm entry equates the expected firm value with entry costs.

Following Lewis and Poilly (2012), we assume that a fraction of firm entries is unsuccessful. The success probability is given by

$$\Psi_t(H_t, H_{t-1}) = 1 - F_{H,t}\left(\frac{H_t}{H_{t-1}}\right) \tag{14}$$

which has the properties $F_H(1) = F'_H(1) = 0, F''_H(1) = \psi > 0$. The free entry condition is then given by

$$f_E\frac{w_t}{A_t} = v_t(\Psi_t + \Psi'_t H_t) + \beta E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-1} v_{t+1}\Psi''_{t+1}H_{t+1}\right] \tag{15}$$

where the left-hand side captures entry costs, and the right-hand side the expected firm value.

Turning to the other side of firms' extensive margin, an incumbent firm decides to stay in the market if its profits are positive, i.e. if $d_t(z) > 0$, and leaves the market if profits are zero or negative. As such, only a subset of firms $\Omega_t \in \Omega$ are actively producing in any given period. The exit decision depends on firms' idiosyncratic productivity. The cutoff level of productivity $\bar{z}_t$ is defined by a zero profit condition given by:

$$\bar{d}_t \equiv d_t(\bar{z}_t) = 0 \tag{16}$$

Firms with $z > \bar{z}_t$ make positive profits and thus remain in the market, whereas low-productivity firms with $z \leq \bar{z}_t$ decide to exit the market. As a result, only relatively more productive firms are actively producing in any period, and exit takes place endogenously. In addition, each firm faces an exogenous exit shock at the end of each period, which occurs with probability δ independently of idiosyncratic productivity. Firm entry and exit thus implies that the total number of firms N_t evolves according to

$$N_t = (1-\delta)(N_{t-1} + \Psi_{t-1}(H_t, H_{t-1})H_{t-1}) \tag{17}$$

where H_t denotes the total number of new firms entering the market in a given period. The number of active or *surviving*[29] firms is given by

$$S_t = (1 - G(\bar{z}_t))N_t \tag{18}$$

which shows that some firms decide to exit the market endogenously due to non-positive profits.

3.2 Households

There is a continuum of infinitely-lived identical and atomistic households. The representative household gains utility from consumption and leisure, and maximizes expected utility U_t given by

$$U_t = E_t\left[\sum_{s=t}^{\infty}\beta^{s-t}\left(\varepsilon_{t+s}^{C} log(C_t) - \chi\frac{L_t^{1+\frac{1}{\eta}}}{1+\frac{1}{\eta}}\right)\right] \tag{19}$$

where C_t is consumption, L_t denotes hours of work supplied by the household and ε_t^C is an exogenous preference (demand) shock following an AR(1) process.[30] The discount factor is given by β, and η is the Frisch elasticity of labor supply to wages.

Consumption is defined as a basket consisting of individual goods/varieties ω over a continuum of goods Ω. In each period, as outlined above, only a subset $\Omega_t \in \Omega$ is available on the market. Consumption preferences follow Dixit and Stiglitz (1977), such that the elasticity of substitution between individual goods is constant. The consumption aggregator is then given by

$$C_t = \left(\int_{\omega\in\Omega_t} c_t(\omega)^{\frac{\theta-1}{\theta}} d\omega\right)^{\frac{\theta}{\theta-1}} \tag{20}$$

where $c_t(\omega)$ is the demand for individual good ω and θ is the constant elasticity of sub-

[29]We use the terms active and surviving firms as synonyms in our analysis.

[30]The per-period utility function follows King et al. (1988). Separable preferences with logarithmic utility over consumption ensure constant labor supply in steady state and the existence of a balanced growth path as income and substitution effects of real wage changes due to productivity growth cancel.

stitution. The optimal price index minimizing overall consumption expenditure is

$$P_t = \left(\int_{\omega \in \Omega_t} p_t(\omega)^{1-\theta} d\omega \right)^{\frac{1}{1-\theta}} \tag{21}$$

where $p_t(\omega)$ is the nominal price of good ω. Demand for individual variety is then obtained as

$$c_t(\omega) = \rho_t(\omega)^{-\theta} C_t \tag{22}$$

which represents the demand schedule faced by firms used above.

Households can invest in (government) bonds and equity shares in a mutual fund of firms. Holding bonds yield a safe gross nominal interest rate R_t in the next period. The mutual fund pays out dividends equivalent to the total profit in the firm sector in each period. In period t, the representative households obtains equity shares x_{t+1} from the universe of firms, which encompasses both incumbent firms and new firms.[31] The real share price is given by v_t, the expected firm value. In addition to interest income and dividend income, the household receives income by selling its initial share holdings and by supplying labor at the real wage w_t. The household allocates total income on consumption and new holdings of bonds and firm equity shares. The budget constraint in real terms can hence be written as

$$C_t + x_{t+1} v_t (N_t + H_t) + B_{t+1} = w_t L_t^S + x_t N_t v_t + x_t S_t \widetilde{d}_t + \frac{R_{t-1}}{\pi_t^C} B_t \tag{23}$$

where B_t are real holdings of bonds, L_t^S is labor supply, $\widetilde{d}_t$ is the average dividend income across active firms and π_t^C denotes the gross consumption-based inflation rate:

$$\pi_t^C = \frac{P_t}{P_{t-1}} \tag{24}$$

The household maximizes expected utility by choosing consumption, labor supply and its portfolio allocation on bonds and equity subject to the budget constraint in Equation (23). The first-order condition with respect to bond holdings is a standard Euler equation given by:

$$1 = \beta E_t \left[\frac{\varepsilon_{t+1}^C}{\varepsilon_t^C} \left(\frac{C_{t+1}}{C_t} \right)^{-1} \frac{R_t}{\pi_{t+1}^C} \right] \tag{25}$$

The optimality condition with respect to share holdings is given by

$$v_t = E_t \left[\Lambda_{t+1} \left(v_{t+1} + \frac{S_t}{N_t} \widetilde{d}_{t+1} \right) \right] \tag{26}$$

where the stochastic discount factor Λ_{t+1} is defined by

$$\Lambda_{t+1} = \beta (1-\delta) E_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{-1} \right] \tag{27}$$

[31] By assumption, the household does not know which firms operate in the next period. As a result, it finances all incumbent and new firms during a given period.

and the fraction of active firms S_t/N_t accounts for the evolution of firms.

The labor market is monopolistically competitive, such that households have some market power and act as price-setters of their wages. The differentiated labor supplied by each households is aggregated by a labor union and hired by firms on a competitive market. Indexing households by j, the labor aggregator is

$$L_t = \left(\int_0^1 \left(L_t^S(j) \right)^{\frac{\theta_W - 1}{\theta_W}} dj \right)^{\frac{\theta_W}{\theta_W - 1}} \tag{28}$$

where θ_W is the constant elasticity of substitution. The demand for individual labor is then given by

$$L_t^S(j) = \left(\frac{w_t(j)}{w_t} \right)^{-\theta_W} L_t \tag{29}$$

In each period, only a fraction of $1 - \lambda_W$ households can reoptimize their wages. The optimal wage w_t^* that is chosen by households that are able to reoptimize satisfies the following first order conditions:

$$g_t = \frac{\theta_W - 1}{\theta_W} \left(w_t^* \right)^{1-\theta_W} w_t^{\theta_W} \varepsilon_t^C C_t^{-1} L_t + \beta \lambda_W E_t \left[\left(\pi_{t+1}^C \frac{w_{t+1}^*}{w_t^*} \right)^{\theta_W - 1} g_{t+1} \right] \tag{30}$$

$$g_t = \chi \left(\frac{w_t}{w_t^*} \right)^{\theta_w \left(1 + \frac{1}{\eta}\right)} L_t^{1+\frac{1}{\eta}} + \beta \lambda_W E_t \left[\left(\pi_{t+1}^C \frac{w_{t+1}^*}{w_t^*} \right)^{\theta_W \left(1+\frac{1}{\eta}\right)} g_{t+1} \right] \tag{31}$$

The real wage then evolves according to:

$$w_t = \left(\int_0^1 \left(w_t(j) \right)^{1-\theta_W} dj \right)^{\frac{1}{1-\theta_W}} = \left(\lambda_W \left(\frac{w_{t-1}}{\pi_t^C} \right)^{1-\theta_W} + (1 - \lambda_W) \left(w_t^* \right)^{1-\theta_W} \right)^{\frac{1}{1-\theta_W}} \tag{32}$$

3.3 Aggregation

Following Melitz (2003) and Ghironi and Melitz (2005), we assume that individual firm productivity is drawn from a Pareto distribution

$$G(z) = 1 - \left(\frac{z_m}{z} \right)^{\kappa} \tag{33}$$

where z_m is the minimum possible productivity level and κ governs the shape and dispersion of the distribution. Since the cutoff level of productivity $\bar{z}_t$ for zero profits varies over the business cycle, the average productivity across surviving firms is time-varying as well and given by:

$$\widetilde{z}_t \equiv \left[\frac{1}{1 - G(\bar{z}_t)} \int_{\bar{z}_t}^{\infty} z^{\theta - 1} dG(z) \right]^{\frac{1}{\theta - 1}} = \bar{z}_t \left[\frac{\kappa}{\kappa - (\theta - 1)} \right]^{\frac{1}{\theta - 1}} \tag{34}$$

Similarly, variables referring to firms with average productivity are denoted with a tilde in the following, i.e. $\widetilde{a}_t \equiv a(\widetilde{z}_t)$ for a generic variable a.

Firm averages: Using the definition of average productivity, the average markup is given by:

$$\widetilde{\mu}_t = \frac{\theta}{(\theta-1)\left(1-\frac{\tau}{2}(\pi_t-1)^2\right)+\tau\left(\pi_t(\pi_t-1)-E_t\left[\Lambda_{t+1}\frac{Y_{t+1}^C}{Y_t^C}\frac{S_t}{S_{t+1}}(\pi_{t+1}-1)\pi_{t+1}\right]\right)} \tag{35}$$

Equation (35) is the non-linear Phillips curve relationship in our model, relating average markups to producer price inflation π. The relationship between producer price inflation π_t and consumption-based inflation rate is given by:

$$\pi_t = \frac{\widetilde{\rho}_t}{\widetilde{\rho}_{t-1}}\pi_t^C \tag{36}$$

As in Bilbiie et al. (2007), one can show that a log-linear version of this Equation (35) reduces to an augmented New Keynesian Phillips curve. In contrast to their model with exogenous exit, the number of *surviving* firms (S) is a crucial determinant of inflation dynamics, instead of the total number of firms (N). We discuss the implications for inflation dynamics in Section 4.4.

The average profit of surviving firms can be written as:

$$\widetilde{d}_t = \left(1-\frac{1}{\widetilde{\mu}_t}-\frac{\tau}{2}(\pi_t-1)^2\right)\widetilde{\rho}_t^{1-\theta}Y_t^C - f\frac{w_t R_t^{\vartheta}}{A_t} \tag{37}$$

The real price, markup, marginal costs and profits of firms with cutoff-productivity level $\bar{z}_t$ are defined similarly. The number of surviving firms is given by

$$S_t = (1-\zeta_t)N_t \tag{38}$$

where the endogenous exit probability ζ_t due to low productivity is:

$$\zeta_t \equiv 1-G(\bar{z}_t) = 1-\left(\frac{z_m}{\bar{z}_t}\right)^{\kappa} \tag{39}$$

The Pareto distribution of individual productivity also implies that output at the cutoff is proportional to average output:

$$\bar{y}_t = \left(\frac{\bar{z}_t}{\widetilde{z}_t}\right)^{\theta}\widetilde{y}_t \tag{40}$$

Market clearing: Equilibrium on the goods market requires that aggregate consumption output equals private consumption plus price adjustment costs:

$$Y_t^C = C_t + S_t\widetilde{pac}_t = \left(1-\frac{\tau}{2}(\pi_t-1)^2\right)^{-1}C_t \tag{41}$$

Equilibrium on asset markets requires

$$x_{t+1} = x_t = 1 \tag{42}$$

and that bonds are in zero net supply:

$$B_{t+1} = B_t = 0 \tag{43}$$

Using these equilibrium conditions in the household budget constraint yields the aggregate accounting identity

$$C_t + v_t H_t = w_t L_t + S_t \widetilde{d}_t \tag{44}$$

which shows that aggregate output (consumption + investment in new firms) is equal to labor and dividend income. Aggregate output and investment are defined as

$$Y_t = C_t + I_t \tag{45}$$

and

$$I_t = v_t H_t \tag{46}$$

Equilibrium on the labor market requires that:

$$L_t = S_t \left(\widetilde{l}_t^C + \frac{f}{A_t} \right) + H_t \frac{v_t}{w_t} \tag{47}$$

where $\widetilde{l}_t^C = l_t^C(\widetilde{z}_t)$ is the labor input used for production of the average firm. Finally, one can show that the price index satisfies

$$\widetilde{\rho}_t = S_t^{\frac{1}{\theta-1}} \tag{48}$$

such that the average real price captures a variety effect that stems from consumers preferences. This implies that aggregate consumption output can be written as:

$$Y_t^C = S_t \widetilde{\rho}_t \widetilde{y}_t = A_t \widetilde{z}_t \widetilde{\rho}_t L_t^C \tag{49}$$

where $L_t^C = S_t \widetilde{l}_t^C$ is total labor input used for production of consumption goods.

Table A2 in the Appendix provides an overview of the equilibrium equations. The model can be closed by specifying the conduct of monetary policy. We assume that the central bank operates according to an interest rate rule given by:

$$log\left(\frac{R_t}{R}\right) = \phi_R \, log\left(\frac{R_{t-1}}{R}\right) + (1-\phi_R)\left[\phi_\pi \, log\left(\frac{\pi_t}{\pi}\right) + \phi_y \, log\left(\frac{Y_t}{Y_{t-1}}\right)\right] + \varepsilon_t^M \tag{50}$$

The central bank thus responds to deviations of producer price inflation from steady state and output growth.[32] ε_t^M is a white-noise monetary policy shock.

[32]As discussed by Bilbiie et al. (2007), a response to welfare-based CPI inflation π^C is not feasible in reality due to infrequent updating of consumption baskets and adjustments for availability of new varieties. Actual CPI data is closer to p_t than P_t. See also Aghion et al. (2019) for a discussion how firm entry and exit raise difficulties for accurately measuring inflation and growth.

4 Monetary Policy in the Model

In this section, we show how monetary policy affects business cycle dynamics and firms' extensive margin in our benchmark model with endogenous entry and exit. We first demonstrate that our benchmark model implies that expansionary monetary policy shocks decrease firm exit and increase firm entry, in line with the empirical evidence. We then discuss the model's implications about average and aggregate productivity. We further outline the role of corporate profits and various model features in replicating the empirical results. Lastly, we compare the transmission of monetary policy in our baseline model to its counterpart with exogenous exit.

4.1 Calibration

The calibration for the following numerical analysis is based on standard values in the literature and estimates for the US economy. In line with the empirical analysis, we interpret periods as quarters and set $\beta = 0.99$. equivalent to an annualized steady state real interest rate of 4 percent. We consider a steady state gross inflation rate of one. Regarding the household preference parameters, we set the Frisch elasticity to labor supply $\eta = 2$ and calibrate χ such that labor supply in steady state is normalized to one.[33]

With respect to the firm parameters, the entry cost f_E and the minimum productivity level z_m are set to unity, without loss of generality. We follow Ghironi and Melitz (2005) and calibrate the elasticity of substitution between goods $\theta = 3.8$ and the shape parameter of the productivity Pareto distribution $\kappa = 3.4$. As in Hamano and Zanetti (2018), we calibrate the steady state fixed costs f and the exogenous exit rate δ to match U.S. exit and entry rates. According to the Business Dynamics Statistics (BDS), average annual exit and entry rates were 10.6% and 12.3% over 1977-2016. Using this entry rate in the firm law of motion (17) implies $\delta = 0.0299$. Together with the overall exit rate, this yields steady state costs of $f = 0.0090$ and a steady state ratio between average and cut-off productivity of $\widetilde{z}/\bar{z} = 1.86$.[34] Following Lewis and Poilly (2012), we calibrate the firm entry cost parameter $\psi = 8.31$. The presence of the working capital channel requires $\vartheta = 1$.

Turning to the parameters referring to nominal rigidities, the elasticity of substitution between differentiated labor is set to $\theta_W = 21$, implying a steady state wage markup of 1.05 as in Christiano et al. (2005). With respect to nominal rigidities, we calibrate the fraction of non-adjusting firms $\lambda_W = 0.75$. The Rotemberg price adjustment parameter τ is set to $\tau = 77$, in line with Bilbiie et al. (2014). The parameters in the monetary policy rule are calibrated as $\phi_R = 0.8$, $\phi_\phi = 1.5$, $\phi_y = 0.5/4$, which are standard values in the literature. Regarding the exogenous shock processes, we calibrate their persistence to $\rho_A = \rho_C = 0.9$, roughly reflecting estimates for the U.S. economy by Del Negro et al. (2015).

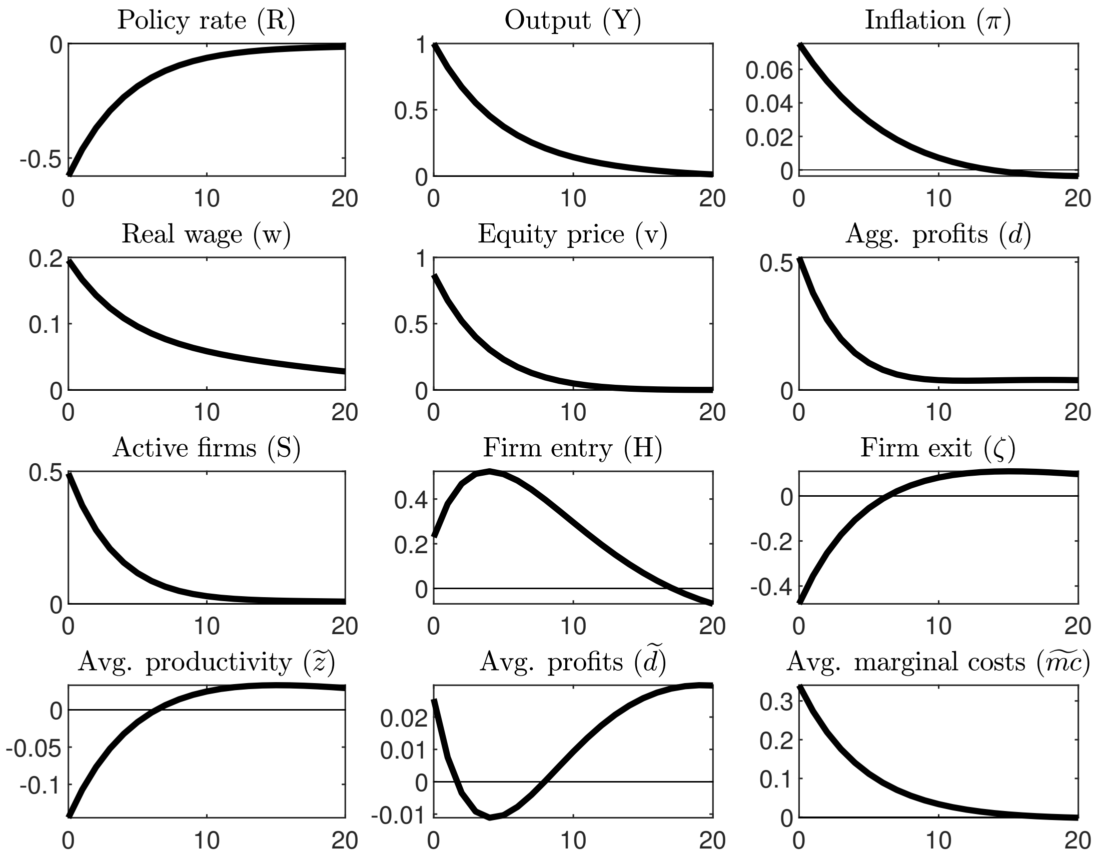
[33] For the benchmark calibration, this implies $\chi = 0.9003$.

[34] The steady state implies that f is given by $f = \beta^{\vartheta} \frac{\kappa-(\theta-1)}{\theta-1} \frac{1-\Lambda}{\Lambda} \left(\frac{S}{N}\right)^{-1} f_E$.

4.2 Monetary Policy Shocks and the Role of Profits

We first analyze how monetary policy shocks are transmitted in the benchmark economy. Figure 2 shows the impulse responses to a monetary policy shock with a negative sign that decreases the nominal interest rate.[35]

Figure 2: Monetary policy shock in the baseline model



Note: Impulse response functions for an expansionary monetary policy shock in the baseline economy. The shock size is calibrated to yield a one-percent increase of output. Inflation, interest rate and exit are shown in percentage-point deviations from steady state, all other variables in percentage deviations.

The monetary policy shock is expansionary in terms of output and increases inflation.[36] As the real interest rate falls, households increase consumption and reduce saving in bonds.

[35]The monetary policy variable in the model is the short-term nominal interest rate on one-period bonds. In the VAR, we use the one-year government constant maturity bond yield as our main indicator of monetary policy. Although the two measures are closely related, they are not identical, and their responses to monetary policy shocks are expected to differ. Therefore, and because we are mainly interested in qualitative implications, we only aim to make empirical and theoretical impulse responses somewhat quantitatively comparable by considering shocks that are associated with roughly identical output effects.

[36]In this graph and in the following, inflation is generally shown in terms of producer prices. Within our closed economy model, producer price inflation corresponds to GDP deflator inflation, which is the variable employed in our empirical exercise.

Firms demand more labor to accommodate the higher demand for consumption goods. The tighter labor market implies increasing real wages. The expansionary effect on GDP is accompanied by a procyclical reponse of the number of active firms; firm entry rises and firm exit falls. Within the baseline model, the response of firms' extensive margin is thus consistent with the empirical evidence. Likewise, corporate profits increase in response to the more favorable conditions, in line with the empirical results.

To understand the economic intuition behind these results within the model, consider the equation for average firm profits, which can be written as:

$$\widetilde{d}_t = \underbrace{\frac{Y_t^C}{S_t}}_{(1)} - \underbrace{\frac{\tau}{2}\left(\frac{\widetilde{p}_t}{\widetilde{p}_{t-1}} - 1\right)^2 \frac{Y_t^C}{S_t}}_{(2)} - \underbrace{w_t \widetilde{l}_t^C}_{(3)} - \underbrace{f\frac{w_t R_t^{\vartheta}}{A_t}}_{(4)} \tag{51}$$

One can decompose the firm profits into various channels: (1) A direct demand channel, (2) a price adjustment cost channel, (3) a labor cost channel and (4) a fixed cost channel. The relative importance of these channels govern how (average) firm profits respond to monetary policy (and other aggregate shocks). In turn, they determine the firm value via the asset pricing formula in (26) as well as corresponding entry and exit decisions.

Regarding the first channel, an expansionary monetary policy shock stimulates aggregate consumption demand by households by decreasing the real interest rate. As a result, demand for the individual variety increases as well. This raises firm profits and firm value directly. The additional demand is – ceteris paribus, i.e. before price changes and general equilibrium effects – distributed proportionally across all firms, such that all firms benefit equally.

The other three channels constitute costs associated with production, thus decreasing profits and reducing the average firm value. The second channel is due to the costs of adjusting prices, which are proportional to real revenues. As such, adjustment costs are positive when firms increase prices. This channel thus reduces profits following monetary policy shocks. However, the quantitative importance of this effect is limited, because adjustment costs are a squared function of price increases (as also argued by Bilbiie et al., 2014). The third channel represents labor costs of production. The higher labor demand and wages imply that firms face a higher wage bill following the monetary expansion, which reduces profits. In contrast, the overall effect of the fourth channel is ambiguous. While wages rise, the borrowing rate decreases one-to-one after the monetary policy shock.

As depicted in Figure 2, aggregate and average corporate profits rise after an expansionary monetary policy shock. This indicates that – in the baseline economy – the direct demand channel (1) dominates the costs channels (2)–(4). As a result, firms become more profitable on impact. This lowers the firm-specific productivity threshold level that guarantees non-negative profits. Incumbent low-productivity firms thus remain active in the market, such that average (aggregate) productivity ($\widetilde{z}_t$) *decreases*. At the same time, the higher corporate profitability increases firms' expected value and thereby render equity investment more attractive to households. This induces more firms to enter the market. Among these new entrants, low-productivity startup firms also make positive profits and produce actively. In general equilibrium, average firm profits only rise marginally because actively producing firms charge higher prices ($\widetilde{\rho}_t$) which *lowers* average profits (because the elasticity of substitution $\theta > 1$), and because they face higher marginal costs (because of lower average productivity).

The expansionary monetary policy shock thus impedes the destruction of less efficient firms along the business cycle, and leads to entry of relatively unproductive firms. As a result, the aggregate productivity of active firms is lower than before.[37] Against this backdrop, we interpret the empirically insignificant response of utilization-adjusted productivity as reflecting these firm dynamics at the extensive margin as well as various other counteracting channels documented in the previous literature (as discussed in the introduction).

Overall, Figure 2 shows that the baseline model successfully replicates the empirical SVAR findings. Following expansionary monetary policy shocks, corporate profits and firm entry increase, whereas less firms exit the market initially. Similar to the empirical results, the model also features a persistent overshooting of firm exit in the medium-run, after around two years. In the model, the monetary stimulus induced economic boom incentivices the entry of new firms, and some of those draw a relatively low idiosyncratic productivity level. While the favorable monetary conditions prevail, these firms are profitable and thus remain in the market. However, as the monetary stimulus and the associated boom fade, the cut-off level of productivity for profitability increases again. As a result, low-productivity firms default and exit the market, leading to an overshooting of firm exit. Through the lens of the model, the exit overshooting thus reflects idiosyncratic productivity levels of firms at the extensive margin.

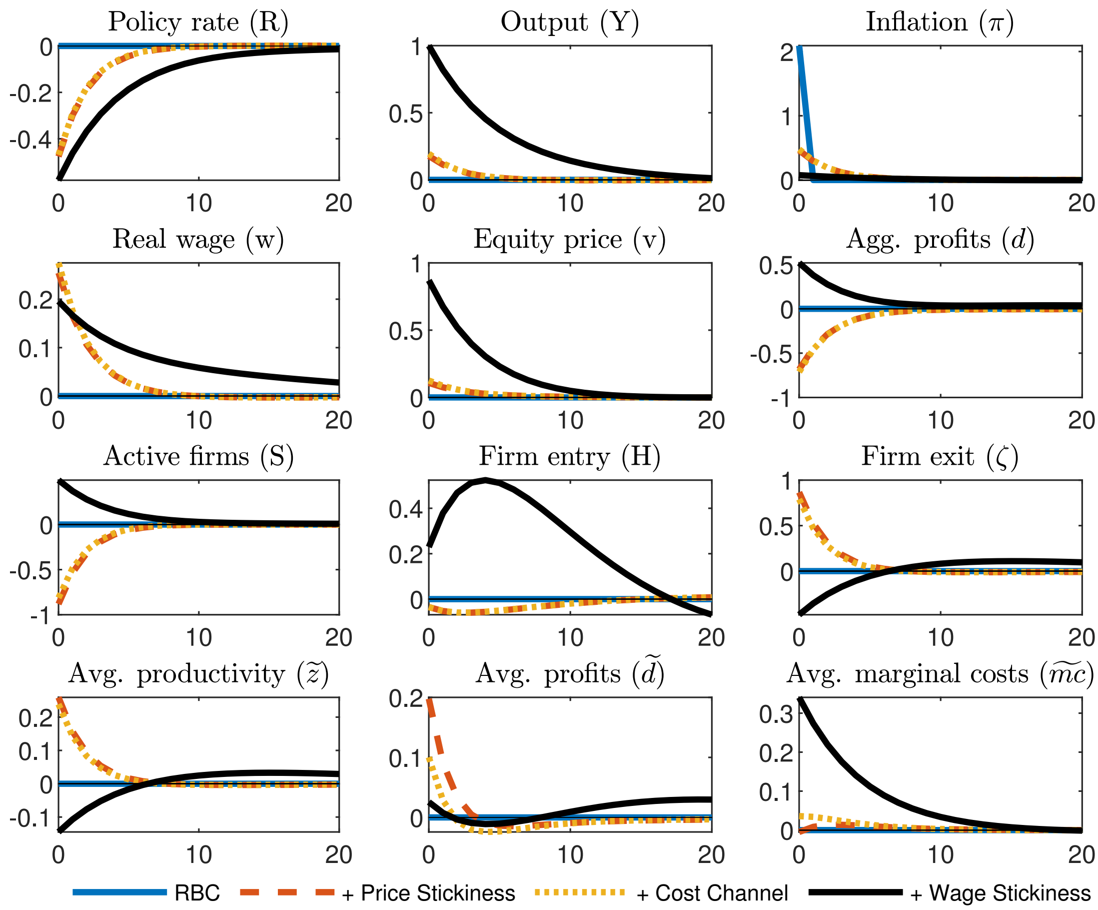
4.3 The Role of Wage Stickiness and Further Model Features

We now investigate the role of the various model features for the model's ability to replicate the empirical findings. To disentangle the effects of the various channels, we compare different variants of the baseline economy in Figure 3. The baseline model is again shown with black solid lines. We start from an RBC model without nominal rigidities and the cost channel (blue lines, $\tau = 0, \lambda_W = 0, \vartheta = 0$). In the absence of nominal rigidities, monetary policy shocks have no real effects. Prices adjust such that the initial equilibrium persists and firm dynamics are unaffected.

A New Keynesian variant with nominal rigidities in prices (red dashed lines, $\tau > 0, \lambda_W = 0, \vartheta = 0$) implies real effects of monetary policy and an associated influence on firms' extensive margin. However, firm exit is procyclical and firm entry is countercyclical in this variant, which contradicts our empirical evidence. To understand this observation, note that the labor cost channel is tightly linked to the markup decision of individual firms (see Equation 13). In the presence of price adjustment costs, optimal markups are inversely related to inflation. As firms increase prices following expansionary monetary policy shocks, markups decrease. The resulting downward pressure on profits makes relatively less productive firms unprofitable, such that they leave the market and firm exit increases. At the same time, the rise in real wages implies that entry costs are higher,

[37] An alternative model-based measure of aggregate TFP is given by $A_t \widetilde{z}_t \widetilde{\rho}_t = A_t \widetilde{z}_t S_t^{\frac{1}{\theta-1}}$. This measure of TFP explicitly accounts for the variety effect from the larger range of available goods. In contrast, empirical consumption baskets typically do not adjust for the availability of new products in the same manner as the welfare-consistent price index in the model (Bilbiie et al., 2014). We hence think of $A_t \widetilde{z}_t$ as being the accurate counterpart to empirical measures of TFP, where $\widetilde{z}_t$ captures the effect of endogenous exit and A_t all other potentially relevant channels.

Figure 3: Monetary policy shock in model variants



Note: Impulse response functions for a monetary policy shock for different model features. The solid line corresponds to the baseline model. The shock size is calibrated to yield a one-percent increase of output in the baseline model. Inflation, interest rate and exit are shown in percentage-point deviations from steady state, all other variables in percentage deviations.

reducing firm entry (as also shown by Bilbiie et al., 2007, Lewis, 2009).[38] In general equilibrium, the fewer surviving firms are on average more productive and profitable.

Additionally introducing the working capital channel (yellow dotted lines, $\tau > 0, \lambda_W = 0, \vartheta = 1$) does not alter these results substantially. As discussed above, the overall effect of the cost channel is ambiguous: While the borrowing rate decreases, real wages increase. Moreover, fixed costs of production constitute a relatively small part of overall costs (f is small). As a result, the counterfactual cyclicality of firm dynamics persists in this model variant. Similarly, the response of aggregate profits is still negative, which is at odds with the empirical results.

Finally adding wage rigidities yields the benchmark model, which successfully replicates the empirical evidence. As discussed in the previous section, the evolution of cor-

[38] The decrease of nominal interest rates implies a fall in the return to bonds. To restore no-arbitrage across different investments, the return to share holdings also decreases slightly. This happens via a slight increase in the equity prices today relative to tomorrow.

porate profits crucially affects firms' extensive margin. In the presence of wage stickiness, the rise of real wages following expansionary monetary policy shocks is inherently weaker. Consequently, the labor cost channel is substantially dampened (as well as the fixed cost channel, to a minor extent). As a result, the direct demand channel dominates following expansionary monetary policy shocks, which implies that firms become more profitable on impact. This reduces exit rates by allowing relatively less productive firms to remain in the market and incentivizes firm entry. We thus conclude that nominal rigidities in wage setting are a crucial ingredient for our model's ability to replicate the empirical findings. The inclusion of wage stickiness in the model is also consistent with our empirical results, which indicate that wages rise only sluggishly following expansionary monetary policy shocks.

4.4 Macroeconomic Dynamics and the Exit Channel

In this section, we analyze the implications of endogenous firm exit for the conduct of monetary policy. From a central bank perspective, the *exit channel* potentially affects the transmission of monetary policy. To isolate the role of the exit channel for the macroeconomic propagation of monetary policy shocks, we compare our benchmark model to an economy where all firms are homogeneous and exit occurs only exogenously. We set idiosyncratic firm productivity $z = 1$ for all firms and abstract from fixed costs of production by setting $f = 0$. As a result, no firm is ever forced to exit the market due to low productivity. This model variant is essentially the one considered by Bilbiie et al. (2007) and Bilbiie et al. (2014), but additionally includes entry frictions and wage rigidities – which, as shown above, are crucial to replicate the empirical results for firm exit and corporate profits.

Figure 4 compares the transmission of an expansionary monetary policy shock across the two models. The impulse responses for the benchmark model (solid black lines) are the same as above, with the shock size calibrated to yield a one-percent increase of output. We consider a shock of the same size in the model with exogenous exit (dashed blue lines).

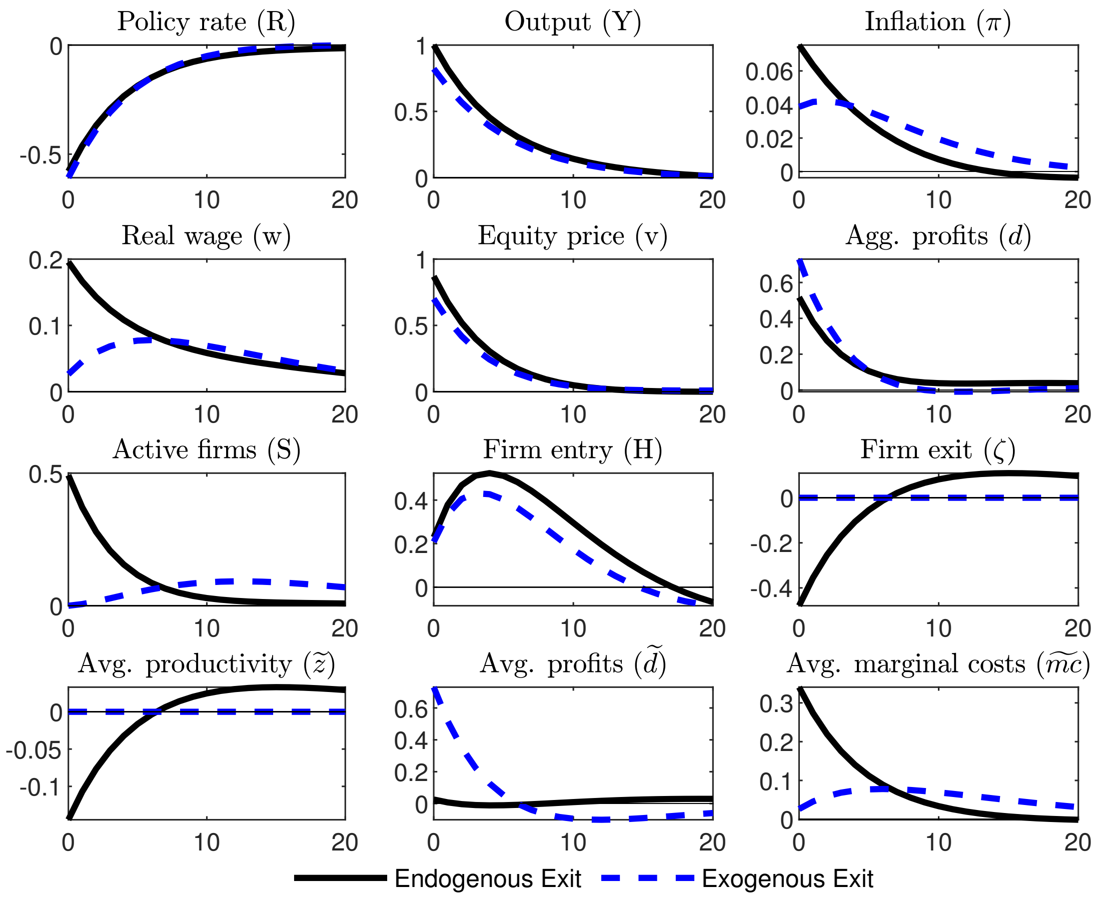
Compared to the model with exogenous exit, the output effect of the monetary policy shocks is slightly higher and more persistent in our benchmark model. Over the first 20 quarters, the exit channel amplifies the cumulative output response by around 19%. In contrast, the cumulative inflationary effect is weaker in the presence of the exit channel, by about 17% over the first 20 quarters. In the first year, however, the inflation response is stronger. Average firm profits increase only slightly, in contrast to the model with exogenous exit. This mirrors a sharp decline in average markups and higher average marginal costs. By construction, endogenous exit (ζ) and average firm-specific productivity ($\widetilde{z}$) do not respond in the model with exogenous exit.

The effects of the exit channel on inflation and markups are best understood by considering the New Keynesian Phillips curve of the model, which is Equation (35). Log-linearizing this equation around a steady state with zero (net) inflation yields

$$\widehat{\pi}_t = -\frac{\theta - 1}{\tau}\widehat{\widetilde{\mu}}_t + \beta(1-\delta)E_t[\widehat{\pi}_{t+1}] \tag{52}$$

where variables with a hat denote log-deviations from steady state. This is the familiar linearized New Keynesian Phillips curve, relating inflation to variations in the (average)

Figure 4: Monetary policy shock and the exit channel



Note: Impulse response functions for an expansionary monetary policy shock in the baseline economy (with endogenous exit, black solid line) and a variant with exogenous exit (blue dashed line). The shock size is calibrated to yield a one-percent increase of output in the baseline economy. Inflation, interest rate and exit are shown in percentage-point deviations from steady state, all other variables in percentage deviations.

firm markup. Using the optimal pricing condition in Equation (10) together with the definitions of marginal costs (5) and the variety effect (48), the markup is given by:

$$\widehat{\widetilde{\mu}}_t = \frac{1}{\theta - 1}\widehat{S}_t - (\widehat{w}_t - \widehat{A}_t - \widehat{\widetilde{z}}_t) \tag{53}$$

Note that in the absence of fluctuations in the number of active firms, this equation is the familiar negative relationship between markups and marginal costs (the expression in brackets) in the baseline New Keynesian model. Substituting (53) into (52) yields:

$$\widehat{\pi}_t = \frac{\theta - 1}{\tau}(\widehat{w}_t - \widehat{A}_t - \widehat{\widetilde{z}}_t) - \frac{1}{\tau}\widehat{S}_t + \beta(1-\delta)E_t[\widehat{\pi}_{t+1}] \tag{54}$$

Equation (54) is a New Keynesian Phillips Curve relating producer price inflation to marginal costs and the number of active firms in the economy. Intuitively, firms price setting crucially depends on their marginal costs. As such, changes in aggregate (A) or firm-specific ($\widetilde{z}$) productivity affect marginal costs and thus inflation. Furthermore, the number of active firms influences relative prices (the price of each good relative to the consumption basket) and thus markups, which translates into an effect on inflation. This may be interpreted as representing the effect of heightened competition. We can further transform this equation by inserting the log-linearized evolution of active firms from (38):

$$\widehat{\pi}_t = \frac{\theta - 1}{\tau} \underbrace{(\widehat{w}_t - \widehat{A}_t - \widehat{\widetilde{z}}_t)}_{\text{marginal costs}} - \frac{1}{\tau} \underbrace{\left(\widehat{N}_t - \frac{\zeta}{1-\zeta}\widehat{\zeta}_t\right)}_{\text{variety effect}} + \beta(1-\delta)E_t[\widehat{\pi}_{t+1}] \tag{55}$$

Equation (55) ties inflation dynamics to marginal costs, the overall number of firms and the endogenous exit rate. This formulation of the New Keynesian Phillips Curve illustrates how the exit channel alters the inflationary effect of monetary policy shocks.

As shown above, expansionary monetary policy shocks increase firm profits and allow unproductive firms to survive. As a consequence, average firm-specific productivity declines, and average marginal costs increase sharply on impact of the shock. The decline in average productivity therefore explains the initially stronger inflation response in the first year. At the same time, the overall number of firms increases and exit rates decline. Via the variety effect, this translates into lower markups and thus lower inflation. After the first year, the variety effect dominates the productivity effect such that the overall inflation response is lower. Interestingly, this shares similarities to the microeconometric findings by Acharya et al. (2020) who document that a rise in zombie credit is associated with disinflation. In this respect, the demand-side and preference-based variety effect in our framework may be interpreted as operating similar to a supply-side competition effect, whereby excess capacity creates downward pressure on prices.

The amplification of the output effect is largely due to higher investment in new firms in the model with endogenous exit. Intuitively, entering the market becomes profitable for firms with relatively low productivity. As a result, investment in new firms and firm entry respond stronger to monetary policy shocks. Over the medium-term, lower inflation and real interest rates also contribute to slightly higher consumption relative to the model with exogenous exit.

Overall, the effect of the exit channel resembles a flatter (medium-term) aggregate supply curve. A given shift of aggregate demand, for example in the form of expansionary monetary policy, translates into higher output and less inflation. The change in macroeconomic dynamics in the presence of the exit channel is, however and of course, not limited to monetary policy shocks. Section B.4 in the Appendix shows that macroeconomic dynamics following aggregate non-policy demand and supply shocks are likewise altered by the exit channel. Furthermore, Section B.5 demonstrates that a new policy trade-off between macroeconomic stabilization and corporate productivity might arise in the presence of the exit channel by comparing different Taylor rules. From a central bank perspective, the exit channel thus potentially calls for different (optimal) policy prescriptions. We leave the detailed investigation of the normative implications to future research, in particular the question whether central banks potentially face a trade-off between macroeconomic stabilization and the cleansing effect of recessions à la Caballero and Hammour (1994).

5 Conclusion

We document a new transmission channel of monetary policy, which we label the "exit channel": monetary conditions affect firms' profitability and thus both sides of firms' extensive margin. In particular, loose monetary conditions stimulate corporate profitability, thereby hampering the exit of relatively less productive firms.

Using U.S. data, we provide empirical evidence that expansionary monetary policy shocks lead to a decline in firm exit rates, whereas entry rates rise as corporate profits grow. As the monetary stimulus fades, firm exit rates overshoot the baseline in the medium run. We rationalize these findings in a structural model with heterogeneous firms and endogenous exit. In the model, expansionary monetary policy stimulates aggregate demand and thereby generates "zombification" of the economy, as firms with low productivity are able to continue operating. As monetary conditions revert to the steady state, these firms become unprofitable, causing exit to overshoot. We demonstrate that nominal rigidities in wage setting are a crucial model ingredient to replicate the empirical findings.

From a positive perspective, the exit channel is equivalent to a flatter aggregate supply curve. Therefore, endogenous firm exit amplifies the effect of monetary policy shocks on output, but reduces the response of inflation. One further avenue for future research is a detailed investigation of the normative implications, as already initiated in contemporaneous work by Hamano and Zanetti (2020). In particular, the adverse productivity effects caused by the exit channel are potentially highly relevant in gauging the optimal degree and time profile of monetary policy in recession times. For example, one might conjecture that the optimal monetary stimulus involves strong front-loading and a rapid normalization to avoid zombification and a prolonged recovery of the economy.

A second interesting issue for future work is the interaction between monetary conditions, firm exit and productivity in the long run. While our analysis is based on a business cycle perspective, the Japanese experience highlights that zombification may be associated with highly persistent economic slowdowns and potentially permanent GDP losses. This raises the question whether a prolonged low interest rate environment may induce the economy to slide into a new long-run equilibrium with lower steady state productivity, growth and inflation. Models with endogenous TFP growth along the lines of Ikeda and Kurozumi (2019) and Queralto (2020) constitute a promising framework in which to analyze this question.

References

Acharya, V. V., Crosignani, M., Eisert, T., and Eufinger, C. (2020). Zombie Credit and (Dis-)Inflation: Evidence from Europe. Mimeo.

Acharya, V. V., Eisert, T., Eufinger, C., and Hirsch, C. (2019). Whatever It Takes: The Real Effects of Unconventional Monetary Policy. *The Review of Financial Studies*, 32(9):3366–3411.

Adalet McGowan, M., Andrews, D., and Millot, V. (2018). The walking dead? Zombie firms and productivity performance in OECD countries. *Economic Policy*, 33(96):685–736.

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J., and Li, H. (2019). Missing Growth from Creative Destruction. *American Economic Review*, 109(8):2795–2822.

Andrews, D. and Petroulakis, F. (2019). Breaking the shackles: Zombie firms, weak banks and depressed restructuring in Europe. *ECB Working Paper*, No. 2240.

Antoni, M., Koetter, M., Müller, S., and Sondershaus, T. (2019). Do asset purchase programmes shape industry dynamics? Evidence from the ECB's SMP on plant entries and exits. *IWH Discussion Papers*, 12/2019.

Baqaee, D. R. and Farhi, E. (2020). Productivity and Misallocation in General Equilibrium. *The Quarterly Journal of Economics*, 135(1):105–163.

Barnichon, R. and Brownlees, C. (2019). Impulse Response Estimation by Smooth Local Projections. *The Review of Economics and Statistics*, 101(3):522–530.

Basu, S., Fernald, J. G., and Kimball, M. S. (2006). Are Technology Improvements Contractionary? *American Economic Review*, 96(5):1418–1448.

Bergin, P. R. and Corsetti, G. (2008). The extensive margin and monetary policy. *Journal of Monetary Economics*, 55(7):1222–1237.

Bergin, P. R., Feng, L., and Lin, C. (2018). Firm Entry and Financial Shocks. *The Economic Journal*, 128(609):510–540.

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

Bilbiie, F. O., Fujiwara, I., and Ghironi, F. (2014). Optimal monetary policy with endogenous entry and product variety. *Journal of Monetary Economics*, 64:1–20.

Bilbiie, F. O., Ghironi, F., and Melitz, M. J. (2007). Monetary Policy and Business Cycles with Endogenous Entry and Product Variety. *NBER Macroeconomics Annual*, 22:299–379.

Bilbiie, F. O., Ghironi, F., and Melitz, M. J. (2012). Endogenous Entry, Product Variety, and Business Cycles. *Journal of Political Economy*, 120(2):304–345.

Bittner, C., Fecht, F., and Georg, C.-P. (2020). Contagious Zombies. Mimeo.

Blattner, L., Farinha, L., and Rebelo, F. (2019). When losses turn into loans: the cost of undercapitalized banks. *ECB Working Paper*, No. 2228.

Caballero, R. J. and Hammour, M. L. (1994). The Cleansing Effect of Recessions. *The American Economic Review*, 84(5):1350–1368.

Caballero, R. J. and Hammour, M. L. (2005). The Cost of Recessions Revisited: A Reverse-Liquidationist View. *The Review of Economic Studies*, 72(2):313–341.

Caballero, R. J., Hoshi, T., and Kashyap, A. K. (2008). Zombie Lending and Depressed Restructuring in Japan. *American Economic Review*, 98(5):1943–77.

Cacciatore, M., Fiori, G., and Ghironi, F. (2016). Market deregulation and optimal monetary policy in a monetary union. *Journal of International Economics*, 99:120–137.

Caldara, D. and Herbst, E. (2019). Monetary Policy, Real Activity, and Credit Spreads: Evidence from Bayesian Proxy SVARs. *American Economic Journal: Macroeconomics*, 11(1):157–92.

Campbell, J. R. (1998). Entry, Exit, Embodied Technology, and Business Cycles. *Review of Economic Dynamics*, 1(2):371–408.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1):1–45.

Cieslak, A. and Schrimpf, A. (2019). Non-monetary news in central bank communication. *Journal of International Economics*, 118:293–315.

Clementi, G. L. and Palazzo, B. (2016). Entry, Exit, Firm Dynamics, and Aggregate Fluctuations. *American Economic Journal: Macroeconomics*, 8(3):1–41.

Colciago, A. and Silvestrini, R. (2020). Monetary policy, productivity, and market concentration. DNB Working Papers 685, Netherlands Central Bank, Research Department.

Del Negro, M., Giannoni, M. P., and Schorfheide, F. (2015). Inflation in the Great Recession and New Keynesian Models. *American Economic Journal: Macroeconomics*, 7(1):168–96.

Dixit, A. K. and Stiglitz, J. E. (1977). Monopolistic Competition and Optimum Product Diversity. *The American Economic Review*, 67(3):297–308.

Evans, C. L. (1992). Productivity shocks and real business cycles. *Journal of Monetary Economics*, 29(2):191–208.

Fernald, J. (2014a). A quarterly, utilization-adjusted series on total factor productivity. *Federal Reserve Bank of San Francisco Working Paper*, 2012-19.

Fernald, J. G. (2014b). Productivity and Potential Output before, during, and after the Great Recession. *NBER Macroeconomics Annual*, 29:1–51.

Garga, V. and Singh, S. R. (2019). Output Hysteresis and Optimal Monetary Policy. Working Paper, University of California, Davis, Department of Economics.

Gertler, M. and Karadi, P. (2015). Monetary Policy Surprises, Credit Costs, and Economic Activity. *American Economic Journal: Macroeconomics*, 7(1):44–76.

Ghironi, F. and Melitz, M. J. (2005). International Trade and Macroeconomic Dynamics with Heterogeneous Firms. *The Quarterly Journal of Economics*, 120(3):865–915.

Gilchrist, S. and Zakrajšek, E. (2012). Credit Spreads and Business Cycle Fluctuations. *The American Economic Review*, 102(4):1692–1720.

Gopinath, G., Kalemli-Özcan, S., Karabarbounis, L., and Villegas-Sanchez, C. (2017). Capital Allocation and Productivity in South Europe. *The Quarterly Journal of Economics*, 132(4):1915–1967.

Gürkaynak, R. S., Sack, B., and Swansonc, E. T. (2005). Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements. *International Journal of Central Banking*, 1:55–93.

Hamano, M. and Zanetti, F. (2017). Endogenous product turnover and macroeconomic dynamics. *Review of Economic Dynamics*, 26:263–279.

Hamano, M. and Zanetti, F. (2018). On Quality and Variety Bias in Aggregate Prices. *Journal of Money, Credit and Banking*, 50(6):1343–1363.

Hamano, M. and Zanetti, F. (2020). Monetary Policy, Firm Heterogeneity, and Product Variety. WINPEC Working Papers E2005, Waseda Institute of Political Economy.

Hamilton, J. D. (2018). Why You Should Never Use the Hodrick-Prescott Filter. *The Review of Economics and Statistics*, 100(5):831–843.

Hopenhayn, H. A. (1992). Entry, Exit, and Firm Dynamics in Long Run Equilibrium. *Econometrica*, 60(5):1127–1150.

Hoshi, T. and Kashyap, A. K. (2004). Japan's Financial Crisis and Economic Stagnation. *Journal of Economic Perspectives*, 18(1):3–26.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Ikeda, D. and Kurozumi, T. (2019). Slow Post-Financial Crisis Recovery and Monetary Policy. *American Economic Journal: Macroeconomics*, 11(4):82–112.

Jaimovich, N. and Floetotto, M. (2008). Firm dynamics, markup variations, and the business cycle. *Journal of Monetary Economics*, 55(7):1238–1252.

Jarociński, M. and Karadi, P. (2020). Deconstructing Monetary Policy Surprises - The Role of Information Shocks. *American Economic Journal: Macroeconomics*, 12(2):1–43.

Jordà, O. (2005). Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review*, 95(1):161–182.

King, R. G., Plosser, C. I., and Rebelo, S. T. (1988). Production, growth and business cycles: II. New directions. *Journal of Monetary Economics*, 21(2):309–341.

Kwon, H. U., Narita, F., and Narita, M. (2015). Resource reallocation and zombie lending in Japan in the 1990s. *Review of Economic Dynamics*, 18(4):709–732.

Lewis, V. (2009). Business Cycle Evidence on Firm Entry. *Macroeconomic Dynamics*, 13(5):605–624.

Lewis, V. (2013). Optimal Monetary Policy and Firm Entry. *Macroeconomic Dynamics*, 17(8):1687–1710.

Lewis, V. and Poilly, C. (2012). Firm entry, markups and the monetary transmission mechanism. *Journal of Monetary Economics*, 59(7):670–685.

Meier, M. and Reinelt, T. (2020). Monetary Policy, Markup Dispersion, and Aggregate TFP. Mimeo.

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6):1695–1725.

Melosi, L. (2017). Signalling effects of monetary policy. *The Review of Economic Studies*, 84(2):853–884.

Midrigan, V. and Xu, D. Y. (2014). Finance and misallocation: Evidence from plant-level data. *American Economic Review*, 104(2):422–58.

Miranda-Agrippino, S. and Ricco, G. (2018). The transmission of monetary policy shocks. Mimeo.

Moll, B. (2014). Productivity losses from financial frictions: Can self-financing undo capital misallocation? *American Economic Review*, 104(10):3186–3221.

Moran, P. and Queralto, A. (2018). Innovation, productivity, and monetary policy. *Journal of Monetary Economics*, 93:24–41.

Nakamura, E. and Steinsson, J. (2018). High-Frequency Identification of Monetary Non-Neutrality: The Information Effect. *The Quarterly Journal of Economics*, 133(3):1283–1330.

Oikawa, K. and Ueda, K. (2018). Reallocation Effects of Monetary Policy. *Research Institute of Economy, Trade and Industry (RIETI) Discussion Papers*, No. 18056. Discussion papers.

Ottonello, P. and Winberry, T. (2018). Financial heterogeneity and the investment channel of monetary policy. Technical report, National Bureau of Economic Research.

Peek, J. and Rosengren, E. S. (2005). Unnatural Selection: Perverse Incentives and the Misallocation of Credit in Japan. *American Economic Review*, 95(4):1144–1166.

Plagborg-Møller, M. and Wolf, C. K. (2019). Local projections and VARs estimate the same impulse responses. Mimeo.

Queralto, A. (2020). A model of slow recoveries from financial crises. *Journal of Monetary Economics*, 114:1 – 25.

Ravenna, F. and Walsh, C. E. (2006). Optimal monetary policy with the cost channel. *Journal of Monetary Economics*, 53(2):199–216.

Restuccia, D. and Rogerson, R. (2008). Policy Distortions and Aggregate Productivity with Heterogeneous Plants. *Review of Economic Dynamics*, 11(4):707–720.

Romer, C. D. and Romer, D. H. (2000). Federal Reserve Information and the Behavior of Interest Rates. *American Economic Review*, 90(3):429–457.

Rossi, L. (2019). The overshooting of firms' destruction, banks and productivity shocks. *European Economic Review*, 113:136–155.

Rotemberg, J. J. (1982). Monopolistic Price Adjustment and Aggregate Output. *The Review of Economic Studies*, 49(4):517–531.

Schivardi, F., Sette, E., and Tabellini, G. (2017). Credit misallocation during the European financial crisis. *Bank of Italy Temi di Discussione (Working Paper)*, No. 1139.

Storz, M., Koetter, M., Setzer, R., and Westphal, A. (2017). Do we want these two to tango? On zombie firms and stressed banks in Europe. *ECB Working Paper*. No. 2104.

Tracey, B. (2019). The real effects of zombie lending in Europe. *Bank of England Working Paper*, No. 783.

Wu, J. C. and Xia, F. D. (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Journal of Money, Credit and Banking*, 48(2-3):253–291.

Appendix

A Empirical Analysis

A.1 Data

Figure A1 shows time-series plots of the data. Firm entry and exit display strong cyclical patterns around US recessions: firm entry is procyclical, while firm exit is countercyclical. Moreover, profits show some pro-cyclical pattern. Aggregate TFP displays some mild signs of procyclicality, while utilization-adjusted TFP and labor productivity evolve rather independently of the cycle. All productivity series fall during the global financial crises. To complement the visual inspection, we present descriptive statistics and investigate average conditional patterns around turning points by means of Burns-Mitchell diagrams.

Table A1 presents descriptive statistics on business cycle fluctuations as measured by the cyclical component of all variables in log-levels using the regression-based filter of Hamilton (2018). Panel (a) reports volatility, relative volatility to the cyclical component of real GDP, persistence and contemporaneous co-movement. Firm entry and exit dynamics are almost three times more volatile than fluctuations in output and profits are about 7 times more volatile. In contrast, aggregate productivity measures are less volatile than output. Firm entry and exit dynamics are the least persistent series. Profits and productivity measures are slightly less persistent than output. Firm entry is procyclical, while exit is countercyclical, in line with the evidence by Campbell (1998) and Jaimovich and Floetotto (2008), who consider a different data set and study an earlier period. Interestingly, aggregate TFP is strongly procyclical while profits, utilization-adjusted TFP and labor productivity hardly co-move with the cycle. In fact, the estimated correlation are insignificant and thus, these series may be considered as a-cyclical.

Panel (b) provides more details on contemporaneous correlations between real GDP, firm dynamics, profits and productivity. While there is no co-movement between firm entry and exit, firm dynamics are strongly correlated with aggregate TFP. A cyclical upswings of firm entry is associated with higher aggregate productivity, while firm exit and productivity tend to move in opposite directions, in line with VAR-based estimates of Rossi (2019). Moreover, profits negatively comove with firm exit and positively comove with productivity measures, while they are unrelated to firm entry.

Firm dynamics are contemporaneously hardly related to pure technological progress. Because of the growth accounting definitions in Basu et al. (2006), this implies that the co-variation of firm dynamics with aggregate TFP is driven by variable capital and labor utilization. Moreover, both firm entry and exit are countercyclical to labor productivity. Roughly speaking, labor productivity rises if workers have more capital, have better skills or if aggregate TFP rises (Fernald, 2014b). Thus, the negative correlation of firm entry suggests that variations in capital and labor quality dominate the positive effects of aggregate TFP. For firm exit, these effects only slightly affect the negative correlation.

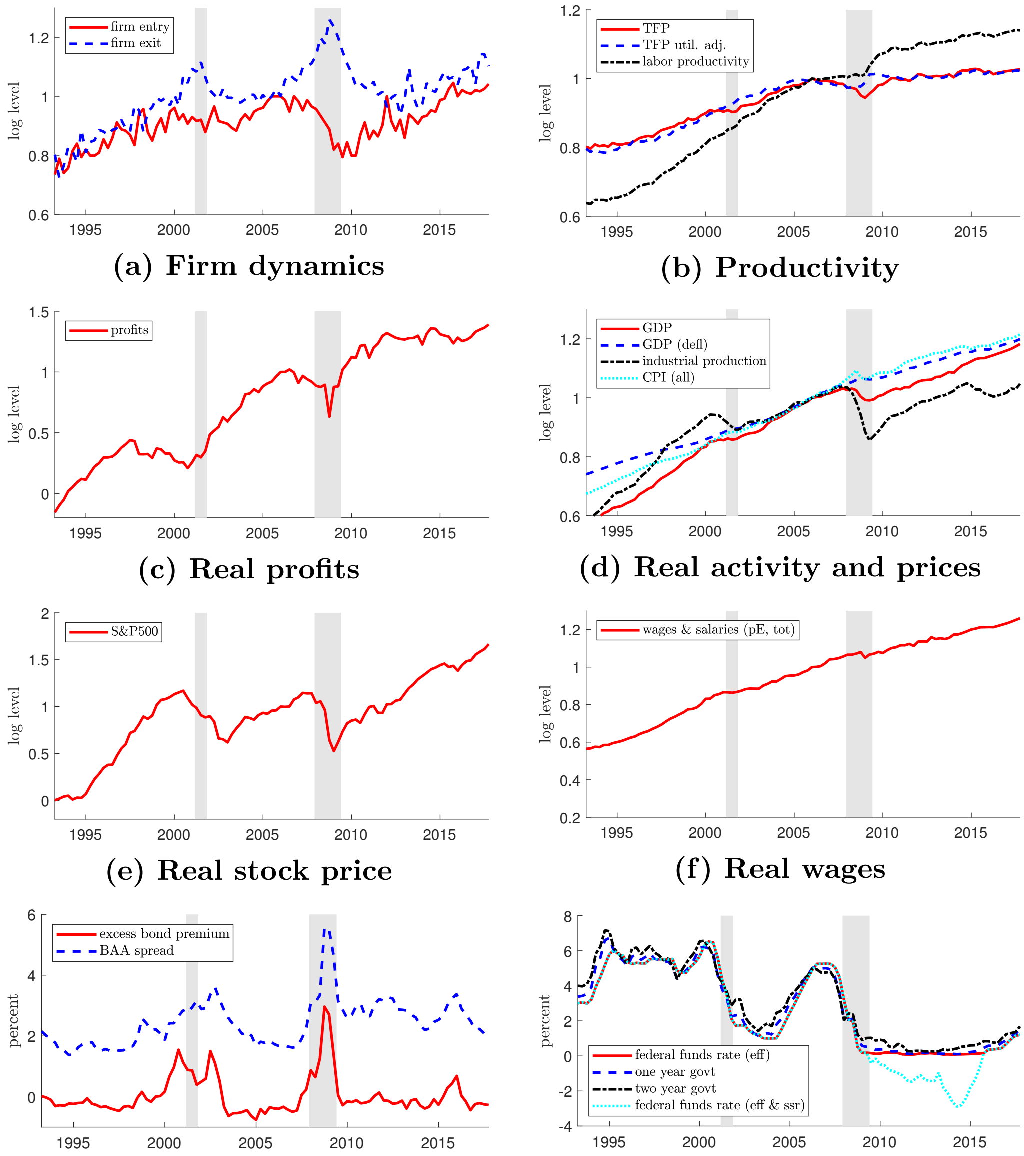
Turning from unconditional to conditional co-movements, Figure A2 shows Burns-Mitchell diagrams of selected variables. These diagrams depict the average behavior of economic time series around the start of US recessions. Chart (a) shows the average behavior of firm dynamics. Firm entry remains high during the expansion, but drops substantially after the turning point of the cycle. Firm exit, on the other hand, starts to increase prior to the start of the recession and peaks after four quarters. During the recovery, firm exit starts to diminish while firm entry remains subdued for a prolonged period.

Chart (b) shows the average dynamics of the productivity measures. Aggregate TFP is procyclical and leading, peaking several quarters before the turning point, while utilization-adjusted TFP and labor productivity show no strong cyclical patterns. Their average growth is uninterrupted during recessions.

Chart (c) show the average dynamics of corporate profits. Profits are a-cyclical to the real activity but start to decline prior to a recession. After the economy reached its through after 4 to 6 quarters and the economy starts to recover, profits increase strongly.

The remaining charts present the average behavior of key macoreconomic and financial variables. Real activity and prices, Chart (d), move as expected. Real activity contracts while prices react rather sluggish. The sluggish behavior of prices is a common feature of more recent recessions, see Figure A1. Equity peaks prior to the turning point in real GDP and substantially declines in the down turn, Chart (e). Financial frictions, Chart (g), are low prior to a recession but increase substantially when the economy dips further into the recession. Chart (h) shows that the policy rate and longer-term interest rates decline in response to subdued economic activity.

Figure A1: Data



(a) Firm dynamics
(b) Productivity
(c) Real profits
(d) Real activity and prices
(e) Real stock price
(f) Real wages
(g) Financial frictions
(h) Interest rates

Note: Time-series plot of the data. All variables are in log-levels, except the proxies of financial frictions and the interest rates. Measures of financial frictions and interests rates are in percent. Shaded gray areas indicate NBER recession dates.

Table A1: Business Cycle Statistics for the US Economy

	Standard deviation	Relative standard deviation	First-order autocorrelation	Contemporaneous correlation with output
Y	2.29	1.00	0.89	1.00
π	1.02	0.45	0.86	0.30
Entry	5.54	2.42	0.75	0.70
Exit	6.71	2.93	0.79	-0.24
Profits	15.32	6.69	0.86	0.00
TFP	1.82	0.79	0.87	0.64
TFPu	1.60	0.70	0.83	-0.13
LP	1.81	0.79	0.82	-0.09
Wages	1.77	0.77	0.79	0.62
Equity	21.57	9.43	0.88	0.81
Interest	2.18	0.95	0.97	0.28
EBP	0.67	0.29	0.83	-0.48

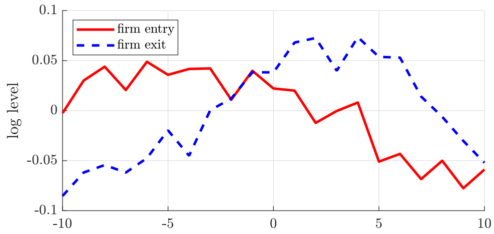
(a) Properties

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Y	(1)	1.00											
π	(2)	0.30	1.00										
Entry	(3)	0.70	0.35	1.00									
Exit	(4)	-0.24	0.10	-0.01	1.00								
Profits	(5)	0.00	0.08	0.03	-0.60	1.00							
TFP	(6)	0.64	0.22	0.36	-0.67	0.58	1.00						
TFPu	(7)	-0.13	-0.17	-0.11	-0.20	0.26	0.36	1.00					
LP	(8)	-0.09	-0.21	-0.29	-0.47	0.51	0.59	0.78	1.00				
Wages	(9)	0.62	0.34	0.46	0.25	-0.57	0.14	-0.29	-0.29	1.00			
Equity	(10)	0.81	0.10	0.55	-0.14	-0.05	0.45	-0.26	-0.15	0.57	1.00		
Interest	(11)	0.28	0.20	0.09	0.04	-0.40	0.01	-0.14	-0.11	0.49	0.06	1.00	
EBP	(12)	-0.48	-0.14	-0.28	0.55	-0.49	-0.63	0.01	-0.18	-0.03	-0.54	-0.03	1.00

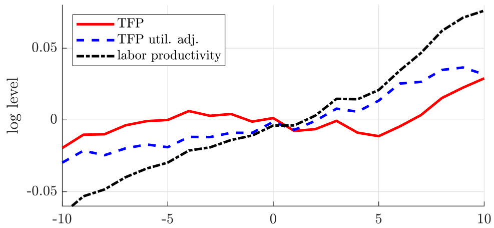
(b) Contemporaneous Correlations

Note: All variables are in log-levels and the cyclical component is estimated using the regression-based filter of Hamilton (2018), except for the interest rate and the excess bond premium. Y is real GDP, π is GDP deflator, Entry is establishment birth, Exit is establishment death, Profits are corporate profits with IVA and CCAdj., TFP is total factor productivity, TFPu is utilization adjusted TFP, LP is labor productivity, Wages are real wages & salaries per employment (total) deflated by the GDP deflator, Equity is the S&P500 stock index deflated by the GDP deflator, Interest is the one-year government bond and EBP is the excess bond premium.

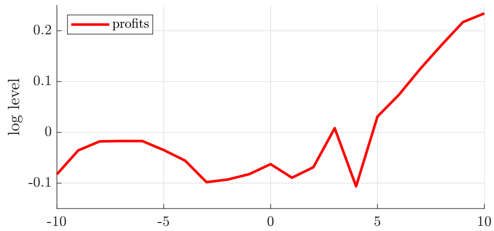
Figure A2: Burns-Mitchell Diagrams



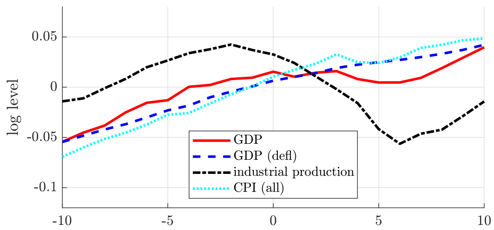
(a) Firm dynamics



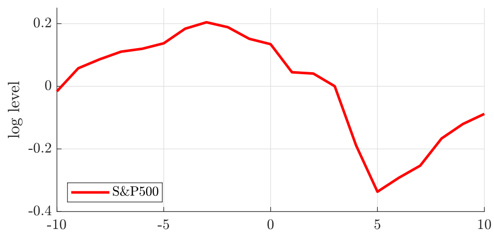
(b) Aggregate productivity



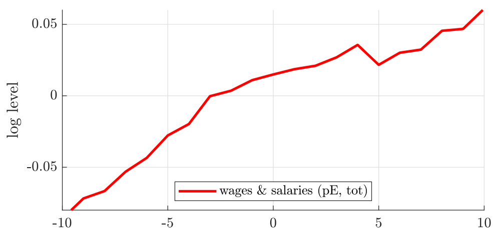
(c) Real profits



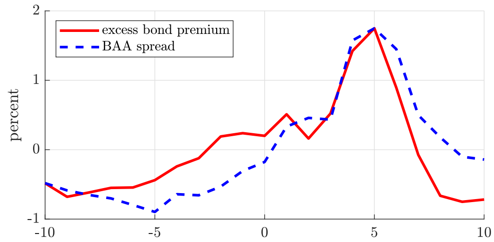
(d) Real activity and prices



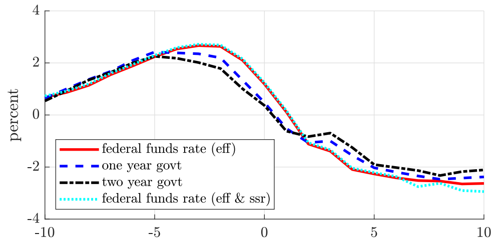
(e) Real stock price



(f) Real wages



(g) Financial frictions



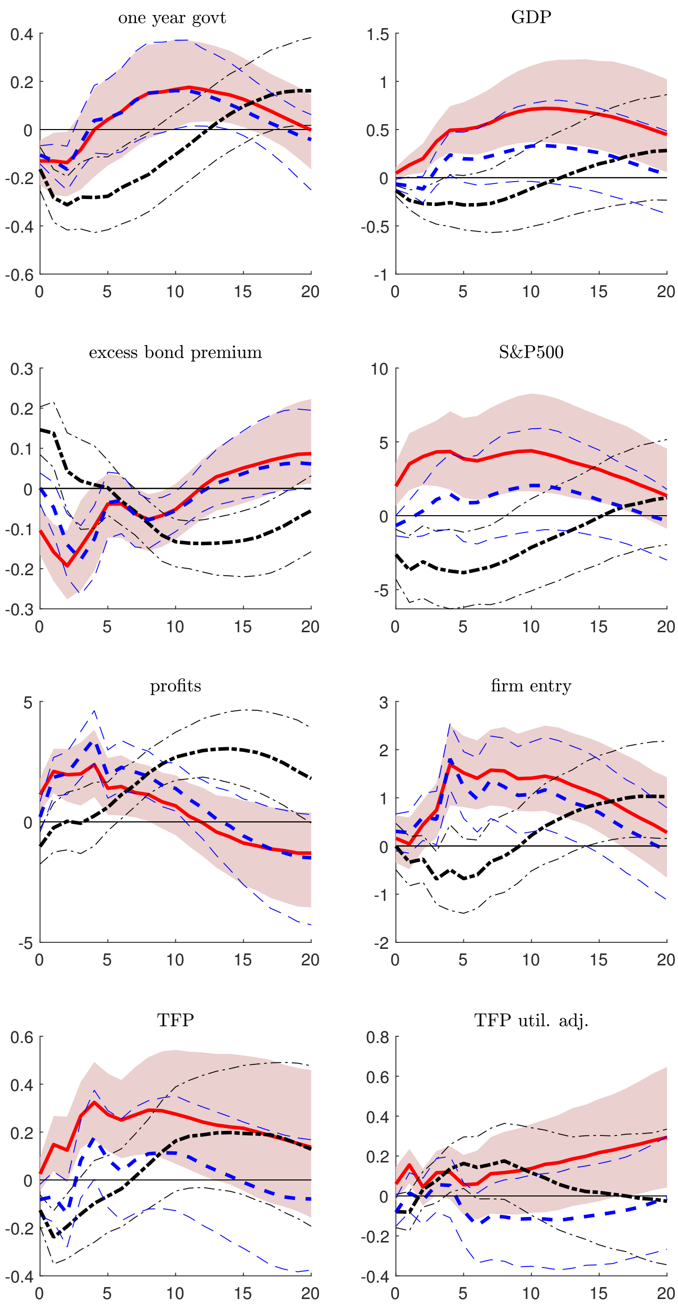
(h) Interest rates

Note: Average behavior of variables around cyclical peaks, as measured by the start of a US recession. $x_t = \frac{1}{M}\sum_{i=1}^{M}\left(y_{i,t} - \frac{1}{21}\sum_{t=-10}^{10} y_{i,t}\right)$ where $y_{i,-10}, y_{i,-9}, ..., y_{i,0}, y_{i,1}, ..., y_{i,10}$, $i = 1, 2, ..., M$ and $y_{i,0}$ is quarter of business cycle peak. All variables enter in log-levels, except interest rates, the excess bond premium and corporate BAA spread which are in percent.

A.2 Importance of Central Bank Information Effect

Figure A3 shows impulse response functions of a monetary policy shock identified by short-run zero-restriction on the interest rate surprises ordered first in the VAR (blue dashed) and those of a central bank information shock (black dashed-dotted).

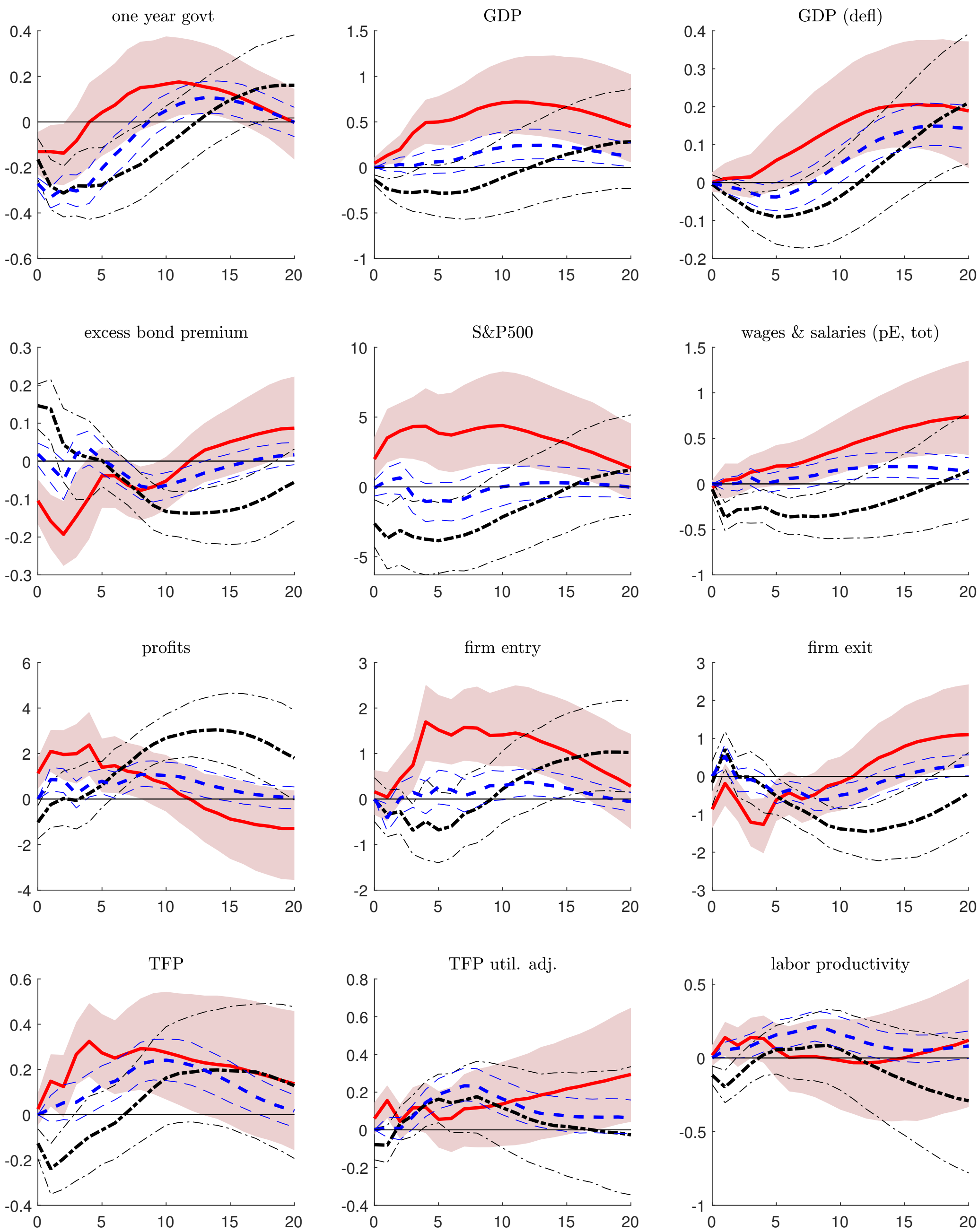
Figure A3: VAR short-run restrictions on interest rate surprises



Note: Impulse response functions to a monetary policy shock of the baseline (red) and identified by zero restriction on the interest rate surprises ordered first (blue) and a central bank information shock identified by sign restrictions (black). The thick lines are the median estimate and the area plots as well as the dashed-lines mark the 68% credible intervals.

Figure A4 shows impulse response functions of a monetary policy shock identified by short-run zero-restriction on real GDP and the prices (blue dashed) and shows those of a central bank information shock (black dashed-dotted).

Figure A4: VAR short-run restrictions on interest rate



Note: Impulse response functions to a monetary policy shock of the baseline (red) and identified by zero restriction on contemporaneous co-movement of GDP and prices for the interest rate (blue) and a central bank information shock identified by sign restrictions (black). The thick lines are the median estimate and the area plots as well as the dashed-lines mark the 68% credible intervals.

A.3 Additional Information on Identification

In this section, we investigate how the estimated impulse response functions are affected by the different data frequency, by the different sample size and by the additional sign restrictions on the low-frequency variables (coupled with the enlargement of the rotation space) as compared to Jarociński and Karadi (2020). Figure A5 shows estimates based on monthly data in Panel (a) and based on quarterly data in Panel (b).[39] Each panel shows the estimates for our considered sample starting in 1993:M4 and the slightly longer sample starting in 1990:M2, as well as for the different identification schemes. Red and black correspond to sign restriction on both high-frequency and low-frequency variables, while blue and cyan correspond to sign restrictions on high-frequency variables only, as in Jarociński and Karadi (2020).

Panel (a) shows that the estimated impulse response functions at the monthly frequency are affected by the different sample size and by the alternative identification restrictions. The monetary impulse identified by sign restrictions on high-frequency variables only and for the sample starting in 1993:M4 (blue dashed) differs substantially from the other median estimates. Particularly, the initial impulse to the interest rate is rather small and disproportionally compensated in the medium term. Also notice that the response of the stock market is also insignificant under this specification.

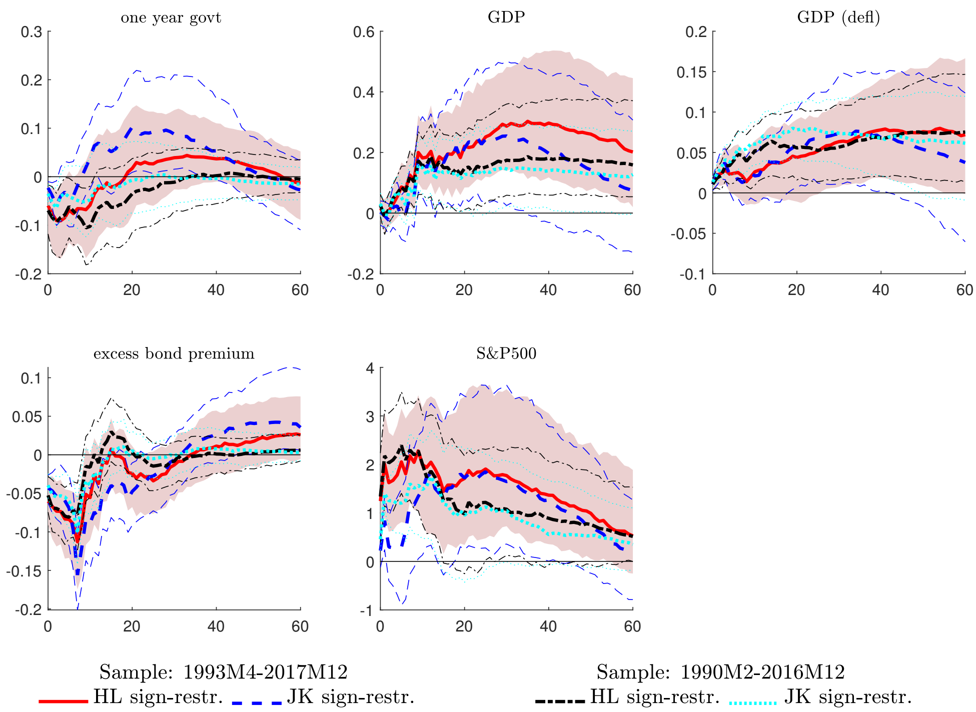
In contrast, the median estimate of the interest rate response based on the longer sample starting in 1990:M2 (cyan dotted) is more comparable to the median estimates obtained under our identification restrictions, which are qualitative similar in both samples. Note that for the longer sample the response of the interest rate and the stock price index are significant and last over several months when using sign restrictions only on high-frequency variables. Apart from that, it should be noted that both identification schemes yield qualitatively similar estimates of the responses of macroeconomic and financial variables to a monetary policy shock.

Turning to Panel (b), the chart shows that median estimates of the interest rate response also differs across different sample size and identification restrictions at the quarterly frequency. However, estimated impulse response functions for a specific sample size and identification scheme are very similar at different data frequencies. Particularly, the quarterly estimates can be interpreted as smoothed version of the monthly estimates. Nevertheless, it should be noted that the stock market response is only very marginally significant when only sign restriction on high frequency variables are imposed.

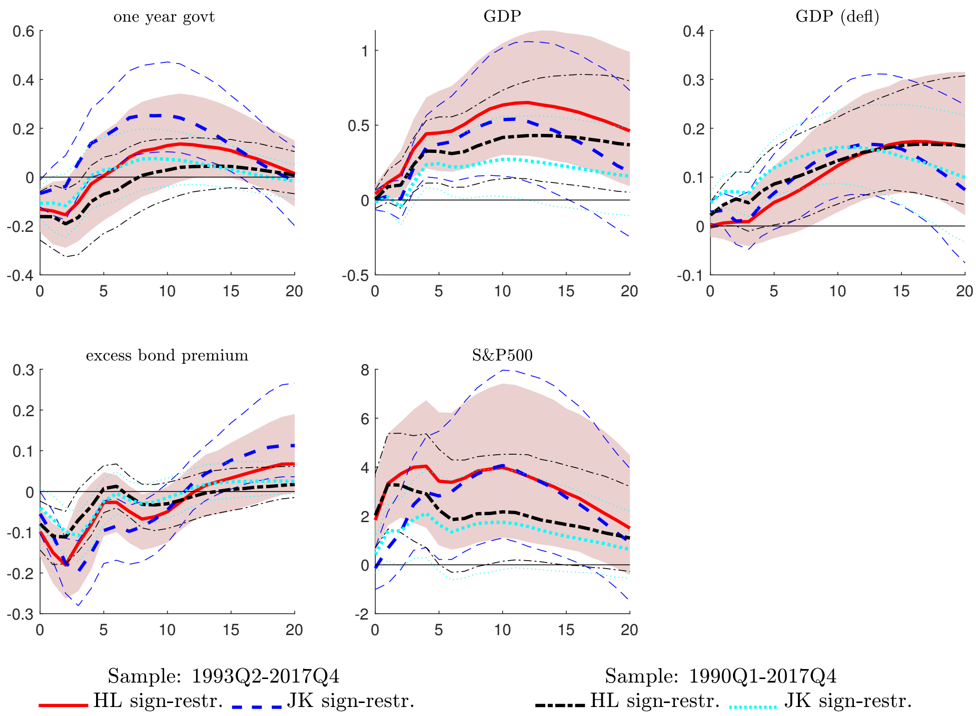
Based on these consideration, we conclude that excluding the sample from 1990:M2 to 1993:M3 from the estimation may obscure the relationship between high-frequency and low-frequency variables. The lack of this data makes it more difficult to identify a plausible monetary transmission channel when structural parameters are identified using sign restriction on high-frequency variables only. By imposing additional restrictions on the low-frequency variables (coupled with the enlarged rotation space), we are able to identify a plausible monetary transmission channel that shares most of the qualitative pattern of the estimated impulse response at monthly frequency and using the longer sample starting in 1990:M2.

[39] We exclude our main variables of interest due to the slightly longer sample.

Figure A5: Monetary transmission at different frequency and sample size



(a) Monthly



(b) Quarterly

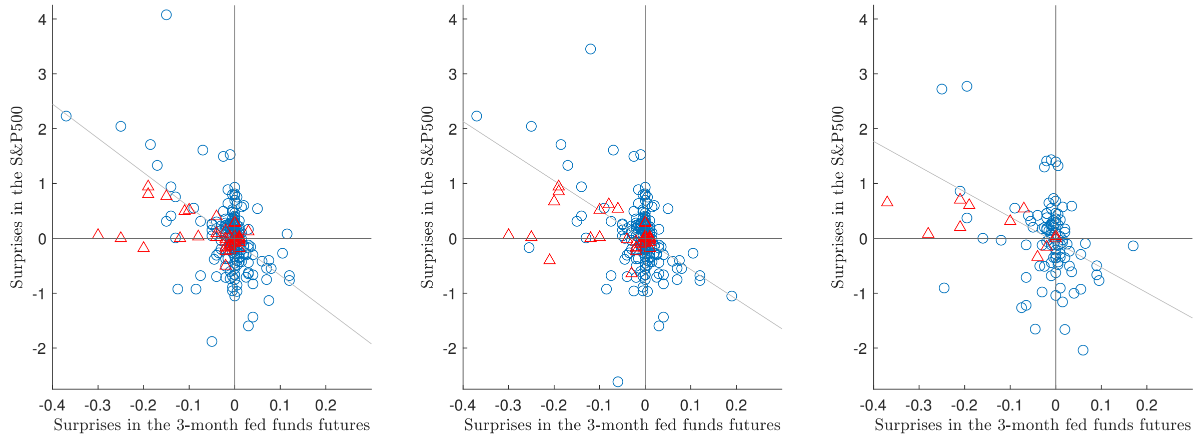
Note: Impulse response functions to a monetary policy shock identified by alternative sign restrictions, for different sample sizes and different data frequency. HL corresponde to the baseline sign restriction and JK correspond to the sign restrictions Jarociński and Karadi (2020). The thick lines are median estimates and the red area as well as the thin lines depict the 68% credible intervals

To further investigate the effects of a different sample size on the relationship between high-frequency surprises and their low-frequency counterparts, Figure A6 depicts scatter plots of interest rate and stock price surprises across (a) intra-daily frequency, (b) monthly frequency, and (c) quarterly frequency for the sample 1993:Q2 – 2017:Q4 in blue dots and the pre-sample 1990:Q1 – 1993:Q1 in red triangles. Two notable features stand out.

First, the pre-sample period from 1990:Q1 until 1993:Q1 features relatively large negative interest rate surprises as well as positive stock market surprises. Thus, the pre-sample is dominated by surprises that classify as a monetary policy shock according to the co-movement restrictions. Note the Fed lowered the interest rate during several intermeeting moves to cushion the effects of the savings and loan crisis on the U.S. economy during this time. Therefore, the absence of this relatively important episode may be the reason why the sign-restriction approach, on high-frequency variables only, lacks the power to identify a reasonably sized monetary impulse for the sample starting in 1993:Q2.

Second, there are fewer large interest rate and stock price surprises at the quarterly frequency as compared to the monthly and the intra-daily frequency. The similarity between monthly and intra-daily frequency can be rationalized by the fact that there is rarely more than one FOMC announcement per month.[40] However, there are several surprises within a given quarter that might potentially offset each other, thus, leading to smaller surprises in the aggregate. This loss in variability might make it more difficult to identify a relationship between high-frequency and low-frequency variables.

Figure A6: Interest Rate and Stock Price Surprises



(a) Intra-daily **(b) Monthly** **(c) Quarterly**

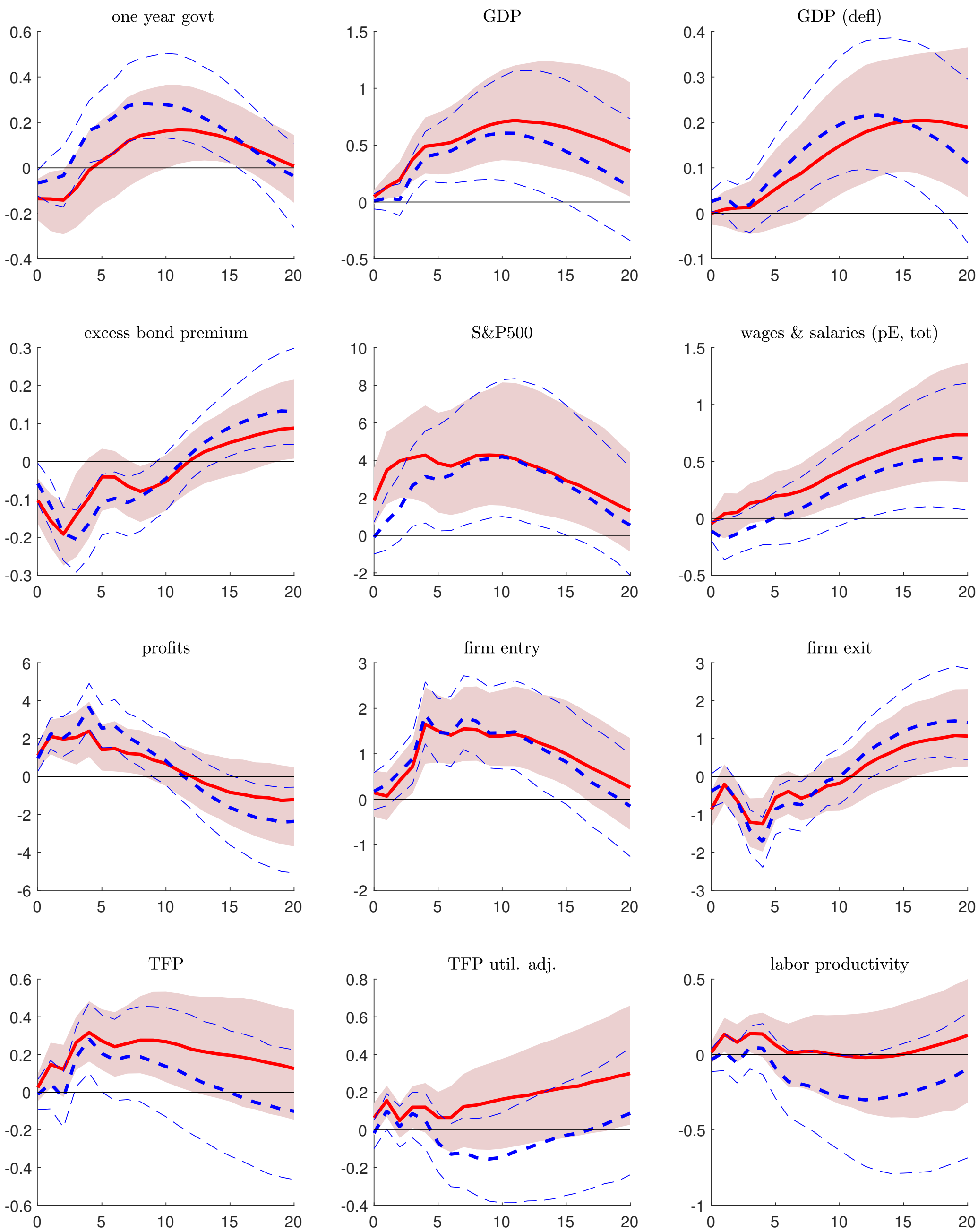
Note: Changes in the 3-month federal funds futures and the S&P500 stock index around FOMC announcements, in percent. For plot(a), each dot represents one FOMC announcement. For plot (b) and (c), each dot represents the sum of intra-daily surprises of FOMC announcements in the current month and quarter, respectively. The grey line is the fitted least squares prediction. Red triangles correspond to the period 1990M2 - 1993M3 and blue dots correspond to 1993M4 until 2017M12.

[40]Since 1994, most FOMC announcement are regularly scheduled meetings and take place at a monthly or six-weeks frequency. The remaining FOMC announcements are unscheduled meetings and conference calls, which are however rare in the sample we consider.

A.4 Robustness

Alternative specifications: Figure A7 shows impulse response functions of a monetary policy shock identified by using sign restrictions of Jarociński and Karadi (2020), that is, only on the co-movement of high-frequency variables (blue dashed).

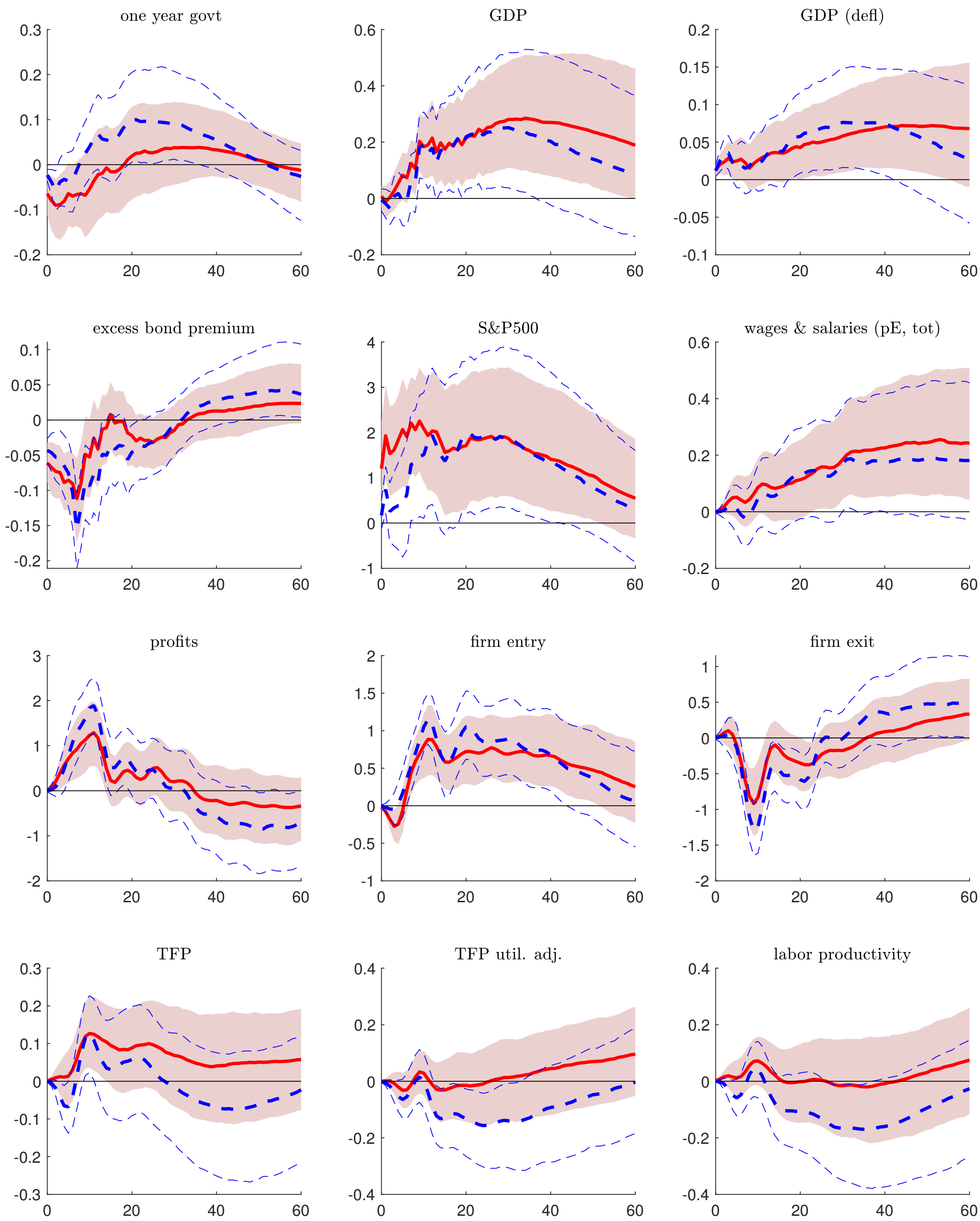
Figure A7: VAR with sign restrictions on high-frequency only



Note: Impulse response functions to a monetary policy shock of the baseline (red) and identified by using sign restrictions on high-frequency variables only (blue dashed). The thick lines are median estimate and the area charts as well as dashed lines are 68% credible intervals.

Figure A8 shows impulse response functions to a monetary policy shock identified by using the baseline sign restrictions (red) and the sign restrictions on high-frequency variables only (blue dashed) for the data set with monthly interpolated data.

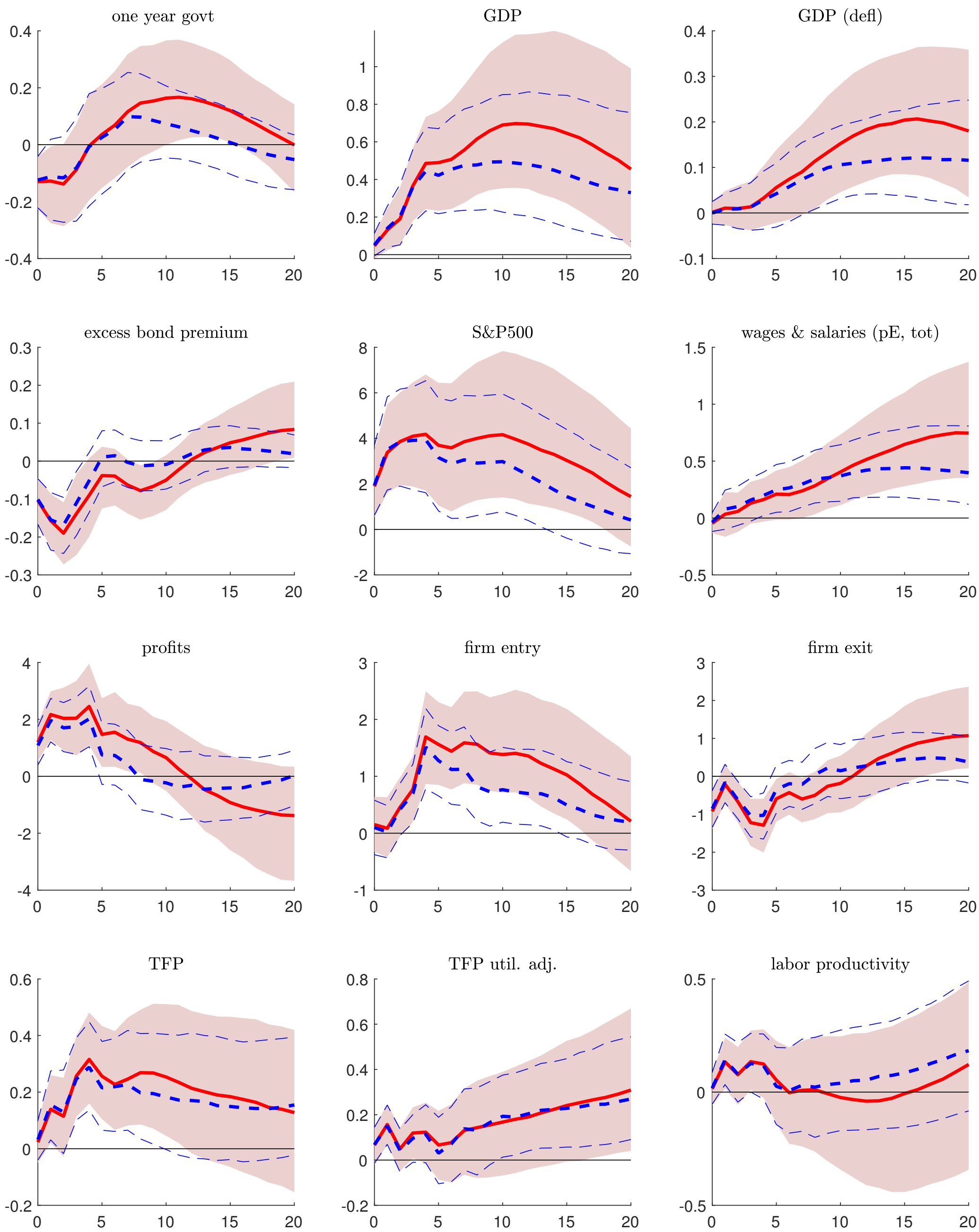
Figure A8: VAR with monthly interpolated time series



Note: Impulse response functions to a monetary policy shock of the baseline (red) and identified by using sign restriction on high-frequency variables only (blue dashed) for the data set at monthly frequency. The thick lines are median estimate and the area charts as well as dashed lines are 68% credible intervals.

Figure A9 shows impulse response functions of a monetary policy shocks identified with sign-restriction on high-frequency and low-frequency variables from an unrestricted VAR (blue dashed).

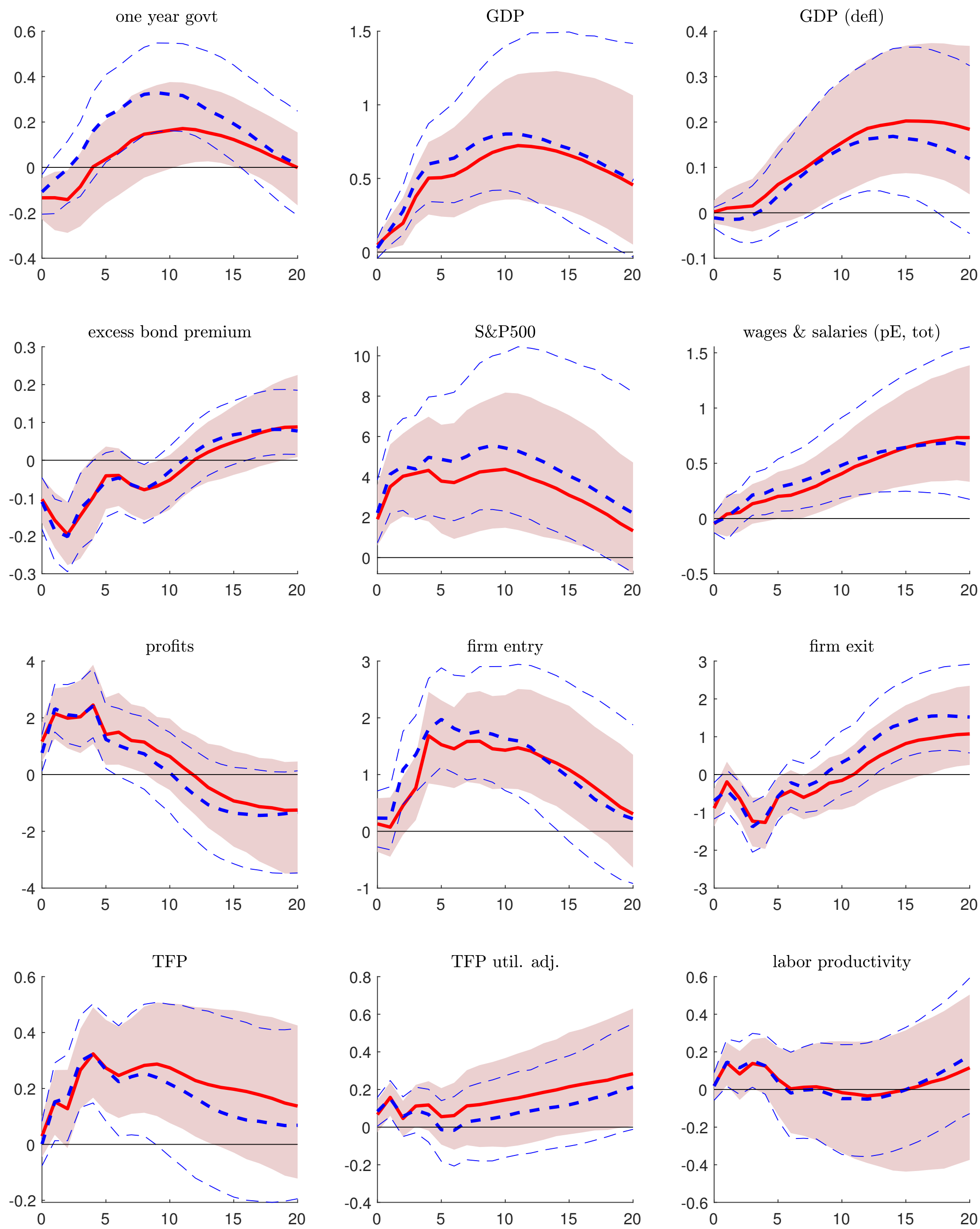
Figure A9: Unrestricted VAR



Note: Impulse response functions to a monetary policy shock of the baseline (red) and from an unrestricted VAR (blue). The thick lines are median estimate and the solid lines are 68% credible intervals.

Figure A10 shows impulse response functions of a monetary policy shocks using only high-frequency surprises from scheduled FOMC announcements (blue dashed).

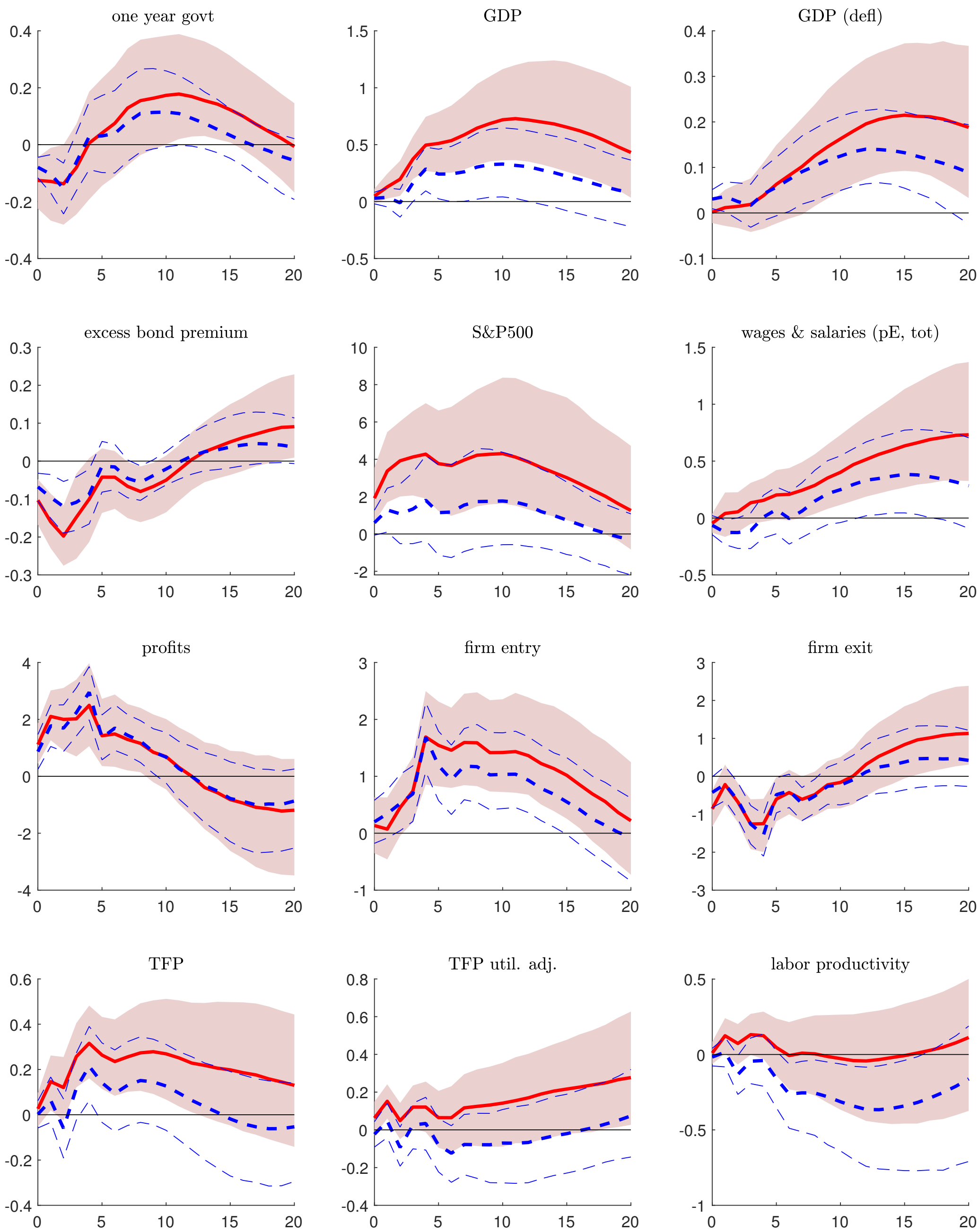
Figure A10: Surprises from scheduled FOMC meetings only



Note: Impulse response functions to a monetary policy shock of the baseline (red) and by using only surprises from scheduled FOMC announcements (blue dashed). The thick lines are median estimate and the solid lines are 68% credible intervals.

Alternative identification procedures: Figure A11 shows impulse response functions to a monetary policy shock identified in VAR with zero restrictions and the poor man's proxy of a monetary policy shock along the lines Jarociński and Karadi (2020).

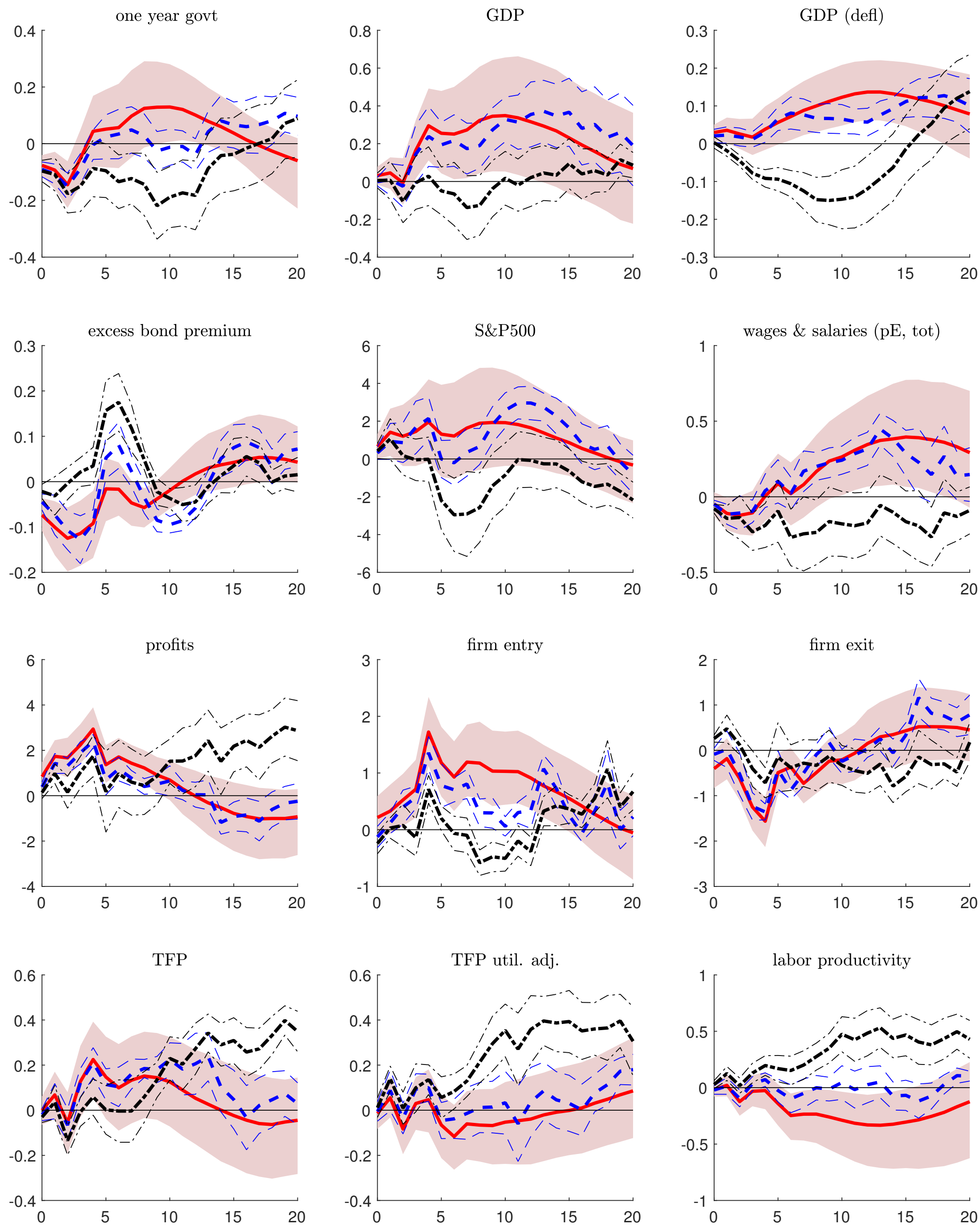
Figure A11: VAR with poor man's proxy



Note: Impulse response functions to a monetary policy shock of the baseline (red) and identified using the poor man's proxy of a monetary policy shock in a VAR with zero-restrictions (blue dashed). The thick lines are median estimate and the thin lines are 68% credible intervals.

Figure A12 shows impulse response functions to a monetary policy shock using the poor man's proxy series in a VAR with zero restrictions (red), in a local projection with a set of macroeconomic and financial controls (blue dashed), and in a local projection without additional controls (black dashed-dotted).

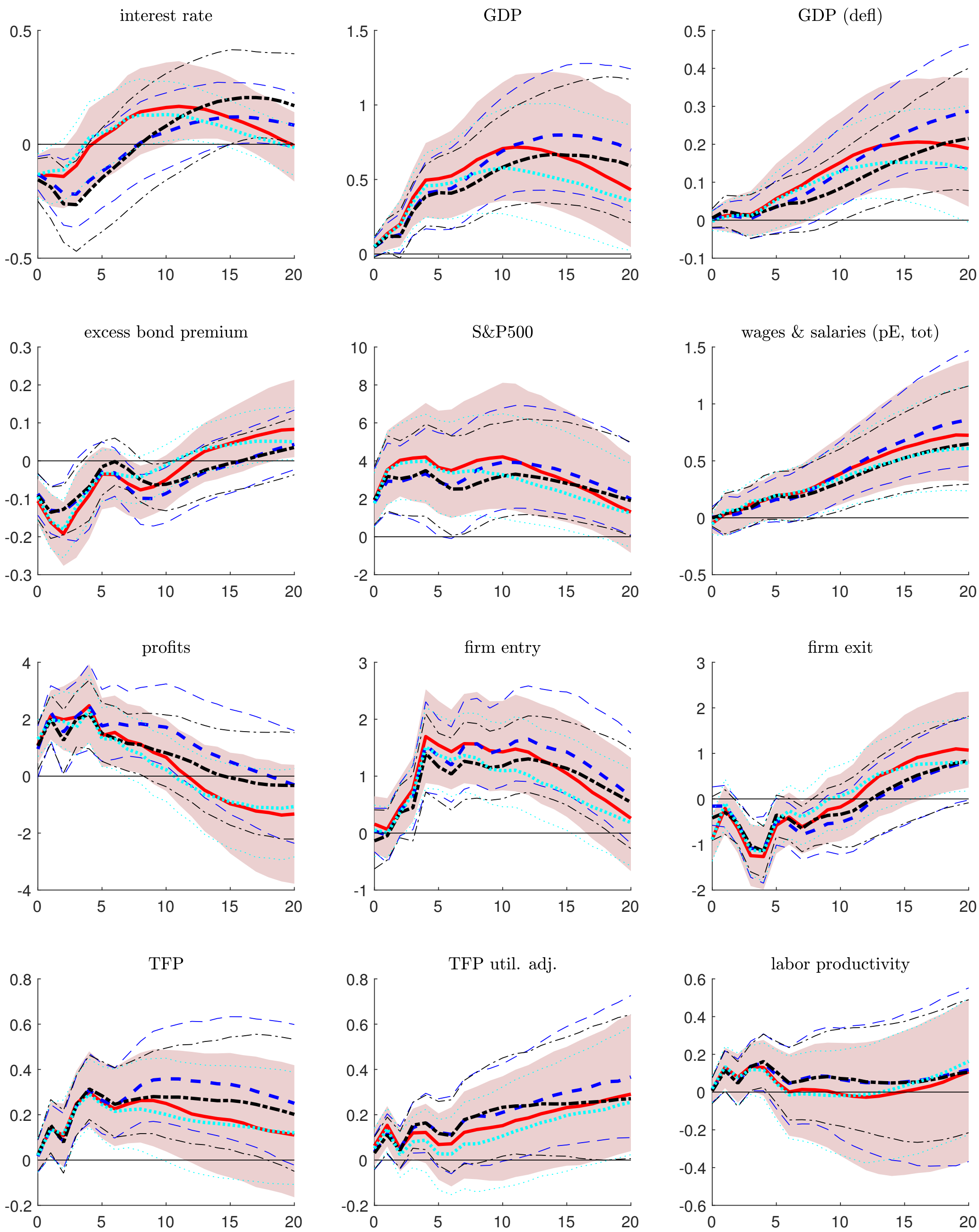
Figure A12: Poor man's proxy: VAR and local projections



Note: Impulse response functions to a monetary policy shock identified using the poor man's proxy in a VAR with zero-restrictions (red), in a local projection with a set of macroeconomic and financial controls (blue dashed) and in a local projection without additional controls (black dashed-dotted). The thick line is the median (point) estimate and the area charts (thin lines) is the 68% credible interval (confidence interval) of the VAR (local projection).

Alternative variables: Figure A13 shows impulse response functions to a monetary policy shock using alternative measures of monetary policy: federal funds rate (blue dashed line), federal funds rate extended by the shadow short rate by Wu and Xia (2016) (black dashed-dotted line) and two year government bond rate (cyan dotted line).

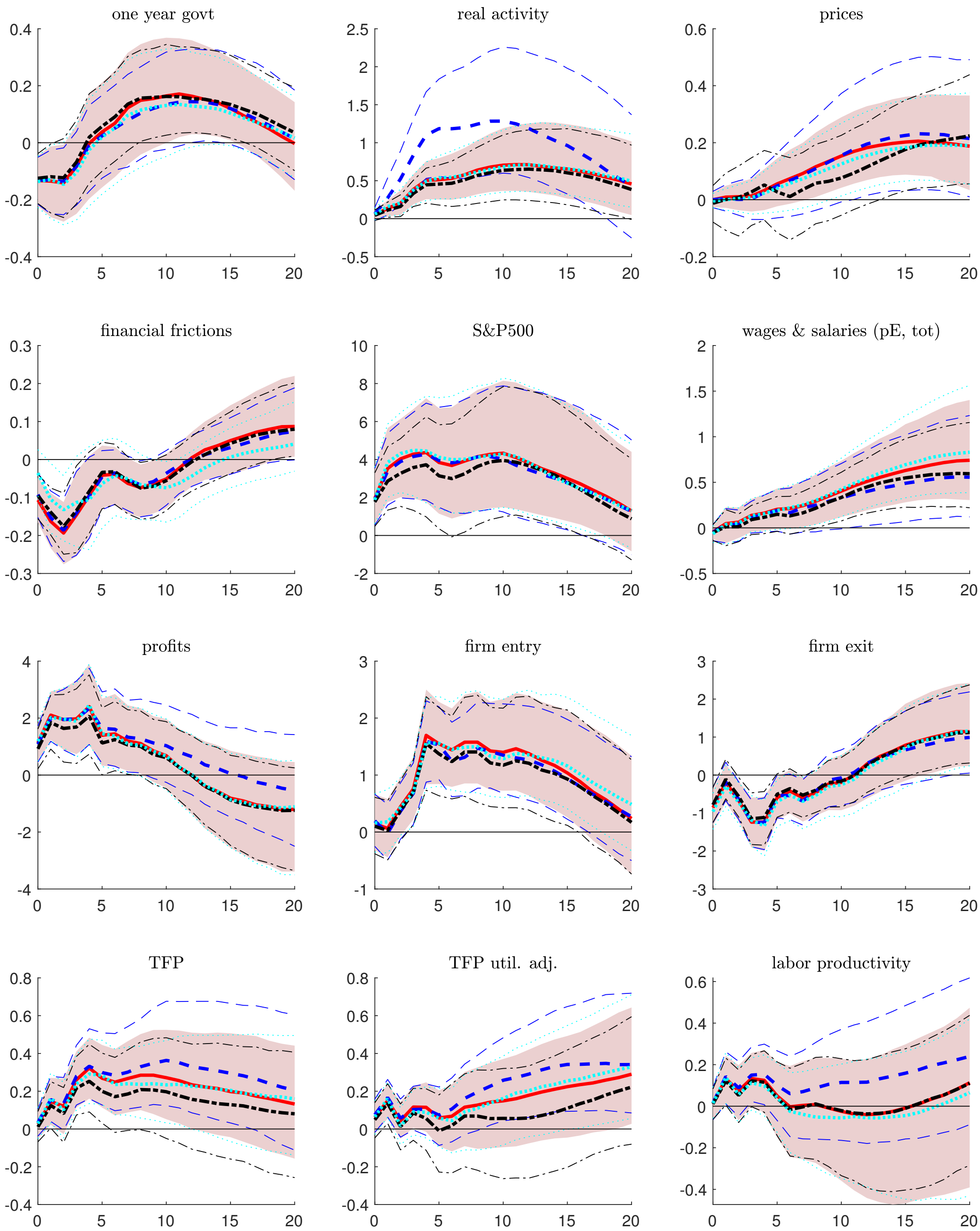
Figure A13: VAR with various interest rate measures



Note: Impulse response functions to a monetary policy shock of the baseline (red) and using alternative measures of monetary policy: federal funds rate (blue dashed), federal funds rate extended by the shadow short rate of Wu and Xia (2016) (black dashed-dotted) and two year government bond (cyan dotted). The thick lines are median estimate and the area charts as well as dashed lines are 68% credible intervals.

Figure A14 shows impulse response functions to a monetary policy shock when we replace real GDP by industrial production (blue dashed line), GDP deflator by consumer price index (black dashed-dotted line), and the excess bond premium by the BAA corporate bond spread (cyan dotted line).

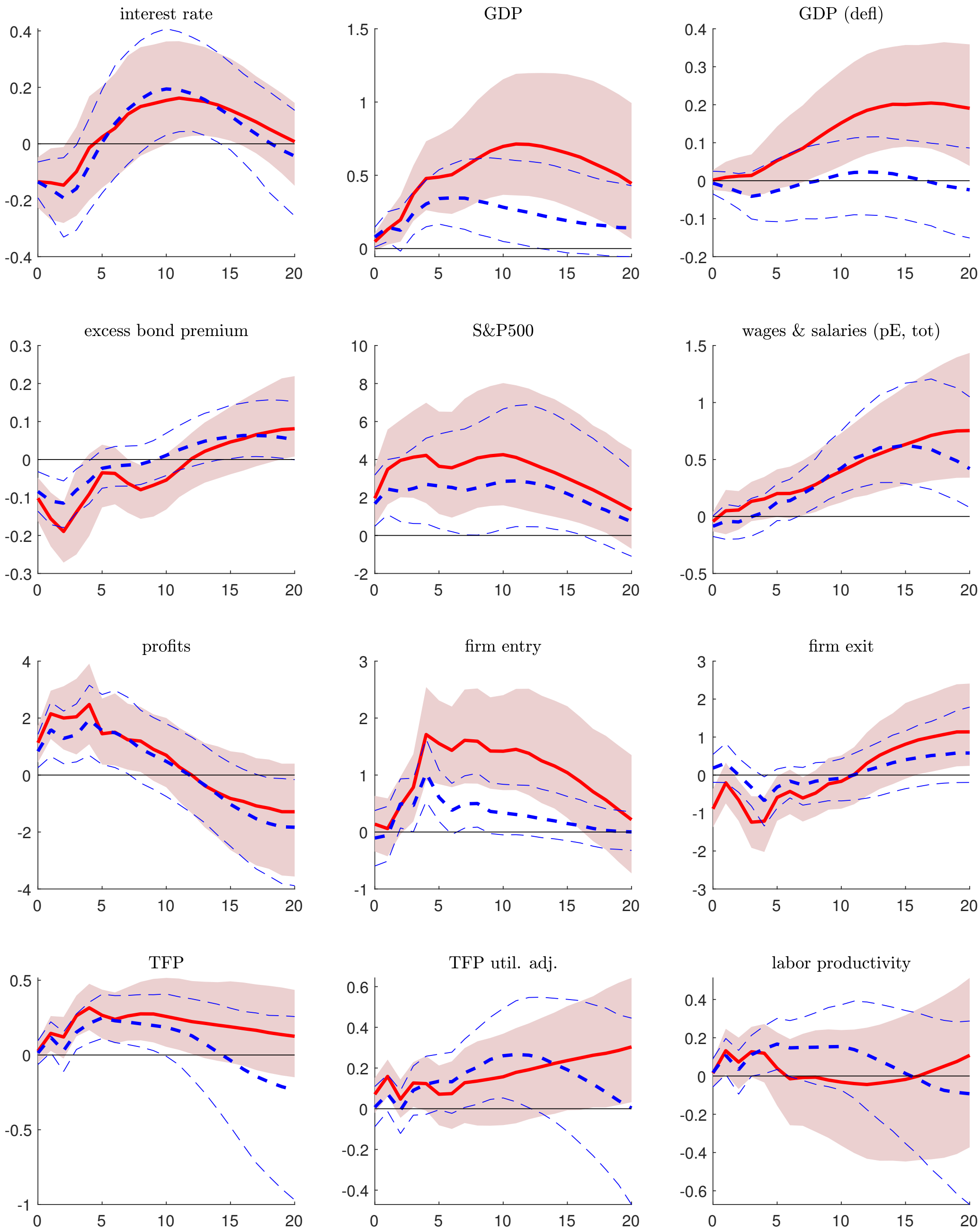
Figure A14: VAR with IP, CPI, BAA



Note: Impulse response functions to a monetary policy shock of the baseline (red) and using alternative measures of activity (blue dashed), prices (black dashed-dotted) and financial frictions (cyan dotted). The thick lines are median estimate and the area charts as well as dashed lines are 68% credible intervals.

Sample splits: Figure A15 shows impulse response functions to a monetary policy shock for the subsample until 2008Q2 and using the federal funds rate as monetary policy indicator (blue dashed). Subsample estimates are obtained from a VAR with zero restriction and a moderately tight Minnestoa prior with $\lambda = 0.7$.

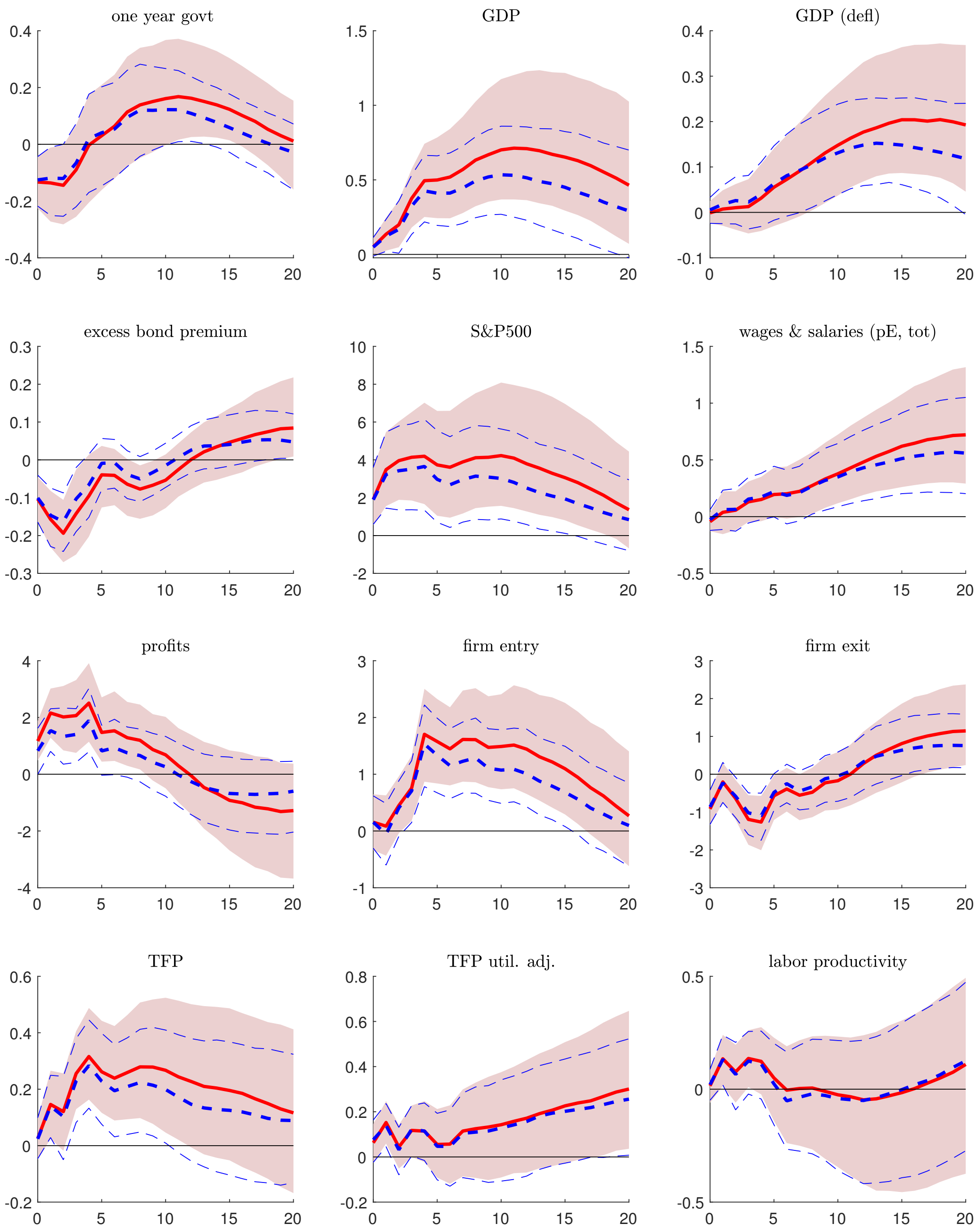
Figure A15: VAR until pre great recession period



Note: Impulse response functions to a monetary policy shock of the baseline (red) and the pre-great recession sample until 2008Q2 (blue-dashed). For the pre-great recession sample, a moderately loose Minnestoa prior is used with overall tightness of $\lambda = 0.7$ and the federal funds rate is used as policy indicator. The thick lines are median estimate and the dashed lines are 68% credible intervals.

Figure A16 shows impulse response functions to a monetary policy shock for the subsample that excludes the interest rate and stock price surprises in the apex of the great recession from 2008Q3 – 2009Q2 (blue dashed).

Figure A16: VAR excluding surprises from apex of great recession



Note: Impulse response functions to a monetary policy shock identified by sign-restrictions of Table 1. The estimates in red and blue correspond to the full sample and the sample excluding the apex of the great recession, ex 2008Q3-2009Q2, respectively. The thick lines are median estimate and the dashed lines are 68% credible intervals.

B Theoretical Analysis

B.1 Equilibrium Equations Baseline Model

Table A2 provides an overview of the equilibrium equations in the benchmark economy. The equilibrium is characterized by 33 endogenous and 3 exogenous variables $(A_t, \varepsilon_t^C, \varepsilon_t^M)$.

Table A2: Equilibrium equations

Firms		
Average pricing	(E1)	$\widetilde{\rho}_t = \widetilde{\mu}_t \widetilde{mc}_t$
Average markup	(E2)	$\widetilde{\mu}_t = \frac{\theta}{(\theta-1)\left(1-\frac{\tau}{2}(\pi_t-1)^2\right)+\tau\left(\pi_t(\pi_t-1)-E_t\left[\Lambda_t \frac{Y_{t+1}^C}{Y_t^C}\frac{S_t}{S_{t+1}}(\pi_{t+1}-1)\pi_{t+1}\right]\right)}$
Average marginal costs	(E3)	$\widetilde{mc}_t = \frac{w_t}{A_t \widetilde{z}_t}$
Real price	(E4)	$\widetilde{\rho}_t = S_t^{\frac{1}{\theta-1}}$
Average survivors' profit	(E5)	$\widetilde{d}_t = \left(1 - \widetilde{\mu}_t^{-1} - \frac{\tau}{2}\left(\pi_t - 1\right)^2\right)\frac{Y_t^C}{S_t} - f\frac{w_t R_t^{\vartheta}}{A_t}$
Entry condition	(E6)	$f_e \frac{w_t}{A_t} = v_t(\Psi_t + \Psi_t' H_t) + \beta E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-1} v_{t+1}\Psi_{t+1}'' H_{t+1}\right]$
Entry success probability	(E7)	$\Psi_t = 1 - g_3\left(exp\left(g_1\left(\frac{H_t}{H_{t-1}}-1\right)\right) + \frac{g_1}{g_2}exp\left(-g_2\left(\frac{H_t}{H_{t-1}}-1\right)\right) - 2\right)$
1st derivative entry success prob.	(E8)	$\Psi_t' = -g_1 g_3\left(exp\left(g_1\left(\frac{H_t}{H_{t-1}}-1\right)\right) - exp\left(-g_2\left(\frac{H_t}{H_{t-1}}-1\right)\right)\right)\frac{1}{H_{t-1}}$
2nd derivative entry success prob.	(E9)	$\Psi_t'' = g_1 g_3\left(exp\left(g_1\left(\frac{H_t}{H_{t-1}}-1\right)\right) - exp\left(-g_2\left(\frac{H_t}{H_{t-1}}-1\right)\right)\right)\frac{H_t}{(H_{t-1})^2}$
Profit at cut-off	(E10)	$\bar{d}_t = \left(1 - \bar{\mu}_t^{-1} - \frac{\tau}{2}\left(\frac{\bar{\rho}_t}{\bar{\rho}_{t-1}}\pi_t^C - 1\right)^2\right)\bar{\rho}_t \bar{y}_t - f\frac{w_t R_t^{\vartheta}}{A_t}$
Exit condition	(E11)	$\bar{d}_t = 0$
Price at cut-off	(E12)	$\bar{\rho}_t = \bar{\mu}_t \bar{mc}_t$
Markup at cut-off	(E13)	$\bar{\mu}_t = \frac{\theta}{(\theta-1)\left(1-\frac{\tau}{2}\left(\frac{\bar{\rho}_t}{\bar{\rho}_{t-1}}\pi_t^C-1\right)^2\right)+\tau\left(\frac{\bar{\rho}_t}{\bar{\rho}_{t-1}}\pi_t^C\left(\frac{\bar{\rho}_t}{\bar{\rho}_{t-1}}\pi_t^C-1\right)-E_t\left[\Lambda_t\frac{\bar{y}_{t+1}}{\bar{y}_t}\left(\frac{\bar{\rho}_{t+1}}{\bar{\rho}_t}\pi_{t+1}^C-1\right)\left(\frac{\bar{\rho}_{t+1}}{\bar{\rho}_t}\right)^2\pi_{t+1}^C\right]\right)}$
Marginal costs at cut-off	(E14)	$\bar{mc}_t = \frac{w_t}{A_t \bar{z}_t}$
Exit rate	(E15)	$\zeta_t = 1 - \left(\frac{z_m}{\bar{z}_t}\right)^{\kappa}$
Average productivity	(E16)	$\widetilde{z}_t = \bar{z}_t\left(\frac{\kappa}{\kappa-(\theta-1)}\right)^{\frac{1}{\theta-1}}$
Surviving firms	(E17)	$S_t = (1-\zeta_t)N_t$
Evolution of firms	(E18)	$N_t = (1-\delta)(N_{t-1} + \Psi_{t-1}H_{t-1})$
Average output	(E19)	$\widetilde{y}_t = \frac{Y_t^C}{\widetilde{\rho}_t S_t}$
Output at the cut-off	(E20)	$\bar{y}_t = \widetilde{y}_t\left(\frac{\bar{z}_t}{\widetilde{z}_t}\right)^{\theta}$
Households		
Wage setting 1st FOC	(E21)	$g_t = \frac{\theta_W - 1}{\theta_W}\left(w_t^*\right)^{1-\theta_W} w_t^{\theta_W}\varepsilon_t^C C_t^{-1} L_t + \beta\lambda_W E_t\left[\left(\pi_{t+1}^C\frac{w_{t+1}^*}{w_t^*}\right)^{\theta_W - 1} g_{t+1}\right]$
Wage setting 2nd FOC	(E22)	$g_t = \chi\left(\frac{w_t}{w_t^*}\right)^{\theta_w\left(1+\frac{1}{\eta}\right)} L_t^{1+\frac{1}{\eta}} + \beta\lambda_W E_t\left[\left(\pi_{t+1}^C\frac{w_{t+1}^*}{w_t^*}\right)^{\theta_W\left(1+\frac{1}{\eta}\right)} g_{t+1}\right]$
Real wage	(E23)	$w_t = \left(\lambda_W\left(\frac{w_{t-1}}{\pi_t^C}\right)^{1-\theta_W} + (1-\lambda_W)\left(w_t^*\right)^{1-\theta_W}\right)^{\frac{1}{1-\theta_W}}$
Euler equation shares	(E24)	$v_t = E_t\left[\Lambda_t(v_{t+1} + \frac{S_t}{N_t}\widetilde{d}_{t+1})\right]$
Euler equation for bonds	(E25)	$1 = \beta E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-1}\frac{R_t}{\pi_{t+1}^C}\right]$
Stochastic discount factor	(E26)	$\Lambda_t = \beta(1-\delta)E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-1}\right]$
CPI inflation	(E27)	$\pi_t^C = \frac{\widetilde{\rho}_{t-1}}{\widetilde{\rho}_t}\pi_t$
Aggregation and Monetary Policy		
Market clearing	(E28)	$Y_t = C_t + v_t H_t$
Accounting	(E29)	$Y_t = w_t L_t + \widetilde{d}_t S_t$
Aggregate consumption output	(E30)	$Y_t^C = \left(1 - \frac{\tau}{2}(\pi_t - 1)^2\right)^{-1} C_t$
Investment	(E31)	$I_t = v_t H_t$
Aggregate profits	(E32)	$d_t = \widetilde{d}_t S_t$
Taylor rule	(E33)	$log\left(\frac{R_t}{R}\right) = \phi_R\, log\left(\frac{R_{t-1}}{R}\right) + (1-\phi_R)\left[\phi_\pi\, log\left(\frac{\pi_t}{\pi}\right) + \phi_y\, log\left(\frac{Y_t}{Y_{t-1}}\right)\right] + \varepsilon_t^M$

B.2 Steady State Baseline Model

To derive the steady state in the benchmark DSGE model, we first normalize technology, labor and inflation in the steady state to one:

$$A = 1 \tag{A1}$$

$$L = 1 \tag{A2}$$

$$\pi = \pi^C = 1 \tag{A3}$$

From the household bond Euler equation (25), we get

$$R = \beta^{-1}\pi \tag{A4}$$

and from the definition of the stochastic discount factor (27):

$$\Lambda = \beta(1-\delta) \tag{A5}$$

Average markup and markups at the cut-off then follow from (11) and (35) as:

$$\widetilde{\mu} = \bar{\mu} = \frac{\theta}{(\theta-1)\left(1-\frac{\tau}{2}(\pi-1)^2\right) + \tau(1-\Lambda)\pi(\pi-1)} \tag{A6}$$

The strategy to derive the remaining steady state values is to obtain an expression for the total number of products N as a function of parameters and the steady state values (A1)-(A6). Starting from the average profit in (37), inserting (41) and (48) and using (A1) yields:

$$\widetilde{d} = \frac{1-\widetilde{\mu}^{-1}-\frac{\tau}{2}(\pi-1)^2}{1-\frac{\tau}{2}(\pi-1)^2}\frac{C}{S} - fwR^{\vartheta} \tag{A7}$$

Rearranging for C yields:

$$C = \frac{1-\frac{\tau}{2}(\pi-1)^2}{1-\widetilde{\mu}^{-1}-\frac{\tau}{2}(\pi-1)^2}\left(\widetilde{d}+fwR^{\vartheta}\right)S \tag{A8}$$

The aggregate resource constraint, obtained by combining (44) and (A2), is given by:

$$C + vH = w + \widetilde{d}S \tag{A9}$$

Inserting (A8) for C gives:

$$\frac{1-\frac{\tau}{2}(\pi-1)^2}{1-\widetilde{\mu}^{-1}-\frac{\tau}{2}(\pi-1)^2}\left(\widetilde{d}+fwR^{\vartheta}\right)S + vH = w + \widetilde{d}S \tag{A10}$$

In steady state, the entry condition (15) implies under the normalization $f_E = 1$, (A1) and the steady state properties of the success probability (14):

$$v = w \tag{A11}$$

Inserting this in (A10) and rearranging yields:

$$1=\left(\frac{\widetilde{\mu}^{-1}}{1-\widetilde{\mu}^{-1}-\frac{\tau}{2}(\pi-1)^2}\frac{\widetilde{d}}{w}+\frac{1-\frac{\tau}{2}(\pi-1)^2}{1-\widetilde{\mu}^{-1}-\frac{\tau}{2}(\pi-1)^2}fR^{\vartheta}\right)S+H \tag{A12}$$

Now, we want to replace the term $\frac{\widetilde{d}}{w}$. Combining (9) at the cut-off and (51), again using (A1), gives:

$$\left(1-\bar{\mu}^{-1}-\frac{\tau}{2}(\pi-1)^2\right)\bar{\rho}^{1-\theta}Y^C=fwR^{\vartheta} \tag{A13}$$

Using (5) and (10) at the cut-off yields:

$$\left(1-\bar{\mu}^{-1}-\frac{\tau}{2}(\pi-1)^2\right)\left(\bar{\mu}\widetilde{mc}\frac{\widetilde{z}}{\bar{z}}\right)^{1-\theta}Y^C=fwR \tag{A14}$$

Inserting (34) and (A6) gives:

$$\left(1-\widetilde{\mu}^{-1}-\frac{\tau}{2}(\pi-1)^2\right)\widetilde{\rho}^{1-\theta}Y^C=f\frac{\kappa}{\kappa-(\theta-1)}wR \tag{A15}$$

Note that the left-hand side is the first term in the average survivors' profit in (37). We can use this to rewrite (A15) as:

$$\widetilde{d}+fwR^{\vartheta}=f\frac{\kappa}{\kappa-(\theta-1)}wR \tag{A16}$$

This can be rewritten as:

$$\frac{\widetilde{d}}{w}=f\frac{\theta-1}{\kappa-(\theta-1)}R^{\vartheta} \tag{A17}$$

This is the term we wanted to replace in (A12), which we can now write as:

$$1=f\left(\frac{\widetilde{\mu}^{-1}}{1-\widetilde{\mu}^{-1}-\frac{\tau}{2}(\pi-1)^2}\frac{\theta-1}{\kappa-(\theta-1)}+\frac{1-\frac{\tau}{2}(\pi-1)^2}{1-\widetilde{\mu}^{-1}-\frac{\tau}{2}(\pi-1)^2}\right)RS+H \tag{A18}$$

Inserting (17) and rearranging yields:

$$N^{-1}=f\left(\frac{\widetilde{\mu}^{-1}}{1-\widetilde{\mu}^{-1}-\frac{\tau}{2}(\pi-1)^2}\frac{\theta-1}{\kappa-(\theta-1)}+\frac{1-\frac{\tau}{2}(\pi-1)^2}{1-\widetilde{\mu}^{-1}-\frac{\tau}{2}(\pi-1)^2}\right)R\frac{S}{N}+\frac{\delta}{1-\delta} \tag{A19}$$

This provides the steady state of the number of firms N, given the endogenous destruction rate S/N. From the Euler equation in (26), we have:

$$1=\Lambda\left(1+\frac{S}{N}\frac{\widetilde{d}}{v}\right) \tag{A20}$$

Again using $v=w$ from (A11) and inserting (A19) yields:

$$1=\Lambda\left(1+f\frac{\theta-1}{\kappa-(\theta-1)}R^{\vartheta}\frac{S}{N}\right) \tag{A21}$$

Rearranging yields:

$$\frac{S}{N} = \frac{1}{fR^{\vartheta}} \frac{\kappa - (\theta - 1)}{\theta - 1} \frac{1 - \Lambda}{\Lambda} \tag{A22}$$

Inserting this into (A19) yields the steady state for the total number of firms:

$$N = \left(\frac{\widetilde{\mu}^{-1} + \left(1 - \frac{\tau}{2}(\pi - 1)^2\right) \frac{\kappa - (\theta - 1)}{\theta - 1}}{1 - \widetilde{\mu}^{-1} - \frac{\tau}{2}(\pi - 1)^2} \frac{1 - \Lambda}{\Lambda} + \frac{\delta}{1 - \delta} \right)^{-1} \tag{A23}$$

The number of active firms follows directly from (A22). The steady state values of all other variables can be solved recursively.

B.3 Model with Exogenous Exit

In the model with exogenous exit, we set $f = 0$ and $z = 1$ for all firms. As a result, all firms are identical and Equations (E10)-(E14) and (E20) – which refer to firms at the cut-off – can be removed from the system. Exogenous exit also implies that the number of (endogenously) surviving firms is equal to the number of total firms

$$S = N \tag{A24}$$

that the endogenous exit probability is equal to 0

$$\zeta = 0 \tag{A25}$$

and that average productivity is equal to 1:

$$\widetilde{z} = 1 \tag{A26}$$

The steady states (A1)-(A6) are valid in this model version as well. Similar to the benchmark model, the strategy is to find the steady state of the total number of products N. The first steps to derive the steady state are identical up to (A12), which is now given by:

$$N^{-1} = \frac{\delta}{1 - \delta} + \frac{\widetilde{\mu}^{-1}}{1 - \widetilde{\mu}^{-1} - \frac{\tau}{2}(\pi - 1)^2} \frac{\widetilde{d}}{w} \tag{A27}$$

From the Euler equation with respect to share holdings (26), we have that:

$$\frac{\widetilde{d}}{w} = \frac{1 - \Lambda}{\Lambda} \tag{A28}$$

Inserting (A28) in (A27) yields the steady state for the total number of firms:

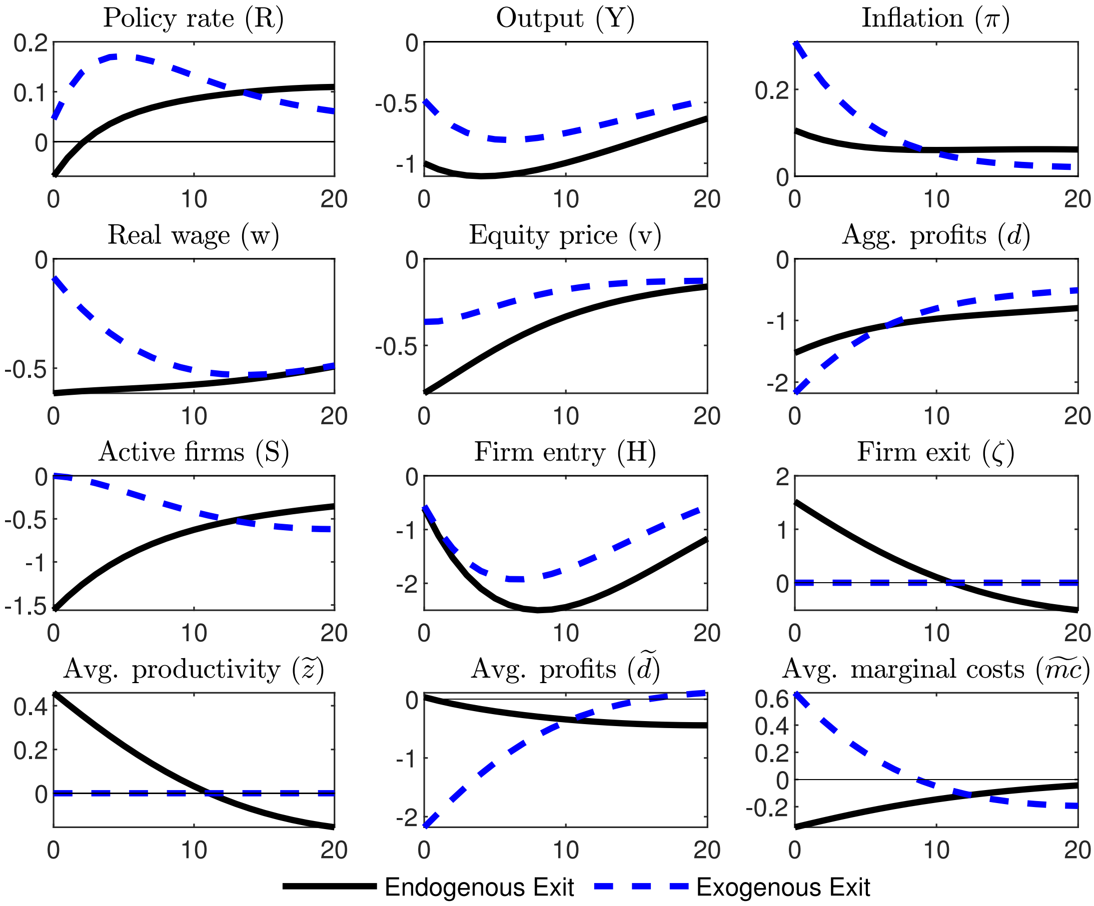
$$N = \left(\frac{\widetilde{\mu}^{-1}}{1 - \widetilde{\mu}^{-1} - \frac{\tau}{2}(\pi - 1)^2} \frac{1 - \Lambda}{\Lambda} + \frac{\delta}{1 - \delta} \right)^{-1} \tag{A29}$$

All remaining steady state values can then be solved recursively.

B.4 Technology and Demand Shocks

In this section, we investigate the role of the exit channel for macroeconomic (non-policy) shocks. Figure A17 shows the transmission of a contractionary technology shock, comparing the benchmark model with endogenous exit to a model where firms' exit probability is exogenous and constant.

Figure A17: Technology shock

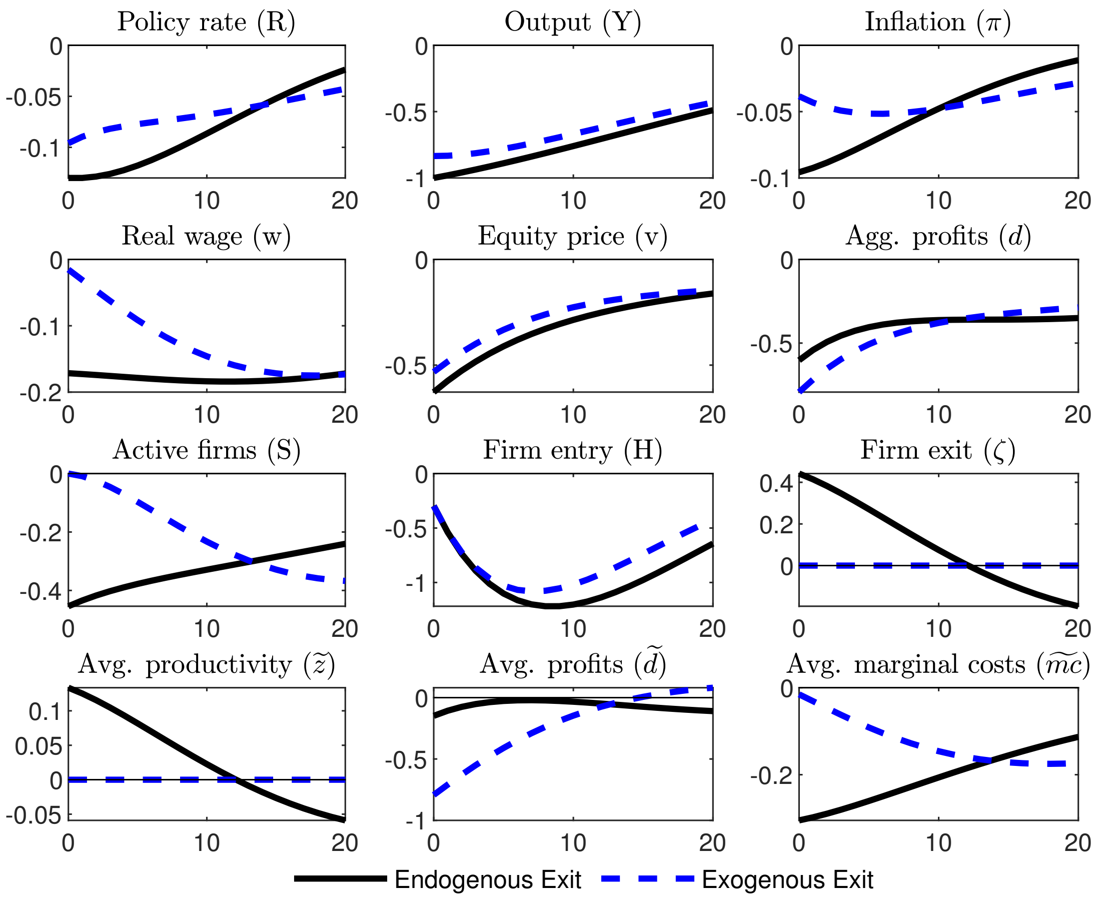


Note: Impulse response functions for a contractionary technology shock with an autoregressive coefficient of 0.9 in the baseline economy (with endogenous exit) and a variant with exogenous exit. The shock size is calibrated to yield a one-percent increase of output in the baseline economy. Inflation, interest rate and exit are shown in percentage-point deviations from steady state, all other variables in percentage deviations.

As described by Hamano and Zanetti (2017) and Rossi (2019), negative technology shocks increase real marginal costs and thus lower expectations of future profits, thereby disincentivizing the entry of new firms. The firm-specific productivity cut-off required for profitability increases, such that more firms exit the market. As a result, the contraction is more pronounced relative to a model with exogenous exit. Only relatively more productive firms are able to survive, causing average productivity to increase initially. As the economy reverts to the initial equilibrium, firm exit drops below baseline, reflecting a decreasing cut-off level of productivity.

Figure A18 shows the transmission of a contractionary preference shock. Similar to the technology shock analyzed above, corporate profits decrease in the recession. Firms' creation and destruction is also notably qualitatively similar to technology shocks. This highlights that the corresponding effect on profits is crucial for the associated reaction of firms' extensive margin, while the source of the aggregate contraction is not essential.

Figure A18: Demand shock

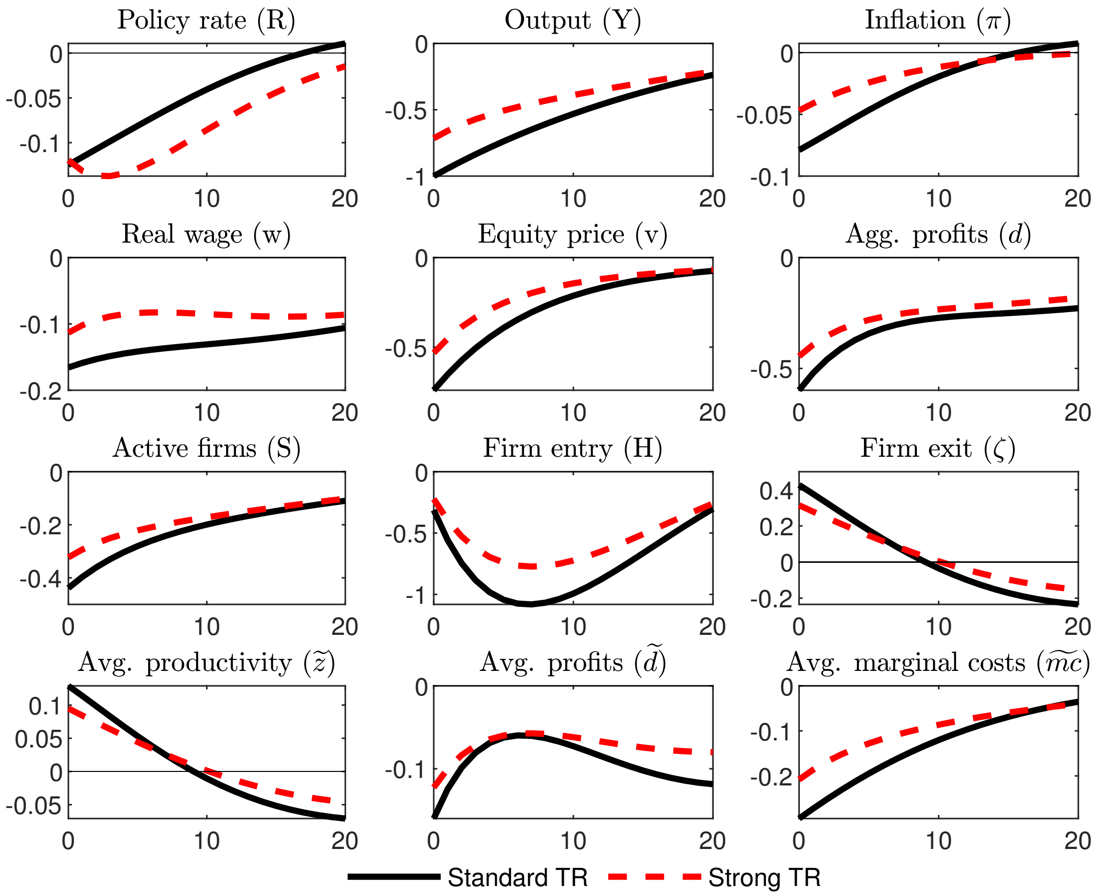


Note: Impulse response functions for a contractionary preference shock with an autoregressive coefficient of 0.9 in the baseline economy (with endogenous exit) and a variant with exogenous exit. The shock size is calibrated to yield a one-percent increase of output in the baseline economy. Inflation, interest rate and exit are shown in percentage-point deviations from steady state, all other variables in percentage deviations.

B.5 Monetary Policy over the Business Cycle

In this section, we investigate the effects of monetary policy over the business cycle. To this end, we analyze how the endogenous response of monetary policy to non-policy shocks alters firm entry and exit. Figure A19 shows the impulse responses for a contractionary demand shock that lowers output in the baseline economy by one percent. We compare two Taylor rules: Our baseline variant with conventional parameter values and a rule with a considerably larger coefficient $\phi_\pi = 5$ (to make the comparison particularly illustrative).

Figure A19: Demand shock



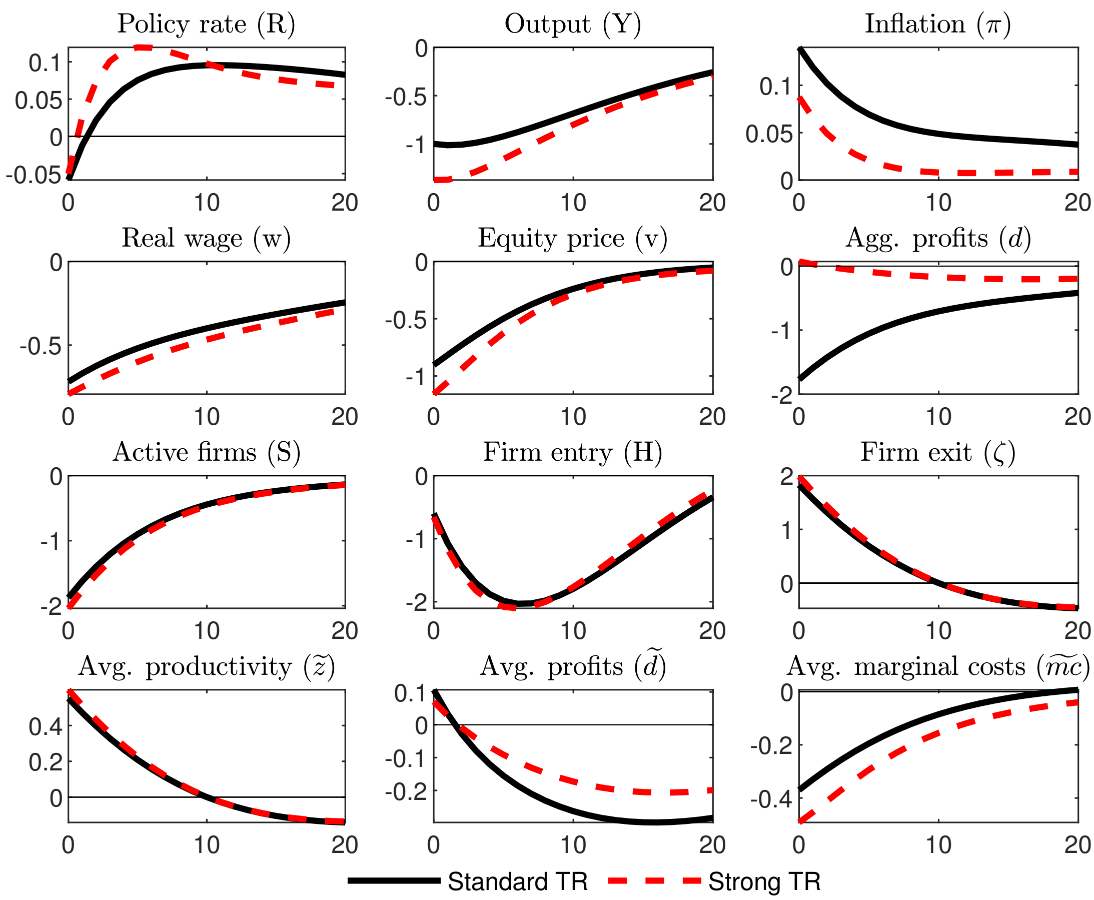
Note: Impulse response functions for a demand shock with an autoregressive coefficient of 0.9 in the baseline economy. The shock size is calibrated to yield a one-percent decrease of output under the baseline Taylor rule. Inflation, interest rate and exit probability are shown in percentage-point deviations from steady state, all other variables in percentage deviations.

Consider first the baseline Taylor rule (solid black lines). The contractionary shock is deflationary. The reduction of economic activity is generated by a fall in both consumption and investment. Labor demand falls, alongside a reduction in the real wage. The decrease of aggregate demand depresses firm profits, such that relatively less productive firms are not profitable anymore. As a consequence, exit rates increase. The reduction in consumption requires increasing real interest rates. No-arbitrage between equity and bonds requires that the return to equity increases as well. This happens via a fall in equity prices (relative to the future) and thus the ex-ante firm value. In turn, firm entry declines. As a result, relatively less productive firms exit the market and firms that remain are more productive. The associated increase in average productivity is consistent with a cleansing effect of recessions Caballero and Hammour (1994).

If the central bank operates according to a stronger Taylor rule (red dashed lines), it dampens the real effects of the demand shock. Relative to the baseline Taylor rule, the contraction of output and inflation is less severe. Under this Taylor rule, the central bank generates a lower increase of real interest rates. In turn, the required fall in equity prices to restore no-arbitrage is also lower. This also stabilizes firm profits and reduces the pressure on relatively unproductive firms to exit the market. As a consequence, firm exit increases substantially less. Equivalently, a more aggressive monetary policy responses along the business cycle dampens the cleansing effect of the demand-driven recession by allowing less productive firms to survive.

Turning to technology-driven recessions, Figure A20 shows the impulse responses following a contractionary aggregate productivity shock.

Figure A20: Technology shock



Note: Impulse response functions for a technology shock with an autoregressive coefficient of 0.9 in the baseline economy. The shock size is calibrated to yield a one-percent decrease of output in the baseline model. Inflation, interest rate and exit probability are shown in percentage-point deviations from steady state, all other variables in percentage deviations.

As shown by Hamano and Zanetti (2017), such shocks are associated with a cleansing effect, similar to the case of demand shock discussed above. Here, the fixed cost channel is particularly important, because these costs are specified in terms of effective labor units. In combination with the direct demand channel, firm's production costs rise strongly after an exogenous drop in productivity. This induces the exit of relatively less productive firms and hampers entry of new firms. In contrast to demand shocks, the technology shock is inflationary such that output and inflation move in opposite directions.

Under the baseline Taylor rule, the motive for stabilizing output dominates initially, such that the central bank sets lower nominal interest rates. This amplifies the inflationary effect. After few quarters, with the economy recovering, the inflation stabilization motive starts to dominate such that the central bank raises nominal interest rates above the steady state. In the medium-run, this contributes to lower aggregate demand. Overall, the impact of monetary policy on the cleansing effect of supply-driven recessions is minor: Firm exit and entry rates are highly similar across the two different Taylor rules, mirrored by barely noticeable differences in average productivity.

To summarize, central banks that aim to stabilize inflation and business cycles inherently affect firm dynamics. Via the extensive firm margin, expansionary monetary policy counteracts the cleansing effect of demand-driven recessions, while supply-driven cleansing is barely affected. When designing monetary policy, these potentially important side-effects should be taken into consideration. This opens the scope for future research to investigate the normative implications of the exit channel, in particular whether there exist trade-offs between macroeconomic stabilization and the cleansing effect of recessions.