EUROPEAN CENTRAL BANK

EUROSYSTEM

Working Paper Series

Matthias Meier, Timo Reinelt

Monetary policy, markup dispersion, and aggregate TFP

ECB – Lamfalussy
Fellowship Programme

No 2427 / June 2020

ECB Lamfalussy Fellowship Programme

This paper has been produced under the ECB Lamfalussy Fellowship programme. This programme was launched in 2003 in the context of the ECB-CFS Research Network on “Capital Markets and Financial Integration in Europe”. It aims at stimulating high-quality research on the structure, integration and performance of the European financial system.

The Fellowship programme is named after Baron Alexandre Lamfalussy, the first President of the European Monetary Institute. Mr Lamfalussy is one of the leading central bankers of his time and one of the main supporters of a single capital market within the European Union.

Each year the programme sponsors five young scholars conducting a research project in the priority areas of the Network. The Lamfalussy Fellows and their projects are chosen by a selection committee composed of Eurosystem experts and academic scholars. Further information about the Network can be found at http://www.eufinancial-system.org and about the Fellowship programme under the menu point “fellowships”.

Abstract

We document three new empirical facts: (i) monetary policy shocks increase the markup dispersion across firms, (ii) they increase the relative markup of firms with stickier prices, and (iii) firms with stickier prices have higher markups. This is consistent with a New Keynesian model in which price rigidity is heterogeneous across firms. In the model, firms with more rigid prices optimally set higher markups and their markups increase by more after monetary policy shocks. The consequent increase in markup dispersion explains a negative aggregate TFP response. In a calibrated model, monetary policy shocks generate substantial fluctuations in aggregate productivity.

Keywords: Monetary policy, markup dispersion, heterogeneous price rigidity, aggregate productivity.

JEL codes: E30, E50.

Non-technical Summary

A long-standing question in macroeconomics is how monetary policy affects real economic activity. While a lot of progress has been made on this question, we argue that an important aspect of how monetary policy transmits is missing in previous research. The missing aspect is related to the empirical regularity that tighter monetary policy lowers aggregate productivity. In other words, if monetary policy raises short-term interest rates, less aggregate output (GDP) is produced from a given amount of aggregate inputs (the capital stock and hours worked). Prior explanations for this regularity have mostly been based on R&D investment. However, lower R&D investments are unlikely to explain short-run fluctuations of aggregate productivity. We propose a novel explanation why monetary policy affects aggregate productivity that is based on markup dispersion and heterogeneity in price rigidity across firms.

Empirically, we use aggregate and firm-level data from the US to document three new facts: (i) Monetary policy shocks increase the markup dispersion across firms, i.e., markups differences across firms widen. (ii) Monetary policy shocks increase the markup of firms that adjust prices less frequently relative to the markup of firms that adjust prices more frequently. (iii) Firms that adjust prices less frequently have higher markups on average. It follows from facts (ii) and (iii) that heterogeneity in price rigidity across firms (partly) explains why markup dispersion increases, fact (i). In a standard New Keynesian model, higher markup dispersion implies misallocation of inputs across firms, which lowers aggregate productivity. In the model environment, equal markups across firms imply firm-specific outputs that maximize aggregate productivity. Any increase in markup dispersion lowers productivity. We show that a sizable share of the aggregate productivity response to monetary policy can be associated to the response in markup dispersion.

Analytically, we show that tighter monetary policy increases markup dispersion if firms with stickier prices charge higher markups. We further show that firms with stickier prices may optimally charge higher markups. The reason is that profits fall more rapidly when markups fall below the profit-maximizing level than if markups rise above that level. This is true locally but also in the limit. When a firm's price is excessively high, demand for its product tends toward zero and so does its profit. When a firm's price is low, the price may not cover marginal costs, in which case it makes losses. If firms adjust prices infrequently or imperfectly to changes in marginal costs, then firms set a higher markup in order to mitigate the downside risk of low future markups.

Quantitatively, we study the effect of monetary policy on aggregate productivity in a New Keynesian model with heterogeneous price rigidity. The calibrated model is consistent with our empirical findings. Firms with stickier prices set higher markups, and monetary policy shocks raise markup dispersion and lower aggregate productivity. The drop of productivity in the model explains about half of the empirical estimate. We further use the model to study a counterfactual scenario in which the central bank attributes all productivity fluctuations, including productivity responses to demand shocks, to supply shocks. In this scenario, monetary policy is substantially less successful in stabilizing demand-driven GDP fluctuations.

1 Introduction

We investigate to what extent cross-sectional heterogeneity in markups and price rigidity are important for monetary transmission. We document three new empirical facts: (i) monetary policy shocks increase the markup dispersion across firms, (ii) monetary policy shocks increase the relative markup of firms that adjust prices less frequently, and (iii) firms that adjust prices less frequently have higher markups. Furthermore, we document that aggregate TFP falls after monetary policy shocks and that a sizable share of this response can be related to the response in markup dispersion. Analytically, we show that firm heterogeneity in price-setting frictions can explain facts (i)–(iii). The fundamental reason is a precautionary price setting motive. In a calibrated New Keynesian model with heterogeneous price rigidity, monetary policy shocks explain substantial fluctuations in markup dispersion and aggregate productivity.

Our empirical analysis builds on quarterly balance-sheet data from the US, which we use to estimate firm-level markups. To capture variation price adjustment frequencies across firms, we use data on price adjustment frequencies in five-digit industries together with data on the firm-specific sales composition across industries. In addition, we construct high-frequency identified monetary policy shocks. Our main empirical finding is that the dispersion of markups within industries significantly increases after monetary policy shocks. The response is persistent and peaks about two years after the shock. Heterogeneity in price stickiness (within industries) can partially explain this response. We document that firms with stickier prices have higher markups on average and increase their markups by more after monetary policy shocks.

The response of markup dispersion is important to understand the transmission of monetary policy shocks. Markup dispersion affects the allocative efficiency of inputs across firms and thereby measured aggregate TFP.[1] Empirically, a one standard deviation monetary policy shock lowers aggregate (utilization-adjusted) TFP by 0.8% (0.4%) at a two-year horizon, compared to a 1% response of aggregate output. Following the approach in Hsieh and Klenow (2009) and Baqaee and Farhi (2020), we show that the estimated increase in markup dispersion accounts for at least half of the utilization-adjusted TFP response at a two-year horizon. At more distant horizons, markup dispersion accounts for a decreasing fraction of the aggregate TFP response.

Analytically, we show that a positive correlation between firm-level markups and the pass-through from marginal costs to prices is sufficient for markup dispersion to increase

[1] Markup (or price) dispersion is well known to lower aggregate TFP in the New Keynesian literature, e.g., Galí (2015), as well as in the macro development literature on factor misallocation, e.g., Hsieh and Klenow (2009).

after monetary policy shocks that lower marginal costs. We further show that such positive correlation can endogenously arise from heterogeneity in price-setting frictions due to precautionary price setting.[2] The types of price-setting frictions we consider are a Calvo (1983) friction, Taylor (1979) staggered price setting, Rotemberg (1982) convex adjustment costs, and Barro (1972) menu costs. In a nutshell, the presence of heterogeneous price-setting frictions gives rise to TFP effects of monetary policy. This implication of heterogeneous price-setting frictions is novel.

Quantitatively, we study the effect of monetary policy shocks on aggregate TFP in a New Keynesian model with heterogeneous price rigidity. We calibrate the heterogeneity in price rigidity across firms to match the within-sector dispersion in price adjustment frequencies documented in Gorodnichenko and Weber (2016). The calibrated model is consistent with our empirical findings. Firms with stickier prices set higher markups and monetary policy shocks raise markup dispersion. A one standard deviation monetary policy shock lowers aggregate TFP by -0.34%. This is more than half of the empirically estimated peak response of utilization-adjusted aggregate TFP. Compared to a model with homogeneous price rigidity, heterogenous rigidity generates additional persistence, which is driven by firms with above-average price rigidity. This is the frequency composition effect described by Carvalho (2006) which can be understood through lower aggregate price selection, see Carvalho and Schwartzman (2015). Our model abstracts from real rigidities and input-output networks, which can amplify shock propagation, see, e.g., Carvalho and Nechio (2016) and Pasten et al. (forthcoming).

To capture the productivity effects of monetary policy in the model, the solution technique plays a crucial role. We use non-linear solution methods to compute the stochastic steady state, to which the economy converges in the presence of uncertainty but absent shocks. In this steady state, firms in the highest quintile of price rigidities charge a 10.9% higher markup than firms in the most flexible quintile. Around this steady state, monetary policy shocks have first-order effects on markup dispersion and aggregate TFP. In contrast, in the determinstic steady state all firms charge the same markup irrespective of their price rigidity, and. Around this steady state, monetary policy shocks have no first-order effects on markup dispersion and aggregate TFP.[3]

Taking the productivity effects of monetary policy into account matters for the effec-

[2]Because the profit function is asymmetric in a firm's price, firms with stronger price setting frictions optimally set higher markups. This resembles the markup response to higher uncertainty in Fernandez-Villaverde et al. (2015).

[3]In the presence of positive trend inflation, markup dispersion is positive even in the deterministic steady state of a homogeneous firm model. However, monetary policy shocks lower markup dispersion and increase aggregate TFP, see Ascari and Sbordone (2014), counterfactual to the empirical evidence.

tiveness of monetary policy in stabilizing the output gap. The traditional view attributes fluctuations in aggregate productivity to exogenous technology shocks. Whereas technology shocks change natural output, monetary policy shocks do not. With a Taylor rule that responds to the output gap, it is important for the monetary authority to distinguish the source of productivity fluctuations. If the monetary authority misattributes endogenous TFP fluctuations to technology shocks, interest rates are adjusted less strongly. In this case, monetary policy shocks have a roughly 20% larger effect on GDP.

The endogenous TFP effects also matter for the response of (aggregate) markups to monetary policy shocks. In the textbook New Keynesian model monetary policy shocks raise markups. However, some empirical evidence points in the opposite direction, see Nekarda and Ramey (2019). If the endogenous TFP response to monetary policy shocks is sufficiently strong, the aggregate marginal costs may rise and hence aggregate markups fall after monetary policy shocks. This argument extends to sector or firm-level markups if price rigidities are heterogeneous within sectors or within firms.

This paper is closely related to four branches of literature. First, a growing literature studies the role of heterogeneous price rigidity for monetary policy. Compared to the case of homogeneous price rigidity, monetary policy shocks have larger and more persistent effects, see Carvalho (2006), Nakamura and Steinsson (2010), Carvalho and Schwartzman (2015), and Pasten et al. (forthcoming), and optimal monetary policy differs, see Aoki (2001), Eusepi et al. (2011), and **?**. We show that such heterogeneity gives rise to productivity effects of monetary policy. Similarly, Baqaee and Farhi (2017) show that negative money supply shocks lower aggregate TFP if sticky-price firms have higher ex-ante markups than flexible-price firms. We provide empirical evidence which supports this transmission channel and show that the rigidity–markup correlation can arise endogenously from differences in price rigidity.

Second, this paper relates to a literature that studies the productivity response to monetary policy, e.g., Christiano et al. (2005), Comin and Gertler (2006), Moran and Queralto (2018), Garga and Singh (2019), and Jordà et al. (2020). We confirm the empirical finding that monetary policy shocks lower aggregate productivity, but provide a novel explanation based on markup dispersion. Previously, Christiano et al. (2005) show that variable utilization and fixed costs explain a relatively small fraction of the aggregate productivity response. Moran and Queralto (2018) and Garga and Singh (2019) show that R&D investment falls after monetary policy shocks, which may ultimately lower productivity. However, it is unclear whether the R&D response has a large impact on aggregate productivity at short horizons. For example, Comin and Hobijn (2010) estimate that new technologies are adopted with an average lag of five years.

Third, our paper relates to a literature on the relation between trend inflation and price dispersion. Whereas we show a positive response of markup dispersion to monetary policy shocks, Nakamura et al. (2018) document that price dispersion is flat across periods of high and low inflation since the 1970s. This suggests that the response of firms to changes in trend inflation differs from their response to the more transitory effects of monetary policy shocks. For example, in a high-inflation environment, managers may schedule more frequent meetings to discuss price changes than in low-inflation environment, see Romer (1990) and Levin and Yun (2007). Focusing on more recent years, characterized by low trend inflation, Sheremirov (2019) documents a positive co-movement between price dispersion (excluding sales) and inflation.

Fourth, this paper relates to a growing literature that studies allocative efficiency over the business cycle. Eisfeldt and Rampini (2006) show that capital misallocation is countercyclical. Fluctuations in allocative efficiency may be driven by various business cycle shocks, e.g., aggregate demand shocks (Basu, 1995), aggregate productivity shocks (Khan and Thomas, 2008), uncertainty shocks (Bloom, 2009), financial shocks (Khan and Thomas, 2013), or supply chain disruptions (Meier, 2020). We relate to this literature by studying the transmission of monetary policy shocks through allocative efficiency. Interestingly, the effects of short-run decreases in interest rates appear to differ in sign from the effects of long-run decreases. Whereas we show that the former lowers misallocation, Gopinath et al. (2017) show that the latter increases misallocation through size-dependent financial frictions. Relatedly, Oikawa and Ueda (2018) study the long-run effects of nominal growth through reallocation across heterogeneous firms.

The remainder of this paper is organized as follows. Section 2 presents the empirical evidence. Section 3 presents analytical results. Section 4 presents a quantitative model with results. Section 5 concludes and an Appendix follows.

2 Empirical Evidence

In this section, we document three new empirical facts: (i) monetary policy shocks increase the markup dispersion across firms, (ii) monetary policy shocks increase the relative markup of firms that adjust prices less frequently, and (iii) firms with higher markups adjust prices less frequently. We further document that aggregate TFP falls after monetary policy shocks and that a sizable share of this response can be linked to the response in markup dispersion.

2.1 Data, identification, and estimation

Firm-level markups. We use quarterly balance-sheet data of publicly-listed US firms from Compustat.[4] We estimate firm-level markups adopting the approach of Hall (1986, 1988) and De Loecker and Warzynski (2012). If firms have a flexible input factor, V_{it}, then cost minimization implies that the markup μ_{it} of firm i in quarter t can be computed as

$$\mu_{it} = \frac{\text{output elasticity of } V_{it}}{\text{revenue share of } V_{it}}. \tag{2.1}$$

For all our main empirical results, we focus on differences of firm-level log markups from their industry-quarter mean. In addition, we assume that firms within the same two-digit industry-quarter have a common output elasticity.[5] Hence, firm-level log markups in deviation from the industry-quarter mean do not depend on the output elasticity. Our main results are therefore not affected by the critique in Bond et al. (2020) of estimating output elasticities from revenue data. Following De Loecker et al. (2020), we assume firms produce using capital and a composite input of labor and materials, which is the flexible factor. We estimate the revenue share as the firm-quarter-specific ratio of costs of goods sold to sales. In Section 2.4, we consider two robustness exercises. First, we allow for firm-specific output elasticities by estimating a four-digit industry-specific translog production technology. Second, we estimate output elasticities using cost shares at the four-digit industry-quarter level.

We consider all industries except public administration, finance, insurance, real estate, and utilities. We drop firm-quarter observations if sales, costs of goods sold, or fixed assets are only reported once in the associated year. We further drop observations if quarterly sales growth is above 100% or below -67% or if real sales are below 1 million USD. We finally drop the bottom and top 5% of the estimated markups. Appendix A.1 provides more details and summary statistics in Table 3. Our results are robust to variations in the data treatment as we discuss toward the end of this section. In our subsequent empirical analysis, we focus on deviations of firm-level log markups from their industry or industry-quarter specific mean. We do so primarily to control for industry-specific characteristics such as competitiveness and production technology.

Price rigidity. We use average industry-level price adjustment frequencies over 2005–2011 based on PPI micro data from Pasten et al. (forthcoming). The data is at the level of five-digit industries. We further use the Compustat segment files, which provides sales and the industry

[4]Compustat data has two central advantages over many other firm-level balance-sheet datasets: First, it is available at quarterly frequency instead of annual (e.g., ASM) or every five years (e.g., Census). This is important to estimate responses to monetary policy shocks. Second, it covers all sectors.

[5]This is consistent with firms in a given industry-quarter using a common Cobb-Douglas technology.

code of business segments within firms. The firm-specific sales composition across industries allows us to compute firm-specific price adjustment frequencies as sales-weighted average of industry-specific price adjustment frequencies. We expect this procedure to underestimate the true extent of heterogeneity across firms, which should bias our subsequent regression coefficients toward zero. For some firms, Compustat segment files are not available and for others they report only one segment per firm. We can construct firm-specific price adjustment frequencies for 42% of firms. For the remaining firms, we use the price adjustment frequency of the five-digit industry they operate in. More details are provided in Appendix A.2.[6] Finally, we define the implied price duration as $-1/\log(1 - \text{price adjustment frequency})$.

Monetary policy shocks. We use high-frequency data of federal fund future prices, which we acquired from the Chicago Mercentile Exchange. We identify monetary policy shocks through changes of the future price in a narrow time window around FOMC announcements. The identifying restrictions are that the risk premium does not change and that no other macroeconomic shock materializes within the time window. We denote the price of a future by f, and by τ the time of a monetary announcement.[7] We use a thirty-minute window around FOMC announcements, as in Gorodnichenko and Weber (2016). Let $\Delta\tau^- = 10$ minutes and $\Delta\tau^+ = 20$ minutes, then monetary policy shocks are

$$\varepsilon_\tau^{\mathrm{MP}} = \mathrm{f}_{\tau+\Delta\tau^+} - \mathrm{f}_{\tau-\Delta\tau^-}. \tag{2.2}$$

To aggregate the shocks to quarterly frequency, we follow Ottonello and Winberry (2020). We assign daily shocks fully to the current quarter if they occur on the first day of the quarter. If they occur within the quarter, we partially assign the shock to the subsequent quarter. This procedure weights shocks across quarters corresponding to the amount of time agents have to respond. Formally, we compute quarterly shocks as

$$\varepsilon_t^{\mathrm{MP}} = \sum_{\tau\in\mathcal{D}(t)} \phi(\tau)\varepsilon_\tau^{\mathrm{MP}} + \sum_{\tau\in\mathcal{D}(t-1)} (1-\phi(\tau))\varepsilon_\tau^{\mathrm{MP}}, \tag{2.3}$$

where $\mathcal{D}(t)$ is the set of days in quarter t and $\phi(\tau)$ = (remaining number of days in quarter t after announcement in τ) / (total number of days in quarter t).

As a baseline, we construct monetary policy shocks from the three-months ahead federal funds future, as in Gertler and Karadi (2015). Our baseline excludes unscheduled meet-

[6]Our results are robust when only using sectoral price adjustment frequencies.

[7]We obtain time and classification of FOMC meetings from Nakamura and Steinsson (2018) and the FRB. We obtain time stamps of the press release from Gorodnichenko and Weber (2016) and Lucca and Moench (2015).

ings and conference calls.[8] Following Nakamura and Steinsson (2018), our baseline further excludes the apex of the financial crisis from 2008Q3 to 2009Q2.[9] The monetary policy shock series covers 1995Q2 through 2018Q3. We discuss alternative monetary policy shocks in Section 2.4. Table 4 in the Appendix reports summary statistics and Figure 6 shows the shock series.

2.2 Markup dispersion and heterogeneous price rigidity

Fact 1: Monetary policy shocks increase markup dispersion

We first estimate the response of markup dispersion to monetary policy shocks. To compute markup dispersion within industry and time, we subtract from firm-level markups the two or four-digit industry s and quarter t specific means. Our baseline measure of markup dispersion is hence the cross-sectional variance $\mathbb{V}_t(\log \mu_{it} - \overline{\log \mu}_{st})$, where $\overline{\log \mu}_{st}$ denotes the mean log markup across all firms in industry s and quarter t. Recall that for computing $\log \mu_{it} - \overline{\log \mu}_{st}$, we do not require an estimate of the output elasticity under our assumption that firms within an industry-quarter have a common output elasticity. Figure 6 in the Appendix plots this measure of markup dispersion over time.[10] To estimate the effects of monetary policy shocks on markup dispersion, we use the local projection

$$y_{t+h} - y_{t-1} = \alpha^h + \beta^h \varepsilon_t^{\mathrm{MP}} + \gamma_0^h \varepsilon_{t-1}^{\mathrm{MP}} + \gamma_1^h (y_{t-1} - y_{t-2}) + u_t^h, \tag{2.4}$$

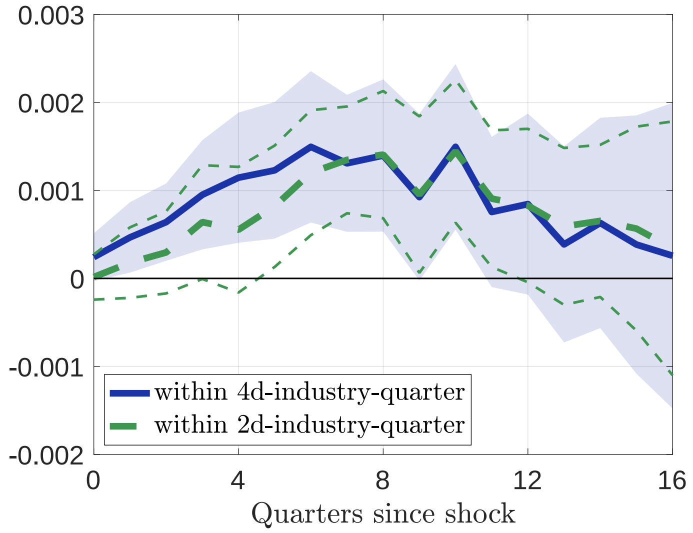
for $h = 0, \ldots, 16$ quarters and where y_t is markup dispersion. Figure 1 shows the response of markup dispersion, captured by the estimates of the coefficients $\{\beta^h\}$. The key finding is that markup dispersion increases significantly and persistently, within two-digit or four-digit industry-quarters. The response of markup dispersion peaks at about two years after the shock and reverts back to zero afterwards. This result proves robust in a large number of dimensions as we discuss in Section 2.4.

[8] Unscheduled meetings and conference calls are often the immediate response to adverse economic developments. Price changes around such meetings may directly reflect these developments, which invalidates the identifying restriction. Non-scheduled meetings are also more likely to communicate private information about the state of the economy. Our results remain broadly robust when including these meetings.

[9] We discard shocks during 2008Q3 to 2009Q2 and we do not regress post-2009Q2 outcomes on pre-2008Q3 shocks. Our results are robust to including this period.

[10] Similar to De Loecker et al. (2020), we find an increasing trend in markup dispersion.

Figure 1: Responses of markup dispersion to monetary policy shocks



Notes: The figure shows the responses of markup dispersion to a one-standard deviation contractionary monetary policy shock, coefficients in β^h in (2.4). The shaded and bordered areas indicate one standard error bands based on the Newey–West estimator.

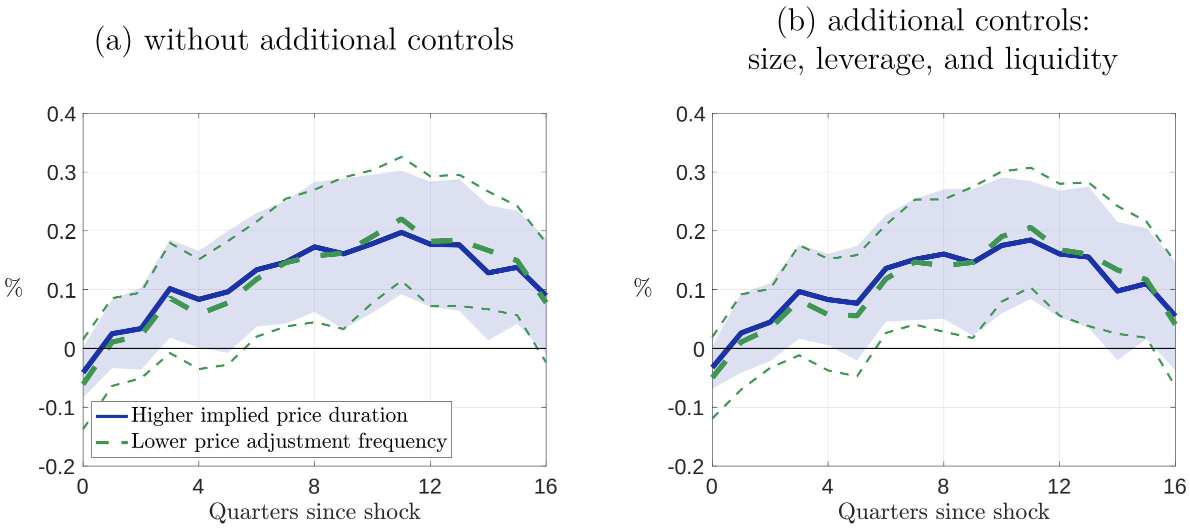
Fact 2: Monetary policy shocks increase the relative markups of firms with stickier prices

We next study the role of heterogeneous price stickiness in explaining the response of markup dispersion. We investigate whether the markup response to monetary policy shocks is stronger for firms with lower average price adjustment frequency. This is not necessarily the case if the average stickiness differs from the stickiness after monetary policy shocks, or if the marginal costs of firms with stickier prices respond differently from other firms.

We estimate panel local projections of firm-level log markups on the interaction between monetary policy shocks and firm-level price rigidity. We measure firm-level price rigidity by the price adjustment frequency or the implied price duration. Let Z_{it} denote a vector of firm-specific characteristics. We consider two specifications for Z_{it}: (i) including one of the two rigidity measures, and (ii) additionally including lags of firm size (log of total assets), leverage (total debt per total assets), and the ratio of liquid assets to total assets.[11] Our selection of controls is motivated by recent work in Ottonello and Winberry (2020) and Jeenas (2019), who study the transmission of monetary policy shocks through financial constraints. We use

[11] We demean the additional firm-level controls by the firm-level mean to focus on within-firm variation.

Figure 2: Relative markup response of firms with stickier prices to monetary policy shocks



Notes: The figures show the relative markup response of firms with a price adjustment frequency one standard deviation below (or with an implied price duration one standard deviation above) the two-digit-industry mean to a one standard deviation monetary policy shock. That is, we plot the appropriately scaled coefficients in B^h that are associated to price rigidity in the panel local projections (2.5). In panel (a), Z_{it} contains only price stickiness. In panel (b), Z_{it} also contains lagged log assets, leverage, and liquidity. The shaded and bordered areas indicate 90% error bands clustered by firm and quarter.

the panel local projection

$$y_{it+h} - y_{it-1} = \alpha_i^h + \alpha_{st}^h + B^h Z_{it}\varepsilon_t^{\mathrm{MP}} + \Gamma^h Z_{it} + \gamma^h(y_{it-1} - y_{it-2}) + u_{it}^h \tag{2.5}$$

for $h = 0, \ldots, 16$ quarters, in which we include two-digit-industry-time and firm fixed effects. To focus on the within-industry variation in the interaction between monetary policy shock and price rigidity, we subtract the corresponding two-digit industry mean from the measure of price rigidity. The main coefficients of interest are the elements of $\{B^h\}$ associated with price rigidity. These capture the relative markup increase for firms with stickier prices. Figure 2 shows the results. The markups of firms with stickier prices increase by significantly more after monetary policy shocks.[12] Firms with a price adjustment frequency one standard deviation above the associated two-digit-industry mean increase their markup by up to 0.2% more. Importantly, the estimates are almost identical when adding controls, see panel (b) of Figure 2.

[12]Using Driscoll–Krayy standard errors yields almost the same confidence bands as in Figure 2.

Fact 3: Firms with stickier prices charge higher markups

Finally, we study the correlation between price rigidity and markup. To compare markups with average price adjustment frequencies and implied price durations for 2005–2011, we compute average firm-level markups over the same time period. Columns (1) and (3) of Table 1 show that firms, which adjust prices less frequently than the two-digit industry average, charge markups significantly above the industry average. While this correlation is consistent with precautionary price setting, it may reflect omitted factors. In columns (2) and (4) we control for firm-specific average size, leverage and liquidity. Even with these controls, the estimated correlations remain of the same sign and statistically significant. In Table 1 we have excluded firms for which price setting frictions are practically irrelevant. In particular firms with a price adjustment frequency above 99% per quarter, which are about 3% of all firms. When including these, the relation between stickiness and markup remains positive, albeit somewhat less significant, see Table 5 in the Appendix.

Table 1: Regressions of markup on price stickiness

	(1)	(2)	(3)	(4)
Implied price duration	0.0538	0.0471		
	(0.0180)	(0.0154)		
Price adjustment frequency			-0.389	-0.333
			(0.100)	(0.0853)
Size		0.0107		0.0113
		(0.00340)		(0.00343)
Leverage		-0.00173		-0.00167
		(0.000649)		(0.000652)
Liquidity		0.553		0.545
		(0.0627)		(0.0640)
2-digit industry FE	Yes	Yes	Yes	Yes
Observations	3870	3867	3870	3867
Adjusted R^2	0.144	0.228	0.149	0.231

Notes: Regressions of firm-level markup on firm-level price adjustment frequency and implied price duration, respectively. Standard errors are clustered at the two-digit industry level and shown in parentheses.

2.3 Aggregate productivity

Fluctuations in markup dispersion can lead to changes in allocative efficiency of inputs across firms and thereby to fluctuations in aggregate TFP. To characterize this link, we build on the seminal work by Hsieh and Klenow (2009) and Baqaee and Farhi (2020). In a model with monopolistic competition and Dixit–Stiglitz aggregation, we can approximately express changes in aggregate TFP as

$$\Delta \log \mathrm{TFP}_t = -\frac{\eta}{2}\Delta \mathbb{V}_t(\log \mu_{it}) \; + \; \Big[\Delta \text{ exogenous productivity}\Big], \tag{2.6}$$

where η is the substitution elasticity between variety goods. The details of the derivation are provided in Appendix D.1.[13] An increase in the variance of log markups by 0.01 lowers aggregate TFP by $\frac{\eta}{2}\%$. Let us provide some intuition. With homogeneous firms, aggregate output is maximal for given aggregate inputs if all firms produce the same quantity, which implies equal markups across firms. If firms face heterogeneous productivity and demand shifts, the efficient allocation of inputs is not homogeneous across firms, but still implies equal markups. Conversely, markup dispersion is associated with an allocation of inputs across firms that implies aggregate TFP losses.

We empirically estimate the aggregate productivity response to monetary policy shocks and compare it with the productivity response implied by equation (2.6) and the response of markup dispersion in Figure 1. We consider aggregate TFP and utilization-adjusted aggregate TFP from Fernald (2014), as well as labor productivity and estimate their responses to monetary policy shocks through equation (2.4).[14] Panel (a) of Figure 3 shows that the responses of all three aggregate productivity measures are significantly negative and persistent. At a two-year horizon, a one-standard deviation monetary policy shock lowers aggregate TFP by 0.8%, labor productivity by 0.6% and utilization-adjusted aggregate TFP by 0.4%. For comparison, we show the responses of interest rates, aggregate output and inputs in Figure 10 in the Appendix. A monetary policy shock of this magnitude raises the federal funds rate by up to 30 basis points and lowers aggregate output by about 1% at a two-year horizon. Aggregate factor inputs respond little and thus aggregate TFP accounts for 50–80% of the output response at a two-year horizon.

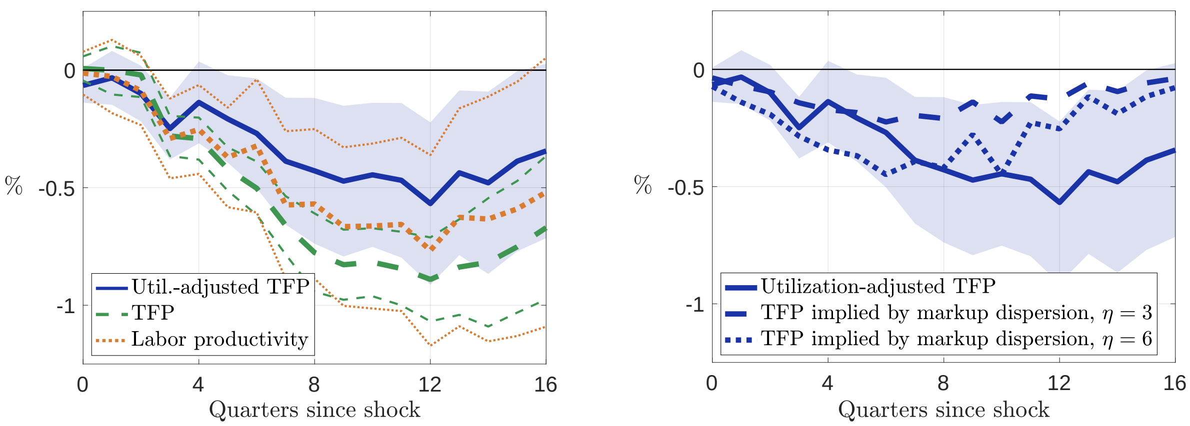
[13]In the calibrated New Keynesian model of Section 4, equation (2.6) is a close approximation to the joint behavior of aggregate TFP and markup dispersion, see Figure 4.

[14]Aggregate TFP is $\Delta \log \mathrm{TFP} = \Delta y - w_k \Delta k - (1 - w_k)\Delta \ell$, with Δy real business output growth, w_k the capital income share, Δk real capital growth (based on separate perpetual inventory methods for 15 types of capital), $\Delta \ell$ the growth of hours worked plus growth in labor composition/quality. Utilization-adjustment follows Basu et al. (2006) and uses hours per worker to proxy factor utilization. Labor productivity is real output per hour in the nonfarm business sector. Figure 6(c) in the Appendix plots the different aggregate productivity series.

Figure 3: Aggregate productivity response to monetary policy shocks

(a) Estimated productivity responses

(b) Implied productivity responses



Notes: Panel (a) shows the responses of aggregate productivity measures to a one-standard deviation contractionary monetary policy shock. Panel (b) shows the imputed response of TFP, implied by the response of markup dispersion within four-digit industry-quarters, according to $\Delta \log \text{TFP}_t = -\frac{\eta}{2} \Delta \mathbb{V}_t(\log \mu_{it})$, see equation (2.6), and using $\eta = 3$ and $\eta = 6$, respectively. Alongside, it shows the empirical response of utilization-adjusted TFP from panel (a). The shaded and bordered areas indicate one standard error bands based on the Newey–West estimator.

Using equation (2.6), we compute the implied TFP response by multiplying the response of markup dispersion by $-\frac{\eta}{2}\%$. In Figure 3(b), we show the implied response for $\eta = 6$, which is the estimate in Christiano et al. (2005), and $\eta = 3$, the assumption in Hsieh and Klenow (2009). The imputed TFP responses closely match the estimated TFP response within the first two years of the shock. This suggests that the response in markup dispersion is quantitatively important to understand the productivity effects of monetary policy.

An alternative explanation why aggregate productivity declines after monetary policy shocks is a reduction in R&D investment. In fact, Figure 8 in the Appendix shows that aggregate R&D expenditures fall after contractionary monetary policy shocks, which reconfirms the findings in Moran and Queralto (2018) and Garga and Singh (2019). Hence, there is scope for R&D to explain part of the aggregate TFP response. What is less clear is how much of the short-run productivity response can be explained by R&D investment. The evidence on technology adoption suggests that R&D has rather medium-run than short-run productivity effects. For example, Comin and Hobijn (2010) estimates an average adoption lag of 5 years. A sluggish effect of R&D investment on aggregate productivity is consistent with the finding in Figure 3(b) that markup dispersion accounts for a relatively small fraction of the TFP response 3–4 years after a monetary policy shock.

2.4 Robustness

Markup estimation. Our baseline specification assumes that firms in the same two-digit industry-quarter have a common output elasticity with respect to labor and materials. Based on De Loecker et al. (2020), we consider two alternatives. First, we estimate a translog production function with four-digit industry-specific coefficients. This gives rise to firm- and time-specific output elasticities. Second, we estimate the output elasticity through the cost share (costs of goods sold divided by total costs) at the four-digit-industry-quarter level. This is a valid estimator of the output elasticity under flexible adjustment of all input factors. All our results are robust to computing markups based on a translog production function or cost shares. Figure 11 (a) and (b) in the Appendix shows the response of markup dispersion to monetary policy shocks under the translog and cost share approach. Figure 12 shows the relative markup response of firms with stickier prices to monetary policy shocks. Table 6 shows the correlation between average markup and price rigidity. In addition, based on Traina (2020) and Basu (2019) we consider as expenses for labor and materials the costs of goods sold plus selling, general, and administrative expenses when estimating markups. Panel (c) of Figure 11 shows that markup dispersion still increases.

Firm-level data treatment. We examine the robustness of our results when tightening or relaxing our baseline data treatment. First, we keep firms with real sales growth above 100% or below -67%. Second, we keep small firms with real quarterly sales below 1 million 2012 USD. Third, instead of dropping the top/bottom 5% of the markup distribution per quarter, we drop the top/bottom 1%. Fourth, we condition on firms with at least 16 quarters of consecutive observations. Figure 13 shows that markup dispersion robustly increases after contractionary monetary policy shocks. Figure 14 shows that the relative markup response to monetary policy shocks is sensitive to removing outliers in the firm-level markups, but robust to other data treatments. Table 7 in the Appendix shows that the correlation between markups and price rigidity is robust across data treatments. A well-known recent trend is the delisting of public firms. We address the concern that this may affect our results in two ways. First, when only considering firms that are in the sample for at least 16 consecutive quarters, we find our results to be robust, as discussed above. Second, we estimate whether the number of firms in the sample responds to monetary policy shocks. Figure 7(b) shows that the response is insignificant and small.

Monetary policy shocks. We show that our results are robust to a variety of alternative monetary policy shock series. Similar to Nakamura and Steinsson (2018), we consider the first principal component of the current/three-month federal funds futures and the 2/3/4-quarters ahead Eurodollar futures. We further address the concern that high-frequency future price changes may not only reflect monetary policy shocks, but also the release of private central bank information about the state of the economy. We apply two alternative strategies to control for such information shocks. First, following Miranda-Agrippino and Ricco (2018), we regress daily monetary policy shocks on internal Greenbook forecasts and revisions for output growth, inflation, and unemployment. Second, following Jarocinski and Karadi (2020), we discard daily monetary policy shocks if the associated high-frequency change in the S&P500 moves in the same direction. We further consider a shock series including unscheduled meetings and conference calls. A different concern may be that unconventional monetary policy drives our result. We address this by setting daily monetary policy shocks at Quantitative Easing (QE) announcements to zero. Figure 15 in the Appendix shows the response of markup dispersion for all monetary policy shock series. Figure 16 in the Appendix shows the robustness of the interaction of firm-level price rigidity with the monetary policy shock for all monetary policy shock series. Figure 17 in the Appendix shows the responses of aggregate productivity for all monetary policy shock series.

LP-IV. We revisit our main results with the LP-IV method (Stock and Watson, 2018). More precisely, we replace the monetary policy shocks $\varepsilon_t^{\mathrm{MP}}$ in equations (2.4) and (2.5) by the quarterly change in the one-year treasury rate and use $\varepsilon_t^{\mathrm{MP}}$ as an instrument. Figure 18 in the Appendix shows that our results are robust to the LP-IV method.

Great Recession. We exclude the apex of the Great Recession from 2008Q3 to 2009Q2 in our baseline estimations. However, our results do not depend on this choice. Moreover, the results are robust to using the Pre-Great Recession period until 2008Q2. Figure 13 and panels (d) and (e) in Figure 14 in the Appendix show that the firm-level heterogeneity in the markup response and the increase in markup dispersion, respectively, after contractionary monetary policy shocks is robust across samples.

TFP measurement. Hall (1986) shows that the Solow residual is misspecified in the presence of market power. He shows that instead of the capital income share w_{kt} as the Solow weight for capital and $1-w_{kt}$ for labor, the correct weights are $\mu_t w_{kt}$ and $1-\mu_t w_{kt}$, where μ_t is the aggregate markup. We therefore examine the response of markup-corrected (utilization-adjusted) aggregate TFP to monetary policy shocks. We use the average markup series from De Loecker et al. (2020) to compute Hall's weights. Figure 9(c) in the Appendix shows that this barely affects our results. We further investigate whether measurement error in quarterly TFP data is responsible for the effect of monetary policy. This problem was flagged for defense spending shocks by Zeev and Pappa (2015). We follow them in re-computing TFP using measurement error corrected quarterly GDP from Aruoba et al. (2016). Figure 9(d) shows that measurement error corrected TFP also falls after monetary policy shocks. In addition, we show that Fernald's (2014) investment-specific and consumption-specific aggregate TFP series significantly falls after contractionary monetary policy shocks, see Figure 9(a) and (b). Notably, the response of investment-specific TFP is significantly stronger than that of consumption-specific TFP.

3 Analytical results

In this section, we characterize a novel monetary transmission mechanism. Monetary policy shocks that lower marginal costs increase markup dispersion if firms with lower pass-through have higher markups. Such a negative correlation between markup and pass-through can arise from firm heterogeneity in price-setting frictions.

3.1 Markup dispersion and the pass-through–markup correlation

Let i index a firm and t time. A firm's markup is $\mu_{it} \equiv P_{it}/(P_t X_t)$, where P_{it} is the firm's price, P_t the aggregate price, and X_t real marginal cost. Define pass-through from marginal cost to price as

$$\rho_{it} \equiv \frac{\partial \log P_{it}}{\partial \log X_t}. \tag{3.1}$$

This is the percentage price change in response to a percentage change in real marginal cost (without conditioning on price adjustment). The correlation between firm-level markup and firm-level pass-through is a key moment for the response of markup dispersion to shocks.

Proposition 1. *If* $\mathrm{Corr}_t(\rho_{it}, \log \mu_{it}) < 0$*, markup dispersion decreases in real marginal costs*

$$\frac{\partial \mathbb{V}_t(\log \mu_{it})}{\partial \log X_t} < 0,$$

and markup dispersion increases if $\mathrm{Corr}_t(\rho_{it}, \log \mu_{it}) > 0$.

Proof: See Appendix D.2.

Contractionary monetary policy shocks that lower real marginal costs increase the dispersion of markups if firms with higher markups have lower pass-through.

3.2 Explaining a negative pass-through-markup correlation

We next show that firm-level heterogeneity in various price-setting frictions can explain a negative correlation between firm-level pass-through and markup.

Consider a risk-neutral investor that sets prices in a monopolistically competitive environment with an isoelastic demand curve[15] and subject to adjustment costs:

$$\max_{\{P_{it+j}\}_{j=0}^{\infty}} \mathbb{E}_t \sum_{j=0}^{\infty} \beta^t \left[\left(\frac{P_{it+j}}{P_{t+j}} - X_{t+j} \right) \left(\frac{P_{it+j}}{P_{t+j}} \right)^{-\eta} Y_{t+j} - \text{adjustment cost}_{it+j} \right] \tag{3.2}$$

Adjustment costs differ across firms and may be deterministic or stochastic. This formulation nests the Calvo (1983) random adjustment, Taylor (1979) staggered price setting, Rotemberg (1982) convex adjustment costs, and Barro (1972) menu costs.

Importantly, the period profit (net of adjustment costs) is asymmetric in the price P_{it} and hence in the markup μ_{it}. Profits fall more rapidly for low markups than for high markups. This gives rise to a precautionary price setting motive: when price adjustment is frictional, firms have an incentive to set a markup above the frictionless optimal markup. Setting a higher markup today provides some insurance against low profits before the next price adjustment (Calvo/Taylor), or lowers the expected costs of future price re-adjustments (Rotemberg/Barro).

To characterize precautionary price setting, we study the problem in partial equilibrium. Analytically solving the non-linear price-setting problem with adjustment costs and aggregate uncertainty in general equilibrium is not feasible. We assume that aggregate price, real marginal costs, and aggregate demand, denoted by (P_t, X_t, Y_t), follow an i.i.d. joint log-normal process around the unconditional means $\bar{P}$, $\bar{X}$, and $\bar{Y}$. The (co-)variances of innovations are σ_k^2 and σ_{kl} for $k, l \in \{p, x, y\}$.

Calvo friction. Consider a Calvo (1983) friction, parametrized by a *firm-specific* price adjustment probability $1 - \theta_i \in (0, 1)$. The profit-maximizing reset price is

$$P_{it}^* = \frac{\eta}{\eta - 1} P_t X_t \frac{\mathbb{E}_t \left[\sum_{j=0}^{\infty} \beta^j \theta_i^j \frac{X_{t+j}}{X_t} \left(\frac{P_{t+j}}{P_t} \right)^{\eta} \frac{Y_{t+j}}{Y_t} \right]}{\mathbb{E}_t \left[\sum_{j=0}^{\infty} \beta^j \theta_i^j \left(\frac{P_{t+j}}{P_t} \right)^{\eta-1} \frac{Y_{t+j}}{Y_t} \right]}, \tag{3.3}$$

and the associated markup is μ_{it}^*. To isolate the role of uncertainty in price setting, we focus on the stochastic steady state, described by the unconditional means $(\bar{P}, \bar{X}, \bar{Y})$. The

[15] An isoelastic demand curve can be derived from a Dixit-Stiglitz aggregator. An alternative is a Kimball (1995) aggregator, which implies that the demand elasticity changes in the firm's relative price. The evidence for Kimball-type demand curves is mixed, however, see Klenow and Willis (2016).

following proposition characterizes the upward price-setting bias as a function of θ_i and establishes a condition under which firms with lower pass-through set higher markups.

Proposition 2. *If $P_t = \bar{P}$, $X_t = \bar{X}$, $Y_t = \bar{Y}$, and $(\eta - 1)\sigma_p^2 + \sigma_{py} + \eta\sigma_{px} + \sigma_{xy} > 0$, the firm sets a markup above the frictionless optimal one and the markup further increases the less likely price re-adjustment is,*

$$\mu_{it}^* > \frac{\eta}{\eta - 1} \quad \textit{and} \quad \frac{\partial \mu_{it}^*}{\partial \theta_i} > 0.$$

Pass-through ρ_{it} is zero with probability θ_i and positive otherwise. Expected pass-through, denoted by $\bar{\rho}_{it}$, of either a transitory or permanent change in X_t, falls monotonically in θ_i,

$$\frac{\partial \bar{\rho}_{it}}{\partial \theta_i} < 0.$$

If the above conditions are satisfied, then $\mathrm{Corr}_t(\rho_{it}, \log \mu_{it}^*) < 0$.

Proof: See Appendix D.3.

A permanent decrease in real marginal costs leads to an permanent increase in the optimal reset price by the same factor. The pass-through is hence one for adjusting firms and zero for non-adjusting firms. A transitory decrease in real marginal costs increases the optimal reset price by less than the marginal cost change if the future reset probability is below one. The pass-through of adjusting firms is below one, and falling in price stickiness.

Staggered price setting. Consider Taylor (1979) staggered price setting and assume that firms adjust asynchronously and at different deterministic frequencies. Staggered price setting is a deterministic variant of the Calvo setup and yields very similar results.

Rotemberg friction. Consider the price setting problem subject to Rotemberg (1982) quadratic price adjustment costs, parametrized by a *firm-specific* cost shifter $\phi_i \geq 0$, i.e., adjustment $\mathrm{cost}_{it} = \frac{\phi_i}{2}\left(\frac{P_{it}}{P_{it-1}} - 1\right)^2$. The first-order condition for P_{it} is

$$[(1-\eta) + \eta X_t]\left(\frac{P_{it}}{P_t}\right)^{-\eta} Y_t = \phi_i \left(\frac{P_{it}}{P_{it-1}} - 1\right)\frac{P_{it}}{P_{it-1}} - \phi_i \mathbb{E}_t\left[\left(\frac{P_{it+1}}{P_{it}} - 1\right)\frac{P_{it+1}}{P_{it}}\right]. \quad (3.4)$$

The following proposition summarizes our analytical results.

Proposition 3. *If* $P_{t-1} = P_t = \bar{P}$, $X_t = \bar{X}$, $Y_t = \bar{Y}$, *and* $\frac{\sigma_{px}}{\sigma_p \sigma_x} > -1$, *then up to a first-order approximation of* (3.4) *around* $\phi_i = 0$, *it holds that*

$$\mu_{it} \geq \frac{\eta}{\eta - 1} \quad \text{and} \quad \frac{\partial \mu_{it}}{\partial \phi_i} \geq 0, \quad \text{with strict inequality if } \phi_i > 0.$$

If in addition $\eta \in (1, \tilde{\eta})$, *where* $\tilde{\eta} = 1 + (\exp\{\frac{3}{2}\sigma_p^2 + \frac{3}{2}\sigma_x^2 + 4\sigma_{px}\} - \exp\{\sigma_{px}\})^{-1}$, *the pass-through, of either a transitory or permanent change in* X_t, *falls monotonically in* ϕ_i,

$$\frac{\partial \rho_{it}}{\partial \phi_i} < 0.$$

If the above conditions are satisfied, then $\mathrm{Corr}_t(\rho_{it}, \log \mu_{it}) < 0$.

Proof: See Appendix D.4.

Menu costs. Consider the price-setting problem subject to firm-specific menu costs. Due to the asymmetry of the profit function, price adjustment is more rapidly triggered for markups below the frictionless optimal markup than above. Thus, a higher reset markup may be optimal to economize on adjustment costs. Analytical results, however, are not available for the fully non-linear menu cost problem. Instead, we investigate this problem quantitatively. We find that markups increase in menu costs, consistent with precautionary price setting. Consequently, the correlation between pass-through and markup is negative. More details on calibration, solution, and results are provided in Appendix E.

4 Quantitative results

In this section, we investigate the quantitative implications of heterogeneous price rigidity in a calibrated New Keynesian model. The model explains almost two thirds of the peak response in utilization-adjusted TFP to monetary policy shocks.

4.1 Model setup

Our model setup builds on Carvalho (2006) and Gorodnichenko and Weber (2016). We discuss the model only briefly and relegate a formal description to Appendix F. An infinitely-lived representative household has additively separable preference in consumption and leisure, and discounts future utility by β. The intertemporal elasticity of substitution for consumption is γ and the Frisch elasticity of labor supply is φ. The consumption good is a Dixit–

Stiglitz aggregate of differentiated goods with constant elasticity of substitution η. In contrast to Carvalho (2006) and the subsequent literature which consider models with cross-sector differences in price rigidity, our model is a one-sector economy, in which price rigidity differs between firms. This speaks more directly to our empirical within-industry evidence. In addition, it seems more plausible to assume equal demand elasticities within than across sectors. There is a continuum of monopolistically competitive intermediate goods firms that produce with a linear technology in labor. Firms can reset their prices with a firm-specific probability $1 - \theta_i$ in any given period. They set prices to maximize the value of the firm to the households. The monetary authority aims to stabilize inflation and the output gap. The output gap is defined as deviations of aggregate output from its natural level, defined as the flexible-price equilibrium output. Monetary policy follows a Taylor rule with interest rate smoothing and is subject to monetary policy shocks, $\nu_t \sim \mathcal{N}(0, \sigma_\nu^2)$.

We expect that some modeling choices dampen the TFP channel of monetary policy, while other choices amplify it. On the one hand, we assume a time-dependent price setting friction. On the other hand, we abstract from input-output networks and real rigidities.

4.2 Calibration and solution

A model period is a quarter. We set the elasticity of substitution between differentiated goods at $\eta = 6$, as estimated in Christiano et al. (2005). This is conservative when compared to $\eta = 21$ in Fernandez-Villaverde et al. (2015), who study precautionary price setting as transmission of uncertainty shocks. A higher η means more curvature in the profit function, hence more precautionary price setting, and larger TFP losses from markup dispersion. We use standard values for the discount factor β and the intertemporal elasticity of substitution γ. We set the former to match an annual real interest rate of 3%, and the latter to a value of 2. We use the estimates in Christiano et al. (2016) for the Taylor rule and set $\rho_r = 0.85$, $\phi_\pi = 1.5$, and $\phi_y = 0.05$.

The parameters which play a key role in this model are the price adjustment frequencies. We assume that there are five equally large groups of firms, indexed by $k \in \{1, \ldots, 5\}$, which differ in their price adjustment frequencies $1 - \theta_k$. We calibrate $\{\theta_k\}$ to match the empirical distribution of within-industry price adjustment frequencies based on Gorodnichenko and Weber (2016). They document mean and standard deviation of monthly price adjustment frequencies for five sectors. We first compute the value-added-weighted average of the means and variances. The monthly mean price adjustment frequency is 0.1315 and the standard deviation is 0.1131. Second, we fit a log-normal distribution to these moments. Third, we compute the mean frequencies within the five quintile groups of the fitted distribution.

Table 2: Calibration

Parameter		Value	Source/Target
Discount factor	β	$1.03^{-1/4}$	Risk-free rate of 3%
Elasticity of intertemporal substitution	γ	2	Standard
Elasticity of substitution between goods	η	6	Christiano et al. (2005)
Interest rate smoothing	ρ_r	0.85	Christiano et al. (2016)
Policy reaction to inflation	ϕ_π	1.5	Christiano et al. (2016)
Policy reaction to output	ϕ_y	0.05	Christiano et al. (2016)
Standard deviation of MP shock	σ_ν	0.00415	30bp effect on nominal rate
Frisch elasticity of labor supply	φ	0.1175	Relative hours response of 11.7%
Distribution of price adjustment frequencies			
Firm type k		Share	Price adjustment frequency $1-\theta_k$
1		0.2	0.0231
2		0.2	0.0678
3		0.2	0.1396
4		0.2	0.2829
5		0.2	0.8470

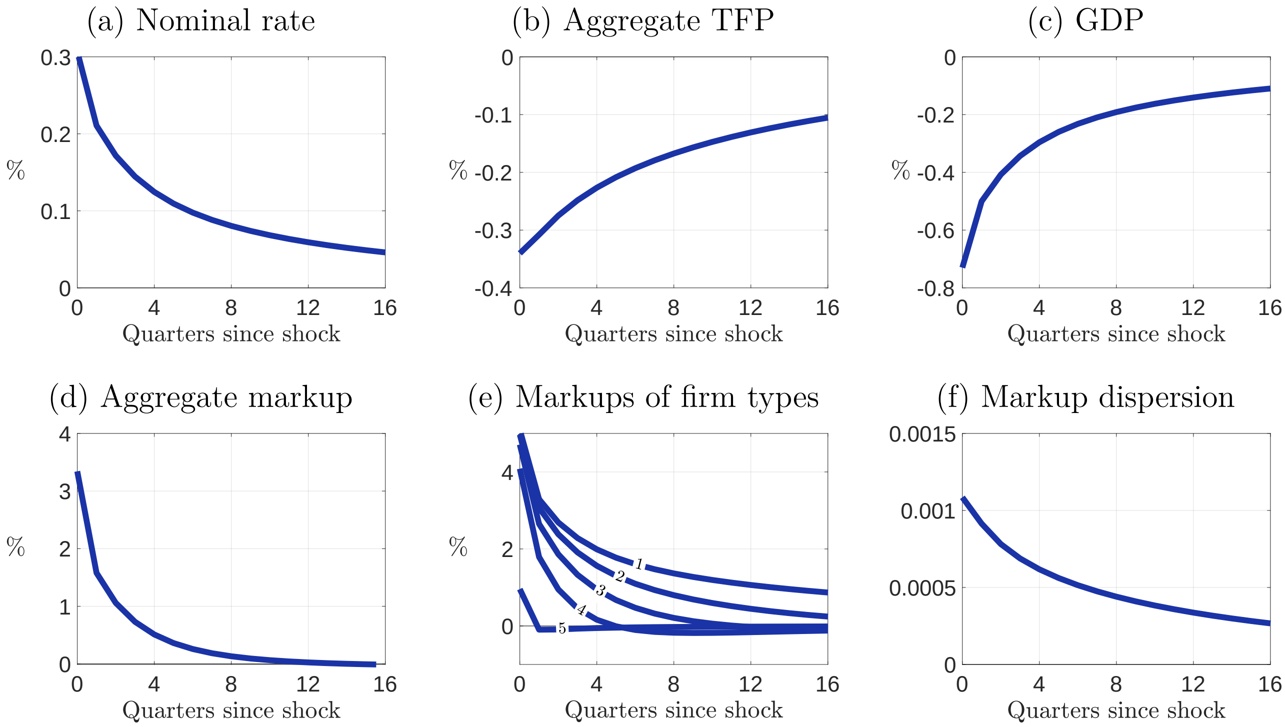
Notes: The distribution of price adjustment frequencies is chosen to match the within-sector distribution reported in Gorodnichenko and Weber (2016).

Finally, we transform the monthly frequencies into quarterly ones to obtain $\{\theta_k\}$.

We calibrate the Frisch elasticity of labor supply internally. The hours response to monetary policy shocks is small on impact, but larger at longer horizons, see Figure 10 in the Appendix. The utilization-adjusted TFP response is immediately negative but has a flatter profile at longer horizons. On average, the two responses have similar magnitude. The average difference of the hours response relative to the response of utilization-adjusted TFP, computed as the mean of $\frac{1\ -\ \text{response of hours in \%}}{1\ -\ \text{response of util-adj. TFP in \%}} - 1$ up to 16 quarters after the shock, is 11.7%. In the model, we compute the relative hours response in the same way and target 11.7% to calibrate the Frisch elasticity. Importantly, we do not directly target the absolute magnitude of the TFP response, but only a relative quantity. The calibrated Frisch elasticity is $\varphi = 0.1175$, which is low compared to the macroeconomics literature, but which is within the range of empirical estimates surveyed by Ashenfelter et al. (2010). The remaining parameter is the standard deviation of monetary policy shocks σ_ν, which we also calibrate internally. The target is the peak nominal interest rate response to a one standard deviation monetary policy shock of 30bp, see Figure 10. This yields $\sigma_\nu = 0.00415$.

For markup dispersion to arise from precautionary price setting, it is important to use an adequate model solution technique. We rely on local solution techniques, but, importantly, solve the model around its stochastic steady state. Whereas markup are the same across

Figure 4: Model responses to monetary policy shocks



Notes: This figure shows responses to a one standard deviation contractionary monetary policy shock. In panel (e), the responses are the average markup responses of the firm types $k = 1, \dots, 5$, where $k = 1$ is the stickiest and $k = 5$ the most flexible type of firms.

firms in the deterministic steady state, differences across firms may exist in the stochastic steady state. We apply the method developed by Meyer-Gohde (2014), which uses a third-order perturbation around the deterministic steady state to compute the stochastic steady state as well as a first-order approximation of the model dynamics around it.[16] In the stochastic steady state, precautionary price setting has large effects. Firms with the most rigid prices have 10.9% higher markups than firms with the most flexible prices. As follows from Proposition 1, the negative correlation between markups and pass-through implies that contractionary monetary policy shocks increase markup dispersion and lower aggregate TFP.

4.3 Results

Figure 4 shows the responses to a one-standard deviation monetary policy shock. The shock depresses aggregate demand and lowers real marginal costs. In response, firms want to

[16] At an earlier stage of this paper, we have also solved the model globally using a time iteration algorithm for the case of two firm types with one of them having perfectly flexible prices. This yields very similar quantitative results compared to using the Meyer-Gohde (2014) algorithm. However, the computational costs of time iteration are exceedingly large for more general setup of heterogeneous price rigidities.

lower their prices. For firms with stickier prices, however, pass-through is lower and on average their markups increase by more. Since firms with stickier prices have higher initial markups, markup dispersion increases. This worsens the allocation of factors across firms and thereby depresses aggregate TFP. The mechanism is quantitatively important. The increase in markup dispersion is about 75% of the peak empirical response, see Figure 1, and the model explains 60% of the peak empirical response in utilization-adjusted TFP, see Figure 3. In addition, the responses show the frequency composition effect described by Carvalho (2006). The firms with flexible prices are quick to adjust. Hence, at longer horizons, the distribution of firms with non-adjusted prices is dominated by the stickier type of firms. This generates additional persistence in the responses.

An important aspect of the monetary transmission channel is the response of aggregate TFP. In contrast, traditional business cycle models assume that fluctuations in aggregate TFP are solely driven by exogenous technology shocks.This motivates us to examine the success of a Taylor rule in stabilizing output if the monetary authority in the model (mis-)perceives the aggregate TFP response to demand shocks as originating from technology shocks. We compute a counterfactual change in natural output supposing the TFP response to monetary policy shocks is driven by technology shocks.[17] Panel (a) in Figure 5 shows the difference between the GDP responses under the counterfactual technology shock and the baseline response.[18] Output drops by up to 0.17 percentage points more if the monetary authority attributes aggregate TFP fluctuations to technology shocks, and the response is markedly more persistent. This finding highlights the importance for the monetary authority to assess the sources of observed aggregate TFP fluctuations.

Panel (b) in Figure 5 shows the response of markup dispersion to a negative technology shock with the size and persistence that matches the endogenous response of TFP to a monetary policy shock.[19] The behavior of markup dispersion helps to discriminate between productivity and monetary policy shocks. It increases after contractionary monetary policy shocks but decreases after contractionary productivity shocks.

The fact that aggregate TFP responds to monetary policy shocks can change the sign of the (aggregate) markup response to monetary policy shocks. This relates to a recent debate. While monetary policy shocks raise markups in a large class of New Keynesian models, recent evidence in Nekarda and Ramey (2019) points in the opposite direction. Following

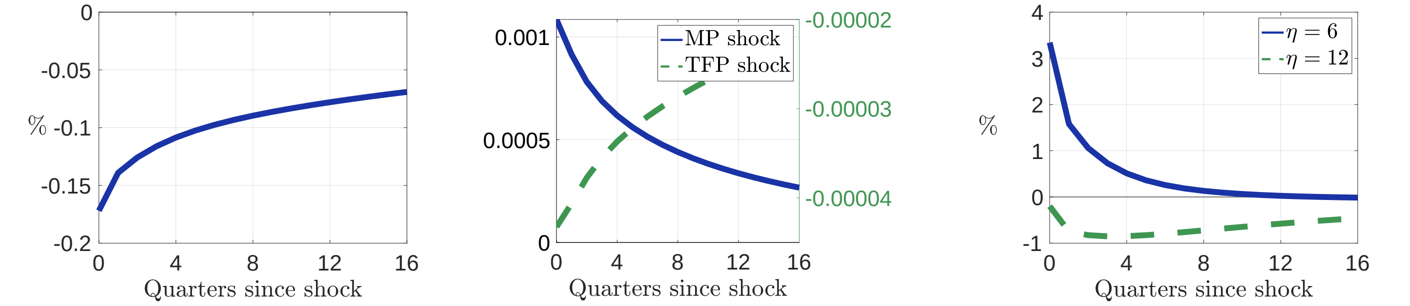
[17]We recalibrate σ_ν to ensure the same interest rate response to a one standard deviation monetary policy shock, but keep all other parameters unchanged.

[18]Figure 20 in the Appendix provides further impulse responses for this counterfactual scenario.

[19]Figure 21 in the Appendix provides further impulse responses for the technology shock.

Figure 5: Additional model results and counterfactuals

(a) Differential GDP response, counterfactual technology shock | (b) Markup dispersion response to technology shock | (c) Aggregate markup response, low vs. high substitution elasticity



Notes: Panel (a) shows the difference between the response to a monetary policy shock in the baseline model and the same model using a Taylor rule in which the output gap is computed by counterfactually assuming the TFP responses are driven by technology shocks. Panel (b) compares the response of markup dispersion to a monetary policy shock (left y-axis) with a technology shock (right y-axis). Panel (c) compares the response of the aggregate markup to a monetary policy shock for two values of the elasticity of substitution between differentiated goods.

Hall (1988), the aggregate markup in our model is

$$\mu_t = \frac{\mathrm{TFP}_t}{W_t/P_t}, \tag{4.1}$$

where W_t/P_t denotes the real wage. In standard New Keynesian models, tighter monetary policy reduces aggregate demand which lowers real marginal costs and, hence, markups increase. In contrast, equation (4.1) shows that the aggregate markup falls if aggregate TFP falls sufficiently strongly in response to tighter monetary policy. This argument extends to sectoral and even firm-level markups, if monetary policy shocks affect TFP at more disaggregated levels. In general equilibrium, an endogenous decline in aggregate TFP will feed back into real marginal costs, which also affects markups.

Panel (c) in Figure 5 shows the aggregate markup response to monetary policy shocks. In our baseline calibration with an elasticity of substitution $\eta = 6$ the aggregate markup raises. In some sense, that is because aggregate TFP does not fall strongly enough. We next compare our baseline results with the results when doubling the elasticity to $\eta = 12$. A larger η increases the misallocation costs of markup dispersion and thus the TFP loss after a monetary policy shock. For $\eta = 12$, the aggregate TFP response is almost twice as large, see Figure 22 in the Appendix. This is sufficient to explain lower aggregate markups after monetary policy shocks. Dynamically, the TFP loss leads to an increase in hours worked, which additionally increases marginal costs and lowers firm-level markups, reinforcing the effect on the aggregate markup.

In the Appendix, we analyze the robustness of our results in two dimensions. First, in Figure 23, we vary the Frisch elasticity of labor supply φ. The markup dispersion and TFP responses are higher for less elastic labor supply and dampened for more elastic labor supply. Second, in Figure 24, we raise the lowest price adjustment frequency, $1-\theta_1$ to the level of the second quintile group $1-\theta_2 = 6.78\%$. This dampens the increase in markup dispersion, and yields an aggregate TFP response of -0.27%.

5 Conclusion

This paper studies a novel monetary transmission mechanism. Monetary policy shocks increase the dispersion of markups across firms if firms with stickier prices have higher pre-shock markups. Increased markup dispersion implies a change in the allocation of inputs across firms, which lowers measured aggregate TFP. Using aggregate and firm-level data, we document three new facts, which are consistent with this mechanism. First, firms that adjust prices less frequently have higher markups. Second, monetary policy shocks increase the relative markup of firms with stickier prices. Third, monetary policy shocks increase the markup dispersion across firms, and lower aggregate productivity. The empirically estimated magnitudes suggest that the response in markup dispersion is quantitatively important to understand the response of aggregate productivity. We show that an explanation for the negative correlation between markup and price stickiness are differences in price stickiness across firms. Firms with stickier prices optimally set higher markups for precautionary reasons. In a calibrated New Keynesian model, heterogeneous price stickiness allows us to explain a large share of the empirically estimated TFP response to monetary policy shocks.

References

Aoki, K. (2001): "Optimal monetary policy responses to relative-price changes," *Journal of Monetary Economics*, 48, 55 – 80.

Aruoba, S. B., F. X. Diebold, J. Nalewaik, F. Schorfheide, and D. Song (2016): "Improving GDP measurement: A measurement-error perspective," *Journal of Econometrics*, 191, 384–397.

Ascari, G. and A. M. Sbordone (2014): "The Macroeconomics of Trend Inflation," *Journal of Economic Literature*, 52, 679–739.

Ashenfelter, O. C., H. Farber, and M. R. Ransom (2010): "Labor market monopsony," *Journal of Labor Economics*, 28, 203–210.

Baqaee, D. R. and E. Farhi (2017): "Productivity and Misallocation in General Equilibrium," *NBER Working Paper Series*.

——— (2020): "Productivity and Misallocation in General Equilibrium," *The Quarterly Journal of Economics*, 135, 105–163.

Barro, R. (1972): "A Theory of Monopolistic Price Adjustment," *Review of Economic Studies*, 39, 17–26.

Basu, S. (1995): "Intermediate Goods and Business Cycles: Implications for Productivity and Welfare," *The American Economic Review*, 85, 512–531.

——— (2019): "Are Price-Cost Markups Rising in the United States? A Discussion of the Evidence," *Journal of Economic Perspectives*, 33, 3–22.

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

Bloom, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77, 623–685.

Bond, S., A. Hashemi, G. Kaplan, and P. Zoch (2020): "Some Unpleasant Markup Arithmetic: Production Function Elasticities and their Estimation from Production Data," mimeo, University of Oxford.

Calvo, G. (1983): "Staggered prices in a utility-maximizing framework," *Journal of Monetary Economics*, 12, 383–398.

Carvalho, C. (2006): "Heterogeneity in Price Stickiness and the Real Effects of Monetary Shocks," *The B.E. Journal of Macroeconomics*, 6, 1–58.

Carvalho, C. and F. Nechio (2016): "Factor Specificity and Real Rigidities," *Review of Economic Dynamics*, 22, 208–222.

Carvalho, C. and F. Schwartzman (2015): "Selection and monetary non-neutrality in time-dependent pricing models," *Journal of Monetary Economics*, 76, 141–156.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2011): "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, 101, 471–75.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113, 1–45.

Christiano, L. J., M. S. Eichenbaum, and M. Trabandt (2016): "Unemployment and Business Cycles," *Econometrica*, 84, 1523–1569.

Comin, D. and M. Gertler (2006): "Medium-Term Business Cycles," *American Economic Review*, 96, 523–551.

Comin, D. and B. Hobijn (2010): "An Exploration of Technology Diffusion," *American Economic Review*, 100, 2031–59.

De Loecker, J., J. Eeckhout, and G. Unger (2020): "The Rise of Market Power and the Macroeconomic Implications," *The Quarterly Journal of Economics*.

De Loecker, J. and F. Warzynski (2012): "Markups and Firm-Level Export Status," *American Economic Review*, 102, 2437–71.

Eisfeldt, A. L. and A. A. Rampini (2006): "Capital reallocation and liquidity," *Journal of Monetary Economics*, 53, 369–399.

Eusepi, S., B. Hobijn, and A. Tambalotti (2011): "CONDI: A Cost-of-Nominal-Distortions Index," *American Economic Journal: Macroeconomics*, 3, 53–91.

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

Fernandez-Villaverde, J., P. Guerron-Quintana, K. Kuester, and J. Rubio-Ramirez (2015): "Fiscal Volatility Shocks and Economic Activity," *American Economic Review*, 105, 3352–84.

Galí, J. (2015): *Monetary Policy, Inflation, and the Business Cycle: an Introduction to the New Keynesian Framework and its Applications*, Princeton University Press.

Garga, V. and S. R. Singh (2019): "Output Hysteresis and Optimal Monetary Policy," Tech. rep., University of California Davis.

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

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

Gorodnichenko, Y. and M. Weber (2016): "Are Sticky Prices Costly? Evidence from the Stock Market," *American Economic Review*, 106, 165–99.

Hall, R. (1986): "Market Structure and Macroeconomic Fluctuations," *Brookings Papers on Economic Activity*, 17, 285–338.

——— (1988): "The Relation between Price and Marginal Cost in U.S. Industry," *Journal of Political Economy*, 96, 921–47.

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

Jarocinski, M. and P. Karadi (2020): "Deconstructing monetary policy surprises: the role of information shocks," *American Economic Journal: Macroeconomics*.

Jeenas, P. (2019): "Firm Balance Sheet Liquidity, Monetary Policy Shocks, and Investment Dynamics," *mimeo*.

Jordà, Ò., S. Singh, and A. M. Taylor (2020): "The long-run effects of monetary policy," *mimeo*.

Khan, A. and J. K. Thomas (2008): "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics," *Econometrica*, 76, 395–436.

——— (2013): "Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity," *Journal of Political Economy*, 121, 1055–1107.

Kimball, M. S. (1995): "The Quantitative Analytics of the Basic Neomonetarist Model," *Journal of Money, Credit and Banking*, 27, 1241–1277.

Klenow, P. J. and J. L. Willis (2016): "Real Rigidities and Nominal Price Changes," *Economica*, 83, 443–472.

Levin, A. and T. Yun (2007): "Reconsidering the natural rate hypothesis in a New Keynesian framework," *Journal of Monetary Economics*, 54, 1344–1365.

Lucca, D. O. and E. Moench (2015): "The pre-FOMC announcement drift," *The Journal of Finance*, 70, 329–371.

Meier, M. (2020): "Supply Chain Disruptions, Time To Build, and the Business Cycle," *mimeo U Mannheim*.

Meyer-Gohde, A. (2014): "Risky linear approximations," mimeo, Goethe University Frankfurt.

Miranda-Agrippino, S. and G. Ricco (2018): "The Transmission of Monetary Policy Shocks," mimeo, University of Warwick.

Moran, P. and A. Queralto (2018): "Innovation, productivity, and monetary policy," *Journal of Monetary Economics*, 93, 24–41, carnegie-Rochester-NYU Conference on Public Policy held at the Stern School of Business at New York University.

Nakamura, E. and J. Steinsson (2010): "Monetary Non-neutrality in a Multisector Menu Cost Model*," *The Quarterly Journal of Economics*, 125, 961–1013.

——— (2018): "High Frequency Identification of Monetary Non-Neutrality: The Information Effect," *Quarterly Journal of Economics*, 133, 1283–1330.

Nakamura, E., J. Steinsson, P. Sun, and D. Villar (2018): "The Elusive Costs of Inflation: Price Dispersion during the U.S. Great Inflation*," *The Quarterly Journal of Economics*, 133, 1933–1980.

Nekarda, C. J. and V. A. Ramey (2019): "The Cyclical Behavior of the Price-Cost Markup," mimeo, University of California San Diego.

Oikawa, K. and K. Ueda (2018): "Reallocation Effects of Monetary Policy," *RIETI Discussion Paper Series 18-E-056*.

Ottonello, P. and T. Winberry (2020): "Financial Heterogeneity and the Investment Channel of Monetary Policy," *mimeo*.

Pasten, E., R. Schoenle, and M. Weber (forthcoming): "The Propagation of Monetary Policy Shocks in a Heterogeneous Production Economy," *Journal of Monetary Economics*.

Romer, D. (1990): "Staggered price setting with endogenous frequency of adjustment," *Economics Letters*, 32, 205–210.

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

Sheremirov, V. (2019): "Price dispersion and inflation: New facts and theoretical implications," *Journal of Monetary Economics*.

Stock, J. H. and M. W. Watson (2018): "Identification and estimation of dynamic causal effects in macroeconomics using external instruments," *The Economic Journal*, 128, 917–948.

Taylor, J. B. (1979): "Staggered Wage Setting in a Macro Model," *American Economic Review*, 69, 108–113.

Traina, J. (2020): "Is Aggregate Market Power Increasing? Production Trends Using Financial Statements," New working paper series, Stigler Center.

Zeev, N. B. and E. Pappa (2015): "Multipliers of unexpected increases in defense spending: An empirical investigation," *Journal of Economic Dynamics and Control*, 57, 205–226.

A Data construction and descriptive statistics

A.1 Firm-level balance sheet data

We use quarterly firm-level balance sheet data of listed US firms for the period 1995Q1 to 2017Q2 from Compustat North America. We delete duplicate firm-quarter observations. Our industry classification is based on the North American Industry Classification System (NAICS). We exclude firms in utilities (2-digit NAICS code 22), finance, insurance, and real estate (52 and 53), and public administration (99). We discard observations of sales (`saleq`), costs of goods sold (`cogsq`), and property, plant, and equipment (net PPE, `ppentq`, and gross PPE, `ppegtq`), that are non-positive. We fill one-quarter gaps in the firm-specific series of these variables by linear interpolation. All variables are deflated using the GDP deflator, except PPE, which is deflated by the investment-specific GDP deflator. We construct a measure of the capital stock of firms using the perpetual inventory method: We initialize $K_{i0} = \texttt{ppegtq}_{i0}$ and recursively compute $K_{it} = K_{it-1} + (\texttt{ppentq}_{it} - \texttt{ppentq}_{it-1})$. We drop firm-quarter observations if sales, costs of goods sold, or fixed assets are only reported once in the associated year. We further drop observations if quarterly sales growth is above 100% or below -67% or if real sales are below 1 million USD. Table 3 shows descriptive statistics for our baseline sample.

Table 3: Summary statistics for Compustat data

	mean	sd	min	max	count
Sales	632.22	3067.46	1.00	132182.15	329173
Fixed assets	987.38	5490.96	0.00	273545.97	326223
Variable costs	439.58	2317.01	0.13	104456.86	329173
Total Assets	2716.05	13374.72	0.00	559922.78	326632

Notes: Summary statistics for Compustat data. All variables are in millions of 2012Q1 US$.

A.2 Data on price rigidity

To maximize firm-level variation in price rigidity, we weight average industry-level price adjustment frequency with firms' industry sales from the Compustat segment files. Industry-level price adjustment frequency is based on Pasten et al. (forthcoming). We define the implied price duration as $-1/\log(1 - \text{price adjustment frequency})$.

We obtain firms' yearly industry sales composition using the operation segments and, if these are not available, the business segments from the Compustat segments file. We drop various types of duplicate observations: In case of exact duplicates, we keep one. In case there are different source dates or more than one accounting month per year, we keep the observation with the newest source dates or the later accounting month, respectively. We drop segment observations for firm-years if the industry code is not reported. If only some segment industry codes are missing, we assign the firm-specific industry code to the segments with missing industry code.

We then compute every firm's average price rigidity over segments weighted by sales. In case we do not observe the five-digit-industry-level price stickiness for all segments or we observe only one segment, we use the five-digit price rigidity measure associated to the firm's general five-digit industry code. Note that even in this case, there is variation across firms within four-digit industries. Our sample comprises 8,091 unique firms. For 1,891 firms (23%), we can compute a segment-based price stickiness level in some year. For firm-years with segment-based price stickiness, the mean (median) number of segments is 2.36 (2) with a standard deviation of 0.67.

A.3 Monetary policy shocks

We compute monetary policy shocks by high-frequency identification as described in Subsection 2.1. Table 4 reports summary statistics for all monetary policy shocks and Figure 6 shows the shock series.

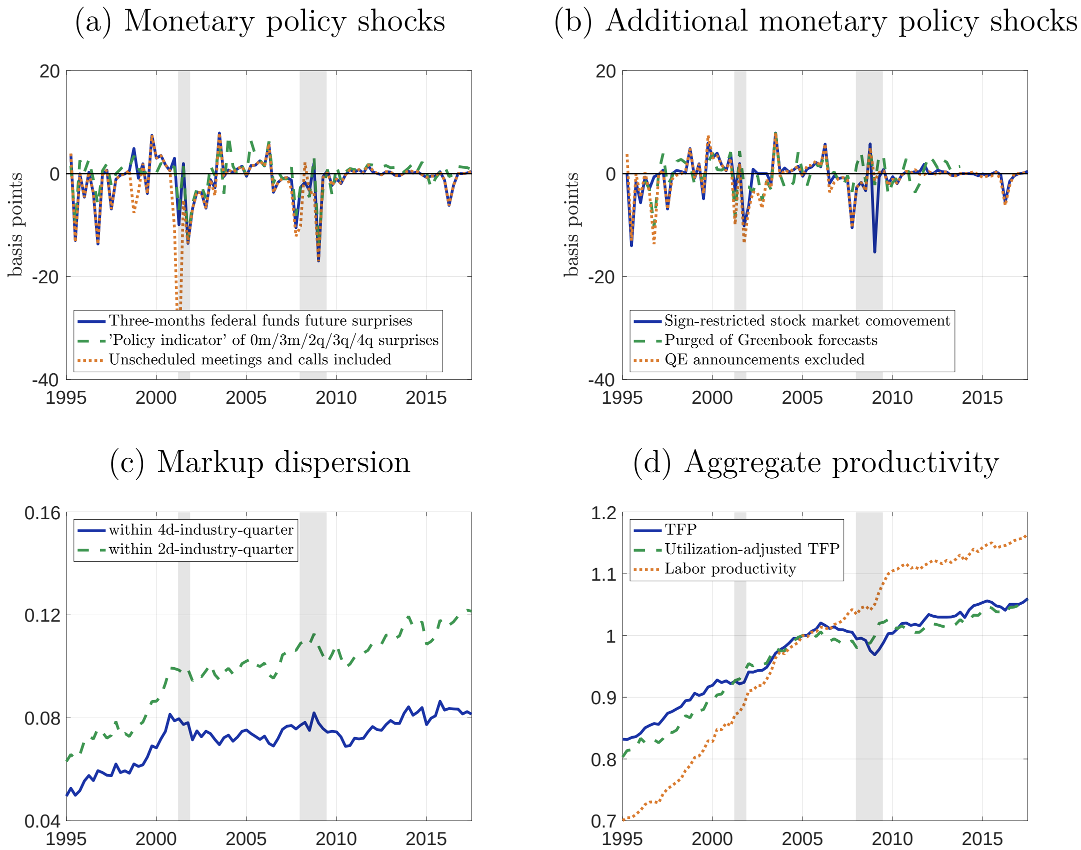
Table 4: Summary statistics of monetary policy shocks

	mean	sd	min	max	count
Three-month Fed funds future surprises	-1.00	4.06	-17.01	7.87	94
... unscheduled meetings and conference calls included	-1.84	5.70	-38.33	7.86	94
... purged of Greenbook forecasts	-0.00	3.10	-10.47	7.98	71
... sign-restricted stock market comovement	-0.52	3.47	-15.27	7.87	94
... QE announcements excluded	-0.83	3.72	-13.71	7.87	94
'Policy indicator' surprise	-0.05	3.43	-14.13	7.45	94

Notes: Summary statistics for monetary policy shocks in basis points.

A.4 Time series plots of monetary policy shocks, markup dispersion, and aggregate productivity

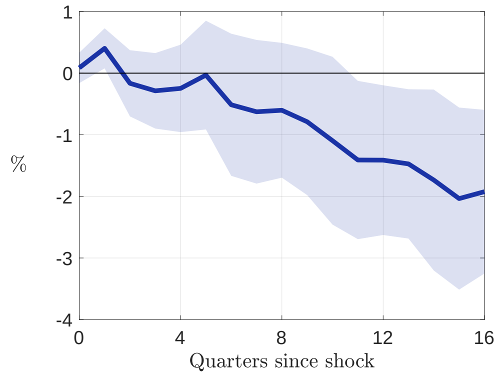
Figure 6: Monetary policy shocks, aggregate productivity, and markup dispersion



Notes: Aggregate productivity (in logs), markup dispersion, and monetary policy shocks are at quarterly frequency. Aggregate (utilization-adjusted) TFP is from Fernald (2014). Labor productivity is from FRED. Markup dispersion is computed from Compustat balance sheet data. Shaded gray areas indicate NBER recession dates.

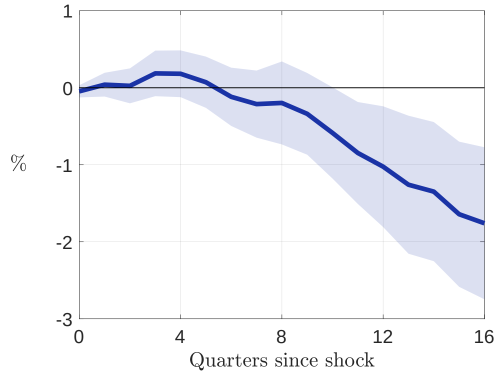
B Additional empirical results

Figure 7: Response of firm-level observations after monetary policy shocks



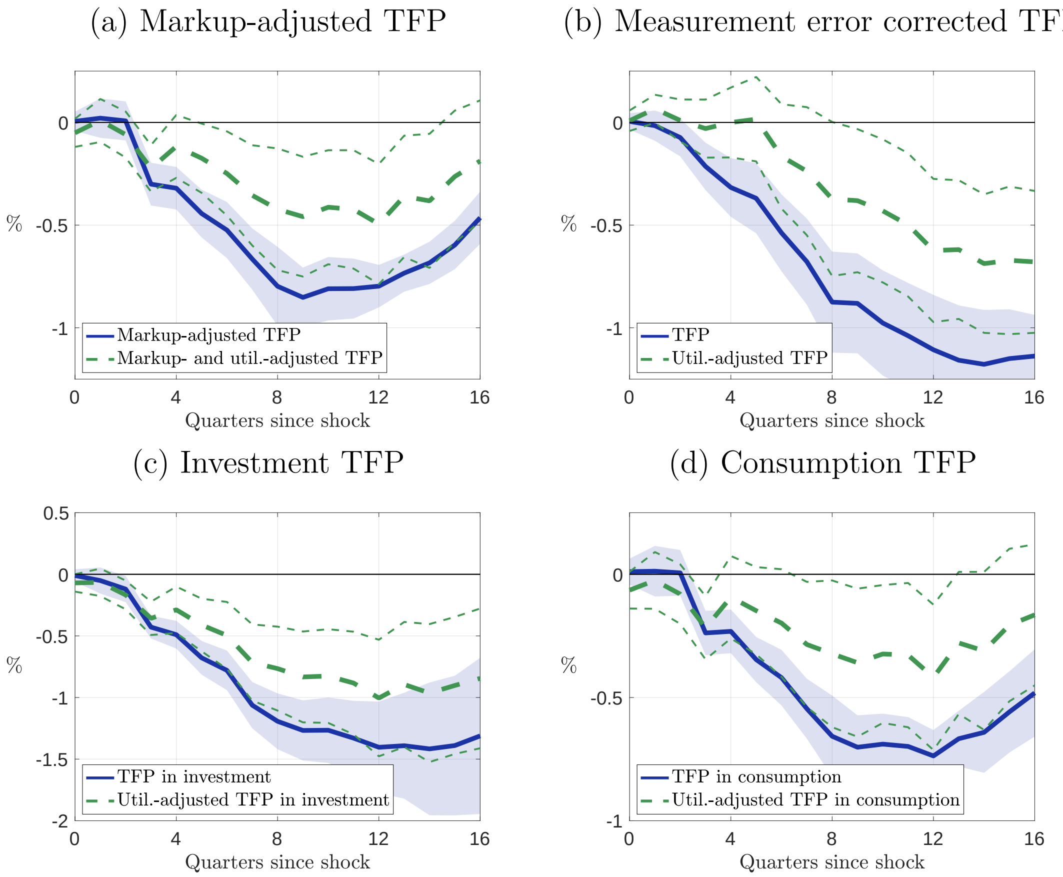
Notes: This figure shows the response of the number of firm-level observations in our sample to monetary policy shocks obtained from local projections as in equation (2.4). The shaded area is a one standard error band based on Newey–West.

Figure 8: Aggregate R&D response to monetary policy shock



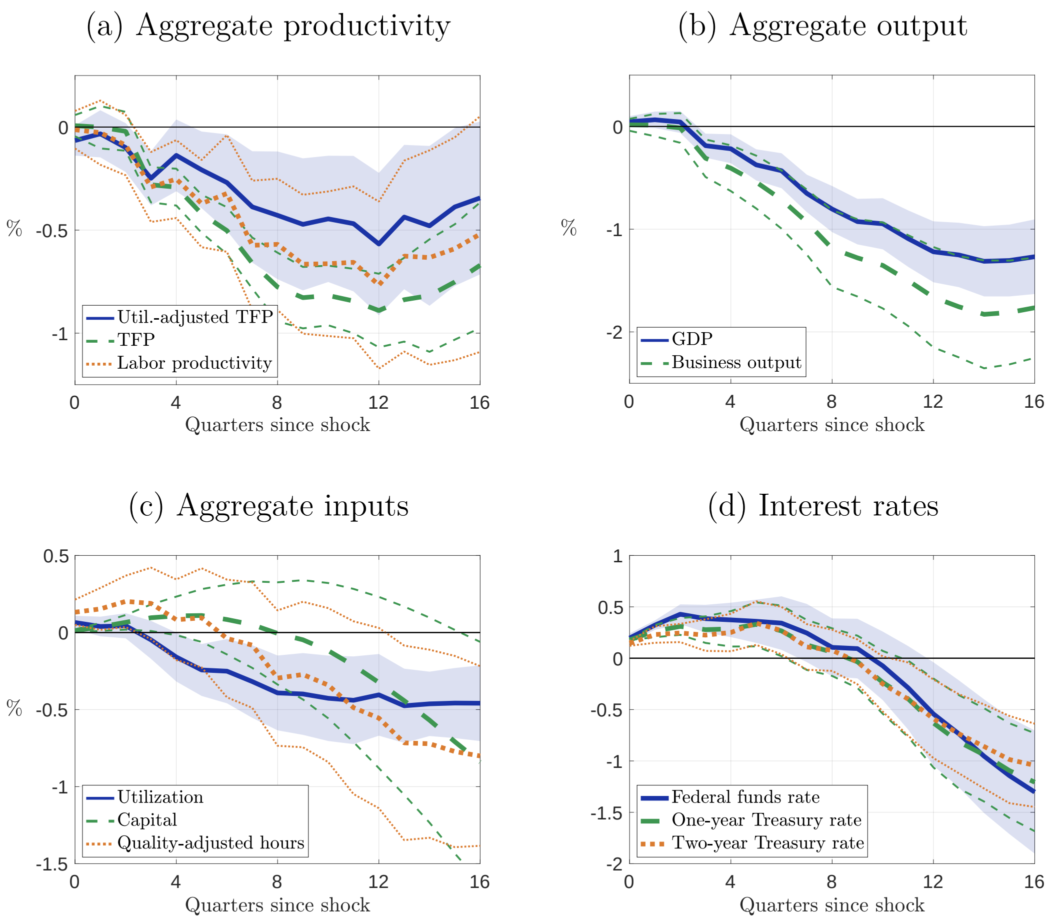
Notes: The plots show the response to a one-standard deviation contractionary monetary policy shock. The shaded area indicate one standard error bands based on the Newey–West estimator.

Figure 9: Further productivity responses



Notes: Responses to monetary policy shocks obtained from local projections as in equation (2.4). Investment-TFP and Consumption-TFP are from Fernald (2014). Markup-corrected TFP is constructed following Hall (1988) using the average markup estimated by De Loecker et al. (2020). Measurement error corrected TFP is constructed using measurement error corrected GDP from Aruoba et al. (2016), total hours from the BLS, and capital stock and output elasticities from Fernald (2014). The utilization-adjusted measure subtracts utilization from Fernald (2014). The shaded and bordered areas indicate one standard error bands based on Newey–West.

Figure 10: Macroeconomic responses to monetary policy shocks

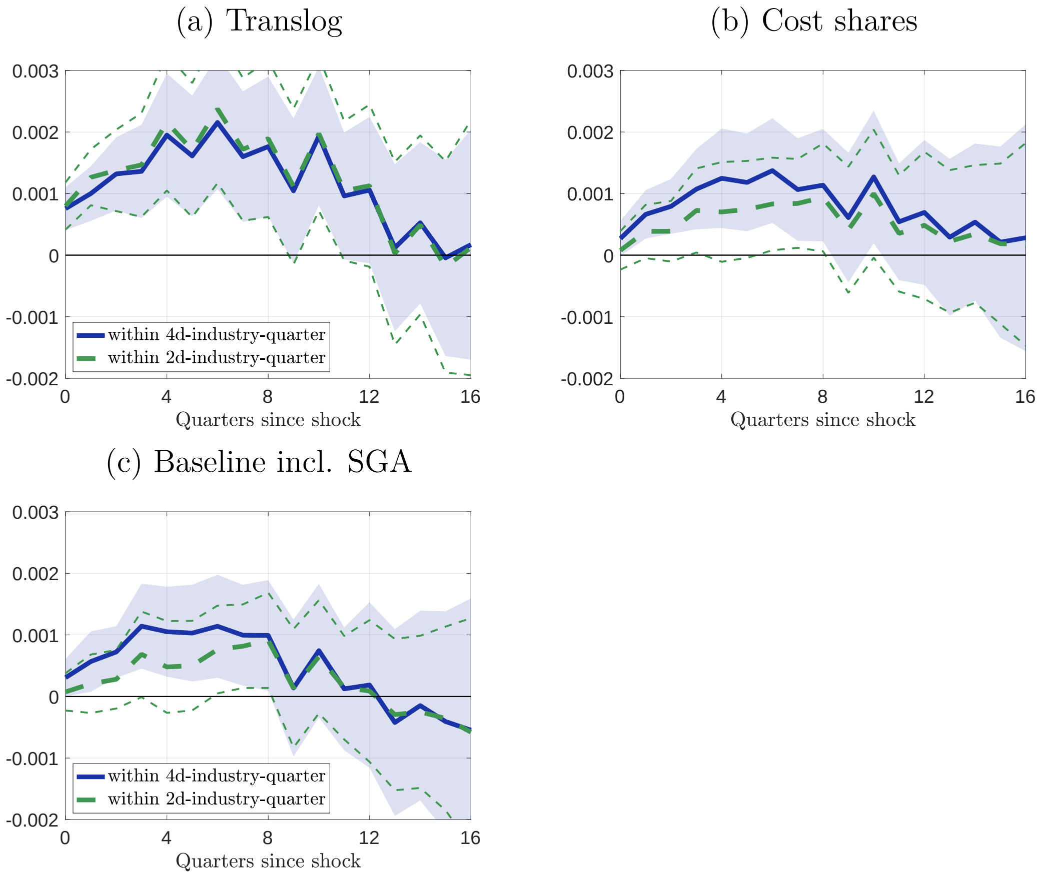


Notes: The plots show the responses to a one-standard deviation contractionary monetary policy shock. The local projections in Panel (d) are estimated in levels rather than differences. The shaded and bordered areas indicate one standard error bands based on the Newey–West estimator.

C Robustness of main findings

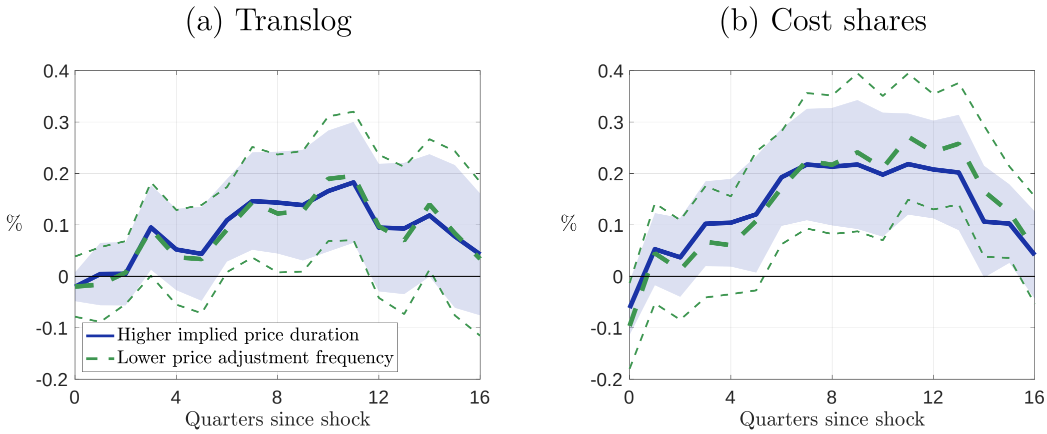
C.1 Results for alternative markup series

Figure 11: Responses of markup dispersion for alternative markup measures



Notes: Responses to monetary policy shocks obtained from local projections as in equation (2.4). Panel (a) uses markups based on an industry-specific translog production function, which gives rise firm-quarter-specific output elasticities. Panel (b) uses markups based on output elasticities estimated as the industry-quarter-specific median cost share. Panel (c) uses markups based on our baseline assumption of common output elasticities within industry-quarters, but measures labor and material expenses using costs of goods sold (COGS) plus selling, general, and administrative expenses (SGA). Markup dispersion is measured within two-digit and four-digit industry-quarters as well as without fixed effects, respectively. The shaded and bordered areas indicate one standard error bands based on Newey–West.

Figure 12: Relative markup response of firms with stickier prices for alternative markup measures



Notes: The figures show the response to a one standard deviation contractionary monetary policy shock of the firm-level markup of firms with a price adjustment frequency one standard deviation below (or with an implied price duration one standard deviation above) from panel local projections as in equation (2.5). The regressions include interactions with lagged log assets, leverage, and liquidity and their interactions with a monetary policy shock. Panel (a) uses markups based on an industry-specific translog production function, which gives rise firm-quarter-specific output elasticities. Panel (b) uses markups based on output elasticities estimated as the industry-quarter-specific median cost share. The shaded and bordered areas indicate 90% error bands clustered by firms and quarters.

Table 5: Regressions of markup on price stickiness including price adjustment frequencies above 99%

	(1)	(2)	(3)	(4)
Implied price duration	0.0435	0.0362		
	(0.0197)	(0.0175)		
Price adjustment frequency			-0.243	-0.184
			(0.143)	(0.134)
Size		0.00923		0.00896
		(0.00361)		(0.00387)
Leverage		-0.00174		-0.00170
		(0.000667)		(0.000673)
Liquidity		0.548		0.548
		(0.0590)		(0.0592)
2-digit industry FE	Yes	Yes	Yes	Yes
Observations	4027	4024	4027	4024
Adjusted R^2	0.106	0.185	0.102	0.180

Notes: Regression of firm-level markup (averaged over 2005–2011) on firm-level price adjustment frequency and implied price duration, respectively. Standard errors are clustered at the two-digit industry level and shown in parentheses.

Table 6: Regressions of markup on price stickiness for alternative markup series

(a) Markups based on translog production function

	(1)	(2)	(3)	(4)
Implied price duration	0.0272	0.0296		
	(0.0106)	(0.0102)		
Price adjustment frequency			-0.210	-0.226
			(0.0731)	(0.0731)
Additional controls	No	Yes	No	Yes
2-digit industry FE	Yes	Yes	Yes	Yes
Observations	3830	3827	3830	3827
Adjusted R^2	0.067	0.123	0.070	0.126

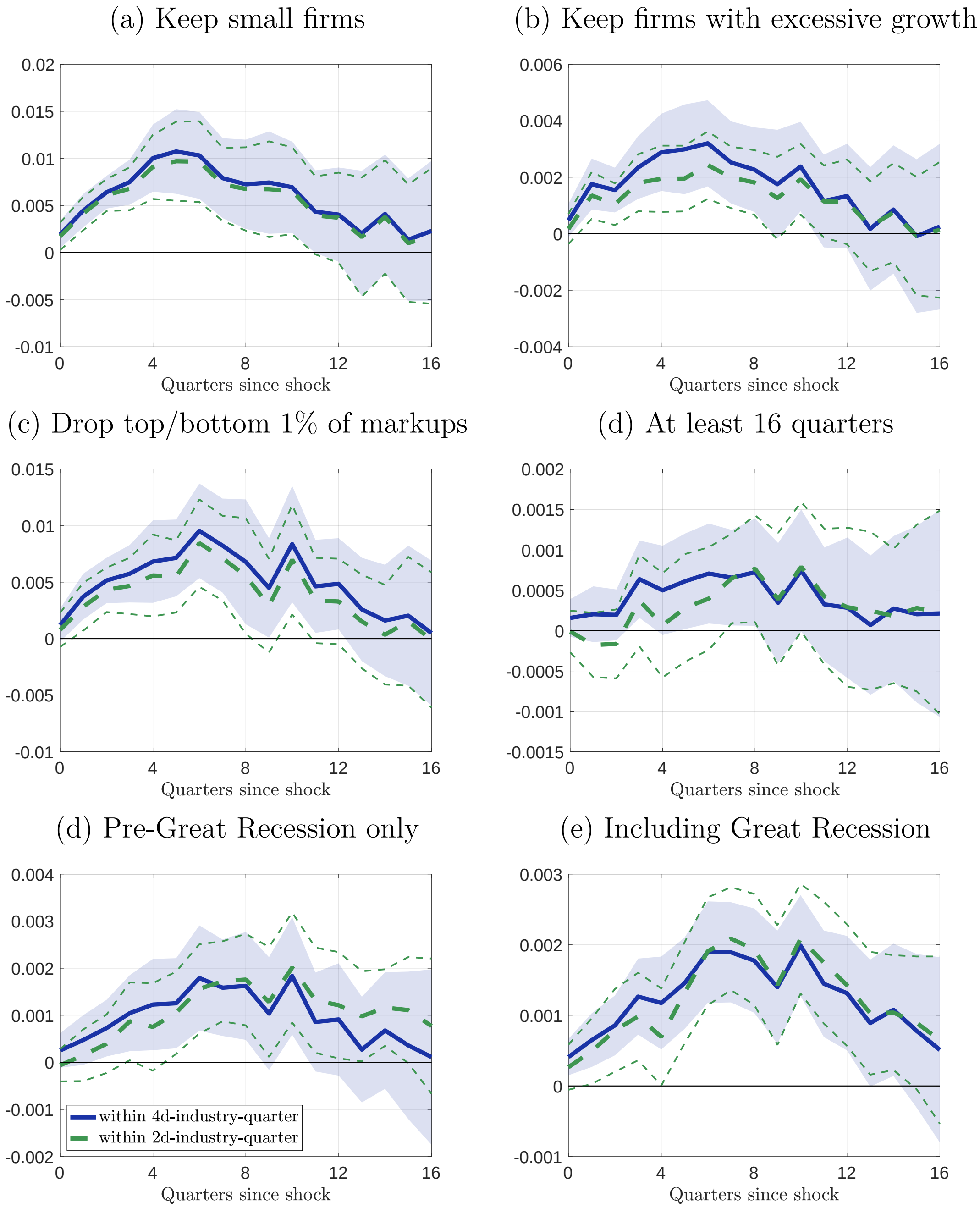
(b) Markups based on cost shares

	(1)	(2)	(3)	(4)
Implied price duration	0.0584	0.0510		
	(0.0177)	(0.0151)		
Price adjustment frequency			-0.431	-0.371
			(0.0931)	(0.0793)
Additional controls	No	Yes	No	Yes
2-digit industry FE	Yes	Yes	Yes	Yes
Observations	3873	3870	3873	3870
Adjusted R^2	0.196	0.271	0.204	0.276

Notes: Regression of firm-level markup (averaged over 2005–2011) on firm-level price adjustment frequency and implied price duration, respectively. Panel (a) uses markups based on an industry-specific translog production function, which gives rise firm-quarter-specific output elasticities. Panel (b) uses markups based on output elasticities estimated as the industry-quarter-specific median cost share. Standard errors are clustered at the two-digit industry level and shown in parentheses.

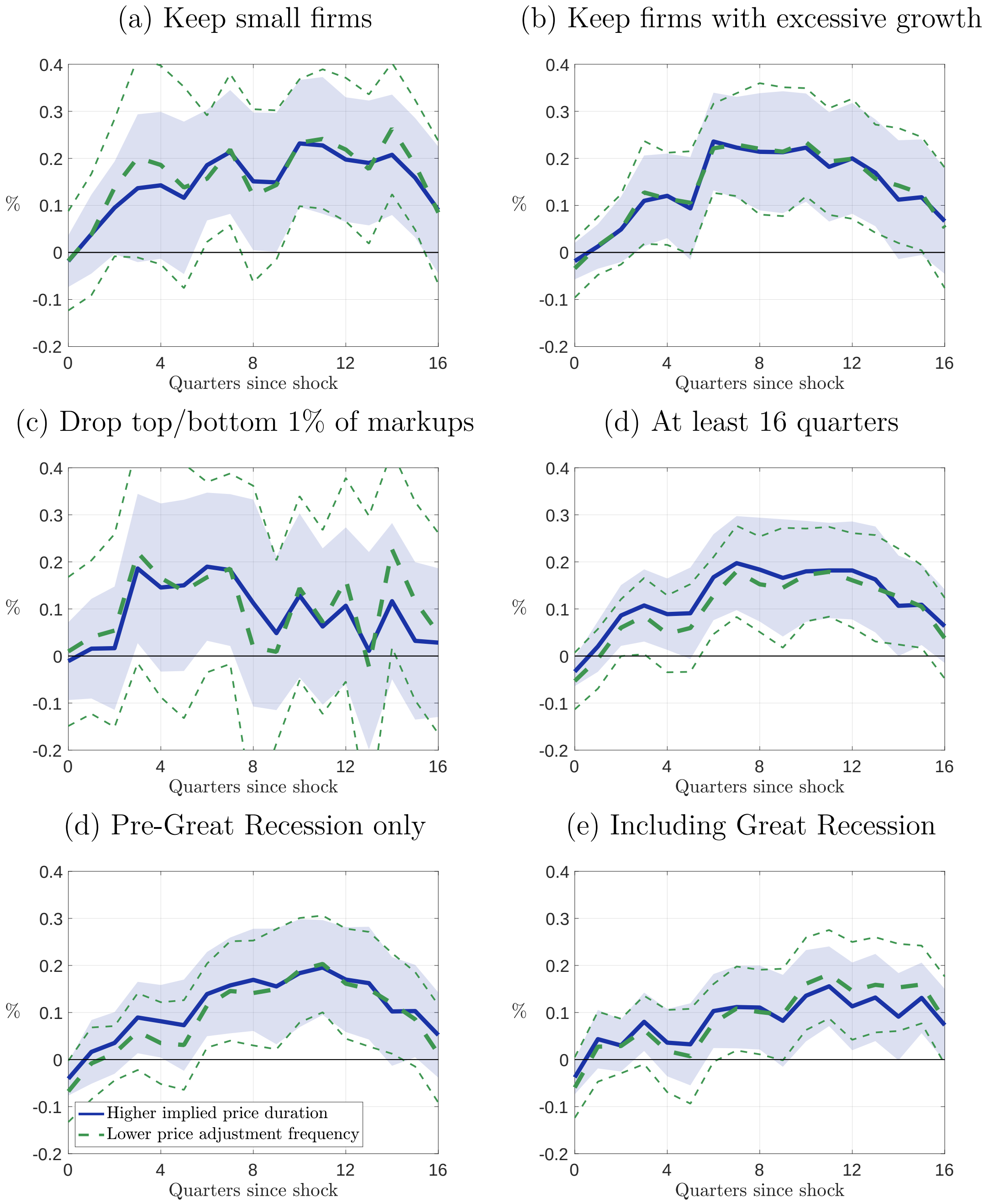
C.2 Results for alternative data treatments

Figure 13: Responses of markup dispersion under alternative data treatments



Notes: Responses to monetary policy shocks obtained from local projections as in equation (2.4). See the notes to Figure 14 for details on the data treatments. The shaded and bordered areas indicate one standard error bands based on Newey–West.

Figure 14: Relative markup response of firms with stickier prices under alternative data treatments



Notes: The figures show the response to a one standard deviation contractionary monetary policy shock of the firm-level markup of firms with a price adjustment frequency one standard deviation below (or with an implied price duration one standard deviation above) from panel local projections as in equation (2.5). The regressions include interactions with lagged log assets, leverage, and liquidity and their interactions with a monetary policy shock. *Keep small firms* does not drop firms with sales below 1 million (in 2012 US$). *Keep firms with excessive growth* does not drop firms with growth above 100% or below -67%. *Drop top/bottom 1%* drops markups in the top/bottom 1% in the quarter instead of 5%. *At least 16 quarters* restricts the sample to firms with at least 16 quarters of consecutive observations. *Pre-Great Recession only* considers only observations before 2008Q3. *Including Great Recession* does not drop the period 2008Q3–2009Q2 from the sample. The shaded and bordered areas indicate 90% error bands clustered by firms and quarters.

Table 7: Regressions of markup on price stickiness under alternative data treatments

(a) Keep small firms

	(1)	(2)	(3)	(4)
Implied price duration	0.0382	0.0447		
	(0.0146)	(0.0166)		
Price adjustment frequency			-0.280	-0.327
			(0.0655)	(0.0792)
Additional controls	No	Yes	No	Yes
2-digit industry FE	Yes	Yes	Yes	Yes
Observations	4308	4300	4308	4300
Adjusted R^2	0.168	0.194	0.170	0.196

(b) Keep firms with excessive growth

	(1)	(2)	(3)	(4)
Implied price duration	0.0485	0.0438		
	(0.0168)	(0.0148)		
Price adjustment frequency			-0.361	-0.317
			(0.0836)	(0.0750)
Additional controls	No	Yes	No	Yes
2-digit industry FE	Yes	Yes	Yes	Yes
Observations	4124	4074	4124	4074
Adjusted R^2	0.131	0.196	0.138	0.199

(c) Drop top/bottom 1% of markups

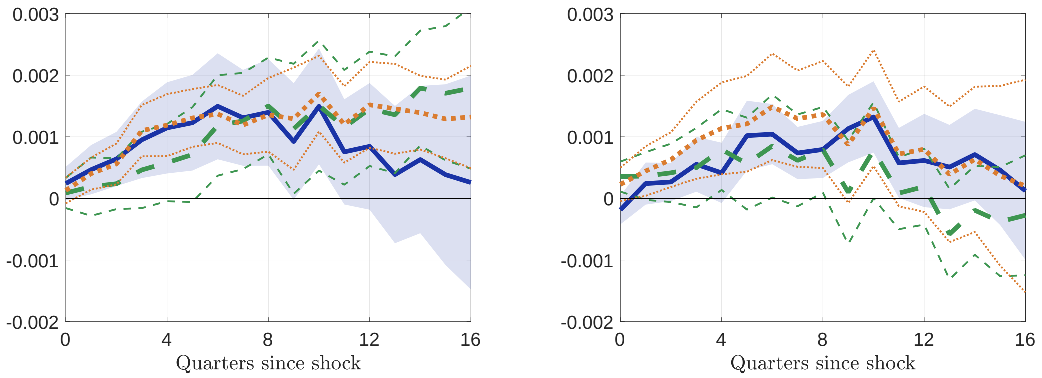
	(1)	(2)	(3)	(4)
Implied price duration	0.0404	0.0432		
	(0.0252)	(0.0277)		
Price adjustment frequency			-0.292	-0.312
			(0.126)	(0.143)
Additional controls	No	Yes	No	Yes
2-digit industry FE	Yes	Yes	Yes	Yes
Observations	4032	4027	4032	4027
Adjusted R^2	0.150	0.159	0.152	0.160

Notes: Regression of firm-level markup (averaged over 2005–2011) on firm-level price adjustment frequency and implied price duration, respectively. *Keep small firms* does not drop firms with sales below 1 million (in 2012 US$). *Keep firms with excessive growth* does not drop firms with growth above 100% or below -67%. *Drop top/bottom 1%* drops markups in the top/bottom 1% in the quarter instead of 5%. Standard errors are clustered at the two-digit industry level and shown in parentheses.

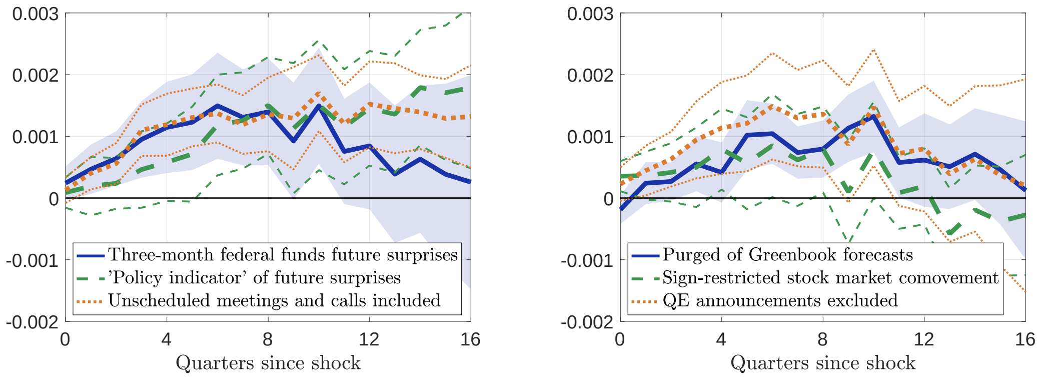
C.3 Results for alternative monetary policy shocks

Figure 15: Responses of markup dispersion for alternative monetary policy shocks

(a) within 2d-industry-quarter

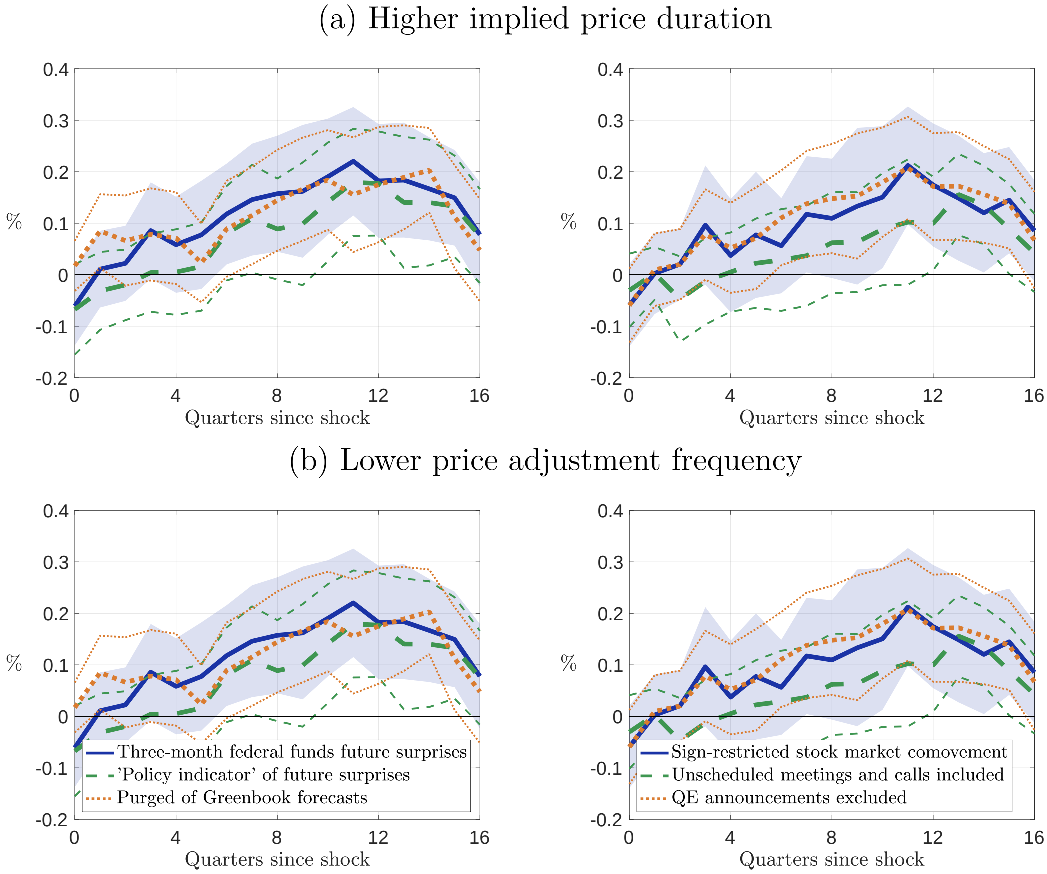


(b) within 4d-industry-quarter



Notes: Responses to monetary policy shocks obtained from local projections as in equation (2.4). The shaded and bordered areas indicate one standard error bands based on Newey–West.

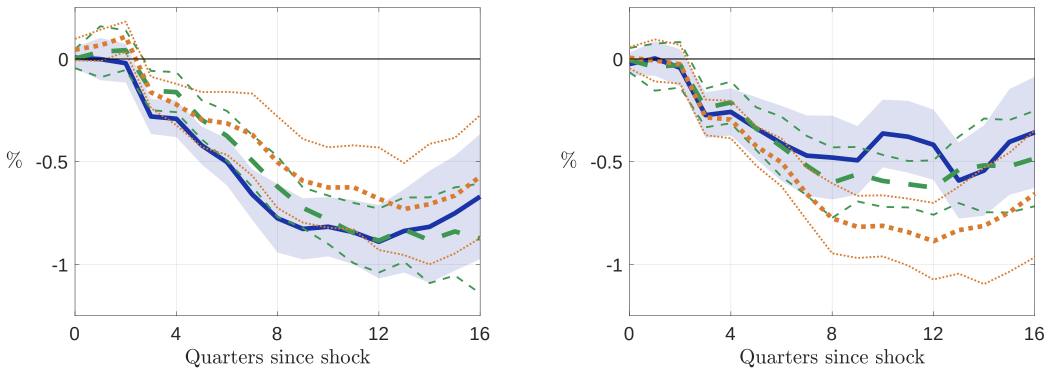
Figure 16: Relative markup response of firms with stickier prices for alternative monetary policy shocks



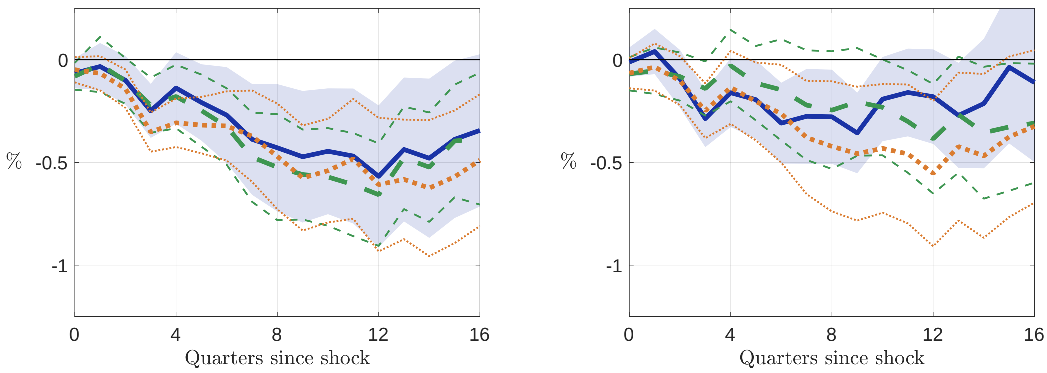
Notes: The figures show the response to a one standard deviation contractionary monetary policy shock of the firm-level markup of firms with a price adjustment frequency one standard deviation below (or with an implied price duration one standard deviation above) from panel local projections as in equation (2.5). The regressions include interactions with lagged log assets, leverage, and liquidity and their interactions with a monetary policy shock. The shaded and bordered areas indicate 90% error bands clustered by firms and quarters.

Figure 17: Aggregate productivity responses for alternative monetary policy shocks

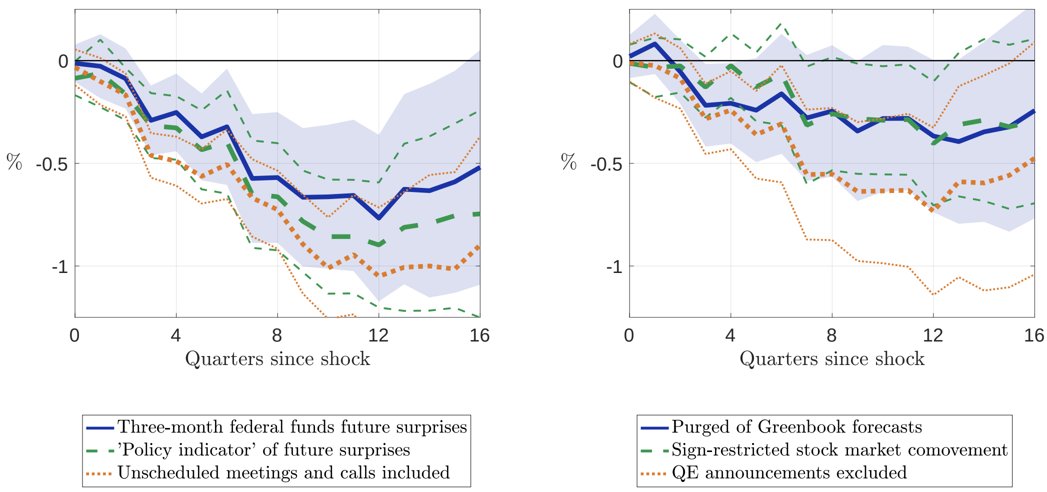
(a) TFP



(b) Utilization-adjusted TFP



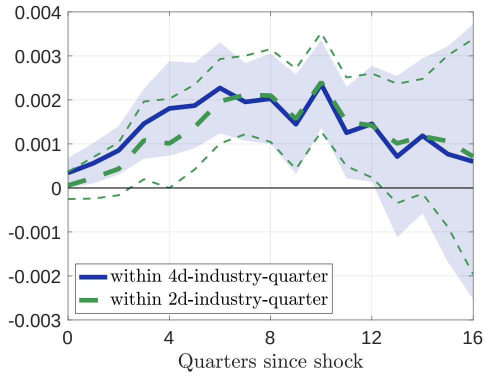
(c) Labor productivity



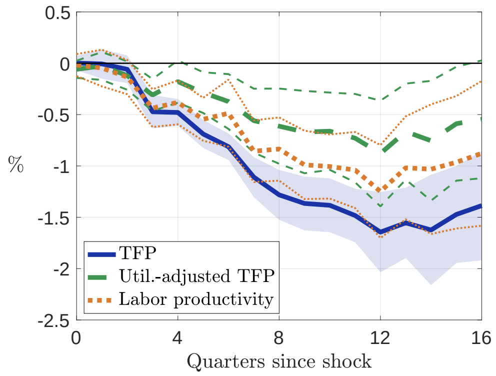
Notes: Responses to monetary policy shocks obtained from local projections as in equation (2.4). The shaded and bordered areas indicate one standard error bands based on Newey–West.

Figure 18: Main results using LP-IV

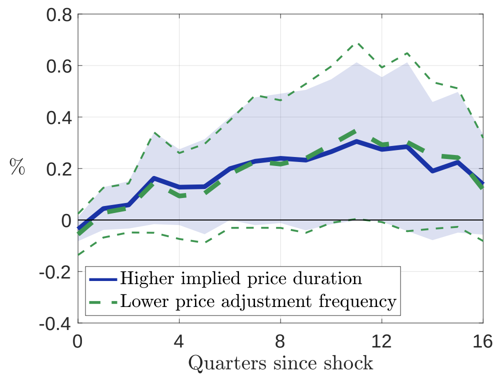
(a) Markup dispersion



(b) Aggregate productivity



(c) Relative markup response of firms with stickier prices



Notes: Responses to monetary policy shocks obtained from local projections with instrumental variables (LP-IV), $y_{t+h} - y_{t-1} = \alpha^h + \beta^h \Delta R_t + \gamma_1^h (y_{t-1} - y_{t-2}) + u_t^h$, and analogues of the panel local projections, where the changes in the one-year Treasury rate, ΔR_t, are instrumented with the monetary policy shocks $\varepsilon_t^{\mathrm{MP}}$. The shaded and bordered areas in panels (a) and (b) indicate a one standard error band based on Newey–West, and in panels (c) they indicate a 90% error band clustered by firms and quarters.

D Proofs

D.1 Markup dispersion and aggregate TFP

Consider a continuum of monopolistically competitive firms that produce variety goods Y_{it}. Firms employ a common constant-returns-to-scale production function $F(\cdot)$ that transforms a vector of inputs L_{it} into output subject to firm-specific productivity shocks $Y_{it} = A_{it}F(L_{it})$. The cost minimization problem yields that firm-specific $X_{it} = X_t/A_{it}$, where X_t denotes a common marginal costs term. Aggregate GDP is the output of a final good producer, which aggregates variety goods using a Dixit–Stiglitz aggregator $Y_t = (\int Y_{it}^{(\eta-1)/\eta} di)^{\eta/(\eta-1)}$. The cost minimization problem of the final good producer yields a demand curve for variety goods $Y_{it} = (P_{it}/P_t)^{-\eta} Y_t$, where P_t is an aggregate price index. Variety good producers choose prices to maximize period profits

$$\max_{P_{it}} (\tau_{it} P_{it} - X_{it}) Y_{it} \qquad \text{s.t.} \quad Y_{it} = (P_{it}/P_t)^{-\eta} Y_t,$$

where τ_{it} is a *markup wedge* in the spirit of Hsieh and Klenow (2009) and Baqaee and Farhi (2020). This wedge may be viewed as a shortcut for price rigidities. Profit maximization yields a markup $\mu_{it} = P_{it}/X_{it} = \frac{1}{\tau_{it}} \frac{\eta}{\eta-1}$. We compute aggregate TFP as a Solow residual by

$$\log \mathrm{TFP}_t = \log \left(\int Y_{it}^{(\eta-1)/\eta} di \right)^{\eta/(\eta-1)} - \log \int \frac{Y_{it}}{A_{it}} di.$$

This Solow residual has a model consistent Solow weight of one for the aggregate cost term. If we (a) apply a second-order approximation to $\log \mathrm{TFP}_t$ in $\log A_{it}$ and $\log \tau_{it}$, or if we (b) assume that A_{it} and τ_{it} are jointly log-normally distributed, we obtain

$$\log \mathrm{TFP}_t = -\frac{\eta}{2} \mathbb{V}_t(\log \mu_{it}) + \mathbb{E}_t(\log A_{it}) + \frac{\eta - 1}{2} \mathbb{V}_t(\log A_{it}).$$

Wedges τ_{it} drive markup dispersion and distort the economy away from allocative efficiency. Firms with high τ_{it} charge lower markups and use more inputs than socially optimal, and vice versa for low τ_{it}. This misallocation across firms results in lower aggregate TFP.

D.2 Proof of Proposition 1

Denote by $\mathbb{V}_t(\cdot)$, $\mathrm{Cov}_t(\cdot)$, $\mathrm{Corr}_t(\cdot)$ respectively the cross-sectional variance, covariance, correlation operator. The cross-sectional variance of the log markup is

$$\mathbb{V}_t(\log \mu_{it}) = \int (\log P_{it} - \log P_t - \log X_t)^2 di - \left[\int (\log P_{it} - \log P_t - \log X_t) di \right]^2.$$

The derivative w.r.t. $\log X_t$ is

$$\frac{\partial \mathbb{V}_t(\log \mu_{it})}{\partial \log X_t} = 2\int \log(\mu_{it})\rho_{it}di - 2\int \log(\mu_{it})di \int \rho_{it}di = 2\text{Cov}_t(\rho_{it}, \log \mu_{it}).$$

Hence, the markup variance falls in $\log X_t$ if $\text{Corr}_t(\rho_{it}, \log \mu_{it}) < 0$. $\square$

D.3 Proof of Proposition 2

We assume that

$$\log \begin{pmatrix} P_t/\bar{P} \\ X_t/\bar{X} \\ Y_t/\bar{Y} \end{pmatrix} \sim \mathcal{N}\left(\begin{bmatrix} -\frac{\sigma_p^2}{2} \\ -\frac{\sigma_x^2}{2} \\ -\frac{\sigma_y^2}{2} \end{bmatrix}, \begin{bmatrix} \sigma_p^2 & & \\ \sigma_{px} & \sigma_x^2 & \\ \sigma_{py} & \sigma_{xy} & \sigma_y^2 \end{bmatrix} \right).$$

Define $\tilde{\theta}_i \equiv \frac{\beta\theta_i}{1-\beta\theta_i}$, as well as

$$\begin{aligned} C_{it} &\equiv \mathbb{E}_t\left[\frac{X_{t+1}}{X_t}\left(\frac{P_{t+1}}{P_t}\right)^{\eta} \frac{Y_{t+1}}{Y_t} \right], \\ D_{it} &\equiv \mathbb{E}_t\left[\left(\frac{P_{t+1}}{P_t}\right)^{\eta-1} \frac{Y_{t+1}}{Y_t} \right], \\ \Psi_{it} &\equiv \frac{1+\tilde{\theta}_i C_{it}}{1+\tilde{\theta}_i D_{it}}, \end{aligned}$$

which allows us to rewrite the first-order condition in (3.3) as

$$P_{it}^* = \frac{\eta}{\eta-1} P_t X_t \Psi_{it}.$$

The terms C_{it} and D_{it} can be simplified

$$C_{it} = \frac{\bar{X}\bar{P}^{\eta}\bar{Y}}{X_t P_t^{\eta} Y_t} \exp\left\{ \eta(\eta-1)\frac{\sigma_p^2}{2} + \eta\sigma_{px} + \eta\sigma_{py} + \sigma_{xy} \right\}, \quad D_{it} = \frac{\bar{P}^{\eta-1}\bar{Y}}{P_t^{\eta-1} Y_t} \exp\left\{ (\eta-1)(\eta-2)\frac{\sigma_p^2}{2} + (\eta-1)\sigma_{py} \right\}.$$

Since $\tilde{\theta}_i \in (0,1)$, we obtain $\Psi_{it} > 1$ when $P_t = \bar{P}$ and $X_t = \bar{X}$, if

$$(\eta-1)\sigma_p^2 + \sigma_{py} + \eta\sigma_{px} + \sigma_{xy} > 0.$$

Under this condition, we obtain $\mu_{it}^* > \frac{\eta}{\eta-1}$. Under the same condition, we further obtain

$$\frac{\partial \Psi_{it}}{\partial \tilde{\theta}_i} = \frac{C_{it} - D_{it}}{(1+\tilde{\theta}_i D_{it})^2} > 0, \quad \text{and hence} \quad \frac{\partial \Psi_{it}}{\partial \theta_i} > 0.$$

We next study the pass-through of a transitory or permanent change in X_t. Consider first a *transitory* change in X_t away from $\bar{X}$. The expected pass-through is

$$\bar{\rho}_{it} = (1-\theta_i)\frac{\partial \log P_{it}}{\partial \log X_t} = (1-\theta_i)\left(1+\Phi_{it}\right), \quad \text{where} \quad \Phi_{it} = \frac{\partial \log \Psi_{it}}{\partial \log X_t}$$

and

$$\Phi_{it} = \frac{\tilde{\theta}_i \frac{\partial C_{it}}{\partial \log X_t}(1+\tilde{\theta}_i D_{it}) - (1+\tilde{\theta}_i C_{it})\tilde{\theta}_i \frac{\partial D_{it}}{\partial \log X_t}}{(1+\tilde{\theta}_i D_{it})^2}\Psi_{it}^{-1} = -\frac{\tilde{\theta}_i C_{it}}{1+\tilde{\theta}_i D_{it}}\Psi_{it}^{-1} = -\frac{\tilde{\theta}_i C_{it}}{1+\tilde{\theta}_i C_{it}} < 0.$$

Hence pass-through becomes

$$\bar{\rho}_{it} = \frac{1-\theta_i}{1+\tilde{\theta}_i C_{it}} \in (0,1).$$

In addition, the pass-through falls in θ_i,

$$\frac{\partial \bar{\rho}_{it}}{\partial \theta_i} = -(1+\Phi_{it}) + (1-\theta_i)\frac{\partial \Phi_{it}}{\partial \theta_i} < 0.$$

We next examine a *permanent* change in X_t, which is a change in $\bar{X}$ (starting in period t). At $P_t = \bar{P}$ and $X_t = \bar{X}$,

$$\frac{\partial \log P_{it}^*}{\partial \log \bar{X}} = 1.$$

Expected pass-through is then $\bar{\rho}_{it} = 1-\theta_i$ and hence $\frac{\partial \bar{\rho}_{it}}{\partial \theta_i} < 0$. □

D.4 Proof of Proposition 3

Let us first define

$$C_{it} = \left(\frac{P_{it}}{P_{i,t-1}} - 1\right)\frac{P_{it}}{P_{i,t-1}},$$
$$D_{it} = \mathbb{E}_t\left[\left(\frac{P_{i,t+1}}{P_{it}} - 1\right)\frac{P_{i,t+1}}{P_{it}}\right],$$

such that we can re-write the first-order condition in equation (3.4) more compactly as

$$(1-\eta)\left(\frac{P_{it}}{P_t}\right)^{1-\eta} Y_t + \eta X_t \left(\frac{P_{it}}{P_t}\right)^{-\eta} Y_t = \phi_i (C_{it} - D_{it}).$$

Further define $\bar{\phi}_i = 0$ and denote by an upper bar any object that is evaluated at $\bar{\phi}_i$, such as the price P_{it}, which is $\bar{P}_{it} = \frac{\eta}{\eta-1} P_t X_t$. In addition,

$$\bar{C}_{it} = \left(\frac{\bar{P}_{it}}{\bar{P}_{i,t-1}} - 1 \right) \frac{\bar{P}_{it}}{\bar{P}_{i,t-1}} = (\Pi_{pt}\Pi_{xt})^2 - \Pi_{pt}\Pi_{xt},$$

$$\bar{D}_{it} = E_t \left[\left(\frac{\bar{P}_{i,t+1}}{\bar{P}_{it}} - 1 \right) \frac{\bar{P}_{i,t+1}}{\bar{P}_{it}} \right] = \frac{\exp\left\{ \frac{3}{2}\sigma_p^2 + \frac{3}{2}\sigma_x^2 + 4\sigma_{px} \right\}}{(\Pi_{pt}\Pi_{xt})^2} - \frac{\exp\{\sigma_{pw}\}}{\Pi_{pt}\Pi_{xt}}.$$

We next use a first-order approximation of the first-order condition at $\bar{\phi}_i$ and with respect to ϕ_i and $\log P_{it}$. Denoting $\mathrm{dlog}P_{it} = \log P_{it} - \log \bar{P}_{it}$ and $\mathrm{d}\phi_i = \phi_i$, we obtain

$$(1-\eta)^2 \left(\frac{P_{it}}{\bar{P}_t} \right)^{1-\eta} Y_t \mathrm{dlog}P_{it} - \eta^2 X_t \left(\frac{\bar{P}_{it}}{P_t} \right)^{-\eta} Y_t \mathrm{dlog}P_{it} = (\bar{C}_{it} - \bar{D}_{it})\mathrm{d}\phi_i.$$

This yields

$$\Psi_{it} \equiv \frac{\mathrm{dlog}P_{it}}{\mathrm{d}\phi_i} = \frac{\bar{D}_{it} - \bar{C}_{it}}{(\eta-1)^\eta \eta^{1-\eta} X_t^{1-\eta} Y_t},$$

and hence $\log P_{it} \approx \log \bar{P}_{it} + \Psi_{it} d\phi_i$. For $\phi_i > 0$, the markup is above the frictionless one if $P_{it} > \bar{P}_{it}$, which holds if $\Psi_{it} > 0$. For $P_t = \bar{P}$ and $X_t = \bar{X}$, $\Psi_{it} > 0$ if

$$\sigma_p^2 + \sigma_x^2 + 2\sigma_{px} > 0,$$

for which a sufficient condition is that the correlation

$$\rho_{px} \equiv \frac{\sigma_{px}}{\sigma_p \sigma_x} > -1.$$

Under the same condition, $\frac{\partial P_{it}}{\partial \phi_i} > 0$.

We next study the pass-through of a transitory or permanent change in X_t. The pass-through is

$$\rho_{it} = 1 + \frac{\partial \Psi_i}{\partial \log X_t} d\phi_i.$$

We next examine the conditions under which pass-through falls in ϕ_i, i.e., conditions under which

$$\frac{\partial \Psi_i}{\partial \log X_t} < 0,$$

which is equivalent to examining the conditions for

$$\frac{\partial \bar{D}_{it}}{\partial \log X_t} - \frac{\partial \bar{C}_{it}}{\partial \log X_t} + (\eta - 1)(\bar{D}_{it} - \bar{C}_{it}) < 0.$$

Consider first a *transitory* change in X_t away from $\bar{X}$,

$$\frac{\partial \bar{C}_{it}}{\partial \log X_t} = 2(\Pi_{pt}\Pi_{xt})^2 - \Pi_{pt}\Pi_{xt},$$
$$\frac{\partial \bar{D}_{it}}{\partial \log X_t} = -2(\Pi_{pt}\Pi_{xt})^{-2} \exp\left\{\frac{3}{2}\sigma_p^2 + \frac{3}{2}\sigma_x^2 + 4\sigma_{px}\right\} + (\Pi_{pt}\Pi_{xt})^{-1} \exp\left\{\sigma_{px}\right\}.$$

For $P_t = \bar{P}$ and $X_t = \bar{X}$, we obtain

$$\frac{\partial \Psi_i}{\partial \log X_t} < 0 \qquad \text{if} \quad \eta < \tilde{\eta}^{\text{transitory}} = 2 + \frac{1 + \exp\left\{\frac{3}{2}\sigma_p^2 + \frac{3}{2}\sigma_x^2 + 4\sigma_{px}\right\}}{\exp\left\{\frac{3}{2}\sigma_p^2 + \frac{3}{2}\sigma_x^2 + 4\sigma_{px}\right\} - \exp\left\{\sigma_{px}\right\}}$$

We next consider a *permanent* change, for which we have

$$\frac{\partial \bar{C}_{it}}{\partial \log X_t} = 2(\Pi_{pt}\Pi_{wt})^2 - \Pi_{pt}\Pi_{wt}, \qquad \frac{\partial \bar{D}_{it}}{\partial \log X_t} = 0.$$

For $P_t = \bar{P}$ and $X_t = \bar{X}$, we obtain

$$\frac{\partial \Psi_i}{\partial \log X_t} < 0 \qquad \text{if} \quad \eta < \tilde{\eta}^{\text{permanent}} = 1 + \frac{1}{\exp\left\{\frac{3}{2}\sigma_p^2 + \frac{3}{2}\sigma_x^2 + 4\sigma_{px}\right\} - \exp\left\{\sigma_{px}\right\}}$$

It always holds that $\eta^{\text{permanent}} < \eta^{\text{transitory}}$ and we define $\tilde{\eta} \equiv \eta^{\text{permanent}}$. □

E Menu cost model

To study the presence of precautionary price setting in menu cost models, we proceed numerically. Consider the partial equilibrium menu cost model

$$V(p, Z) = \mathbb{E}_\xi[\max\{V^A(Z) - \xi, V^N(Z)\}]$$
$$V^A(Z) = \max_{p^*}\left\{\left(\frac{p^*}{P} - X\right)\left(\frac{p^*}{P}\right)^{-\eta} + \beta\mathbb{E}_Z\left[V(p^*, Z')\right]\right\}$$
$$V^N(p, Z) = \left(\frac{p}{P} - X\right)\left(\frac{p}{P}\right)^{-\eta} + \beta\mathbb{E}_Z\left[V(p, Z')\right]$$

where p is the price a firm sets and Z denote a vector of the aggregate state variables price level (P), aggregate demand (Y), and marginal costs (X). The firm chooses to adjust prices in the presence of the menu cost ξ.

We set $\eta = 6$ and $\beta = 1.03^{-1/4}$. We solve the model using value function iteration with off-grid interpolation with respect to p using cubic splines as basis function. To solve accurately for differences in p^* that arise from small differences in ξ requires a fine grid for both p and Z. To alleviate the numerical challenge, we assume ξ is stochastic and drawn from an iid exponential distribution, parametrized by $\bar{\xi}$. Results change only little when using a uniform distribution.

We assume 200 grid points on a log-spaced grid for p. To capture aggregate uncertainty in Z, we first estimate a first-order Markov process for Z in the data and then discretize it using a Tauchen procedure. In the univariate case, when only allowing for inflation uncertainty, the precautionary price setting was accurately captured starting from about 49 grid points for Z. Discretizing a three-variate VAR with 49 grid points for each variable is costly. Even more importantly, the state space, on which to solve the model, becomes very large. We therefore proceed with the univariate case. We estimate an AR(1) on quarterly post-1984 data of the log CPI and apply the Tauchen method with 49 grid points.

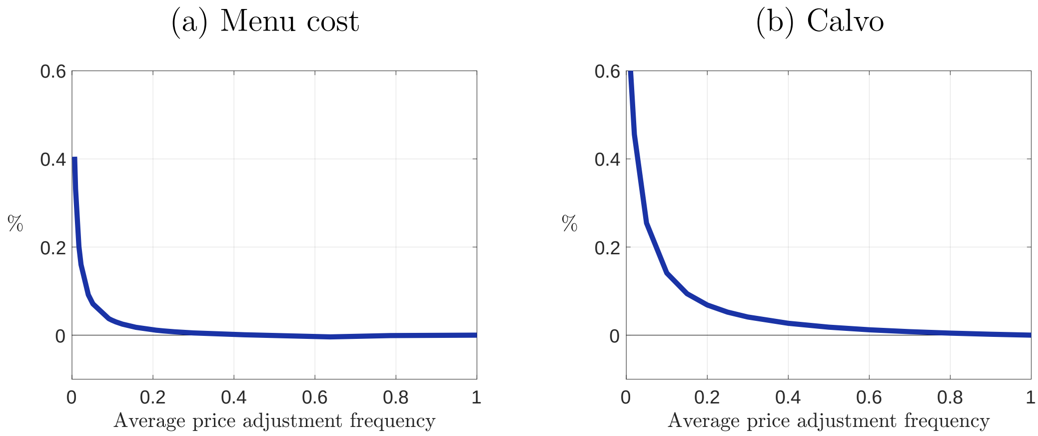
We solve the stationary equilibrium of the menu cost and Calvo model for a vector of different $\bar{\xi}$, which imply different equilibrium price adjustment frequencies. Figure 19 plots the price setting policy p^* at the unconditional mean of Z for different average price adjustment frequencies. We compare menu costs in panel (a) with Calvo in panel (b). The figures shows that precautionary price setting exists and is amplified by the degree of price-setting friction in a menu cost environment. Compared to Calvo, menu costs generate somewhat muted precautionary price setting.

F Details on the Quantitative New Keynesian Model

This section presents details on the quantitative New Keynesian model in Section 4. We assume a representative infinitely-lived household who maximizes

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^{1-\frac{1}{\gamma}}}{1-\frac{1}{\gamma}} - \frac{N_t^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}} \right),$$

Figure 19: Precautionary price setting under menu costs and Calvo



Notes: The figures show percentage difference between the dynamic optimal price relative to the frictionless optimal one.

subject to the budget constraints $P_tC_t + R_t^{-1}B_t \leq B_{t-1} + W_tN_t + D_t$ for all t, where C_t is aggregate consumption, P_t an aggregate price index, B_t denotes one-period discount bounds purchased at price R_t^{-1}, N_t employment, W_t the nominal wage, and D_t aggregate dividends. We impose the solvency constraint $\lim_{T\to\infty} \mathbb{E}_t[\Lambda_{t,T}\frac{B_T}{P_T}] \geq 0$ for all t, where $\Lambda_{t,T} = \beta^{T-t}(C_T/C_t)^{-\frac{1}{\gamma}}$ is the stochastic discount factor. The final output good Y_t is produced with a Dixit–Stiglitz aggregator

$$Y_t = \left(\int_0^1 Y_{it}^{\frac{\eta-1}{\eta}}\,\mathrm{d}i\right)^{\frac{\eta}{\eta-1}},$$

where η is the elasticity of substitution between differentiated goods $\{Y_{it}\}$. Intermediate goods are with the technology $Y_{it} = A_tN_{it}$, where A_t is a common technology shifter, which follows $\log A_t = \rho_a \log A_{t-1} + \varepsilon_{a,t}$ and $\varepsilon_{a,t} \sim \mathcal{N}(0, \sigma_a^2)$ are technology shocks. Final good aggregation implies an isoelastic demand schedule for intermediate goods given by $Y_{it} = (P_{it}/P_t)^{-\eta}Y_t$, where $P_t = (\int_0^1 P_{it}^{1-\eta}di)^{1/(1-\eta)}$ denotes the aggregate price index and P_{it} the firm-level price. Firms may reset their prices P_{it} with firm-specific probability $1 - \theta_i$. The price setting policy maximizes the value of the firm to its shareholder,

$$\max_{P_{it}} \sum_{j=0}^{\infty} \theta_i^j \mathbb{E}_t \left[\frac{\Lambda_{t,t+j}}{P_{t+j}}\left(\frac{P_{it}}{P_{t+j}} - W_{t+j}\right)\left(\frac{P_{it}}{P_{t+j}}\right)^{-\eta} Y_{t+j}\right].$$

The firm type k-specific price index is

$$P_{kt} = \left[(1-\theta_k)\tilde{P}_{kt}^{1-\eta} + \theta_k P_{kt-1}^{1-\eta}\right]^{\frac{1}{1-\eta}}$$

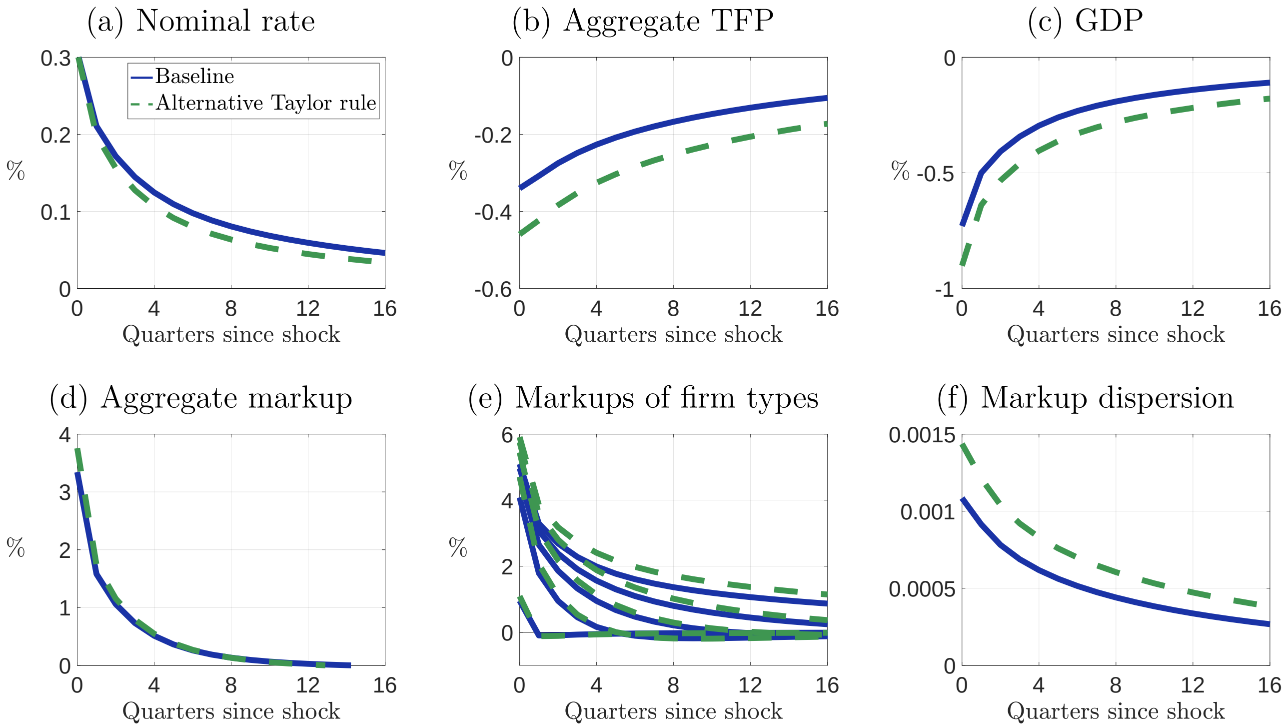
where $\tilde{P}_{kt}$ is the optimal reset price of a firm of type k. The monetary authority aims to stabilize inflation (P_t/P_{t-1}) and fluctuations in output, Y_t, around its natural level, denoted $\tilde{Y}_t$, by following the Taylor-type rule, subject to monetary policy shocks ν_t,

$$R_t = R_{t-1}^{\rho_r}\left[\frac{1}{\beta}\left(\frac{P_t}{P_{t-1}}\right)^{\phi_\pi}\left(\frac{Y_t}{\tilde{Y}_t}\right)^{\phi_y}\right]^{1-\rho_r} \exp\{\nu_t\}, \quad \nu_t \sim \mathcal{N}(0, \sigma_\nu^2).$$

The competitive equilibrium is defined by firm-level allocations $\{Y_{it}, N_{it}\}_{t=0}^{\infty}$ and prices $\{P_{it}\}_{t=0}^{\infty}$ for all i, and aggregate allocations and prices $\{Y_t, N_t, P_t, R_t\}_{t=0}^{\infty}$ such that households and firms maximize their objective functions, the monetary authority follows the policy rule. The final goods market clears, $Y_t = C_t$, and the labor market clears, $N_t = \int_0^1 N_{it}\mathrm{d}i$, in every period t.

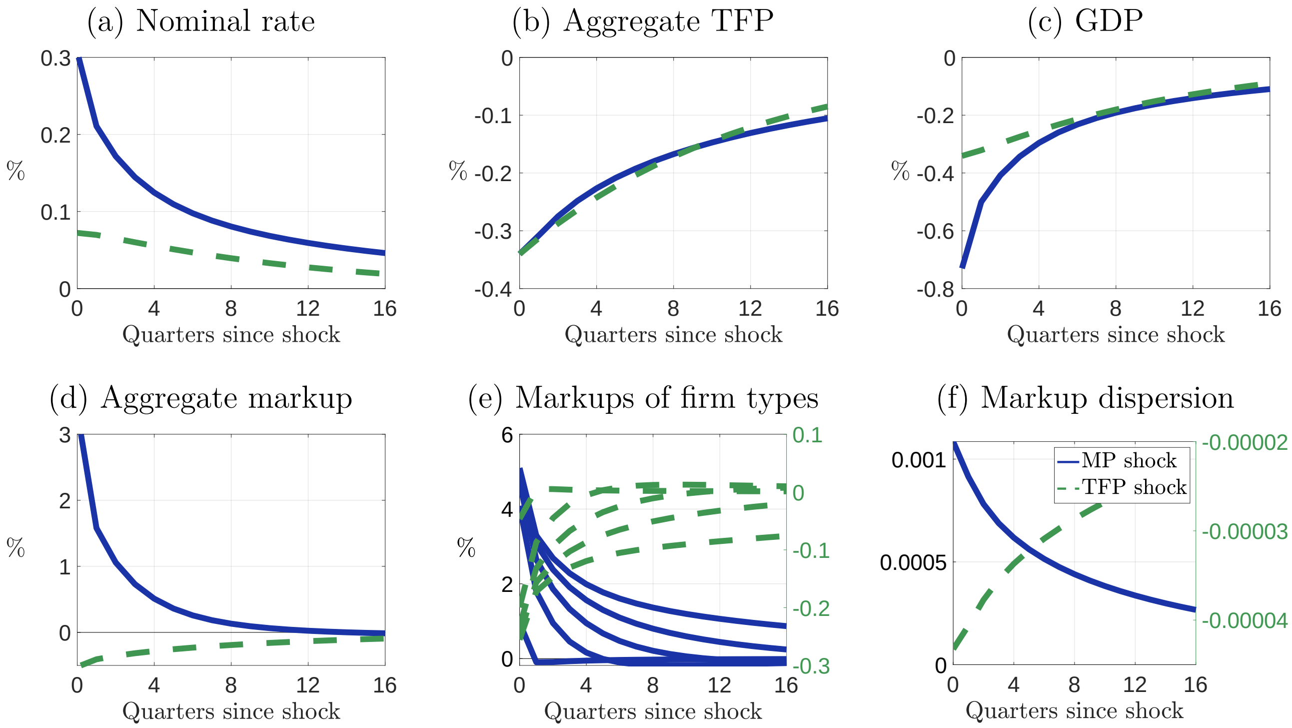
G Additional Model Results

Figure 20: Model responses to monetary policy shocks under alternative Taylor rule



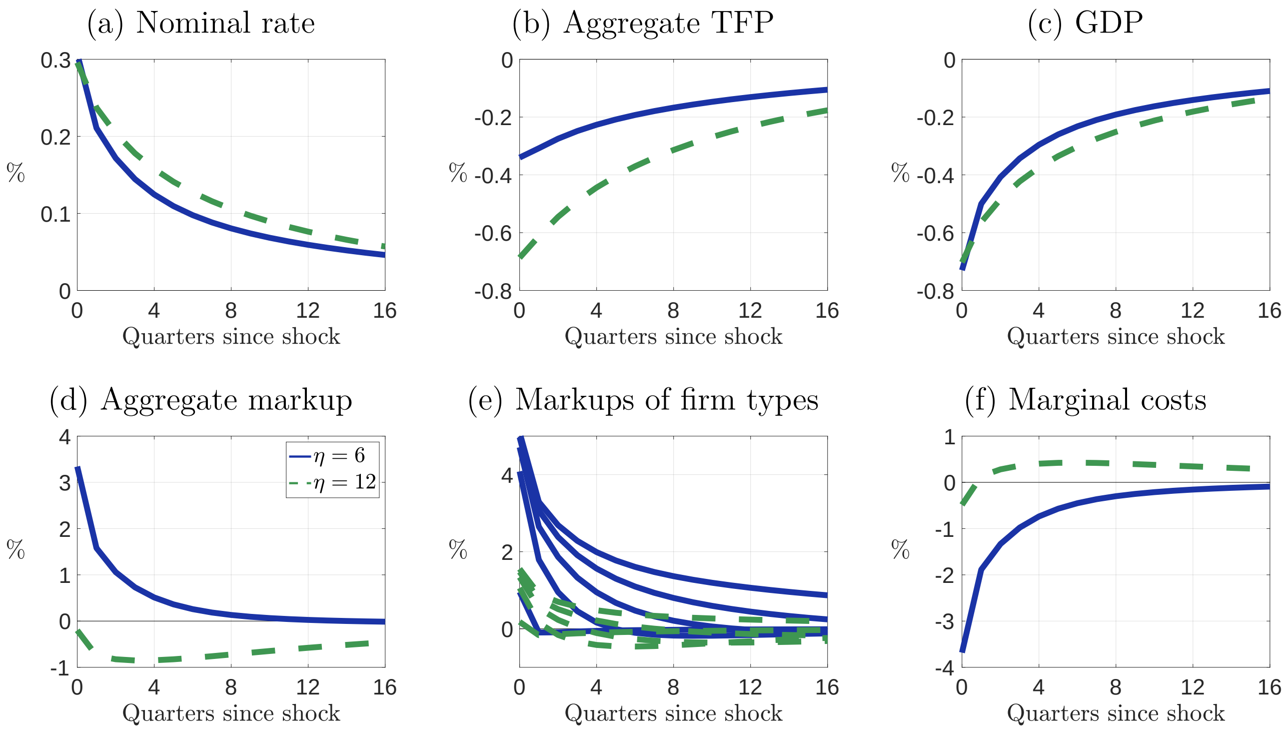
Notes: This figure shows impulse responses to a one standard deviation monetary policy shock. *Baseline* corresponds to the model in the main text. In particular, the central bank follows a Taylor rule, which reacts to fluctuation in the output gap. The gap is defined relative to natural output (the level prevailing under flexible prices), which is unchanged after monetary policy shocks. *Alternative Taylor rule* corresponds to a setup in which the central bank computes natural output based on the observed movements in aggregate TFP. Consequently, natural output is perceived to react to monetary policy shocks, which leads to a different policy response. The standard deviation of monetary policy shocks σ_ν is re-calibrated to match the response of the nominal rate of 30bp.

Figure 21: Model responses to technology shocks



Notes: This figure compares the impulse responses to a one standard deviation monetary policy shock to those to a technology shock. The persistence of TFP and the technology shock size are calibrated to match the shape of the TFP response to monetary policy shocks.

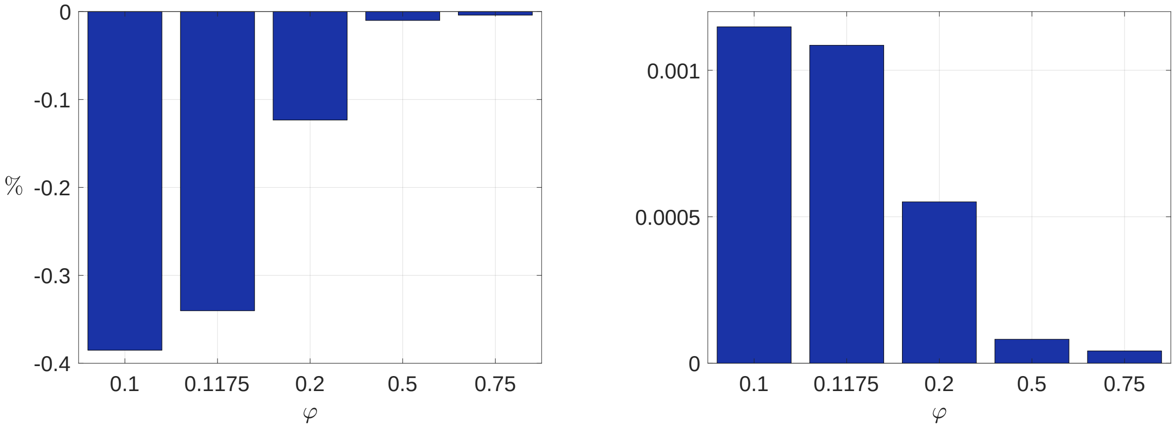
Figure 22: Model responses to monetary policy shocks when varying the elasticity of substitution



Notes: This figure shows impulse responses to a one standard deviation monetary policy shock for two values of the elasticity of substitution between variety goods η. The value 6 corresponds to our baseline calibration and the value 12 corresponds to an intermediate value of elasticities considered in the literature (e.g., Fernandez-Villaverde et al., 2015). The standard deviation of monetary policy shocks σ_ν is re-calibrated to match the response of the nominal rate of 30bp.

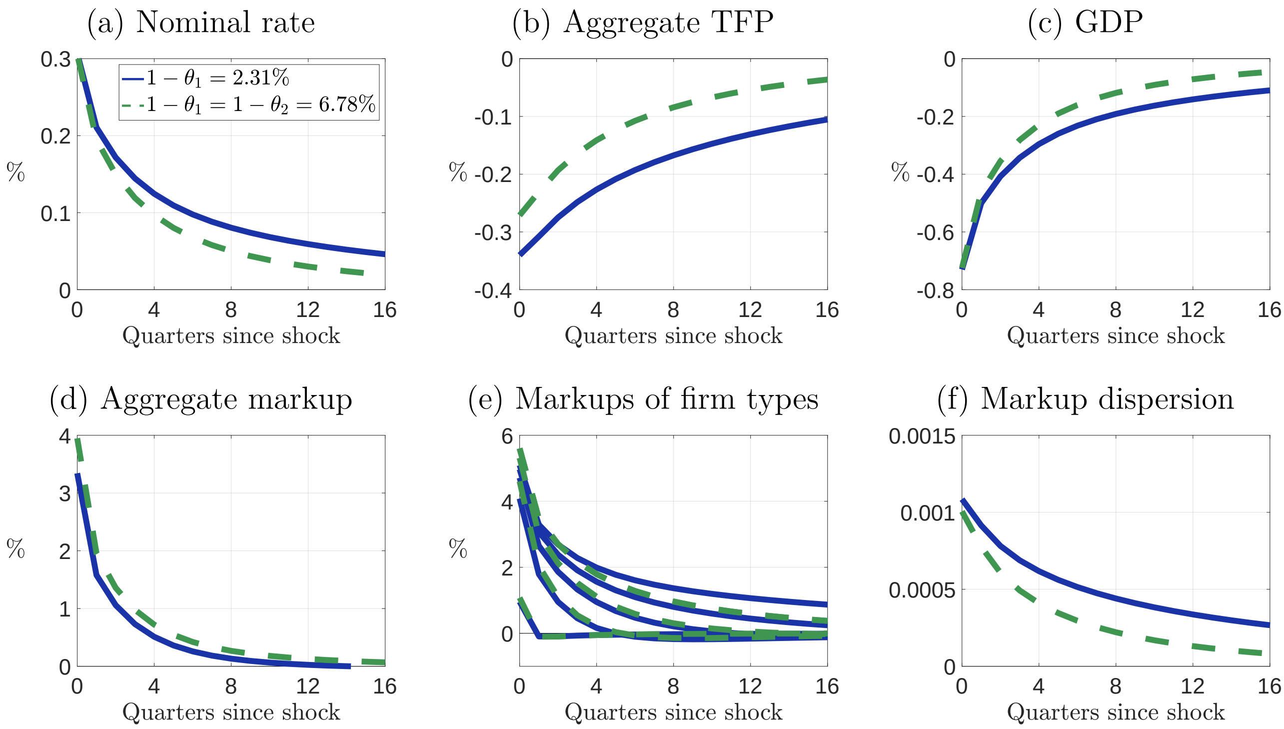
Figure 23: Model responses to monetary policy shocks when varying the Frisch elasticity

(a) On-impact aggregate TFP response (b) On-impact markup dispersion response



Notes: This figure plots the size of on-impact impulse responses of TFP and markup dispersion to a one standard deviation monetary policy shock under different calibrations of the Frisch elasticity of labor supply φ. The value 0.1 corresponds to the lower end of short-run elasticity estimates surveyed by Ashenfelter et al. (2010). The value 0.1175 corresponds to our baseline calibration, which matches the contribution of TFP to the GDP response and is in the range of Ashenfelter et al. (2010). The value 0.5 corresponds to a lower end and an intermediate value (which is suspected to be potentially upward-biased) in Chetty et al. (2011). The standard deviation of monetary policy shocks σ_ν is re-calibrated to match the response of the nominal rate of 30bp.

Figure 24: Model responses to monetary policy shocks when stickiest firm type is more flexible



Notes: This figure compares impulse responses to a one standard deviation monetary policy shock in the baseline calibration to the setting in which we reduce the largest price rigidity, θ_1, to the second-largest price rigidity, θ_2. The standard deviation of monetary policy shocks σ_ν is re-calibrated to match the response of the nominal rate of 30bp.

Acknowledgements

A previous version of this paper was titled "The misallocation channel of monetary policy". We owe special thanks to Klaus Adam, Antonio Ciccone, Konrad Stahl, Venky Venkateswaran, as well as numerous seminar participants and discussants for helpful comments and suggestions. We thank Michael Weber for generously sharing data on price adjustment frequencies with us. Funding by the German Research Foundation (DFG) through CRC TR 224 (Project C02) is gratefully acknowledged. Timo Reinelt gratefully acknowledges financial support from Stiftung Geld & Währung. Matthias Meier gratefully acknowledges financial support from the UniCredit & Universities Foundation. The contribution of Matthias Meier has been prepared under the Lamfalussy Fellowship Program sponsored by the ECB. Any views expressed are only those of the authors and do not necessarily represent the views of the ECB or the Eurosystem.

Matthias Meier
Universität Mannheim, Mannheim, Germany; email: m.meier@uni-mannheim.de

Timo Reinelt
Universität Mannheim, Mannheim, Germany; email: timo.reinelt@gess.uni-mannheim.de

Postal address 60640 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Website www.ecb.europa.eu

PDF ISBN 978-92-899-4070-2 ISSN 1725-2806 doi:10.2866/201533 QB-AR-20-079-EN-N