Four Essays in Macroeconomics

Inauguraldissertation zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

Timo Reinelt

2022

Abteilungssprecher:	Prof. Klaus Adam, Ph.D.
Referent:	Prof. Dr. Matthias Meier
Koreferent:	Prof. Klaus Adam, Ph.D.
Vorsitzender der Disputation:	Prof. Dr. Antonio Ciccone
Tag der Disputation:	2.11.2022

Acknowledgements

This dissertation would not have been possible without the support from many people, whom I would like to thank. First and foremost, I am extremely grateful to my advisors Matthias Meier and Klaus Adam, who provided exceptional support and guidance. Their comments and advice, as well as the collaboration on joint projects, have been most valuable.

I thank Antonio Ciccone for becoming part of the dissertation committee and for providing valuable feedback during early stages of my research.

Parts of the research in this thesis are joint work with Luca Dedola, Joachim Jungherr, Chiara Osbat, Oliver Pfäuti, and Immo Schott. I am very grateful for the collaboration with them.

I greatly benefited from the feedback and guidance from various faculty members, in particular from Antoine Camous, Husnu Dalgic, Harald Fadinger, Andreas Gulyas, Anne Hannusch, Krzysztof Pytka, David Koll, Konrad Stahl, Michèle Tertilt, and Minchul Yum. I am also grateful for the influence and support from previous teachers, especially Martin Biewen, Joachim Grammig, Sophocles Mavroeidis, and Johannes Pfeifer.

This thesis was written within the Center for Doctoral Studies in Economics and the Graduate School of Economics and Social Sciences at the University of Mannheim. I am extremely grateful to have had the opportunity to pursue my Ph.D. in this environment. I am grateful for the administrative support from Golareh Khalilpour, Ulrich Kehl, Marion Lehnert, Caroline Mohr, Susanne Putz, and Sylvia Rosenkranz. I gratefully acknowledge financial support from Stiftung Geld & Währung and Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), also through CRC TR 224 (Project C02). Previous support from the Studienstiftung (German National Academic Foundation) and the DAAD (German Academic Exchange Foundation) was instrumental to reach my Ph.D. studies.

I am very happy to have spent the time in Mannheim with a great group of fellow Ph.D. students, and I thank especially Andrey Alexandrov, Moritz Drechsel-Grau, Fabian Greimel, Niklas Garnadt, Johannes Gönsch, Lukas Henkel, Maximilian Jager, Alexander Matusche, Michaela Paffenholz, Johannes Wacks, and Frederick Zadow for their collegiality.

I am grateful to Alexander Haas and Lukas Hensel for the long-lasting friendship and their academic and non-academic support.

Finally, I thank my family and especially my wife Julia for her love and support.

Contents

Introduction

This dissertation contains four essays in macroeconomics and monetary economics. The chapters consider topics in firm dynamics, price setting, financial frictions, and subjective beliefs, to understand their implications for business cycles and monetary policy.

Chapter 1, which is joint work with Matthias Meier, studies the role of factor misallocation and price rigidity for the transmission of monetary policy. It is known from previous research that monetary policy not only affects the aggregate economy through employment and investment, but also through aggregate total factor productivity (TFP). The chapter shows that the heterogeneous response of firms' markups, by generating changes in markup dispersion across firms, can account for a significant fraction of the aggregate TFP response in the first two years after a monetary policy shock. This effect can be explained by heterogeneity in price rigidity across firms if firms have a precautionary price setting motive. Further evidence shows that firms with more rigid prices set higher markups on average, consistent with this explanation. The chapter studies the mechanism and its implications in a quantitative New Keynesian model. When solved around the stochastic steady state, the model demonstrates that firms with more rigid prices set higher markups in general equilibrium. Additionally, the model generates quantitatively relevant responses of markup dispersion and aggregate TFP after monetary policy shocks. These effects are important for monetary policy making. Interest rates would respond less aggressively, leading to larger macroeconomic fluctuations, if the monetary authority ignored the endogeneity of aggregate TFP.

Chapter 2, which is joint work with Klaus Adam and Oliver Pfäuti, studies the implications of housing price dynamics and falling natural rates of interest for optimal monetary policy. The chapter shows that housing price expectations deviate from the full-information rational expectations benchmark. A comprehensive analysis of household survey expectations shows that beliefs about future housing prices are adjusted too sluggishly, that housing price growth expectations covary positively with market valuation while actual housing price growth covaries negatively with market valuation, and that housing price growth expectations initially underreact and subsequently overreact to observed price increases. These deviations can be generated by weak forms of capital gain extrapolation as an equilibrium outcome. Additionally, such belief formation connects the secular decline in natural rates of interest with higher volatility of housing prices, as observed in a number of advanced economies including the United States. By embedding this belief formation into a New Keynesian model with a lower-bound constraint on nominal interest rates, the chapter shows that lower average natural rates of interest increase the volatility of housing prices and thereby the volatility of the natural rate. In the presence of capital gain extrapolation, the optimal inflation target increases considerably in response to a fall in the average natural rate. This is due to the increased volatility of the natural rate and cost-push shocks, which jointly cause the lower bound on the nominal rate to become more restrictive. In constrast, under rational expectations, optimal monetary policy would prescribe an average inflation rate close to zero at all levels of the natural rate of interest.

Chapter 3, which is joint work with Joachim Jungherr, Matthias Meier, and Immo Schott, studies the role of debt maturity in the transmission of monetary policy. Novel empirical evidence shows that firms' investment is more responsive to monetary policy when a higher fraction of their debt matures. After a tightening of monetary policy, investment, borrowing, sales, and employment all fall by more for firms with high shares of maturing debt. The chapter develops a heterogeneous firm New Keynesian model with financial frictions and endogenous debt maturity to understand the macroeconomic implications of this result. Debt maturity matters for monetary policy because of roll-over risk and debt overhang. Firms with more maturing debt have larger roll-over needs and are therefore more exposed to fluctuations in the real interest rate. Moreover, these firms also have higher default risk and therefore react more strongly to changes in the real burden of outstanding nominal debt. The model generates the rich heterogeneity in firm financing choices found in the data, including the heterogeneity in debt maturity, and it rationalizes the empirical evidence that firms with more maturing debt respond more strongly to changes in interest rates. Compared to existing models without long-term debt or without heterogeneity in debt maturity, the considered model implies larger aggregate effects of monetary policy. The maturity of debt and its distribution across firms are shown to be key for this result.

Chapter 4, which is joint work with Chiara Osbat and Luca Dedola, studies firms' price setting responses to changes in the corporate tax rate. A model of a currency area shows that individual firms, located in small open economies with minimal trading frictions, optimally increase prices in response to higher corporate tax rates if these firms have market power. Empirically, the response of product-level retail prices to tax changes is estimated using variation in tax rates across time and space in Germany, where municipalities set the local business tax once a year. Based on 1,058 tax changes between 2013 and 2017, a one percentage point tax increase is estimated to result in a 0.4% increase in firms' retail prices on average. This finding suggests that manufacturers may in fact use their market power to shield profits from corporate taxes by adjusting prices. The chapter also explores various dimensions of heterogeneity in pass-through, including producer size, market shares, and retail store types. While producer heterogeneity does not seem to matter, the significant pass-through of corporate taxes to consumer prices is mostly due to price changes in supermarkets and hypermarkets. The findings imply that not only are the redistributive effects of corporate taxation ambiguous, but aggregate changes in corporate taxes may have also have implications for monetary policy through their effect on consumer prices.

Chapter 1

Monetary Policy, Markup Dispersion, and Aggregate TFP*

Motivated by empirical evidence that monetary policy affects aggregate TFP, we study the role of markup dispersion for monetary transmission. Empirically, we show that the response of markup dispersion to monetary policy shocks can account for a significant fraction of the aggregate TFP response in the first two years after the shock. Analytically, we show that heterogeneous price rigidity can explain the response of markup dispersion if firms have a precautionary price setting motive, which is present in common New Keynesian environments. We provide empirical evidence on the relationship between markups and price rigidity in support of this explanation. Finally, we study the mechanism and its implications in a quantitative model.

1.1 Introduction

We revisit one of the long-standing questions in macroeconomics: What are the channels through which monetary policy affects real economic outcomes? Our paper is motivated by empirical evidence that monetary policy shocks have sizable effects on measured aggregate productivity.[1] A potential explanation for fluctuations in measured aggregate TFP is changing resource misallocation across firms. The TFP-misallocation link has been widely studied in the macro-development literature (e.g., Hsieh and Klenow, 2009), and is well understood in the New Keynesian literature. While in New Keynesian models, misallocation is commonly captured by price dispersion, our preferred empirical measure of misallocation is dispersion in markups across firms. Markup dispersion is price dispersion when controlling for differences in marginal costs across firms.

We study the role of markup dispersion for monetary transmission by asking

*Joint work with Matthias Meier.

[1]Using US data, we document that monetary policy shocks lower measured aggregate productivity, which reconfirms the evidence in Evans and Santos (2002), Christiano, Eichenbaum, and Evans (2005), Moran and Queralto (2018), Garga and Singh (2021), and Jordà, Singh, and Taylor (2020).

two questions: First, does markup dispersion respond to monetary policy shocks? Using US data, we document a significant response of markup dispersion, which can account for a significant fraction of the aggregate TFP response up to two years after the shock. Second, what explains the response of markup dispersion? We show analytically that heterogeneity in price setting frictions – in an otherwise standard New Keynesian framework – can explain the response of markup dispersion. The fundamental reason is that firms with stickier prices have a stronger precautionary price setting motive. This channel has testable implications, which, as we show, are supported empirically. Finally, we study the mechanism and its implications in a quantitative model.

We estimate the response of markup dispersion to monetary policy shocks based on quarterly balance-sheet data and high-frequency identified monetary policy shocks. A central contribution of this paper is to show that the dispersion of markups across firms (within industries) significantly increases after contractionary monetary policy shocks and decreases after expansionary monetary policy shocks. The response is persistent and peaks about two years after the shock. We establish this empirical pattern for a host of markup measures, following, amongst others, De Loecker and Warzynski (2012) and Gutiérrez and Philippon (2017). To translate the estimated response of markup dispersion into an aggregate TFP response, we follow Hsieh and Klenow (2009) and Baqaee and Farhi (2020). The response of markup dispersion implies a response in aggregate TFP between -0.2% and -0.4% two years after a one standard deviation contractionary monetary policy shock. For comparison, the directly estimated empirical response of utilization-adjusted aggregate TFP is -0.4% at a two-year horizon. At more distant horizons, markup dispersion accounts for a decreasing fraction of the aggregate TFP response.

Our evidence sheds new light on the TFP effects of monetary policy. Strikingly, the estimated response of markup dispersion cannot be explained by a large class of New Keynesian models, at least when solved with standard perturbation methods. In many New Keynesian models, including medium-scale models (e.g., Christiano et al., 2005) and models with heterogeneous price rigidity (e.g., Carvalho, 2006), markup dispersion does not respond to monetary policy shocks up to a first-order approximation around the deterministic steady state. In the second-order approximation, markup dispersion responds, but counterfactually increases in response to both positive and negative shocks. In models with trend inflation (e.g., Ascari and Sbordone, 2014), markup dispersion decreases after contractionary and increases after expansionary monetary shocks, which contradicts our empirical evidence.

What can explain the response of markup dispersion to monetary policy shocks instead? We propose a novel mechanism that arises from heterogeneity in the severity of price setting frictions across firms. A sufficient condition for higher markup dispersion after a monetary tightening is that firms with higher markups have lower pass-through from marginal costs to prices, i.e., relatively strong price setting frictions. A contractionary monetary shock that lowers marginal costs increases the relative markup of low pass-through firms, which increases markup dispersion. Analogously, expansionary monetary shocks that raise marginal costs will lower markup dispersion. We show that a negative correlation between firm-level markup and pass-through can arise endogenously from heterogeneity in price-setting frictions. 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. The intuition for this negative correlation is a precautionary price setting motive. The firm profit function in the common New Keynesian environment is asymmetric, i.e., it penalizes markups below more than markups above the static optimal one. A higher reset markup provides insurance against low profits before the next price adjustment opportunity (Calvo/Taylor), or lowers the expected costs of future price re-adjustments (Rotemberg/Barro).[2] To summarize, heterogeneous price-setting frictions imply markup dispersion and hence TFP effects of monetary policy. Importantly, precautionary price setting is absent in the deterministic steady state. By extension, our transmission mechanism is absent in model with heterogeneous price-setting frictions when solved around the deterministic steady state.

We empirically test two implications of this transmission mechanism. First, precautionary price setting implies that firms with stickier prices charge higher markups. Second, the markups of firms with stickier prices should increase by relatively more. A caveat is that we do not observe firm-specific price adjustment frequencies. Instead, we capture variation in price adjustment frequencies across firms using price adjustment frequencies in five-digit industries together with the firm-specific sales composition across industries. We find that firms with stickier prices indeed have higher markups on average and increase their markups by more after monetary policy shocks. These two results hold when controlling for two-digit sector fixed effects, firm size, leverage, and liquidity.

Finally, we study the mechanism and its implications in a quantitative New Keynesian model with heterogeneous price rigidity. To capture precautionary price setting, we use non-linear solution methods to solve the model dynamics around the stochastic steady state, to which the economy converges in the presence of uncertainty but absent of shocks. We find that indeed firms with stickier prices set higher markups on average, and monetary policy shocks raise markup dispersion. Quantitatively, a one standard deviation contractionary monetary policy shock lowers aggregate TFP by -0.34%. We use the model to study two implications of our mechanism. Whereas a contractionary monetary shock increases aggregate markups in many New Keynesian models, the empirically estimated responses of aggregate markups in Nekarda and Ramey (2020) have the opposite sign. In our model, the aggregate markup falls if contractionary monetary shocks lower aggregate TFP sufficiently strongly. This argument extends to sector or firm-level markups if price rigidities are heterogeneous within sectors or firms such that sector or firm-level TFP responds to monetary policy. We further analyze the effectiveness of monetary policy when the endogenous TFP effects are ignored by

[2]Relatedly, in a setup with homogeneous price setting frictions Fernandez-Villaverde, Guerron-Quintana, Kuester, and Rubio-Ramirez (2015) study precautionary price setting as a channel through which higher uncertainty leads to higher markups.

the monetary authority. If the monetary authority attributes all TFP fluctuations to technology shocks, interest rates are adjusted less aggressively and monetary policy shocks lead to larger GDP fluctuations.

Related literature. This paper is closely related to four branches of the literature. First, a growing literature studies the positive and normative implications of heterogeneous price rigidity, see, e.g., Aoki (2001), Carvalho (2006), Nakamura and Steinsson (2010), Eusepi, Hobijn, and Tambalotti (2011), Carvalho and Schwartzman (2015), Castro Cienfuegos and Loria (2017), Pasten, Schoenle, and Weber (2020), and Rubbo (2020). 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 exogenously 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 effects of monetary policy, e.g., Evans and Santos (2002), Christiano et al. (2005), Comin and Gertler (2006), Moran and Queralto (2018), Garga and Singh (2021), 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. In terms of alternative explanations, 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 (2021) show that R&D investment falls after monetary policy shocks, which may ultimately lower productivity. However, it is unclear whether the R&D response can explain a large response of aggregate productivity at short horizons. For example, Comin and Mestieri (2018) show that recent technologies are adopted with an average lag of five years. Conversely, price rigidities are a more natural candidate for the effects at shorter horizons.

Third, our paper relates to a literature on the relation between inflation and price dispersion. Whereas we show that contractionary monetary policy shocks raise markup dispersion, Nakamura, Steinsson, Sun, and Villar (2018) document flat price dispersion across periods of high and low inflation since the 1970s. This suggests that long-lived changes in inflation have different effects than short-lived monetary policy shocks. For example, when trend inflation increases managers may schedule more frequent meetings to discuss price changes (Levin and Yun, 2007), while monetary policy shocks are less likely to trigger such responses.

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 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- versus long-run changes in interest rates on allocative efficiency seem to differ in sign. Whereas we show that short-run expansionary monetary policy decreases misallocation, Gopinath, Kalemli-Özcan, Karabarbounis, and Villegas-Sanchez (2017) show that, in the case of Southern Europe, persistently lower interest rates have increased misallocation. Relatedly, Oikawa and Ueda (2018) study the long-run effects of nominal growth through reallocation across heterogeneous firms.

1.2 Evidence on markup dispersion and TFP

In this section, we present novel empirical evidence that monetary policy shocks increase the markup dispersion across firms. We further show that aggregate TFP falls after monetary policy shocks and that a sizable share of this response can be accounted for by the response of markup dispersion.

1.2.1 Data

Firm-level markups. We use quarterly balance sheet data of publicly-listed US firms from Compustat. We estimate markups through a variety of methods. Our baseline method is the ratio estimator pioneered by Hall (1986) and more recently used in De Loecker and Warzynski (2012), De Loecker, Eeckhout, and Unger (2020), Flynn, Traina, and Gandhi (2019) and Traina (2020). We further consider markups using the accounting profits and user cost approaches in Gutiérrez and Philippon (2017), Basu (2019) and Baqaee and Farhi (2020).

The ratio estimator of the markup can be obtained from the cost minimization problem. With a flexible input V_{it}, the markup μ_{it} of firm i in quarter t can be computed as

$$\mu_{it} = \frac{\text{output elasticity of } V_{it}}{\text{revenue share of } V_{it}}. \tag{1.1}$$

We assume that firms in the same two-digit-industry and quarter have a common output elasticity. All our subsequent empirical analysis focuses on differences of firm-level log markups from their industry-quarter average. Under our assumption, these markup differences do not depend on the output elasticities. Hence, our empirical results are not affected by challenges to identify output elasticities from revenue data, as recently emphasized by Bond, Hashemi, Kaplan, and Zoch (2021).[3] By controlling for industry-quarter fixed effects in log markups, we also difference out industry and time-specific characteristics such as differences in competitiveness and production technology.

Formally, we define differences of firm-level log markups from their industry-time average as $\hat{\mu}_{it} \equiv \log \mu_{it} - \frac{1}{N_{st}} \sum_{j \in \mathcal{J}_{st}} \log \mu_{jt}$, where $\mathcal{J}_{st}$ is the set of firms j in industry s, quarter t, and N_{st} is the cardinality of $\mathcal{J}_{st}$. Following De Loecker et al. (2020) we assume firms produce output using capital and a composite input of

[3] Our baseline approach assumes the ratio estimator to be valid in principle. This excludes the case when the input is not perfectly flexible, or when its choice affects demand, see Bond et al. (2021). We also consider non-ratio estimators of markups, see Section 2.4.

labor and materials, with the latter the flexible factor. We estimate the revenue share as the firm-quarter-specific ratio of costs of goods sold (`cogsq` in Compustat) to sales (`saleq`).

We further consider a host of alternative markup estimation methods in Section 2.4 below. First, we construct (non-ratio estimator) markups through an accounting profit approach and a user cost approach, following Gutiérrez and Philippon (2017) and Baqaee and Farhi (2020). Second, following Traina (2020), we add selling, general and administrative expenses (SGA) to the costs of goods sold in the baseline markup measure. Third, we estimate a four-digit industry-specific translog production technology, which implies variation in output elasticities within industry and time. Fourth, we estimate four-digit industry-quarter specific output elasticities through cost shares.

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 reported only 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 1.A.1 provides more details and summary statistics in Table 1.A.1. Our results are robust to alternative data treatments as we discuss toward the end of this section.

Monetary policy shocks. Using high-frequency data of federal fund future prices, 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.[4] 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

$$\varepsilon_\tau^{\mathrm{MP}} = \mathrm{f}_{\tau+\Delta\tau^+} - \mathrm{f}_{\tau-\Delta\tau^-}. \tag{1.2}$$

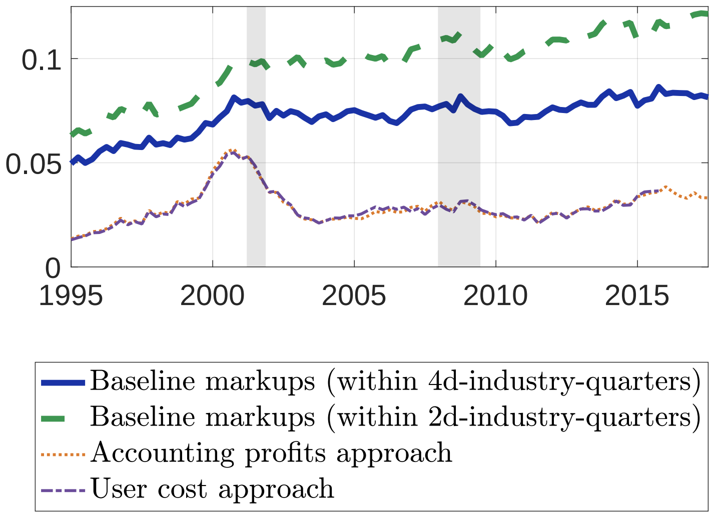
To aggregate the shocks to quarterly frequency, we follow Ottonello and Winberry (2020). We assign daily shocks fully to the current quarter if they occur on the first day of the quarter. If they occur within the quarter, we partially assign the shock to the subsequent quarter. This procedure weights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

$$\varepsilon_t^{\mathrm{MP}} = \sum_{\tau\in\mathcal{D}(t)} \phi(\tau)\varepsilon_\tau^{\mathrm{MP}} + \sum_{\tau\in\mathcal{D}(t-1)} (1-\phi(\tau))\varepsilon_\tau^{\mathrm{MP}}, \tag{1.3}$$

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).

[4] 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).

Figure 1.1: Evolution of markup dispersion



Notes: The figure shows the evolution of markup dispersion for different markup measures from 1995Q1 to 2017Q3. Markup dispersion is the variance of log markups across firms, $\mathbb{V}_t(\hat{\mu}_{it})$, where $\hat{\mu}_{it}$ is the difference of a firm's log markup from the mean log markup across firms in the same industry-quarter. Baseline markups are constructed according to equation (1.1) assuming a common output elasticity for firms in the same 2d-industry-quarter. Further details on the accounting profits and user cost approaches are provided in Section 1.2.4.

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 meetings and conference calls.[5] Following Nakamura and Steinsson (2018), our baseline further excludes the apex of the financial crisis from 2008Q3 to 2009Q2.[6] The monetary policy shock series covers 1995Q2 through 2017Q3. We discuss alternative monetary policy shocks in Section 1.2.4. Table 1.A.2 in the Appendix reports summary statistics and Figure 1.A.1 (a) and (b) shows the shock series.

1.2.2 Markup dispersion

We estimate the response of markup dispersion to monetary policy shocks. Our baseline measure of markup dispersion is the cross-sectional variance $\mathbb{V}_t(\hat{\mu}_{it})$, where $\hat{\mu}_{it}$ denotes firm-level log markups in deviation from their respective industry-quarter mean. Recall that our baseline estimator of $\hat{\mu}_{it}$ does not depend on an estimator of the output elasticity under our assumption that firms within a two-digit industry-quarter have a common output elasticity. Figure 1.1 shows time

[5]Unscheduled meetings and conference calls often occur after adverse economic developments. Price changes around such meetings may reflect these developments, invalidating the identifying restriction. Our results remain broadly robust when including these meetings.

[6]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.

series of markup dispersion for our baseline ratio estimator within four-digit-industry-quarters, the same estimator but within two-digit-industry-quarters, and for markups based on account profits and user costs. Figure 1.A.1 (c) in the Appendix shows time series for further alternative markup dispersion, notably the ratio estimator when including SGA, the translog-based markups, and markups based on cost shares.

To estimate the effects of monetary policy shocks on markup dispersion, we use the following local projection for $h = 0, \dots, 16$ quarters and where y_t is markup dispersion:

$$y_{t+h} - y_{t-1} = \alpha^h + \beta^h \varepsilon_t^{\mathrm{MP}} + \gamma_0^h \varepsilon_{t-1}^{\mathrm{MP}} + \gamma_1^h (y_{t-1} - y_{t-2}) + u_t^h \tag{1.4}$$

The central empirical finding of this paper is shown in panel (a) of Figure 1.2, which plots the response of markup dispersion, captured by the estimates of coefficients β^h. The key finding is that markup dispersion increases significantly and persistently. The response of markup dispersion peaks at about two years after the shock and reverts back to zero afterwards. Whether we compute markup dispersion within two-digit or four-digit industry-quarters changes this result by little.

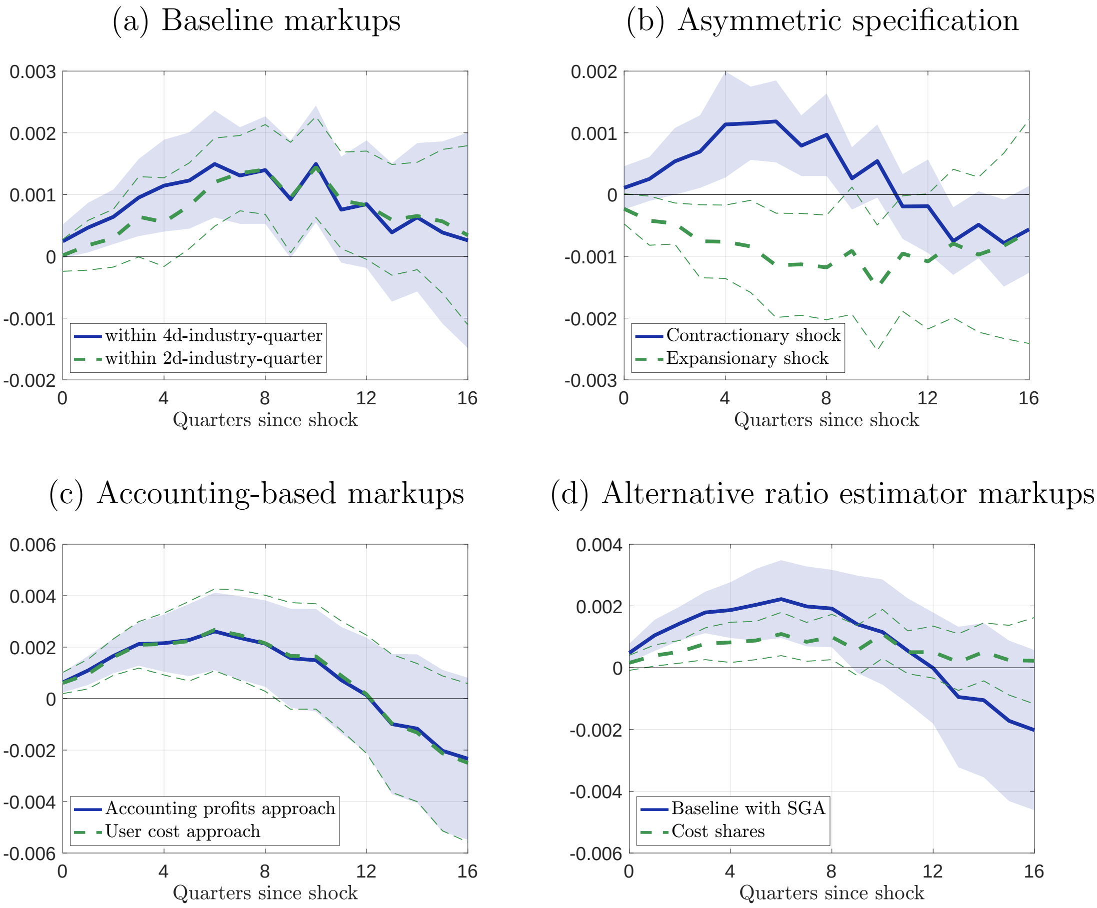
The specification of (1.4) implicitly assumes that the effects of monetary policy shocks are symmetric in the sign of the shock. However, in a large class of New Keynesian models, solved via a second-order approximation, markup dispersion increases in response to both positive and negative shocks, cf. Figure 1.H.5 in the Appendix. So to investigate whether markup dispersion responds asymmetrically to shocks of different sign, we separately estimate the separate effects of contractionary and expansionary monetary policy shocks. To be precise, we replace $\varepsilon_t^{\mathrm{MP}}$ by the two sign-dependent shocks in specification (1.4). Panel (b) of Figure 1.2 shows the sign-dependent responses of (within 4-digit industry-quarter) markup dispersion. The evidence suggests that the responses are indeed symmetric in shock sign. While contractionary monetary policy shocks significantly increase markup dispersion, expansionary shocks significantly lower markup dispersion. In addition, the estimated magnitudes are comparable across shock sign. The results in panel (a) and (b) prove robust in a large number of dimensions, including alternative measures of markups, as we discuss in Section 1.2.4.

1.2.3 Aggregate productivity

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

$$\Delta \log \mathrm{TFP}_t = -\frac{\eta}{2} \Delta \mathbb{V}_t(\log \mu_{it}) \; + \; \Big[\Delta \text{ exogenous productivity}\Big], \tag{1.5}$$

Figure 1.2: Responses of markup dispersion to monetary policy shocks



Notes: Panel (a) shows the responses of markup dispersion to a one standard deviation monetary policy shock, coefficients β^h in (1.4). Panel (b) shows the sign-dependent responses of markup dispersion to a one standard deviation contractionary and expansionary monetary policy shock, respectively. Panel (c) shows the response of markup dispersion using the *accounting profits* and *user cost approach*, respectively. Panel (d) shows the response of markup dispersion using the *baseline with SGA* approach that adds SGA to the costs of goods sold, and the *cost shares* approach that uses a ratio estimator of four-digit-industry-quarter-specific cost shares as output elasticities, and constructs markup dispersion within two-digit-industry-quarters. The shaded and bordered areas indicate one standard error bands based on the Newey–West estimator.

where η is the substitution elasticity between variety goods. The details of the derivation are provided in Appendix 1.E.1.[7] An increase in the variance of log markups by 0.01 lowers aggregate TFP by $\frac{\eta}{2}$%. To provide some intuition for this link, first suppose firms are homogeneous. Aggregate output is maximal for given aggregate inputs if all firms produce the same quantity, which implies equal markups across firms. If instead firms have 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

[7] In the calibrated New Keynesian model of Section 1.4, equation (1.5) closely matches the comovement of aggregate TFP and markup dispersion, *cf.* Figure 1.5 (b) and (f).

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 implied productivity response according to equation (1.5) and the estimated response of markup dispersion in Figure 1.2(a). 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 (1.4).[8] Panel (a) of Figure 1.3 shows that the responses of all three aggregate productivity measures are significantly and persistently negative. 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, a monetary policy shock of the same magnitude raises the federal funds rate by up to 30 basis points and lowers aggregate output by about 1% at a two-year horizon, see Figure 1.B.2 in the Appendix. However, aggregate factor inputs respond little and thus aggregate TFP accounts for 50–80% of the output response at a two-year horizon.

We compute the implied TFP response by multiplying the estimated response of markup dispersion with $-\frac{\eta}{2}\%$. Panel (b) of Figure 1.3 shows the implied response for $\eta = 6$, which corresponds to 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 of 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 1.B.1 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 (2021). Hence, there is scope for R&D to explain part of the aggregate TFP response. However, it is less clear 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 Mestieri (2018) estimate an average adoption lag of 5 years for recent technologies. A sluggish effect of R&D investment on aggregate productivity is consistent with the finding in Figure 1.3 (b) that markup dispersion accounts for a relatively small fraction of the TFP response 3–4 years after a monetary policy shock.

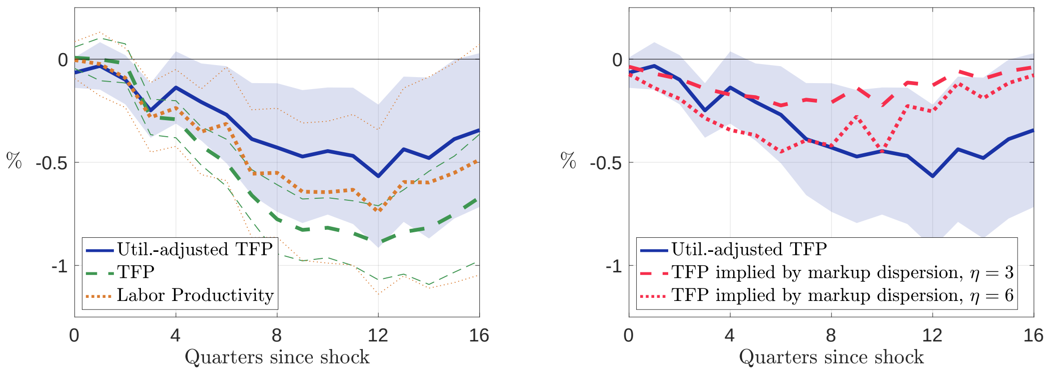
1.2.4 Robustness

Markup estimation. We investigate the robustness of our empirical findings by considering a host of alternative markup measures. Our baseline results are robust to using these alternative markups. First, we construct (non-ratio estimator)

[8] 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, $\Delta \ell$ the growth of hours worked plus growth in labor quality. Utilization-adjustment follows Basu, Fernald, and Kimball (2006). Labor productivity is real output per hour in the nonfarm business sector. Figure 1.A.1 (d) in the Appendix shows the different aggregate productivity time series.

Figure 1.3: 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 (1.5), 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.

markups through an accounting profit approach and a user cost approach, following Gutiérrez and Philippon (2017) and Baqaee and Farhi (2020). The accounting profit approach uses operating income after depreciation, which is sales (`saleq`) minus costs of goods sold (`cogsq`), selling, general and administrative expenses (`xsgaq`), and depreciation and amortization (`dpq`). We compute markups from these accounting profits via $(\text{accounting profit})_{it} = \left(1 - \mu_{it}^{-1}\right) \texttt{saleq}_{it}$.

For the user cost approach, we additionally subtract the firm's capital costs (excluding depreciation) from accounting profits as in Baqaee and Farhi (2020). We construct firm-level capital stocks k_{it} via a perpetual inventory method to property, plant and equipment, see Appendix 1.A.1. The user cost of capital is $r_t = r_t^f + RP_{jt} - (1 - \delta_{jt})\Pi_{jt+1}^K$, where r^f is the risk-free real rate, RP_j the industry-specific risk premium, δ_j the industry-specific BEA depreciation rate, and Π_j^K is the industry-specific growth in the relative price of capital, based on data in Gutiérrez and Philippon (2017).[9] In general, the size of capital costs relative to total costs is modest with an average of 3.2%. This may explain the small differences between the accounting profit and user cost approaches.

Second, we construct a ratio estimator which adds selling, general and administrative expenses (SGA) to the costs of goods sold, following Traina (2020). Third, we estimate a four-digit-industry-specific translog production technology, which implies firm-quarter-specific output elasticities as in De Loecker et al. (2020). We then compute markups by combining output elasticities with revenue shares according to equation (1.1). Fourth, we compute four-digit-industry-quarter-specific

[9]The Gutiérrez and Philippon (2017) user cost is at annual frequency; we divide through by four to arrive at a quarterly rate. The data from Gutiérrez and Philippon (2017) ends in 2015, so that the time sample of user cost approach markups is shorter.

cost shares to estimate output elasticities. Specifically, we follow De Loecker et al. (2020) and compute the industry-quarter median of costs of goods sold plus 3% of the capital stock (which approximates the user cost of capital by an annual rate of 12% that includes risk premium and depreciation) divided by sales. This is a valid estimator of the output elasticity if all factors are flexible.

Our results are robust to computing markups based on these alternative measures.[10] Figure 1.2 (c) shows the response of markup dispersion within four-digit-industry-quarters to monetary policy shocks when using the accounting profits and user cost approach. Figure 1.2 (d) shows the markup dispersion response when including SGA as well as for the cost share approach. In the Appendix we show additional results. Figure 1.C.1 shows the responses of all alternative markup dispersion measures within two-digit- and four-digit-industry quarters. Figure 1.C.2 shows the responses of all markup dispersion measures conditional on the sign of the monetary policy shock.

Firm-level data treatment. We show the robustness of our results under alternative data treatments. 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 1.C.3 shows that markup dispersion robustly increases after contractionary monetary policy shocks. Figure 1.C.4 shows the responses of markup dispersion remain symmetric in the sign of the monetary policy shock. 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 1.C.5 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. High-frequency future price changes may release private central bank information about the state of the economy. To control for such information effects we employ two alternative strategies. First, following Miranda-Agrippino and Ricco (2021), we regress daily monetary policy shocks on internal Greenbook forecasts and revisions for output growth, inflation, and unemployment. Second, following Jarociński and Karadi (2020), we discard daily monetary policy shocks if the associated high-frequency change in the S&P500 moves in the same direction. To address the concern that

[10]For the comparability of our results across markup measures, we include only firms in the robustness checks for which the baseline markup is non-missing after the data treatment steps. Additionally we trim the alternative markups at the 1% and 99% quantiles of the quarterly markup distributions.

unconventional monetary policy may drive our result, we set daily monetary policy shocks at Quantitative Easing (QE) announcements to zero. Figure 1.C.6 in the Appendix shows the response of markup dispersion for all monetary policy shock series. Figure 1.C.7 shows the sign-dependent responses of markup dispersion to monetary policy shock. Figure 1.C.8 in the Appendix shows the responses of aggregate productivity for all monetary policy shock series.

Great Recession. We exclude the apex of the Great Recession from 2008Q3 to 2009Q2 in our baseline estimations. However, our results are robust to using the Pre-Great Recession period until 2008Q2, see panels (d) and (e) of Figures 1.C.3 and 1.C.4 in the Appendix.

LP-IV. To revisit our main results with the LP-IV method, we replace the shocks $\varepsilon_t^{\mathrm{MP}}$ by the quarterly change in the one-year treasury rate and use $\varepsilon_t^{\mathrm{MP}}$ as an instrument. Figure 1.C.9 (a) and (b) in the Appendix shows that our results are robust to the LP-IV method.

Proxy SVAR. Additionally, we revisit our main results through a proxy SVAR model following Gertler and Karadi (2015).[11] Figure 1.C.10 in the Appendix shows the responses to monetary policy shocks in a VAR, including the one-year rate, (log) industrial production, (log) CPI, the excess bond premium of Gilchrist and Zakrajšek (2012), (log) TFP and the baseline measure of markup dispersion (within four-digit industry-quarters). At a horizon between 1 and 5 quarters after the shock, the responses of TFP and markup dispersion are similar to our local projection results.

TFP measurement. Hall (1986) shows that the Solow residual is misspecified in the presence of market power. Hall shows that the correct Solow weights are not the income share for capital w_{kt} and labor $1-w_{kt}$, but instead $\mu_t w_{kt}$ and $1-\mu_t w_{kt}$, where μ_t is the aggregate markup. We 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 1.C.11 (a) in the Appendix shows that the TFP response is barely different from Figure 1.3 (a). In response to expansionary monetary shocks, Figure 1.C.12 shows a significant increase of TFP, while the response to contractionary shocks is insignificant. We further investigate whether measurement error in quarterly TFP data is responsible for the effects of monetary policy. This problem was flagged for defense spending shocks by Zeev and Pappa (2015). We follow them in recomputing TFP using measurement error corrected quarterly GDP from Aruoba, Diebold, Nalewaik, Schorfheide, and Song (2016). Figure 1.C.11 (b) shows that measurement error corrected TFP also falls after monetary policy shocks. We further show that Fernald's (2014) investment-specific and consumption-specific TFP

[11] In contrast to the proxy SVAR model, both our baseline LP approach in (1.4) and the LP-IV approach are robust to non-invertibility, see Plagborg-Møller and Wolf (2021).

significantly falls after contractionary monetary policy shocks, see Figure 1.C.11 (c) and (d).

1.3 Heterogeneous price setting frictions

In this section, we characterize a novel mechanism through which firm heterogeneity in price setting frictions may explain why markup dispersion increases in response to contractionary monetary policy shocks, and decreases after expansionary ones. In addition, we provide empirical evidence in support of this mechanism, and discuss alternative mechanisms.

1.3.1 Sufficient condition

We first propose a sufficient condition for monetary policy shocks, which lower real marginal costs, to increase the dispersion of markups across firms. 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. Let pass-through from marginal cost to price be defined as

$$\rho_{it} \equiv \frac{\partial \log P_{it}}{\partial \log X_t}. \tag{1.6}$$

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 1.1 *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 1.E.2.

Contractionary monetary policy shocks that lower real marginal costs increase the dispersion of markups if firms with higher markups have lower pass-through. While we focus on monetary policy shocks in this paper, in principle any shock that lowers real marginal costs will raise markup dispersion as long as markups and pass-through are negatively correlated across firms.

1.3.2 Precautionary price setting

We next show that firm-level heterogeneity in the severity of various price-setting frictions may explain a negative correlation between firm-level pass-through and markup. It follows from Proposition 1.1 that heterogeneous price-setting frictions can explain why contractionary monetary policy shocks raise markup dispersion.

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

$$\max_{\{P_{it+j}\}_{j=0}^{\infty}} \mathbb{E}_t \sum_{j=0}^{\infty} \beta^t \left[\left(\frac{P_{it+j}}{P_{t+j}} - X_{t+j} \right) \left(\frac{P_{it+j}}{P_{t+j}} \right)^{-\eta} Y_{t+j} - \text{adjustment cost}_{it+j} \right] \tag{1.7}$$

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 opportunity (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

$$P_{it}^* = \frac{\eta}{\eta - 1} P_t X_t \frac{\mathbb{E}_t \left[\sum_{j=0}^{\infty} \beta^j \theta_i^j \frac{X_{t+j}}{X_t} \left(\frac{P_{t+j}}{P_t} \right)^{\eta} \frac{Y_{t+j}}{Y_t} \right]}{\mathbb{E}_t \left[\sum_{j=0}^{\infty} \beta^j \theta_i^j \left(\frac{P_{t+j}}{P_t} \right)^{\eta-1} \frac{Y_{t+j}}{Y_t} \right]}, \tag{1.8}$$

and we denote the associated markup by μ_{it}^*. To isolate the role of uncertainty in price setting, we focus on the dynamics around the stochastic steady state, which is described by the unconditional means $(\bar{P}, \bar{X}, \bar{Y})$. The following proposition characterizes the precautionary upward price-setting bias – relative to the frictionless environment – as a function of θ_i, and establishes a condition under which firms with lower pass-through set higher markups.

Proposition 1.2 *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 1.E.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 hence less than 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 $\text{cost}_{it} = \frac{\phi_i}{2}\left(\frac{P_{it}}{P_{it-1}} - 1\right)^2$. The first-order condition for P_{it} is

$$\left[(1-\eta)\frac{P_{it}}{P_t} + \eta X_t\right]\left(\frac{P_{it}}{P_t}\right)^{-\eta} Y_t = \phi_i \left(\frac{P_{it}}{P_{it-1}} - 1\right)\frac{P_{it}}{P_{it-1}} - \phi_i \beta \mathbb{E}_t\left[\left(\frac{P_{it+1}}{P_{it}} - 1\right)\frac{P_{it+1}}{P_{it}}\right]. \tag{1.9}$$

The following proposition summarizes our analytical results.

Proposition 1.3 *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* (1.9) *around $\phi_i = 0$, it holds that*

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

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 1.E.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 1.F.

1.3.3 Empirical evidence for the mechanism

We corroborate the mechanism by considering two testable implications. First, firms with higher markups adjust prices less frequently. Second, monetary policy shocks increase the relative markup of firms that adjust prices less frequently. We show that both implications are supported empirically.

For the subsequent empirical analysis, we use data on price adjustment frequencies together with the data described in Section 1.2. We observe average price adjustment frequencies over 2005–2011 for five-digit industries, computed in Pasten et al. (2020) from PPI micro data.[12] We further use the Compustat segment files, which provide sales and industry codes 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 we expect will bias our subsequent regression coefficients toward zero because of attenuation bias.[13] 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 25% of firms. For the remaining firms, we use the price adjustment frequency of the five-digit industry they operate in.[14] More details are provided in Appendix 1.A.4. To measure price rigidity, we consider both the price adjustment frequency and the implied price duration, defined as $-1/\log(1 - \text{price adjustment frequency})$.

Testable implication 1: Firms with stickier prices charge higher markups. We provide empirical evidence that firms with stickier prices tend to charge higher markups. 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 1.1 show that firms, which have more rigid prices than other firms in the same two-digit industry, charge markups significantly above the industry average. The correlation is statistically

[12] We thank Michael Weber for generously sharing the data with us.

[13] A sufficient condition for downward bias is that the error in the measured firm-specific price adjustment frequencies is independent of the true unobserved firm-specific price adjustment frequencies.

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

Table 1.1: Markups and price stickiness

(a) Regressions of markups on implied price duration

	log(Markup)			
	Baseline		Accounting profits	User cost approach
Implied price duration	0.0537 (0.0180)	0.0472 (0.0155)	0.00706 (0.00300)	0.00882 (0.00344)
Additional controls	No	Yes	Yes	Yes
2-digit industry FE	Yes	Yes	Yes	Yes
Observations	3857	3857	3806	3798
Adjusted R^2	0.145	0.228	0.237	0.184

(b) Regressions of markups on price adjustment frequency

	log(Markup)			
	Baseline		Accounting profits	User cost approach
Price adjustment frequency	-0.391 (0.0999)	-0.336 (0.0860)	-0.0501 (0.0199)	-0.0600 (0.0214)
Additional controls	No	Yes	Yes	Yes
2-digit industry FE	Yes	Yes	Yes	Yes
Observations	3857	3857	3806	3798
Adjusted R^2	0.151	0.231	0.237	0.184

Notes: Regressions of firm-level markup on firm-level price adjustment frequency and implied price duration, respectively. The regressions with additional controls include firm-level size, liquidity, and leverage as regressors. Standard errors are clustered at the two-digit industry level and shown in parentheses.

significant for both implied price duration and price adjustment frequency as measures of price rigidity. While this correlation is consistent with precautionary price setting, it may reflect omitted factors. In columns (2) and (4) we control for firm-specific size, leverage, and liquidity, all averages over 2005–2011. The conditional correlations remain of the same sign and statistically significant at the 1% level. In Table 1.1 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 1.D.1 in the Appendix. Note that we have not considered four-digit industry FE, because for many firms our measure of rigidity is based on the five-digit industry average, which limits the variation in rigidity measures within four-digit industries.

Testable implication 2: Monetary policy shocks increase the relative markups of firms with stickier prices. We investigate whether contractionary monetary policy shocks increase the relative markup of firms with stickier prices. 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 (all in deviation from their firm-level mean). Our selection of controls is motivated by recent work in Ottonello and Winberry (2020), who study the transmission of monetary policy shocks through financial constraints. We use the panel local projection

$$y_{it+h} - y_{it-1} = \alpha_i^h + \alpha_{st}^h + B^h Z_{it}\varepsilon_t^{\mathrm{MP}} + \Gamma^h Z_{it} + \gamma^h(y_{it-1} - y_{it-2}) + u_{it}^h \qquad (1.10)$$

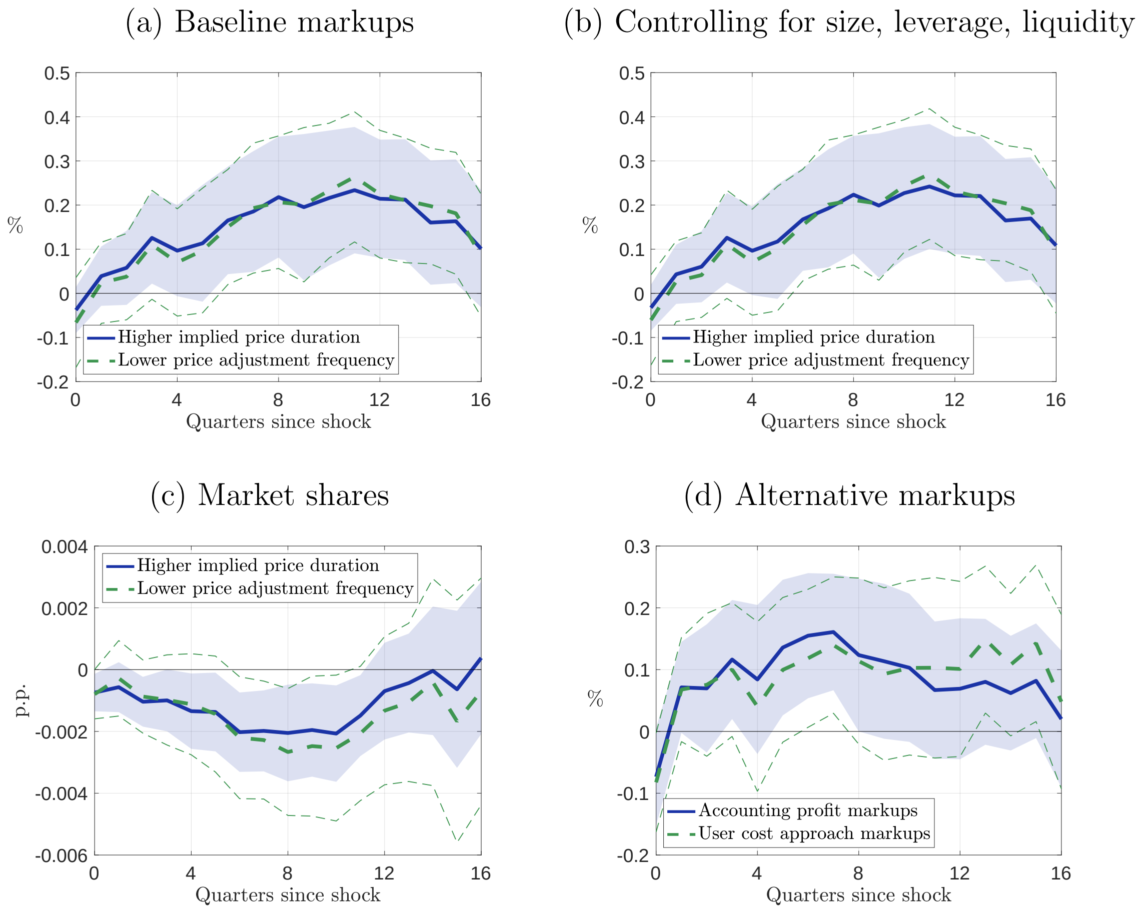
for $h = 0, \ldots, 16$ quarters, in which we include two-digit-industry-time and firm fixed effects. To focus on the within-industry variation in the interaction between monetary policy shock and price rigidity, we subtract the corresponding two-digit industry mean from the measure of price rigidity. The main coefficients of interest are the coefficients in $\{B^h\}$ associated with price rigidity. These capture the relative markup increase for firms with stickier prices. Figure 1.4 shows the results. The markups of firms with stickier prices increase by significantly more after monetary policy shocks.[15] 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 1.4. We additionally investigate the relative size response of firms with stickier prices. In particular, we consider firm-level sales market shares at the two-digit-industry-quarter level. As a relative increase in markup implies relatively lower demand, we expect that firms with stickier prices become relatively smaller after contractionary monetary policy shocks. Indeed, we find that firms with stickier prices lose market share after contractionary monetary policy shocks, as can be seen in panel (c) of Figure 1.4.[16]

Robustness. Our findings are robust along various dimensions, similar to Section 1.2.4. Figure 1.4 (d) shows the differential markup response of firms with more sticky prices based on the accounting profits and user costs approach. We show further robustness checks in Appendix 1.D.

[15] Driscoll–Krayy standard errors yield almost the same confidence bands as in Figure 1.4.

[16] The response of dispersion in firm-level market shares increases after monetary policy shocks, similar to markup dispersion, see Appendix Figure 1.C.1 (f).

Figure 1.4: Relative markup and market share responses of firms with stickier prices



Notes: The figures show the relative responses 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 (1.10). In panel (a), Z_{it} contains only price stickiness. In panels (b)–(d), Z_{it} also contains lagged log assets, leverage, and liquidity. Panel (d) uses implied price duration as measure of price rigidity. The shaded and bordered areas indicate 90% error bands two-way clustered by firm and quarter.

1.3.4 Alternative mechanisms

A key condition to explain the response of markup dispersion to monetary policy shocks is a negative correlation between firm-level markups and pass-through (Proposition 1). We show that firm heterogeneity in price setting frictions can explain this correlation and we provide empirical evidence in support of this explanation. However, this does not preclude other mechanisms. In the following, we discuss three alternative mechanisms.

First, a non-isoelastic demand system as proposed by Kimball (1995) can explain a negative correlation between markup and pass-through and thus the response of markup dispersion.[17] Indeed, recent work by Baqaee, Farhi, and Sangani (2021) shows that under Kimball preferences (also applied, e.g., by Edmond, Midrigan, and Xu, 2021), firms with a higher market share may have higher markups and lower pass-through. Even in the absence of heterogeneous price setting frictions, this environment can qualitatively explain our empirically estimated response of markup dispersion to monetary policy shocks. Second, a negative correlation between markup and pass-through can arise in an environment with oligopolistic competition and different elasticities of substitution across and within sectors, as proposed by Atkeson and Burstein (2008). Third, heterogeneity in pass-through across firms can arise from financial frictions. For example, markup dispersion may increase if contractionary monetary policy shocks increase by more the financing costs of firms with lower markups.

1.4 Quantitative example

In this section, we investigate the transmission mechanism and its implications in a New Keynesian model with heterogeneous price rigidity.

1.4.1 Model setup

Our model setup builds on Carvalho (2006), Kara (2015), and Gorodnichenko and Weber (2016). We discuss the model only briefly and relegate a formal description to Appendix 1.G. 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.

The economy is populated by five types of monopolistically competitive intermediate goods firms. There is an equal mass of firms of each type. All firms produce differentiated output goods with the same linear technology in labor. The only ex-ante difference across firms is the exogenous price adjustment probability $1 - \theta_k$, which is specific to type k. Firms set prices to maximize the value of the firm to the households. 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. 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, $\nu_t \sim \mathcal{N}(0, \sigma_\nu^2)$.

[17] The evidence for Kimball-type demand curves is mixed, see Klenow and Willis (2016).

Table 1.2: Calibration

Parameter		Value	Source/Target
Discount factor	β	$1.03^{-1/4}$	Risk-free real 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	0.00411	30bp nominal rate increase
Frisch elasticity of labor supply	φ	0.1135	Relative hours response
Distribution of price adjustment frequencies			
Firm type k		Share	Price flexibility $(1-\theta_k)$
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 (price flexibility) is chosen to match the within-sector distribution reported in Gorodnichenko and Weber (2016).

1.4.2 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, Eichenbaum, and Trabandt (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. For the five types of firms, we calibrate θ_k for $k = 1, \ldots, 5$ 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. 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 1.B.2 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 response of utilization-adjusted TFP relative to the hours response, computed as the mean of $\frac{1\ -\ \text{response of util-adj. TFP in \%}}{1\ -\ \text{response of hours in \%}} - 1$ 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 $\varphi = 0.1135$, which is low compared to the macroeconomics literature, but which is within the range of empirical estimates surveyed by Ashenfelter, Farber, and Ransom (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 1.B.2. This yields $\sigma_\nu = 0.00411$.

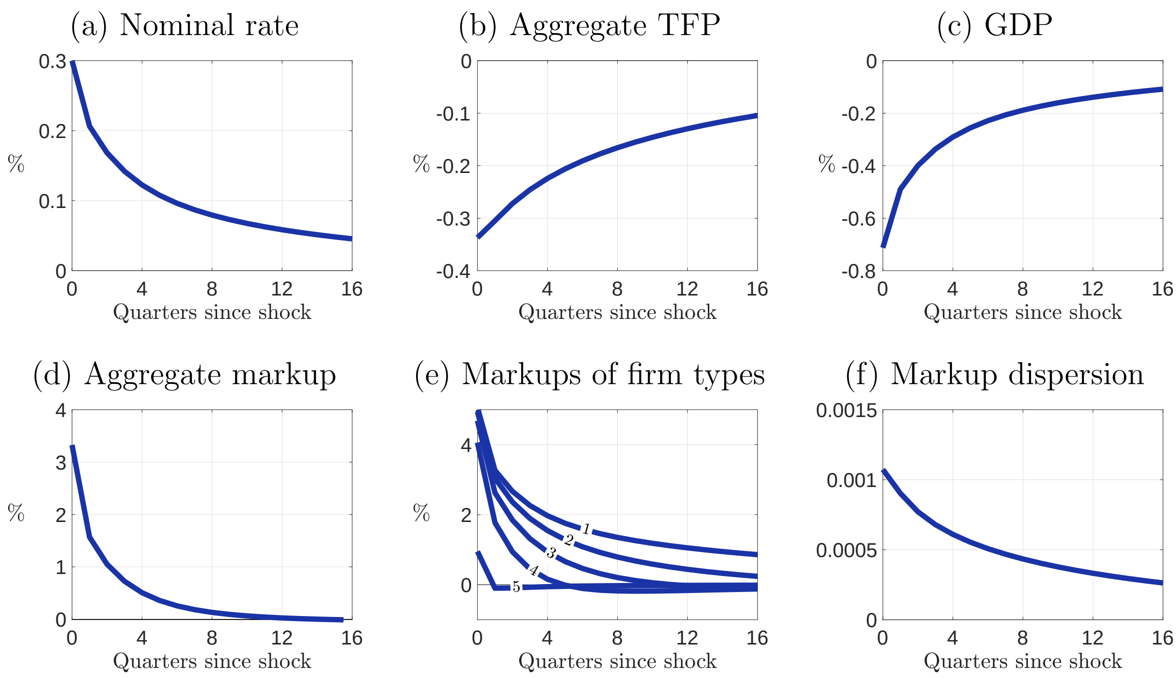
For markup dispersion to arise from precautionary price setting, it is important to use an adequate model solution technique. We rely on local solution techniques, but, importantly, solve the model around its stochastic steady state. Whereas markup are the same across firms in the deterministic steady 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. In the stochastic steady state, precautionary price setting has large effects. Firms with the most rigid prices have 11.5% higher markups than firms with the most flexible prices.[18] As follows from Proposition 1.1, the negative correlation between markups and pass-through implies that contractionary monetary policy shocks increase markup dispersion and lower aggregate TFP.

1.4.3 Results

Figure 1.5 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 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 1.2, and the model explains 60% of the peak empirical response in utilization-adjusted TFP, see Figure 1.3. In addition, the responses show the frequency composition effect described by

[18] The only source of uncertainty in the stochastic steady state are monetary policy shocks. In principle, considering multiple shocks may increase or decrease the precautionary price setting motive. As Proposition 2 shows, precautionary price setting depends on the co-movement of prices, marginal costs, and aggregate demand. A sufficient condition for precautionary price setting is that all covariances between these variables are positive. This is commonly satisfied by monetary policy shocks.

Figure 1.5: Model responses to monetary policy shocks



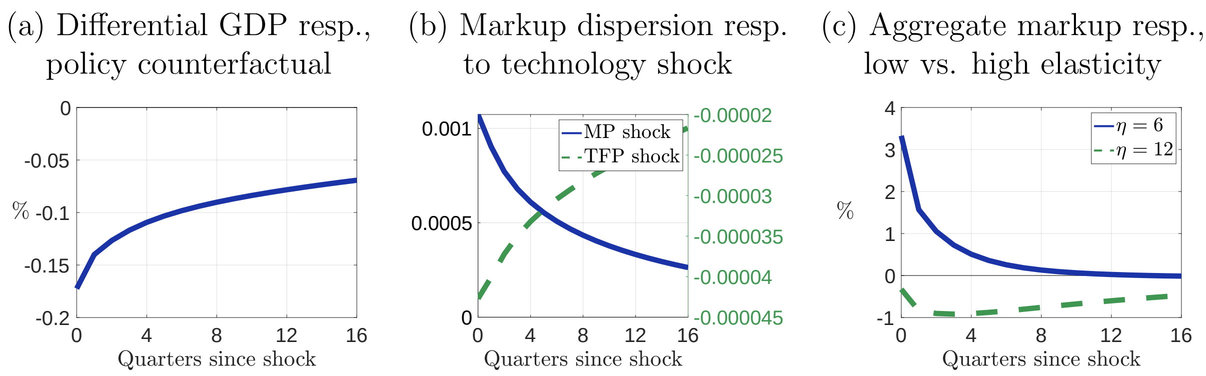
Notes: This figure shows responses to a one standard deviation contractionary monetary policy shock. In panel (e), the responses are the average markup responses of the firm types $k = 1, \ldots, 5$, where $k = 1$ is the stickiest and $k = 5$ the most flexible type of firms.

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.

In the model, contractionary monetary policy shocks raise markup dispersion and expansionary shocks lower markup dispersion, consistent with our empirical evidence. This response of markup dispersion critically depends on solving the model around the stochastic steady state, which allows us to capture precautionary price setting. In contrast, the deterministic steady state is characterized by zero markup dispersion. If we solve the model using a second-order approximation around the deterministic steady state, markup dispersion increases in response to both expansionary and contractionary monetary policy shocks, and irrespective of whether price rigidity is heterogeneous or homogeneous, see Figure 1.H.5 in the Appendix.

Even when capturing precautionary price setting, contractionary monetary policy shocks do not necessarily increase markup dispersion outside a local neighborhood around the stochastic steady state. After sufficiently large expansionary monetary policy shocks, markups of stickier firms may fall below the markups of more flexible firms. At this point, contractionary monetary policy may lower markup dispersion. We study the behavior of the model away from the stochastic steady state using a stochastic simulation of the model. The estimated response of markup dispersion on simulated data is similar and only somewhat smaller than

Figure 1.6: Policy counterfactual and additional model results



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.

the baseline response in Figure 1.5, see Appendix 1.H.1 for details.

An important aspect of the monetary transmission channel in our model 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. Specifically, we construct a policy counterfactual, in which the only counterfactual element is natural output, and thus the output gap in the Taylor rule. Whereas model-consistent natural output responds to aggregate technology shocks but not to monetary policy shocks, counterfactual natural output responds to all changes in aggregate TFP.

We then compare the effects of a monetary policy shock in the baseline and counterfactual model.[19] Panel (a) in Figure 1.6 shows the difference between the response of GDP in the counterfactual versus the baseline response.[20] 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. In the counterfactual, the output gap response is dampened, which implies a less aggressive response of (systematic) monetary policy. This is similar to a lower Taylor coefficient on the output gap, and hence output falls by more. For further details and discussion, see Appendix 1.H.2.

Panel (b) in Figure 1.6 shows the response of markup dispersion to a negative technology shock with the size and persistence that matches the endogenous

[19] We ensure the same interest rate response (30 bp) in baseline and counterfactual, by scaling up the size of the shock to 1.147 standard deviations in the counterfactual.

[20] Figure 1.H.2 in the Appendix provides further responses for this counterfactual scenario.

response of TFP to a monetary policy shock.[21] 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. So, to avoid the cost of misattributing changes in aggregate TFP to technology shocks, the monetary authority could monitor changes in markup dispersion.

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 (2020) points in the opposite direction. Following Hall (1986), the aggregate markup in our model is

$$\mu_t = \frac{\text{TFP}_t}{W_t/P_t}, \tag{1.11}$$

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 (1.11) 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 1.6 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 1.H.4 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.

To investigate the robustness of our quantitative results, we analyze the effects of monetary policy shocks in a number of model variations, including a model with real rigidities, a model with Rotemberg price adjustment, and a model with trend inflation, see Appendix 1.I.

1.5 Conclusion

This paper studies how markup dispersion matters for monetary transmission. Monetary policy shocks increase the dispersion of markups across firms if firms

[21] Figure 1.H.3 in the Appendix provides further responses for the technology shock.

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. We show that our novel mechanism has implications for monetary policy and for the markup response to monetary policy shocks.

Appendices for Chapter 1

1.A Data construction and descriptive statistics

1.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. We delete duplicate firm-quarter observations. We use the NAICS industry classification and exclude firms in utilities (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`) and total assets (`atq`) that are weakly negative. We fill missing values of depreciation and amortization (`dpq`), selling, general and administrative expenses (`xsgaq`), debt in current liabilities (`dlcq`), long-term debt (`dlttq`) and cash and short-term investments (`cheq`) by zero. We discard observations of these same variables if they are strictly negative. 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 $K_{it_0} = \texttt{ppegtq}_{it_0}$ 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 1.A.1 shows descriptive statistics for our baseline sample.

Table 1.A.1: Summary statistics for Compustat data

	mean	sd	min	max	count
Sales	636.99	3045.50	1.00	132182.15	332308
Fixed assets	1021.37	5554.55	0.00	273536.00	329311
Variable costs	441.86	2290.12	0.16	104456.86	332308
Total assets	2773.18	13447.23	0.00	559922.78	331105

Notes: All variables are in millions of 2012Q1 US$.

1.A.2 Monetary policy shocks

We construct high-frequency identified monetary policy shocks as described in Subsection 1.2.1. Table 1.A.2 reports summary statistics for shock series and Figure 1.A.1 shows the time series.

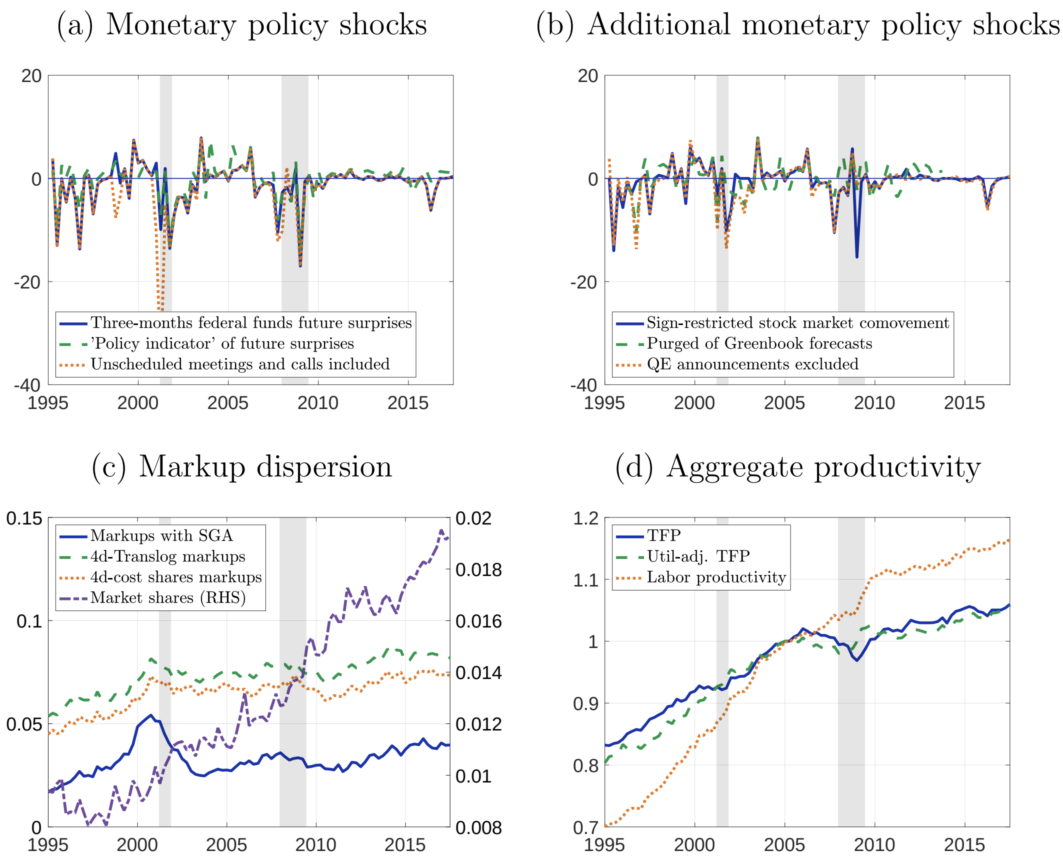
Table 1.A.2: Summary statistics of monetary policy shocks

	mean	sd	min	max	count
Three-month Fed funds future surprises	-1.00	4.06	-17.01	7.87	94
... unscheduled meetings and 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 shocks	-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: Monetary policy shocks in basis points.

1.A.3 Time series plots

Figure 1.A.1: Monetary policy shocks, aggregate productivity, and markup dispersion



(a) Monetary policy shocks

(b) Additional monetary policy shocks

(c) Markup dispersion

(d) Aggregate productivity

Notes: Panel (a) and (b) show monetary policy shock series. Panel (c) plots markup dispersion measures within four-digit-industry-quarters in addition those in Figure 1.1. Productivity measures in panel (d) are in logs and normalized to 1 in 2005Q1. Shaded gray areas indicate NBER recession dates.

1.A.4 Data on price rigidity

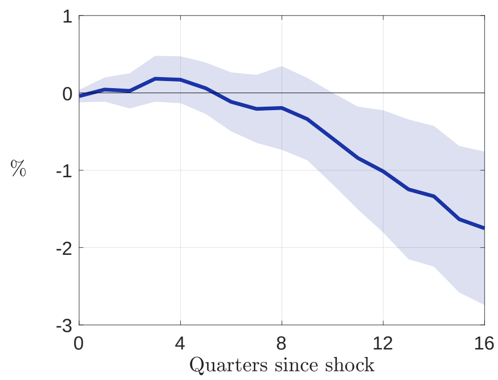
To maximize firm-level variation in price rigidity, we weight average industry-level price adjustment frequency with firms' industry sales from the Compustat segment files. Industry-level price adjustment frequency is based on Pasten et al. (2020). We define the implied price duration as $-1/\log(1 - \text{price adjustment frequency})$.

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 10,956 unique firms. For 2,685 firms (25%), 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 2.14 (2) with a standard deviation of 0.41.

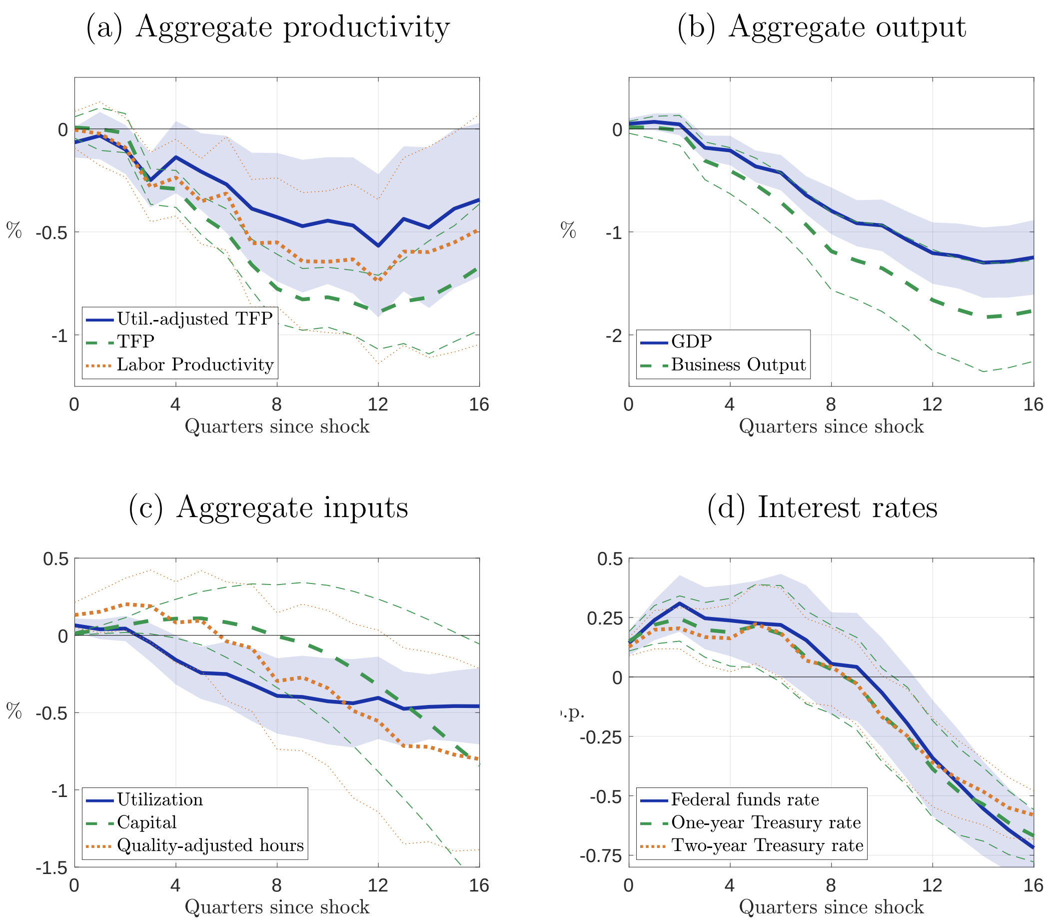
1.B Additional empirical results

Figure 1.B.1: Aggregate R&D response to monetary policy shock



Notes: This figure shows the response of aggregate R&D investment to monetary policy shocks obtained from local projections as in equation (1.4). The shaded area indicate one standard error bands based on the Newey–West estimator.

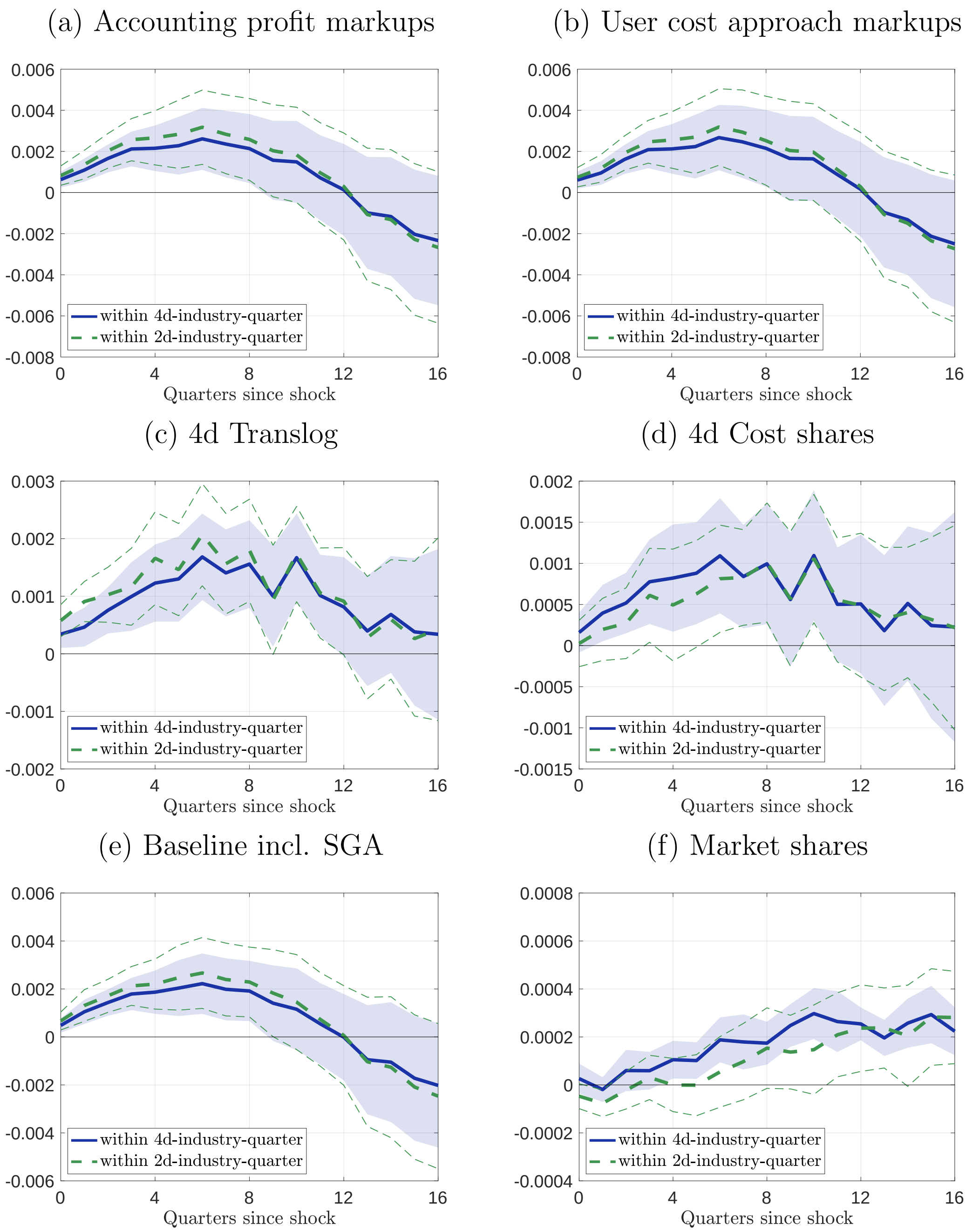
Figure 1.B.2: Macroeconomic responses to monetary policy shocks



Notes: This figure shows macroeconomic responses to monetary policy shocks obtained from local projections as in equation (1.4). The local projections in Panel (d) are estimated in levels rather than log differences. The shaded and bordered areas indicate one standard error bands based on the Newey–West estimator.

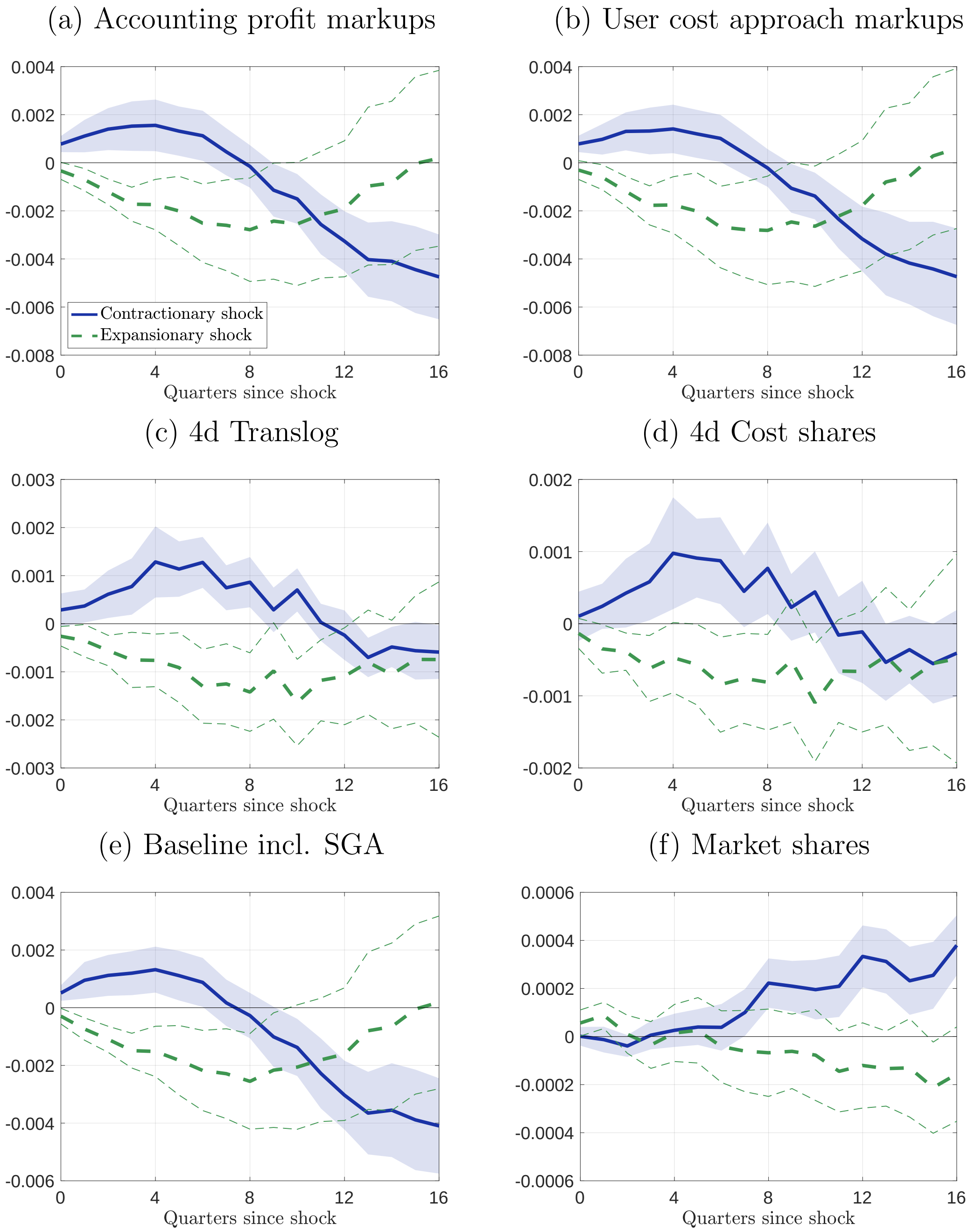
1.C Robustness of evidence in Section 2

Figure 1.C.1: Responses of markup dispersion



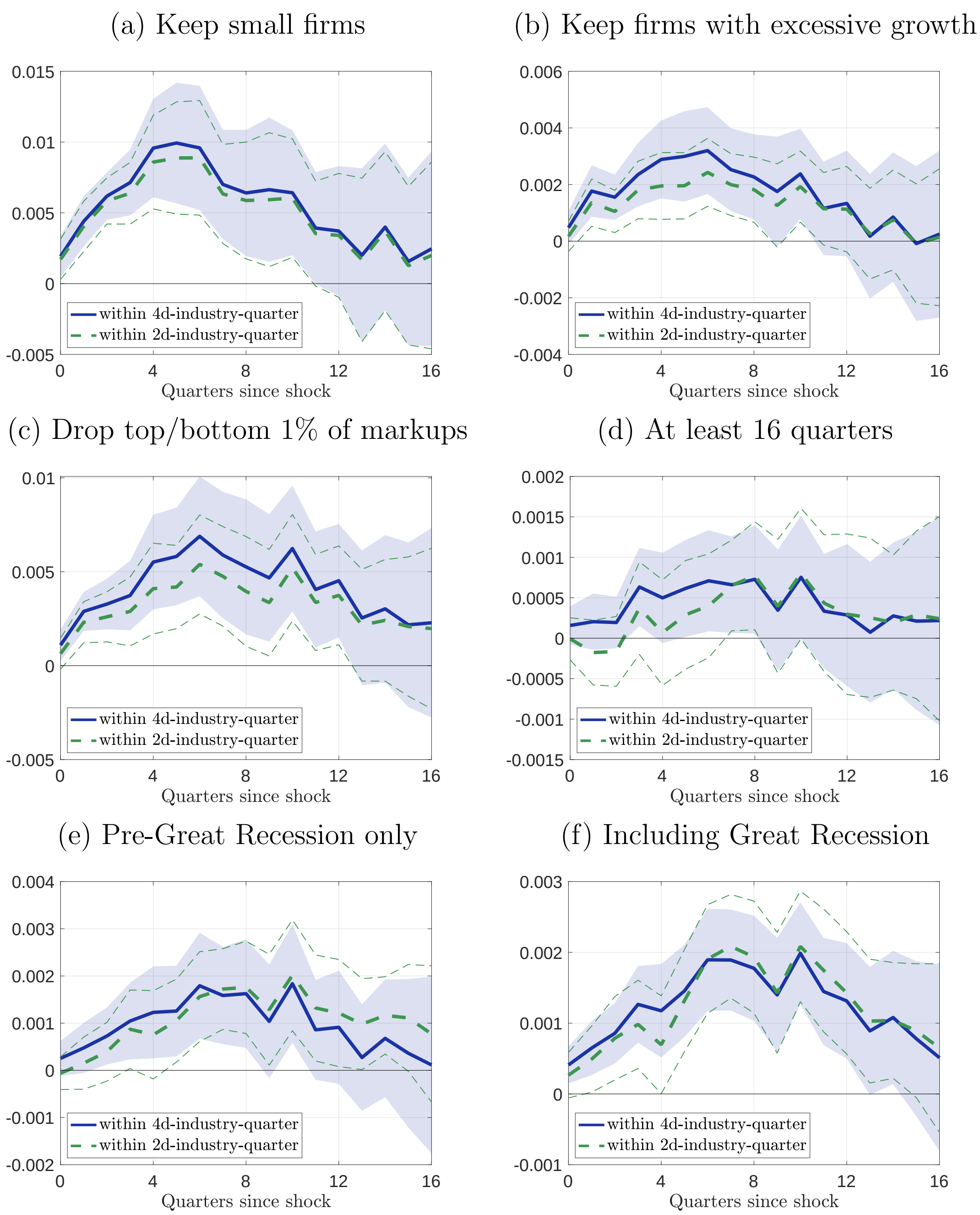
Notes: This figure shows the responses of markup dispersion to monetary policy shocks obtained from local projections as in equation (1.4). Markup dispersion is measured within two-digit and four-digit industry-quarters based on different markup measures, see Section 1.2.4 for details. Market share dispersion is computed as the variance of firm-level sales over total sales within two-digit and four-digit industry-quarters. The shaded and bordered areas indicate one standard error bands based on Newey–West.

Figure 1.C.2: Asymmetric responses of markup dispersion



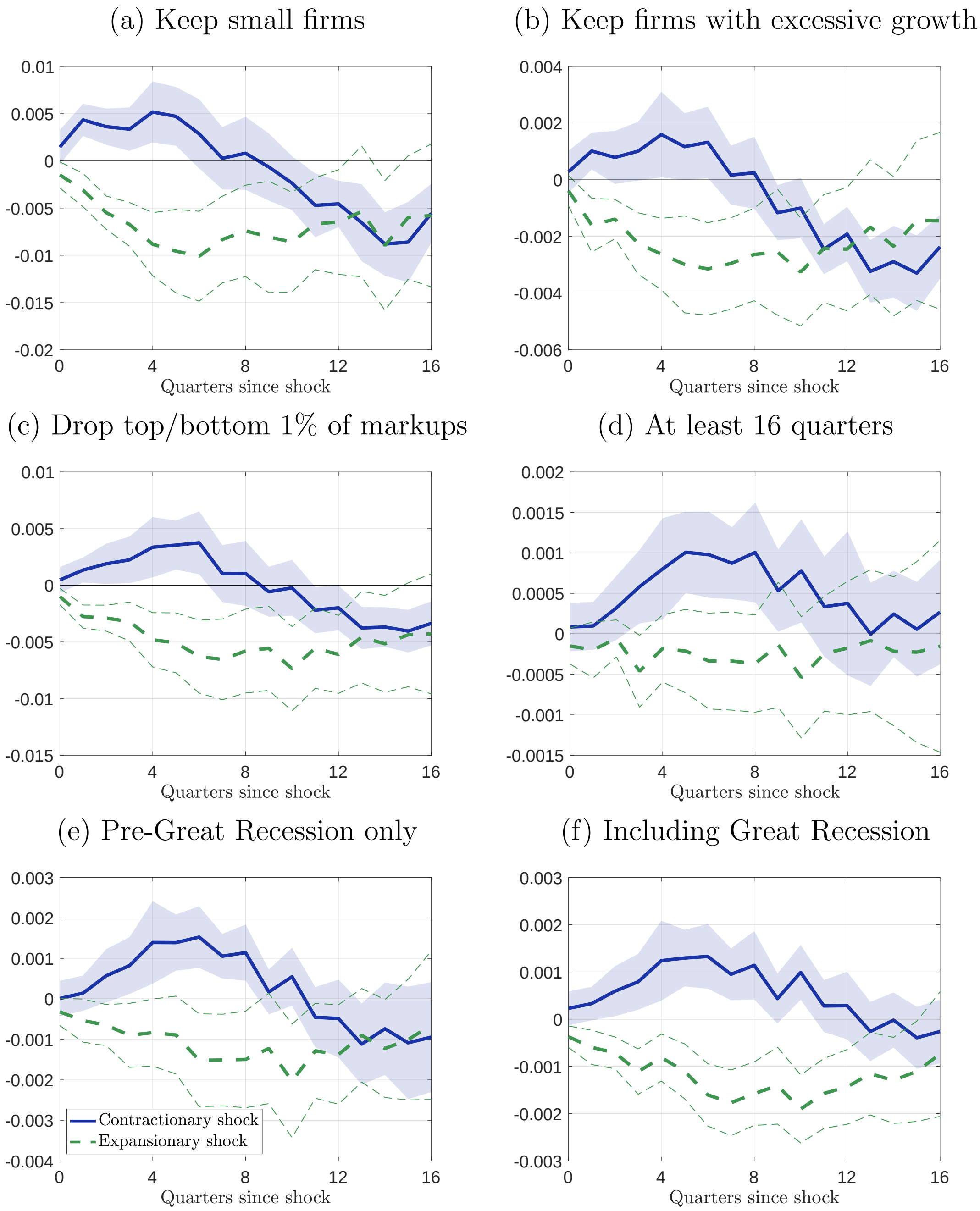
Notes: This figure shows the asymmetric responses to monetary policy shocks obtained from local projections extending specification (1.4) to separately estimate the response to positive and negative shocks. Markup dispersion is measured within two-digit-industry-quarters. Panels (a)–(e) use various markup measures and panel (f) uses market shares within two-digit-industry-quarters, see Section 1.2.4 for details. The shaded and bordered areas indicate one standard error bands based on Newey–West.

Figure 1.C.3: Responses of markup dispersion under alternative data treatments



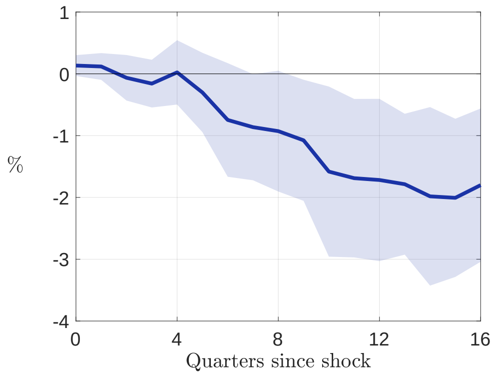
Notes: This figure shows the responses to monetary policy shocks obtained from local projections as in equation (1.4). Markup dispersion is measured within two-digit and four-digit industry-quarters using the baseline markup measure. See Section 1.2.4 for details on the different data treatments. The shaded and bordered areas indicate one standard error bands based on Newey–West.

Figure 1.C.4: Asymmetric markup dispersion responses for alternative data treatments



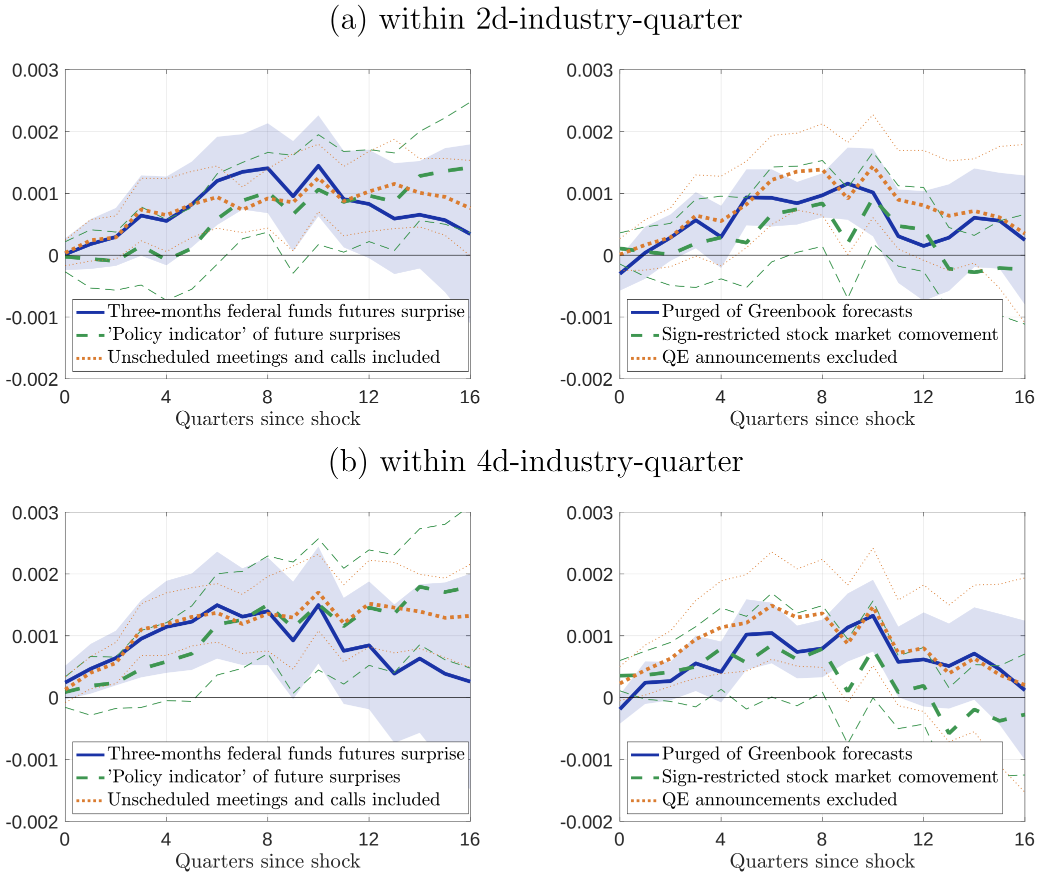
Notes: This figure shows the responses to monetary policy shocks obtained from local projections extending specification (1.4) to separately estimate the response to positive and negative shocks. Markup dispersion is measured within four-digit industry-quarters using the baseline markup measure. See Section 1.2.4 for details on the different data treatments. The shaded and bordered areas indicate one standard error bands based on Newey–West.

Figure 1.C.5: 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 (1.4). The shaded area is a one standard error band based on Newey–West.

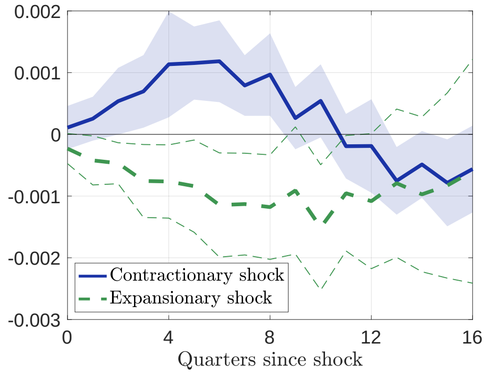
Figure 1.C.6: Responses of markup dispersion for alternative monetary policy shocks



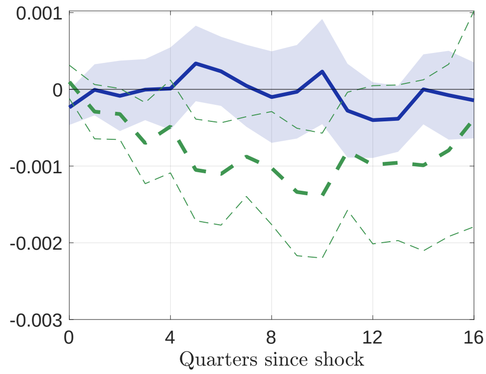
Notes: This figure shows the responses to monetary policy shocks obtained from local projections as in equation (1.4). Markup dispersion is measured within four-digit industry-quarters using the baseline markup measure. See Section 1.2.4 for details on the different monetary policy shocks. The shaded and bordered areas indicate one standard error bands based on Newey–West.

Figure 1.C.7: Asymmetric markup dispersion responses for alternative monetary policy shocks

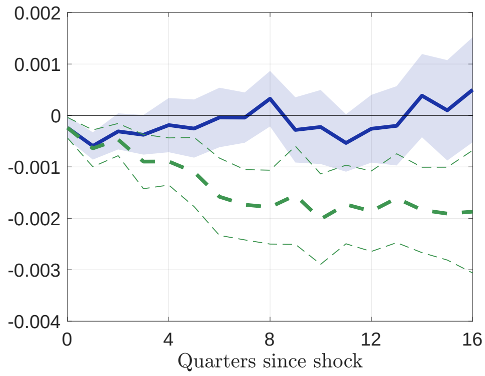
(a) Baseline



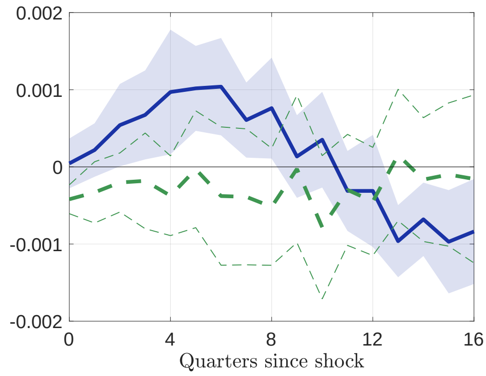
(b) Purged of Greenbook forecasts



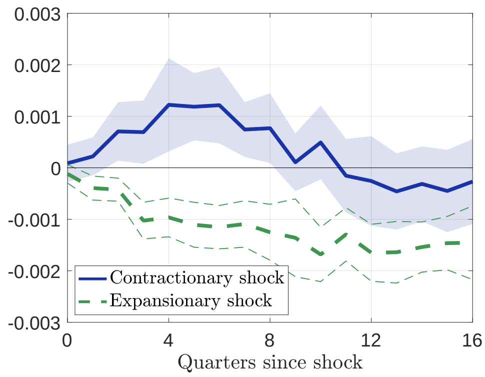
(c) 'Policy indicator' of future surprises



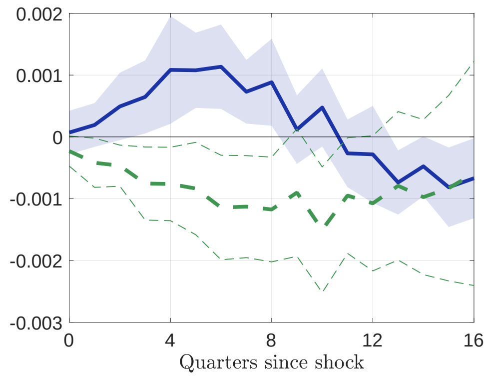
(d) Sign-restricted stock market comovement



(e) Unscheduled meetings and calls included



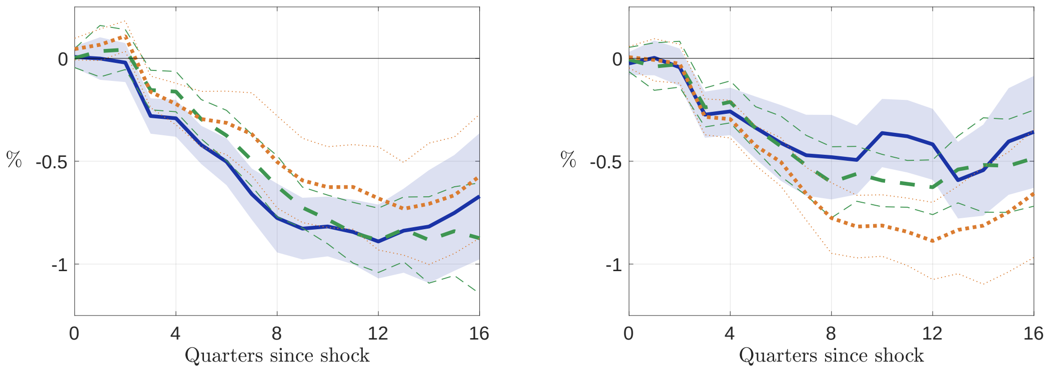
(f) QE announcements excluded



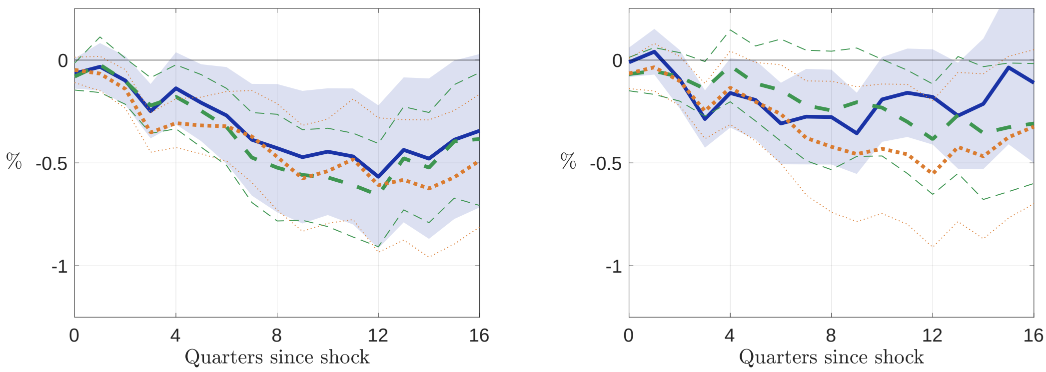
Notes: This figure shows the asymmetric responses to monetary policy shocks obtained from local projections extending specification (1.4) to separately estimate the response to positive and negative shocks. Markup dispersion is measured within four-digit industry-quarters using the baseline markup measure. See Section 1.2.4 for details on the different monetary policy shocks. The shaded and bordered areas indicate one standard error bands based on Newey–West.

Figure 1.C.8: Aggregate productivity responses for alternative monetary policy shocks

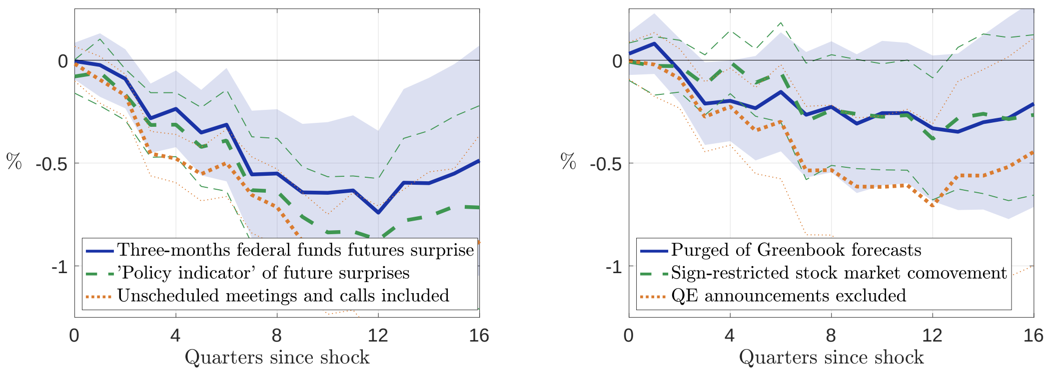
(a) TFP



(b) Utilization-adjusted TFP

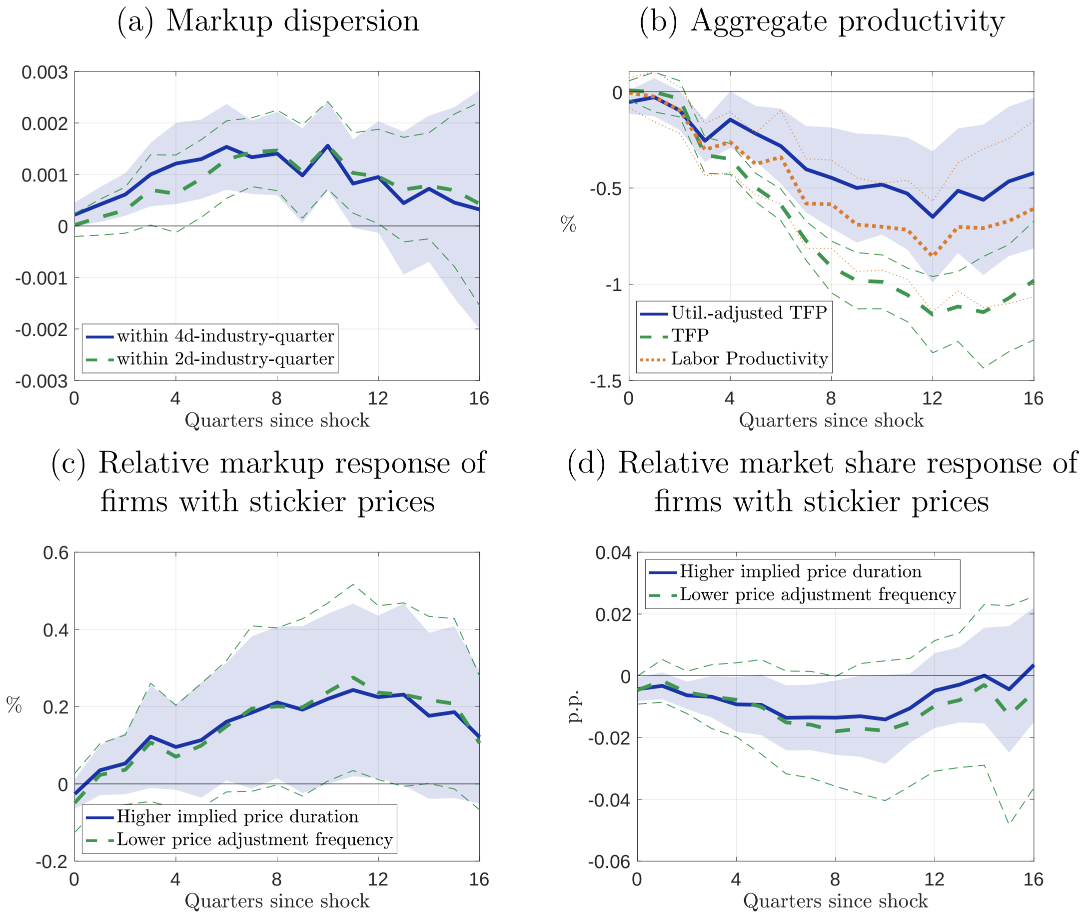


(c) Labor productivity



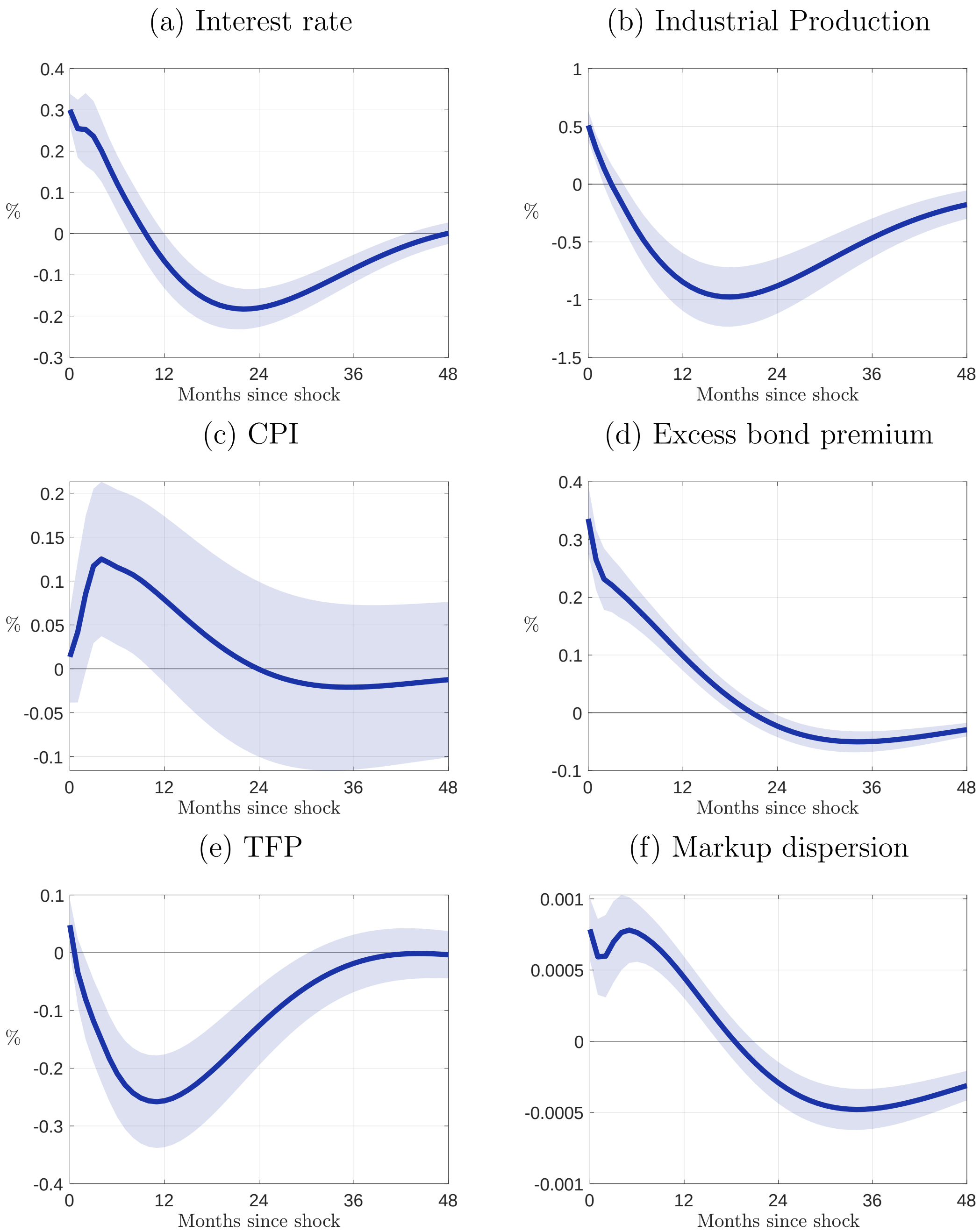
Notes: This figure shows the responses of aggregate productivity to monetary policy shocks obtained from local projections as in equation (1.4). The shaded and bordered areas indicate one standard error bands based on Newey–West.

Figure 1.C.9: Main results using LP-IV



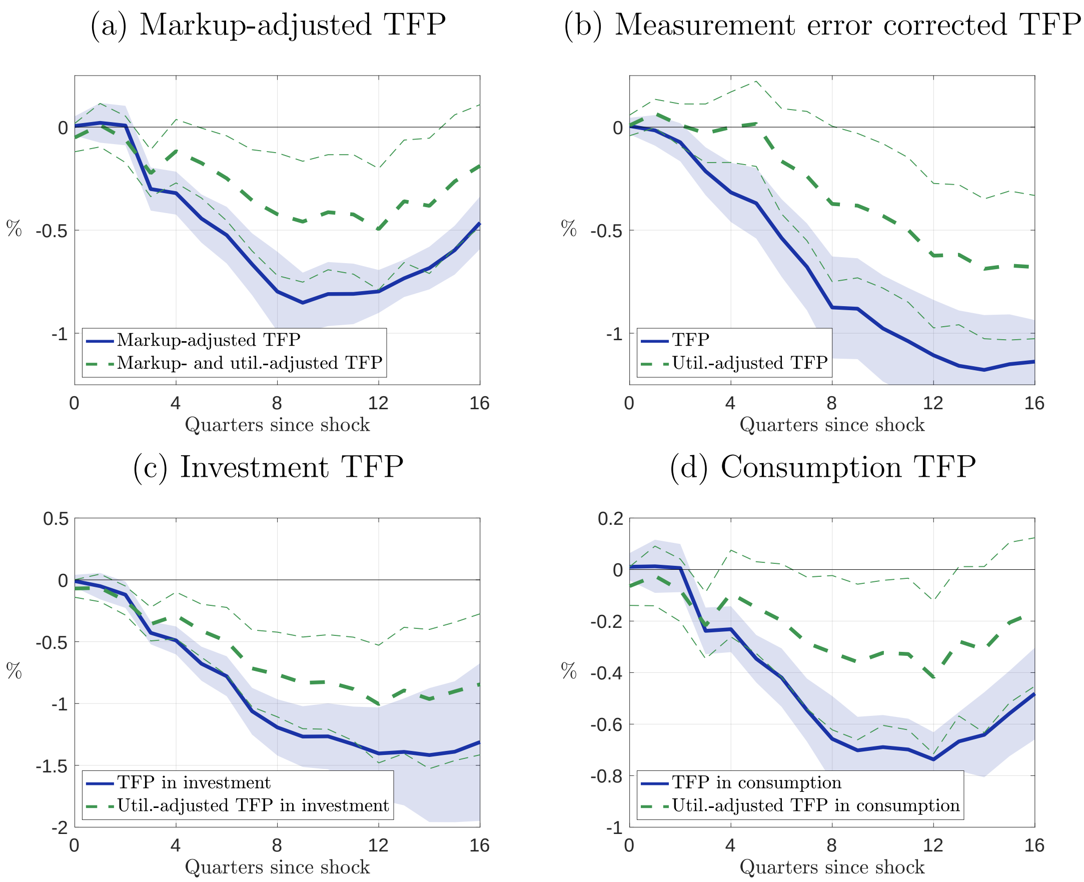
Notes: This figure shows the responses to monetary policy shocks obtained from local projections with instrumental variables (LP-IV), $y_{t+h} - y_{t-1} = \alpha^h + \beta^h \Delta R_t + \gamma_1^h(y_{t-1} - y_{t-2}) + u_t^h$, (in panels (a) and (b)) and analogues of the panel local projections (in panels (c) and (d)), where the changes in the one-year Treasury rate, ΔR_t, (and the interactions thereof with price stickiness, respectively) are instrumented with the monetary policy shocks $\varepsilon_t^{\text{MP}}$ (and the interactions of monetary policy shocks with price stickiness, respectively). The impulse responses are normalized such that they correspond to a 30bp peak increase in the one-year rate. The shaded and bordered areas in panels (a) and (b) indicate a one standard error band based on Newey–West, and in panels (c) and (d) they indicate a 90% error band two-way clustered by firms and quarters.

Figure 1.C.10: Proxy SVAR results



Notes: This figures shows the responses to a monetary policy shock, which raises the interest rate by 30bp, based on proxy SVAR similar to Gertler and Karadi (2015). The VAR is estimated at monthly frequency with three lags, including the one-year rate, (log) industrial production, (log) CPI, the excess bond premium of Gilchrist and Zakrajšek (2012), (log) TFP and the baseline measure of markup dispersion (within four-digit-industry-quarters). TFP and markup dispersion are interpolated to monthly frequency using the procedure of Chow and Lin (1971). Shaded areas are one-standard error bands from a wild bootstrap-after-bootstrap.

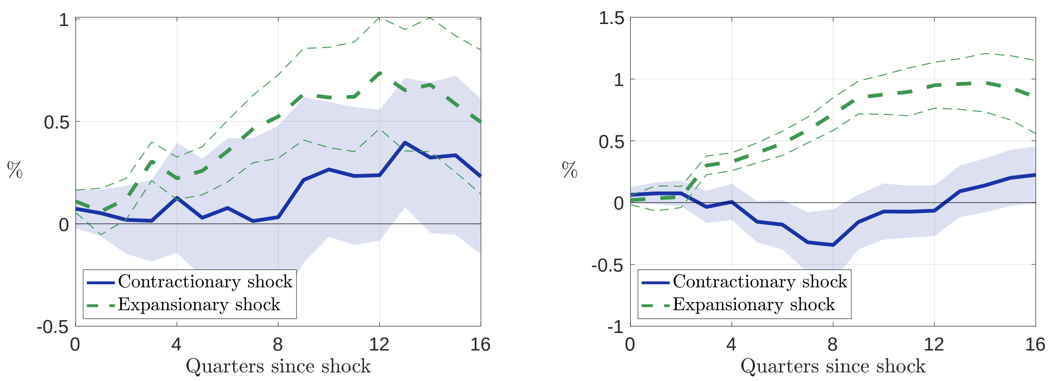
Figure 1.C.11: Further productivity responses



Notes: Responses to monetary policy shocks obtained from local projections as in equation (1.4). Investment-TFP and Consumption-TFP are from Fernald (2014). Markup-corrected TFP is constructed following Hall (1986) 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 1.C.12: Asymmetric responses of (util.-adjusted) TFP to monetary policy shocks

(a) Util.-adjusted TFP (b) TFP



Notes: This figure shows the responses of productivity to monetary policy shocks obtained from local projections extending specification (1.4) to separately estimate the response to positive and negative shocks. TFP and utilization-adjusted TFP are from Fernald (2014). The shaded and bordered areas indicate one standard error bands based on Newey–West.

1.D Robustness of evidence in Section 3

We first consider alternative markup estimates based on the accounting profits and user costs approach. Table 1.D.2 shows the correlation between average markup and price rigidity. While Figure 1.4 (d) in the paper shows the relative markup response of firms with stickier prices to monetary policy shocks, we also estimate the relative markup responses when markups are based on cost shares, translog technology, or the baseline including SGA, see Figure 1.D.1. Second, we consider the role of alternative data treatments. Table 1.D.3 shows that the correlation between markups and price rigidity is robust across data treatments. Figure 1.D.2 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. Third, we consider alternative monetary policy shock series, see Figure 1.D.3. Fourth, we consider an LP-IV setup as described in Section 1.2.4, see Figure 1.C.9 (c). Finally, we include the apex of the Great Recession, see Figure 1.D.2 (d) and (e).

Table 1.D.1: Regressions of markup on price stickiness incl. all price adjustment frequencies

(a) Regressions of markups on implied price duration

	log(Markup)			
	Baseline		Accounting profits	User cost approach
Implied price duration	0.0434 (0.0197)	0.0363 (0.0176)	0.00683 (0.00291)	0.00927 (0.00334)
Additional controls	No	Yes	Yes	Yes
2-digit industry FE	Yes	Yes	Yes	Yes
Observations	4014	4014	3960	3951
Adjusted R^2	0.107	0.184	0.240	0.166

(b) Regressions of markups on price adjustment frequency

	log(Markup)			
	Baseline		Accounting profits	User cost approach
Price adjustment frequency	-0.244 (0.144)	-0.186 (0.134)	-0.0444 (0.0180)	-0.0630 (0.0193)
Additional controls	No	Yes	Yes	Yes
2-digit industry FE	Yes	Yes	Yes	Yes
Observations	4014	4014	3960	3951
Adjusted R^2	0.103	0.180	0.240	0.166

Notes: Regression of firm-level markup (averaged over 2005–2011) on firm-level price adjustment frequency and implied price duration, respectively, when including firms with price adjustment frequencies above 99%. Standard errors are clustered at the two-digit industry level and shown in parentheses.

Table 1.D.2: Regressions of markup on price stickiness for alternative markup series

(a) Regressions of markups on implied price duration

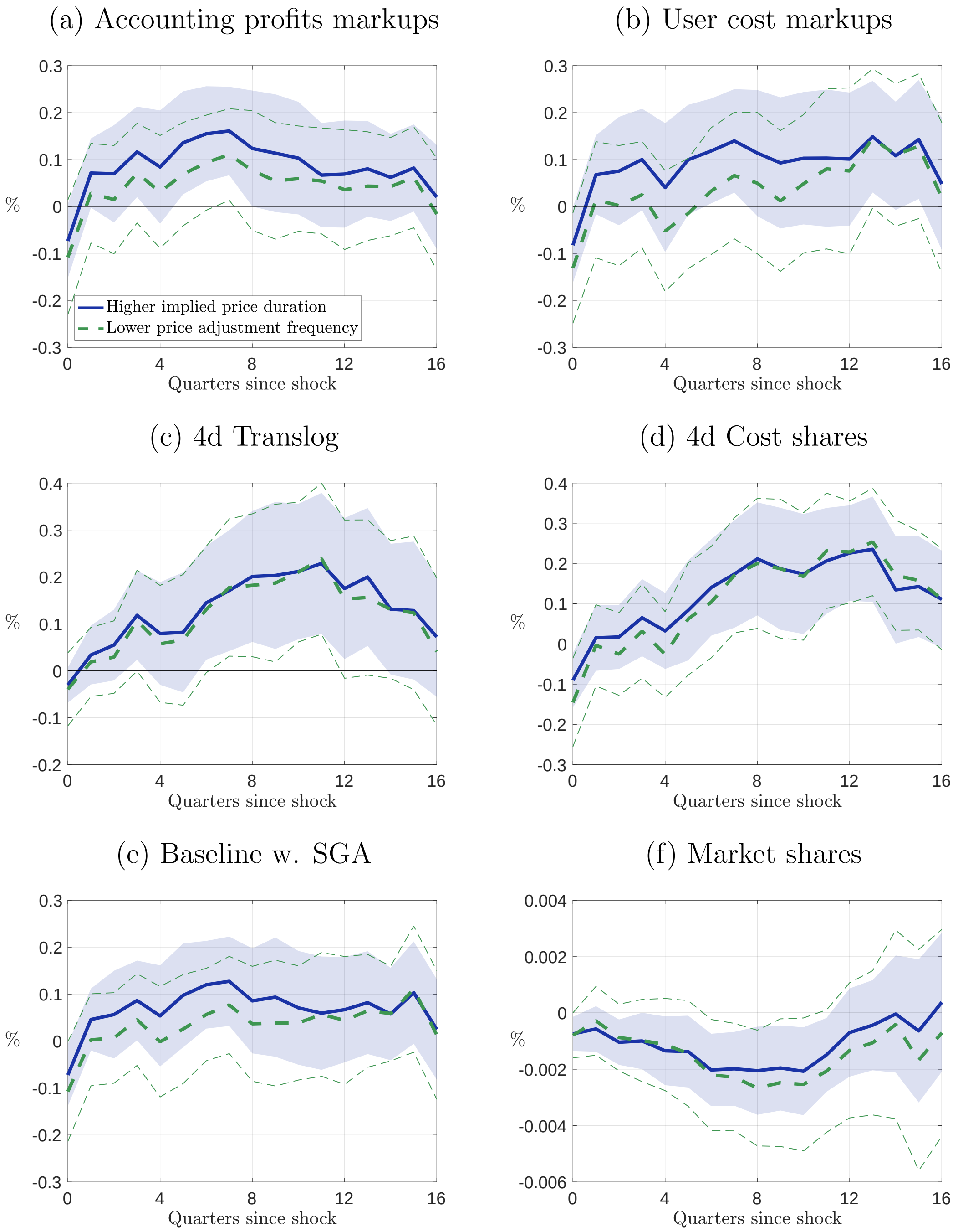
	log(Markup)		
	4d Translog	4d cost shares	Baseline incl. SGA
Implied price duration	0.0270	0.0511	0.00698
	(0.0112)	(0.0144)	(0.00337)
Additional controls	Yes	Yes	Yes
2-digit industry FE	Yes	Yes	Yes
Observations	3786	3826	3812
Adjusted R^2	0.125	0.288	0.262

(b) Regressions of markups on price adjustment frequency

	log(Markup)		
	4d Translog	4d cost shares	Baseline incl. SGA
Price adjustment frequency	-0.194	-0.370	-0.0425
	(0.0917)	(0.0759)	(0.0269)
Additional controls	Yes	Yes	Yes
2-digit industry FE	Yes	Yes	Yes
Observations	3786	3826	3812
Adjusted R^2	0.127	0.292	0.262

Notes: Regression of firm-level markup (averaged over 2005–2011) on firm-level price adjustment frequency and implied price duration, respectively. For details on the different markup measures, see Section 1.2.4. Standard errors are clustered at the two-digit industry level and shown in parentheses.

Figure 1.D.1: Relative markup and market share 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 (log) firm-level markup (or market share) of firms with a price adjustment frequency one standard deviation below mean (or with an implied price duration one standard deviation above mean) from panel local projections as in equation (1.10). Panels (a)–(e) use different markup measures and panel (f) uses market shares within two-digit-industry-quarters; see Section 1.2.4 for details. The shaded and bordered areas indicate 90% error bands clustered by firms and quarters.

Table 1.D.3: Regressions of markup on price stickiness under alternative data treatments

(a) Keep small firms

	log(Markup)			
Implied price duration	0.0417 (0.0143)	0.0485 (0.0162)		
Price adjustment frequency			-0.301 (0.0681)	-0.352 (0.0810)
Additional controls	No	Yes	No	Yes
2-digit industry FE	Yes	Yes	Yes	Yes
Observations	4395	4389	4395	4389
Adjusted R^2	0.167	0.195	0.169	0.197

(b) Keep firms with excessive growth

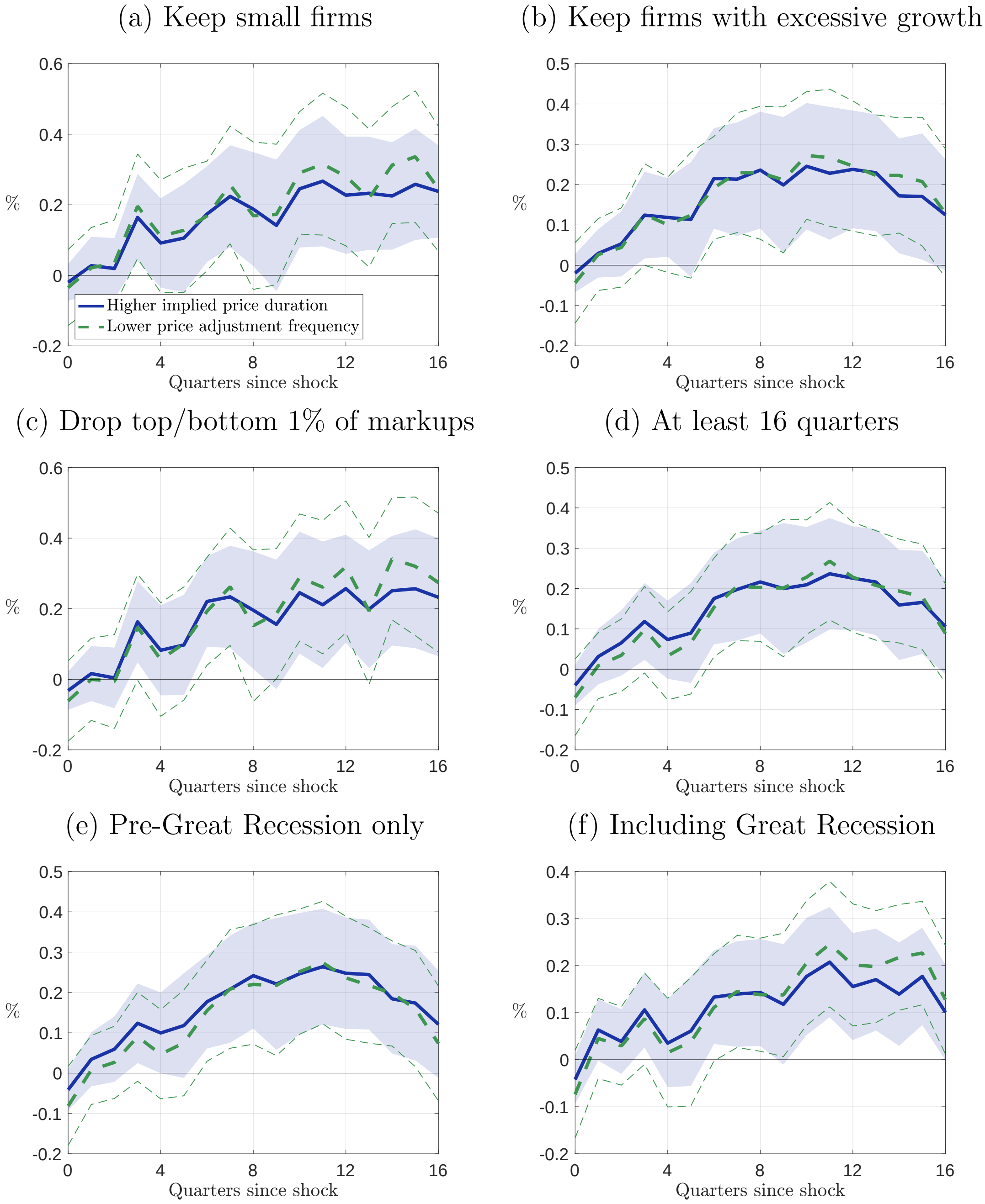
	log(Markup)			
Implied price duration	0.0522 (0.0161)	0.0473 (0.0142)		
Price adjustment frequency			-0.385 (0.0854)	-0.337 (0.0768)
Additional controls	No	Yes	No	Yes
2-digit industry FE	Yes	Yes	Yes	Yes
Observations	4208	4160	4208	4160
Adjusted R^2	0.128	0.193	0.134	0.196

(c) Drop top/bottom 1% of markups

	log(Markup)			
Implied price duration	0.0487 (0.0218)	0.0494 (0.0232)		
Price adjustment frequency			-0.362 (0.104)	-0.366 (0.120)
Additional controls	No	Yes	No	Yes
2-digit industry FE	Yes	Yes	Yes	Yes
Observations	4070	4070	4070	4070
Adjusted R^2	0.150	0.159	0.153	0.162

Notes: Regression of baseline firm-level markup (averaged over 2005–2011) on firm-level price adjustment frequency and implied price duration, respectively. See Section 1.2.4 for details on the different data treatments. Standard errors are clustered at the two-digit industry level and shown in parentheses.

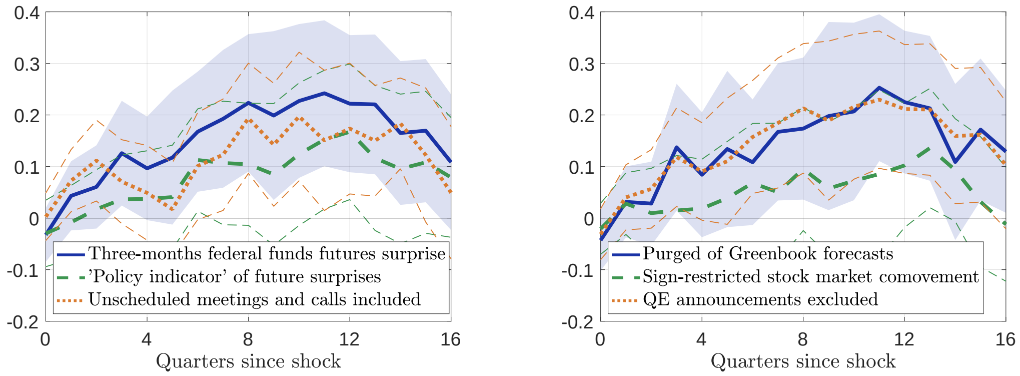
Figure 1.D.2: Relative markup response of firms with stickier prices under alternative data treatments



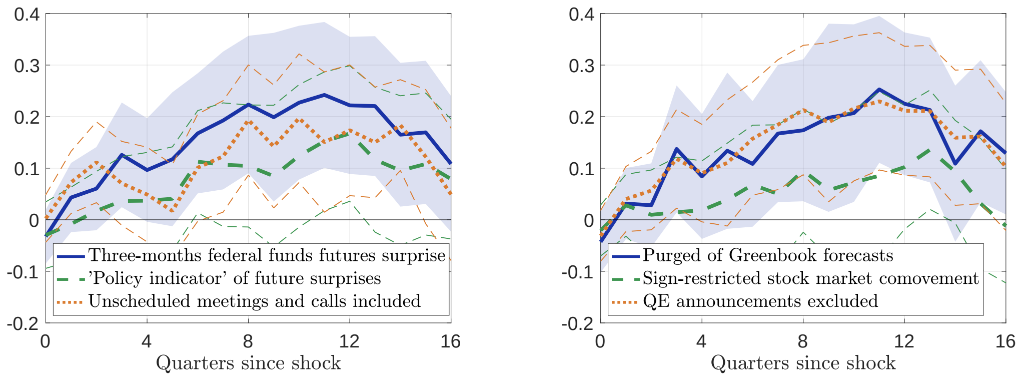
Notes: The figures show the response to a one standard deviation contractionary monetary policy shock of the firm-level markup of firms with a price adjustment frequency one standard deviation below mean (or with an implied price duration one standard deviation above mean) from panel local projections as in equation (1.10). The regressions include interactions with lagged log assets, leverage, and liquidity and their interactions with the monetary policy shock. See Section 1.2.4 for details on the different data treatments. The shaded and bordered areas indicate 90% error bands clustered by firms and quarters.

Figure 1.D.3: Relative markup response of firms with stickier prices for alternative monetary policy shocks

(a) Higher implied price duration



(b) Lower price adjustment frequency



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 mean (or with an implied price duration one standard deviation above mean) from panel local projections as in equation (1.10). The regressions include interactions with lagged log assets, leverage, and liquidity and their interactions with the monetary policy shock. See Section 1.2.4 for details on the different monetary policy shocks. The shaded and bordered areas indicate 90% error bands clustered by firms and quarters.

1.E Proofs

1.E.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 a firm-specific marginal cost $X_{it} = X_t/A_{it}$, where X_t denotes a common marginal cost 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

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

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.

1.E.2 Proof of Proposition 1.1

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

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

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

1.E.3 Proof of Proposition 1.2

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 (1.8) as

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

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

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

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

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

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

1.E.4 Proof of Proposition 1.3

Let us first define

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

such that we can re-write the first-order condition in equation (1.9) 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,

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

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 $\bar{X}$,

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

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

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

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

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

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

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

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

1.F 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

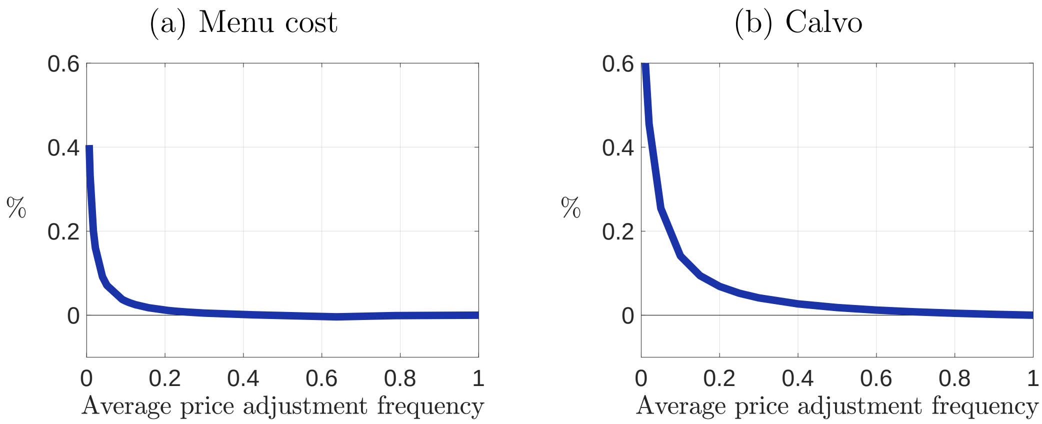
$$
\begin{aligned}
V(p, Z) &= \mathbb{E}_{\xi}[\max\{V^A(Z) - \xi, V^N(Z)\}] \\
V^A(Z) &= \max_{p^*} \left\{ \left(\frac{p^*}{P} - X \right) \left(\frac{p^*}{P} \right)^{-\eta} + \beta \mathbb{E}_Z \left[V(p^*, Z') \right] \right\} \\
V^N(p, Z) &= \left(\frac{p}{P} - X \right) \left(\frac{p}{P} \right)^{-\eta}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

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.

Figure 1.F.1: 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.

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 1.F.1 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.

1.G Details on the quantitative model

1.G.1 Preferences and technologies

We assume a representative infinitely-lived household who maximizes

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

subject to the budget constraints $P_t C_t + R_t^{-1} B_t \leq B_{t-1} + W_t N_t + D_t$ 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 $\{Y_{it}\}$. Each intermediate good i is produced by a monopolistically competitive intermediate good firm i. The unit measure of differentiated goods is split equally across K different types of intermediate goods firms producing the differentiated goods. The firm types are indexed by $k = 1, \ldots, K$ and firms are ordered according to their type such that firms indexed $i \in [0, 1/K)$ belong to type $k = 1$ and firms indexed $i \in ((k-1)/K, k/K]$ belong to type $k = 2, \ldots, K$. Firms across the K types are ex-ante identical except for differences in their exogenous price reset probability $1 - \theta_k$. Intermediate goods are produced with technology $Y_{it} = A_t N_{it}$, 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. Real marginal costs are hence $mc_t = w_t/A_t$. Final good aggregation implies an isoelastic demand schedule for intermediate goods given by $Y_{it} = (P_{it}/P_t)^{-\eta} Y_t$, where P_{it} is the firm-level price and P_t the aggregate price index

$$P_t = \left[\int_0^1 P_{it}^{1-\eta} di \right]^{\frac{1}{1-\eta}} = \left[\frac{1}{K} \sum_{k=1}^{K} P_{kt}^{1-\eta} \right]^{\frac{1}{1-\eta}}.$$

P_{kt} denotes the firm type-k specific price index

$$P_{kt} = \left[\int_{(k-1)/K}^{k/K} P_{it}^{1-\eta} di\right]^{\frac{1}{1-\eta}} = \left[(1-\theta_k)\tilde{P}_{kt}^{1-\eta} + \theta_k P_{kt-1}^{1-\eta}\right]^{\frac{1}{1-\eta}},$$

where $\tilde{P}_{kt}$ is the optimal reset price, which solves maximizes the value of the firm to its shareholder

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

where μ_t is the marginal utility of consumption $\mu_t = C_t^{-1/\gamma}$. The monetary authority follows a Taylor rule to stabilize inflation, $\Pi_t = P_t/P_{t-1}$, and fluctuations in output, Y_t, around its natural level, denoted $\tilde{Y}_t$, subject to monetary policy shocks ν_t,

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

1.G.2 Equilibrium conditions

$$\frac{\tilde{P}_{kt}}{P_t} = \frac{\eta}{\eta-1}\frac{\mathbb{E}_t \sum_{j=0}^{\infty}\theta_k^j\beta^j\mu_{t+j}Y_{t+j}\Pi_{t,t+j}^{\eta}mc_{t+j}}{\mathbb{E}_t \sum_{j=0}^{\infty}\theta_k^j\beta^j\mu_{t+j}Y_{t+j}\Pi_{t,t+j}^{\eta-1}} \quad \forall k = 1,\dots,K \qquad \text{(reset price)}$$

$$1 = \frac{1}{K}\sum_{k=1}^{K}\left(\frac{P_{kt}}{P_t}\right)^{1-\eta} \qquad \text{(aggregate price index)}$$

$$\frac{P_{kt}}{P_t} = \left[(1-\theta_k)\left(\frac{\tilde{P}_{kt}}{P_t}\right)^{1-\eta} + \theta_k \Pi_t^{\eta-1}\left(\frac{P_{kt-1}}{P_{t-1}}\right)^{1-\eta}\right]^{\frac{1}{1-\eta}} \quad \forall k = 1,\dots,K \qquad \text{(type } k \text{ price index)}$$

$$S_t = \frac{1}{K}\sum_{k=1}^{K} S_{kt} \qquad \text{(aggregate price dispersion)}$$

$$S_{kt} = (1-\theta_k)\left(\frac{\tilde{P}_{kt}}{P_t}\right)^{-\eta} + \theta_k \Pi_t^{\eta} S_{kt-1} \quad \forall k = 1,\dots,K \qquad \text{(type } k \text{ price dispersion)}$$

$$Y_t = \frac{A_t}{S_t}N_t \qquad \text{(aggregate output)}$$

$$TFP_t = \frac{Y_t}{N_t} \qquad \text{(TFP)}$$

$$mc_t = \frac{w_t}{A_t} \qquad \text{(marginal cost)}$$

$$N_t^{\frac{1}{\varphi}} C_t^{\frac{1}{\gamma}} = w_t \qquad \text{(intratemporal optimality)}$$

$$C_t^{-\frac{1}{\gamma}} = \beta \mathbb{E}_t \left[\frac{R_t}{\Pi_{t+1}} C_{t+1}^{-\frac{1}{\gamma}} \right] \qquad \text{(intertemporal optimality)}$$

$$C_t = Y_t \qquad \text{(resource constraint)}$$

$$R_t = R_{t-1}^{\rho_r} \left[\frac{1}{\beta} (\Pi_t)^{\phi_\pi} \left(\frac{Y_t}{\tilde{Y}_t} \right)^{\phi_y} \right]^{1-\rho_r} \nu_t \qquad \text{(Taylor rule)}$$

1.H Additional model results

1.H.1 Stochastic simulation of the model

In our baseline model, contractionary monetary policy shocks raise markup dispersion and expansionary shocks lower markup dispersion. This response of markup dispersion critically depends on solving the model around the stochastic steady state, which allows us to capture precautionary price setting. However, even when capturing precautionary price setting, contractionary monetary policy shocks do not increase markup dispersion in all states of the world outside a local neighborhood around the stochastic steady state. In particular, after sufficiently large expansionary monetary policy shocks, the average markup of stickier firms may fall below the average markup of more flexible firms. At this point, a contractionary monetary policy shock may lower markup dispersion.

We investigate this possibility through a large stochastic simulation of our model. We simulate 50,000 firms for 10,000 periods and find that the average markup of the stickiest quintile of firms is below the average markup of the most flexible quintile of firms in 11.1% of the periods.

We further investigate by how much this occasional inverted order of markups affects the average response of markup dispersion to monetary policy shocks. We compute a time series of markup dispersion and project it on the simulated monetary policy shock series using our empirical framework in equation (1.4). The estimated average response of markup dispersion is similar but a bit smaller than the baseline response, see Figure 1.H.1.

1.H.2 Policy counterfactual

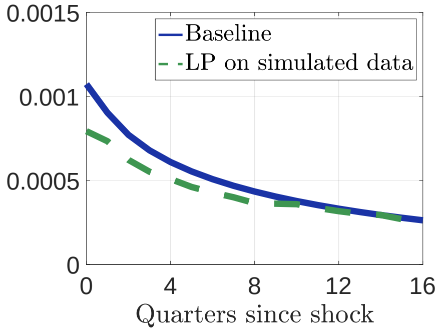
In the following, we provide some details on a counterfactual experiment, in which the monetary authority in the model (mis-)perceives the aggregate TFP response to monetary policy shocks as originating from technology shocks.

The natural level of output in the absence of price setting frictions is

$$\tilde{Y}_t = \left[\frac{\eta}{\eta - 1} A_t^{1+\frac{1}{\varphi}} \right]^{\frac{1}{\frac{1}{\varphi}+\frac{1}{\gamma}}}.$$

In the counterfactual, we assume the monetary authority mis-attributes observed fluctuations in TFP to exogenous productivity shocks A_t, which we implement by

Figure 1.H.1: Response of markup dispersion



Notes: *Baseline* replicates the response of markup dispersion to a one standard deviation monetary policy shock implied by the model solution and as shown in Figure 1.5. The dashed line is the response of markup dispersion when applying our empirical setup in equation (1.4) to a large stochastic simulation of the model. We simulate the markups of 50,000 firms, equally distributed across the five groups of price stickiness, over 10,000 periods.

using the following counterfactual (cf) natural output definition

$$\tilde{Y}_t^{\mathrm{cf}} = \left[\frac{\eta}{\eta - 1} TFP_t^{1+\frac{1}{\varphi}}\right]^{\frac{1}{\frac{1}{\varphi}+\frac{1}{\gamma}}} .$$

Since monetary policy shocks lower TFP_t, counterfactual natural output $\tilde{Y}_t^{\mathrm{cf}}$ falls, while natural output $\tilde{Y}_t$ remains constant.

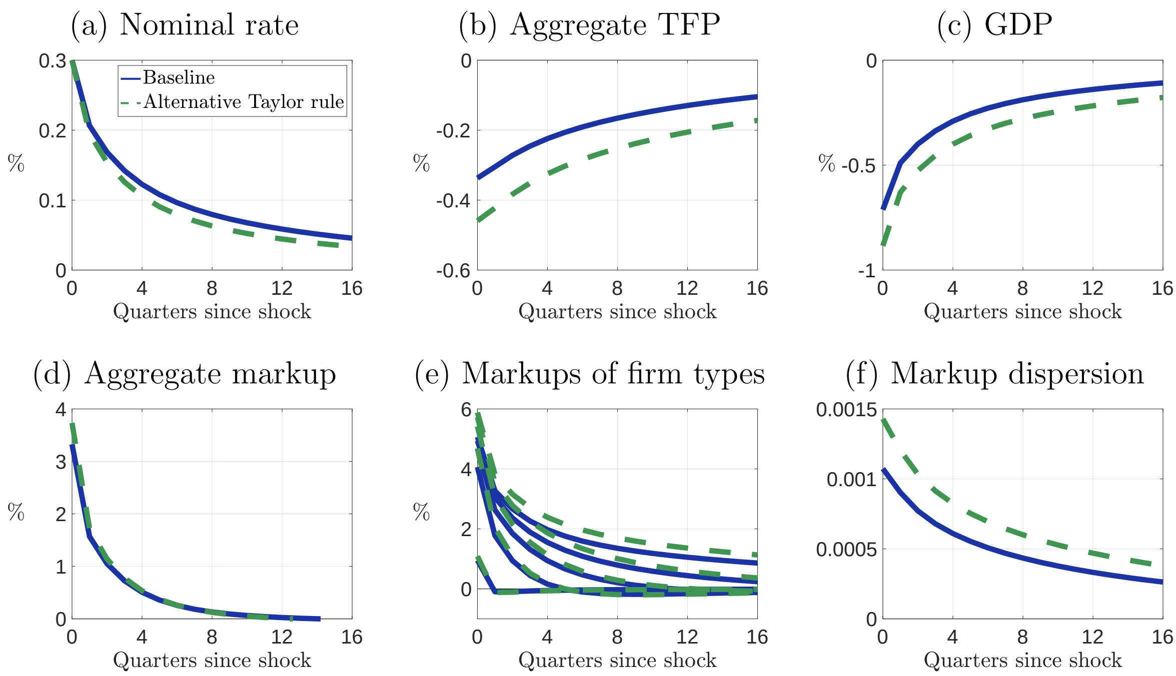
Let us define the systematic component of monetary policy in the baseline and counterfactual Taylor rule as

$$\bar{R}_t = \frac{1}{\beta} (\Pi_t)^{\phi_\pi} \left(\frac{Y_t}{\tilde{Y}_t}\right)^{\phi_y}, \qquad \text{and} \quad \bar{R}_t^{\mathrm{cf}} = \frac{1}{\beta} (\Pi_t)^{\phi_\pi} \left(\frac{Y_t}{\tilde{Y}_t^{\mathrm{cf}}}\right)^{\phi_y}.$$

The systematic component of monetary policy sets a lower nominal interest rate in response to lower inflation and output gaps. In the counterfactual, the responsiveness of the systematic component to lower output gaps is dampened because $\tilde{Y}_t^{\mathrm{cf}}$ falls as well. The counterfactual Taylor rule is hence similar to a Taylor rule with a smaller coefficient ϕ_y. This may lead to large output and inflation responses to monetary policy shocks. In addition, because the response of Y_t to a shock converges more quickly to the response of $\tilde{Y}_t^{\mathrm{cf}}$ than to $\tilde{Y}_t$, we can think of the implicit ϕ_y in the counterfactual as falling in the forecast horizon. This explains the more persistent effects in the counterfactual.

To quantify the implications of counterfactual natural output, we compare the macroeconomic effects of monetary policy shocks that raise the nominal interest rate by 30 bp in the baseline and counterfactual model. We keep all model parameters, including the variance of monetary policy shock unchanged in the counterfactual, but scale up the size of the shock in the counterfactual such that the nominal interest rate increases by 30bp. The required size of the monetary policy shock in the counterfactual corresponds to $1.16 \cdot \sigma_\nu$. Figure 1.H.2 compares

Figure 1.H.2: Model responses to monetary policy shocks under alternative Taylor rule

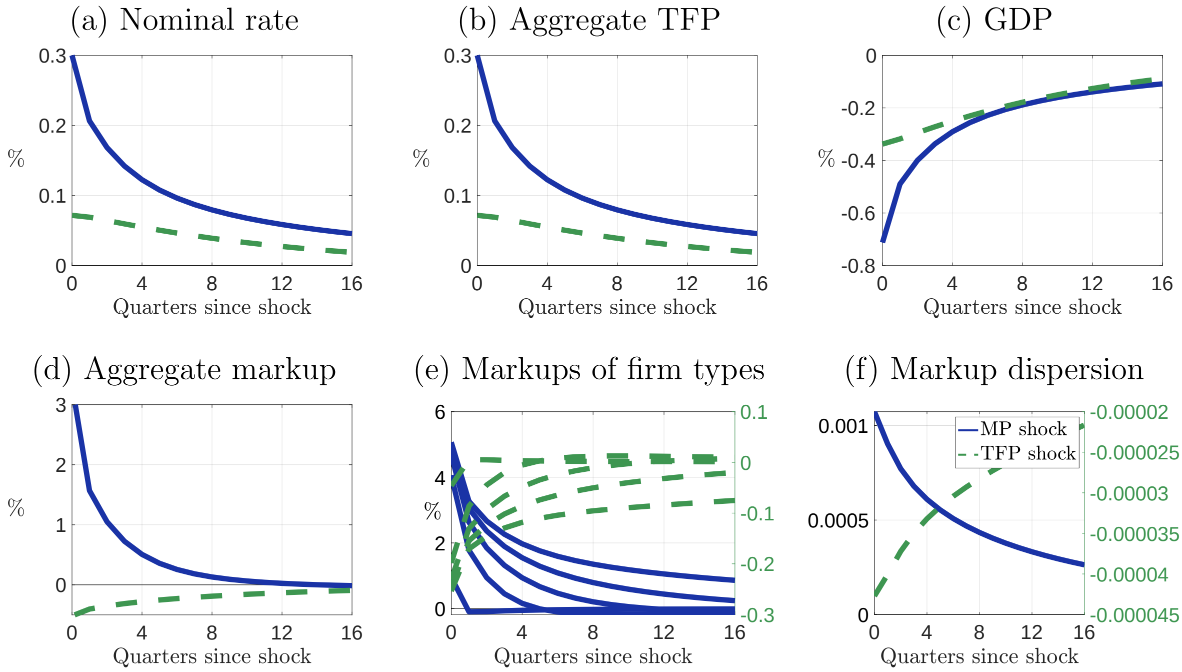


Notes: This figure shows impulse responses to a one standard deviation monetary policy shock. *Baseline* corresponds to the model in the main text. In particular, the monetary authority follows a Taylor rule, which reacts to fluctuation in the output gap. The gap is defined relative to natural output (the level prevailing under flexible prices), which is unchanged after monetary policy shocks. *Alternative Taylor rule* describes a policy counterfactual in which the monetary authority computes natural output as if the responses of aggregate TFP to monetary policy shocks were driven by technology shocks.

the responses to a monetary policy shock in the baseline model and the counterfactual exercise. Both GDP and markup dispersion respond by more on impact and more persistently in the counterfactual.

1.H.3 Responses to technology shocks

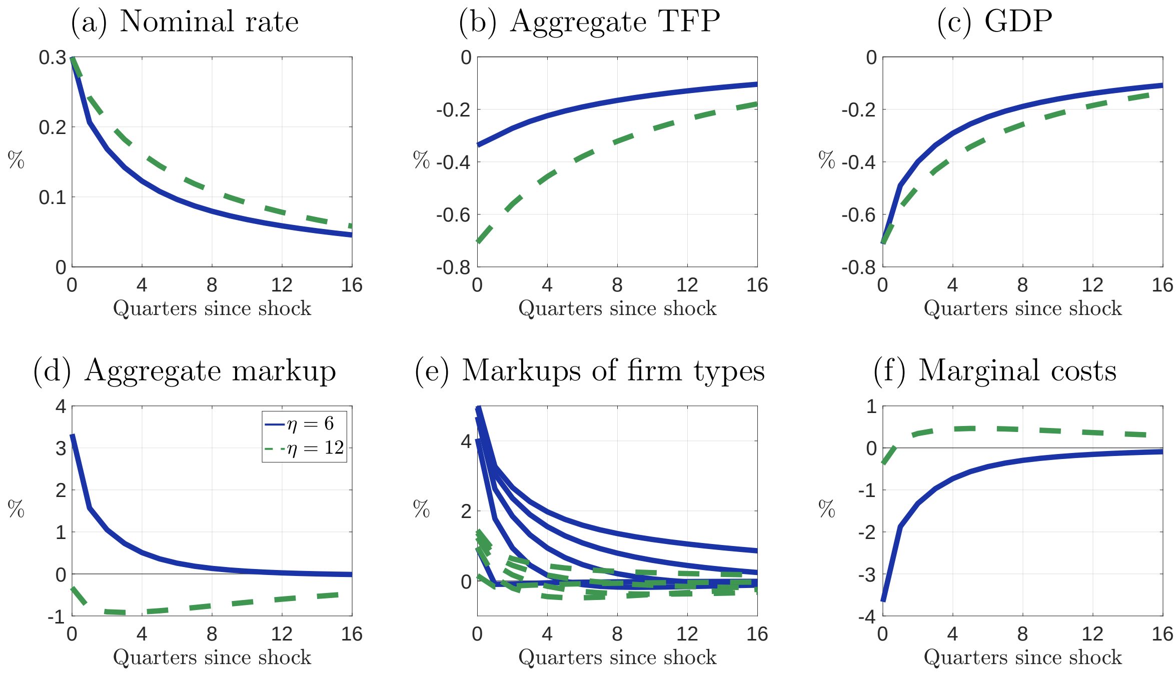
Figure 1.H.3: Model responses to technology shocks



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.

1.H.4 Varying the elasticity of substitution

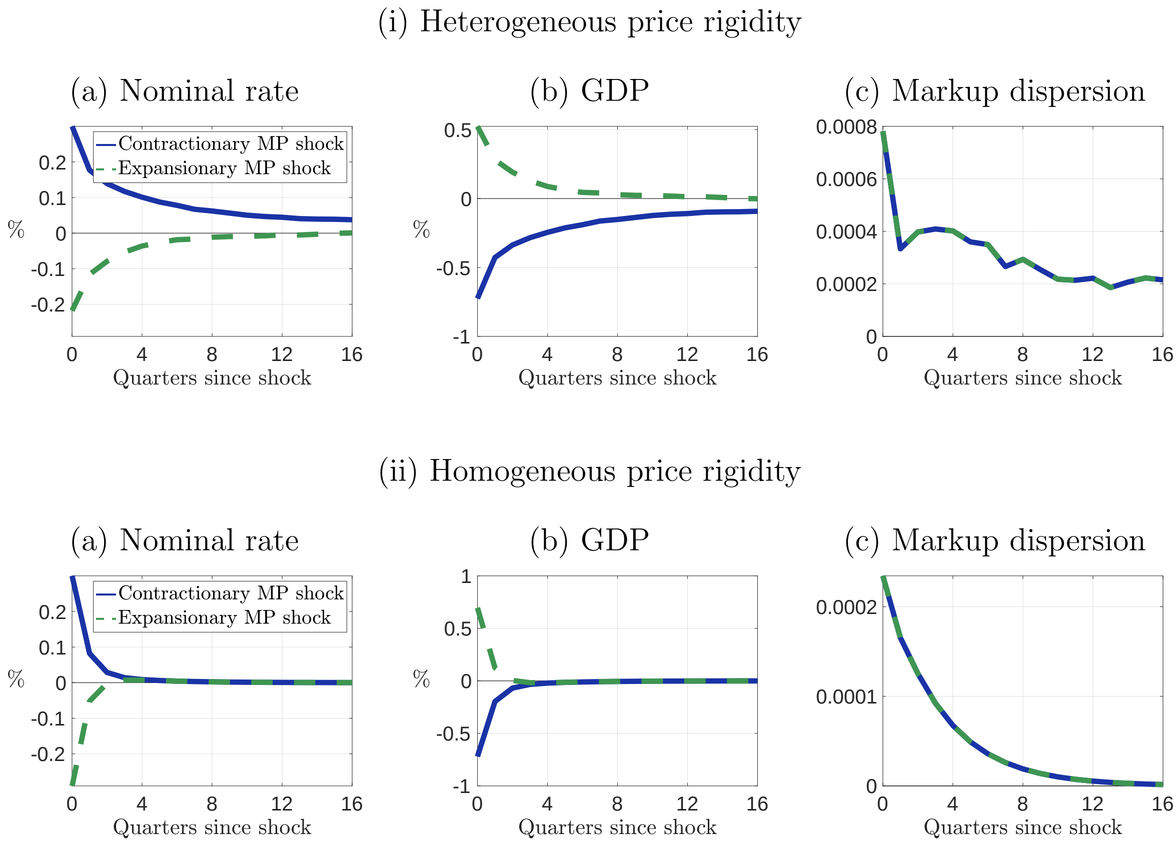
Figure 1.H.4: Model responses to monetary policy shocks when varying the elasticity of substitution



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 is re-calibrated to match the response of the nominal rate of 30bp.

1.H.5 Solution around the deterministic steady state

Figure 1.H.5: Model responses to monetary policy shocks around deterministic steady state



Notes: This figure shows responses to a one standard deviation monetary policy shock, when solving the model through a second-order approximation around the deterministic steady state.

1.I Robustness to model variations

To investigate the robustness of our quantitative results, we analyze the effects of monetary policy shocks in a number of model variations. These include a model with real rigidities, a model with Rotemberg price adjustment, and a model with trend inflation.

1.I.1 Real rigidities (firm-specific labor)

We model real rigidities via firm-specific labor. In particular, households supply differentiated labor, which is firm-specific and immobile across firms. In this section we show that under a condition similar to the one in Proposition 2, firms with more rigid prices optimally set higher prices. Proposition 1 then suggests that contractionary monetary policy shocks raise markup dispersion in this model. In a variant of the baseline model, in which labor is firm-specific, we find that the presence of such real rigidity amplifies the TFP effects of monetary policy.

We integrate the assumptions on differentiated labor supply from Woodford (2003, ch. 3) into our model of heterogeneous price rigidity. Households supply differentiated labor inputs N_{it} specific to each differentiated goods producer with preferences

$$E_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^{1-\frac{1}{\gamma}}}{1-\frac{1}{\gamma}} - \int_0^K \frac{N_{it}^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}} di \right).$$

Wages may differ across the differentiated labor inputs and we assume that the producers of differentiated goods take wages as given.

Compared to the model with homogeneous labor supply, described by its equilibrium conditions in Appendix 1.G.2, the main difference is the marginal cost, which is now firm-specific. For a firm of type k, which adjusts its price in period t, period $t+j$ marginal costs (conditional on no price re-adjustment after period t) are given by

$$mc_{k,t+j|t} = \frac{w_{k,t+j|t}}{A_{t+j}}.$$

We assume that firm k takes the wage $w_{k,t}$ as given. Using the intratemporal optimality condition, production technology, and demand curve, we obtain

$$mc_{k,t+j|t} = \left(\left(\frac{\tilde{P}_{kt}}{P_t} \frac{P_t}{P_{t+j}} \right)^{-\eta} \frac{Y_{t+j}}{A_{t+j}} \right)^{\frac{1}{\varphi}} \frac{C_{t+j}^{\frac{1}{\gamma}}}{A_{t+j}} = \left(\frac{\tilde{P}_{kt}}{P_t} \right)^{-\frac{\eta}{\varphi}} \Pi_{t,t+j}^{\frac{\eta}{\varphi}} \tilde{mc}_{t+j}.$$

where $\tilde{mc}_{t+j}$ collects aggregate variables in period $t+j$. The optimal reset price of a type k firm is

$$\frac{\tilde{P}_{kt}}{P_t} = \frac{\eta}{\eta-1} \frac{E_t \sum_{j=0}^{\infty} \theta_k^j \beta^j \mu_{t+j} Y_{t+j} \Pi_{t,t+j}^{\eta} mc_{kt+j|t}}{E_t \sum_{j=0}^{\infty} \theta_k^j \beta^j \mu_{t+j} Y_{t+j} \Pi_{t,t+j}^{\eta-1}},$$

which we can rewrite, using the above expression for $mc_{kt+j|t}$, as

$$\left(\frac{\tilde{P}_{kt}}{P_t}\right)^{1+\frac{\eta}{\varphi}} = \frac{\eta}{\eta-1}\frac{E_t\sum_{j=0}^{\infty}\theta_k^j\beta^j\mu_{t+j}Y_{t+j}\Pi_{t,t+j}^{\eta+\frac{\eta}{\varphi}}\tilde{mc}_{t+j}}{E_t\sum_{j=0}^{\infty}\theta_k^j\beta^j\mu_{t+j}Y_{t+j}\Pi_{t,t+j}^{\eta-1}}.$$

This reset price may feature precautionary price setting. In fact, when applying the partial equilibrium framework used in Proposition 2, one can show that the reset price increases in θ_k if

$$\left[\tilde{\eta}(\tilde{\eta}-1)-(\eta-1)(\eta-2)\right]\sigma_p^2 + \left[\tilde{\eta}-(\eta-1)\right]\sigma_{py} + \tilde{\eta}\sigma_{px} + \sigma_{xy} > 0,$$

where $\tilde{\eta} = \eta + \frac{\eta}{\varphi}$. Since marginal costs fall in the reset price, the markup also increases in θ_k if the above condition is met. This condition is similar to the setup with homogeneous labor. A sufficient condition for precautionary price setting is that aggregate prices have positive variance, and that the covariances between prices, aggregate demand, and the aggregate component of real marginal costs, are non-negative.

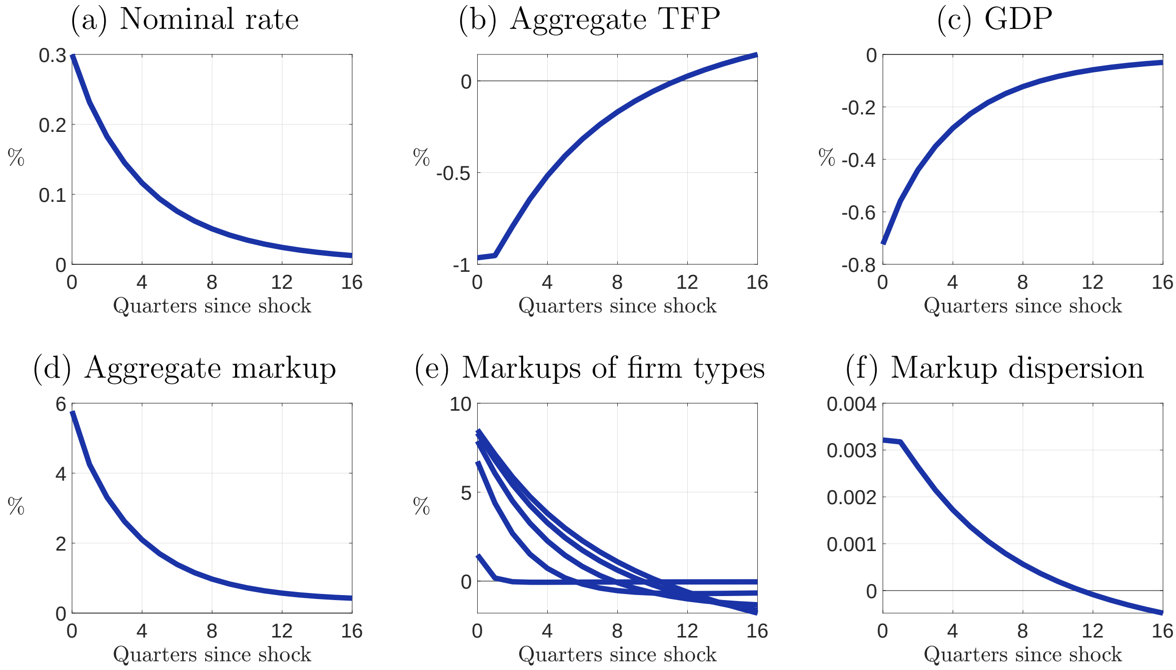
Importantly, if the above condition is satisfied, markups are negatively correlated with pass-through. Hence, Proposition 1 implies that markup dispersion increases in response to monetary policy shocks that lower real marginal costs. Aggregate TFP, computed as the Solow residual of an aggregate production function, then falls. The derivation of the Solow residual is identical to the model with homogeneous labor supply (and independent from assumptions on labor supply). If we aggregate labor input across firms, substitute in the production technology for differentiated goods ($Y_i = AN_i$), and the CES demand curve for differentiated goods, we obtain

$$N_t = \int N_{it}di = \int \frac{Y_{it}}{A_t}di = \int \frac{1}{A_t}\left(\frac{P_{it}}{P_t}\right)^{-\eta} Y_t di = \frac{Y_t}{A_t}\underbrace{\int\left(\frac{P_{it}}{P_t}\right)^{-\eta} di}_{=S_t}$$

and hence $Y_t = \frac{A_t}{S_t}N_t = TFP_tN_t$. Figure 1.I.1 below quantifies the response of markup dispersion and aggregate TFP to monetary policy shocks. For comparability with the baseline (homogeneous labor) model, we keep all model parameters unchanged except for the standard deviation of monetary policy shocks, which we re-calibrate to imply a 30bp increase of the nominal interest rate.

Finally, we next quantify the response of markup dispersion and aggregate TFP to monetary policy shocks in general equilibrium. For comparability with the baseline (homogeneous labor) model, we keep all model parameters unchanged except for the standard deviation of monetary policy shocks, which we re-calibrate to imply a 30bp increase of the nominal interest rate. Figure 1.I.1 shows that the response of markup dispersion and aggregate TFP to the shock is strongly amplified compared to the baseline model. The peak decline in aggregate TFP is 0.96% compared to 0.34% in the baseline model.

Figure 1.I.1: Model responses to monetary policy shocks when assuming specific labor

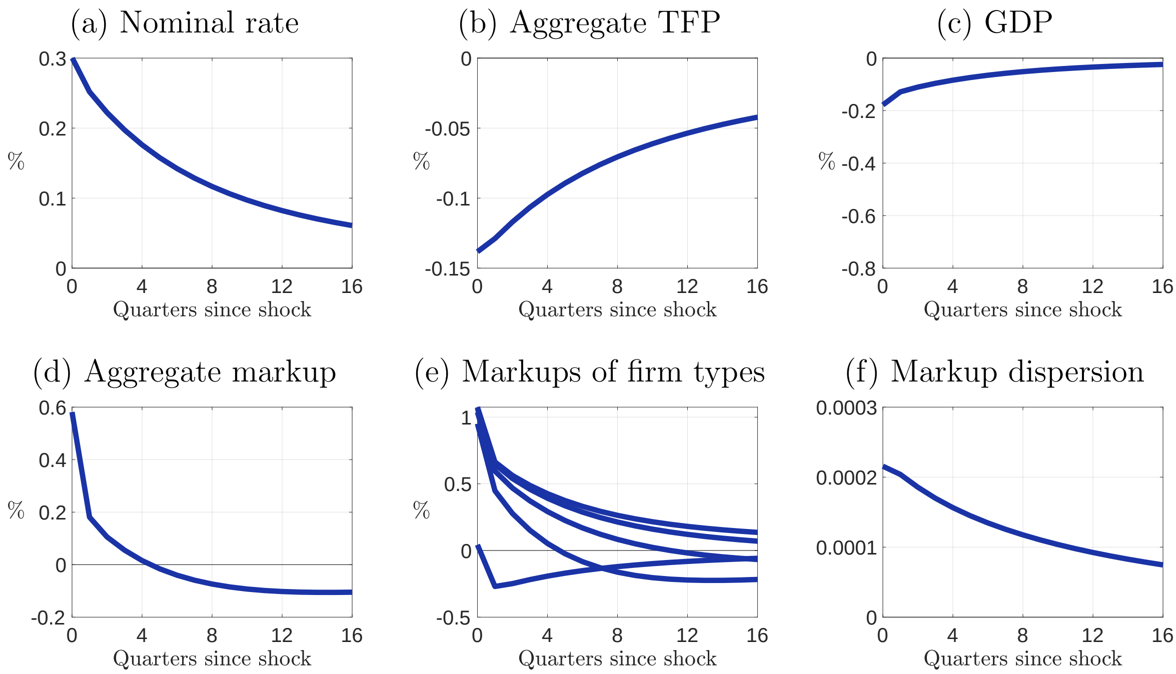


Notes: This figure shows responses to a one standard deviation contractionary monetary policy shock.

1.I.2 Rotemberg adjustment costs

In the Rotemberg version of our model, we assume price adjustment costs of the form $\frac{\phi_i}{2}\left(\frac{P_{it}}{P_{it-1}}-1\right)^2$, as in Section 3.2. We assume ϕ_i differs across 5 quintile groups of firms similar to our Calvo model. We calibrate the ϕ_i for $i=1,\ldots,5$ to match the markups (in the stochastic steady state) of our baseline Calvo model. In addition, we re-calibrate the standard deviation of monetary policy shocks to match a 30 bp response of the nominal interest rate. We leave all other parameters unchanged. The Rotemberg model can exactly match the steady state markups of the baseline model. For example, firms in the most rigid quintile set 11.5% higher markups than firms in the most flexible quintile. This shows that quantitatively strong precautionary price setting motives are not per se limited to the Calvo model. At the same time, the calibrated ϕ_i are not unreasonably large in the sense that monetary policy shocks do not have larger real effects than our baseline Calvo model. In fact, monetary policy shocks generate a 2/3 smaller GDP and a 1/3 smaller TFP response, see Figure 1.I.2.

Figure 1.I.2: Model responses to monetary policy shocks when assuming Rotemberg adjustment costs

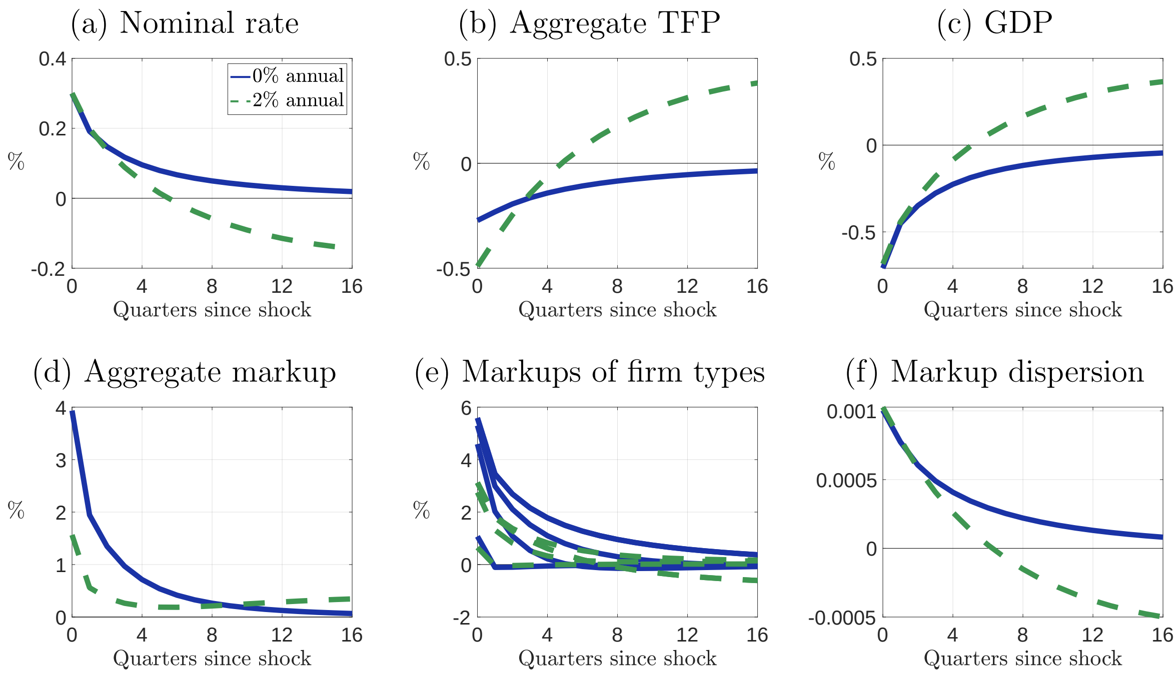


Notes: This figure shows responses to a one standard deviation contractionary monetary policy shock.

1.I.3 Trend inflation

We extend our baseline Calvo model by deterministic trend inflation as in Ascari and Sbordone (2014). We assume an annualized trend inflation of 2%. In this model, we only recalibrate the standard deviation of monetary policy shocks to match the 30 bp response of the nominal interest rate. We leave all other parameters unchanged. We find amplified markup differences across firms with differently rigid prices. In the stochastic steady state, firms with the most rigid prices have 19% higher markups than firms with the most flexible prices. Figure 1.I.3 provides the impulse responses to a one-standard deviation monetary policy shock. On impact, monetary policy shocks generate an even larger increase in markup dispersion and thus drop in aggregate TFP.

Figure 1.I.3: Model responses to monetary policy shocks when assuming trend inflation



Notes: This figure shows responses to a one standard deviation contractionary monetary policy shock. We set θ_1 to the value of θ_2, because the price setting problem is not well defined for the baseline value of θ_1 when annual trend inflation is 2%.

Chapter 2

Subjective Housing Price Expectations, Falling Natural Rates, and the Optimal Inflation Target*

U.S. households' housing price expectations deviate systematically from rational expectations: (i) expectations are updated on average too sluggishly; (ii) following housing price changes, expectations initially underreact but subsequently overreact; (iii) households are overly optimistic (pessimistic) about capital gains when the price-to-rent ratio is high (low). We show that weak forms of capital gain extrapolation allow to simultaneously replicate the behavior of housing prices and these deviations from rational expectations as an equilibrium outcome. Embedding capital gain extrapolation into a sticky price model featuring a lower-bound constraint on nominal interest rates, we show that lower natural rates of interest increase the volatility of housing prices and thereby the volatility of the natural rate of interest. This exacerbates the relevance of the lower bound constraint and causes the optimal inflation target to increase strongly as the natural rate falls.

2.1 Introduction

The large and sustained booms and busts in housing prices in advanced economies are often attributed to households holding excessively optimistic or pessimistic beliefs about future housing prices (Piazzesi and Schneider (2006), Kaplan, Mitman, and Violante (2020)). This view is supported by a nascent literature that documents puzzling facts about the behavior of housing price expectations. Survey measures of expected future housing prices have been found to be influenced by past housing price changes, but appear to underreact to these changes, and they also miss the tendency of housing prices to mean revert over time (Kuchler and Zafar (2019), Case, Shiller, and Thompson (2012), Ma (2020) and Armona, Fuster, and Zafar (2018)).

Documenting in which ways households' housing price expectations deviate

*Joint work with Klaus Adam and Oliver Pfäuti.

from the rational expectations benchmark is an important task but remains in itself uninformative about how important the observed deviations are for economic outcomes in housing markets and for the conduct of monetary policy. Understanding these features requires a structural equilibrium model that quantitatively replicates how households' expectations deviate from rational expectations. Constructing such an equilibrium model, calibrating it to the behavior of household beliefs in survey data, and understanding its implications for the optimal design of monetary policy is the main objective of the present paper. To the best of our knowledge, it is the first paper pursuing this task.

We begin our analysis by comprehensively quantifying the dimensions along which households' housing price expectations deviate from the full-information rational expectations benchmark. To this end, we consider the Michigan Survey of Consumers, which provides the longest available time series of quantitative housing price expectations for the United States, covering the years 2007-2021.

We document three dimensions along which household expectations deviate from rational expectations. First, expectations about future housing prices are revised too sluggishly over time, a feature that housing price expectations share with other household expectations (Coibion and Gorodnichenko (2015)). Second, households' capital gain expectations covary positively with market valuation, i.e., the price-to-rent ratio, while actual future capital gains covary negatively with market valuation. We show that the difference is striking, highly statistically significant, and in line with findings on investor expectations in stock markets (Adam, Marcet, and Beutel (2017)).[1] Third, in a dynamic sense, households' capital gain expectations initially underreact to observed capital gains, i.e., households are too pessimistic in the first few quarters following a positive capital gain, but later on overreact, i.e., households hold too optimistic expectations after about twelve quarters. The pattern of initial underreaction and subsequent overreaction is similarly present in other macroeconomic expectations, see Angeletos, Huo, and Sastry (2020).

While the first and third deviation from rational housing price expectations have been documented before using different data sets, see Armona et al. (2018), the second deviation from rational expectations is new to the housing literature. We quantify here all three deviations using a single data set, so as to obtain a coherent set of quantitative targets for our structural equilibrium model with subjective housing expectations.

Equipped with these facts, we construct first a simple housing model with optimizing households that hold subjective beliefs about housing price behavior. Bayesian belief updating implies that households weakly extrapolate past capital gains into the future. The model reproduces – as an equilibrium outcome – important patterns of the behavior of U.S. housing prices, in particular, the large and protracted swings in the price-to-rent ratio over time, as well as the three dimensions mentioned above along which household expectations deviate from the rational expectations benchmark. The quantitative fit is surprisingly good, despite the simplicity of the model.

[1]For stock markets, Adam, Matveev, and Nagel (2021) show that this cannot be explained by investors reporting risk-adjusted expectations.

The simple model generates two important insights. First, it shows that the standard deviation for the price-to-rent ratio would be much lower in the presence of rational housing expectations. This suggests that the observed volatility of housing prices is to a significant extent due to the presence of subjective beliefs. This lends credence to the view that the observed deviations from rational expectations substantially contribute to booms and busts in housing markets.

Second, the simple model connects the secular decline in natural rates of interest with higher volatility of housing prices. Specifically, the model predicts that lower real interest rates imply larger effects of belief fluctuations on equilibrium housing prices. This prediction does not emerge in the presence of rational housing price expectations, but is consistent with the data. We show that in a number of advanced economies, including the United States, the volatility of housing prices has increased considerably at the same time as the level of the natural rate of interest has fallen.

The most important objective of paper is to understand the monetary policy implications generated by a setting where households (weakly) extrapolate capital gains into the future. We are particularly interested in the optimal policy response to increased housing price volatility that is induced by falling natural rates of interest in a setting where policy rates cannot move into negative territory. To this end, we introduce capital gain extrapolation into an otherwise standard New Keynesian model featuring a housing sector and a lower bound constraint on nominal interest rates.

The sticky price model has a number of attractive features. First, it shares the implications for housing price behavior and household beliefs with the simpler model considered before and thus quantitatively replicates the patterns of belief deviations and housing prices. Second, it is immune to the critique by Barsky, House, and Kimball (2007) regarding the behavior of sticky price models featuring durable goods. In line with the data, the model implies that housing demand reacts more strongly to monetary disturbances than non-housing demand, despite the fact that housing prices are fully flexible. Third, the model introduces subjective housing beliefs in a way that monetary policy is unable to manipulate household beliefs to its own advantage. This allows for a meaningful discussion of Ramsey optimal monetary policy in the presence of subjective beliefs. Finally, the model makes a minimal departure from rational expectations: expectations about non-housing related variables are rational and all agents maximize given their (subjective) beliefs about the future.

To gain analytic insights, we derive a linear-quadratic approximation to the optimal policy problem and show how it is affected by the presence of subjective housing beliefs. We find that housing price gaps, i.e., deviations of housing prices from their efficient level, affect the economy via two channels. First, inefficiently high housing prices, driven by capital gain optimism, give rise to negative cost-push terms in the Phillips curve.[2] This feature allows the model to potentially generate a non-inflationary housing boom. Yet, a second channel is more important: rising

[2]Conversely, inefficiently low housing prices, driven by capital gain pessimism, cause positive cost-push terms.

housing price volatility increases the volatility of the natural rate of interest. Since increased housing price volatility is itself triggered by a fall in the average level of the natural rate, this dramatically exacerbates the lower-bound problem for a monetary policy authority confronted with falling natural rates.

The natural rate is affected by housing prices, because higher housing prices make it optimal to allocate more resources to housing investment. This exerts positive pressure on the output gap and counteracting these – so as to keep the output gap stable – requires policy to increase the real interest rate. Under rational expectations, housing prices never deviate from their efficient value, so that policy never has to work against inefficient investment pressures. With rational expectations, the volatility of the natural rate is thus independent of the average level of the natural rate.

These contrasting predictions of the model under rational and subjective housing beliefs also lead to rather different policy messages on how the optimal inflation target, i.e., the average inflation rate implied by optimal monetary policy, should respond to a fall in the natural rate of interest. Under rational expectations, the optimal inflation target is nearly invariant to the average level of the natural rate.

In the presence of capital gain extrapolation, the optimal inflation target increases considerably in response to a fall in the average natural rate. This is due to the increased volatility in the natural rate and cost-push shocks, which causes the lower bound on the nominal rate to become more restrictive. A more restrictive lower bound forces monetary policy to rely more strongly on promising future inflation in order to lower the real interest rate. This increases the average inflation rate under optimal policy. For our calibrated model, we find that the optimal inflation target should increase approximately by one third of a percent in response to a one percent fall in the natural rate with the increase becoming non-linear for very low levels of the natural rate.

We also investigate the optimal policy response to housing demand shocks. While inflation and the output gap do not respond to these shocks under rational expectations, capital gains induced by housing demand shocks get amplified by capital gain extrapolation and thereby generate persistent housing price gaps to which monetary policy optimally responds. Housing price gaps, however, generate opposing effects. On the one hand, inefficiently high housing prices generate negative cost-push pressures, which calls for a decrease in the policy rate; on the other hand, inefficiently high housing prices trigger a housing investment boom, which puts upward pressure on the output gap. Counteracting this second effect requires hiking policy rates.

In our calibrated model, the second effect quantitatively dominates. Optimal monetary policy thus 'leans against' housing price movements, but the optimal strength of the reaction depends on the direction of the shock. Following a positive housing preference shock, the increase in the interest rate (nominal and real) is more pronounced than the interest rate decrease following a negative housing demand shock. The presence of the lower-bound constraint thus attenuates the degree to which monetary policy leans against negative housing demand shocks.

We also consider whether macroprudential policies could address the housing market inefficiencies generated by capital gain extrapolation. We do so by considering housing taxes that might be levied on households in order to insulate monetary policymakers from the fluctuations in the housing price gap. We find that the required taxes would have to be large and very volatile. For our calibrated model, taxes must often exceed 20% of the rental value of housing per period and also often require equally sized or even larger housing subsidies. It appears somewhat unlikely that any of the existing macroprudential tools are capable of generating effects of such magnitude. And to the best of our knowledge, none of the available macroprudential tools allows subsidizing private sector behavior. Less aggressive tax policies turn out to be considerably less effective in bringing down the volatility of the housing price gap. This suggests that macroprudential policies are unable to substantially reduce the monetary policy trade-offs arising from subjective housing price expectations.

This paper is related to work by Andrade, Galí, Le Bihan, and Matheron (2019, 2021) who study how the optimal inflation target depends on the natural rate of interest in a setting with a lower bound constraint. In line with our findings, they show that an increase in the inflation target is a promising approach to deal with the lower-bound problem. While their work considers optimized Taylor rules in a medium-scale sticky price model without a housing sector and rational expectations, the present paper studies Ramsey optimal policy in a model featuring a housing sector and subjective housing expectations.

A number of papers consider Ramsey optimal policy in the presence of a lower-bound constraint, but also abstract from housing markets and the presence of subjective beliefs (Eggertsson and Woodford (2003), Adam and Billi (2006), Coibion, Gorodnichenko, and Wieland (2012)). This literature finds that lower bound episodes tend to be short and infrequent under optimal policy, so that average inflation is very close to zero under optimal policy. The present paper shows that this conclusion is substantially altered in the presence of subjective housing price expectations.

Optimal monetary policy with subjective beliefs has previously been analyzed in Caines and Winkler (2021) and Adam and Woodford (2021). These papers abstract from the lower-bound constraint and consider different belief setups that are not calibrated to replicate patterns of deviations from rational housing price expectations as observed in survey data.[3] We show that taking into account the existence of a lower-bound constraint on nominal rates is quantitatively important for understanding how the optimal inflation target responds to lower natural rates.

The rest of the paper is structured as follows. Section 2.2 documents how survey expectations about future housing prices deviate from rational expectations. Section 2.3 presents a simple housing model in which households extrapolate capital gains. It shows how this simple model can jointly replicate in equilibrium the behavior of housing prices and the pattern of deviations from rational expectations. Section 2.4 then presents the full housing model with sticky prices,

[3]Adam and Woodford (2021) consider 'worst-case' belief distortions, while Caines and Winkler (2021) consider a setting with 'conditionally model-consistent beliefs'. Both setups generate deviations from rational expectations for variables other than housing prices.

subjective housing beliefs, and a lower-bound constraint on nominal rates. Section 2.5 derives a quadratic approximation to the monetary policy problem, which allows obtaining important analytic insights into the new economic forces arising from the presence of subjective housing price beliefs. We calibrate the model in Section 2.6 and present quantitative results about the optimal inflation target and the optimal policy response to housing shocks in Section 2.7. Section 2.8 discusses macroprudential policies and Section 2.9 concludes.

2.2 Properties of housing price expectations

This section documents that households' housing price expectations deviate in systematic ways from the full-information rational expectations (RE) benchmark. We consider three rationality tests that have recently been proposed in the literature (Coibion and Gorodnichenko (2015), Adam et al. (2017) and Angeletos et al. (2020)). These tests cover different dimensions along which subjective expectations deviate from RE.

We measure housing prices using the S&P/Case-Shiller U.S. National Home Price Index and let q_t denote the quarterly average of the monthly housing price index. We consider both nominal and real housing prices with real housing prices being obtained by deflating nominal housing prices with the CPI.[4][5]

Expectations about housing capital gains are taken from the Michigan household survey. The survey provides subjective expectations about nominal four-quarter-ahead housing price growth, $E_t^{\mathcal{P}}[q_{t+4}/q_t]$, for the period 2007-2021. The survey also provides housing price growth expectations over the next five years. We focus on the shorter horizon because these expectations determine housing prices according to our model. The shorter horizon also allows performing a dynamic decomposition of forecast errors over time in response to realized capital gains.[6]

We consider both mean and median household expectations.[7] When considering real housing price expectations, we deflate the nominal mean (median) capital gain expectations with the mean (median) inflation expectation over the same period, as obtained from the Michigan survey.[8]

[4]The simplified model in the next section makes predictions about real housing prices only, while the survey data contains information about nominal capital gain expectations. This leads us to consider nominal and real housing prices.

[5]We use the "Consumer Price Index for All Urban Consumers: All Items in U.S. City Average"obtained from FRED.

[6]Data limitations make such a decomposition difficult for five-year-ahead forecasts: with only 15 years of data, the dynamic decompositions become largely insignificant. Appendix 2.A.1 shows, however, that all other patterns documented below are equally present in five-year-ahead expectations.

[7]Analyzing the dynamics of individual expectations over time is difficult because households in the Michigan survey are sampled at most twice. In general, cross-sectional disagreement between households might partly reflect heterogeneous information on the part of households, see Kohlhas and Walther (2021).

[8]As is well-known, these inflation expectations feature an upward bias relative to actual inflation outcomes. This, however, will not be the source of rejection of the RE hypothesis:

Table 2.1: Sluggish adjustment of housing price expectations

	Mean Expectations	Median Expectations
Nominal Housing Prices		
$\widehat{b}^{CG}$	2.22***	2.85***
	(0.507)	(0.513)
Real Housing Prices		
$\widehat{b}^{CG}$	2.00***	2.47***
	(0.332)	(0.366)

Notes: This figure shows the empirical estimates of regression (2.1) for nominal and real housings price and considers mean and median expectations. The reported standard errors are robust with respect to heteroskedasticity and serial correlation (Newey–West with four lags). Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Sluggish updating about the expected housing price level. We start by documenting that the mean/median household expectation about the future level of housing prices is updated too sluggishly. This can be tested following the approach of Coibion and Gorodnichenko (2015), which involves considering regressions of the form

$$q_{t+4} - E_t^{\mathcal{P}}\left[q_{t+4}\right] = a^{CG} + b^{CG} \cdot \left(E_t^{\mathcal{P}}\left[q_{t+4}\right] - E_{t-1}^{\mathcal{P}}\left[q_{t+3}\right]\right) + \varepsilon_t. \tag{2.1}$$

The regression projects forecast errors about the future housing price level on past forecast revisions. Under the RE hypothesis, information that is contained in agents' information set, i.e., past forecasts and their revisions, should not predict future forecast errors ($H_0 : b^{CG} = 0$).

We estimate equation (2.1) for nominal and real capital gains, using mean and median expectations, respectively. Expectations of the future house price level are computed as $E_t^{\mathcal{P}}\left[q_{t+4}\right] = E_t^{\mathcal{P}}\left[q_{t+4}/q_t\right] q_t$, where $E_t^{\mathcal{P}}\left[q_{t+4}/q_t\right]$ denotes the capital gain expectations from the Michigan survey and q_t the S&P/Case-Shiller Index.[9]

Table 2.1 reports the estimated b^{CG} from regression (2.1). We find that $\widehat{b}^{CG} > 0$, which is inconsistent with the RE hypothesis. The regression coefficient is positive and statistically significant at the 1% level in all considered specifications. This implies that the mean/median agent updates beliefs too sluggishly: future realizations move (on average) by more than what is suggested by past forecast revisions. The magnitude of the estimates is also large in economic terms: a coefficient estimate of two suggests that forecast revisions should approximately be three times as strong than they actually are.

Overall, sluggish belief updating is consistent with previous findings on the behavior of survey expectations about output, inflation and unemployment (Coibion

all our tests focus on the cyclical properties of capital gain expectations and eliminate mean differences between forecasts and realizations using appropriate regression constants that will not be used in our rationality tests.

[9] When considering real housing prices, nominal capital gain expectations from the Michigan survey are deflated using the subjective (mean or median) inflation expectations from the Michigan survey.

and Gorodnichenko (2015), Angeletos et al. (2020), Kohlhas and Walther (2021)). Furthermore, Bordalo, Gennaioli, Ma, and Shleifer (2020) provide evidence of sluggish belief adjustment in consensus forecasts for other housing variables, such as residential investment and new housing starts.

Appendix 2.A.2 shows that our findings are robust to using an instrumental-variable approach for estimating regression (2.1), in which forecast revisions are instrumented with monetary policy shocks obtained via high-frequency identification. Appendix 2.A.3 shows that similar results emerge when using capital gains and expected capital gains in equation 2.1 instead of the level and expected level of the housing price.

Opposing cyclicality of actual and expected capital gains. Our second test documents the different cyclicality of actual and expected capital gains in housing markets. Differences between the cyclicality of actual and expected capital gains have previously been documented for stock markets, where actual and expected stock market capital gains covary differently with the price-to-dividend ratio (Greenwood and Shleifer (2014), Adam et al. (2017)). We consider here the cyclicality of expected and actual capital gains in the housing market with the price-to-rent ratio PR:

$$E_t^{\mathcal{P}}\left[\frac{q_{t+4}}{q_t}\right] = a + c \cdot PR_{t-1} + u_t \tag{2.2}$$

$$\frac{q_{t+4}}{q_t} = \mathbf{a} + \mathbf{c} \cdot PR_{t-1} + \mathbf{u}_t. \tag{2.3}$$

The rational expectations hypothesis implies $H_0 : c = \mathbf{c}$, whenever the agents' information set contains the past price-to-rent ratio as an observable.[10] Since the predictor variable used on the right-hand side of the preceding regressions equations is highly persistent, we correct for small sample bias in the coefficient estimates (Stambaugh (1999)).[11]

Table 2.2 reports the regression results. It shows that expected capital gains are positively associated with the PR-ratio, while realized capital gains are negatively associated. Expected capital gains are pro-cyclical, i.e., are high when market valuation is high, while realized capital gains are counter-cyclical, i.e., are low when market valuation is high. This pattern of is akin to the one documented in stock markets.

Quantitatively, the results imply that a two standard deviation increase of the PR-ratio by 15.5 units increases the mean household expectations about four-quarter-ahead real capital gains by around 0.5%. Actual four-quarter ahead capital gains, however, fall by around 1.5%, so that the forecast error is approximately 2%.

[10] In the regressions, we use the lagged PR-ratio, PR_{t-1}, instead of the current value, because the PR-ratio is computed using the average price over a quarter. In Adam et al. (2017) the price-to-dividend ratio was computed using the beginning of quarter stock price, which allowed using the current value in the regression.

[11] The small sample bias correction in Table 2.2 follows the same approach as the one in Table 1A in Adam et al. (2017).

Table 2.2: Cyclicality of expected vs. actual capital gains

	$\hat{c}$ (in %)	$\hat{\mathbf{c}}$ (in %)	bias (in %) $-E(\hat{\mathbf{c}} - \hat{c})$	p-value $H_0 : c = \mathbf{c}$
Nominal Housing Prices				
Mean Expectations	0.033	-0.102	0.006	0.000
	(0.008)	(0.007)		
Median Expectations	0.014	-0.102	0.009	0.000
	(0.001)	(0.007)		
Real Housing Prices				
Mean Expectations	0.030	-0.113	-0.003	0.000
	(0.017)	(0.009)		
Median Expectations	0.010	-0.113	0.006	0.000
	(0.004)	(0.009)		

Notes: $\widehat{c}$ is the estimate of c in equation (2.2) and $\widehat{\mathbf{c}}$ the estimate of $\mathbf{c}$ in equation (2.3). The Stambaugh (1999) small sample bias correction is reported in the second-to-last column and the last column reports the p-values for the null hypothesis $c = \mathbf{c}$. Newey–West standard errors using four lags are in parentheses.

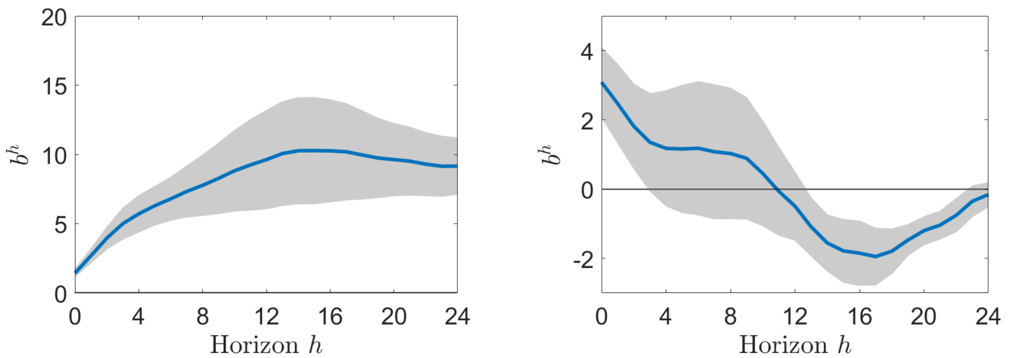
The last column in Table 2.2 performs a test of the rational expectations hypothesis that the cyclicality of actual and expected returns are equal ($H_0 : c = \mathbf{c}$). The test corrects for small sample bias, which is reported in the second to last column. We find that the difference in the cyclicality of actual and expected capital gains is highly statistically significant in all cases. Appendix 2.A.4 shows that similar results are obtained when first subtracting equation (2.2) from (2.3) and estimating the resulting equation with forecast errors on the left-hand side, as in Kohlhas and Walther (2021) who do not consider housing related variables.

Initial under- and subsequent over-reaction of housing price expectations. While the results in Table 2.1 show that households adjust short-term housing price beliefs on average too sluggishly, the results in Table 2.2 indicate over-optimism (over-pessimism) in housing price expectations when the current market valuation is high (low), which points to some form of overreaction to past housing price increases. It turns out that both patterns can be jointly understood by considering the dynamic response of actual and expected capital gains to housing price changes.

Following the approach in Angeletos et al. (2020), who analyze forecast errors about unemployment and inflation, we investigate how capital gains and forecast errors about these capital gains evolve over time in response to realized capital gains.[12] Provided households observe realized capital gains, the RE hypothesis implies that it should not be possible to predict future forecast errors with current

[12]These dynamic responses are well-defined in econometric terms, even if they cannot be given a structural interpretation, because past capital gains are likely driven by a combination of past shocks.

Figure 2.1: Dynamic responses to a realized capital gain



Notes: Panel (a) shows the dynamic response of cumulative real capital gains at horizon h to a one standard deviation innovation in the housing capital gain. Panel (b) reports the dynamic response of housing-price forecast errors at horizon h of one-year ahead expectations to a one standard deviation innovation in the housing capital gain. Positive (negative) values indicate that realized capital gains exceed (fall short of) expected capital gains. The shaded area shows the 90%-confidence intervals, standard errors are robust with respect to serial correlation and heteroskedasticity (Newey–West with $h+1$ lags).

capital gains.

We estimate the dynamic responses using local projections (Jorda (2005)) of the form

$$X_{t+h} = a^h + b^h \frac{q_{t-1}}{q_{t-2}} + u_t^h, \tag{2.4}$$

where the left-hand side variable X_{t+h} is either the cumulative capital gain (q_{t+h+4}/q_t), or the forecast error about the four-quarter-ahead capital gain $(q_{t+h+4}/q_{t+h} - E^{\mathcal{P}}_{t+h}[q_{t+h+4}/q_{t+h}])$, and u_t^h a serially correlated and heteroskedastic error term. Note that forecast errors are positive when households are overly pessimistic about capital gains and negative if households are overly optimistic.

Figure 2.1 reports the estimated coefficients b^h from local projection (2.4). Panel (a) depicts the response of cumulative capital gains. It shows that the initial capital gains is not only persistent, but increases further over time, reaching a plateau after around twelve quarters. Given the high serial correlation displayed by capital gains in housing markets, this feature is perhaps not too surprising.

Panel (b) depicts the dynamic response of forecast errors. Forecast errors are initially positive but later on – once cumulative capital gains reach their plateau – become negative before eventually disappearing. The positive values initial periods indicates that agents' expectations react too sluggishly: realized capital gains are persistently larger than the expected gains. This also implies an underreaction in terms of the expected level of housing prices. Subsequently, when all actual capital gains have materialized and housing prices start to slightly mean revert, agents are too optimistic about future capital gains. This aligns well with the prior observation that capital gain expectations display the wrong cyclicality with housing market valuation.[13] It also implies that households entirely miss the mean-reversion in capital gains: forecast errors turn negative once housing prices level off and start to slightly mean-revert. This pattern is consistent with the experimental evidence provided in Armona et al. (2018).

In Appendix 2.A.5, we show that the nominal forecast error responses look very similar. Likewise, using median expectations instead of mean expectations makes no noticeable difference of the results. In Appendix 2.A.6, we show all our results obtained thus far are robust to excluding the Corona Virus period, i.e., to letting the sample period end in the last quarter of 2019.

Analysis using regional data. As is well known, housing prices often display considerable regional variation across the United States. We thus check whether the three deviations from the RE documented above are also present in regional housing prices and housing price beliefs. Appendix 2.A.7 uses regional housing price indices and exploits local information contained in the Michigan survey that allows grouping survey respondents into four different U.S. regions (North East, North Central/Midwest, South, and West). Repeating the above analyses at the regional level, it shows that one obtains quantitatively similar results.

The next section presents a simple housing model that can quantitatively replicate the forecast error deviations documented in this section.

2.3 Simple model with capital gain extrapolation

This section presents a bare-bones housing model in which households (weakly) extrapolate past capital gains. The model makes equilibrium predictions for the joint dynamics of housing prices and housing price beliefs. Housing prices in the model depend on households' housing price beliefs, with the latter being influenced by past housing price behavior. We show that equilibrium dynamics of housing prices and housing price beliefs quantitatively replicate key features of housing price behavior in the U.S., as well as the deviations from rational expectations documented in the previous section. The simple model also predicts that low levels of the natural rate of interest give rise to increased housing price volatility. As we show, this prediction is consistent with the evolution of natural rates and housing prices in advanced economies over the past decades.

The full model in Section 2.4 additionally features nominal rigidities, a lower bound constraint on nominal rates, generalized preferences, and endogenous production of consumption goods and housing. The present section abstracts from these features, but nevertheless shares its implications for housing price behavior and housing price beliefs with the full model.

[13] Since rents move only very slowly over time, changes in housing prices capture changes in the price-to-rent ratio rather well.

The household problem. There is a measure one of identical households.[14] Households are internally rational, as in Adam and Marcet (2011), i.e., maximize utility holding potentially subjective beliefs about variables beyond their control. The representative household chooses consumption C_t, housing units to own D_t, and housing units to rent D_t^R, to maximize

$$\max_{\left\{C_t \geq 0, D_t \in [0, D^{\max}], D_t^R \geq 0\right\}_{t=0}^{\infty}} \quad E_t^{\mathcal{P}} \sum_{t=0}^{\infty} \beta^t \left[C_t + \xi_t^d \left(D_t + D_t^R\right)\right]$$
$$\text{s.t.:} \quad C_t + \left(D_t - (1-\delta) D_{t-1}\right) q_t + R_t D_t^R = Y_t \text{ for all } t \geq 0,$$

where Y_t is an exogenous (and sufficiently large) endowment, q_t the real price of housing, R_t the real rental price and $\delta > 0$ the housing depreciation rate. Rental units and housing units owned are perfect substitutes and $\xi_t^d \geq 0$ denotes a housing preference shock. The household's subjective probability measure $\mathcal{P}$ allows for subjective housing price beliefs. For simplicity, we assume beliefs about other variables beyond the household's control, $\{Y_t, \xi_t^d, R_t\}_{t=1}^{\infty}$, to be rational. The latter assumption is not important for the results derived in this section.

Housing choices are subject to a short-selling constraint $D_t \geq 0$, which is standard, and to a long constraint $D_t \leq D^{\max}$. The latter insures existence of optimal plans in the presence of distorted housing beliefs. The long constraint is chosen such that it will never bind in equilibrium, i.e., $D^{\max} > D$, where D denotes the exogenously fixed housing supply. Without loss of generality, rental units are assumed to be in zero net supply.

The household first-order conditions imply that rents are given by

$$R_t = \xi_t^d \tag{2.1}$$

and that equilibrium housing prices satisfy[15]

$$q_t = \xi_t^d + \beta(1-\delta) E_t^{\mathcal{P}} q_{t+1}. \tag{2.2}$$

Capital gain extrapolation. We now introduce subjective price beliefs that give rise to capital gain extrapolation, using the setup in Adam and Woodford (2012). Importantly, the precise details generating capital gain extrapolation are not essential for the results in this section and we could have used alternative belief assumptions, e.g., learning from life-time experience as in Nagel and Xu (2018) and Malmendier and Nagel (2011, 2015), or could have directly assumed extrapolative behavior as in Barberis, Greenwood, Jin, and Shleifer (2015).

Households perceive capital gains to evolve according to

$$\frac{q_t}{q_{t-1}} = b_t + \varepsilon_t, \tag{2.3}$$

[14] The fact that households are identical is not common knowledge among households.

[15] This holds true in equilibrium because $0 < D < D^{\max}$. For the household, however, first-order conditions may hold only with inequality under the subjectively optimal plans, due to the presence of short and long constraints. The latter explains why rational households can hold price expectations that differ from the discounted sum of future rents, see Adam and Marcet (2011) for details and Adam and Nagel (2022) for related arguments.

where $\varepsilon_t \sim iiN(0, \sigma_\varepsilon^2)$ is a transitory component of capital gains and b_t a persistent component, which itself evolves according to $b_t = b_{t-1} + \nu_t$, with $\nu_t \sim iiN(0, \sigma_\nu^2)$.[16] Households observe the realized capital gains (q_t/q_{t-1}) and use Bayesian belief updating to decompose observed capital gains into their persistent and transitory components. With conjugate prior beliefs, the subjective conditional one-step-ahead capital gain expectations

$$\beta_t \equiv E_t^{\mathcal{P}}\left(q_{t+1}/q_t\right) \tag{2.4}$$

evolve according to

$$\beta_t = \min\left\{\beta_{t-1} + \frac{1}{\alpha}\left(\frac{q_{t-1}}{q_{t-2}} - \beta_{t-1}\right), \beta^U\right\}, \tag{2.5}$$

where $1/\alpha$ is the Kalman gain determining how strongly households' capital gain expectations respond to past capital gain surprises.[17] The Kalman gain thus captures the degree to which past capital gain surprises are extrapolated into the future. The upper bound β^U on the beliefs in equation (2.5) is there to insure that capital gain optimism is bounded from above, so as to keep subjectively expected utility finite.[18]

Figure 2.1 illustrates the relationship between belief revisions and forecast errors implied by equation (2.5) using the Michigan survey data. The figure plots on the vertical axis a measure of the quarterly revision in capital gain expectations, $\left(E_t^{\mathcal{P}}\left(q_{t+4}/q_t\right)\right)^{\frac{1}{4}} - \left(E_{t-1}^{\mathcal{P}}\left(q_{t+3}/q_{t-1}\right)\right)^{\frac{1}{4}}$, and on the horizontal axis a measure of the forecast error in quarterly capital gains, $\frac{q_t}{q_{t-1}} - \left(E_{t-1}^{\mathcal{P}}\left(q_{t+3}/q_{t-1}\right)\right)^{\frac{1}{4}}$, for all quarters in the Michigan survey. Consistent with equation (2.5), there is a clear positive and approximately linear relationship between capital gain surprises and belief revisions in Figure 2.1. The most notable deviations from the regression line are the ones around the Great Recession (2008Q3 and 2009Q2) and the Covid Recession (2020Q2 and 2020Q3).

Equilibrium dynamics of housing prices and capital gain expectations. From equation (2.2) and the definition of subjective beliefs β_t it follows that the equilibrium housing price is given by

$$q_t = \frac{1}{1 - \beta(1-\delta)\beta_t}\xi_t^d, \tag{2.6}$$

where β_t evolves according to (2.5). Equations (2.5) and (2.6) thus jointly characterize the equilibrium dynamics of housing prices and subjective beliefs. From equations (2.1) and (2.6) follows that the equilibrium price-to-rent ratio is given by

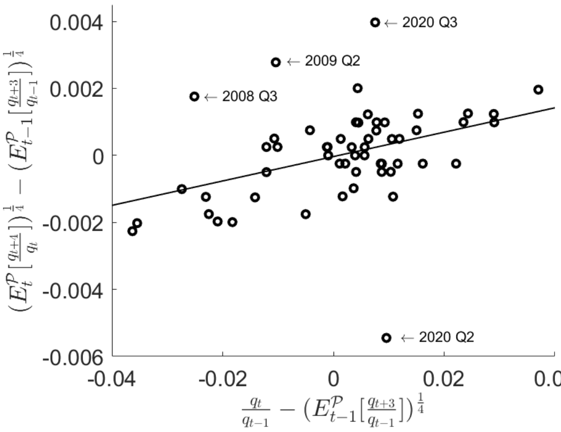
$$PR_t \equiv \frac{q_t}{R_t} = \frac{1}{1 - \beta(1-\delta)\beta_t}. \tag{2.7}$$

[16] In the full model in Section 2.4, we will assume the same beliefs for risk-adjusted house price growth. With risk neutrality, the two coincide.

[17] The (steady-state) Kalman gain depends on the subjectively perceived values for $(\sigma_\varepsilon^2, \sigma_\nu^2)$.

[18] The bound can be interpreted as a short-cut for a truncated prior support or b_t. The bounding function in (2.5) is a special case of the bounding function used in Adam and Woodford (2012), obtained by setting $\beta^L = \beta^U$.

Figure 2.1: Capital gain surprises and revisions



Notes: This figure plots the capital gain surprises against capital gain revisions in the Michigan survey (2007-2021), along with a linear regression line.

Calibration. The simple model just described can generate empirically plausible housing price behavior and the resulting housing price beliefs quantitatively match the deviations from rational expectations presented in the previous section. The calibration in this section is identical to the one for the full model, with the exception for the standard deviation of the innovations to housing preferences.[19] We consider housing preference shocks evolving according to

$$\log \xi_t^d = (1 - \rho_\xi) \log \underline{\xi}^d + \rho_\xi \log \xi_{t-1}^d + \varepsilon_t^d, \tag{2.8}$$

where $\varepsilon_t^d \sim iiN$ satisfies $E[e^{\varepsilon_t^d}] = 1$. Following Adam and Woodford (2021), we set $\rho_\xi = 0.99$ and $\delta = 0.03/4$. The standard deviation of ε_t^d is set to 0.0067, so that the model replicates the empirical standard deviation of the price-to-rent ratio, expressed in percent deviation from its mean, over the period for which we have survey data on housing expectations (2007-2021). The average value of the housing preference $\underline{\xi}^d > 0$ is irrelevant, as we are only concerned with moments characterizing cyclical properties (deviations from average values).

For the subjective belief process, we completely tie our hands and set $1/\alpha = 0.007$, which is the value estimated in Adam and Woodford (2012) using stock market expectations. The low value for the Kalman gain implies that agents extrapolate only weakly, as they believe most of the realized capital gains being due to transitory components. The value for the upper belief bound β^U is set as in the full model, where it matches the maximum observed deviation of the price-to-rent ratio from its mean. Finally, the quarterly discount factor β is set such that the real interest rate is equal to 0.75%, which is the average value of the estimated U.S. natural rate over the period 2007-2021, according to estimates using the approach of Holston, Laubach, and Williams (2017).

[19]This is so because the present section matches moments for a different time period than the full model, i.e., the period for which we have subjective expectations data (2007-2021).

Table 2.1: Housing price moments: data vs. model

	Data	Subjective Belief Model	RE Model
$std\,(PR_t)$	8.6	8.6	2.69
$corr(PR_t, PR_{t-1})$	0.99	0.99	0.99
$std(q_t/q_{t-1})$	0.06	0.04	0.003
$corr(q_t/q_{t-1}, q_{t-1}/q_{t-2})$	0.97	0.94	-0.01

Notes: The table reports the standard deviation and first-order autocorrelation of price-to-rent ratios and capital gains in the data, for the model under subjective housing beliefs and the model under rational expectations.

Housing price behavior. Table 2.1 illustrates that the subjective belief model replicates surprisingly well the empirical behavior of the price-to-rent ratio and of capital gains. While the standard deviation of the price-to-rent ratio is a targeted moment, all other moments are untargeted. The model matches very well the high quarterly autocorrelation of the price-to-rent ratio and the fairly high quarterly autocorrelation of capital gains. It undershoots somewhat the standard deviation of quarterly capital gains, illustrating that it features perhaps too little high-frequency variation in prices.[20]

Table 2.1 also reports the rational expectations (RE) outcome using the same calibration as for the subjective belief model. It shows that the about 70% of the fluctuations in the price-dividend ratio in the subjective belief model is due to capital gain extrapolation. Adam and Woodford (2012) explain how capital gain extrapolation generates momentum and mean reversion in prices and thus contributes to asset price volatility.

While the ability of capital gain extrapolation to increase the price volatility is well-known, we now turn to the new question of whether the model with capital gain extrapolation matches the structure of forecast errors documented in Section 2.2.

Belief revisions and forecast errors. The simple model quantitatively matches the three deviations from rational housing price expectations documented in Section 2.2.

Table 2.2 reports the outcomes of population regressions of equations (2.1), (2.2) and (2.3) for the calibrated subjective belief model. The results shows the model matches sluggish updating about expected housing prices ($b^{CG} > 0$) and the opposing cyclicality of actual and expected capital gains ($c > 0$ and $\mathbf{c} < 0$). For better comparison, Table 2.2 also reports also the empirical estimates of the corresponding coefficients from Tables 2.1 and 2.2. The magnitude of the coefficients generated by the model closely match the ones obtained using survey data, with the exception that the model underpredicts the counter-cyclicality of actual capital gains.

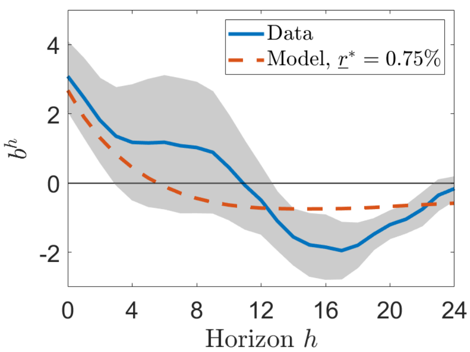
[20]This could easily be remedied by adding some iid shocks, say iid shocks to the discount factor β.

Table 2.2: Patterns of deviations from rational expectations: data vs. model

	Subjective Belief Model	Data	
		Mean Expect.	Median Expect.
b^{CG} from (2.1)	2.09	1.68	2.12
		(0.355)	(0.394)
c (in %) from (2.2)	0.03	0.030	0.010
		(0.172)	(0.043)
$\mathbf{c}$ (in %) from (2.3)	-0.063	-0.113	-0.113
		(0.009)	(0.009)

Notes: This table shows the model-implied regression coefficients of regressions (2.1), (2.2) and (2.3) for a natural rate of 0.75% (annualized) in the first column and the empirical results (for real housing prices) in the second and third column.

Figure 2.2: Dynamic forecast error response: data vs. model



Notes: The figure shows impulse-response functions of housing-price forecast errors of one-year ahead expectations to a one standard deviation innovation in the housing capital gain from the model and the data. The shaded area shows the 90%-confidence intervals of the empirical estimates, standard errors are robust with respect to serial correlation and heteroskedasticity (Newey–West with $h+1$ lags).

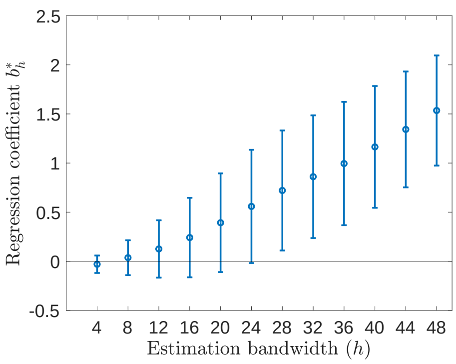
Figure 2.2 shows that the simple model is able to match the dynamic response of forecast errors documented empirically in Figure 2.1(b). We compute model-implied forecast errors as $FE^{model}_{t+h} = \frac{q_{t+4+h}}{q_{t+h}} - (\beta_{t+h})^4$ and compute the population local projections (2.4). Consistent with the data, the model generates initial underprediction of capital gains (over-pessimism) and subsequently overprediction (over-optimism).

Appendix 2.B.1 reports the dynamic forecast error responses for the model and in the data about the expected housing price level (rather than the expected capital gain). It shows that the model matches equally well the patterns of forecasts errors about the future housing price level.

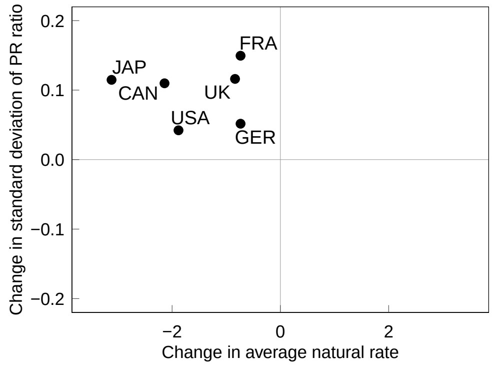
Falling natural rates and rising housing price volatility. The simple housing model also predicts that falling natural rates of interest will give rise to higher volatility for housing prices. Such a relationship between natural rates and hous-

Figure 2.3: The natural rate and housing price fluctuations

(a) Relationship between lower U.S. natural rates & housing price volatility



(b) Changes in the natural rate & housing price volatility in advanced economies



Notes: Panel (a) reports the regression coefficient b_h^* from equation (2.9) together with 68% Newey–West error bands using h lags. Panel (b) plots the pre-/post-1990 changes in the average natural rate against the changes in the volatility of the price-to-rent ratio for different advanced economies. The volatilities of the price-to-rent ratios in the pre-/post-1990 periods are the standard deviations relative to the period-specific mean values.

ing volatility is present in the data. It can be seen by considering regressions of the form

$$Std(PR_{t-\frac{h}{2}}, ..., PR_{t+\frac{h}{2}}) = a_h^* - b_h^* \cdot r_t^* + u_{t,h}, \tag{2.9}$$

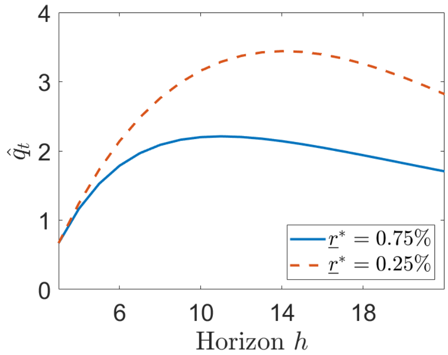
where r_t^* denotes the natural rate of interest from Holston et al. (2017) and $Std(PR_{t-\frac{h}{2}}, ..., PR_{t+\frac{h}{2}})$ the standard deviation of the price rent ratio using a window of $h+1$ quarters centered around period t. Under the standard assumption that the natural rate of interest is only a function of exogenous fundamentals, the regression coefficients b_h^* can be interpreted as capturing a causal relationship.

Panel (a) in Figure 2.3 reports the coefficients b_h^* for the United States using various estimation bandwidths h. While narrow bandwidths generate insignificant outcomes, most likely due to the difficulty associated with reliably estimating the standard deviation of the PR-ratio, the coefficient becomes positive and significant for larger bandwidth and is quite large when using a bandwidth of 48 quarters (+/- 2 years). We can thus conclude that the standard deviation of U.S. housing prices is rising as the natural rate falls.

Panel (b) in Figure 2.3 shows that this relationship is also present in other countries: it plots the change in the average level of the natural rate from the period before 1990 to the period after 1990 for the U.S., Canada, France, Germany, and the United Kingdom, against the change in the standard deviation of the price-to-rent ratio over the same periods. To take possible shifts in the mean of the PR-ratio over time into account, e.g., due to falling real interest rates, the standard deviation of the PR-ratio is computed in each of the two sub-periods for the *percent* deviation of the PR-ratio from its period-specific mean.[21] In all six

[21] The empirical results become even stronger if one considers instead the absolute standard

Figure 2.4: Model-implied housing price response for different natural rates



advanced economies, the PR-ratio has become more volatile as the average level of the natural rate has declined.

Equations (2.5) and (2.6) reveal how housing prices in our simple model are affected by the level of the natural rate of interest. The natural rate of interest is given by $\underline{r}^* = 1/\beta - 1$ and only depends on the discount factor $\beta \in (0, 1)$. A discount factor closer to one thus lowers the natural rate of interest.

Figure 2.4 illustrates that the model in fact generates a negative relationship between the level of the natural rate and housing price volatility.[22] It presents the impulse response of real housing prices to a positive housing preference shock ξ_t^d, which is the only shock driving housing prices in the model. It considers this response for the calibrated level of the natural rate of 0.75% and for a lower natural rate level equal to 0.25%.[23] The key message of Figure 2.4 is that housing prices respond stronger to housing demand shocks when natural rates are lower: the same shock gives rise to an approximately 75% stronger housing price response when the natural rate is at its lower level.

This surprising model outcome can be explained as follows. The capital gain increase triggered by the fundamental shock in the initial period leads to an upward revision of capital gain expectations. Equation (2.6) implies, however, that these higher capital gain expectations produce larger realized capital gains, the higher is the value for β, i.e., the lower is the natural rate of interest. Higher realized capital gains produce stronger upward revisions in beliefs in the future and thus feed stronger capital gains in the subsequent period. Through this feedback loop, low natural rates generate more momentum in housing price changes following fundamental shocks, allowing the model to replicate the relationship between natural rates and the volatility of housing prices.

---

deviation of the PR-ratio.

[22]The full model presented later on will also be able to quantiatively replicate this relationship, see section 2.6.

[23]To account for the higher housing price levels associated with lower natural rates, we show impulse responses in terms of percent deviations from their respective steady state values. The model-implied response for the PR-ratio to a housing preference shock looks very similar and is shown in Appendix 2.B.2.

2.4 Full model with capital gain extrapolation

This section studies the monetary policy implications of falling natural rates of interest and rising housing price volatility. To this end, we embed capital gain extrapolation into a sticky price model with a housing sector. The model features endogenous production of consumption goods and housing and generalizes the setup in Adam and Woodford (2021) by allowing for belief distortions that are not absolutely continuous with respect to the beliefs held by the policymaker. This permits analyzing the subjective housing beliefs as in equation (2.3), which give rise to capital gain extrapolation and deviations from rational expectations matching patterns in the survey data. In addition, we consider a lower-bound constraint on nominal rates, which we show to be quantitatively important for understanding how the optimal inflation target responds to lower natural rates in the presence of subjective housing beliefs.

We consider an economy populated by internally rational decision makers (Adam and Marcet (2011)): households maximize utility and firms maximize profits, but both do so using a potentially subjective probability measure $\mathcal{P}$, which assigns probabilities to all external variables, i.e., to all variables that are beyond agents' control. These variables include fundamental shocks, as well as competitive market prices (wages, goods prices, housing prices and rents). The setup delivers rational expectations in the special case where $\mathcal{P}$ is the objective probability measure.

The economy is made up of identical infinitely-lived households, each of which maximizes the following objective function[24]

$$U \equiv E_0^{\mathcal{P}} \sum_{t=0}^{\infty} \beta^t \left[\tilde{u}(C_t;\xi_t) - \int_0^1 \tilde{v}(H_t(j);\xi_t)dj + \tilde{\omega}(D_t + D_t^R;\xi_t) \right], \tag{2.1}$$

subject to the sequence of flow budget constraints

$$\begin{aligned} & C_t + B_t + (D_t - (1-\delta)D_{t-1}) \frac{q_t^u}{\tilde{u}_C(C_t;\xi_t)} + k_t + R_t D_t^R = \\ & \tilde{d}(k_t;\xi_t)\frac{q_t^u}{\tilde{u}_C(C_t;\xi_t)} + \int_0^1 w_t(j)H_t(j)dj + \frac{B_{t-1}}{\Pi_t}(1+i_{t-1}) + \frac{\Sigma_t}{P_t} + \frac{T_t}{P_t}, \end{aligned} \tag{2.2}$$

where C_t is an aggregate consumption good, $H_t(j)$ is the quantity supplied of labor of type j and $w_t(j)$ the associated real wage, D_t the stock of owned houses, D_t^R the units of rented houses, $\delta \in [0,1]$ the housing depreciation rate, and q_t^u the real price of houses in marginal utility units, defined as

$$q_t^u \equiv q_t \tilde{u}_C(C_t;\xi_t),$$

where q_t is the real house price in units of consumption.[25] The variable q_t^u provides a measure of whether housing is currently expensive or inexpensive, in units that

[24]It cannot be common knowledge to households that they are representative whenever $\mathcal{P}$ deviates from the rational measure.

[25]In Section 2.3, q_t^u and q_t coincide due to risk-neutrality.

are particularly relevant for determining housing demand. The variable k_t denotes investment in new houses and $\tilde{d}(k_t;\xi_t)$ the resulting production of new houses.[26] $B_t \equiv \widetilde{B}_t/P_t$ denotes the real value of nominal government bond holdings $\widetilde{B}_t$ and P_t the nominal price of consumption. $\Pi_t = P_t/P_{t-1}$ is the inflation rate, i_t the nominal interest rate, R_t the real rental rate for housing units, and ξ_t is a vector of exogenous disturbances, which may induce random shifts in the functions $\tilde{u}$, $\tilde{v}$, $\tilde{\omega}$ and $\tilde{d}$. T_t denotes nominal lump sum transfers (taxes if negative) from the government and Σ_t nominal profits accruing to households from the ownership of firms.

Households discount future payoffs at the rate $\beta \in (0,1)$. Since our model is formulated in terms of growth-detrended variables, the discount rate β jointly captures the time preference rate $\widetilde{\beta} \in (0,1)$ and the steady-state growth rate of marginal utility. Letting $g_c \geq 0$ denote the steady-state growth rate of consumption in non-detrended terms, we have

$$\beta \equiv \widetilde{\beta}\frac{\tilde{u}_C(C(1+g_c);\underline{\xi})}{\tilde{u}_C(C;\underline{\xi})}, \tag{2.3}$$

where $\underline{\xi}$ denotes the steady state value of the disturbance ξ_t. When the growth rate g_c of the economy falls, the discount rate β increases because marginal utility falls less strongly. We can thus capture a fall in the trend growth rate of the economy simply via an increase in the time discount rate β. Declining trend growth causes the steady-state real interest rate and thus the average natural rate of interest to fall, which is in line with the estimates provided in Holston et al. (2017) (see Appendix 2.G).

The aggregate consumption good is a Dixit-Stiglitz aggregate of each of a continuum of differentiated goods,

$$C_t \equiv \left[\int_0^1 c_t(i)^{\frac{\eta-1}{\eta}}\,di\right]^{\frac{\eta}{\eta-1}}, \tag{2.4}$$

with an elasticity of substitution $\eta > 1$. We further assume isoelastic functional forms

$$\begin{aligned}
\tilde{u}(C_t;\xi_t) &\equiv \frac{C_t^{1-\tilde{\sigma}^{-1}}\bar{C}_t^{\tilde{\sigma}^{-1}}}{1-\tilde{\sigma}^{-1}},\\
\tilde{v}(H_t(j);\xi_t) &\equiv \frac{\lambda}{1+\nu}\left(H_t(j)\right)^{1+\nu}\bar{H}_t^{-\nu},\\
\tilde{\omega}(D_t+D_t^R;\xi_t) &\equiv \xi_t^d\left(D_t+D_t^R\right),\\
\tilde{d}(k_t;\xi_t) &\equiv \frac{A_t^d}{\tilde{\alpha}}k_t^{\tilde{\alpha}},
\end{aligned} \tag{2.5}$$

where $\tilde{\sigma},\nu > 0$, $\tilde{\alpha} \in (0,1)$ and $\{\bar{C}_t,\bar{H}_t,\xi_t^d,A_t^d\}$ are bounded, exogenous and positive disturbance processes which are among the exogenous disturbances included in the vector ξ_t.

[26] We consolidate housing production into the household budget constraint. It would be equivalent to have instead a separate housing production sector that is owned by households.

Our specification includes two housing-related disturbances, namely ξ_t^d, which captures shocks to housing preferences, and A_t^d, which captures shocks to the productivity in the construction of new houses. We impose linearity in the utility function (2.5), because it greatly facilitates the characterization of optimal policy, with rented and owned housing units being perfect substitutes. Introducing a weight on rental units relative to housing units would allow us to perfectly match the average price-to-rent ratio we observe in the data. However, since this does not change any other results, we abstract from such a scaling parameter and assign equal weight to housing and renting in the utility.

Each differentiated good is supplied by a single monopolistically competitive producer; there is a common technology for the production of all goods, in which (industry-specific) labor is the only variable input,

$$y_t(i) = A_t f(h_t(i)) = A_t h_t(i)^{1/\phi}, \tag{2.6}$$

where A_t is an exogenously varying technology factor, and $\phi > 1$. The Dixit-Stiglitz preferences (2.4) imply that the quantity demanded of each individual good i will equal[27]

$$y_t(i) = Y_t \left(\frac{p_t(i)}{P_t} \right)^{-\eta}, \tag{2.7}$$

where Y_t is the total demand for the composite good defined in (2.4), $p_t(i)$ is the price of the individual good, and P_t is the price index,

$$P_t \equiv \left[\int_0^1 p_t(i)^{1-\eta} di \right]^{\frac{1}{1-\eta}}, \tag{2.8}$$

corresponding to the minimum cost for which a unit of the composite good can be purchased in period t. Total demand is given by

$$Y_t = C_t + k_t + g_t Y_t, \tag{2.9}$$

where g_t is the share of the total amount of composite goods purchased by the government, treated here as an exogenous disturbance process.

2.4.1 Household optimality conditions

Internally rational households choose state-contingent sequences for the choice variables $\{C_t, H_t(j), D_t, D_t^R, k_t, B_t\}$ so as to maximize (2.1), subject to the budget constraints (2.2), taking as given their beliefs about the processes

$$\{P_t, w_t(j), q_t^u, R_t, i_t, \Sigma_t/P_t, T_t/P_t\},$$

as determined by the (subjective) measure $\mathcal{P}$.

[27] In addition to assuming that household utility depends only on the quantity obtained of C_t, we assume that the government also cares only about the quantity obtained of the composite good defined by (2.4), and that it seeks to obtain this good through a minimum-cost combination of purchases of individual goods.

We shall be particularly interested in the policy implications generated by subjective housing price beliefs. To insure that an optimum exists in the presence of potentially distorted beliefs about the housing price q_t^u, we require housing choices to lie in some compact choice set $D_t \in [0, D^{\max}]$, as discussed in Section 2.3, where the upper bound can be arbitrarily large.

The first order conditions give rise to an optimal labor supply relation

$$w_t(j) = \frac{\tilde{v}_H(H_t(j); \xi_t)}{\tilde{u}_C(C_t; \xi_t)}, \tag{2.10}$$

a consumption Euler equation

$$\tilde{u}_C(C_t; \xi_t) = \beta E_t^{\mathcal{P}} \left[\tilde{u}_C(C_{t+1}; \xi_{t+1}) \frac{1 + i_t}{P_{t+1}/P_t} \right], \tag{2.11}$$

an equation characterizing optimal investment in new houses

$$k_t = \left(A_t^d q_t^u \frac{C_t^{\tilde{\sigma}^{-1}}}{\bar{C}_t^{\tilde{\sigma}^{-1}}} \right)^{\frac{1}{1-\tilde{\alpha}}}, \tag{2.12}$$

an optimality condition for rental units

$$\xi_t^d = R_t \tilde{u}_C(C_t, \xi_t), \tag{2.13}$$

and a set of conditions determining the optimal housing demand D_t:

$$\begin{array}{ll} q_t^u < \xi_t^d + \beta(1-\delta) E_t^{\mathcal{P}} q_{t+1}^u & \text{if } D_t = D^{\max} \\ q_t^u = \xi_t^d + \beta(1-\delta) E_t^{\mathcal{P}} q_{t+1}^u & \text{if } D_t \in (0, D^{\max}) \\ q_t^u > \xi_t^d + \beta(1-\delta) E_t^{\mathcal{P}} q_{t+1}^u & \text{if } D_t = 0. \end{array} \tag{2.14}$$

With rational expectations, the upper and lower holding bounds never bind.[28] Since we are interested in how the presence of belief distortions about future housing values affect equilibrium outcomes, the bounds in equation (2.14) can potentially bind under the *subjectively* optimal plans. This explains why an internally rational household can hold subjective housing price expectations, even if she holds rational expectations about the preference shocks ξ_t^d in equation (2.14).

Forward-iterating on equation (2.11), which holds with equality under all belief-specifications, delivers a present-value formulation of the consumption Euler equation

$$\tilde{u}_C(C_t; \xi_t) = \lim_{T \to \infty} E_t^{\mathcal{P}} \left[\tilde{u}_C(C_T; \xi_T) \beta^T \prod_{k=0}^{T-t} \frac{1 + i_{t+k}}{P_{t+k+1}/P_{t+k}} \right], \tag{2.15}$$

which will be convenient to work with, especially under subjective belief specifications. Household choices must also satisfy the transversality constraint

$$\lim_{T \to \infty} \beta^T E_t^{\mathcal{P}} \left[\tilde{u}_C(C_T; \xi_T) B_T + D_T q_T^u \right] = 0. \tag{2.16}$$

Optimal household behavior under potentially distorted beliefs is jointly characterized by equations (2.10) and (2.12)-(2.16).

[28]The upper bound $D^{\max}$ has been chosen sufficiently large for this to be true. The lower bound is never reached because the housing production function satisfies Inada conditions.

2.4.2 Optimal price setting by firms

The producers in each industry fix the prices of their goods in monetary units for a random interval of time, as in the model of staggered pricing introduced by Calvo (1983) and Yun (1996). Producers use the representative households' subjectively optimal consumption plans to discount profits and are assumed to know the product demand function (2.7). They need to formulate beliefs about the future price levels P_T, industry-specific wages $w_T(j)$, aggregate demand Y_T, and productivity A_T.

Let $0 \leq \alpha < 1$ be the fraction of prices that remain unchanged in any period. A supplier i in industry j that changes its price in period t chooses its new price $p_t(i)$ to maximize

$$E_t^{\mathcal{P}} \sum_{T=t}^{\infty} \alpha^{T-t} Q_{t,T} \Pi \left(p_t(i), P_T, w_T(j), Y_T, A_T\right), \tag{2.17}$$

where $E_t^{\mathcal{P}}$ denotes the expectations of price setters conditional on time t information, which are identical to the expectations held by consumers. Firms discount random nominal income in period T using households' subjective stochastic discount factor $Q_{t,T}$, which is given by

$$Q_{t,T} = \beta^{T-t} \frac{\tilde{u}_C\left(C_T, \xi_T\right)}{\tilde{u}_C\left(C_t, \xi_t\right)} \frac{P_t}{P_T}.$$

The term α^{T-t} in equation (2.17) captures the probability that a price chosen in period t will not have been revised by period T, and the function $\Pi\left(p_t(i), ...\right)$ indicates the nominal profits of the firm in period t, as discussed next.

Profits are equal to after-tax sales revenues net of the wage bill. Sales revenues are determined by the demand function (2.7), so that (nominal) after-tax revenue equals

$$\left(1 - \tau_t\right) p_t(i) Y_t \left(\frac{p_t(i)}{P_t}\right)^{-\eta}.$$

Here τ_t is a proportional tax on sales revenues in period t, $\{\tau_t\}$ is treated as an exogenous disturbance process, taken as given by the monetary policymaker. We assume that τ_t fluctuates over a small interval around a non-zero steady state level $\underline{\tau}$. We allow for exogenous variations in the tax rate in order to include the possibility of "pure cost-push shocks" that affect the equilibrium pricing behavior while implying no change in the efficient allocation of resources.

The labor demand of firm i at a given industry-specific wage $w_t(j)$ can be written as

$$h_t(i) = \left(\frac{Y_t}{A_t}\right)^{\phi} p_t(i)^{-\eta\phi} P_t^{\eta\phi}, \tag{2.18}$$

which follows from (2.6) and (2.7). Using this, the nominal wage bill is given by

$$P_t w_t(j) h_t(i) = P_t w_t(j) \left(\frac{Y_t}{A_t}\right)^{\phi} p_t(i)^{-\eta\phi} P_t^{\eta\phi}.$$

Subtracting the nominal wage bill from the above expression for nominal after tax revenue, we obtain the function $\Pi\left(p_t(i), P_T, w_T(j), Y_T, A_T\right)$ used in (2.17).

Each of the suppliers that revise their prices in period t chooses the same new price p_t^*, that maximizes (2.17). The first-order condition with respect to $p_t(i)$ is given by[29]

$$E_t^{\mathcal{P}} \sum_{T=t}^{\infty} \alpha^{T-t} Q_{t,T} \Pi_1\left(p_t(i), P_T, w_T(j), Y_T, A_T\right) = 0.$$

The equilibrium choice p_t^*, which is the same for each firm i in industry j, is the solution to this equation. Letting p_t^j denote the price charged by firms in industry j at time t, we have $p_t^j = p_t^*$ in periods in which industry j resets its prices and $p_t^j = p_{t-1}^j$ otherwise.

Under the assumed isoelastic functional forms, the optimal choice has a closed-form solution

$$\left(\frac{p_t^*}{P_t}\right)^{1+\eta(\phi-1)} = \frac{E_t^{\mathcal{P}} \sum_{T=t}^{\infty} \alpha^{T-t} Q_{t,T} \frac{\eta}{\eta-1} \phi w_T(j) \left(\frac{Y_T}{A_T}\right)^{\phi} \left(\frac{P_T}{P_t}\right)^{\eta\phi+1}}{E_t^{\mathcal{P}} \sum_{T=t}^{\infty} \alpha^{T-t} Q_{t,T} (1-\tau_T) Y_T \left(\frac{P_T}{P_t}\right)^{\eta}}. \tag{2.19}$$

The price index evolves according to a law of motion

$$P_t = \left[(1-\alpha)\, p_t^{*1-\eta} + \alpha P_{t-1}^{1-\eta}\right]^{\frac{1}{1-\eta}}, \tag{2.20}$$

as a consequence of (2.8). The equilibrium inflation in any period is characterized by

$$\left(\frac{P_t}{P_{t-1}}\right)^{\eta-1} = \frac{1-(1-\alpha)\left(\frac{p_t^*}{P_t}\right)^{1-\eta}}{\alpha}. \tag{2.21}$$

The welfare loss from price adjustment frictions can be captured by price dispersion, which is defined as

$$\Delta_t \equiv \int_0^1 \left(\frac{p_t^j}{P_t}\right)^{-\eta(1+\omega)} dj \geq 1, \tag{2.22}$$

where

$$\omega \equiv \phi(1+\nu) - 1 > 0$$

is the elasticity of real marginal cost in an industry with respect to industry output.

Using equation (2.20) together with the fact that the relative prices of the industries that do not change their prices in period t remain the same, one can derive a law of motion for the price dispersion term Δ_t of the form

$$\Delta_t = h(\Delta_{t-1}, P_t/P_{t-1}), \tag{2.23}$$

[29] Note that supplier i's profits in (2.17) are a concave function of the quantity sold $y_t(i)$, since revenues are proportional to $y_t(i)^{\frac{\eta-1}{\eta}}$ and hence concave in $y_t(i)$, while costs are convex in $y_t(i)$. Moreover, since $y_t(i)$ is proportional to $p_t(i)^{-\eta}$, the profit function is also concave in $p_t(i)^{-\eta}$. The first-order condition for the optimal choice of the price $p_t(i)$ is the same as the one with respect to $p_t(i)^{-\eta}$; hence the first-order condition with respect to $p_t(i)$ is both necessary and sufficient for an optimum.

with

$$h(\Delta_t, P_t/P_{t-1}) \;\equiv\; \alpha\Delta_t \left(\frac{P_t}{P_{t-1}}\right)^{\eta(1+\omega)} + (1-\alpha)\left(\frac{1-\alpha\left(\frac{P_t}{P_{t-1}}\right)^{\eta-1}}{1-\alpha}\right)^{\frac{\eta(1+\omega)}{\eta-1}}.$$

As is commonly done, we assume that the initial degree of price dispersion is small ($\Delta_{-1} \sim O(2)$).

Equations (2.19), (2.21), and (2.23) jointly define a short-run aggregate supply relation between inflation, output and house prices (via the aggregate demand equation (2.9) and (2.12)), given the current disturbances ξ_t, and expectations regarding future wages, prices, output, consumption and disturbances. Equation (2.23) describes the evolution of the costs of price dispersion over time.

For future reference, we remark that all firms together make total profits equal to

$$\frac{\Sigma_t}{P_t} = (1-\tau_t)Y_t - w_t H_t, \tag{2.24}$$

where $w_t H_t = \int_0^1 w_t(j) H_t(j) dj$.

2.4.3 Government budget constraint and market clearing conditions

The government consumes goods $g_t Y_t$, imposes a sales tax τ_t, issues nominal bonds $\widetilde{B}_t \equiv P_t B_t$, and pays lump sum transfers T_t to households. The government budget constraint is given by

$$B_t = B_{t-1}\frac{1+i_{t-1}}{P_t/P_{t-1}} + \frac{T_t}{P_t} + (g_t - \tau_t)\, Y_t.$$

For simplicity, we assume that lump sum transfers (taxes if negative) are set such that they keep real government debt constant at some initial level B_{-1}. This implies that government transfers are given by

$$\frac{T_t}{P_t} = -(g_t - \tau_t)Y_t + B_{t-1}\left(1 - \frac{1+i_{t-1}}{P_t/P_{t-1}}\right). \tag{2.25}$$

Using (2.9) and (2.12), one can express the market clearing condition for the consumption/investment good as

$$Y_t = \frac{C_t + \Omega_t C_t^{\frac{\tilde{\sigma}^{-1}}{1-\tilde{\alpha}}}}{1-g_t}, \tag{2.26}$$

where

$$\Omega_t \equiv \left(A_t^d \bar{C}_t^{-\tilde{\sigma}^{-1}} q_t^u\right)^{\frac{1}{1-\tilde{\alpha}}} > 0 \tag{2.27}$$

is a term that depends on exogenous shocks and belief distortions in the housing market only, see equation (2.14). The previous two equations implicitly define a function

$$C_t = C(Y_t, q_t^u, \xi_t), \tag{2.28}$$

which delivers the market clearing consumption level, for a given output level Y_t, given housing prices q_t^u and given exogenous disturbances ξ_t.

The market clearing condition for housing is

$$D_t = (1-\delta)D_{t-1} + \tilde{d}(k_t; \xi_t), \tag{2.29}$$

and rental market clearing requires

$$D_t^R = 0. \tag{2.30}$$

Labor market clearing requires that the supply of labor of type j in (2.10) is equal to labor demand of industry j, which is given by (2.18), as all firms in the industry charge the same price. This delivers

$$w_t(j) = \frac{\tilde{v}_H(H_t(j); \xi_t)}{\tilde{u}_C(C_t; \xi_t)} = \frac{\lambda\,(H_t(j))^{\nu}\,\bar{H}_t^{-\nu}}{C_t^{-\tilde{\sigma}^{-1}}\bar{C}_t^{\tilde{\sigma}^{-1}}} = \lambda \frac{\bar{H}_t^{-\nu}}{\bar{C}_t^{\tilde{\sigma}^{-1}}} \left(\frac{Y_t}{A_t}\right)^{\nu\phi} C_t^{\tilde{\sigma}^{-1}} \left(\frac{p_t^j}{P_t}\right)^{-\nu\eta\phi}, \tag{2.31}$$

where $p_t^j = p_t^*$ in periods where industry j can adjust prices and $p_t^j = p_{t-1}^j$ otherwise.

2.4.4 Equilibrium and Ramsey problem with subjective beliefs

We now define the equilibrium in the presence of subjective beliefs, as well as the nonlinear Ramsey problem characterizing the monetary policymaker's optimization problem in the presence of subjective beliefs.

We start by defining an *Internally Rational Expectations Equilibrium (IREE)*, which is a generalization of the notion of a Rational Expectations Equilibrium (REE) to settings with subjective private sector beliefs:

Definition 2.1 *An internally rational expectations equilibrium (IREE) is a bounded stochastic process for* $\{Y_t, C_t, k_t, D_t, \{w_t(j)\}, p_t^*, P_t, \Delta_t, q_t^u, i_t\}_{t=0}^{\infty}$ *satisfying the aggregate supply equations (2.19), and (2.21), the law of motion for the evolution of price distortions (2.23), the household optimality conditions (2.12), (2.14), (2.15), and the market clearing conditions (2.26), (2.29) and (2.31) for all* j.

The equilibrium features ten variables (counting the continuum of wages as a single variable) that must satisfy nine conditions, leaving one degree of freedom to be determined by monetary policy.[30] In the special case with rational beliefs ($E_t^{\mathcal{P}}[\cdot] = E_t[\cdot]$), the IREE is a Rational Expectations Equilibrium (REE).

[30]The transversality condition (2.16) must also be satisfied in equilibrium, but is not imposed as an equilibrium condition, as it will hold for all belief specifications considered below.

Given the equilibrium outcome, the remaining model variables can be determined as follows. Equilibrium profits are given by equation (2.24) and equilibrium taxes by equation (2.25). Equilibrium labor supply $H_t(j)$ follows from equation (2.10) for each labor type j. Equilibrium bond holdings satisfy $B_t = B_{-1}$ and equilibrium inflation is $\Pi_t \equiv P_t/P_{t-1}$. Equilibrium rental units are given by equation (2.30) and equilibrium rental prices by equation (2.13).

The Ramsey problem allows the policymaker to choose the sequence of policy rates, prices and allocations to maximize household utility, subject to the constraint that prices and allocations constitute an IREE. The policymaker thereby maximizes household utility under rational expectations, i.e., under a probability measure that is different from the one entertained by households, whenever the latter hold distorted beliefs. Benigno and Paciello (2014) refer to such a policymaker as a 'paternalistic' policymaker. The non-linear Ramsey problem is spelled out in Appendix 2.C. To gain economic insights into the forces shaping the policy problem, the next section considers a quadratic approximation to the nonlinear problem.

2.5 The monetary policy problem: analytic insights

This section derives analytic insights into the monetary policy problem. In particular, it presents a quadratic approximation to the policymaker's Ramsey problem that highlights the new economic forces arising from the presence of capital gain extrapolation.[31] It shows how subjective capital gain expectations shift the Phillips curve and affect the natural rate of interest in the IS equation.

The quadratic approximation derived below is valid for two alternative belief settings.[32] The first setting is standard and assumes rational expectations. While constituting a useful benchmark, the assumption of rational housing price expectations is strongly rejected by the survey evidence in Section 2.2.

The second setting considers subjective housing beliefs. In particular, it considers capital gain extrapolation according to equations (2.3)-(2.5) introduced in the simple model in Section 2.3, but with the variable q_t being replaced by q_t^u. The latter implies that households extrapolate capital gains in units of marginal utility rather than in units of consumption. Specifying subjective beliefs in units of marginal utility leaves the ability of the learning rule to replicate the survey

[31] The nonlinear problem can be found in Appendix 2.C. The quadratic approximation delivers a valid second-order approximation to the problem, whenever (i) the steady-state Lagrange multipliers associated with the nonlinear constraints are of order $O(1)$, which is the case when the steady state output distortion $\Theta \equiv \log\left(\frac{\eta}{\eta-1}\frac{1-g}{1-\tau}\right)$ is of order $O(1)$, and (ii) the gap between the steady-state interest rate and the lower bound, i.e., $\frac{1}{\beta}-1$, is also of $O(1)$, i.e., when steady state real interest rates/natural rates are low.

[32] Recall from our earlier discussion that firms must hold beliefs about future values of P_t, $w_t(j)$, Y_t and that households must hold beliefs about future values of $(P_t, w_t(j), q_t^u, R_t, i_t, \Sigma_t/P_t, T_t/P_t)$. Both actors must additionally hold beliefs about the fundamental shocks entering their decision problem.

evidence unchanged[33], but has three advantages.

First, the dynamics of housing prices in units of marginal utility is unaffected by monetary policy, even if housing prices in units of consumption do depend on policy. As a result, the object about which agents learn does not depend on policy. The policymaker thus cannot 'manipulate' households' subjective housing price beliefs in a way to achieve outcomes that are potentially better than under rational expectations.[34] In addition, it allows side-stepping the otherwise thorny issue of how the learning rule should respond to the conduct of monetary policy.

Second, the belief setup allows replicating the fact that housing demand and investment responds more strongly to monetary policy disturbances than non-housing demand, thereby avoiding the pitfalls described in Barsky et al. (2007). Appendix 2.D shows that in response to an exogenous shift in the path of nominal interest rates i, the change in housing investment and consumption satisfies at all times

$$\frac{d \log k_t}{d\boldsymbol{i}} = \frac{1}{1-\tilde{\alpha}} \frac{1}{\tilde{\sigma}} \cdot \frac{d \log C_t}{d\boldsymbol{i}}, \tag{2.1}$$

where $1/(1-\tilde{\alpha})$ is the price elasticity of housing supply and $1/\tilde{\sigma}$ the coefficient of relative risk aversion in consumption. The calibrated model considered later on features $1/((1-\tilde{\alpha})\tilde{\sigma}) > 1$.[35]

Third, the belief specification greatly simplifies the algebra involved in deriving the second-order approximation to the Ramsey problem, because it allows for a relatively straightforward determination of the equilibrium path of subjectively optimal consumption choices.

Overall, we wish to consider a minimal deviation from rational expectations, therefore keep expectations about all other variables rational to the extent possible.[36] Finally, to insure that households' subjectively optimal plans satisfy the transversality condition, we assume that household hold rational capital gain expectations in the very long run, i.e., after some arbitrarily large but finite period $\bar{T} < \infty$.[37] We then consider the policy problem with subjective beliefs in periods $t \ll \overline{T}$.

---

[33]This is so because we consider log consumption preferences which imply that contributions from fluctuations in marginal utility are orders of magnitude smaller than those generated by subjective beliefs.

[34]This is a key distinction to the setups analyzed in Molnar and Santoro (2014), Mele, Molnar, and Santoro (2020), and Caines and Winkler (2021).

[35]The calibration use log utility in consumption ($1/\tilde{\sigma} = 1$) and a supply elasticity of $1/(1-\tilde{\alpha}) = 5$.

[36]In particular, household continue to hold rational expectations about all other prices, i.e., about $\{P_t, w_t(j), i_t\}$ and firms hold rational expectations about $\{P_t, w_t(j), Y_t\}$. Furthermore, all actors continue to hold rational expectations about the exogenous fundamentals. Beliefs about profits and lump sum taxes, $\{\Sigma_t/P_t, T_t/P_t\}$ continue to be determined by equations (2.24) and (2.25), evaluated with rational output expectations and the state-contingent optimal choices for $\{H_t, k_t, B_t\}$. Rental price expectations, however, cannot be kept rational: they need to satisfy equation (2.13), which shows that they are influenced by the subjectively optimal consumption plans implied by equation (2.15).

[37]Appendix 2.E shows that this is sufficient to insure that subjectively optimal plans satisfy the transversality constraint (2.16).

For the two belief settings just described, the quadratic approximation of the Ramsey problem is given by[38]

$$\max_{\{\pi_t, y_t^{gap}, \widehat{q}_t^u, i_t \geq \underline{i}\}} -E_0 \sum_{t=0}^{\infty} \beta^t \frac{1}{2} \left(\Lambda_\pi \pi_t^2 + \Lambda_y \left(y_t^{gap}\right)^2 + \Lambda_q \left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right)^2 \right) \tag{2.2}$$

s.t.:

$$\pi_t = \beta E_t \pi_{t+1} + \kappa_y y_t^{gap} + \kappa_q \left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right) + u_t \tag{2.3}$$

$$y_t^{gap} = \lim_{T\to\infty} E_t y_T^{gap} - \varphi E_t \sum_{k=0}^{\infty} \left(i_{t+k} - \pi_{t+1+k} - r_{t+k}^{*,RE} \right) - \frac{C_q}{C_Y} \left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right) \tag{2.4}$$

for $t \geq 0$ as well as equations determining $\left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right)$
and initial pre-commitments,

where $\pi_t = \log \Pi_t$ denotes inflation and y_t^{gap} the output gap, which is defined as $y_t^{gap} = \log Y_t - \log Y_t^*$, with Y_t^* denoting the efficient level of output, as defined in equation (2.F.1) in Appendix 2.F. The housing price gap $\widehat{q}_t^u - \widehat{q}_t^{u*}$ is the difference between the housing price $\widehat{q}_t^u = \log q_t^u$ and its efficient welfare-maximizing level $\widehat{q}_t^{u*}$, which is given by[39]

$$q_t^{u*} = \overline{\xi}_t^d, \tag{2.5}$$

where $\overline{\xi}_t^d \equiv \sum_{T=t}^{\infty} E_t[(1-\delta)^{T-t} \beta^{T-t} \xi_T^d]$.

The policymaker's objective (2.2) involves the standard terms of squared inflation and the squared output gap, but also depends on the squared housing price gap. The latter arises because any deviation of housing prices from their efficient level distorts – for a given level of the output gap – housing investment, as we explain below. The equilibrium value of the housing price gap will depend on the belief specification and will be discussed in detail in the next two sections.

Constraint (2.3) is the New Keynesian Phillips Curve and depends on the housing price gap. The coefficients $\kappa_q < 0$ and $\kappa_y > 0$ are defined in Appendix 2.F.3 and imply that positive housing price gaps exert *negative* cost-push effects: high housing prices increase housing investment and – for a given output gap – decrease non-housing consumption. The latter raises the marginal utility of non-housing consumption and thereby depresses wages and marginal costs. This allows the model to potentially produce a non-inflationary boom in housing prices and housing investment. The mark-up disturbance u_t is a function of exogenous disturbances only.

Constraint (2.4) is the linearized and forward-iterated IS equation. A key new insight here is that the IS equation also depends on the housing price gap. This implies that the housing price gap affects the natural rate of interest, as discussed in detail below. The coefficients $C_q < 0$ and $C_Y > 0$ are the derivatives of the function $C(\cdot)$ defined in (2.28) with respect to q^u and Y, respectively, evaluated

[38]See Appendix 2.F for a derivation.

[39]See the derivation in Appendix 2.H.4. All variables in the approximation are expressed in terms of log deviations from the efficient steady state.

at the efficient steady state. The variable $r_t^{*,RE}$ in equation (2.4) denotes the natural interest rate under RE and is a function of exogenous disturbances only.[40] The long-run output gap expectations $\lim_T E_t y_T^{gap}$ in equation (2.4) are the ones associated with a setting in which agents hold rational housing expectations.[41]

Note, that the policymaker's choice of the nominal interest rate i_t is subject to an effective lower bound $i_t \geq \underline{i}$, where the bound $\underline{i} < 0$ is expressed in terms of deviations from the interest rate in a zero-inflation steady state. For the special case with a zero lower bound, we have $\underline{i} = -(1-\beta)/\beta$. In the absence of a lower bound constraint or when economic shocks never cause the bound to become binding, the IS equation (2.4) can be dropped from the policy problem.

Interestingly, the expectations showing up in the monetary policy problem (2.F.11) are all rational. The way subjective housing price expectations affect the monetary policy problem are thus fully captured through their effects on the housing price gap. The next two sections determine the housing price gaps under rational and subjective beliefs and what they imply for optimal policy.

2.5.1 Rational housing price expectations

With fully rational expectations we have

$$\widehat{q}_t^{u,RE} = \widehat{q}_t^{u*}, \tag{2.6}$$

which shows that the housing price *gap* is zero at all times, independently of monetary policy and independently of the economic disturbances hitting the economy.[42] Under RE, the Ramsey problem with a lower bound constraint (2.2) is thus isomorphic to the Ramsey problem in a standard New Keynesian model without a housing sector, as considered for instance in Adam and Billi (2006). This result may appear surprising because monetary policy decisions do affect the housing price in units of consumption $\widehat{q}_t$. Yet, as the policy problem (2.2) makes clear, it is only the housing price gap in units of marginal utility, $\widehat{q}_t^u - \widehat{q}_t^{u*}$, that is relevant from a welfare perspective. Under RE, the presence of a housing sector thus generates no fundamentally new economic insights into the monetary policy problem.[43]

The RE setup also has difficulties in making a connection between the average natural rate of interest and the volatility of the price-to-rent (PR) ratio. Under RE, the equilibrium PR-ratio is

$$PR_t^{RE} = \frac{q_t^u}{\xi_t^d}, \tag{2.7}$$

[40] More precisely, $r_t^{*,RE}$ is the real interest rate consistent with the optimal consumption level in a setting with flexible prices and fully rational expectations, see Appendix 2.F.4 for details.

[41] Recall that housing expectations are assumed rational in the long-run in both belief settings.

[42] See Appendix 2.H.1 for proofs on the results about housing prices and the price-rent ratio presented in this section.

[43] The inclusion of a housing sector only affects the definition of the output gap, which now also depends on housing sector disturbances.

which to a first-order approximation is given by

$$\widehat{PR}_t^{RE} = Z \cdot \widehat{\xi}_t^d, \tag{2.8}$$

with $Z \equiv \beta(1-\delta)(\rho_\xi - 1)/(1-\beta(1-\delta)\rho_\xi)$. Equation (2.8) shows that the PR-ratio displays persistent variation under RE, if and only if housing demand shocks $\widehat{\xi}_t^d$ are persistent. In fact, replicating the high quarterly auto-correlation of the PR-ratio in Table 2.1 requires choosing a shock persistence ρ_ξ very close to one. Yet, in the limit $\rho_\xi \to 1$, the derivative $\partial Z/\partial\beta$ uniformly converges to zero for all $\beta \in [0,1]$. This implies that the volatility of the PR ratio will be largely independent of the natural rate of interest when housing demand shocks are sufficiently persistent. Under RE, there is thus no quantitatively important relationship between the average natural rate of interest and the volatility of the PR ratio, unlike in the case with capital gain extrapolation.

Given equation (2.6), the IS equation (2.4) implies that setting

$$i_t - E_t\pi_{t+1} = r_t^{*,RE} \quad \text{for all } t \geq 0 \tag{2.9}$$

is consistent with a constant output gap, i.e.,

$$y_t^{gap} = \lim_T E_t y_T^{gap} \quad \text{for all } t \geq 0.$$

This justifies our interpretation of $r_t^{*,RE}$ as the natural rate of interest under RE.[44] It also shows that the volatility of the natural rate of interest is independent of the average value of the natural rate under RE. This will cease to be the case under subjective housing beliefs.

2.5.2 Subjective housing price expectations

This section discusses three new economic forces showing up in the monetary policy problem in the presence of subjective housing price beliefs. It shows (i) how housing price fluctuations are affected by the average level of the natural rate of interest, (ii) how these fluctuations affect the volatility of the natural rate of interest, and (iii) how these fluctuations distort the allocation of output.

Housing prices under subjective beliefs are jointly determined by equations (2.5) and (2.6), where q_t should again be replaced by q_t^u. Since these equations do not depend on policy, the policymaker can treat the housing price gap as exogenous, as is the case with RE.[45] Yet, the housing price gap will now generally differ from zero, as the housing price gap can become positive or negative depending on the degree of capital gain optimism/pessimism.

[44] In the presence of a lower bound constraint on nominal rates, it might not be feasible to implement (2.9) at all times.

[45] This does not imply that the housing price q_t is invariant to monetary policy: monetary can determine how variations in q_t^u get split up into variations of the housing price q_t and variations in marginal utility $\widetilde{u}_C(C_t;\xi_t)$.

The average natural rate and housing price volatility. With subjective housing price expectations, the equilibrium housing price is given by[46]

$$q_t^{u,\mathcal{P}} = \frac{1}{1-\beta(1-\delta)\beta_t}\xi_t^d \tag{2.10}$$

and the price-to-rent ratio by

$$PR_t^{\mathcal{P}} = \frac{q_t^{u,\mathcal{P}}}{\xi_t^d}. \tag{2.11}$$

For the limit with persistent housing demand shocks ($\rho_\xi \to 1$), we can derive the first-order approximation

$$\widehat{q}_t^{u,\mathcal{P}} = \widehat{q}_t^{u,RE} + (\beta_t - 1)\frac{\beta(1-\delta)}{1-\beta(1-\delta)\beta_t}\left(1+\widehat{\xi}_t^d\right), \tag{2.12}$$

which decomposes the equilibrium housing price into its RE value plus a contribution coming from the presence of subjective beliefs. We then also have

$$E_t^{\mathcal{P}}\left[\widehat{q}_{t+1}^{u,\mathcal{P}}\right] = E_t\left[\widehat{q}_{t+1}^{u,RE}\right] + (\beta_t - 1)\left[1+\frac{\beta(1-\delta)}{1-\beta(1-\delta)\beta_t}\left(1+\widehat{\xi}_t^d\right)\right], \tag{2.13}$$

which shows that subjective housing price expectations are equal to their RE equilibrium value whenever expected capital gains are equal to one ($\beta_t = 1$). Capital gain extrapolation, however, will induce fluctuations of β_t around one and thus drive a wedge between the housing price under learning and RE.[47]

As explained for the simple model in Section 2.3, lower values for the average natural rate (discount factors β closer to one), will induce stronger fluctuations in capital gain expectations (β_t), because housing prices are more sensitive to belief revisions, see equation (2.10). Lower average values for the natural rates will thus be associated with increased fluctuations in housing prices and the PR-ratio, in line with empirical evidence presented in Section 2.3.

Housing price fluctuations and the natural rate of interest. The presence of non-zero housing price gaps also affects the natural rate of interest. This can be seen by considering a policy that sets real interest rates equal to the RE natural real rate ($r_t^{*,RE}$). Such a policy now ceases to deliver a constant output gap, instead implies

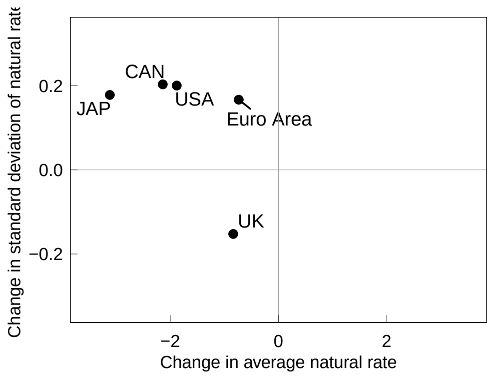
$$y_t^{gap} = \lim_T E_t y_T^{gap} - \frac{C_q}{C_Y}\left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right). \tag{2.14}$$

Since $C_q/C_Y < 0$, a positive (negative) housing price gap is then associated with a positive (negative) output gap: high housing prices stimulate housing investment and thereby output. Since the output expansion is inefficient, the policymaker

[46] See Appendix 2.H.1 for a derivation of this and subsequent results, including the generalized expressions for the case with $\rho_\xi < 1$.

[47] In the limit where the Kalman gain ($1/\alpha$) in the updating equation (2.5) approaches zero, the model with capital gain extrapolation converges to the RE model.

Figure 2.1: Changes in the average natural rate vs. changes in the volatility of the natural rate



Notes: This figure plots the pre-/post-1990 changes in the average natural rates against the changes in the natural rate volatility for several advanced economies. The volatilities of the natural rates in the pre-/post-1990 periods are the standard deviations of the linearly detrended series.

might find it optimal to *lean against housing prices.* The extent to which this is optimal will be explored quantitatively in Section 2.7 below.

The following lemma derives the natural rate $r_t^{*,\mathcal{P}}$ for our setting with subjective housing price beliefs:[48]

Lemma 1 *Let the natural rate of interest under subjective beliefs be given by*

$$r_t^{*,\mathcal{P}} \equiv r_t^{*,RE} - \frac{1}{\varphi}\frac{C_q}{C_Y}\left(\left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right) - E_t\left(\widehat{q}_{t+1}^u - \widehat{q}_{t+1}^{u*}\right)\right) \qquad \textit{for all } t. \tag{2.15}$$

When real interest rates are equal to $r_t^{*,\mathcal{P}}$ for all $t \geq 0$, then the IS equation (2.4) is consistent with

$$y_t^{gap} = \lim_T E_t y_T^{gap} \qquad \textit{for all } t. \tag{2.16}$$

The proof can be found in Appendix 2.H.1. Equation (2.15) generalizes the natural interest rate definition under RE to a setting with potentially subjective beliefs. In the special case with a constant housing price gap, we have $r_t^{*,\mathcal{P}} = r_t^{*,RE}$, even when the constant housing price gap differs from zero. This shows that the natural rate under subjective beliefs differs from it RE value if and only if the housing price gap is expected to go up or down. Since $C_q/C_Y < 0$, the natural rate will exceed (fall short of) its RE level, when the current housing price gap is higher (lower) than tomorrow's (expected) gap.

Since fluctuations in housing prices become larger when the average natural rate falls, the expected changes in the housing price gap will also become more volatile. A lower average level of the natural rate is thus not only associated with more volatile housing prices but also with more volatile natural rates of interest.

[48]As is the case with RE, it will generally not be optimal (or not even feasible) to set interest rates equal to the natural rate at all times due to the presence of a lower bound constraint on nominal rates

Figure 2.1 shows that this model prediction is consistent with the data. The figure plots the changes in the average level of the natural rate from the period before 1990 to the period after 1990 on the horizontal axis and the corresponding change in the natural rate *volatility* on the vertical axis. The volatilities of the natural rates in the pre-/post-1990 periods are the standard deviations of the linearly detrended series. The figure is again based on the estimates in Holston et al. (2017). While the level of the natural rate decreased, the volatility of it increased in four out of the five advanced economies. Appendix 2.G discusses the robustness of these results.

Housing price fluctuations and the misallocation of output. We now show that fluctuations in the housing price gap distort the allocation of output between its alternative uses, i.e., between housing investment and non-housing consumption. The housing investment gap, i.e., the difference between actual investment $\widehat{k}_t$ and its efficient level $\widehat{k}_t^*$, is – to a first-order approximation – given by

$$\widehat{k}_t - \widehat{k}_t^* = \frac{\tilde{\sigma}^{-1} C_Y}{1-\tilde{\alpha}} y_t^{gap} + \frac{1+\tilde{\sigma}^{-1} C_q}{1-\tilde{\alpha}} \left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right). \tag{2.17}$$

Under rational expectations, the housing price gap is zero and the investment gap is only distorted to the extent that the output gap is not closed. Additional output then gets allocated in constant proportions to housing investment and non-housing consumption, as $(\tilde{\sigma}^{-1} C_Y)/(1-\tilde{\alpha}) > 0$. In the presence of subjective beliefs, however, an additional distortion arises: the housing investment gap is then also driven by the housing price gap. Given the calibration considered later on, we have $(1+\tilde{\sigma}^{-1} C_q)/(1-\tilde{\alpha}) > 0$, so that a positive housing price gap ($\widehat{q}_t^u - \widehat{q}_t^{u*} > 0$) reinforces the investment distortions generated by a positive output gap.[49] This explains why the squared housing price gap shows up in the policymaker's objective function (2.2). While monetary policy cannot affect the housing price gap within our belief setup, it is the case that larger housing price gap fluctuations, as induced by lower natural rates, contribute to increased welfare losses.

2.6 Model calibration

To explore the quantitative implications for monetary policy arising from the presence of capital gain extrapolation, we consider a calibrated model. The calibration strategy consists of choosing a set of standard parameter values previously considered in the literature and of matching salient features of the behavior of natural interest rates and housing prices in the United States in the pre-1990 period. We then test the model by considering its predictions for the lower natural rate levels observed in the post-1990 period up to 2021. We compare across long time spans of 30 years each to obtain more reliable estimates of housing price volatility, which is difficult to estimate given the high degree of serial correlation of housing prices.

[49]This distortion in the allocation of output between housing investment and non-housing consumption is present independently of other frictions such as sticky prices or the lower-bound constraint on nominal interest rates.

Table 2.1: Model calibration

Parameter	Value	Source/Target
Preferences and technology		
β	0.9917	Average U.S. natural rate pre 1990
φ	1	Adam and Billi (2006)
κ_y	0.057	Adam and Billi (2006)
$\frac{\Lambda_y}{\Lambda_\pi}$	0.007	Adam and Billi (2006)
κ_q	-0.0023	Adam and Woodford (2021)
$\frac{C_q}{C_Y}$	-0.29633	Adam and Woodford (2021)
δ	0.03/4	Adam and Woodford (2021)
Exogenous shock processes		
ρ_{r^*}	0.8	Adam and Billi (2006)
σ_{r^*}	0.2940% (RE) 0.1394% (subj. beliefs)	Adam and Billi (2006)
ρ_ξ	0.99	Quarterly autocorr. of the PR of 0.99
σ_{ξ^d}	0.0233 (RE) 0.0165 (subj. beliefs)	Std. dev. of price-to-rent ratio pre 1990
Subjective belief parameters		
α	1/0.007	Adam, Marcet and Nicolini (2016)
β^U	1.0031	Max. rel. deviation of PR from mean

Calibration to the pre-1990 period. Table 2.1 summarizes the model parameterization. The quarterly discount factor β is chosen such that the steady-state natural rate equals the pre-1990 average of the U.S. natural rate of 3.34%, as estimated by Holston et al. (2017). The interest rate elasticity of output φ, the slope of the Phillips curve κ_y, and the welfare weight $\frac{\Lambda_y}{\Lambda_\pi}$ are taken from Table 2 in Adam and Billi (2006). The Phillips Curve coefficient κ_q and the ratio C_q/C_y are set as in Adam and Woodford (2021).[50]

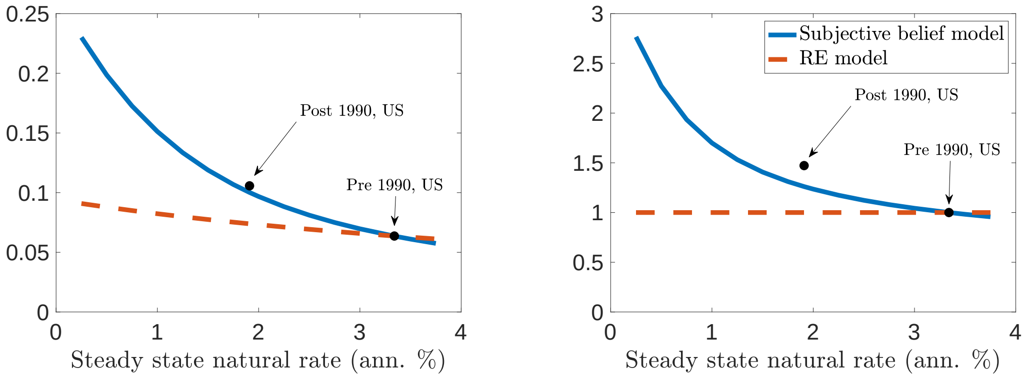
We now discuss the parameterization of the exogenous shock processes. The persistence of the housing preference shock ρ_ξ is set such that the RE model captures the high serial autocorrelation of the PR ratio in the data. The standard deviation of the innovations to the housing preferences σ_ξ are set such that the rational expectations and subjective belief models both replicate the pre-1990 standard deviation of the PR-ratio. For the subjective belief model, this is achieved by simulating equations (2.5) and (2.7), which requires specifying the belief updating parameters α and β^U. We set $\alpha = 1/0.007$ following Adam and Woodford (2012) and determine σ_{ξ^d} and β^U jointly such that (i) we match the volatility of the price-to-rent ratio and (ii) the simulated data matches the maximum deviation of the price-to-rent ratio in the data from its sample mean. The latter statistic identifies β^u. This procedure yields $\beta^U = 1.0031$ and $\sigma_{\xi^d} = 0.0165$. Housing demand distur-

[50]The calibration target for the ratio C_q/C_y is the ratio of residential fixed investment over the sum of nonresidential fixed investment and personal consumption expenditure, which is on average approximately equal to 6.3% in the US. This and the remaining parameters then imply $\kappa_q = -0.0023$, see Appendix 2.H.8 for details.

Figure 2.1: Standard deviation of price-to-rent ratio and natural rate

(a) Standard deviation of PR-ratio (relative to corresponding mean)

(b) Standard deviation of natural rate relative to case $\underline{r}^{*,RE} = 3.34\%$



Notes: This figure plots, for different steady state levels of the natural rate, the standard deviation of the price-to-rent ratio (relative to its mean) and the standard deviation of the natural rate.

bances are less volatile than under RE because fluctuations in subjective beliefs contribute to the fluctuations in housing prices. In fact, the calibration implies that about 50% of housing price fluctuations are due to subjective beliefs.

We consider the natural rate process

$$r_t^{*,RE} = \rho_{r*} r_{t-1}^{*,RE} + \varepsilon_t^r, \tag{2.1}$$

where $\varepsilon_t^r \sim iiN(0, \sigma_{r^*}^2)$. For the RE model, we set ρ_{r^*} and σ_{r^*} equal to the values in Adam and Billi (2006). For the subjective believe model, we use the same value for ρ_{r^*} but choose σ_{r^*} such that the generalized natural rate for the subjective belief model, defined in equation (2.15), has the same volatility as the natural rate in the RE model. This yields $\sigma_{r^*,RE} = 0.1393\%$, which is lower than under RE, because fluctuations in the housing price gap contribute to fluctuations in the natural rate in the presence of subjective beliefs. To economize on the number of state variables in the policy problem, we abstract from the presence of mark-up shocks.[51]

Evaluation of the model in the post-1990 period. Figure 2.1 illustrates the predictions of the RE model (dashed line) and subjective belief model (solid line) for the standard deviation of the price-to-rent ratio (panel a) and the standard deviation of the natural rate of interest (panel b). The panels depict these outcomes, which are independent of monetary policy, on the vertical axis for various levels of the steady-state natural rate on the horizontal axis. Variations in the steady-state level of the natural rate are achieved via appropriate variations in the

[51] Adam and Billi (2006) show that mark-up shocks are too small and display too little persistence to cause the lower-bound constraint to become binding.

discount factor.[52] The dots in Figure 2.1 report the average values for the pre- and post-1990 U.S. sample periods, where the average natural rate was equal to 3.34% and 1.91%, respectively.[53]

Since the model has been calibrated to the pre-1990 period, the RE and subjective belief model both match the pre-1990 data points in Figure 2.1. The subjective belief model also performs quite well in matching the post-1990 outcomes, despite the fact that these outcomes are not calibration targets. In particular, the standard deviation of the price-to-rent ratio and the standard deviation of the natural rate endogenously increase as the natural rate falls, with the magnitudes roughly matching the increase observed in the data. In contrast, the RE model produces no increase in the volatility of the natural rate and only a weak increase in the volatility of the price-to-rent ratio, for reasons discussed in Section 2.5.1. Matching the increase in housing price volatility under RE requires increasing the volatility of housing demand shocks. Since such an increase is irrelevant for monetary policy under RE, we leave the volatility of housing preference shocks unchanged. Similarly, matching the increase in the natural rate volatility under RE would require increasing σ_{r^*}. We will consider such increases when discussing our quantitative results.

2.7 Quantitative implications for monetary policy

This section illustrates the quantitative implications of falling natural rates and rising housing price volatility for the conduct of optimal monetary policy. It starts by determining the implications of falling natural rates for the optimal inflation target, i.e., for the average inflation rate implied by optimal monetary policy. It then illustrates the dramatically different optimal response to housing demand shocks under subjective and objective housing beliefs. Details of the nonlinear numerical solution procedure underlying the results in this section can be found in Appendix 2.H.6.

2.7.1 The optimal inflation target

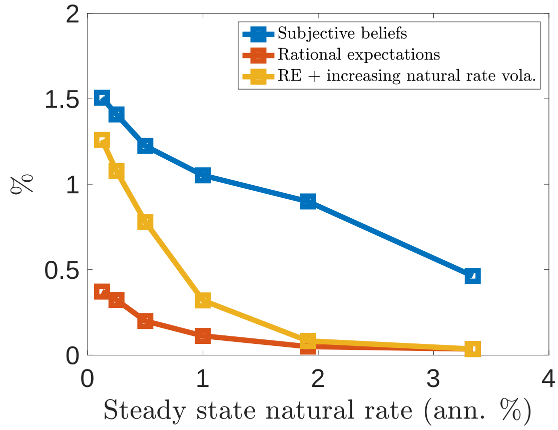
Figure 2.1 depicts the optimal inflation target for different steady-state levels of the natural rate of interest, i.e., the average inflation rate implied by optimal monetary policy. It shows the optimal target for the setup with subjective housing beliefs (upper line), for the case with rational housing price beliefs (lower line), and for a third case that we discuss below.

We find that the optimal target is close to zero, whenever housing expectations are rational. This holds quite independently of the average level of the natural rate, confirming earlier findings in Adam and Billi (2006) who considered the value

[52] As discussed before, variations in the discount factor may be driven by variations in the long-term growth rate and/or by variations in time-preferences.

[53] The reported increase in the standard deviation of the natural rate is again based on the estimates in Holston et al. (2017).

Figure 2.1: Average inflation under optimal monetary policy



Notes: The figure reports the optimal inflation target for different average levels of the natural rate in the presence of a zero lower bound constraint. The red line depicts the optimal target for the case with rational housing price beliefs and the blue line the one with subjective housing price beliefs. The yellow line shows the optimal average inflation under RE where the exogenous volatility of the natural rate is adjusted such that it matches the endogenous volatility increase under subjective beliefs.

for the average natural rate at the upper end of the range shown in Figure 2.1. This may appear surprising given that it is optimal for monetary policy to promise future inflation, so as to lower real interest rates, whenever adverse natural rate shocks cause the lower-bound constraint on nominal rates to bind. While the lower bound is reached more often when the average natural rate is low, inflation promises still have to be made relatively infrequently and can be quite modest. Hence, they do not significantly affect the average rate of inflation.

This result differs quite substantially from the ones reported in Andrade et al. (2019), who find that the optimal target should move up approximately one-to-one with a fall in the natural rate under rational expectations. Besides that Andrade et al. (2019) consider a medium-scale sticky price model without housing, the main difference to our approach is that they study Taylor rules with optimized intercepts rather than optimal monetary policy. As shown in Coibion et al. (2012) it makes a big difference for the optimal inflation target whether the monetary policy maker follows a Taylor rule or Ramsey optimal policy.

While lower natural rates trigger (slightly) larger housing price fluctuations under rational expectations, increased volatility is fully efficient and does not affect the natural rate of interest. Under rational expectations, the optimal inflation target is thus unaffected by housing price fluctuations, including for very low levels of the natural rate.

The upper line in Figure 2.1 shows that the situation is quite different with subjective housing beliefs. The optimal inflation target is overall substantially higher and also reacts more strongly to a fall in the average natural rate of interest. In fact, a fall in the steady-state natural rate from its pre-1990 average (3.34%) to its post-1990 average (1.9%) causes the optimal inflation target to increase by almost 0.5%. The corresponding increase under rational expectations is less than

0.05%. This difference is due to the fact that the endogenous volatility component of the natural rate increases once the natural rate drops. This reinforces the stringency of the zero lower bound, but is an effect that is absent under RE. It requires that the central bank engages more often in inflation promises, as it faces the lower bound constraint.

The optimal inflation target with subjective housing beliefs is substantially higher than the optimal target with RE, even at the pre-1990 average level of the natural rate. This is the case although the volatility of the natural rate is calibrated at this point to be equal across both models. This is due to two reasons: First, fluctuations in the housing price gap also generate cost-push term in the Phillips curve. Second, belief fluctuations induce more persistent variations in the natural rate than the exogenous natural rate shocks. This puts further upward pressure on the optimal inflation rate, as it requires larger and more persistent inflation promises by the central bank.

To illustrate this last point, the middle line in Figure 2.1 depicts the optimal inflation rate under rational expectations, when we set the volatility of the (exogenous) natural rate in the RE model such that it matches the volatility of the natural rate in the subjective belief model, for each considered level of the natural rate. While the optimal inflation rate increases relative to the benchmark RE setting, the level of the optimal inflation target still falls short of the one implied by subjective beliefs.

2.7.2 Leaning against housing demand shocks

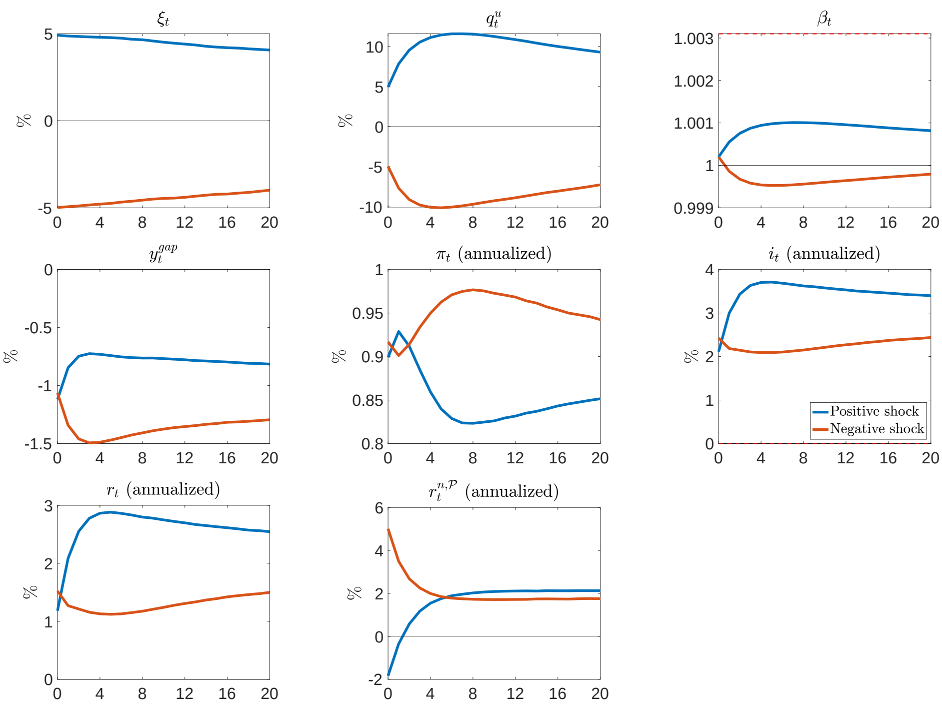
We now examine the optimal monetary policy response to housing demand shocks. Under RE, housing demand shocks affect the housing price and the efficient housing price identically, so that the housing price gap remains at zero. As a result, neither the output gap nor inflation respond to housing demand shocks. In contrast, it becomes optimal to "lean against" housing demand shocks in the presence of subjective beliefs. Yet, due to the lower bound constraint, the optimal response to positive and negative housing demand shocks displays considerable asymmetry.

The top row in Figure 2.2 shows the response of housing-related variables to a persistent positive/negative housing demand shock of 5%.[54] On impact, the shock triggers capital gains of an equal amount, which then trigger belief revisions that fuel further upward movements of the housing price in the same direction. The positive shock, for instance, pushes housing prices up by about 5% on impact, with belief momentum generating approximately another 5% in the first six quarters after the shock. This causes the housing price gap to become significantly positive (not shown in the figure). Once actual housing price increases start to fall short of the expected housing price increases, the housing boom reverts direction.

Higher housing prices push up housing investment, which causes upward pressure on the output gap. Optimal monetary policy leans strongly against the

[54]We initialize the economy at its ergodic mean and then hit the economy with a one-time shock of three standard deviations. We then average the subsequent response over the possible future shock realizations. We assume a steady-state natural rate equal to its post-1990 mean (1.91%).

Figure 2.2: Impulse responses to a housing preference shock



Notes: The figure reports the average impulse responses of the economy under subjective beliefs (at $\underline{r}^* = 1.91\%$) after a three-standard-deviation housing demand shock. The blue lines show the responses after a positive shock and the red lines after a negative shock.

housing price and increases nominal and real interest rates. It does so despite the fact that the natural rate of interest falls in response to the shock. The policy response causes a fall in inflation, which is amplified by the fact that the increase in housing prices and investment increases the marginal utility of consumption, hence, dampens wages and marginal costs. A positive housing demand shock thus results – in the presence of subjective housing beliefs – in a disinflationary housing boom episode under optimal monetary policy.

The policy response to a positive housing demand shock is much stronger than that to a negative housing demand shock. In particular, nominal and real interest rates fall considerably less following a negative shock realization. This is so because a negative housing price gap is inflationary and inflation is already high to start with. Negative housing demand shocks thus move inflation further away from its optimal level of zero.[55] Yet, policy still "leans against" the housing price decrease: real interest rates fall despite the fact that the natural rate increases.

The fact that leaning against housing prices can be optimal in the presence of capital gain extrapolation is in line with results in Caines and Winkler (2021), who consider a setting with 'conditionally model consistent beliefs' in which expectations differ for many variables from rational expectations, and with results in

[55] While the output gap is moved closer to its optimal level, the weight on the output gap in the welfare function is two orders so magnitude smaller than that on inflation, see Table 2.1.

Adam and Woodford (2021), who consider a setting where the policymaker fears 'worst-case' belief distortions about inflation and housing price expectations. As none of these papers consider a lower-bound constraint, the policy response to positive and negative shocks is symmetric in their settings.

2.8 The role of macroprudential policy

It is often argued that macroprudential policies can be used to stabilize financial markets and that this would allow monetary policy to ignore disturbances coming from the housing sector, see Svensson (2018) for a prominent exposition of this view. In this section, we evaluate the quantitative plausibility of this view within our setup with subjective housing beliefs.

We show below that fully eliminating fluctuations in the housing price gap requires imposing large and volatile macroprudential taxes. None of the macroprudential instruments thus far available in advanced economies appear suited to achieve economic effects anywhere near the required size. In addition, it is often necessary for macroprudential policy to pay substantial subsidies. To the best of our knowledge, none of the available macroprudential instruments acts in a way that subsidizes actions by economic actors. Less aggressive policies, that aim at only partly eliminating the housing price gap, still require considerable tax volatility, because fluctuations in subjective beliefs turn out not to be independent of tax policy pursued.

We analyze the issue by considering a setup in which the policymaker can tax or subsidize the ownership of housing. While actual macroprudential policies often operate via constraints imposed on the banking sector, their ultimate effect is to make housing more or less expensive to households. For this reason, we consider taxes and subsidies at the household level.

Specifically, we analyze a proportional and time-varying tax τ_t^D that is applied to the rental value of housing in every period t. A household owning D_t units of houses, then has to pay taxes of

$$\tau_t^D D_t R_t \tag{2.1}$$

units of consumption.[56] We find this specification more plausible than a policy that taxes the market value of housing, as it is difficult to determine market values in real time. A setup that taxes the physical housing units, i.e., where taxes are equal to $\tau_t^D D_t$, delivers very similar results, but is analytically more cumbersome. Furthermore, the tax setup in equation (2.1) is equivalent to a setup where taxes directly affect household utility, i.e., where the utility contribution from owning houses would instead be given by $\xi_t^d \left(1 - \tau_t^D\right) D_t$ and no monetary taxes would have to be paid. We prefer the formulation in equation (2.1) because it allows expressing taxes in monetary units.

In the presence of these taxes, housing prices under subjective beliefs are given

[56] To keep the rest of the model unchanged, the household also needs to expect lump sum tax rebates that are equal to the amount of subjectively expected tax payments.

by

$$q_t^u = \frac{\left(1-\tau_t^D\right)\xi_t^d}{1-\beta(1-\delta)\beta_t}, \tag{2.2}$$

and the housing-price *gap* in percentage deviations from the steady state (where $\underline{\tau}^D = 0$) is

$$\begin{aligned}\widehat{q}_t^u - \widehat{q}_t^{u*} =& \frac{(1-\beta(1-\delta))\left(1-\tau_t^D\right)\widehat{\xi}_t^d}{1-\beta(1-\delta)\beta_t} + \frac{\beta(1-\delta)(\beta_t-1)}{1-\beta(1-\delta)\beta_t} - \frac{1-\beta(1-\delta)}{1-\beta(1-\delta)\beta_t}\tau_t^D \\ &- \frac{1-\beta(1-\delta)}{1-\beta(1-\delta)\rho_\xi}\widehat{\xi}_t^d. \end{aligned} \tag{2.3}$$

The previous equation shows that macroprudential policy must eliminate housing price gap fluctuations that are due to housing demand shocks ($\widehat{\xi}_t^d$) and due to fluctuations in subjective capital gain expectations (β_t). Doing so requires setting the tax according to

$$\tau_t^{D*} = \frac{\beta(1-\delta)}{1+\widehat{\xi}_t^d}\left[\frac{(\beta_t-\rho_\xi)}{1-\beta(1-\delta)\rho_\xi}\widehat{\xi}_t^d + \frac{1}{1-\beta(1-\delta)}(\beta_t-1)\right]. \tag{2.4}$$

To understand what the preceding equation implies for the behavior of taxes, one has to take into account that the fluctuations in subjective beliefs (β_t) depend themselves on the tax: the tax influences housing prices, see equation (2.2), and thus – via capital gain extrapolation – the evolution of subjective beliefs.

To analyze the behavior of taxes, we consider the calibrated subjective belief model from Section 2.6 for the case where the average natural rate is equal to its post-1990 average (1.9%). We consider also intermediate forms of taxation that do not aim at fully eliminating the housing gap, by specifying taxes as

$$\tau_t^D = \lambda^D \tau_t^{D*},$$

where $\lambda^D \in [0,1]$ is a sensitivity parameter. Our prior setup assumed $\lambda^D = 0$, while fully eliminating the housing price gap using macroprudential policy requires setting $\lambda^D = 1$. We then simulate the dynamics of housing prices, beliefs and taxes for alternative values of λ^D.

Table 2.1 reports the main outcomes. It shows that a higher tax sensitivity (λ^D) steadily reduces the standard deviation of the housing price gap (second column). However, the standard deviation of taxes has to steadily increase. For a policy that fully eliminates the housing price gap ($\lambda^D = 1$), the standard deviation of taxes is a staggering 8% of the rental value of housing. Taxes reach maximum values up to 24% and minimum values deeply in negative territory, with subsidies above 40% of the rental value. These taxes fully stabilize the housing price gap but still induce substantial variation in subjective beliefs. The latter explains why taxes have to remain rather volatile. Intermediate policies, say ones that set $\lambda^D = 0.4$, substantially reduce the volatility of the housing gap, but still require rather volatile taxes and often very large subsidies.

Given the outcomes in Table 2.1, we conclude that the currently available macroprudential instruments will unlikely be able to insulate the monetary authority from disturbances in the housing sector arising from capital gain extrapolation.

Table 2.1: Taxes and housing price fluctuations for alternative tax sensitivities λ^D

Tax sensitivity λ^D	Housing price gap $\widehat{q}_t^u - \widehat{q}_t^{u*}$	Housing taxes τ_t^D		
Value	Std. dev.	Std. dev.	Max	Min
0.0	14.2%	0.0%	0.0%	0.0%
0.2	9.8%	2.4%	7.0%	-12.1%
0.4	6.4%	4.2%	13.8%	-21.7%
0.6	3.7%	5.7%	18.0%	-30.0%
0.8	1.7%	7.0%	21.3%	-36.2%
1.0	0.0%	8.0%	23.9%	-41.8%

Notes: The table reports the standard deviation of the housing gap, $\widehat{q}_t^u - \widehat{q}_t^{u*}$, as well as the standard deviation, minimum value and maximum value of the macroprudential tax τ^D, for different tax sensitivities λ^D.

2.9 Conclusion

This paper documents systematic deviations from rational housing price expectations and constructs a structural equilibrium model that jointly replicates the behavior of housing prices and the patterns of deviations from rational expectations. The model shows that subjective housing price beliefs significantly contribute to housing price fluctuations and that lower natural rates of interest generate increased volatility for housing prices and the natural rate.

Optimal monetary policy responds to falling and more volatile natural rates by implementing higher average rates of inflation. Monetary policy should also lean against housing price fluctuations induced by housing demand shocks, with reactions to housing price increases being more forceful than the reaction to housing price downturns. None of these features is optimal if households hold rational housing price expectations. This highlights the importance of basing policy advice on economic models featuring empirically plausible specifications for household beliefs.

Appendices for Chapter 2

2.A Additional results for Section 2.2

2.A.1 Five-year-ahead capital gain expectations

While for our baseline results in Section 2.2 we focus on short-term housing price expectations, our findings equally hold for medium-term five-year-ahead expectations. We estimate the five-year analogue of regression (2.1) as follows:

$$q_{t+20} - E_t^{\mathcal{P}}\left[q_{t+20}\right] = a^{CG} + b^{CG} \cdot \left(E_t^{\mathcal{P}}\left[q_{t+20}\right] - E_{t-1}^{\mathcal{P}}\left[q_{t+19}\right]\right) + \varepsilon_t. \tag{2.A.1}$$

Table 2.A.1 reports the estimates of b^{CG} showing that five-year expectations are updated sluggishly.

Table 2.A.1: Sluggish adjustment of five-year-ahead housing price expectations

	Mean Expectations	Median Expectations
$\widehat{b}^{CG}$	6.95***	6.89***
	(1.703)	(1.680)

Notes: This table reports the empirical estimates of regression (2.A.1) using nominal housing-price expectations. The reported standard errors are robust with respect to heteroskedasticity and serial correlation (Newey–West with four lags). Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

We also run five-year-ahead versions of the regressions (2.2) and (2.3):

$$E_t^{\mathcal{P}}\left[\frac{q_{t+20}}{q_t}\right] = a + c \cdot PR_{t-1} + u_t \tag{2.A.2}$$

$$\frac{q_{t+20}}{q_t} = \mathbf{a} + \mathbf{c} \cdot PR_{t-1} + \mathbf{u}_t. \tag{2.A.3}$$

Table 2.A.2 shows that five-year-ahead capital gain expectations covary positively with the price-to-rent ratio, whereas actual capital gains covary negatively.

Table 2.A.2: Expected vs. actual capital gains using five-year-ahead housing price expectations

	$\hat{c}$ (in %)	$\hat{\mathbf{c}}$ (in %)	bias (in %) $-E(\hat{\mathbf{c}} - \hat{c})$	p-value $H_0 : c = \mathbf{c}$
Mean Expectations	0.045	-1.889	0.0159	0.000
	(0.0001)	(0.01997)		
Median Expectations	0.044	-1.889	0.0155	0.000
	(0.00024)	(0.01997)		

Notes: $\widehat{c}$ is the estimate of c in equation (2.A.2) and $\widehat{\mathbf{c}}$ the estimate of $\mathbf{c}$ in equation (2.A.3). The Stambaugh (1999) small sample bias correction is reported in the second-to-last column and the last column reports the p-values for the null hypothesis $c = \mathbf{c}$. Newey–West standard errors using four lags in parentheses.

2.A.2 IV estimation of sluggish belief updating

To insure that the results obtained from regression (2.1) in Section 2.2 are not driven by forecast revisions being correlated with the error term, we follow Coibion and Gorodnichenko (2015) by adopting an Instrumental Variable approach. Specifically, we consider monetary policy shocks as an instrument for forecast revisions. We identify daily monetary policy shocks as changes of the current-month federal funds future in a 30-minute window around scheduled FOMC announcements (following the approach in Gürkaynak, Sack, and Swanson (2005) and Gorodnichenko and Weber (2016)). We then aggregate shocks to quarterly frequency by assigning daily shocks partly to the current quarter and partly to the consecutive quarter, based on the number of remaining days in the current quarter. Table 2.A.3 reports the results of the IV regression. The coefficients are positive and statistically significant, with point estimates that are even larger than the ones reported in Section 2.2.

Table 2.A.3: Instrumental variable regression

	Mean Expectations	Median Expectations
Nominal Housing Prices		
$\widehat{b}^{CG}$	2.85^{**}	3.84^{***}
	(1.259)	(1.497)
First-stage F-statistic	21.88	17.78
Real Housing Prices		
$\widehat{b}^{CG}$	2.62^{***}	3.45^{***}
	(0.745)	(0.649)
First-stage F-statistic	44.49	34.13

Notes: $\widehat{b}^{CG}$ report the results from regression (2.1), instrumenting forecast revisions using monetary policy shocks, obtained via high-frequency identification. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

2.A.3 Sluggish adjustment of capital gain expectations

Regression (2.1) in Section 2.2 studies sluggish adjustment of expectations about the housing price level. Similar results can be obtained when considering expectations about capital gains. *Specification 1* in Table 2.A.4 reports the regression coefficient when one replaces actual and expected housing price levels on the left-hand side of equation (2.1) with actual and expected capital gains. The coefficient estimates remain positive and highly statistically significant. *Specification 2* in Table 2.A.4 reports results when replacing expectations about housing price levels with expectations about capital gains on the right-hand side of equation (2.1) and *Specification 3* reports results when replacing levels by (actual and expected) capital gains on both sides of equation (2.1). The coefficient estimates remain positive, but the significance levels are lower for Specifications 2 and 3.

Table 2.A.4: Sluggish adjustment of housing price growth expectations

	Mean Expectations	Median Expectations
Specification 1		
Nominal Housing Prices		
$\widehat{b}^{CG}$	0.023***	0.030***
	(0.005)	(0.005)
Real Housing Prices		
$\widehat{b}^{CG}$	0.024***	0.031***
	(0.004)	(0.004)
Specification 2		
Nominal Housing Prices		
$\widehat{b}^{CG}$	492*	182
	(279)	(210)
Real Housing Prices		
$\widehat{b}^{CG}$	302*	158
	(164)	(168)
Specification 3		
Nominal Housing Prices		
$\widehat{b}^{CG}$	5.20*	2.16
	(2.896)	(2.06)
Real Housing Prices		
$\widehat{b}^{CG}$	3.23*	2.06
	(1.678)	(1.835)

Notes: This table shows the results of regression (2.1) in terms of house-price growth rates instead of house-price levels. *Specification 1* denotes the case in which we replace housing-price levels with capital gains on the left-hand side of regression (2.1), *Specification 2* the case in which we replace the right-hand side and *Specification 3* denotes the case in which we replace levels with capital gains on both sides of regression (2.1). The reported standard errors are robust with respect to heteroskedasticity and serial correlation (Newey–West with four lags). Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

2.A.4 Cyclicality of housing price forecast errors

A similar version of the test from Adam et al. (2017) presented in Section 2.2, which considers the cyclicality of expected gains, is proposed by Kohlhas and Walther (2021). In this case, we regress forecast errors about housing prices on the price-to-rent ratio. Formally, we estimate

$$\frac{q_{t+4}}{q_t} - E_t^{\mathcal{P}}\left[\frac{q_{t+4}}{q_t}\right] = \alpha + \gamma \cdot PR_{t-1} + \varepsilon_t. \tag{2.A.4}$$

Table 2.A.5 shows the results. We find a negative and statistically significant coefficient in all cases. Thus, consumers tend to become too optimistic (pessimistic) when they observe high (low) housing valuations, inconsistent with rational expectations.

Table 2.A.5: Forecast errors and price-to-rent ratios

	Mean Expectations	Median Expectations
Nominal Housing Prices		
$\widehat{\gamma}$	-0.5^{***}	-0.5^{***}
	(0.09)	(0.10)
Real Housing Prices		
$\widehat{\gamma}$	-0.5^{***}	-0.5^{***}
	(0.08)	(0.10)

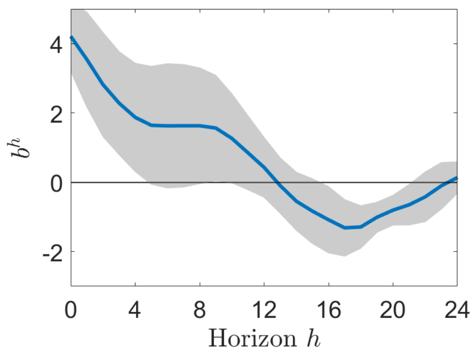
Notes: This table shows the results of regression (2.A.4), whereas the estimated regression coefficients (and standard errors) are multiplied by one hundred for better readability. The reported standard errors are robust with respect to heteroskedasticity and serial correlation (Newey–West with four lags). Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

2.A.5 Dynamics of forecast errors with median and nominal housing price expectations

Figure 2.A.1 shows alternative specifications of the dynamic forecast error responses presented in Section 2.2. Panel (a) presents the response of forecast errors for nominal housing prices. Panel (b) shows the response of forecast errors for real housing prices (as in Section 2.2) but considering median expectations. The figure shows that these responses are very close to the baseline specification shown in Section 2.2.

Figure 2.A.1: Dynamic Forecast error response to realized capital gains

(a) Nominal (mean) Expectations (b) Median (real) expectations



Notes: Panel (a) shows impulse-response functions of nominal capital gain forecast errors to a one standard deviation innovation in the housing capital gain. Panel (b) shows the impulse-response functions of median (real) capital gain forecast errors of one-year ahead expectations to a one standard deviation innovation in the housing capital gain. The shaded area shows the 90%-confidence intervals, standard errors are robust with respect to serial correlation and heteroskedasticity (Newey–West with $h+1$ lags).

2.A.6 Results when excluding the Corona virus period

The empirical results reported in Section 2.2 are based on the entire period for which household-survey expectations are available, i.e., 2007-2021. This section reports results obtained when ending the sample in 2019, thereby excluding the recent Corona Virus crisis period. This is motivated by the fact that the two largest outliers in Figure 2.1 fall into the period after 2019. Tables 2.A.6 and 2.A.7 show, however, that our results are qualitatively and quantitatively robust to excluding observations from the years 2020 and 2021.

Table 2.A.6: Sluggish adjustment of housing price expectations: excluding coronavirus crisis

	Mean Expectations	Median Expectations
Nominal Housing Prices		
$\widehat{b}^{CG}$	2.18***	2.80***
	(0.503)	(0.502)
Real Housing Prices		
$\widehat{b}^{CG}$	1.97***	2.43***
	(0.332)	(0.360)

Notes: This table shows the results of regression (2.1) excluding the coronavirus crisis, i.e., we exclude the years 2020 and 2021. The reported standard errors are robust with respect to heteroskedasticity and serial correlation (Newey–West with four lags). Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2.A.7: Expected vs. actual capital gains: excluding coronavirus crisis

	$\hat{c}$ (in %)	$\hat{\mathbf{c}}$ (in %)	bias (in %) $-E(\hat{\mathbf{c}} - \hat{c})$	p-value $H_0 : c = \mathbf{c}$
Nominal Housing Prices				
Mean Expectations	0.058	-0.065	0.0036	0.000
	(0.0066)	(0.0126)		
Median Expectations	0.018	-0.065	0.0118	0.042
	(0.0010)	(0.0126)		
Real Housing Prices				
Mean Expectations	0.0614	-0.0483	-0.0009	0.000
	(0.0136)	(0.0090)		
Median Expectations	0.196	-0.483	0.076	0.017
	(0.0034)	(0.0090)		

Notes: This table shows the results of regressions (2.2) and (2.3) excluding the coronavirus period, i.e., we exclude the years 2020 and 2021. $\widehat{c}$ is the estimate of c in equation (2.2) and $\widehat{\mathbf{c}}$ the estimate of $\mathbf{c}$ in equation (2.3). The small sample bias correction is reported in the second to last column and the last column reports the p-values for the null hypothesis $c = \mathbf{c}$ in the fifth column. Newey–West standard errors using four lags in parentheses.

2.A.7 Regional housing prices and expectations

This appendix considers regional variation in housing prices and housing price expectations. This is possible because the Michigan Survey reports the location of respondents using four different regions: West, North East, North Central (or Midwest) and South. While the Case-Shiller Price Index is not available at this level of regional disaggregation, we can construct a regional housing price index using the Case-Shiller Index that is available for twenty large U.S. cities. Following the definition of the regions in the Michigan Survey, we assign the twenty cities to the four regions and then aggregate city price indices to a regional index using two alternative approaches. The first approach weighs cities by population (as of 2019) within each region, while the second approach uses equal weights for all cities within a region.

Table 2.A.8 lists all twenty cities, the region to which we allocate them and their regional population weights.[57] As in our baseline approach using aggregate data, we deflate nominal housing price indices by the aggregate CPI to obtain a real housing price index. We obtain real housing price expectations by deflating nominal (mean) expectations with region-specific (mean) inflation expectations.

Table 2.A.8: Regions, cities and their weights

City	Region	Weight	City	Region	Weight
Denver	West	$\frac{0.705}{10.595}$	Chicago	North Central	$\frac{2.71}{4.189}$
Las Vegas	West	$\frac{0.634}{10.595}$	Cleveland	North Central	$\frac{0.385}{4.189}$
Los Angeles	West	$\frac{3.97}{10.595}$	Detroit	North Central	$\frac{0.674}{4.189}$
Phoenix	West	$\frac{1.633}{10.595}$	Minneapolis	North Central	$\frac{0.42}{4.189}$
Portland	West	$\frac{0.645}{10.595}$	Atlanta	South	$\frac{0.488}{4.209}$
San Diego	West	$\frac{1.41}{10.595}$	Charlotte	South	$\frac{0.857}{4.209}$
San Francisco	West	$\frac{0.874}{10.595}$	Dallas	South	$\frac{1.331}{4.209}$
Seattle	West	$\frac{0.724}{10.595}$	Miami	South	$\frac{0.454}{4.209}$
Boston	North East	$\frac{0.68}{9.1}$	Tampa	South	$\frac{0.387}{4.209}$
New York	North East	$\frac{8.42}{9.1}$	Washington DC	South	$\frac{0.692}{4.209}$

Notes: This table lists the twenty cities for which the Case-Shiller Home Price Index is available, the region to which the cities are allocated based on the Michigan Survey and their respective weights within region.

Table 2.A.9 reports the region-specific estimates of b^{CG} in regression equation (2.1). All point estimates are positive with magnitudes that are broadly in line with the estimates at the national level. Furthermore, all regional estimates are significant at the 1% level. This shows that households update expectations sluggishly in all regions, consistent with the findings reported for the national level reported in the main text.

[57] The weights are calculated as the ratio of the population in the considered city, divided by the sum of populations in all cities in the respective region.

Table 2.A.9: Sluggish adjustment of housing price expectations across regions

	Weighted	Unweighted
$\widehat{b}^{CG,W}$	2.00***	1.95***
	(0.411)	(0.374)
$\widehat{b}^{CG,NE}$	1.24***	1.15***
	(0.385)	(0.441)
$\widehat{b}^{CG,NC}$	1.97***	1.95***
	(0.461)	(0.459)
$\widehat{b}^{CG,S}$	1.74***	1.94***
	(0.385)	(0.393)

Notes: This table shows the results of regression (2.1) using regional housing prices and expectations. The superscripts W, NE, NC and S denote the regions West, North East, North Central (or Midwest) and South, respectively. The reported standard errors are robust with respect to heteroskedasticity and serial correlation (Newey–West with four lags). Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2.A.10 reports the region-specific estimates of c and $\mathbf{c}$ from regressions (2.2) and (2.3). Since regional price-to-rent ratios are not available, the regression uses real housing prices on the right-hand side. In line with our findings at the aggregate level, we find $c > 0$ and $\mathbf{c} < 0$ in all the regions with the differences being largely highly statistically significant.

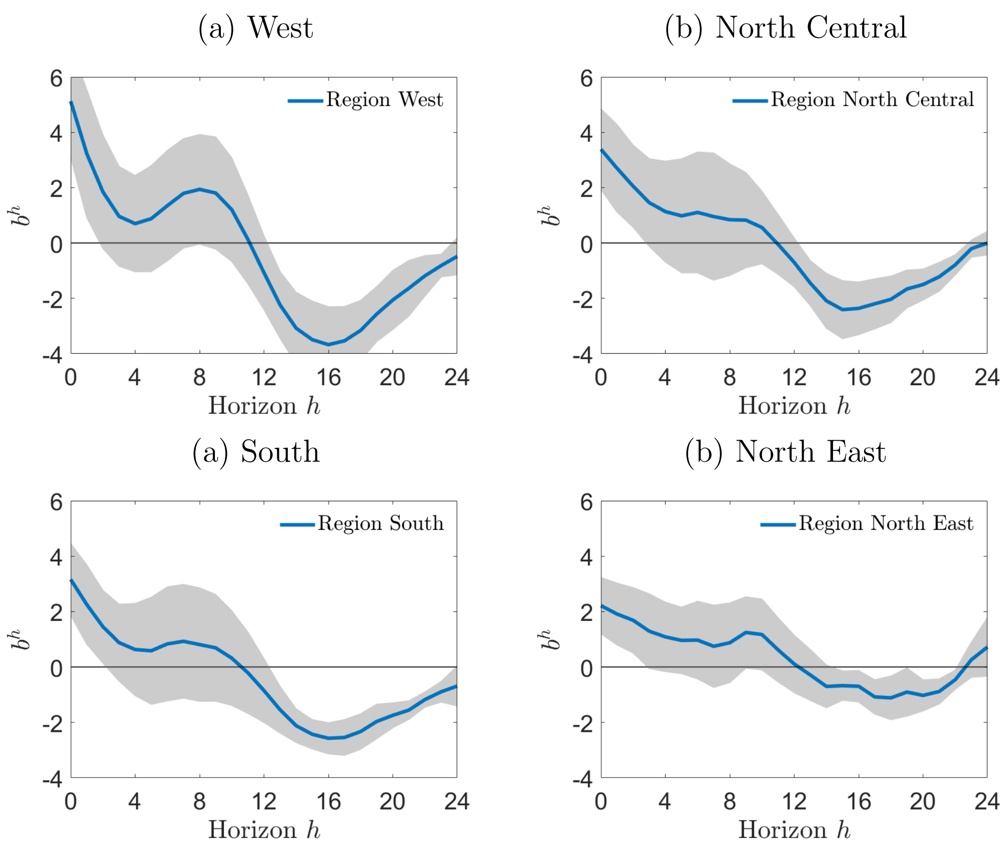
Figure 2.A.2 shows the dynamic forecast errors responses to a one standard deviation innovation in the real housing capital gain in each of the four regions. Figure 2.A.3 shows the results for the case in which the cities within regions are equally weighted. In line with the findings reported in the main text, households' housing capital gain expectations initially underreact but overshoot after some time.

Table 2.A.10: Expected vs. actual capital gains across regions

	$\hat{c}$ (in %)	$\hat{\mathbf{c}}$ (in %)	bias (in %) $-E(\hat{\mathbf{c}} - \hat{c})$	p-value $H_0 : c = \mathbf{c}$
West				
Population-weighted	0.109	-0.216	0.090	0.083
	(0.0036)	(0.1360)		
Equally weighted	0.132	-0132	0.137	0.183
	(0.0034)	(0.1197)		
North Central				
Population-weighted	0.045	-0.544	0.008	0.000
	(0.0089)	(0.0256)		
Equally weighted	0.088	-0.458	0.0191	0.000
	(0.0118)	(0.0769)		
North East				
Population-weighted	0.013	-0.474	0.001	0.000
	(0.0089)	(0.0072)		
Equally weighted	0.126	-0.315	0.023	0.000
	(0.0187)	(0.0838)		
South				
Population-weighted	0.210	-0.008	0.137	0.144
	(0.0023)	(0.1067)		
Equally weighted	0.163	-0.238	0.055	0.014
	(0.0044)	(0.1250)		

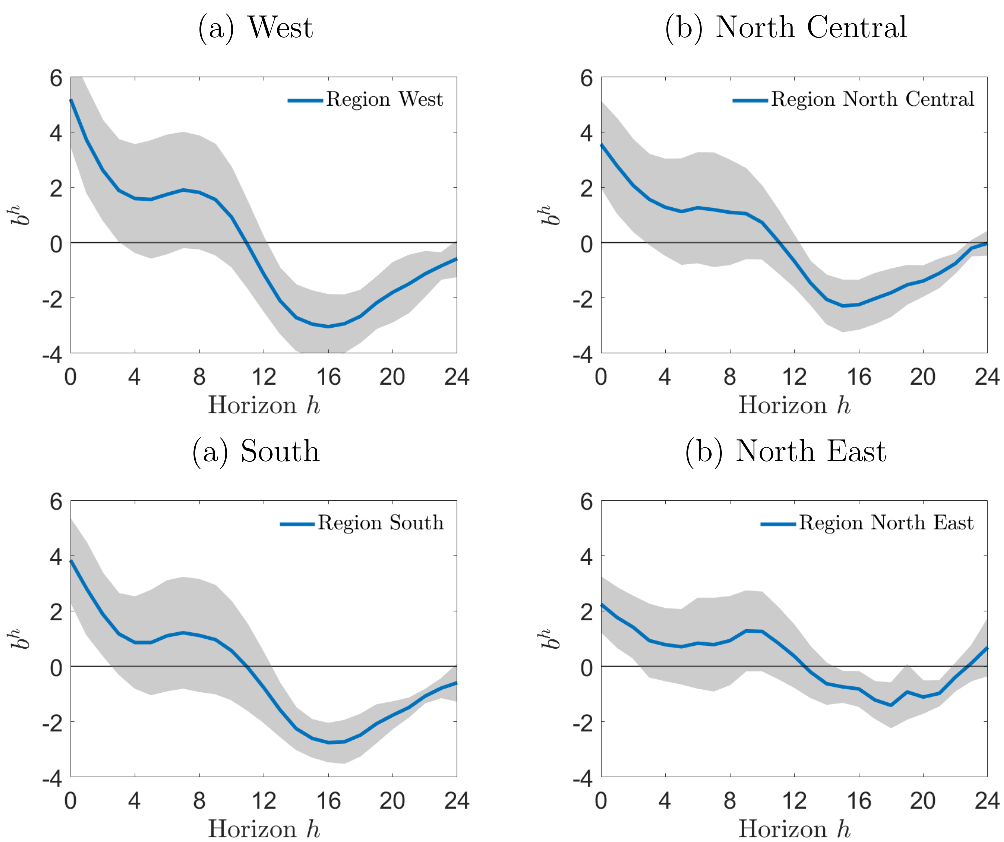
Notes: This table shows the results of regressions (2.2) and (2.3) for different regions. $\widehat{c}$ is the estimate of c in equation (2.2) and $\widehat{\mathbf{c}}$ the estimate of $\mathbf{c}$ in equation (2.3). The small sample bias correction is reported in the second to last column and the last column reports the p-values for the null hypothesis $c = \mathbf{c}$ in the fifth column. Newey–West standard errors using four lags in parentheses.

Figure 2.A.2: Regional dynamic forecast error responses to realized capital gains (population-weighted city housing price index)



Notes: The figure shows the dynamic response of real capital gain forecast errors across the four different regions (in which cities' housing indices are weighted by their population share) to a one standard deviation innovation in the housing capital gain. The shaded area shows the 90%-confidence intervals, standard errors are robust with respect to serial correlation and heteroskedasticity (Newey–West with $h + 1$ lags).

Figure 2.A.3: Regional dynamic forecast error response to realized capital gains (equal weights)



Notes: The figure shows the dynamic response of real capital gain forecast errors across the four different regions (in which cities are equally weighted) to a one standard deviation innovation in the housing capital gain. The shaded area shows the 90%-confidence intervals, standard errors are robust with respect to serial correlation and heteroskedasticity (Newey–West with $h+1$ lags).

2.B Additional results for Section 2.3

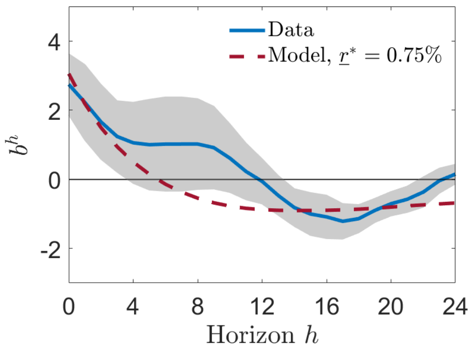
2.B.1 Dynamic forecast error eesponses: Housing price level

Figure 2.B.1 shows that the simple housing model also not only matches the empirical dynamic forecast error response about capital gains well, but also does a good job in matching the forecast errors about the level of future housing prices. The results are obtained by defining the forecast error X_{t+h} in equation (2.4) as

$$X_{t+h} \equiv q_{t+4+h} - E^{\mathcal{P}}_{t+h}\left[q_{t+4+h}\right] \tag{2.B.1}$$

and estimating the resulting local projections in the data and the population local projection for the model. Figure 2.B.1 shows that households' expectations about the future level of housing prices initially undershoot and subsequently overshoot, as is the case with expected capital gains.

Figure 2.B.1: Dynamic forecast error responses: housing price levels

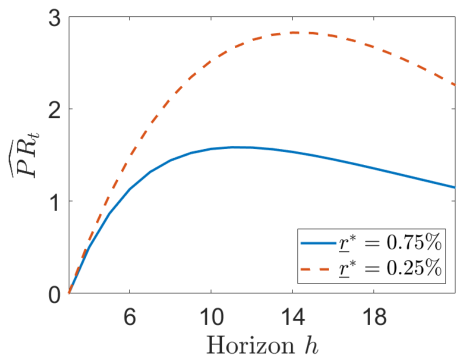


Notes: The figure shows impulse-response functions of housing-price level forecast errors of one-year ahead expectations to a one standard deviation innovation in the housing capital gain from the data and in the data. The shaded area shows the 90%-confidence intervals of the empirical estimates, standard errors are robust with respect to serial correlation and heteroskedasticity (Newey–West with $h + 1$ lags).

2.B.2 Model response of the PR-ratio to housing demand shocks

Section 2.3 shows that real housing prices are more sensitive to housing demand shocks at lower levels of the natural rate. Figure 2.B.2 illustrates that the same holds true for the model-implied price-to-rent ratio. The figure depicts the structural impulse response of the price-to-rent ratio (in percent deviations from steady state) to a one standard deviation housing-preference shock. It shows the response for a natural rate of 0.75% (blue line) and 0.25% (red line). The IRFs for the price-to-rent ration look very similar to the ones for real housing prices, shown in Figure 2.3(a).

Figure 2.B.2: Impulse response functions



Notes: This figure shows the structural impulse response functions of the price-to-rent ratio (in percent deviations from steady state) to a one standard deviation housing-preference shock for different natural rates.

2.C The nonlinear optimal policy problem

We shall consider Ramsey optimal policies in which the policymaker chooses the sequence of policy rates, prices, and allocations to maximize rationally expected household utility, subject to the constraint that prices and allocations constitute an Internally Rational Expectations Equilibrium. Note that the policymaker maximizes utility under a probability measure that is different from the one entertained by households, whenever the latter hold subjective beliefs. Benigno and Paciello (2014) refer to such a policymaker as being 'paternalistic'.

The objective of the policymaker is to maximize household utility. Using equation (2.7) to express the relative quantities demanded of the differentiated goods each period as a function of their relative prices and the linear dependence of utility on the stock of assets, we can write the utility flow to the representative household in the form

$$u(Y_t, q_t^u; \xi_t) - v(Y_t; \xi_t)\Delta_t + \bar{\xi}_t^d \frac{A_t^d}{\tilde{\alpha}} k_t^{\tilde{\alpha}},$$

with

$$\begin{aligned} u(Y_t, q_t^u; \xi_t) &\equiv \tilde{u}(C(Y_t, q_t^u, \xi_t); \xi_t) \\ v(y_t^j; \xi_t) &\equiv \tilde{v}(f^{-1}(y_t^j/A_t); \xi_t), \end{aligned}$$

where Δ_t, defined in equation (2.22), captures the misallocations from price dispersion. The term

$$\bar{\xi}_t^d \equiv \sum_{T=t}^{\infty} E_t[(1-\delta)^{T-t}\, \beta^{T-t} \xi_T^d]$$

captures the present value contribution from new housing investment. We can use (2.12) and (2.28) to express k_t in terms of Y_t, q_t^u and exogenous shocks. Hence, we can express the policy maker's objective of maximizing (2.1) under rational expectations, as maximizing

$$U = E_0 \sum_{t=0}^{\infty} \beta^t U(Y_t, \Delta_t, q_t^u; \xi_t),$$

where the flow utility is given by

$$\begin{aligned} U(Y_t, \Delta_t, q_t^u; \xi_t) \equiv & \frac{\bar{C}_t^{\tilde{\sigma}^{-1}} C(Y_t, q_t^u, \xi_t)^{1-\tilde{\sigma}^{-1}}}{1-\tilde{\sigma}^{-1}} \\ & - \frac{\lambda}{1+\nu} \bar{H}_t^{-\nu} \left(\frac{Y_t}{A_t}\right)^{1+\omega} \Delta_t \\ & + \frac{A_t^d \bar{\xi}_t^d}{\tilde{\alpha}} \Omega(q_t^u, \xi_t)^{\tilde{\alpha}}\; C(Y_t, q_t^u, \xi_t)^{\frac{\tilde{\alpha}}{1-\tilde{\alpha}} \tilde{\sigma}^{-1}}, \end{aligned} \tag{2.C.1}$$

which is a monotonically decreasing function of Δ given Y, q^u and ξ, and where $\Omega(q^u, \xi)$ is the function defined in (2.27). The only endogenous variables that are relevant for evaluating the policymaker's objective function are thus Y_t, Δ_t and q_t^u.

The non-linear optimal monetary policy problem is then given by

$$\max_{\{Y_t, q_t^u, p_t^*, w_t(j), P_t, \Delta_t, i_t \geq 0\}} E_0 \sum_{t=0}^{\infty} \beta^t U(Y_t, \Delta_t, q_t^u; \xi_t) \tag{2.C.2}$$

subject to

$$\left(\frac{p_t^*}{P_t}\right)^{1+\eta(\phi-1)} = \frac{E_t^{\mathcal{P}} \sum_{T=t}^{\infty} (\alpha)^{T-t} Q_{t,T} \frac{\eta}{\eta-1} \phi w_T(j) \left(\frac{Y_T}{A_T}\right)^{\phi} \left(\frac{P_T}{P_t}\right)^{\eta\phi+1}}{E_t^{\mathcal{P}} \sum_{T=t}^{\infty} (\alpha)^{T-t} Q_{t,T}(1-\tau_T) Y_T \left(\frac{P_T}{P_t}\right)^{\eta}} \tag{2.C.3}$$

$$w_t(j) = \lambda \frac{\bar{H}_t^{-\nu}}{\bar{C}_t^{\tilde{\sigma}^{-1}}} \left(\frac{Y_t}{A_t}\right)^{\phi\nu} C\left(Y_t, q_t^u, \xi_t\right)^{\tilde{\sigma}^{-1}} \left(\frac{p_t^*}{P_t}\right)^{-\eta\phi\nu} \tag{2.C.4}$$

$$(P_t/P_{t-1})^{\eta-1} = \frac{1-(1-\alpha)\left(\frac{p_t^*}{P_t}\right)^{1-\eta}}{\alpha} \tag{2.C.5}$$

$$\Delta_t = h(\Delta_{t-1}, P_t/P_{t-1}) \tag{2.C.6}$$

$$\tilde{u}_C(C(Y_t, q_t^u, \xi_t); \xi_t) = \lim_{T\to\infty} E_t^{\mathcal{P}} \left[\tilde{u}_C(C_T; \xi_T) \beta^T \prod_{k=0}^{T-t} \frac{1+i_{t+k}}{P_{t+k+1}/P_{t+k}}\right] \tag{2.C.7}$$

$$q_t^u = \xi_t^d + \beta(1-\delta) E_t^{\mathcal{P}} q_{t+1}^u, \tag{2.C.8}$$

where the initial price level P_{-1} and initial price dispersion Δ_{-1} are given. Equation (2.C.4) insures that wages clear current labor markets. Similarly, by setting $C_t = C(Y_t, q_t^u, \xi_t)$ on the left-hand side of the consumption Euler equation (2.C.7), we impose market clearing for output goods in period t. Similarly, setting q_t^u equal to the value defined in (2.C.8) insures market clearing in the housing market.[58] Firms' subjective expectations about future wages and households' subjectively optimal consumption plans for the future, however, will generally not be consistent with labor market or goods market clearing in the future in all subjectively perceived contingencies, when beliefs deviate from rational ones.

To be able to analyze the policy problem further, it is necessary to be more specific about the beliefs $\mathcal{P}$ entertained by households and firms.

2.D Derivation of equation (2.1)

Recall the definition of q_t^u which implies

$$\log q_t^u = \log q_t + \log \widetilde{u}_c(C_t; \xi_t)$$

Under the considered belief setup in which agents learn about risk-adjusted capital gains, the dynamics of risk-adjusted capital gains and beliefs are independent of

[58] This holds as long as $D^{\max}$ is chosen sufficiently large, such that it never binds along the equilibrium path.

monetary policy. The response of $\log q_t^u$ to a unexpected change in the path of nominal rates $\boldsymbol{i}$ is thus $\frac{d\log q_t^u}{d\boldsymbol{i}} = 0$, so that

$$\begin{aligned}\frac{d\log q_t}{d\boldsymbol{i}} &= -\frac{d\log \widetilde{u}_c(C_t;\xi_t)}{d\boldsymbol{i}} \\ &= -\frac{d\log \widetilde{u}_c(C_t;\xi_t)}{d\log C_t}\frac{d\log C_t}{d\boldsymbol{i}} \\ &= -\frac{\widetilde{u}_{cc}(C_t;\xi_t)C_t}{\widetilde{u}_c(C_t;\xi_t)}\frac{d\log C_t}{d\boldsymbol{i}} \\ &= \frac{1}{\widetilde{\sigma}}\frac{d\log C_t}{d\boldsymbol{i}} \end{aligned} \tag{2.D.1}$$

The optimal housing supply equation (2.12) can be written as

$$\log k_t = \frac{1}{1-\tilde{\alpha}}\left(\log A_t^d + \log q_t\right).$$

Taking derivatives with respect to $\boldsymbol{i}$ in the previous equation and using (2.D.1) delivers (2.1).

2.E Assumptions about long-run beliefs

To insure that the subjectively optimal consumption plans satisfy the transversality condition (2.16), we impose that equation (2.3) describes subjective housing price beliefs for an arbitrarily long but finite amount of time $t < \bar{T} < \infty$ and that households hold rational expectations in the long-run, i.e. for all periods $t \geq \overline{T}$. Agents thus perceive

$$q_t^u = q_t^{u*} \quad \text{for all } t \geq \bar{T}, \mathcal{P} \text{ almost surely,}$$

where $q_t^{u*} = \bar{\xi}_t^d \equiv \sum_{T=t}^{\infty} E_t[(1-\delta)^{T-t}\,\beta^{T-t}\xi_T^d]$ is the rational expectations housing price. Appendix 2.H.3 shows that this assumption is sufficient to insure that the transversality condition is satisfied. The transversality condition may also hold under weaker conditions, but actually showing this turns out to be difficult. The fact that agents will eventually hold rational housing and rental price expectations could be interpreted as agents learning to make rational predictions in the long-run.

2.F Quadratic approximation of the policy problem

This appendix derives the linear-quadratic approximation to the nonlinear policy problem in Appendix 2.C.

2.F.1 Optimal dynamics and the housing price gap

It will be convenient to determine the welfare-maximizing level of output and the welfare-maximizing housing price under flexible prices, so as to express output and housing prices in terms of gaps relative to these maximizing values. We thus define (Y_t^*, q_t^{u*}) as the values (Y_t, q_t^u) that maximize $U(Y_t, 1, q_t^u; \xi_t)$, which are implicitly defined by[59]

$$U_Y(Y_t^*, 1, q_t^{u*}; \xi_t) = U_{q^u}(Y_t^*, 1, q_t^{u*}; \xi_t) = 0.$$

In particular, we have

$$q_t^{u*} = \bar{\xi}_t^d, \tag{2.F.1}$$

as shown in Appendix 2.H.4. We have

$$\widehat{q}_t^{u,RE} = \widehat{q}_t^{u*}, \tag{2.F.2}$$

which shows that housing price fluctuations are indeed efficient under RE.

The output gap is defined as

$$y_t^{gap} \equiv \log(Y_t) - \log(Y_t^*) = \widehat{y}_t - \widehat{y}_t^*, \tag{2.F.3}$$

i.e. the log-difference of output from its dynamically optimal value.

Under subjective beliefs, it follows from equations (2.12) and the linearization of (2.F.1) (see Appendix 2.H.1 below) that

$$\widehat{q}_t^{u,\mathcal{P}} - \widehat{q}_t^{u*} = \left(\frac{1-\beta(1-\delta)}{1-\beta(1-\delta)\beta_t} - \frac{1-\beta(1-\delta)}{1-\beta(1-\delta)\rho_\xi} \right) \widehat{\xi}_t^d + \frac{\beta(1-\delta)(\beta_t - 1)}{1-\beta(1-\delta)\beta_t}. \tag{2.F.4}$$

Again, for the case where $\beta_t = 1$ and with persistent housing demand shocks ($\rho_\xi \to 1$), the housing price gap under subjective beliefs is equal to the housing price gap under RE. Belief fluctuations, however, now contribute to fluctuations in the housing price gap.

For the real housing price gap, $\widehat{q}_t - \widehat{q}_t^*$, this implies

$$\widehat{q}_t - \widehat{q}_t^* = \left(1 + \widetilde{\sigma}^{-1} C_q\right) \left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right) + \widetilde{\sigma}^{-1} C_Y y_t^{gap}. \tag{2.F.5}$$

2.F.2 Quadratically approximated welfare objective

A second-order approximation to the utility function delivers

$$\frac{1}{2} U_{\widehat{Y}\widehat{Y}} \left(\widehat{y}_t - \widehat{y}_t^*\right)^2 + \frac{1}{2} U_{\widehat{q}^u \widehat{q}^u} \left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right)^2 + \frac{1}{2} \underline{\gamma}^* h_{22} \pi_t^2 + t.i.p.,$$

where $t.i.p.$ denotes terms independent of policy and $\underline{\gamma}^*$ is the Lagrange multiplier associated with equation (2.C.6) at the optimal steady state. See Appendix 2.H.5

[59] The optimal path for $\{Y_t^*, q_t^{u*}\}$ can then be used to determine optimal dynamics for the remaining variables. In particular, equation (2.28) determines C_t^*, equation (2.12) determines k_t^* and thus D_t^*, and equation (2.6) determines H_t^*.

for a detailed derivation. The dependence of the objective function on inflation follows from a second-order approximation of the constraint (2.C.6), which allows expressing the second-order utility losses associated with price distortions Δ_t as a function of squared inflation terms.

Since the fluctuations in the housing price gap, $\widehat{q}_t^u - \widehat{q}_t^{u*}$, are either constant (with RE) or determined independently of policy (under subjective beliefs, see Equation (2.F.4)), the endogenous part of the loss function can be written as

$$\sum_{t=0}^{\infty} \beta^t \frac{1}{2} \left(\Lambda_\pi \pi_t^2 + \Lambda_y \left(y_t^{gap} \right)^2 \right).$$

2.F.3 New Keynesian Phillips curve

Linearizing Equations (2.C.3)-(2.C.5) delivers the linearized Phillips curve. The condition for the equilibrium wage (2.C.4) in period T in industry j in which firms last updated their prices in period t is given by

$$w_T(j) = \tilde{w}_T(j) \left(\frac{p_t^j}{P_t} \right)^{-\eta\phi\nu} \left(\frac{P_T}{P_t} \right)^{\eta\phi\nu},$$

where

$$\tilde{w}_T(j) \equiv \lambda \frac{\bar{H}_T^{-\nu}}{\bar{C}_T^{\tilde{\sigma}^{-1}}} \left(\frac{Y_T}{A_T} \right)^{\phi\nu} C\left(Y_T, q_T^u, \xi_T \right)^{\tilde{\sigma}^{-1}}.$$

Since firms' expectations about $w_T(j)$ and P_T are rational, their expectations about $\tilde{w}_T(j)$ are rational as well. Using the expression for $w_T(j)$, noting that $p_t(i) = p_t^j = p_t^*$, and writing out $Q_{t,T}$, it follows that

$$\left(\frac{p_t^*}{P_t} \right) = \left(\frac{E_t^{\mathcal{P}} \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \frac{\eta}{\eta-1} \phi \bar{C}_T^{\tilde{\sigma}^{-1}} C_T^{-\tilde{\sigma}^{-1}} \tilde{w}_T(j) \left(\frac{Y_T}{A_T} \right)^{\phi} \left(\frac{P_T}{P_t} \right)^{\eta(1+\omega)}}{E_t^{\mathcal{P}} \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \bar{C}_T^{\tilde{\sigma}^{-1}} C_T^{-\tilde{\sigma}^{-1}} (1-\tau_T) Y_T \left(\frac{P_T}{P_t} \right)^{\eta-1}} \right)^{\frac{1}{1+\omega\eta}}. \tag{2.F.6}$$

Log-linearizing equation (2.F.6) delivers[60]

$$\widehat{p}_t^* - \widehat{P}_t = \frac{1-\alpha\beta}{1+\omega\eta} \left\{ \widehat{\tilde{w}}_t(j) + \phi \left(\widehat{y}_t - \widehat{A}_t \right) - \widehat{\tau}_t - \widehat{y}_t + \alpha\beta E_t^{\mathcal{P}} \left[\frac{1+\omega\eta}{1-\alpha\beta} \left(\widehat{p}_{t+1}^* - \widehat{P}_{t+1} + \pi_{t+1} \right) \right] \right\}. \tag{2.F.7}$$

[60]This follows from the the fact that in steady state, we have $p^* = P$, so that

$$\frac{\eta}{\eta-1} \phi \bar{C}^{\tilde{\sigma}^{-1}} C^{-\tilde{\sigma}^{-1}} \tilde{w}(j) \left(\frac{Y}{A} \right)^{\phi} = \bar{C}^{\tilde{\sigma}^{-1}} C^{-\tilde{\sigma}^{-1}} (1-\tau) Y.$$

The steady state value of the numerator in (2.F.6) is thus given by $\frac{1}{1-\alpha\beta} \frac{\eta}{\eta-1} \phi \bar{C}^{\tilde{\sigma}^{-1}} C^{-\tilde{\sigma}^{-1}} \tilde{w}(j) \left(\frac{Y}{A} \right)^{\phi}$ and the steady state value of the denominator by $\frac{1}{1-\alpha\beta} \bar{C}^{\tilde{\sigma}^{-1}} C^{-\tilde{\sigma}^{-1}} (1-\tau) Y$.

As the expectation in (2.F.7) is only about variables about which the private agents hold rational expectations, we can replace $E_t^{\mathcal{P}}[\cdot]$ with $E_t[\cdot]$.[61] Therefore, (2.C.5) can be used in period t and $t+1$, which in its linearized form is given by

$$\widehat{p}_t^* - \widehat{P}_t = \frac{\alpha}{1-\alpha}\pi_t.$$

Substituting $\widehat{\tilde{w}}_t(j)$ by the linearized version of the equilibrium condition (2.C.4) delivers the linearized New Keynesian Phillips Curve:

$$\pi_t = \kappa_y y_t^{gap} + \kappa_q \left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right) + \beta E_t \pi_{t+1} + u_t, \tag{2.F.8}$$

where the coefficients κ are given by

$$\begin{aligned}
\kappa_y &= \frac{1-\alpha}{\alpha}\frac{1-\alpha\beta}{1+\omega\eta}\left(k_y - f_y\right) > 0 \\
\kappa_q &= -\frac{1-\alpha}{\alpha}\frac{1-\alpha\beta}{1+\omega\eta} f_q < 0,
\end{aligned}$$

with $k_y = \partial \log k / \partial \log y$, $f_y = \partial \log f / \partial \log y$, $f_q = \partial \log f / \partial \log q^u$, such that

$$\begin{aligned}
k_y - f_y &= \omega + \tilde{\sigma}^{-1} \frac{\left(1-\underline{g}\right)\underline{Y}}{\underline{C} + \frac{\tilde{\sigma}^{-1}}{1-\tilde{\alpha}}\underline{k}} = \omega + \tilde{\sigma}^{-1} C_Y > 0 \\
f_q &= \tilde{\sigma}^{-1} \frac{\frac{\underline{k}}{1-\tilde{\alpha}}}{\underline{C} + \frac{\tilde{\sigma}^{-1}}{1-\tilde{\alpha}}\underline{k}} = -\tilde{\sigma}^{-1} C_q > 0,
\end{aligned}$$

where $C_q \equiv \frac{q^u}{C}\frac{\partial C}{\partial q^u}$ and $C_Y \equiv \frac{Y}{C}\frac{\partial C}{\partial Y}$, and where the functions

$$f\left(Y, q^u; \xi\right) \equiv \left(1-\tau\right) \bar{C}^{\tilde{\sigma}^{-1}} Y C\left(Y, q^u; \xi\right)^{-\tilde{\sigma}^{-1}}$$

and $k\left(y; \xi\right) \equiv \frac{\eta}{\eta-1}\lambda\phi\frac{\bar{H}^{-\nu}}{A^{1+\omega}}Y^{1+\omega}$ are the same as in Adam and Woodford (2021), for the current period in which markets clear and the internally rational agents observe this.

The cost-push shock u_t is given by

$$u_t = \frac{\left(1-\alpha\right)\left(1-\alpha\beta\right)}{\alpha\left(1+\omega\eta\right)}\left(\Theta + \widehat{\tau}_t - \widehat{g}_t\right),$$

where

$$\begin{aligned}
\widehat{\tau}_t &= -\log\left(\frac{1-\tau_t}{1-\bar{\tau}_t}\right) \\
\widehat{g}_t &= -\log\left(\frac{1-g_t}{1-\bar{g}_t}\right)
\end{aligned}$$

define deviations of τ_t and g_t from their second-best steady state values.

[61] The subjective consumption plans showing up in the stochastic discount factor drop out at this order of approximation.

As in the standard New Keynesian model, a linearization of (2.C.6) implies that the state variable Δ_t is zero to first order under the maintained assumption that initial price dispersion satisfies $\Delta_{-1} \sim O(2)$. This constraint, together with the assumption that the Lagrange multipliers are of order $O(1)$, thus drops out of the quadratic formulation of the optimal policy problem. The second-order approximation of (2.C.6) is, however, important to express the quadratic approximation of utility in terms of inflation.

2.F.4 Linearized IS Equation with potentially non-rational housing price beliefs

We here linearize the constraint (2.C.7). One difficulty with this constraint is that it features the limiting expectations of the subjectively optimal consumption plan on the right hand side. Generally, this would require solving for the subjectively optimal consumption paths, which is generally difficult.

Under our beliefs specifications, housing prices beliefs are rational in the limit. This insures that we do not have to solve for the subjectively optimal consumption plan, instead can derive the IS equation directly in terms of the output gap.

We can now define the natural rate of interest:

Definition 2.F.1 *The natural rate $r_t^{*,RE}$ is the one implied by the consumption Euler equation (2.11) or (2.C.7), rational expectations, and the welfare-maximizing consumption levels under flexible prices $\{C_t^*\}$. It satisfies*

$$\tilde{u}_C(C_t^*;\xi_t) = \beta E_t \left[u_C(C_{t+1}^*;\xi_t)(1 + r_{t+k}^{*,RE}) \right]. \tag{2.F.9}$$

Using the previous definition, we obtain the linearized Euler equation under potentially subjective housing prices beliefs:

Lemma 2 *For the considered belief specifications, the log-linearized household optimality condition (2.C.7) implies for all t*

$$y_t^{gap} = \lim_T E_t y_T^{gap} - E_t \left(\sum_{k=0}^{\infty} \varphi \left(i_{t+k} - \pi_{t+1+k} - r_{t+k}^{*,RE} \right) \right) - \frac{C_q}{C_Y} \left(\widehat{q}_t^u - \widehat{q}_t^{u*} \right), \tag{2.F.10}$$

where $\lim_T E_t y_T^{gap}$ is the (rational) long-run expectation of the output gap, and $\varphi \equiv -\frac{\tilde{u}_c}{\tilde{u}_{cc} C} \frac{1}{C_Y} > 0$. The coefficients $C_q < 0$ and $C_Y > 0$ are the ones defined in the derivation of the linearized Phillips Curve.

The proof can be found in Appendix 2.H.2

2.F.5 Lagrangian formulation of the approximated Ramsey problem

Collecting results from the previous sections, we obtain the following Lagrangian formulation of the Ramsey problem

$$
\begin{aligned}
&\max_{\{\pi_t, y_t^{gap}, i_t \geq \underline{i}\}} \min_{\{\varphi_t, \lambda_t\}} && (2.F.11)\\
&E_0 \sum_{t=0}^{\infty} \beta^t \Bigg\{ -\frac{1}{2} \left(\Lambda_\pi \pi_t^2 + \Lambda_y \left(y_t^{gap} \right)^2 \right) \\
&+ \varphi_t \left[\pi_t - \kappa_y y_t^{gap} - \kappa_q \left(\widehat{q}_t^u - \widehat{q}_t^{u*} \right) - u_t - \beta E_t \pi_{t+1} \right] && (2.F.12)\\
&+ \lambda_t \left[y_t^{gap} - \lim_T E_t y_T^{gap} + \varphi E_t \sum_{k=0}^{\infty} \left(i_{t+k} - \pi_{t+1+k} - r_{t+k}^{*,RE} \right) + \frac{C_q}{C_Y} \left(\widehat{q}_t^u - \widehat{q}_t^{u*} \right) \right] \\
& && (2.F.13)\\
&- \varphi_{-1} \pi_0 - \lambda_{-1} \left(\varphi \pi_0 - y_0^{gap} - \frac{C_q}{C_Y} \left(\widehat{q}_0^u - \widehat{q}_0^{u*} \right) \right) \Bigg\},
\end{aligned}
$$

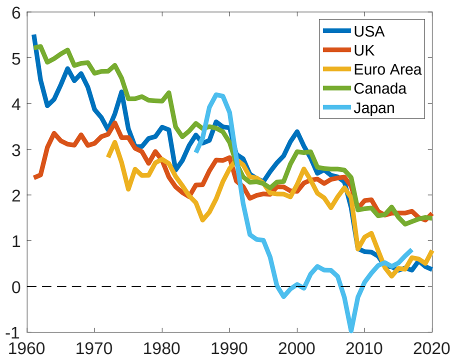
where the process for $(\widehat{q}_t^u - \widehat{q}_t^{u*})$ can be treated as exogenous for the purpose of monetary policy and where the initial Lagrange multipliers $(\varphi_{-1}, \lambda_{-1})$ capture initial pre-commitments. In order to numerically solve the optimal policy problem in (2.F.11), we recursify the problem as proposed in Marcet and Marimon (2019) and solve for the associated value functions and optimal policies. Details of the recursive formulation can be found in Appendix 2.H.6.

2.G The volatility of the PR-ratio and the natural rate

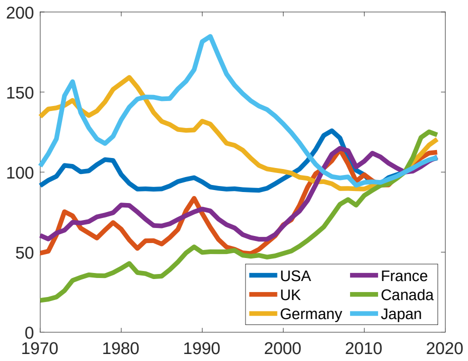
Figure 2.G.1 shows the evolution of natural rates of interest and price-to-rent ratios the U.S., Canada, France, Germany, and the United Kingdom, which we use in Section 2.2. The natural rates are estimated by Holston et al. (2017) and Fujiwara, Iwasaki, Muto, Nishizaki, and Sudo (2016). The price-to-rent ratios are taken from the OECD. We convert the quarterly series of natural rates to annual series by taking arithmetic averages and the quarterly series or PR-ratios to annual series by taking harmonic averages.

Figure 2.G.1: Natural rates and price-to-rent ratios

(a) Natural Rates of Interest



(b) Price-to-rent ratios



Notes: This figure shows the evolution of the natural rate of interest (left panel) and price-to-rent ratios (right panel) for different advanced economies over the period 1961-2020 and 1970-2019, respectively.

2-Housing-Figures 2.3(b) and 2.1 in the Section 2.2 document that the fall in the level of the natural rates of interest across several advanced economies was accompanied by an increase in the volatility of the price-to-rent ratio and in the volatility of natural rates. These trends are consistent with the subjective belief model, outlined in Sections 2.3 and 2.4.

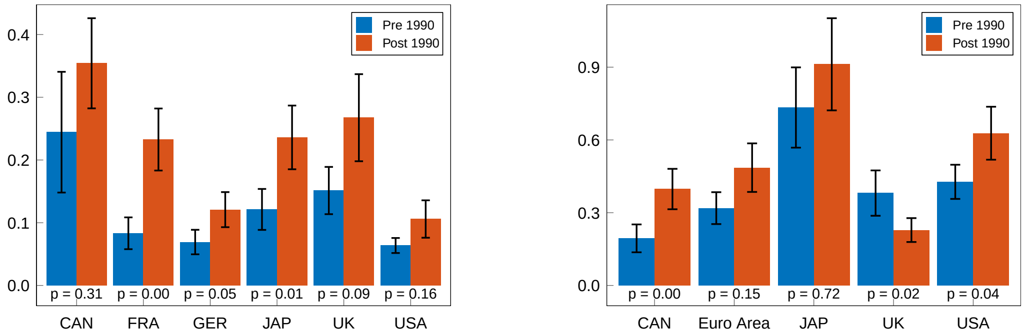
Figure 2.G.2 plots the volatility of the price-to-rent ratio (left panel) and the standard deviation of the natural rate (right panel), respectively before 1990 (blue bars) and after 1990 (red bars), along with 90% confidence bands. The reported volatilities of the price-to-rent ratios are the standard deviations relative to the period-specific mean values, in line with the model. The reported volatilities of the natural rates of interest are the standard deviations of the fluctuations around a linear time trend, in order to isolate high-frequency volatility that can be related to natural rate fluctuations in the model around a fixed steady state value of the natural rate. Figure 2.G.4 shows the volatility price-to-rent ratio using the same linear detrending approach. The p-values below the respective bars are for the null hypothesis of no change in the volatility. The increase in the volatility of the PR-ratio and the natural rate were statistically significant in most of the advanced economies. The evidence is not always statistically significant due to the high

serial correlation of the price-to-rent ratio and the natural rate, which makes it difficult to estimate standard deviations precisely. Figure 2.G.3 shows that the reported volatility increases are not driven by the exact point where we split the data, instead looks often similar for other split points.

Figure 2.G.2: Volatility of the PR-ratio and natural rates pre and post 1990.

(a) Standard Deviation of the Price-to-Rent Ratio Pre and Post 1990

(b) Standard Deviation of Natural Rate Pre and Post 1990

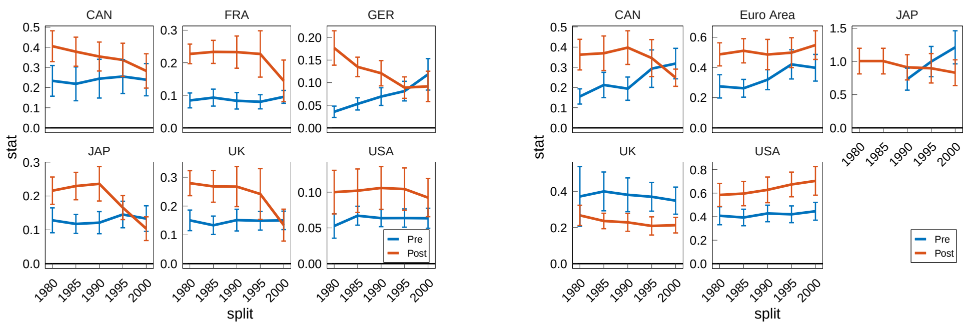


Notes: The black lines denote the 90%-confidence bands. The p-value corresponds to the test whether or not the values changed from pre to post 1990. The reported volatilities of the price-to-rent ratios are the standard deviations relative to the period-specific mean values. The reported volatilities of the natural rates of interest are the standard deviations of the fluctuations around a linear time trend.

Figure 2.G.3: Robustness of housing and natural rate volatility increases with different sample splits

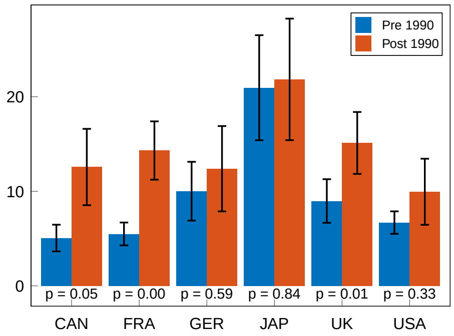
(a) Volatility of the Price-to-Rent ratios

(b) Volatility of natural rates



Notes: Panel (a) shows the standard deviation of the price-to-rent ratio, and panel (b) shows the standard deviation of the natural rate for different advanced economies, computed for varied subsamples. The blue lines show the estimates for the pre-period, and the red lines for the post-period, when the sample is split at the year marked on the horizontal axis. The whiskers denote 90%-confidence bands.

Figure 2.G.4: Standard deviation of the detrended PR-ratio pre and post 1990



Notes: The black lines denote the 90%-confidence bands. The p-value corresponds to the test whether or not the values changed from pre to post 1990.

2.H Proofs

2.H.1 Proofs for Section 2.5

Proof 2.H.1 (Proof of Results in Section 2.5.1) *Result (2.6) follows from iterating forward on (2.14). Log linearizing (2.6), we have*

$$\widehat{q}^u_t = \widehat{\overline{\xi}}^d_t,$$

and log-linearizing (2.8) delivers

$$\widehat{\xi}^d_t = \rho_\xi \widehat{\xi}^d_{t-1} + \varepsilon^d_t.$$

Since the steady-state value of $\overline{\xi}^d$ *is*

$$\overline{\xi}^d = \frac{\underline{\xi}^d}{1-\beta(1-\delta)},$$

the log-linearization of $\overline{\xi}^d_t$ *delivers*

$$\begin{aligned}
\widehat{\overline{\xi}}^d_t &= (1-\beta(1-\delta))\left[\widehat{\xi}^d_t + \beta(1-\delta)E_t\widehat{\xi}^d_{t+1} + ...\right] \\
&= (1-\beta(1-\delta))\left[\widehat{\xi}^d_t + \beta(1-\delta)\rho_\xi\widehat{\xi}^d_t + ...\right] \\
&= (1-\beta(1-\delta))\sum_{T=t}^{\infty}(\beta(1-\delta)\rho_\xi)^{T-t}\widehat{\xi}^d_t \\
&= \widehat{\xi}^d_t\frac{1-\beta(1-\delta)}{1-\beta(1-\delta)\rho_\xi}.
\end{aligned}$$

The results for the price-to rent ration follow by noticing that equation (2.13) implies

$$PR_t \equiv \frac{q_t}{R_t} = \frac{q^u_t}{\xi^d_t}. \tag{2.H.1}$$

Proof 2.H.2 (Proof of Results in Section 2.5.2) *From equation (2.14), which has to hold with equality in equilibrium, and equation (2.4) we get*

$$q_t^{u,\mathcal{P}} = \frac{1}{1-\beta(1-\delta)\beta_t}\xi_t^d$$

The percent deviation of housing prices from the steady state, in which $\beta_t = 1$ *and* $\xi_t^d = \underline{\xi}^d$*, is then given by*

$$\begin{aligned}
\widehat{q}_t^{u,\mathcal{P}} &= \frac{\frac{1}{1-\beta(1-\delta)\beta_t}\xi_t^d - \frac{1}{1-\beta(1-\delta)}\underline{\xi}^d}{\frac{1}{1-\beta(1-\delta)}\underline{\xi}^d} \\
&= \frac{1-\beta(1-\delta)}{1-\beta(1-\delta)\beta_t}\frac{\xi_t^d}{\underline{\xi}^d} - 1 \\
&= \frac{1-\beta(1-\delta)}{1-\beta(1-\delta)\beta_t}\left(1+\widehat{\xi}_t^d\right) - 1 \\
&= \frac{1-\beta(1-\delta)}{1-\beta(1-\delta)\beta_t}\widehat{\xi}_t^d + \frac{\beta(1-\delta)(\beta_t-1)}{1-\beta(1-\delta)\beta_t} \qquad (2.H.2)
\end{aligned}$$

Note, that we can decompose the housing price under subjective beliefs into the housing price under RE and terms that are driven by beliefs:

$$\widehat{q}_t^{u,\mathcal{P}} = \widehat{q}_t^{u,RE} + \frac{\beta(1-\delta)(\beta_t-1)}{1-\beta(1-\delta)\beta_t} + \frac{(1-\beta(1-\delta))(\beta(1-\delta)(\beta_t-\rho_\xi))}{(1-\beta(1-\delta)\beta_t)(1-\beta(1-\delta)\rho_\xi)}\widehat{\xi}_t^d. \quad (2.H.3)$$

Note, that

$$E_t^{\mathcal{P}}\left[q_{t+1}^{u,\mathcal{P}}\right] = \beta_t q_t^{u,\mathcal{P}}.$$

Therefore, a log-linear approximation around the optimal steady state, in which $\underline{\beta} = 1$*, yields*

$$E_t^{\mathcal{P}}\left[\widehat{q}_{t+1}^{u,\mathcal{P}}\right] = \widehat{q}_t^{u,\mathcal{P}} + (\beta_t - 1).$$

From this, we can add and subtract on the right-hand side

$$E_t\left[\widehat{q}_{t+1}^{u,RE}\right] = \rho_\xi\widehat{\xi}_t^d\frac{1-\beta(1-\delta)}{1-\beta(1-\delta)\rho_\xi},$$

which, after plugging in the expression from (2.H.2)*, delivers*

$$\begin{aligned}
E_t^{\mathcal{P}}\left[\widehat{q}_{t+1}^{u,\mathcal{P}}\right] &= E_t\left[\widehat{q}_{t+1}^{u,RE}\right] + (\beta_t-1)\left[1+\frac{\beta(1-\delta)}{1-\beta(1-\delta)\beta_t}\right] \\
&\quad + (1-\beta(1-\delta)\rho_\xi - (1-\beta(1-\delta)\beta_t)\rho_\xi)\frac{(1-\beta(1-\delta))}{(1-\beta(1-\delta)\beta_t)(1-\beta(1-\delta)\rho_\xi)}\widehat{\xi}_t^d.
\end{aligned}$$

In the limit $\rho_\xi \to 1$*, this boils down to*

$$E_t^{\mathcal{P}}\left[\widehat{q}_{t+1}^{u,\mathcal{P}}\right] = E_t\left[\widehat{q}_{t+1}^{u,RE}\right] + (\beta_t-1)\left[1+\frac{\beta(1-\delta)}{1-\beta(1-\delta)\beta_t}\left(1+\widehat{\xi}_t^d\right)\right]$$

Log-linearizing equation (2.H.1)*, which holds true independent of the belief specification, yields*

$$\widehat{PR}_t^{\mathcal{P}} = \widehat{q}_t^{u,\mathcal{P}} - \widehat{\xi}_t^d.$$

Proof 2.H.3 (Proof of Lemma 1) *Under the proposed policy that sets $i_t - E_t\pi_{t+1}$ equal to the natural rate defined in equation (2.15), we have*

$$\begin{aligned}
y_t^{gap} &= \lim_T E_t y_T^{gap} - E_t\left(\sum_{k=0}^{\infty}\varphi\left(i_{t+k} - \pi_{t+1+k} - r_{t+k}^{*,RE}\right)\right) - \frac{C_q}{C_Y}\left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right) \\
&= \lim_T E_t y_T^{gap} - E_t\left(\sum_{k=0}^{\infty}\varphi\left(r_{t+k}^{*,RE} - \frac{1}{\varphi}\frac{C_q}{C_Y}\left(\left(\widehat{q}_{t+k}^u - \widehat{q}_{t+k}^{u*}\right)\right.\right.\right. \\
&\quad \left.\left.\left. - E_{t+k}\left(\widehat{q}_{t+k+1}^u - \widehat{q}_{t+k+1}^{u*}\right)\right) - r_{t+k}^{*,RE}\right)\right) - \frac{C_q}{C_Y}\left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right) \\
&= \lim_T E_t y_T^{gap} + E_t\left(\sum_{k=0}^{\infty}\left(\frac{C_q}{C_Y}\left(\left(\widehat{q}_{t+k}^u - \widehat{q}_{t+k}^{u*}\right) - \left(\widehat{q}_{t+k+1}^u - \widehat{q}_{t+k+1}^{u*}\right)\right)\right)\right) - \frac{C_q}{C_Y}\left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right) \\
&= \lim_T E_t y_T^{gap} + E_t\left(\frac{C_q}{C_Y}\left(\left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right) - \lim_k E_t\left(\widehat{q}_{t+k+1}^u - \widehat{q}_{t+k+1}^{u*}\right)\right)\right) - \frac{C_q}{C_Y}\left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right) \\
&= \lim_T E_t y_T^{gap} + \left(\frac{C_q}{C_Y}\left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right)\right) - \frac{C_q}{C_Y}\left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right) \\
&= \lim_T E_t y_T^{gap},
\end{aligned}$$

which proves that with this policy, the output gap is indeed constant, and $r^{*,\mathcal{P}}$ is the real rate that implies a constant output gap.

2.H.2 Log-linearized Euler equation

Proof 2.H.4 (Proof of Lemma 2) *Log-linearizing equation (2.C.7) around the optimal steady state delivers*

$$\tilde{u}_{CC}C\widehat{c}_t + \tilde{u}_{C\xi}\underline{\xi}\widehat{\xi}_t = E_t^{\mathcal{P}}\sum_{k=0}^{\infty}\tilde{u}_C\left(i_{t+k} - \pi_{t+1+k}\right) + \lim_{T\to\infty}E_t^{\mathcal{P}}\left(\tilde{u}_{CC}C\widehat{c}_T + \tilde{u}_{C\xi}\underline{\xi}\widehat{\xi}_T\right),$$

and log-linearizing (2.F.9) gives

$$\tilde{u}_{CC}C\widehat{c}_t^* + \tilde{u}_{C\xi}\underline{\xi}\widehat{\xi}_t = E_t\sum_{k=0}^{\infty}\tilde{u}_C r_{t+k}^{*,RE} + \lim_{T\to\infty}E_t\left(\tilde{u}_{CC}C\widehat{c}_T^* + \tilde{u}_{C\xi}\underline{\xi}\widehat{\xi}_T\right).$$

Subtracting the previous equation from (2.H.4) delivers

$$\widehat{c}_t - \widehat{c}_t^* = E_t^{\mathcal{P}}\sum_{k=0}^{\infty}\frac{\tilde{u}_C}{\tilde{u}_{CC}C}\left(i_{t+k} - \pi_{t+1+k} - r_{t+k}^{*,RE}\right) + \lim_{T\to\infty}E_t^{\mathcal{P}}\left(\widehat{c}_{T+1} - \widehat{c}_{T+1}^*\right), \quad (2.H.4)$$

where we used $E_t^{\mathcal{P}}\xi_T = E_t\xi_T$ and $E_t^{\mathcal{P}}\widehat{c}_{T+1}^* = E_t\widehat{c}_{T+1}^*$, which hold because agents hold rational expectations about fundamentals.

In all periods in which the subjectively optimal plan is consistent with market clearing in the goods sector, the plan satisfies equation (2.28). Log-linearizing equation (2.28) delivers

$$\widehat{c}_t = C_Y \widehat{y}_t + C_q \widehat{q}^u_t + C_\xi \widehat{\xi}_t, \tag{2.H.5}$$

where $\widehat{\xi}_t$ *is a vector of exogenous disturbances (involving* $A^d_t, \bar{C}_t, g_t$*). Evaluating this equation at the optimal dynamics defines the optimal consumption gap* $\widehat{c}^*_t$*:*

$$\widehat{c}^*_t \equiv C_Y \widehat{y}^*_t + C_q \widehat{q}^{u*}_t + C_\xi \widehat{\xi}_t.$$

Subtracting the previous equation from (2.H.5) delivers

$$\begin{aligned} \widehat{c}_t - \widehat{c}^*_t &= C_Y \left(\widehat{y}_t - \widehat{y}^*_t\right) + C_q \left(\widehat{q}^u_t - \widehat{q}^{u*}_t\right) \\ &= C_Y y^{gap}_t + C_q \left(\widehat{q}^u_t - \widehat{q}^{u*}_t\right) \end{aligned} \tag{2.H.6}$$

Since the current consumption market in period t clears, equation (2.H.6) holds in period t and can be used to substitute the consumption gap on the left-hand side of equation (2.H.4). Similarly, since housing price expectations are rational in the limit, the consumption market also clears in the limit under the subjectively optimal plans, i.e., equation (2.28) holds for $t \geq T'$*. We can thus use equation (2.H.6) also to substitute the consumption gap on the r.h.s. of equation (2.H.4). Using the fact that housing price expectations are rational in the limit (*$\lim_T E^P_t \left(\widehat{q}^u_t - \widehat{q}^{u*}_t\right) = 0$*), we obtain*

$$y^{gap}_t = \lim_T E^{\mathcal{P}}_t y^{gap}_T - E_t \left(\sum_{k=0}^{\infty} \varphi \left(i_{t+k} - \pi_{t+1+k} - r^{*,RE}_{t+k} \right) \right) - \frac{C_q}{C_Y} \left(\widehat{q}^u_t - \widehat{q}^{u*}_t\right).$$

Since we assumed that agents' beliefs about profits and taxes are given by equations (2.24) and (2.25), respectively, evaluated using rational income expectations, the household holds rational expectations about total income. This can be seen by substituting (2.24) and (2.25) into the budget constraint (2.2). We thus have $\lim_T E^{\mathcal{P}}_t y^{gap}_T = \lim_T E_t y^{gap}_T$ *in the previous equation, which delivers (2.F.10).*

2.H.3 Transversality condition subjective housing price beliefs

This appendix shows that under the considered subjective belief specifications, the optimal plans satisfy the transversality constraint (2.16). Since $D_t \in [0, D^{\max}]$ and $E^{\mathcal{P}}_t q^u_T = E_t \bar{\xi}^d_T$ for $T \geq T'$, we have $\lim_{T\to\infty} \beta^T E^{\mathcal{P}}_t \left(D_T q^u_T\right) = 0$. We thus only need to show that $\lim_{T\to\infty} \beta^T E^{\mathcal{P}}_t \frac{\bar{C}_T^{\tilde{\sigma}^{-1}}}{C_T^{\tilde{\sigma}^{-1}}} B_T = 0$. Combining the budget constraint (2.2) with (2.24) and (2.25) we obtain

$$C_t + B_t + \left(D_t - (1-\delta) D_{t-1} - \tilde{d}(k_t; \xi_t) \right) q^u_t \frac{C_t^{\tilde{\sigma}^{-1}}}{\bar{C}_t^{\tilde{\sigma}^{-1}}} + k_t = (1 - g_t) Y_t + B_{t-1}.$$

For $t \geq T'$ the subjectively optimal plans satisfy market clearing in the housing market, i.e.,

$$D_t - (1-\delta)D_{t-1} - \tilde{d}(k_t;\xi_t) = 0$$

so that the budget constraint implies

$$C_t + B_t + k_t = (1-g_t)\,Y_t + B_{t-1}. \tag{2.H.7}$$

Furthermore, for $t \geq T'$ subjectively optimal plans also satisfy market clearing for consumption goods, i.e.,

$$C_t + k_t = (1-g_t)\,Y_t.$$

It thus follows that the subjectively optimal debt level B_t in the budget constraint (2.H.7) is constant under the subjectively optimal plan, after period $t \geq T'$. Furthermore, the expectations about Y_t in the budget constraint (2.H.7) is rational under the assumed lump sum transfer expectations, so that the household's subjective consumption expectations are the same as in a rational expectations equilibrium. (The subjectively optimal investment decisions k_t are driven by rational housing price expectations). Since the limit expectations $\bar{C}_T^{\tilde{\sigma}^{-1}}/C_T^{\tilde{\sigma}^{-1}}$ are bounded in the rational expectations equilibrium, it follows that $\lim_{T\to\infty}\beta^T E_t^{\mathcal{P}}\frac{\bar{C}_T^{\tilde{\sigma}^{-1}}}{C_T^{\tilde{\sigma}^{-1}}}B_T = 0$.

2.H.4 Optimal house price absent price rigidities

The following derivation closely follows Adam and Woodford (2021). We obtain $U_{q^u}\left(Y_t,\Delta_t,q_t^u,\xi_t\right)$ from differentiating equation (2.C.1) with respect to q_t^u and set it equal to 0:

$$\begin{aligned}
U_{q^u}\left(Y_t,\Delta_t,q_t^u,\xi_t\right) = {} & \bar{C}_t^{\tilde{\sigma}^{-1}} C_{q^u}\left(Y_t,q_t^u,\xi_t\right) C\left(Y_t,q_t^u,\xi_t\right)^{-\tilde{\sigma}^{-1}} \\
& + A_t^d \bar{\xi}_t^d \frac{\partial\Omega\left(q_t^u,\xi_t\right)}{\partial q_t^u}\Omega\left(q_t^u,\xi_t\right)^{\tilde{\alpha}-1} C\left(Y_t,q_t^u,\xi_t\right)^{\frac{\tilde{\alpha}}{1-\tilde{\alpha}}\tilde{\sigma}^{-1}} \\
& + \frac{\tilde{\sigma}}{1-\tilde{\alpha}} A_t^d \bar{\xi}_t^d \Omega\left(q_t^u,\xi_t\right)^{\tilde{\alpha}} C\left(Y_t,q_t^u,\xi_t\right)^{\frac{\tilde{\alpha}}{1-\tilde{\alpha}}\tilde{\sigma}^{-1}-1} C_{q^u}\left(Y_t,q_t^u,\xi_t\right) = 0,
\end{aligned}$$

where

$$\frac{\partial\Omega\left(q_t^u,\xi_t\right)}{\partial q_t^u} = \frac{1}{q_t^u}\frac{1}{1-\tilde{\alpha}}\Omega\left(q_t^u,\xi_t\right),$$

and when defining $\chi \equiv \frac{\tilde{\sigma}^{-1}}{1-\tilde{\alpha}} - 1$, we get

$$C_{q^u}(Y_t,q_t^u;\xi_t) \equiv \frac{\partial C(Y_t,q_t^u;\xi_t)}{\partial q^u} = \frac{-\frac{1}{q_t^u}\frac{1}{1-\tilde{\alpha}}\Omega(q_t^u,\xi_t)C(Y_t,q_t^u,\xi_t)^{\chi+1}}{1+(1+\chi)\,\Omega(q_t^u,\xi_t)C(Y_t,q_t^u,\xi_t)^{\chi}}.$$

Taking everything together, we get

$$U_{q^u}\left(Y_t,\Delta_t,q_t^u,\xi_t\right) = \frac{\frac{1}{q_t^u}\frac{1}{1-\tilde{\alpha}}\Omega(q_t^u,\xi_t)C(Y_t,q_t^u,\xi_t)^{\chi+1}}{1+(1+\chi)\,\Omega(q_t^u,\xi_t)C(Y_t,q_t^u,\xi_t)^{\chi}}\bar{C}_t^{\tilde{\sigma}-1}\left(\frac{\bar{\xi}_t^d}{q_t^u}-1\right).$$

In order for U_{q^u} to be zero, we need to have that

$$q_t^{u*} = \bar{\xi}_t^d,$$

as stated in equation (2.F.1).

2.H.5 Quadratically approximated welfare objective

This derivation follows Adam and Woodford (2021). In the optimal steady state, we have $U_Y = U_{q^u} = U_{Yq^u} = 0$, as well as $U_\Delta + \underline{\gamma}\,(\beta h_1 - 1) = 0$. Given the assumption $\Delta_{-1} \sim O(2)$, it follows $\Delta_t \sim O(2)$ for all $t \geq 0$. Additionally, we have $h_2 \equiv \frac{\partial h(\Delta,\Pi)}{\partial \Pi} = 0$ at the optimal steady state. Therefore, a second-order approximation of the contribution of the variables $(Y_t, \Delta_t, q_t^u, \Pi_t, \xi_t)$ to the utility of the household yields

$$\frac{1}{2} U_{\widehat{Y}\widehat{Y}}\left(\widehat{y}_t - \widehat{y}_t^*\right) + \frac{1}{2} U_{\widehat{q}^u\widehat{q}^u}\left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right) + \frac{1}{2}\underline{\gamma}^* h_{22}\pi_t^2 + t.i.p.,$$

where $t.i.p.$ contains all terms independent of policy. Under rational expectations, we have that $\left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right) = 0$ and is thus constant and independent of (monetary) policy. Under subjective beliefs, $\left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right)$ is purely driven by beliefs β_t and housing demand shocks ξ_t^d, see equation (2.F.4), both independent of policy. Therefore, we include $\frac{1}{2} U_{\widehat{q}^u\widehat{q}^u}\left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right)$ in $t.i.p.$.

The term $U_{\widehat{Y}\widehat{Y}}$ is given by $U_{\widehat{Y}\widehat{Y}} \equiv Y\frac{\partial}{\partial Y}\left(U_{\widehat{Y}}\right) \equiv Y\frac{\partial}{\partial Y}\left(YU_Y\right) = \underline{Y}^* U_Y + \left(\underline{Y}^*\right)^2 U_{YY}$. At the optimal steady state, we have

$$\begin{aligned}
\Lambda_\pi &= -\frac{1}{2}\underline{\gamma}^* h_{22} > 0 \\
\Lambda_y &= -\frac{1}{2}\left(\underline{Y}^*\right)^2 U_{YY} > 0,
\end{aligned}$$

where

$$\begin{aligned}
U_{YY} &= -\tilde{\sigma}^{-1}\left(1 - \underline{g}\right)\underline{\bar{C}}^{\tilde{\sigma}^{-1}} C\left(\underline{Y}, \underline{q}^u, \underline{\xi}\right)^{-\tilde{\sigma}^{-1}-1} C_Y \frac{\underline{Y}^*}{C\left(\underline{Y}, \underline{q}^u, \underline{\xi}\right)} \\
&\quad - \frac{\lambda}{1+\nu}\left(1+\omega\right)\omega\frac{\underline{\bar{H}}^{-\nu}}{\underline{A}^{1+\omega}}\underline{Y}^{\omega-1} < 0 \\
h_{22} &= \frac{\alpha\eta\left(1+\omega\right)\left(1+\omega\eta\right)}{1-\alpha} > 0 \\
\underline{\gamma}^* &= \frac{U_\Delta}{1-\alpha\beta} < 0,
\end{aligned}$$

with

$$U_\Delta = -\frac{\underline{Y}^*\left(1-\underline{g}\right)}{1+\omega}\left(\frac{\underline{\bar{C}}^{\tilde{\sigma}^{-1}}}{C\left(\underline{Y}^*, \underline{q}^{u*}, \underline{\xi}\right)}\right)^{\tilde{\sigma}^{-1}} < 0.$$

2.H.6 Recursified optimal policy problem with lower bound

We numerically solve the quadratically approximated optimal policy problem with forward-looking constraints (2.F.11). While it would be preferable to solve the fully nonlinear Ramsey problem, as spelled out in Appendix 2.C, this is computationally not feasible with sufficient degree of numerical accuracy because the problem features 9 state variables and an occasionally binding constraint. The quadratically approximated problem features 2 state variables less because price dispersion Δ_t is to first order independent of policy and because the Phillips curve reduces from a system involving two forward-looking infinite sums, see equation (2.F.6), to a system involving only a single infinite sum, see (2.F.7).

Eggertsson and Singh (2020) compare the exact solution of the standard New Keynesian model with lower bound to the solution of the linear-quadratic approximation with lower bound and show that the quantitative deviations are modest, even for extreme shocks of the size capturing the 2008 recession in the U.S..

To obtain a recursive problem, we apply the approach of Marcet and Marimon (2019) to the problem with forward-looking constraints (2.F.11). We thereby assume that the Lagrangian defined by problem (2.F.11) satisfies the usual duality properties that allow interchanging the order of maximization and minimization, which we verify ex-post using the computed value function. We set the terminal value function for $t = T'$ to its RE value function $W^{RE}(\cdot)$. For $t \leq T'$ we have a value function $W_t(\cdot)$ satisfying the following recursion:

$$
\begin{aligned}
& W_t(\varphi_{t-1}, \mu_{t-1}, u_t, r_t^{*,RE}, \beta_t, \xi_t^d, q_{t-1}^u) \\
= \; & \max_{\left(\pi_t, y_t^{gap}, i_t \geq \underline{i}\right)} \min_{(\varphi_t, \lambda_t)} -\frac{1}{2}\left(\Lambda_\pi \pi_t^2 + \Lambda_y \left(y_t^{gap}\right)^2\right) \\
& + \left(\varphi_t - \varphi_{t-1}\right)\pi_t - \varphi_t\left(\kappa_y y_t^{gap} + \kappa_q\left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right) + u_t\right) \\
& + \lambda_t\left[y_t^{gap} - \lim_T E_t y_T^{gap} + \varphi\left(i_t - E_t \sum_{k=0}^{\infty} r_{t+k}^{*,RE}\right) + \frac{C_q}{C_Y}\left(\widehat{q}_t^u - \widehat{q}_t^{u*}\right)\right] \\
& + \mu_{t-1}\varphi\left(i_t - \pi_t\right) + \gamma_t\left(i_t - \underline{i}\right) \\
& + \beta E_t\left[W_{t+1}(\varphi_t, \underbrace{\beta^{-1}\left(\lambda_t + \mu_{t-1}\right)}_{=\mu_t}, u_{t+1}, r_{t+1}^{*,RE}, \beta_{t+1}, \xi_{t+1}^d, q_t^u\right]
\end{aligned}
\tag{2.H.8}
$$

where the next period state variables (β_{t+1}, q_t^u) are determined by equations (2.5) and (2.10) and $(\widehat{q}_t^u - \widehat{q}_t^{u*})$ is determined by equation (2.F.4). Here we assume that $r_t^{*,RE}$ follows a Markov process, such that the term $E_t \sum_{k=0}^{\infty} r_{t+k}^{*,RE}$ showing up in the current-period return can be expressed as a function of the current state $r_t^{*,RE}$. The future state variables $(\varphi_t, \mu_t, \beta_{t+1}, q_t^u)$ are predetermined in period t. The expectation about the continuation value is thus only over the exogenous states $(u_{t+1}, r_{t+1}^{*,RE}, \xi_{t+1}^d)$. The endogenous state variable φ_{t-1} is simply the lagged Lagrange multiplier on the New Keynesian Phillips curve with housing. The endogenous state variable μ_{t-1} is given for all $t \geq 0$ by

$$
\mu_t = \beta^{-(t+1)}\left(\lambda_0 + \mu_{-1}\right) + \beta^{-t}\lambda_1 + ... + \beta^{-1}\lambda_t.
$$

The initial values (φ_{-1}, μ_{-1}) are given at time zero and equal to zero in the case of time-zero-optimal monetary policy.

For periods $t < T'$, where T' is the period from which housing price expectations are rational and the lower bound constraint ceases to bind, the value functions depend on time, thereafter they are time-invariant. Likewise for sufficiently large T', the value functions $W_t(\cdot)$ and $W_{t+1}(\cdot)$ will become very similar.

We can numerically solve for the value function $W_t(\cdot)$ by value function iteration, starting with $W_{T'}$ which is the value function associated with the linear-quadratic problem with RE.

2.H.7 Optimal targeting rule

Differentiating (2.H.8) with respect to $\{\pi_t, y_t^{gap}, i_t\}$ yields:

$$\begin{aligned}
\frac{\partial W_t}{\partial \pi_t} &= -\Lambda_\pi \pi_t + (\varphi_t - \varphi_{t-1}) - \mu_{t-1}\varphi = 0 \\
\frac{\partial W_t}{\partial y_t^{gap}} &= -\Lambda_y y_t^{gap} - \varphi_t \kappa_y + \lambda_t = 0 \\
\frac{\partial W_t}{\partial i_t} &= \gamma_t + \lambda_t \varphi + \mu_{t-1}\varphi = 0 \text{ and } \gamma_t \left(i_t - \underline{i}\right) = 0.
\end{aligned}$$

Combining these first-order conditions, we can derive the following targeting rule which characterizes optimal monetary policy

$$\Lambda_\pi \pi_t + \frac{\Lambda_y}{\kappa_y}\left(y_t^{gap} - y_{t-1}^{gap}\right) + \frac{\lambda_{t-1}}{\kappa_y} + \mu_{t-1}\left(\varphi + \frac{1}{\kappa_y}\right) + \frac{\gamma_t}{\varphi\kappa_y} = 0,$$

where γ_t is the Lagrange multiplier associated with the lower bound on interest rates. If the lower bound on the nominal interest rate does not bind in the current period, we have $\gamma_t = 0$. Furthermore, if the lower bound has not been binding up to period t, the IS equation has not posed a constraint for the monetary policymaker. Thus, $\lambda_{t-1} = \lambda_{t-k} = 0$ for all $k = 0, 1, ..., t$. For an initial value of $\mu_{-1} = 0$, it follows that $\mu_{t-1} = 0$. The targeting rule then collapses to

$$\Lambda_\pi \pi_t + \frac{\Lambda_y}{\kappa_y}\left(y_t^{gap} - y_{t-1}^{gap}\right) = 0,$$

which is the same as in Clarida, Galí, and Gertler (1999).

The Lagrange multiplier $\gamma_t \leq 0$ captures the cost of a currently binding lower bound. If $\gamma_t < 0$, the optimal policy requires a compensation in the form of a positive output gap or inflation. The multipliers λ_{t-1} and μ_{t-1} capture promises from past commitments when the lower bound was binding.

Another way to express equation (2.H.9) is to write it as

$$\Lambda_\pi \pi_t + \frac{\Lambda_y}{\kappa_y}\left(y_t^{gap} - y_{t-1}^{gap}\right) + \frac{1}{\varphi\kappa_y}\left[\gamma_t - \frac{1+\beta+\varphi\kappa_y}{\beta}\gamma_{t-1} + \frac{\gamma_{t-2}}{\beta}\right] = 0. \qquad (2.H.9)$$

House prices do not enter the optimal target criterion directly but larger fluctuations in house prices make the lower bound bind more often and for a longer period of time. The optimal policy, thus, requires larger compensations in terms of positive output gaps and inflation. To implement this, the nominal interest rate needs to be kept longer at the lower bound.

2.H.8 Calibration of C_q/C_Y

To calibrate C_q/C_Y, the ratio of the consumption elasticities to housing prices and income, respectively, note that from appendix "Second-Order Conditions for Optimal Allocation" in Adam and Woodford (2021), we have

$$C_{q^u}(Y_t, q_t^u; \xi_t) \equiv \frac{\partial C(Y_t, q_t^u; \xi_t)}{\partial q^u} = \frac{-\frac{1}{q_t^u}\frac{1}{1-\widetilde{\alpha}}\Omega(q_t^u, \xi_t)C(Y_t, q_t^u, \xi_t)^{\chi+1}}{1+(1+\chi)\,\Omega(q_t^u, \xi_t)C(Y_t, q_t^u, \xi_t)^{\chi}}$$

where $\chi \equiv \frac{\tilde{\sigma}^{-1}}{1-\tilde{\alpha}} - 1$. In our formulation, we have defined

$$C_q \equiv \frac{\partial C(Y_t, q_t^u; \xi_t)}{\partial \ln q_t^u} = \frac{\partial C(Y_t, q_t^u; \xi_t)}{\partial q_t^u}\frac{\partial q_t^u}{\partial \ln q_t^u} = C_{q^u}(Y_t, q_t^u; \xi_t)\frac{q_t^u}{C_t}$$

so that we have

$$C_q = -\frac{\frac{1}{1-\widetilde{\alpha}}\Omega(q_t^u, \xi_t)C(Y_t, q_t^u, \xi_t)^{\chi+1}}{C(Y_t, q_t^u, \xi_t) + (1+\chi)\,\Omega(q_t^u, \xi_t)C(Y_t, q_t^u, \xi_t)^{\chi+1}}.$$

From the appendix in Adam and Woodford (2021) we also have

$$C_Y(Y_t, q_t^u, \xi_t) \equiv \frac{\partial C_Y(Y_t, q_t^u, \xi_t)}{\partial Y_t} = \frac{1-g_t}{1+\Omega(q_t^u, \xi_t)\,(1+\chi)\,C(Y_t, q_t^u, \xi_t)^{\chi}}$$

so that in our notation

$$C_Y \equiv \frac{\partial C_Y(Y_t, q_t^u, \xi_t)}{\partial \ln Y_t} = \frac{(1-g_t)\,Y_t}{C(Y_t, q_t^u, \xi_t) + \Omega(q_t^u, \xi_t)\,(1+\chi)\,C(Y_t, q_t^u, \xi_t)^{\chi+1}}.$$

We then have

$$\frac{C_q}{C_Y} = \frac{\frac{-\frac{1}{1-\widetilde{\alpha}}\Omega(q_t^u,\xi_t)C(Y_t,q_t^u,\xi_t)^{\chi+1}}{C(Y_t,q_t^u,\xi_t)+(1+\chi)\Omega(q_t^u,\xi_t)C(Y_t,q_t^u,\xi_t)^{\chi+1}}}{\frac{(1-g_t)Y_t}{C(Y_t,q_t^u,\xi_t)+\Omega(q_t^u,\xi_t)(1+\chi)C(Y_t,q_t^u,\xi_t)^{\chi+1}}} = -\frac{1}{1-\widetilde{\alpha}}\frac{\Omega(q_t^u, \xi_t)C(Y_t, q_t^u, \xi_t)^{\chi+1}}{(1-g_t)\,Y_t}.$$

In the steady state, we have $\overline{Y}(1-\overline{g}) = \overline{C} + \overline{\Omega C}^{\chi+1}$, which says that privately consumed output $\overline{Y}(1-\overline{g})$ is divided up into consumption $\overline{C}$ and resources invested in the housing sector, $\overline{\Omega C}^{1+\chi}$. We thus have that

$$\frac{\overline{\Omega C}^{\chi+1}}{\overline{Y}(1-\overline{g})} = 1 - \frac{\overline{C}}{\overline{Y}(1-\overline{g})} = 1 - \frac{\overline{C}}{\overline{C}+\overline{\Omega C}^{\chi+1}} = 1 - \frac{1}{1+\overline{\Omega C}^{\chi}}.$$

Following Adam and Woodford (2021), we set this to the share of housing investment to total consumption, $\overline{\Omega C}^{\chi}$, equal to 6.3%, so that in steady state we have

$$\frac{C_q}{C_Y} = -\frac{1}{1-\widetilde{\alpha}}\left(1 - \frac{1}{1.063}.\right)$$

Finally, following Adam and Woodford (2021), we set the long-run elasticity of housing supply equal to five, which implies $\widetilde{\alpha} = 0.8$, so that

$$\frac{C_q}{C_Y} = -5\left(1 - \frac{1}{1.063}\right) \approx -0.29633.$$

From this, it follows that

$$C_Y = \frac{(1-g)Y}{C + (1+\chi)\,\Omega C^{\chi+1}} = \frac{C+k}{C + \frac{\widetilde{\sigma}-1}{1-\tilde{\alpha}}k} = \frac{1+\frac{k}{C}}{1+\frac{\widetilde{\sigma}-1}{1-\tilde{\alpha}}\frac{k}{C}} = \frac{1+0.063}{1+5\cdot 0.063} = 0.80836$$

and $C_q = -0.29633 \cdot 0.80836 = -0.23954$.

Chapter 3

Corporate Debt Maturity Matters For Monetary Policy*

We provide novel empirical evidence that firms' investment is more responsive to monetary policy when a higher fraction of their debt matures. In a heterogeneous firm New Keynesian model with financial frictions and endogenous debt maturity, two channels explain this finding: (i) Firms with more maturing debt have larger roll-over needs and are therefore more exposed to fluctuations in the real interest rate (*roll-over risk*); (ii) these firms also have higher default risk and therefore react more strongly to changes in the real burden of outstanding nominal debt (*debt overhang*). In comparison to existing models, we show that a model which accounts for the maturity of debt and its distribution across firms implies larger aggregate effects of monetary policy.

3.1 Introduction

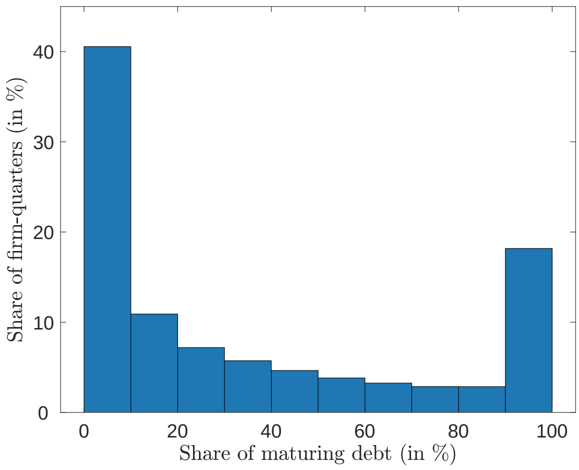
Debt is the main source of external firm financing and plays a key role for investment. But not all debt is created equal. While a part of debt comes due in the short-run, a large share is issued with long maturities and need not be repaid until years in the future. Figure 3.1 shows the distribution of debt maturity across listed U.S. firms. While for many firms only a small fraction of debt matures within the next year, in almost a fifth of firm-quarters this fraction amounts to ninety percent or more. In this paper, we show that this heterogeneity matters for the real effects of monetary policy.

We begin by providing novel empirical evidence that firms respond more strongly to monetary policy shocks when a higher fraction of their debt matures. This result holds both across and within firms and is robust to a wide set of controls and specifications. After a tightening of monetary policy, investment, borrowing, sales, and employment all fall by more for firms with high shares of maturing debt.

To understand the macroeconomic implications of this result, we develop a heterogeneous firm New Keynesian model with financial frictions and endogenous

*Joint work with Joachim Jungherr, Matthias Meier, and Immo Schott.

Figure 3.1: Share of debt maturing within the next year



Notes: The figure shows the distribution of the share of debt which matures within the next twelve months across all firm-quarters of listed U.S. non-financial firms for 1995Q1–2017Q4 from Compustat.

debt maturity. Debt maturity matters for monetary policy because of *roll-over risk* and *debt overhang*. Roll-over needs make firms with higher shares of maturing debt more sensitive to changes in interest rates. Long-term debt insures firms against roll-over risk but creates debt overhang. When tighter monetary policy increases the real burden of outstanding nominal long-term debt, this leads to higher default risk and lower investment.

The model generates the rich heterogeneity in firm financing choices found in the data, including the heterogeneity in debt maturity. Importantly, the model rationalizes the empirical evidence that firms with higher shares of maturing debt respond more strongly to monetary policy shocks. Given this ability to replicate key non-targeted micro moments, we study the model's macroeconomic implications. Compared to existing models, our model implies larger aggregate effects of monetary policy. The maturity of debt and its distribution across firms are key for this result.

In our empirical analysis, we combine balance sheet data of listed U.S. firms with detailed bond-level information about outstanding debt and its maturity. This allows us to construct the precise distribution of bond maturity across firms and time. We complement this data with high-frequency identified monetary policy shocks and estimate their effect on firm-level outcomes using panel local projections. The main result of our empirical analysis is that firms' investment is more responsive to monetary policy if a larger fraction of their debt matures at the time of a shock. This result is statistically and economically significant. After a typical contractionary monetary policy shock, firms with a one-standard deviation higher maturing bond share experience a persistent additional reduction of their capital stock which peaks at 0.2% eight quarters after the shock. Assuming an annual investment-to-capital ratio of 10%, this corresponds to a reduction of investment of 1%. A higher maturing bond share is also associated with similar-sized reductions in debt, sales, and employment. These results are robust to controlling for permanent differences across firms as well as various time-varying firm characteristics such as size, leverage, and liquidity.

To rationalize the empirical evidence and to study the implications for the aggregate effects of monetary policy, we develop a heterogeneous firm New Keynesian model with financial frictions and endogenous debt maturity. In the model, firms finance investment using equity and nominal debt. Debt has a tax advantage relative to equity but introduces the risk of costly default. Firms can choose a mix of short-term and long-term debt. Long-term debt saves roll-over costs but creates a debt overhang problem which increases future default risk.

We calibrate the model to empirical moments which characterize investment and financing choices of listed U.S. firms. Because the effects of debt overhang are more distortive for firms with higher default risk, these firms choose to borrow at shorter maturities. Through this mechanism, the model generates the empirical fact that smaller and younger firms pay higher credit spreads and have higher maturing debt shares.

Importantly, the model explains our main empirical finding: a higher share of maturing debt at the time of a monetary policy shock is associated with a stronger response of firm capital. Both roll-over risk and debt overhang contribute to this result: (1.) Firms with more maturing debt roll over more debt and therefore experience a higher pass-through of interest rate changes to cash flow. This influences firms' need to raise costly outside financing and thereby affects the firm-specific costs of capital. (2.) Firms with more maturing debt have higher default risk and therefore react more strongly to fluctuations in the real burden of outstanding nominal debt. For these firms, both default risk and investment respond more strongly to surprise changes in interest rates and inflation.

The model generates over two thirds of the peak empirical differential capital response associated with the maturing bond share. As in the data, the model produces a hump-shaped response: the initial effect is small and builds up over time. In addition, the model rationalizes the empirical role of the maturing bond share for the firm-level responses of debt, sales, and employment. We show that debt overhang is quantitatively more important in generating these results than roll-over risk.

Finally, we use our model to study the implications of these micro-level results for the macroeconomic effects of monetary policy. To this end, we compare our model to two alternative versions of our model. In the first one, we abstract from cross-sectional differences in debt maturity by assuming that all debt is short-term. This is the standard assumption in many quantitative macro models (e.g., Bernanke, Gertler, and Gilchrist, 1999; Ottonello and Winberry, 2020). In the second alternative economy we allow firms to choose the maturity of their debt, but assume that all firms are ex-ante identical (as in Gomes, Jermann, and Schmid, 2016). Our results show that both long-term debt and heterogeneity amplify the effects of monetary policy shocks on GDP, investment, and inflation. We conclude that the maturity of firm debt and its distribution are important for the aggregate effects of monetary policy.

Related literature. This paper provides an empirical and theoretical analysis of the role of debt maturity for the transmission of monetary policy. It thereby contributes to three related strands of the literature.

First, our work contributes to empirical studies of how debt maturity shapes firms' investment response to aggregate shocks. Duchin, Ozbas, and Sensoy (2010) and Almeida, Campello, Laranjeira, and Weisbenner (2012) show that firms with more maturing debt at the onset of the Financial Crisis of 2007–2008 reduced investment by more.[1] Similarly, higher shares of maturing debt are associated with stronger investment declines during the Great Depression 1929–1933 (Benmelech, Frydman, and Papanikolaou, 2019) and during the 2010–2012 European sovereign debt crisis (Kalemli-Ozcan, Laeven, and Moreno, 2018; Buera and Karmakar, 2021). We complement these event studies of financial crises by providing evidence on how debt maturity shapes the investment response to monetary policy shocks.

A second related group of empirical papers studies the role of firm financing in explaining heterogeneous effects of monetary policy across firms. Important empirical covariates of firms' response to monetary policy shocks are size (Gertler and Gilchrist, 1994), leverage (Anderson and Cesa-Bianchi, 2020; Ottonello and Winberry, 2020), age (Cloyne, Ferreira, Froemel, and Surico, 2020), liquidity (Jeenas, 2019; Greenwald, Krainer, and Paul, 2021), the share of floating-rate debt (Ippolito, Ozdagli, and Perez-Orive, 2018; Gurkaynak, Karasoy Can, and Lee, 2021), and the share of bond financing (Darmouni, Giesecke, and Rodnyansky, 2021). To this literature, we contribute the result that not only the *level* of debt (or leverage) is important, but also the precise *timing* of when this debt comes due.[2]

Third, the theoretical contribution of this paper is to develop a heterogeneous firm New Keynesian model with financial frictions and endogenous debt maturity. Existing quantitative models do not account for differences in debt maturity across firms. Gomes et al. (2016) study the role of nominal long-term debt for monetary policy using a representative firm setup with exogenous debt maturity. Our heterogeneous firm model accounts for the distribution of debt maturity across firms. In a short-term debt model without equity issuance, Ottonello and Winberry (2020) show that firms with low net worth and high leverage react less to monetary policy shocks. In our model the value of firm assets in place is a key determinant of leverage and investment as well. By allowing firms to choose both short-term debt and long-term debt, we study an additional dimension of firm heterogeneity and show its quantitative importance for monetary policy.

Starting with Bernanke et al. (1999), the theoretical literature on the role of financial frictions in generating cross-sectional differences in firm-level responses to aggregate shocks includes important contributions by Cooley and Quadrini

[1] Chodorow-Reich and Falato (2021) highlight the role of covenant violations in determining the effective maturity of bank loans during the 2007–2008 Financial Crisis.

[2] Fabiani, Falasconi, and Heineken (2022) show that monetary policy shocks affect the maturity structure of firms' new borrowing. Deng and Fang (2022) use Compustat data and find that firms with a higher share of long-term debt are less responsive to monetary policy. We show that Compustat data on debt maturity is not precise enough to yield robust and statistically significant results. Detailed bond-level information is crucial for precisely estimating the role of debt maturity for monetary policy.

(2006), Covas and Den Haan (2012), Khan and Thomas (2013), Gilchrist, Sim, and Zakrajšek (2014), Khan, Senga, and Thomas (2016), Begenau and Salomao (2018), Crouzet (2018), Arellano, Bai, and Kehoe (2019), and Arellano, Bai, and Bocola (2020). Because firms issue only one-period debt in these models, all firms have identical exposure to roll-over risk and no significant exposure to debt overhang.[3][4]

The paper is organized as follows. In Section 3.2, we describe the data set, the estimation strategy, and the empirical results. Section 3.3 develops the heterogeneous firm New Keynesian model with financial frictions and endogenous debt maturity. We characterize equilibrium firm behavior in Section 3.4 highlighting the role of roll-over risk and debt overhang for firms' investment response to monetary policy. Section 3.5 presents results from the quantitative model, compares them to the data, and studies the role of debt maturity for the cross-sectional and aggregate effects of monetary policy. Concluding remarks follow.

3.2 Empirical evidence

In this section, we show that firms respond more strongly to monetary policy shocks when a higher fraction of their debt matures.

3.2.1 Data

Our empirical analysis uses detailed bond-level information in combination with firm-level balance sheet data and high-frequency identified monetary policy shocks.

Bond-level data. We obtain comprehensive bond-level information from the Mergent Fixed Income Securities Database (FISD). This database contains key characteristics of publicly-offered U.S. corporate bonds such as their issue date, maturity date, amount issued, principal, and coupon. It also records reductions

[3] Net worth is the only financial state variable in one-period debt models. If firms are allowed to issue long-term debt, the existing stock of previously issued debt enters the firm problem as additional state variable. For quantitative models which explore the implications of long-term debt for firm financing and investment, see also Crouzet (2017), Caggese, Gutierrez, and Pérez-Orive (2019), Perla, Pflueger, and Szkup (2020), Poeschl (2020), Reiter and Zessner-Spitzenberg (2020), Xiang (2020), Gomes and Schmid (2021), Jermann and Xiang (2021), Jungherr and Schott (2021), Karabarbounis and Macnamara (2021), and Jungherr and Schott (2022). None of these models studies the role of debt maturity for monetary policy. Deng and Fang (2022) study exogenous changes in the real interest rate in a partial equilibrium model with debt maturity. For continuous-time approaches to modeling debt maturity in corporate finance, see Admati, Demarzo, Hellwig, and Pfleiderer (2018), Crouzet and Tourre (2021), Dangl and Zechner (2021), or DeMarzo and He (2021). Also related is the sovereign debt literature on risky long-term debt (e.g., Arellano and Ramanarayanan, 2012; Chatterjee and Eyigungor, 2012; Hatchondo, Martinez, and Sosa-Padilla, 2016; Aguiar, Amador, Hopenhayn, and Werning, 2019; Bocola and Dovis, 2019; Aguiar and Amador, 2020).

[4] A related quantitative literature explores the role of household heterogeneity for the transmission of monetary policy (e.g., Gornemann, Kuester, and Nakajima, 2016; Kaplan, Moll, and Violante, 2018; Auclert, 2019; Bayer, Lütticke, Pham-Dao, and Tjaden, 2019; Wong, 2019; Berger, Milbradt, Tourre, and Vavra, 2021; Eichenbaum, Rebelo, and Wong, 2022).

in the amount of outstanding bonds between issuance and maturity, as well as the reason for the reduction, e.g., a call, reorganization, or default. Our empirical analysis focuses on fixed-coupon non-callable bonds, which account for the majority of the value of maturing bonds.[5] Appendix 3.A.1 provides further details on the bond-level data.

Firm-level data. We merge the FISD bond-level information with quarterly firm-level balance sheet data from Compustat. This is not a straightforward task. First, the firm identifiers frequently change over time (e.g., after changes in the company name). Second, the bond debtor may change due to mergers and acquisitions. To map bonds to firms, we use information from CRSP and the Thomson Reuters M&A database. Appendix 3.A.2 provides further details.

We exclude firms in the public administration, finance, insurance, real estate, and utilities sectors. We further exclude firm-quarters in which no bond is outstanding or maturing. This means that we are focusing on the subset of listed U.S. firms which issue corporate bonds. Even though this is a relatively small subset of firms, it contains the largest U.S. companies. Bond-issuing Compustat firms account for 66% of total sales in Compustat and 67% of total fixed assets.

A key variable in our empirical analysis is the maturing bond share

$$\mathcal{M}_{it} = \frac{(\text{maturing bonds})_{it}}{\text{debt}_{it-1}} \times 100, \tag{3.1}$$

where $(\text{maturing bonds})_{it}$ is the value of bonds of firm i that mature in quarter t, and debt_{it-1} is the average total debt of firm i over the preceding four quarters from $t-1$ to $t-4$.[6]

Monetary policy shocks. We use high-frequency changes in the prices of federal funds futures around FOMC meetings to identify monetary policy shocks. Our baseline shocks are based on the three-months ahead federal funds future within 30-minute event windows, as in Gertler and Karadi (2015). We exclude unscheduled FOMC meetings and conference calls. This helps to mitigate the problem that monetary surprises may convey private central bank information about the state of the economy (Meier and Reinelt, 2020). Following Jarociński and Karadi (2020), we further use sign restrictions to separate information effects from conventional monetary policy shocks. Finally, we aggregate the daily shocks to quarterly frequency. Daily shocks are assigned fully to the current quarter if they occur on the first day of the quarter. If they occur within the quarter, they are partially assigned to the current and subsequent quarter (Gorodnichenko and Weber, 2016). The monetary policy shock series covers 1995Q2 through 2018Q3.[7]

[5] In Section 3.2.4, we discuss separate results for callable and variable-coupon bonds.

[6] We use the backward-looking four-quarter moving average of debt in the denominator to smooth out firm-specific seasonal factors and other transitory fluctuations. See Section 3.2.4 for a sensitivity analysis using alternative denominators for the maturing bond share.

[7] In addition to this baseline shock series, we consider various alternative series in Section 3.2.4.

Table 3.1: Descriptive statistics

	Mean	Sd	Min	Max	Obs
Capital growth *(in log points)*	0.78	3.94	-40.52	72.81	35,533
Maturing bond share $\mathcal{M}_{it}$ *(in % of debt)*	0.19	1.77	0.00	67.18	35,533
Leverage *(debt/assets in %)*	34.01	18.47	0.00	151.49	35,533
Liquidity *(cash/assets in %)*	7.59	8.41	0.00	72.64	35,532
Total assets *(in bln. 2005 US$)*	13.48	26.34	0.03	188.75	35,533
Sales growth *(in log points)*	0.76	17.75	-90.51	95.58	35,478
Average bond maturity *(in years)*	9.02	6.18	0.08	99.83	35,533
Monetary policy shocks *(in basis points)*	-0.52	3.47	-15.27	7.87	94

Notes: This table provides descriptive statistics for bond-issuing firms from 1995Q2 through 2018Q3. For details on the definition of variables, see Appendix 3.A.3.

Descriptive statistics. Table 3.1 reports descriptive statistics of key observables used in our empirical analysis. Our sample consists of 35,533 firm-quarter observations from 1995Q2 through 2018Q3. The primary outcome variable in our analysis is capital. We construct firm-level capital stock series by applying a perpetual inventory method to fixed assets in the balance sheet data.[8] Our empirical analysis emphasizes the role of the maturing bond share $\mathcal{M}_{it}$. Corporate bonds have long maturities with an average remaining time to maturity of 9 years, and they constitute more than 60% of total debt in our sample. The average value of $\mathcal{M}_{it}$ is 0.19% and the standard deviation is 1.77%. For firm-quarters in which bonds mature, the average of $\mathcal{M}_{it}$ is 7.64% and the standard deviation is 8.37%. Table 3.1 also documents the distribution of various firm-level control variables used in our analysis: leverage, liquidity, total assets, sales growth, and average bond maturity. Finally, Table 3.1 documents the distribution of the (baseline) monetary policy shock time series. The mean is approximately zero and the standard deviation 3.47 basis points. A one-standard deviation monetary policy shock leads to a 30 basis point increase in the federal funds rate (Meier and Reinelt, 2020).

3.2.2 Investment response to monetary policy shocks

We use panel local projections to investigate the role of the maturing bond share for firms' investment response to monetary policy shocks.

Baseline local projection. We start with a parsimonious baseline specification. Formally, we estimate

$$\Delta^{h+1} \log k_{it+h} = \beta_0^h \mathcal{M}_{it} + \beta_1^h \mathcal{M}_{it}\varepsilon_t^{\mathrm{mp}} + \beta_2^h \mathcal{M}_{it}\Delta\mathrm{gdp}_{t-1} + \delta_i^h + \delta_{st}^h + \nu_{it+h}^h, \quad (3.2)$$

[8]For details on the perpetual inventory method, see Appendix 3.A.3. Our results are robust to using deflated fixed assets instead of using the perpetual inventory method, see Section 3.2.4.

for $h = 0, \ldots, 16$ quarters. On the left-hand side, k_{it} denotes the real capital stock of firm i in quarter t and $\Delta^{h+1} \log k_{it+h} = \log k_{it+h} - \log k_{it-1}$ is the cumulative capital growth between $t-1$ and $t+h$. On the right-hand side, δ_i^h and δ_{st}^h are firm and sector-quarter fixed effects, $\varepsilon_t^{\mathrm{mp}}$ is the monetary policy shock, and $\Delta\mathrm{gdp}_{t-1}$ is lagged real GDP growth.[9]

Figure 3.1 presents the main empirical result of our paper. In panel (a), we show the estimated β_1^h coefficients. These capture the differential response of capital growth for firms which have a higher maturing bond share $\mathcal{M}_{it}$ at the time of a contractionary monetary policy shock. The figure shows that capital growth falls relatively more for firms that have a larger maturing bond share in the quarter of the shock. The shaded area is a 95% confidence band based on standard errors that are two-way clustered by firms and quarters. The differential response is statistically different from zero at the 5% significance level at horizons between six and eleven quarters after the shock. Given that the average capital growth response is negative, this means that firms with more maturing bonds are more responsive to monetary policy shocks.[10]

The estimated coefficients β_1^h in panel (a) of Figure 3.1 are standardized to reflect the differential response of firms which have a one standard deviation higher $\mathcal{M}_{it}$ at the time of a one standard deviation contractionary monetary policy shock. For instance, the estimate $\beta_1^8 = -0.21$ means an additional 0.21 percentage points reduction of capital growth over eight quarters ($\Delta^{8+1} \log k_{it+8}$). Given an annual investment-capital ratio of 10%, this translates into a reduction of investment by 1% between quarter $t-1$ and quarter $t+8$.[11]

Extended local projection. Debt maturity is endogenous and varies systematically across firms. Even within firms, the time series variation of debt maturity may be related to other firm observables. We next show that the results described above are highly robust to focusing on the within-firm variation in $\mathcal{M}_{it}$ over time, and to including a set of time-varying firm-level control variables. Formally, we estimate the extended specification

$$\begin{aligned}\Delta^{h+1} \log k_{it+h} = {} & \beta_0^h \left(\mathcal{M}_{it} - \overline{\mathcal{M}}_i\right) + \beta_1^h \left(\mathcal{M}_{it} - \overline{\mathcal{M}}_i\right) \varepsilon_t^{\mathrm{mp}} + \beta_2^h \left(\mathcal{M}_{it} - \overline{\mathcal{M}}_i\right) \Delta\mathrm{gdp}_{t-1} \\ & + \Gamma_0^h Z_{it-1} + \Gamma_1^h Z_{it-1} \varepsilon_t^{\mathrm{mp}} + \Gamma_2^h Z_{it-1} \Delta\mathrm{gdp}_{t-1} + \delta_i^h + \delta_{st}^h + \nu_{it+h}^h,\end{aligned} \tag{3.3}$$

where $\mathcal{M}_{it} - \overline{\mathcal{M}}_i$ is the deviation of $\mathcal{M}_{it}$ from its firm-specific average $\overline{\mathcal{M}}_i$, and Z_{it-1} is a vector of control variables. Z_{it-1} includes leverage, liquidity, average

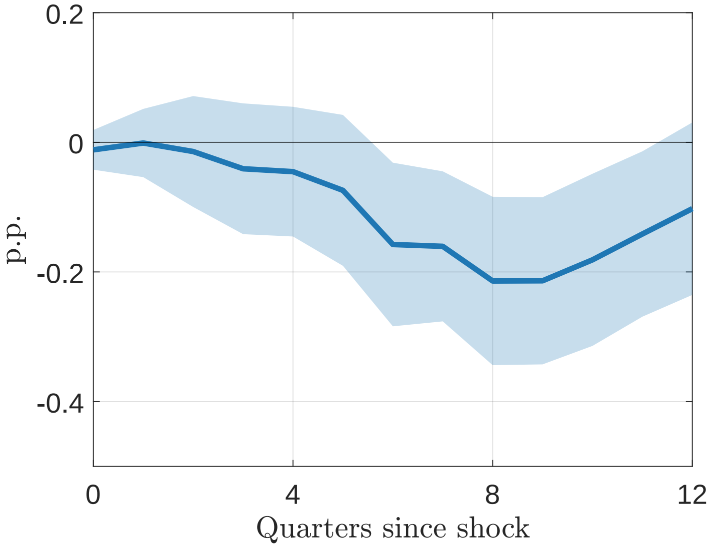
[9] We include the interaction between $\mathcal{M}_{it}$ and $\Delta\mathrm{gdp}_{t-1}$ to control for differences in capital growth cyclicality across firms and time. For our main finding, including this interaction marginally lowers the standard errors of β_1^h but is not important for our conclusions.

[10] The average response of capital growth is shown in Figure 3.B.1 in the Appendix.

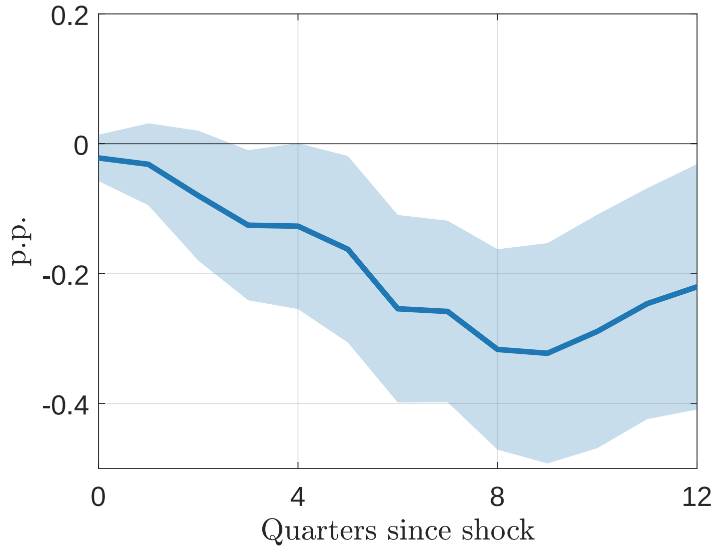
[11] We use the law of motion of capital over a nine-quarter horizon: $K_{t+8} = (1-\delta)K_{t-1} + I_{t+8}$, where δ and I_{t+8} denote depreciation and investment between quarter $t-1$ and $t+8$. If capital growth increases relative to the stationary case ($\overline{I}_{t+8} = \delta K_{t-1}$), this implies an increase of investment by $(I_{t+8} - \overline{I}_{t+8})/\overline{I}_{t+8} = (K_{t+8} - K_{t-1})/\delta K_{t-1}$. Given $(K_{t+8} - K_{t-1})/K_{t-1} = 0.21$ and $\delta = 0.21$ (consistent with 10% annual depreciation), this implies: $(I_{t+8} - \overline{I}_{t+8})/\overline{I}_{t+8} = 1$.

Figure 3.1: Differential investment response associated with higher $\mathcal{M}_{it}$

(a) Baseline specification



(b) Extended specification



Notes: Panel (a) shows the estimated β_1^h coefficients using the baseline specification in equation (3.2). Panel (b) shows the estimated β_1^h coefficients using the extended specification in equation (3.3), where Z_{it-1} includes leverage, liquidity, assets, sales growth, and average maturity (all demeaned). The β_1^h estimates are standardized to capture the differential response (approx. in p.p.) to a one standard deviation increase in $\varepsilon_t^{\mathrm{mp}}$ associated with a one standard deviation higher $\mathcal{M}_{it}$ (in panel (a)) and a one standard deviation higher $(\mathcal{M}_{it} - \overline{\mathcal{M}}_i)$ (in panel (b)). Shaded areas indicate 95% confidence bands two-way clustered by firms and quarters.

maturity of outstanding bonds, real sales growth, and log real total assets (all in deviation from their respective firm-specific averages).

Panel (b) of Figure 3.1 shows the estimated β_1^h coefficients. The estimates conform with the finding in panel (a). The response of capital growth is more negative for firms that have a larger share of maturing bonds relative to the firm-level average share of maturing bonds, and conditional on other control variables. Compared to panel (a), the estimates shown in panel (b) tend to be larger (e.g., $\beta_1^8 = -0.32$) and more precisely estimated.[12]

3.2.3 Response of debt, sales, and other inputs

We next explore whether the share of maturing bonds is important for other firm responses besides investment. Specifically, we estimate the differential responses of firm-level debt, sales, employment, and cost of goods sold using the local projection in equation (3.3).[13] We focus on within-firm variation and use the same controls as in panel (b) of Figure 3.1.[14]

Panel (a) of Figure 4 shows the differential debt response. After a contractionary monetary policy shock, debt grows by less for firms with a larger maturing bond share at the time of the shock. At a two-year horizon, the differential decline in debt growth is 0.40 p.p. This difference is statistically different from zero at significance levels between five and ten percent at horizons between three and eight quarters after the shock. The finding suggests that in periods of tighter monetary policy firms with maturing bonds refinance a smaller fraction of their maturing bonds.

Panel (b) of Figure 3.2 shows that sales growth declines by more for firms with a larger maturing bond share. A caveat here is that the differential sales response is estimated relatively imprecisely. Panels (c) and (d) show the differential responses of employment and cost of goods sold, where the latter measures total expenses for materials, intermediate inputs, labor, and energy. Both employment and cost of goods sold decline by more if $\mathcal{M}_{it}$ is larger at the time of the monetary policy shock. These estimates are statistically different from zero at significance levels between five and ten percent around eight quarters after the shock. Overall, the evidence in Figure 3.2 shows that a high maturing bond share not only shapes the response of capital, but also that of other firm-level outcomes.

3.2.4 Additional results

We conclude the empirical analysis with additional results and robustness exercises.

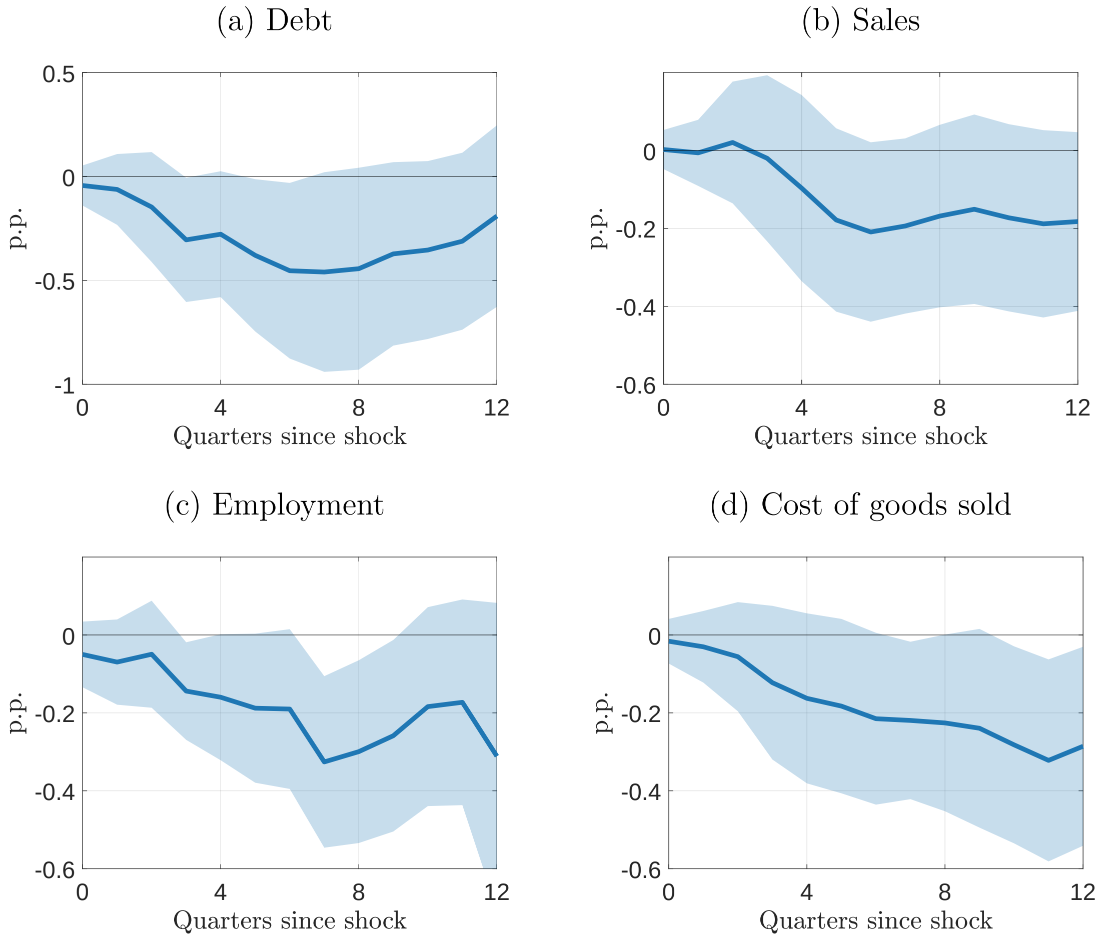
---

[12] For a list of coefficients in the baseline and extended specification, see Appendix 3-DebtMaturity-Tables 3.B.1 and 3.B.2.

[13] Debt, sales, and cost of goods sold are backward-looking four-quarter moving averages to smooth out firm-specific seasonal factors and other transitory fluctuations. Annual employment data is imputed at quarterly frequency using quarterly data on cost of goods sold. For further details, see Appendix 3.A.3.

[14] Figure 3.B.2 in the Appendix provides the corresponding estimates for the baseline specification in (3.2).

Figure 3.2: Differential response of other variables associated with higher $\mathcal{M}_{it}$



Notes: The figure shows the estimated β_1^h coefficients using the extended specification in equation (3.3), but where the left-hand side is $\Delta^{h+1}\log(\text{debt})_{it+h}$ in panel (a), $\Delta^{h+1}\log(\text{sales})_{it+h}$ in panel (b), $\Delta^{h+1}\log(\text{employment})_{it+h}$ in panel (c), and $\Delta^{h+1}\log(\text{cost of goods sold})_{it+h}$ in panel (d). In all panels, Z_{it-1} includes leverage, liquidity, assets, sales growth, and average maturity (all demeaned). The β_1^h estimates are standardized to capture the differential response (approx. in p.p.) to a one standard deviation increase in $\varepsilon_t^{\text{mp}}$ associated with a one standard deviation higher $(\mathcal{M}_{it} - \overline{\mathcal{M}}_i)$. Shaded areas indicate 95% confidence bands two-way clustered by firms and quarters.

Timing of maturity. Our empirical analysis uses detailed FISD bond-level information which allows us to measure the amount of maturing bonds in a given quarter. Figure 3.B.3 in the Appendix shows the importance of measuring the precise timing of maturity relative to monetary policy shocks. In a quasi-Placebo exercise, we replace $\mathcal{M}_{it}$ with $\mathcal{M}_{it-1}$, the maturing bond share in the quarter preceding the monetary policy shock. In the baseline specification, the differential investment response associated with $\mathcal{M}_{it-1}$ is small and insignificant. In the extended specification with additional control variables, the differential response even turns positive several quarters after the shock. These findings underline the importance of using precise information on the timing of maturity.

Maturing debt share in Compustat. In contrast to FISD data, Compustat only provides information on maturing debt within a twelve-month window and does not distinguish between bonds and bank loans. In Figure 3.B.4, we replicate our empirical analysis using Compustat maturity data. Let $\widetilde{\mathcal{M}}_{it}$ denote the Compustat share of total maturing debt within the next twelve months. We show that the differential investment response associated with $\widetilde{\mathcal{M}}_{it}$ (instead of $\mathcal{M}_{it}$) is very imprecisely estimated. This shows the benefit of using FISD data to precisely measure bond maturity.

Callable and variable-coupon bonds. Our main results are based on the maturity of non-callable fixed coupon bonds. A concern with callable bonds is that firm-level conditions which determine the decision to call a bond before maturity may affect the estimates associated to $\mathcal{M}_{it}$. Panel (a) of Figure 3.B.5 shows the β_1^h estimates when constructing $\mathcal{M}_{it}$ using only the maturing amount of callable bonds. The estimated coefficients are insignificant. When combining the maturing amount of callable and non-callable bonds, the estimates are close to our baseline results, see panel (b). We also consider variable-coupon bonds. Panel (c) of Figure 3.B.5 shows the estimated β_1^h coefficients when constructing $\mathcal{M}_{it}$ using only the maturing amount of variable-coupon bonds. The estimates are insignificant. A potential reason for this result is the relatively low number and value of variable-coupon bonds. We observe four times more fixed-coupon bonds than variable-coupon bonds in our sample. When combining the maturing amount of fixed-coupon and variable-coupon bonds, the estimates are close to our baseline results, see panel (d).

Denominators in $\mathcal{M}_{it}$. Equation (3.1) defines the maturing bond share $\mathcal{M}_{it}$ as the ratio of maturing bonds over the backward-looking four-quarter average of total debt. We consider three alternative measures, for which we replace total debt in the denominator with capital, sales, or assets. Panels (a)-(c) of Figure 3.B.6 show the associated β_1^h estimates. In panel (d), we show the β_1^h estimates when using the simple lagged level of debt, capital, sales, and assets, respectively. Our main finding is robust to these alternative definitions of $\mathcal{M}_{it}$.

Unobserved firm characteristics. Our main empirical result is robust to controlling for permanent differences in the maturing bond share across firms and a broad set of other variables. To investigate the potential role of unobserved variables, we follow the approach in Cinelli and Hazlett (2020) which provides a necessary condition for an estimate to be purely spurious in the sense that the true coefficient is zero. If we consider $h = 8$ in Figure 3.1 (b) where $\beta_1^8 = -0.32$, unobserved variables would need to explain at least 36% of the residual variance in capital growth and in the interaction between $\mathcal{M}_{it}$ and $\varepsilon_t^{\text{mp}}$. For comparison, all controls and fixed effects included in specification (3.3) explain 21% of the variance in the interaction between $\mathcal{M}_{it}$ and $\varepsilon_t^{\text{mp}}$. Unobserved variables would thus need to explain more residual variance than all controls and fixed effects included in our extended specification.

Firm age. Recent research has highlighted the role of firm age for understanding the investment response to monetary policy shocks (e.g., Cloyne et al., 2020). Figure 3.B.7 in the Appendix shows that our main finding is not affected by controlling for firm age.

Book value of capital. Our main finding in Figure 3.1 is based on firm-level capital stocks, constructed using a perpetual inventory method. If we instead measure capital by deflated net fixed assets, we obtain similarly significant, but larger, estimates of β_1^h, see Figure 3.B.8.

Monetary policy shocks. Our main findings are robust to a variety of alternative monetary policy shock series. Our baseline shock series is based on changes in the three-months-ahead federal funds future around regular FOMC meetings, sign-restricted following Jarociński and Karadi (2020). Figure 3.B.9 compares these baseline shocks with changes in the 2-quarter, 3-quarter, and 4-quarter ahead eurodollar futures, using either the observed future price changes or the sign-restricted price changes.

Dummy specification. Our baseline specification includes a linear interaction between monetary policy shocks and the maturing bond share. Alternatively, we consider a modification of (3.2), in which monetary policy shocks are interacted with a dummy variable that is one if the maturing bond share is above a certain threshold. As thresholds, we consider 0% and 15%. Figure 3.B.10 shows that this leads to similar conclusions.

Great Recession and ZLB. We study the sensitivity of our results with respect to different time periods by excluding the Great Recession period or the post-Great Recession period, which is largely characterized by a binding effective zero lower bound on monetary policy. Figure 3.B.11 shows the β_1^h estimates when using monetary policy shocks until the height of the Great Recession in 2008Q2, or when excluding 2008Q3–2009Q2 from the sample. The results show that our findings are robust to varying the time sample of our analysis.

3.3 Model

The previous section established empirically that firms' investment response to monetary policy shocks is larger when a higher fraction of their debt matures. To understand the implications of this result for the aggregate effects of monetary policy, we develop a heterogeneous firm New Keynesian model with financial frictions and endogenous debt maturity.

At the heart of the model is a continuum of heterogeneous production firms which produce output using capital and labor. Capital is financed through equity and nominal debt. Debt has a tax advantage relative to equity but introduces the risk of costly default. Firms can choose a mix of short-term debt and long-term debt. Long-term debt saves roll-over costs but generates debt overhang which increases future leverage and default risk.

In addition, the economy consists of retail firms, capital producers, a representative household, and a government. Retail firms buy undifferentiated goods from production firms, turn them into differentiated retail goods and sell them to a final goods sector. Capital producers convert final goods into capital. The representative household works, consumes final goods, and saves by buying equity and debt securities issued by production firms. The government collects a corporate income tax and conducts monetary policy by setting the nominal riskless interest rate.

3.3.1 Production firms

A production firm i enters period t with productivity z_{it} and capital k_{it}. It chooses labor l_{it} to produce an amount y_{it} of undifferentiated output:

$$y_{it} = z_{it} \left(k_{it}^{\psi} l_{it}^{1-\psi} \right)^{\zeta}, \quad \text{with} \quad \zeta, \psi \in (0,1). \tag{3.1}$$

Earnings before interest and taxes are

$$\max_{l_{it}} \quad p_t y_{it} - w_t l_{it} + (\varepsilon_{it} - \delta) Q_t k_{it} - f, \tag{3.2}$$

where p_t is the price of undifferentiated output, w_t is the wage rate, δ is the depreciation rate, Q_t is the price of capital goods, and f is a fixed cost of production. All prices (p_t, w_t, Q_t) are expressed in terms of time t final goods. The firm-specific capital quality shock ε_{it} is i.i.d. with mean zero and continuous probability distribution $\varphi(\varepsilon_{it})$. The shock is realized after production has taken place. An example of a negative capital quality shock is an unforeseen change in technology or consumer demand which reduces the value of existing firm-specific capital.

After the realization of ε_{it}, firms decide whether to pay current debt obligations. There are two types of debt instruments.

Definition 3.1 ***Short-term debt.*** *A short-term bond is a promise to pay one unit of currency in period t together with a nominal coupon c. The quantity of nominal short-term bonds outstanding at the beginning of period t is B_{it}^S.*

Definition 3.2 ***Long-term debt.*** *A long-term bond is a promise to pay a fraction $\gamma \in (0,1)$ of the principal in period t together with a nominal coupon c. In period $t+1$, a fraction $1-\gamma$ of the bond remains outstanding. Firms pay the fraction γ of the remaining principal together with a coupon $(1-\gamma)c$, and so on. The quantity of nominal long-term bonds outstanding at the beginning of period t is B_{it}^L.*

This computationally tractable specification of long-term debt goes back to Leland (1994). Long-term debt payments decay geometrically over time. The maturity parameter γ controls the speed of decay. In the following, we use the real face value of short-term debt and long-term debt (expressed in terms of time $t-1$ final goods): $b_{it}^S \equiv B_{it}^S/P_{t-1}$ and $b_{it}^L \equiv B_{it}^L/P_{t-1}$, where P_{t-1} denotes the price of final goods in period $t-1$.

Firm earnings are taxed at rate τ. Debt coupon payments are tax deductible. After production, taxation, and payment of current debt obligations, the real market value of firm assets is

$$q_{it} = Q_t k_{it} - \frac{b_{it}^S}{\pi_t} - \frac{\gamma b_{it}^L}{\pi_t} + (1-\tau)\left[A_{it}k_{it}^\alpha + (\varepsilon_{it}-\delta)Q_t k_{it} - f - \frac{c(b_{it}^S + b_{it}^L)}{\pi_t}\right], \quad (3.3)$$

where the real face value of nominal short-term and long-term debt depends on (gross) inflation $\pi_t \equiv P_t/P_{t-1}$, and $A_{it}k_{it}^\alpha = \max_{l_{it}}\{p_t y_{it} - w_t l_{it}\}$, with $A_{it} = A(z_{it}, p_t, w_t)$ and $\alpha \in (0,1)$ (see Appendix 3.C.1 for details). The fact that coupon payments are tax deductible lowers total tax payments by the amount $\tau c(b_{it}^S + b_{it}^L)/\pi_t$. This is the benefit of debt. The downside is that firms cannot commit to paying their debt obligations.

Definition 3.3 ***Default.*** *Shareholders are protected by limited liability. They are free to default and hand over a firm's assets to creditors for liquidation. Default is costly. Creditors only recover a fraction $1-\xi$ of firm assets.*

A defaulting firm exits the economy. In addition, there is exogenous exit with probability κ. In this case, the firm repurchases any outstanding long-term debt at market value and pays out all remaining firm assets to shareholders. Continuing firms draw next period's productivity level z_{it+1} from the probability distribution $\Pi(z_{it+1}|z_{it})$.

At the end of period t, next period's capital stock k_{it+1} is financed through retained earnings, outside equity, and by selling new short- and long-term bonds. A firm that sells new short-term bonds of (real) face value b_{it+1}^S at price p_{it}^S raises $b_{it+1}^S p_{it}^S$ on the bond market. Selling new long-term bonds of real value $b_{it+1}^L - (1-\gamma)b_{it}^L/\pi_t$ at price p_{it}^L raises $(b_{it+1}^L - (1-\gamma)b_{it}^L/\pi_t)p_{it}^L$. The market value of next period's capital is accordingly

$$Q_t k_{it+1} = q_{it} + e_{it} + b_{it+1}^S p_{it}^S + \left(b_{it+1}^L - \frac{(1-\gamma)b_{it}^L}{\pi_t}\right)p_{it}^L, \quad (3.4)$$

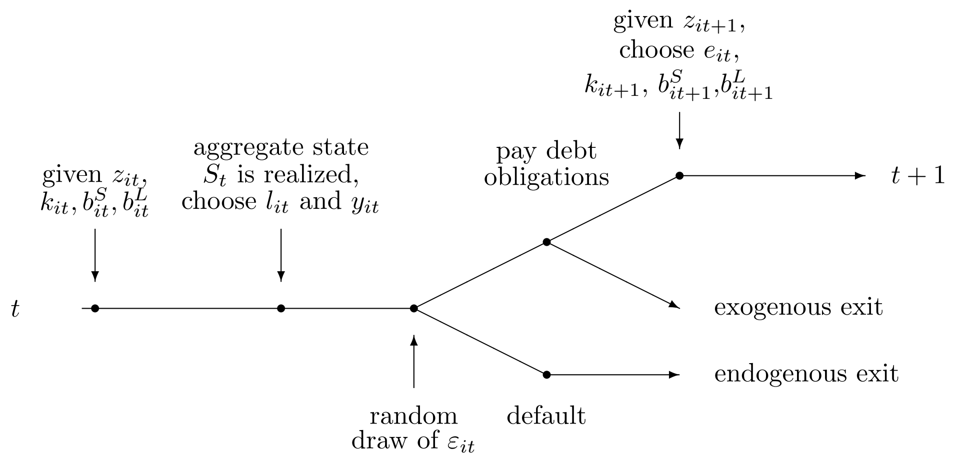
where e_{it} denotes net issuance of outside equity. A negative value of e_{it} indicates a dividend payment from a firm to its shareholders. Whereas dividend payouts are costless, issuing equity and debt is costly.[15]

Definition 3.4 ***Equity issuance cost.*** *Firms pay a quadratic issuance cost whenever they raise outside equity. Net dividend payouts ($e_{it} < 0$) are costless. Equity issuance costs $G(e_{it})$ are given by*

$$G(e_{it}) = \nu \cdot (\max\{e_{it}, 0\})^2. \quad (3.5)$$

[15]Equity and debt issuance costs capture underwriting fees charged by investment banks to firms. Equity issuance costs may also capture costs from adverse selection on the stock market (*cf.* Myers and Majluf, 1984). Altınkılıç and Hansen (2000) provide empirical evidence of increasing marginal issuance costs of equity and debt.

Figure 3.1: Timing



Definition 3.5 ***Debt issuance cost.*** *Firms pay a quadratic issuance cost for selling new short- and long-term debt. Repurchasing outstanding long-term debt (by choosing* $b^L_{it+1} < (1-\gamma)b^L_{it}/\pi_t$*) is costless. Total debt issuance costs* $H(b^S_{it+1}, b^L_{it+1}, b^L_{it}/\pi_t)$ *are therefore*

$$H\left(b^S_{it+1}, b^L_{it+1}, \frac{b^L_{it}}{\pi_t}\right) = \eta \cdot \left(b^S_{it+1} + \max\left\{b^L_{it+1} - \frac{(1-\gamma)b^L_{it}}{\pi_t}, 0\right\}\right)^2. \tag{3.6}$$

Short-term debt needs to be constantly rolled over which implies high issuance costs. Long-term debt matures slowly over time and therefore allows maintaining a given stock of debt at a lower level of bond issuance per period. This saves debt issuance costs.

Value functions. The timing of the firm problem is summarized in Figure 3.1. A firm enters period t with an idiosyncratic state $x_{it} \equiv (z_{it}, k_{it}, b^S_{it}, b^L_{it})$. Given the aggregate state S_t (defined below), it chooses labor demand l_{it} and produces output y_{it}. After the idiosyncratic capital quality shock ε_{it} is realized, the firm decides whether to default. Negative realizations of ε_{it} can generate losses that absent default must be borne by shareholders through lower dividends or higher equity injections. Limited liability creates an upper bound on the losses that shareholders are willing to bear. Let $W_t(x_{it}, \varepsilon_{it}; S_t)$ denote shareholder value conditional on servicing all current debt obligations. Default is optimal if and only if $W_t(x_{it}, \varepsilon_{it}; S_t) < 0$. After the realization of ε_{it}, shareholder value is therefore given by

$$V_t(x_{it}, \varepsilon_{it}; S_t) = \max\left\{0, W_t(x_{it}, \varepsilon_{it}; S_t)\right\}. \tag{3.7}$$

The value of servicing current debt obligations $W_t(x_{it}, \varepsilon_{it}; S_t)$ includes the possibility of exogenous exit:

$$W_t(x_{it}, \varepsilon_{it}; S_t) = (1-\kappa)\mathbb{E}_{z_{it+1}|z_{it}} W_t^C(x_{it}, \varepsilon_{it}, z_{it+1}; S_t) + \kappa \left(q_{it} - \frac{(1-\gamma)b_{it}^L}{\pi_t} \mathbb{E}_{z_{it+1}|z_{it}} p_{it}^L \right) \tag{3.8}$$

With probability κ, a non-defaulting firm exits exogenously. In this case, it repurchases all outstanding long-term debt and pays out remaining firm assets $q_{it} - ((1-\gamma)b_{it}^L/\pi_t)p_{it}^L$ to shareholders. With probability $1-\kappa$, the firm stays active and chooses e_{it}, k_{it+1}, b_{it+1}^S, b_{it+1}^L with associated continuation value $W_t^C(x_{it}, \varepsilon_{it}, z_{it+1}; S_t)$:

$$\begin{aligned} W_t^C(x_{it}, \varepsilon_{it}, z_{it+1}; S_t) = \max_{\substack{e_{it} \geq \underline{e}, k_{it+1}, \\ b_{it+1}^S, b_{it+1}^L}} & -e_{it} - G(e_{it}) - H\left(b_{it+1}^S, b_{it+1}^L, \frac{b_{it}^L}{\pi_t} \right) + \\ & \mathbb{E}_{S_{t+1}|S_t} \Lambda_{t,t+1} \int_{\varepsilon_{it+1}} V_{t+1}(x_{it+1}, \varepsilon_{it+1}; S_{t+1}) \varphi(\varepsilon_{it+1}) d\varepsilon_{it+1} \end{aligned} \tag{3.9}$$

Because all firms are owned by the representative household, firms optimize using the household's stochastic discount factor $\Lambda_{t,t+1}$. In (3.9), equity issuance e_{it} is pinned down through the cash flow constraint (3.4): $e_{it} = Q_t k_{it+1} - q_{it} - b_{it+1}^S p_{it}^S - (b_{it+1}^L - (1-\gamma)b_{it}^L/\pi_t)p_{it}^L$. A firm's choice of e_{it} is bounded from below: $e_{it} \geq \underline{e}$, where $\underline{e} < 0$ sets an upper limit for dividend payments.[16]

3.3.2 Creditors

A firm's choice of capital k_{it+1}, short-term debt b_{it+1}^S, and long-term debt b_{it+1}^L crucially depends on the two bond prices p_{it}^S and p_{it}^L set by creditors. Low bond prices imply high credit spreads which increase a firm's cost of capital. If a firm does not default in period $t+1$, short-term creditors receive a real amount $(1+c)b_{it+1}^S/\pi_{t+1}$, and long-term creditors are paid $(\gamma+c)b_{it+1}^L/\pi_{t+1}$. In case of default, the value of firm assets is

$$\underline{q}_{it+1} \equiv Q_{t+1}k_{it+1} + (1-\tau)\left[p_{t+1}y_{it+1} - w_{t+1}l_{it+1} + (\varepsilon_{it+1} - \delta)Q_{t+1}k_{it+1} - f \right]. \tag{3.10}$$

At this point, creditors liquidate the defaulting firm's assets and receive $(1-\xi)\underline{q}_{it+1}$.

Creditors are perfectly competitive. Because ultimately all debt is held by the representative household, bonds are priced using the stochastic discount factor $\Lambda_{t,t+1}$. Short- and long-term debt have equal seniority. The break-even price of

[16] If the stock of previously issued outstanding debt $(1-\gamma)b_{it}^L/\pi_t$ is sufficiently large, a firm may find it optimal to choose a corner solution and pay out the entire asset value of the firm as dividend: $e_{it} = -q_{it}$. In practice, it is illegal to pay dividends which substantially exceed firm earnings and deplete a firm's stock of capital. We choose the value of the constraint $\underline{e}$ such that it rules out this corner solution but is not binding in equilibrium. The exact value of $\underline{e}$ does not affect equilibrium variables.

nominal short-term debt is therefore

$$p_{it}^S = \mathbb{E}_{S_{t+1}|S_t} \Lambda_{t,t+1} \int_{\varepsilon_{it+1}} \left[(1 - \mathcal{D}_{it+1}) \frac{1+c}{\pi_{t+1}} + \mathcal{D}_{it+1} \frac{(1-\xi)\underline{q}_{it+1}}{b_{it+1}^S + b_{it+1}^L} \right] \varphi(\varepsilon_{it+1}) d\varepsilon_{it+1}, \tag{3.11}$$

where the indicator function $\mathcal{D}_{it+1}$ is one if and only if the firm defaults in period $t+1$, i.e., if $W_{t+1}(x_{it+1}, \varepsilon_{it+1}; S_{t+1}) < 0$. The probability of default in $t+1$ depends on the firm's future state $x_{it+1} = (z_{it+1}, k_{it+1}, b_{it+1}^S, b_{it+1}^L)$. Low values of capital k_{it+1} and high values of short-term debt b_{it+1}^S and long-term debt b_{it+1}^L tend to increase the risk of default. Whereas the price of short-term debt p_{it}^S only depends on the probability distribution of variables in $t+1$, today's price of long-term debt p_{it}^L also depends on the future price of long-term debt:

$$\begin{aligned} p_{it}^L = \mathbb{E}_{S_{t+1}|S_t} \Lambda_{t,t+1} \int_{\varepsilon_{it+1}} \Bigg[& (1 - \mathcal{D}_{it+1}) \frac{\gamma + c + (1-\gamma) \mathbb{E}_{z_{it+2}|z_{it+1}} g_{t+1}(x_{it+1}, \varepsilon_{it+1}, z_{it+2}; S_{t+1})}{\pi_{t+1}} \\ & + \mathcal{D}_{it+1} \frac{(1-\xi)\underline{q}_{it+1}}{b_{it+1}^S + b_{it+1}^L} \Bigg] \varphi(\varepsilon_{it+1}) d\varepsilon_{it+1}. \end{aligned} \tag{3.12}$$

If the firm does not default in period $t+1$, it repays a fraction γ of outstanding long-term debt plus the coupon c. A fraction $1-\gamma$ of debt remains outstanding at price $p_{it+1}^L = g_{t+1}(x_{it+1}, \varepsilon_{it+1}, z_{it+2}; S_{t+1})$. Because this price depends on future firm behavior, it is a function of the future state of the firm.

3.3.3 Retail firms

The remainder of the model setup closely follows Bernanke et al. (1999) and Ottonello and Winberry (2020). Nominal rigidities are introduced through a unit mass of retail firms which buy undifferentiated goods from production firms and sell them as differentiated varieties to the final goods sector. Retail firms are subject to Rotemberg-style quadratic costs of price adjustment. The resulting New Keynesian Phillips Curve is

$$1 - \rho(1 - p_t) - \lambda \pi_t(\pi_t - 1) + \mathbb{E}_{S_{t+1}|S_t} \Lambda_{t,t+1} \lambda \frac{Y_{t+1}}{Y_t} \pi_{t+1}(\pi_{t+1} - 1) = 0, \tag{3.13}$$

where $\rho > 1$ is the elasticity of substitution over differentiated varieties, and λ is a price adjustment cost parameter (see Appendix 3.C.1 for a detailed derivation). Equation (3.13) relates retailers' markup $1/p_t$ to contemporaneous inflation π_t as well as to expected future inflation π_{t+1} and expected real output growth Y_{t+1}/Y_t. After a positive shock to aggregate demand, the relative price of undifferentiated production goods p_t increases and the markup $1/p_t$ falls. Retailers respond by raising prices which increases inflation through (3.13). A higher value of the price adjustment cost parameter λ dampens the contemporary response of inflation.

3.3.4 Capital producers

There is a representative capital good producer who adjusts the aggregate stock of capital using an amount I_t of final goods with decreasing returns (determined by $\phi > 1$):

$$K_{t+1} = \Phi\left(\frac{I_t}{K_t}\right)K_t + (1-\delta)K_t, \quad \text{where} \quad \Phi\left(\frac{I_t}{K_t}\right) = \frac{\delta^{\frac{1}{\phi}}}{1-\frac{1}{\phi}}\left(\frac{I_t}{K_t}\right)^{1-\frac{1}{\phi}} - \frac{\delta}{\phi-1}. \tag{3.14}$$

Profit maximization pins down the price of capital goods:

$$Q_t = \left(\frac{\frac{I_t}{K_t}}{\delta}\right)^{\frac{1}{\phi}} \tag{3.15}$$

3.3.5 Government and monetary policy

The government levies a corporate income tax and pays out the proceeds to the representative household as a lump-sum transfer. In addition, the government conducts monetary policy by setting the nominal riskless interest rate i_t according to the Taylor rule:

$$1 + i_t = \frac{1}{\beta}\pi_t^{\varphi^{\text{mp}}} e^{\eta_t^{\text{mp}}}, \tag{3.16}$$

where $\beta \in (0,1)$ is the representative households' discount rate. The parameter φ^{mp} is the inflation weight of the reaction function, and the stochastic component η_t^{mp} is driven by monetary shocks $\varepsilon_t^{\text{mp}}$ following

$$\eta_t^{\text{mp}} = \rho^{\text{mp}}\eta_{t-1}^{\text{mp}} + \varepsilon_t^{\text{mp}}, \quad \text{with} \quad \varepsilon_t^{\text{mp}} \sim N(0, \sigma_{\text{mp}}^2). \tag{3.17}$$

3.3.6 Households

We close the model by introducing a representative household that owns all equity and debt claims issued by production firms and receives all income in the economy including profits by retail firms and capital producers. Government revenue from taxation is paid out to the household as a lump-sum transfer. The household works and consumes final goods. It saves by buying equity and debt securities issued by production firms.

Future utility is discounted at rate β. We assume additive-separable preferences over consumption C_t and labor L_t. Period utility is

$$\log(C_t) - \frac{L_t^{1+\theta}}{1+\theta}, \quad \text{with} \quad \theta > 0. \tag{3.18}$$

The stochastic discount factor of the representative household is

$$\Lambda_{t,t+1} = \beta\frac{C_t}{C_{t+1}}. \tag{3.19}$$

3.3.7 General equilibrium

A firm maximizes shareholder value (3.9) subject to the firm's cash flow constraint (3.4) and creditors' bond pricing equations (3.11) and (3.12). Because we assume that firms cannot commit to future actions, they must take their own future behavior as given and choose today's policy as a best response. In other words, firms play a game against their future selves. As in Klein, Krusell, and Ríos-Rull (2008), we restrict attention to the Markov perfect equilibrium, i.e., we consider policy rules which are functions of the payoff-relevant state variables. The time-consistent policy is a fixed point in which future firm policies coincide with today's firm policies.

The value function $W_t^C(x_{it}, \varepsilon_{it}, z_{it+1}; S_t)$ can be computed recursively, where W_t^C depends on the firm's idiosyncratic state $x_{it} = (z_{it}, k_{it}, b_{it}^S, b_{it}^L)$, the realization of the firm's capital quality shock ε_{it}, next period's firm productivity z_{it+1}, and the aggregate state S_t. Time subscripts are dropped in the recursive formulation. At the end of each period, the firm chooses a policy vector $\phi(x, \varepsilon, z'; S) = \{e, k', b^{S'}, b^{L'}\}$ which solves

$$W^C(x, \varepsilon, z'; S) = \max_{\phi(x,\varepsilon,z';S)=\left\{\substack{e\geq\underline{e},k', \\ b^{S'},b^{L'}}\right\}} \Bigg\{ -e - G(e) - H\left(b^{S'}, b^{L'}, \frac{b^L}{\pi}\right) + \mathbb{E}_{S'|S}\Lambda \int_{\varepsilon'} V(x', \varepsilon'; S')\varphi(\varepsilon')d\varepsilon' \Bigg\} \tag{3.20}$$

subject to:

$$\begin{aligned}
e &= Qk' - q(x, \varepsilon; S) - b^{S'}p^S - \left(b^{L'} - \frac{(1-\gamma)b^L}{\pi}\right)p^L \\
q(x, \varepsilon; S) &= Qk - \frac{b^S}{\pi} - \frac{\gamma b^L}{\pi} + (1-\tau)\left[Ak^\alpha + (\varepsilon - \delta)Qk - f - \frac{c(b^S + b^L)}{\pi}\right] \\
V(x', \varepsilon'; S') &= \max\Big\{0, W(x', \varepsilon'; S')\Big\} \\
W(x', \varepsilon'; S') &= (1-\kappa)\mathbb{E}_{z''|z'}W^C(x', \varepsilon', z''; S') + \kappa\left(q(x', \varepsilon'; S') - \frac{(1-\gamma)b^{L'}}{\pi'}\mathbb{E}_{z''|z'}p^{L'}\right),
\end{aligned}$$

where bond prices p^S and p^L are determined by (3.11) and (3.12). Given a firm policy $\phi(x, \varepsilon, z'; S) = \{e, k', b^{S'}, b^{L'}\}$, the continuum of production firms is characterized by the distribution $\mu(x)$ with law of motion

$$\mu(x') = \int_x \int_\varepsilon \mathcal{I}(k', b^{S'}, b^{L'}, x, \varepsilon, z'; S)\left[1 - \mathcal{D}(x, \varepsilon; S)\right]\varphi(\varepsilon)d\varepsilon(1-\kappa)\Pi(z'|z)\mu(x)dx + \mathcal{E}(x'; S), \tag{3.21}$$

where the indicator function $\mathcal{I}(k', b^{S'}, b^{L'}, x, \varepsilon, z'; S) = 1$ if $\{k', b^{S'}, b^{L'}\}$ corresponds to the firm's choice $\phi(x, \varepsilon, z'; S) = \{e, k', b^{S'}, b^{L'}\}$. Firms exit the economy endogenously because of default, $\mathcal{D}(x, \varepsilon; S) = 1$, and exogenously at rate κ. The function $\mathcal{E}(x'; S)$ is equal to the mass of entrants starting in state x'. The total mass of firms is always equal to one because in each period the total mass of entrants equals the time-varying mass of exiting firms.

Definition 3.6 *Given the aggregate state $S = (\mu(x), \eta^{mp})$, the equilibrium consists of* (i) *value functions $V(x, \varepsilon; S)$, $W(x, \varepsilon; S)$, and $W^C(x, \varepsilon, z'; S)$,* (ii) *a policy vector $\phi(x, \varepsilon, z'; S) = \{e, k', b^{S'}, b^{L'}\}$,* (iii) *bond price functions p^S and p^L,* (iv) *household consumption C and aggregate labor supply L,* (v) *aggregate prices p, Q, w,* (vi) *a nominal interest rate i, inflation π, a real interest rate r, and a stochastic discount factor Λ, such that:*

1. *Production firms: The value functions $V(x, \varepsilon; S)$, $W(x, \varepsilon; S)$, $W^C(x, \varepsilon, z'; S)$, and policy functions $\phi(x, \varepsilon, z'; S) = \{e, k', b^{S'}, b^{L'}\}$ solve the firm problem* (3.20).

2. *Creditors: p^S and p^L are given by* (3.11) *and* (3.12).

3. *Retail firms: p and π follow the New Keynesian Phillips curve* (3.13).

4. *Capital producers: The price of capital Q is given by* (3.15).

5. *Households: The representative household chooses C and L optimally: $(1 + r)^{-1} = \mathbb{E}_{S'|S}\Lambda$, $(1 + i)^{-1} = \mathbb{E}_{S'|S}\Lambda/\pi'$, and $w = L^{\theta}C$.*

6. *Government: The nominal interest rate i follows the Taylor rule* (3.16).

7. *Firm distribution: $\mu(x') = \Gamma(\mu(x); S)$ as in* (3.21).

8. *Market clearing: The labor market, the final goods market, and the market for capital goods clear (see Appendix 3.C.1 for details).*

3.4 Characterization

In this section, we first describe how production firms choose capital, leverage, and debt maturity. We then explain how firms' investment responses to monetary policy shocks depend on debt maturity.

3.4.1 First-order conditions

The problem of a production firm (3.20) can be expressed in terms of three choice variables: the scale of production k' and the amounts of short-term debt $b^{S'}$ and long-term debt $b^{L'}$. We characterize the equilibrium behavior of firms in terms of the three associated first-order conditions. For simplicity, we discuss these optimality conditions assuming that there is no exogenous exit ($\kappa = 0$). See Appendix 3.C.2 for the general case and detailed derivations.

Capital. The first-order condition with respect to capital k' is

$$\left[1+\frac{\partial G(e)}{\partial e}\right]\left[-Q+b^{S'}\frac{\partial p^S}{\partial k'}+\left(b^{L'}-\frac{(1-\gamma)b^L}{\pi}\right)\frac{\partial p^L}{\partial k'}\right]$$
$$+\mathbb{E}_{S'|S}\Lambda\int_{\varepsilon'}[1-\mathcal{D}(x',\varepsilon';S')]\frac{\partial q(x',\varepsilon';S')}{\partial k'}\mathbb{E}_{z''|z'}\left(1+\frac{\partial G(e')}{\partial e'}\right)\varphi(\varepsilon')d\varepsilon'=0. \quad (3.1)$$

This equation can be decomposed into the costs and benefits of capital. For given choices of $b^{S'}$ and $b^{L'}$, an increase in capital k' must be financed through an equity injection into the firm (see equation 3.4). The marginal cost of capital therefore depends on the price of capital Q and the marginal equity issuance cost $\partial G(e)/\partial e$, shown on the first line of (3.1). The marginal benefit of capital consists of two parts. The first one is direct: capital increases production and raises future assets $q(x',\varepsilon';S')$, as shown on the second line of (3.1). If default is avoided, higher assets reduce the need for future equity issuance or increase future dividends. The second benefit is indirect. If capital reduces default risk, it increases bond prices and bond market revenue, $\partial p^S/\partial k' > 0$ and $\partial p^L/\partial k' > 0$ on the first line of (3.1).

A firm's past choices of debt issuance and debt maturity are important for this indirect benefit of capital. As shown on the first line of (3.1), the benefit is falling in the amount of previously issued long-term debt $(1-\gamma)b^L/\pi$. A higher long-term bond price p^L benefits shareholders only to the extent that it increases the firm's revenue from selling new long-term debt. The fact that a lower default risk also increases the market value of existing long-term debt is not internalized by the firm. In this way, a larger existing stock of debt can reduce firm investment. This is the classic *debt overhang* effect described in Myers (1977).

Short-term debt. The first-order condition for short-term debt $b^{S'}$ is

$$\left[1+\frac{\partial G(e)}{\partial e}\right]\left[p^S+b^{S'}\frac{\partial p^S}{\partial b^{S'}}+\left(b^{L'}-\frac{(1-\gamma)b^L}{\pi}\right)\frac{\partial p^L}{\partial b^{S'}}\right]-\frac{\partial H(b^{S'},b^{L'},\frac{b^L}{\pi})}{\partial b^{S'}}$$
$$+\mathbb{E}_{S'|S}\Lambda\int_{\varepsilon'}[1-\mathcal{D}(x',\varepsilon';S')]\frac{\partial q(x',\varepsilon';S')}{\partial b^{S'}}\mathbb{E}_{z''|z'}\left(1+\frac{\partial G(e')}{\partial e'}\right)\varphi(\varepsilon')d\varepsilon'=0. \quad (3.2)$$

For given choices of k' and $b^{L'}$, selling additional short-term debt is beneficial because it reduces the need for costly equity issuance by $[1 + \partial G(e)/\partial e)] \cdot p^S$. This is shown on the first line of (3.2). The costs of short-term debt consist of debt issuance costs $H(\cdot)$ and higher default risk which reduces bond market revenue, i.e., $\partial p^S/\partial b^{S'} < 0$ and $\partial p^L/\partial b^{S'} < 0$. For each short-term bond sold, the firm promises a payment of $(1 + c)/\pi'$ which reduces future assets, captured by $\partial q(x',\varepsilon';S')/\partial b^{S'} < 0$ on the second line of (3.2). The bond price p^S fully reflects the coupon c promised to creditors, but because it is tax deductible it only reduces $q(x',\varepsilon';S')$ by $(1-\tau)c$. This is the tax benefit of debt.

A larger stock of previously issued long-term debt $(1-\gamma)b^L/\pi$ lowers bond market revenue. As can be seen from the first line of (3.2), this reduces the impact of changes in p^L caused by additional short-term debt $b^{S'}$. The firm disregards the fact that an increase in default risk lowers the market value of existing long-term debt. In this way, debt overhang increases firms' incentive to issue additional debt.[17]

Long-term debt. Finally, the first-order condition with respect to $b^{L'}$ is

$$\left[1+\frac{\partial G(e)}{\partial e}\right]\left[p^L+b^{S'}\frac{\partial p^S}{\partial b^{L'}}+\left(b^{L'}-\frac{(1-\gamma)b^L}{\pi}\right)\frac{\partial p^L}{\partial b^{L'}}\right]-\frac{\partial H(b^{S'},b^{L'},\frac{b^L}{\pi})}{\partial b^{L'}}$$
$$+\mathbb{E}_{S'|S}\Lambda\int_{\varepsilon'}[1-\mathcal{D}(x',\varepsilon';S')]\mathbb{E}_{z''|z'}\left[\left(\frac{\partial q(x',\varepsilon';S')}{\partial b^{L'}}-\frac{1-\gamma}{\pi'}\cdot g(x',\varepsilon',z'';S')\right)\right.$$
$$\left.\left(1+\frac{\partial G(e')}{\partial e'}\right)-\frac{\partial H(b^{S''},b^{L''},\frac{b^{L'}}{\pi'})}{\partial b^{L'}}\right]\varphi(\varepsilon')d\varepsilon'=0. \tag{3.3}$$

Similar to short-term debt, selling additional long-term debt reduces the need for costly equity issuance by $[1+\partial G(e)/\partial e)]\cdot p^L$. At the same time, it increases a firm's default risk and lowers bond market revenue, $\partial p^S/\partial b^{L'}<0$ and $\partial p^L/\partial b^{L'}<0$. In addition, the firm incurs the marginal debt issuance cost $\partial H(b^{S'},b^{L'},b^L/\pi)/\partial b^{L'}>0$. This is shown on the first line of (3.3). Different from short-term debt, a long-term bond only promises a payment of $(\gamma+c)/\pi'$ next period, a fraction γ of the principal plus a coupon. The associated reduction of future assets $q(x',\varepsilon';S')$ on the third line of (3.3) is therefore smaller. However, the fact that a fraction $1-\gamma$ of long-term debt remains outstanding lowers future bond market revenue by $(1-\gamma)/\pi'\cdot g(x',\varepsilon',z'';S')$.

The main benefit of issuing long-term debt is that it reduces future debt issuance costs, shown as $\partial H(b^{S''},b^{L''},b^{L'}/\pi')/\partial b^{L'}<0$ on the third line of (3.3). The downside is that it creates debt overhang. Whereas an increase in $b^{S'}$ affects p^L only through next period's default risk, an increase in $b^{L'}$ also affects p^L through its effect on future choices of capital k'', short-term debt $b^{S''}$, and long-term debt $b^{L''}$. As discussed above, a higher future stock of outstanding long-term debt generates debt overhang which can lead to reduced investment and higher borrowing. This increases future leverage and default risk and thereby has an additional negative effect on today's bond price p^L.

Debt overhang is a commitment problem. When selling long-term debt, shareholders would like to promise low future values of leverage and default risk because this would increase today's bond price p^L. However, this promise is not credible. After long-term debt is sold, the firm continues to internalize the benefits of higher leverage. Yet a part of the associated costs is borne by existing creditors. As creditors have rational expectations, p^L correctly anticipates the effects of debt overhang on future firm behavior. Shareholders therefore face a commitment problem: leverage is higher ex-post than optimal ex-ante (see Jungherr and Schott, 2021).[18]

[17]In the sovereign debt literature (e.g., Hatchondo et al., 2016) this incentive to increase indebtedness at the expense of existing creditors is known as *debt dilution*. In corporate finance, the term *debt dilution* is sometimes used to describe the specific situation that a larger number of creditors must share a given liquidation value of a bankrupt firm. The mechanism described above is at work even if the liquidation value is zero or if existing debt is fully prioritized (as in Bizer and DeMarzo, 1992).

[18]A large literature documents the empirical use and effects of seniority structures, secured

3.4.2 Debt maturity and the investment effects of monetary policy

We now return to the question that lies at the heart of this paper: How does debt maturity affect the firm-level investment response to monetary policy shocks? This section explains that in our model debt maturity matters because of two channels: *roll-over risk* and *debt overhang*. Ceteris paribus, roll-over risk generates larger investment responses for firms which borrow at shorter maturities whereas debt overhang implies the opposite.

Consider a contractionary monetary policy shock which raises the real interest rate. While this increases the cost of capital for all firms, differences in debt maturity generate heterogeneous investment responses. In the model, the maturing bond share is

$$\mathcal{M} = \frac{b^S + \gamma b^L}{b^S + b^L}. \tag{3.4}$$

It measures the share of a firm's total debt that is due in the current period, i.e., short-term debt plus a fraction γ of outstanding long-term debt. For a given amount of total debt, firms with lower $\mathcal{M}$ borrow at longer maturities, roll over less debt per period, and choose next period's capital k' in the presence of a higher amount of outstanding long-term debt $(1-\gamma)b^L/\pi$. Outstanding long-term debt enters the firm problem through equity issuance e, i.e., the cash flow from shareholders to the firm:

$$e = Qk' - q - \underbrace{\left(b^{S'} p^S + \left(b^{L'} - \frac{(1-\gamma)b^L}{\pi}\right) p^L\right)}_{\text{bond market revenue}}. \tag{3.5}$$

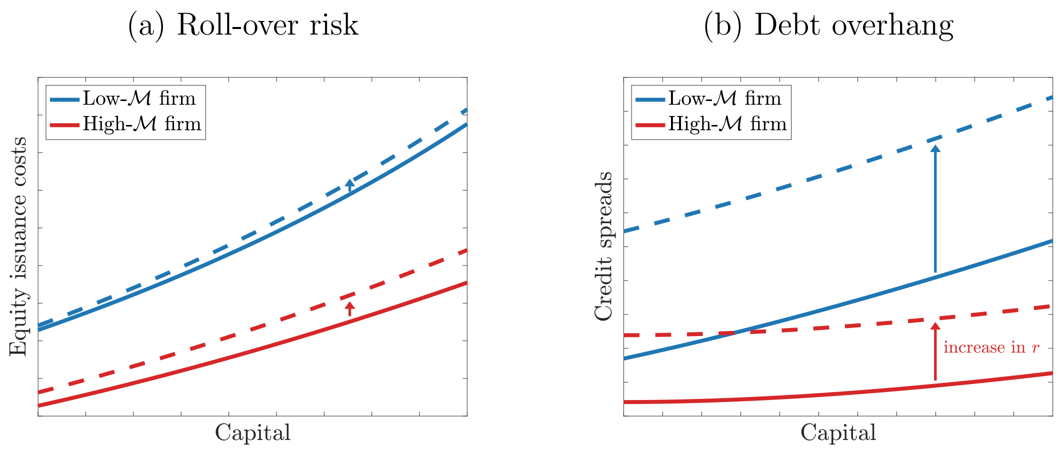
The role of $\mathcal{M}$ in generating heterogeneity in firms' investment responses can be decomposed into two channels, roll-over risk and debt overhang. These channels are illustrated in Figure 3.1.

Roll-over risk. Panel (a) of Figure 3.1 shows that a contractionary monetary policy shock increases equity issuance costs by more for firms with a higher maturing bond share $\mathcal{M}$. Because these firms have higher roll-over needs, they face a larger reduction in bond market revenue. Ceteris paribus, this requires higher equity issuance which increases the cost of capital. Roll-over risk can therefore generate a larger investment reduction for high-$\mathcal{M}$ firms.

More precisely, the figure shows equity issuance costs, $G(e)$, as a function of capital k'. The red solid line plots $G(e)$ for a high-$\mathcal{M}$ firm, the blue line is drawn for a firm with low $\mathcal{M}$. Firm assets q and leverage are identical across firms and held constant as k' increases. Because an increase in capital is partly financed

assets, and debt covenants aimed at mitigating conflicts of interest between existing creditors and shareholders (e.g., Green, 2018; Drechsel, 2019; Greenwald, 2019; Adler, 2020; Benmelech, Kumar, and Rajan, 2020; Ivashina and Vallee, 2020; Chodorow-Reich and Falato, 2021; Lian and Ma, 2021). Empirically, these contracting features are less common for bonds than they are for bank loans, and their usage is increasing with default risk.

Figure 3.1: Debt maturity and the effects of a contractionary monetary policy shock



Note: Panel (a) shows equity issuance costs $G(e)$ as a function of capital k', panel (b) shows firm-specific credit spreads. Credit spreads are a maturity-weighted average over the short-term spread and long-term spread, see Appendix 3.D.1. Solid lines represent the steady state, dashed lines show values after an unexpected increase in the real interest rate r. Blue lines show a firm with a low maturing bond share $\mathcal{M}$ (i.e., high $(1-\gamma)b^L/\pi$ and low $\mathcal{M}' = (b^{S'}+\gamma b^{L'})/(b^{S'}+b^{L'})$), red lines show a high-$\mathcal{M}$ firm. Both firms have identical productivity z and assets q. Leverage $(b^{S'}+b^{L'})/k'$ is identical across firms and held constant as k' increases. For the blue dashed line in panel (b), both leverage and r are increased for the low-$\mathcal{M}$ firm.

through additional equity, equity issuance costs are increasing in capital for both firms.[19]

The dashed lines show the effect of an increase in the real interest rate r. A higher real rate implies a lower stochastic discount factor Λ and lower bond prices p^S and p^L for both firms. Because the high-$\mathcal{M}$ firm rolls over more debt per period, the pass-through of bond price changes to bond market revenue and cash flow is higher. For a given choice of capital, this implies a larger increase in equity issuance. With increasing equity issuance costs, this raises $\partial G(e)/\partial e$ and thereby the marginal cost of capital in first-order condition (3.1). Through this mechanism, a higher $\mathcal{M}$ exposes firms to roll-over risk and generates a larger investment response to changes in the real rate. Long-term debt lowers $\mathcal{M}$ and thereby provides insurance against roll-over risk.

Increasing marginal equity issuance costs are a necessary condition for roll-over risk to have an effect on investment. If equity issuance costs were linear or zero, current cash flow and existing assets q would not appear in firms' first-order condition for capital. Differences in $\mathcal{M}$ would still imply different effects of interest rate changes on cash flow and dividends, but those differences would not affect the marginal cost of capital.

In addition to increasing the real interest rate, a contractionary monetary policy shock also reduces inflation π and thereby increases the real value of outstand-

[19] Equity issuance costs are higher for the low-$\mathcal{M}$ firm. Because this firm has the same amount of assets q but a higher amount of outstanding debt $(1-\gamma)b^L/\pi$, its bond market revenue is lower. To obtain a given amount of capital k' it therefore needs to issue more equity.

ing nominal long-term debt $(1-\gamma)b^L/\pi$. This effect is known as *Fisherian debt deflation*. In our model, this effect further reduces the roll-over needs of low-$\mathcal{M}$ firms and therefore amplifies firms' heterogeneous exposure to roll-over risk.

Debt overhang. Panel (b) of Figure 3.1 shows that a contractionary monetary policy shock leads to a larger increase in credit spreads for firms with a low maturing bond share $\mathcal{M}$. Because these firms have higher amounts of previously issued long-term debt $(1-\gamma)b^L/\pi$, debt overhang generates a larger increase in default risk and credit spreads in response to the shock. This increase in default risk and credit spreads lowers investment. In this way, debt overhang can generate a larger investment response for low-$\mathcal{M}$ firms.

The figure shows credit spreads as a function of capital k'. For the high-$\mathcal{M}$ firm, credit spreads display little variation in k'. This is because the high-$\mathcal{M}$ firm mainly relies on short-term debt whose credit spread only depends on next period's default risk. As leverage is held constant in Figure 3.1, next period's default risk varies very little in k'. Credit spreads increase more rapidly in k' for the low-$\mathcal{M}$ firm. This firm has a higher share of long-term debt whose credit spread also depends on default risk in future periods. Future default risk increases in k' because a higher value of k' implies a higher future stock of outstanding long-term debt $(1-\gamma)b^{L'}/\pi'$. Through debt overhang, this increases future leverage and default risk and thereby already raises the long-term credit spread today.

The dashed lines show the effect of an increase in the real interest rate r. The discounted net present value of future firm earnings falls, while the amount of previously issued long-term debt $(1-\gamma)b^L/\pi$ remains unchanged (or even rises if inflation π falls). Firms' incentive to increase leverage at the expense of existing creditors becomes stronger. This debt overhang effect is larger for the low-$\mathcal{M}$ firm with a higher amount of outstanding long-term debt $(1-\gamma)b^L/\pi$. In panel (b), this is illustrated through a larger relative increase in leverage for the low-$\mathcal{M}$ firm. Its default risk and credit spreads increase by more, which drives up the firm's cost of capital. Through this mechanism, a lower $\mathcal{M}$ exposes firms to debt overhang and generates a bigger investment response to changes in the real rate.[20]

Inflation π falls after a contractionary monetary policy shock. This raises the real burden of outstanding nominal long-term debt $(1-\gamma)b^L/\pi$. Low-$\mathcal{M}$ firms have higher amounts of outstanding long-term debt and are therefore more strongly affected by the increase in the real value of their nominal debt. Through debt overhang, this generates larger increases in default risk and credit spreads. In this way, Fisherian debt deflation amplifies firms' heterogeneous exposure to debt overhang.

3.5 Quantitative analysis

The previous section showed that the role of debt maturity for firms' investment response is theoretically ambiguous. We therefore proceed with a quantitative

[20] The amplification of aggregate shocks through debt overhang is studied in more detail in Gomes et al. (2016) and Jungherr and Schott (2022).

analysis. Our calibrated model replicates several targeted and non-targeted moments that characterize financing choices of U.S. listed firms. The model also rationalizes the empirical result that firms with higher shares of maturing debt react more strongly to monetary policy shocks. At the aggregate level, we show that both long-term debt and heterogeneity amplify the effects of monetary policy.

3.5.1 Solution method

We use value function iteration and interpolation to compute the Markov perfect equilibrium of our model. There are three key challenges. The first is the dimensionality of the state space. The variables (z, k, b^S, b^L) describe the firm's idiosyncratic state at the beginning of the period. Together with S and ε, they determine a firm's default decision. Firms decide about investment and financing at the end of the period after the realization of z'. The state in (3.20) is therefore given by $(z, k, b^S, b^L, \varepsilon, z'; S)$. To solve the model, we exploit the fact that this information can be summarized in the reduced state vector $(q, b, z'; S)$ which includes firm assets $q = q(z, k, b^S, b^L, \varepsilon; S)$ and outstanding long-term debt $b = (1-\gamma)b^L$.

The second difficulty is finding the equilibrium price of risky long-term debt, p^L. Optimal firm behavior depends on p^L, which itself depends on current and future firm behavior. A firm that cannot commit to future actions must take into account how today's choices will affect its own future behavior and thereby today's bond price p^L. We solve this fixed point problem by computing the solution to a finite-horizon problem. Starting from a final date, we iterate backward until all firm-level quantities and bond prices have converged. We then use the first-period equilibrium firm policy and bond prices as the equilibrium of the infinite-horizon problem. This means that we iterate simultaneously on the value function and the long-term bond price (as in Hatchondo and Martinez, 2009). The presence of the idiosyncratic i.i.d. capital quality shock ε with continuous probability distribution $\varphi(\varepsilon)$ facilitates the computation of p^L (*cf.* Chatterjee and Eyigungor, 2012).

The third challenge is that the aggregate state of our general equilibrium model includes the time-varying firm distribution. We follow Reiter (2009) in first computing a fully non-linear global solution of the steady state with idiosyncratic firm-level uncertainty but without aggregate shocks. We then use a numerical first-order perturbation method (as in Schmitt-Grohé and Uribe, 2004) to approximate the dynamics of the model and its endogenous firm distribution around the steady state in response to aggregate shocks.

3.5.2 Calibration

A number of parameters can be set externally using standard values from the literature on firm dynamics and New Keynesian business cycle models. The remaining parameters are internally calibrated.

Externally set parameters. The model period is one quarter. We set $\beta = 0.99$ which implies a quarterly steady state real interest rate of $r^* = 1.01\%$. In the

Table 3.1: Externally set parameters

Parameter	β	c	θ	ζ	ψ	δ	γ	τ	ρ	φ^{mp}	ρ^{mp}	λ	ϕ
Value	0.99	0.01	0.5	0.75	0.33	0.025	0.05	0.4	10	1.25	0.5	90	4

steady state of the model, inflation is zero and the nominal interest rate i is equal to the real rate. The debt coupon is fixed at $c = r^*$ which implies that the steady state equilibrium prices of riskless short-term and long-term bonds are both equal to one. The preference parameter θ is chosen to match a Frisch elasticity of 2 as in Arellano et al. (2019).

The production technology parameters ζ and ψ are taken from Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2018). The quarterly depreciation rate δ is 2.5%. We follow Gomes et al. (2016) in setting the tax rate $\tau = 0.4$ and the repayment rate of long-term debt $\gamma = 0.05$.[21] The choice of γ implies a Macaulay duration of $(1+r^*)/(\gamma+r^*) = 16.8$ quarters or 4.2 years. This is a conservative choice relative to the average duration of 6.5 years calculated by Gilchrist and Zakrajšek (2012) for a sample of U.S. corporate bonds with remaining term to maturity above one year.

As in Kaplan et al. (2018), we set the elasticity of substitution for retail good varieties to $\rho = 10$ (implying a steady state markup of 11 percent) and the Taylor rule parameters to $\varphi^{\mathrm{mp}} = 1.25$ and $\rho^{\mathrm{mp}} = 0.5$. The price adjustment cost parameter λ and the parameter of the capital goods technology ϕ are taken from Ottonello and Winberry (2020). The parameters generate a slope of the Phillips Curve of $\rho/\lambda = 0.1$ as in Kaplan et al. (2018), and a response of aggregate investment to monetary policy shocks which is roughly twice as large as that of aggregate output (Christiano et al., 2005). All externally set parameters are summarized in Table 3.1.

Internally calibrated parameters. The probability distribution of the firm-specific capital quality shock ε is normal with zero mean and standard deviation σ_ε. Firm-level productivity z follows a productivity ladder with discrete support $\{Z_1, ..., Z_j, ..., Z_J\}$, where $\log Z_1 = -\bar{z}$ and $\log Z_J = +\bar{z}$. Entrants start at the lowest productivity level $z^e = Z_1$ (with zero assets, $q = 0$, and zero debt, $b = 0$). For an incumbent firm with $z = Z_j$, the probability to become more productive next period is given by $1 - \rho_z$:

$$z' = \begin{cases} Z_j & \text{with probability } \rho_z \\ Z_{\min\{j+1,J\}} & \text{with probability } 1 - \rho_z \end{cases} \tag{3.1}$$

Once a firm has reached the highest productivity level Z_J, it remains there until it defaults or exits the economy exogenously. This productivity process has two

[21] The parameter τ should be thought of as capturing additional benefits of using debt over equity besides the actual tax benefit of debt and equity issuance costs (e.g., limiting agency frictions between firm managers and shareholders as in Arellano et al., 2019).

Table 3.2: Internally calibrated parameters

Parameter	Value	Target	Data	Model
σ_ε	0.66	Average firm leverage *(in %)*	34.4	29.3
ξ	0.90	Average credit spread on long-term debt *(in %)*	3.1	3.3
η	0.0045	Average share of debt due within a year *(in %)*	30.5	30.7
ν	0.0005	Average equity issuance *(in %)*	11.4	14.6
ρ_z	0.983	Average firm capital growth *(in %)*	1.0	1.2
$\bar{z}$	0.184	Std. of firm capital growth *(in %)*	8.3	9.7
κ	0.0151	Firm exit rate *(in %)*	2.2	2.3
f	0.274	Steady state value of firm entry	-	0

Notes: The data sample is 1995-2017. Firm-level data on leverage (debt/assets), the share of debt due within a year, equity issuance (relative to assets), and capital growth is from Compustat. Firm-level credit spreads are computed using data from Compustat and FISD. The exit rate is from Ottonello and Winberry (2020). See Appendix 3.D.1 and 3.D.2 for details.

desirable features. First, it captures the positive skewness of empirical firm growth (Decker, Haltiwanger, Jarmin, and Miranda, 2014). Second, it facilitates the computation of the Markov perfect equilibrium.[22]

We internally calibrate eight parameters: σ_ε, ξ, η, ν, ρ_z, $\bar{z}$, κ, and f. Their values are chosen to match key empirical moments which are informative about the financing and investment behavior of firms. Firm-level data on leverage, equity issuance, and capital growth comes from Compustat. Credit spreads are calculated by combining firm-level credit ratings with rating-specific corporate bond spreads, following Arellano et al. (2019). To discipline firms' maturity choices in the model, we use Compustat information on the share of total debt (bonds and loans) due within a year (*cf.* Figure 3.1). While the FISD data used in Section 3.2 contains more precise information on maturity within a quarter, it is only available for a subset of Compustat firms.

The internal calibration is summarized in Table 3.2. While the model is highly non-linear and all parameters are jointly identified, we provide some intuition for their identification. Average leverage depends on the standard deviation of the capital quality shock σ_ε because higher earnings volatility induces firms to reduce leverage in order to contain the risk of default. The average credit spread is directly affected by the default cost ξ. The average maturing debt share pins down the debt issuance cost parameter η because higher debt issuance costs make short-term debt less attractive. The equity issuance cost parameter ν targets equity issuance relative to firm assets. The parameters ρ_z and $\bar{z}$ are important for matching the empirical moments of firm-level capital growth. The probability of exogenous exit κ affects the total rate of exit (endogenous and exogenous). Finally, the fixed cost of production f is chosen such that the steady state value of firm entry is zero.

[22]If a firm's amount of outstanding long-term debt $(1-\gamma)b^L/\pi$ is sufficiently high, large negative shocks to z' would cause the dividend payout constraint $e \geq \underline{e}$ in (3.20) to bind for any value of $\underline{e}$. The productivity process described above avoids this problem.

Table 3.2 shows that the model matches the data well. Average firm leverage and the maturing debt share are both about 30%. The average annual credit spread on long-term debt is close to 3 percent. Even though the value of the equity issuance cost parameter ν is smaller than the debt issuance cost parameter η, aggregate equity issuance costs exceed aggregate debt issuance costs (0.12% vs. 0.05% of GDP). The model generates a quarterly default rate of 0.8%. Although untargeted, the default rate is very close to the corresponding values of 0.8% in Bernanke et al. (1999) and the 1.0% in Moody's expected default frequency across rated and unrated Compustat firms reported by Hovakimian, Kayhan, and Titman (2011).

3.5.3 Steady state results

As we show in this section, the steady state of the calibrated model replicates how empirical firm financing choices vary by size and by age. One important fact in the data is that smaller and younger firms pay higher credit spreads and have larger shares of maturing debt. The model generates this result. It will play an important role for the cross-sectional effects of monetary policy in Section 3.5.4 below.

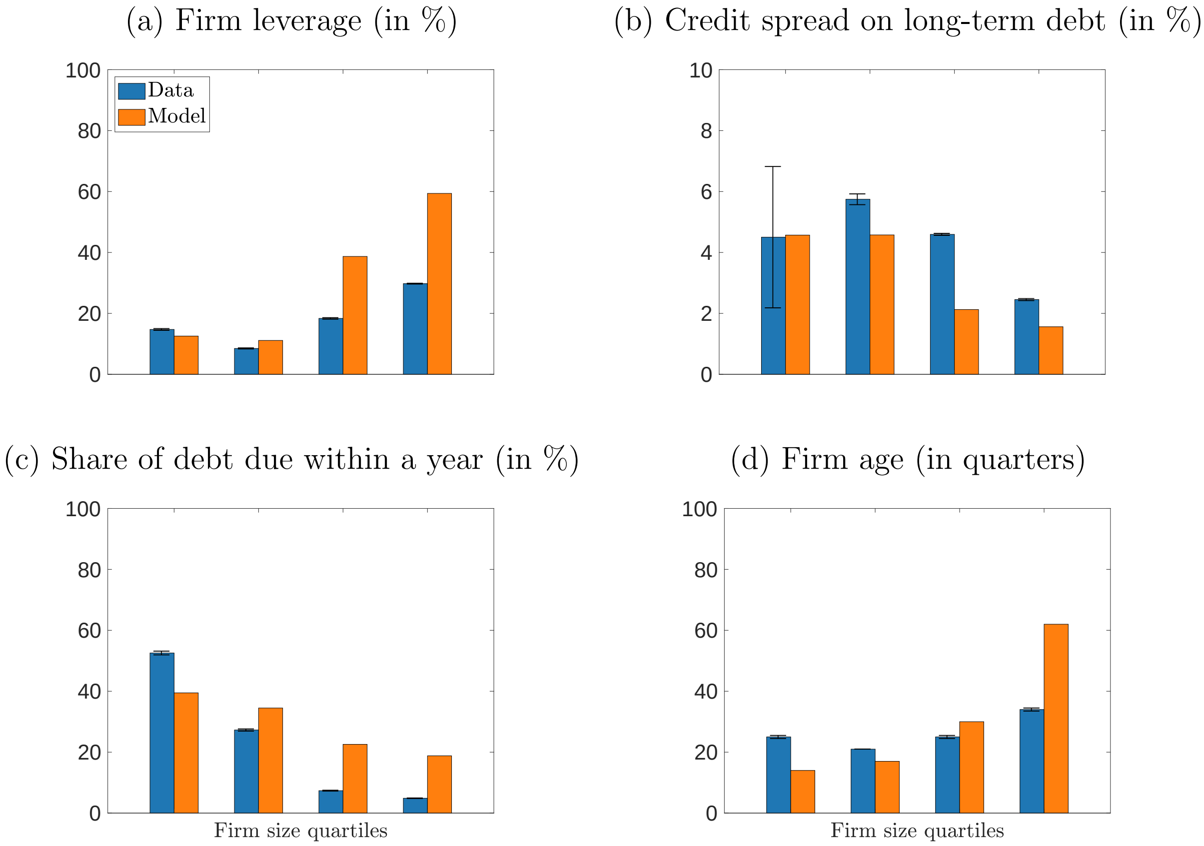
Figure 3.1 shows leverage, credit spreads, and the maturing debt share across quartiles of the firm size distribution. Blue bars indicate empirical values (with 95% confidence intervals). Orange bars show the corresponding moments in the model. In the data, leverage increases with firm size. Smaller firms firms pay higher credit spreads and have larger shares of maturing debt per period. The last panel shows that larger firms are older.

The model replicates these empirical patterns. Differences in firm productivity are key for this result. Low productivity firms choose a smaller scale of production. The fixed cost of production f implies that smaller firms are less profitable and therefore have higher default risk for given amounts of leverage. As a consequence, smaller firms pay higher credit spreads and choose lower amounts of leverage (see Appendix 3.D.3 for the policy functions of the calibrated model).

Panel (c) shows that the model also replicates the fact that the maturing debt share is higher for smaller firms. An advantage of long-term debt common to all firms is that it reduces future debt issuance costs. A disadvantage of issuing long-term debt is that it lowers today's long-term bond price because debt overhang will lead to higher future leverage and default risk (*cf.* Section 3.4.1). This negative effect of long-term debt on today's long-term bond price $\partial p^L / \partial b^{L'} < 0$ is stronger for smaller firms. Smaller firms have a higher default risk, which implies that their long-term bond price is more sensitive to changes in future firm behavior (see Figure 3.D.3 in Appendix 3.D.3). As a consequence, the costs of debt overhang are higher for them. Through this mechanism, the model can explain why smaller firms borrow at shorter maturities and therefore have higher shares of maturing debt.

A brief comparison with Figure 3.1 in Section 3.4.2 is in order here. In Figure 3.1, we compared two firms with exogenous differences in maturing debt shares. Debt overhang was larger for the firm with a lower maturing debt share. In the

Figure 3.1: Firm variables conditional on size



Notes: For each variable, median values are shown by size quartile. The data sample is 1995–2017. Firm-level data on size (total assets), leverage, the share of debt due within a year, and age (quarters since initial public offering) is from Compustat. Firm-level credit spreads are computed using data from Compustat and FISD. Empirical median values are shown with 95% confidence intervals. Confidence intervals are large for the bottom quartile in panel (b) because small firms are often unrated which means that we are unable to assign credit spreads to them. Model moments are computed from the stationary distribution of the model. See Appendix 3.D.1 and 3.D.2 for details.

quantitative model, debt maturity is endogenous. As Figure 3.1 shows, firms' maturity choice responds to differences in the costs of debt overhang. Debt overhang is a larger problem for firms with higher default risk. As a consequence, high-default risk firms choose to borrow at shorter maturities and therefore have higher maturing debt shares. This result will play an important role for the cross-sectional effects of monetary policy discussed below.

Finally, panel (d) of Figure 3.1 shows that the model also replicates the positive empirical relationship between firm age and size. Average firm productivity increases with age in the model. Older firms therefore choose higher values of capital and are larger. Additional results on the co-movement of age with leverage, credit spreads, and debt maturity are shown in Figure 3.D.4 in the Appendix. In the data, leverage is increasing in age whereas credit spreads and the maturing debt share are falling. The model replicates these untargeted patterns.

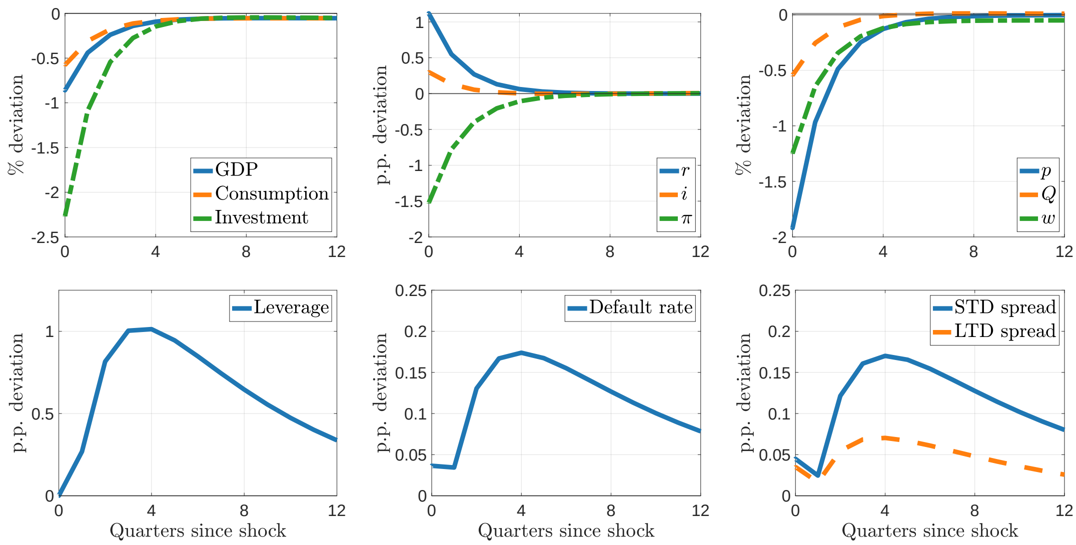
3.5.4 Aggregate effects of monetary policy shocks

The previous section showed that the model successfully replicates key cross-sectional facts about the financing choices of U.S. public firms. The model thus provides an appropriate quantitative framework for studying the role of debt maturity for the aggregate and heterogeneous effects of monetary policy. We begin by showing the model's aggregate implications.

Figure 3.2 shows the aggregate effects of an unexpected one-standard deviation (30bp) increase in the nominal interest rate i caused by a monetary policy shock ($\varepsilon_t^{\mathrm{mp}}$ in equation (3.17)). GDP, consumption, and investment all fall in response to the shock. The real interest rate r increases by more than the nominal rate because inflation π falls. The associated decline in aggregate demand causes a reduction in the price of undifferentiated output p. This reduces firms' demand for capital and labor and decreases the wage w and the price of capital goods Q.

The second row of Figure 3.2 shows key financial variables. The increase in the real interest rate reduces firm value while lower inflation π increases the real burden of outstanding nominal long-term debt $(1-\gamma)b^L/\pi$. As a result, firms accept an increase in leverage and default risk. Short-term credit spreads respond more strongly than long-term spreads because the price of short-term debt only depends on next period's default risk while the long-term bond price depends on default risk in all future periods.[23]

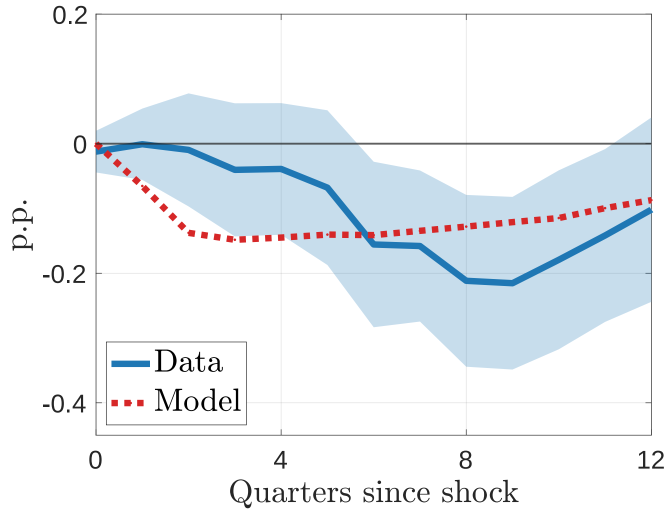
Figure 3.2: Aggregate response to a contractionary monetary policy shock



Notes: The real interest rate r, the nominal rate i, and inflation π are annualized. Leverage is aggregate firm debt over aggregate firm capital. The default rate is annual. The short-term credit spread (*STD spread*) and the long-term credit spread (*LTD spread*) are cross-sectional averages. See Appendix 3.D.1 for details.

[23]The model result that credit spreads rise after a contractionary monetary policy shock is consistent with empirical results in Gertler and Karadi (2015).

Figure 3.3: Differential capital growth response associated with $\mathcal{M}_{it}$



Notes: The red dotted line shows the estimated β_1^h coefficients based on equation (3.2) using simulated model data. The β_1^h estimates are standardized to capture the differential cumulative capital growth response (in p.p.) to a one standard deviation (30bp) increase in the nominal interest rate i associated with a one standard deviation higher $\mathcal{M}_{it}$. The blue solid line shows the empirical estimates from Figure 3.1(a) together with 95% confidence bands.

3.5.5 Heterogenous effects of monetary policy shocks

Our empirical analysis showed that firms with a higher share of maturing debt are more responsive to monetary policy shocks. In this section, we show that our model replicates this result.

Local projection on simulated model data. To compare the model with the empirical evidence, we run the model counterpart of the baseline local projection (3.2) on simulated data generated by our model. We estimate:

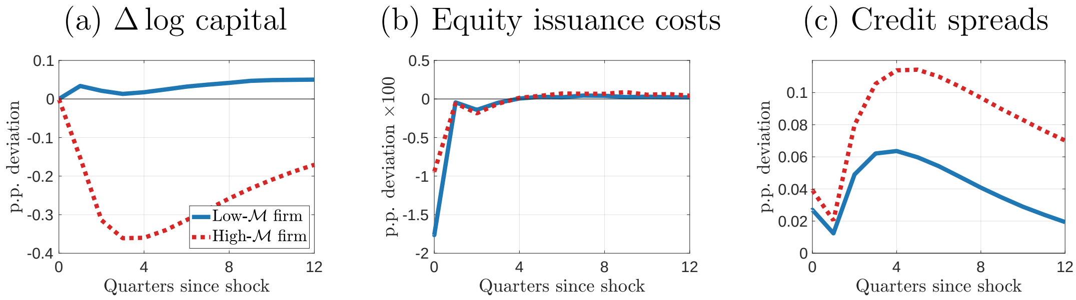
$$\Delta^{h+1} \log k_{it+h} = \beta_0^h \mathcal{M}_{it} + \beta_1^h \mathcal{M}_{it} \varepsilon_t^{\text{mp}} + \delta_i^h + \delta_t^h + \nu_{it+h}^h, \tag{3.2}$$

where δ_i^h and δ_t^h are firm- and quarter-fixed effects, and $\mathcal{M}_{it}$ is the maturing bond share as defined in (3.4).[24] Figure 3.3 shows the estimated β_1^h coefficients in the model (red dotted line) and in the data (blue solid line, *cf.* Figure 3.1(a)). The estimates in Figure 3.3 are standardized to measure the differential response associated with a one standard deviation higher $\mathcal{M}_{it}$ at the time of an unexpected one standard deviation (30bp) increase in the nominal interest rate i.

As in the data, β_1^h is negative at all time horizons: A higher $\mathcal{M}_{it}$ implies a larger negative capital response. The model accounts for 69% of the peak empirical effect. Similar to the empirical results, the differential effect on firm investment is

[24] As in the empirical specification, we use average total debt over the preceding four quarters as the denominator for $\mathcal{M}_{it}$. All model results are virtually indistinguishable when using the current level of debt as the denominator instead.

Figure 3.4: Heterogeneous responses to a contractionary monetary policy shock



Notes: The panels show the effect of an unexpected one-standard deviation (30bp) increase in the nominal interest rate i for firms below and above the median maturing bond share $\mathcal{M}$ at the time of the shock. Panel (a) shows average firm-level changes in log capital. Panel (b) shows average equity issuance costs (relative to steady state capital). Panel (c) shows average credit spreads.

initially small and builds up over time, reaching its peak three quarters after the shock. The persistence generated by the model is high: Twelve quarters after the shock, 59% of the peak differential effect is still present.

The model also replicates the empirical role of $\mathcal{M}_{it}$ for the response of other firm variables. Figure 3.D.5 in the Appendix shows that a higher $\mathcal{M}_{it}$ at the time of the shock is associated with larger reductions in sales, employment, and debt relative to pre-shock capital. These model results are in line with the empirical findings of Figure 3.2.

Monetary transmission and $\mathcal{M}_{it}$. The model rationalizes the main empirical result of the paper: a higher share of maturing debt at the time of a monetary policy shock is associated with a stronger effect on firm investment. Figure 3.4 shows that both roll-over risk and debt overhang contribute to this result.

The figure shows average responses of firms whose maturing bond share is above or below the median at the time of the shock. Panel (a) shows that high-$\mathcal{M}$ firms sharply reduce investment after the contractionary monetary policy shock. In contrast, low-$\mathcal{M}$ firms slightly increase capital as they benefit from lower factor prices w and Q. As in Figure 3.3, the difference between the two firm groups builds up over time and peaks several quarters quarters after the shock.

Roll-over risk: Panel (b) shows that equity issuance costs fall for both groups of firms after the contractionary shock. However, the decline is smaller for high-$\mathcal{M}$ firms. As explained in Section 3.4.2, high-$\mathcal{M}$ firms have higher roll-over needs which generates a higher pass-through of lower bond prices to bond market revenue and cash flow. This cash shortfall requires higher equity issuance compared to low-$\mathcal{M}$ firms. Equity issuance costs therefore fall by less for high-$\mathcal{M}$ firms which increases their cost of capital relative to low-$\mathcal{M}$ firms and contributes to a larger reduction in investment. However, the differential impact on equity issuance is short-lived.

Debt overhang: Panel (c) shows credit spreads by firm group. Different from

Figure 3.1 in Section 3.4.2, credit spreads increase by more for high-$\mathcal{M}$ firms. The reason for this seeming contradiction is that debt maturity is an endogenous response to the firm-specific costs of debt overhang. Because debt overhang is more severe for firms with higher default risk, they choose to borrow at shorter maturities and therefore have higher maturing bond shares $\mathcal{M}$.

After a contractionary monetary policy shock, many firms reduce capital while the real burden of outstanding nominal long-term debt $(1-\gamma)b^L/\pi$ grows through debt deflation. It is feasible to keep leverage, default risk, and credit spreads constant after a reduction in capital but this would require repurchasing some of the now outsized stock of previously issued long-term debt. And while these repurchases would need to be financed by shareholders, they would to a large extent benefit existing creditors. The size of this externality is larger for firms with higher default risk because the market value of their debt is more sensitive to firm behavior. As a result, debt overhang drives up default risk and credit spreads by more for high-default risk firms despite their higher $\mathcal{M}$.[25]

The differential impact on credit spreads is long-lived, peaking four quarters after the shock. Debt deflation is an important reason for this persistence. The decline in inflation leads to a gradual build-up in the real burden of outstanding nominal debt. This amplifies firms' heterogeneous exposure to debt overhang and is key for the high degrees of persistence displayed in both panel (a) of Figure 3.4 and in Figure 3.3.

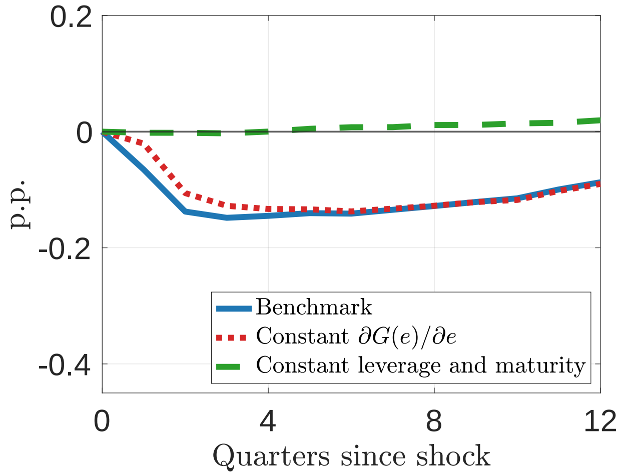
Decomposing the transmission channels. Roll-over risk and debt overhang both contribute to the result that investment falls by more for high-$\mathcal{M}$ firms after a contractionary monetary policy shock. To assess the two channels' relative quantitative importance, we simulate two counterfactual economies. We find that debt overhang is more important than roll-over risk for explaining the persistent differential investment effect associated with $\mathcal{M}$.

Constant marginal equity issuance costs: In the first counterfactual economy, for every firm state we hold marginal equity issuance costs $\partial G(e)/\partial e$ fixed at steady state values. This exercise is motivated by our analysis in Section 3.4.2, where we showed that roll-over risk only affects investment through changes in marginal equity issuance costs. By keeping marginal equity issuance costs at their steady state levels, we eliminate relative changes in the cost of capital that stem from different responses in equity issuance across firm groups.

Figure 3.5 compares the results from the local projection (3.2) using data from our benchmark model and the two counterfactual economies. The blue solid line reprints the estimates from the benchmark model. The red dotted line shows the β_1^h coefficients from the model with constant marginal equity issuance costs. The difference to our benchmark model is modest and short-lived. One reason for this result is that the cash flow effect stemming from different exposure to interest rate changes is small and does not generate much heterogeneity across

[25]Because high-$\mathcal{M}$ firms face a higher default risk, a given increase in leverage causes a larger increase in their default risk compared to low-$\mathcal{M}$ firms. This explains why credit spreads grow by more for high-$\mathcal{M}$ firms even though debt increases by less, as shown in Figure 3.D.5.

Figure 3.5: Counterfactuals: Differential capital growth response associated with $\mathcal{M}_{it}$



Notes: The blue solid line shows the estimated β_1^h coefficients based on equation (3.2) using simulated model data (*cf.* Figure 3.3). The β_1^h estimates are standardized to capture the differential cumulative capital growth response (in p.p.) to a one standard deviation (30bp) increase in the nominal interest rate i associated with a one standard deviation higher $\mathcal{M}_{it}$. The red dotted line shows the corresponding value in a counterfactual economy with fixed marginal equity issuance costs. The green dashed line shows the corresponding value in a counterfactual economy with fixed leverage and debt maturity.

firms. A simple back-of-the-envelope calculation shows that a 30bp increase in the nominal interest rate only produces differences in cash flow of less than 0.04% of firm capital. Because interest rates revert back to their long-run mean quickly after the monetary policy shock, this effect is short-lived.[26]

Constant leverage and debt maturity: In the second counterfactual economy, we remove the effects of debt overhang on firms' investment response to monetary policy shocks. To do so, for every firm state we fix leverage $(b^{S'} + b^{L'})/k'$ and debt maturity $b^{L'}/(b^{S'} + b^{L'})$ at the respective steady state value. This is motivated by Section 3.4.2, where we described that debt overhang affects investment through the impact of firms' financing choices on default risk. In the counterfactual economy, firms cannot adjust their leverage and maturity choices in response to a monetary policy shock. As debt deflation increases the real burden of outstanding nominal debt, firms must keep leverage constant by raising outside equity or by reducing dividends.

The green dashed line in Figure 3.5 shows the β_1^h coefficients estimated using data from this counterfactual economy. The difference between high- and low-$\mathcal{M}$ firms' capital response disappears at all time horizons. Once firms' financing

[26]The standard deviation of the maturing bond share across firms is 13.1%. Assuming a leverage ratio of 30% and a real interest rate increase of one percentage point (as in Figure 3.2), a one-standard deviation higher value of $\mathcal{M}_{it}$ increases the fall in bond market revenue by $13.1\% \times 30\% \times 1\% = 0.039\%$ relative to firm capital. This calculation abstracts from changes in credit spreads caused by the monetary policy shock.

structure is held fixed, default risk increases homogeneously across firms. This prevents credit spreads and the cost of capital from increasing more for high-$\mathcal{M}$ firms, as shown in Figure 3.D.6 in the Appendix. We conclude that debt overhang is the key channel for explaining persistent differences in the response of capital and credit spreads across firms.

3.5.6 Aggregate implications of heterogeneous debt maturity

In this section, we study the importance of heterogeneous debt maturity for the aggregate effects of monetary policy. We find that both long-term debt and heterogeneity amplify the aggregate effects of monetary policy shocks.

Model without long-term debt. To highlight the role of debt maturity for the aggregate effects of monetary policy, we first compare our benchmark model to an alternative economy in which firms can only issue short-term debt, but not long-term debt. This is the case in most macro models with firm-level financial frictions (e.g., Bernanke et al., 1999; Ottonello and Winberry, 2020). Because firms are only allowed to issue short-term debt, there is no heterogeneity in debt maturity in this economy. In all other respects, the setup is identical to the benchmark model with endogenous debt maturity described above.[27]

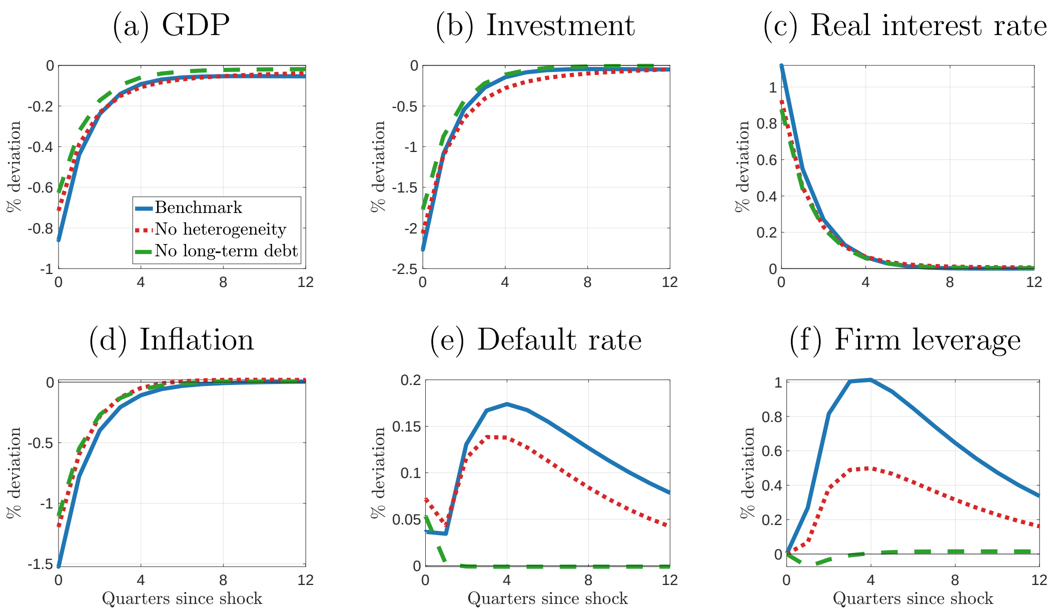
Figure 3.6 compares the aggregate effects of a contractionary monetary policy shock in our benchmark model (blue, solid lines) to two alternative economies. The green dashed lines show results for the short-term debt model. Although on impact the nominal interest rate increases by 30bp in all economies, the effects are very different. The negative GDP response is about 27% smaller in the short-term debt model (-0.63 p.p., compared to -0.86 p.p. in the benchmark economy). Investment and inflation also respond by less in the alternative model without long-term debt.

The reason for these dampened aggregate effects is that leverage and the default rate hardly react to the contractionary shock, as shown by the green dashed lines in panels (e) and (f). In the absence of long-term debt, there is no debt overhang. When firms decide on their leverage and default risk, no existing stock of previously issued long-term debt distorts their incentives. Because default risk and credit spreads move very little in the short-term debt model, the cost of capital increases by less compared to the benchmark economy which results in lower financial amplification.[28]

[27]To parameterize the short-term debt model, we set $\gamma = 1$ and re-calibrate model parameters to match the same empirical targets as above (*cf.* Table 3.2). Details are provided in Appendix 3.D.5.

[28]As a matter of fact, the model without long-term debt displays financial dampening relative to a model without financial frictions, as shown in Figure 3.D.7 in the Appendix. In the benchmark model, the effects of monetary policy are amplified relative to the frictionless case through strongly counter-cyclical default rates and credit spreads.

Figure 3.6: Aggregate response to monetary policy shock: Model comparison



Notes: The real interest rate r, inflation π, and default rates are annualized. Leverage is aggregate firm debt over aggregate firm capital. The blue solid lines are identical to those in Figure 3.2. The green dashed lines come from an alternative economy without long-term debt. The red dotted lines are from an alternative economy in which firms are ex-ante identical at the start of each period.

Model without heterogeneity. Different from existing models of long-term debt (e.g., Gomes et al., 2016), our model generates a realistic degree of debt heterogeneity across firms. In our second alternative economy, we study the quantitative importance of this heterogeneity. To do so, we solve an alternative model in which all firms are ex-ante identical every period. The setup is otherwise identical to the benchmark model with firm heterogeneity. In particular, both models include a debt maturity choice.[29]

The red dotted lines in Figure 3.6 show the response of the model without heterogeneity to the contractionary monetary policy shock. Compared to the benchmark model, the negative GDP response is about 16% smaller (-0.72 p.p.). The initial responses of investment and inflation are dampened as well. Even though debt overhang is present in both economies and the steady state averages of leverage, credit spreads, and debt maturity are the same, an identical increase in the nominal rate causes very different model responses with and without firm heterogeneity.

An important reason for these distinct outcomes is the difference in persistence generated by the two models. While on impact GDP and investment respond by more in the benchmark model, they also revert back more quickly to their unconditional long-run averages. The persistence is lower in the benchmark model because it has an endogenous firm distribution: As default rates increase after a

[29] We calibrate the model without heterogeneity to the same empirical targets as above (*cf.* Table 3.2). Details are provided in Appendix 3.D.5.

contractionary shock, defaulting firms are replaced by new firms that enter without existing debt. This reduces the average stock of outstanding long-term debt in the benchmark economy and lowers the negative impact of debt overhang, speeding up the recovery. This effect is absent in the economy without heterogeneity because all firms are ex-ante identical and the (degenerate) firm distribution is not time-varying.

Differences in persistence are important because of intertemporal substitution. The shorter-lived reduction of wages in the benchmark economy strengthens the substitution effect on labor supply and allocates labor away from periods of low wages (King and Rebelo, 1999). The resulting larger drop in output and consumption implies that the real interest rate increases by more in order to balance households' desire for consumption smoothing. In this way, lower persistence contributes to the larger initial decline in GDP in the benchmark economy. The alternative model without heterogeneity over-predicts the persistence of debt overhang and therefore understates the aggregate effects of monetary policy.

3.6 Conclusion

More than two decades after the first seminal contributions introduced frictional firm financing into quantitative dynamic models of the macroeconomy (e.g., Bernanke et al., 1999), the contemporaneous literature offers new insights by focusing on debt heterogeneity.[30] As part of this broader research agenda, our paper documents the vast amount of heterogeneity in U.S. public firms' maturity choices. The maturity dimension of debt heterogeneity is typically absent from standard one-period-debt macro models.

We showed that heterogeneous debt maturity matters for monetary policy. We used micro data to show that firms respond more strongly to monetary policy shocks when a higher fraction of their debt matures. We then developed a heterogeneous firm New Keynesian model with financial frictions and endogenous debt maturity. The model accounts for the maturity of debt and its distribution across firms. It replicates the empirical result that firms with higher shares of maturing debt react more strongly to monetary policy shocks. At the aggregate level, we showed that both long-term debt and heterogeneity amplify the effects of monetary policy shocks on GDP, investment, and inflation. We conclude that the maturity of firm debt and its distribution are important for the aggregate effects of monetary policy.

These results raise new questions for the conduct of systematic monetary policy. How should central banks' policy response to shocks take debt maturity into account? When facing a trade-off between stabilizing output and inflation, the important role of debt overhang and debt deflation suggests that a given surprise increase in inflation can achieve a larger reduction in the output gap. The

[30]For instance, recent contributions study differences between bonds and loans (Crouzet, 2018; Darmouni et al., 2021), between floating-rate debt and fixed-rate debt (Ippolito et al., 2018; Gurkaynak et al., 2021), or between credit lines and term loans (Greenwald et al., 2021).

model developed in this paper provides a quantitative framework for studying this question.

Another natural application of our framework is to study the consequences of unconventional monetary policy and quantitative easing. The persistent decline in the term structure of interest rates during the ten years following the Great Recession had different implications for firms borrowing at short and long maturities. Our results highlighted systematic differences between these firm groups. A rigorous analysis of the aggregate effects of quantitative easing therefore requires a model of heterogeneous debt maturity. We hope that the results presented in this paper provide a useful starting point for addressing these open questions.

One interesting empirical finding was that the precise timing of bond maturity can make a difference for firms' investment response to monetary policy shocks. An open question is whether non-convex adjustment costs induce firms to be more responsive to aggregate shocks at times of re-financing. While conceptually and computationally demanding, introducing non-convex adjustment costs to a framework of endogenous debt maturity and default could yield additional valuable insights.[31]

Appendices for Chapter 3

3.A Data construction

3.A.1 Bond-level data

From Mergent FISD we obtain detailed bond-level data for bonds that mature between 1995Q2 and 2018Q3. The initial sample contains 304,868 bonds denominated in US$. In this sample, the total value of bonds at issue date amounts to 70.6 trillion (tn) US$ and the total value of bonds at maturity date is 57.7tn US$. The main reason why the value changes between issue date and maturity date is (partial) calls.

We construct a sample of comparable bonds by dropping the following types of bonds: convertible (number of bonds: 3,217; value at issue date: 698bn US$; value at maturity date: 292bn US$), convertible on call (322; 83bn; 37bn), exchangeable (32,105; 790bn; 752bn), (yankee) bonds issued by foreign entities (44,035; 8.8tn; 8.3tn), and bonds that mature less than one year after issuance (55,280; 22.3tn; 21.9tn). These bond types are not mutually exclusive and partially overlap. Dropping these type of bonds leaves us with a sample of 220,253 bonds with a value at issue date of 38.4tn US$ and a value at maturity date of 26.9tn US$. Of these bonds, we focus on fixed-coupon, non-callable bonds (61,642; 17.4tn; 17.1tn), which account for the majority of the value of bonds at maturity date. We further analyze bonds that are callable (140,598; 16.0tn; 4.9tn) or have a variable coupon (43,450; 7.1tn; 5.6tn).

[31] For recent contributions on aggregate implications of lumpy firm-level adjustment, see Koby and Wolf (2020), Baley and Blanco (2021), and Winberry (2021).

We then create a monthly panel of bonds which tracks the outstanding amount – the par value computed as number of bonds issued times principal amount – over the lifetime of a bond. Mergent FISD further records (only) the most recent action taken on a bond before maturity. An action can involve a reduction in the amount outstanding before maturity, e.g., due to a a call, reorganization, or default. In this case, the data records the date, amount, and reason of reductions in the amount outstanding that occur before maturity, e.g., due to a a call, reorganization, or default. Among the total sample of bonds, about half record an action, while for only 5% of non-callable bonds an action is recorded. We use those records to adjust the outstanding amount in our bond panel. When the bond matures at its scheduled maturity date, we use the remaining amount of the bond at maturity as maturing amount.

3.A.2 Linking bonds and firms

To match bonds to the debtor firm in every period over the bond's lifetime, we proceed in three steps. First, we construct a mapping from `gvkey`, the Compustat firm identifier, to the historical firm `cusip`. A firm `cusip` identifier is contained in the bond `cusip` identifier, which allows us to match bonds to firms. However, the bond `cusip` contains an identifier of a firm valid at the time of issuance. Because these firm `cusips` frequently change over time (for a given firm), we need to identify the historic firm `cusip` identifier valid in a given time period. To link `gvkey` and historical firm `cusip`, we combine the Compustat–CRSP link table (linking `gvkey` and `permno`, a firm identifier in CRSP) with CRSP, which links `permno` and historical firm `cusip`. The Compustat–CRSP link contains the start and end dates for which `gvkey-permno` links are valid. We only use links which are classified as reliable, coded "C" or "P" in the link table. We join this link table with the CRSP data and keep records that fall within link validity. For few `cusips` we have a link to more than one `gvkey`, which may arise due to the presence of subsidiary firms in CRSP. Among these ambiguous links, we drop links from `cusip` to `gvkey` with missing sales in Compustat. For the remaining ambiguous links we keep the `gvkey` link to the firm with the largest sales.

Second, we cannot simply match the bond panel to the firm panel by using the historical `cusip` in both panels. In the bond panel, the historical firm `cusip`, encoded in the bond `cusip`, is the firm `cusip` at the time of bond issuance. In contrast, the firm panel records the historical firm `cusip` as the one valid in a given period, which may change over time. Reasons for changes in the historical `cusip` are changes in the firm name or the firm trading symbol. To match firm and bond panel, we use the so-called *header* firm `cusip` associated to the bond's initial *historical* firm `cusip`. The header `cusip` is the latest observed `cusip` in a firm's history. The mapping between header `cusips` and historical `cusips` over time is provided in CRSP data. We match the header `cusip` to both the firm and the bond panel. The link between bond and firm panel along the header `cusip` is ambiguous in a small number of cases. We delete those bonds for which no link to `gvkey` is available in the Compustat—CRSP table and drop the bonds with remaining ambiguous links. Given the header `cusip` of the bond issuer, we can attach the historical `cusip` series throughout the lifetime of the bond using the same mapping. If the debtor firm of the bond does not change (e.g., because of M&A), this procedure correctly identifies the bond debtor over the lifetime of the bond.

Third, we account for M&A events. The Thomson–Reuters SDC database records events at which firms - as identified by historical `cusip` - are merged or acquired by another firm, also identified by historical `cusip`. This allows us to change a bond's firm identifier to the identifier of the acquiring firm. We prepare the SDC data as follows. We do not consider M&A events for which no date is reported, the M&A status is not reported as completed, the target firm is classified as a subsidiary, or if the acquiring firm does not buy the target firm fully. If an M&A event is associated to multiple buyers, we drop buyers that do not have associated gvkeys as per the Compustat—CRSP link table and drop remaining events of this sort entirely. With this data at hand, we merge M&A events to the bond panel. For bond-months in which the creditor was subject to an M&A event, we replace the historical firm `cusip` associated to the bond by the acquiring firm's `cusip` from the M&A date going forward. Because the acquiring firm may have changed its `cusip` after the M&A event, we need to repeat the steps outlined above to find the actual evolution of the historical `cusip` for the new creditor firm. Having done so, we search for additional M&A events that may have happened after the first M&A event, now with the first acquiring firm being the target firm. We repeat this procedure until we find no M&A events that would imply a change in the `cusip` identifier.

3.A.3 Variables

Capital growth. We construct capital stock series either using a perpetual inventory method (PIM) or deflated book values. Both are based on net property, plants, and equipment (PPE, `ppentq` in Compustat), and we exclude firm-quarters with negative values of net PPE. For the PIM, we first identify investment spells for which net PPE is observed without gaps. If the gap is only a single quarter, we impute net PPE via linear interpolation. We exclude a small number of one-quarter capital spikes. These are quarters in which the real absolute growth rate of PPE exceeds 50% and is followed by a reversal in the opposite direction of more than 50% in the following quarter. For the first period of every investment spell we initialize capital by (deflated) gross PPE (`ppegtq`). For all subsequent quarters of the same spell we compute capital by adding the first difference in (deflated) net PPE to capital of the previous quarter. To construct deflated book values we simply deflate net PPE by the CPI. For both measures of capital, we only consider firm-quarters of firms for which at least 40 quarters of capital are observed, similar to Ottonello and Winberry (2020). We trim the cumulative capital growth rates at the top and bottom 1% of the distribution.

Maturing bond share. We compute the maturing bond share $\mathcal{M}_{it}$ defined in (3.1) by dividing the total par value of maturing bonds of firm i in quarter t by average total debt of firm i over the preceding four quarters from $t-1$ to

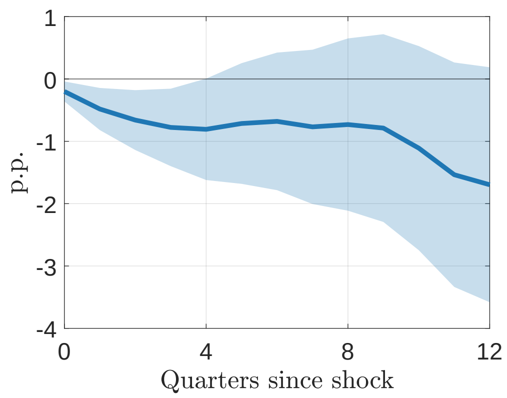
$t-4$. Total debt is based on current and long-term liabilities (`dlcq`+`dlttq`). We smooth out firm-specific seasonal factors and other transitory fluctuations by using the backward-looking four-quarter moving average of debt. We trim the maturing bond share at 100%. Analogous to capital growth, we only consider firm-quarters for firms with at least 40 quarters of observed maturing bond shares. The alternative denominators for $\mathcal{M}_{it}$ we consider are total debt at the end of period $t-1$, as well as capital, sales, and assets (both as backward-looking four-quarter moving averages and as simple lagged values), see Section 3.2.4.

Control variables. The list of control variables includes leverage, liquidity, average maturity, sales growth, and log assets. Leverage is total debt (`dlcq`+`dlttq`) divided by assets (`atq`). Liquidity is cash and short-term investments (`cheq`) divided by assets (`atq`). Average maturity is the average remaining maturity across outstanding bonds for firm i in quarter t, weighted by the par value of the outstanding bonds. Sales growth is the growth rate of deflated sales (`saleq`). Log assets is the natural logarithm of deflated assets (`atq`). All control variables are winsorized at the top and bottom 0.5% of the distribution. We measure firm age as the time since a firm's entry into the Compustat database. For this we complement quarterly Compustat data with annual Compustat data, as some firms initially only issue annual statements.

Other outcomes. In Figures 3.2 and 3.B.2, we consider growth in debt, sales, employment, and cost of goods sold as outcomes. We use total debt (`dlcq`+`dlttq`), sales (`saleq`), and costs of goods (based on `cogsq`), all deflated. We smooth out firm-specific seasonal factors and other transitory fluctuations by using the backward-looking four-quarter moving average of debt, sales, and cost of goods sold. We then estimate local projections on the log differences of these smoothed variables. This yields similar results as Smooth Local Projections proposed by Barnichon and Brownlees (2019). Employment is only recorded annually in Compustat. We construct quarterly firm-level employment via the Chow and Lin (1971) method by combining annual employment and quarterly cost of goods sold. We use `cogsq` because it contains employment expenses, which means quarterly variation in `cogsq` should be informative about employment. We trim the cumulative growth rates of debt, sales, employment, and cost of goods sold at the top and bottom 1% of the distribution.

3.B Additional empirical results

Figure 3.B.1: Average response of capital growth



Notes: The figure shows the estimated β_1^h coefficients in the local projection $\Delta^{h+1} \log k_{it+h} = \alpha_i^h + \alpha_{sq}^h + \beta_1^h \varepsilon_t^{\text{mp}} + \Gamma_1^h Y_{t-1} + \nu_{it}^h$, where α_i^h and α_{sq}^h are firm and sector-fiscal quarter fixed effects, $\varepsilon_t^{\text{mp}}$ is a monetary policy shock, and Y_{t-1} a vector of macroeconomic control variables including four lags of real GDP growth and CPI inflation. The β_1^h estimates are standardized to capture the response to a one standard deviation increase in $\varepsilon_t^{\text{mp}}$. Shaded areas indicate 95% confidence bands two-way clustered by firms and quarters.

Table 3.B.1: Full list of coefficients in baseline local projection for selected forecast horizons h

	$\Delta^{h+1} \log k_{it+h}$			
	$h = 0$	$h = 4$	$h = 8$	$h = 12$
$\mathcal{M}_{it}$	0.0140	0.00278	0.0743	0.238**
	(0.0238)	(0.0858)	(0.0951)	(0.0967)
$\mathcal{M}_{it}$ × MP shock	-0.0116	-0.0453	-0.214***	-0.102
	(0.0156)	(0.0511)	(0.0663)	(0.0679)
$\mathcal{M}_{it}$ × GDP growth	-0.0331	-0.0244	-0.000532	-0.231
	(0.0348)	(0.0971)	(0.161)	(0.154)
Firm FE	Yes	Yes	Yes	Yes
Industry-quarter FE	Yes	Yes	Yes	Yes
R^2	.15	.26	.33	.38
N	35,499	35,113	33,583	31,691

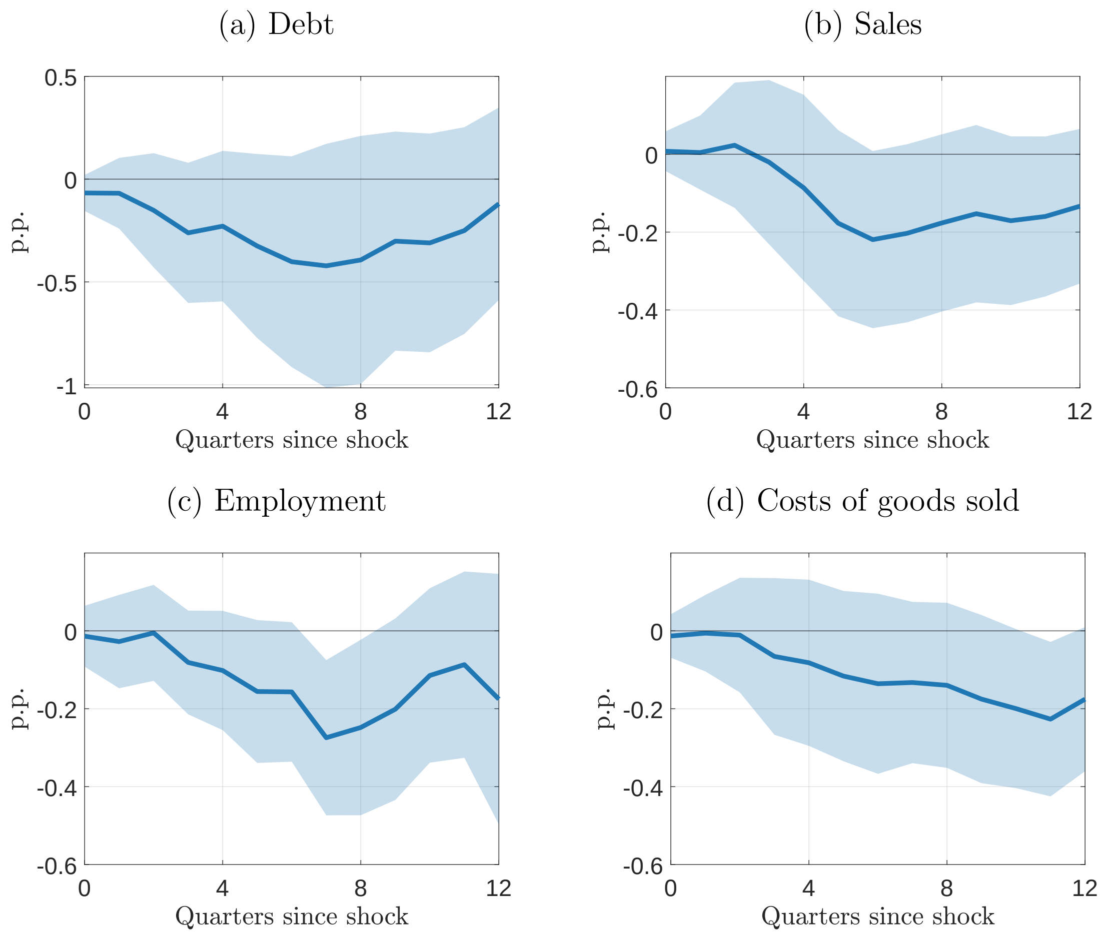
Notes: The table shows all estimated coefficients from the baseline local projection (3.2). The coefficient estimates are standardized to capture the effects of a one standard deviation change in $\mathcal{M}_{it}$, a one standard deviation change in the monetary policy shock, and a 1 p.p. change in GDP growth. Standard errors (in parentheses) are clustered by firm and quarter.

Table 3.B.2: Full list of coefficients in extended local projection for selected forecast horizons h

	$\Delta^{h+1} \log k_{it+h}$			
	$h = 0$	$h = 4$	$h = 8$	$h = 12$
$\mathcal{M}_{it}$	-0.0148	-0.124	-0.137	-0.0272
	(0.0243)	(0.0834)	(0.0861)	(0.109)
$\mathcal{M}_{it}$ × MP shock	-0.0219	-0.127*	-0.317***	-0.220**
	(0.0182)	(0.0652)	(0.0788)	(0.0965)
$\mathcal{M}_{it}$ × GDP growth	0.00539	0.196**	0.367**	0.222
	(0.0385)	(0.0958)	(0.140)	(0.160)
Leverage	-0.284**	-2.333***	-3.392***	-4.246***
	(0.130)	(0.579)	(1.017)	(1.230)
Leverage × MP shock	-0.0360	-0.102	0.0479	0.309**
	(0.0445)	(0.268)	(0.280)	(0.150)
Leverage × GDP growth	-0.212*	-0.620	-0.956	-0.827
	(0.122)	(0.374)	(0.692)	(0.764)
Liquidity	0.516***	1.162**	2.370***	2.835***
	(0.104)	(0.472)	(0.752)	(0.919)
Liquidity × MP shock	0.120**	-0.0549	-0.0434	0.151
	(0.0584)	(0.154)	(0.248)	(0.314)
Liquidity × GDP growth	-0.160*	0.361	-0.373	-0.122
	(0.0830)	(0.386)	(0.631)	(0.639)
Sales growth	0.0999	0.929***	0.803***	1.007***
	(0.0686)	(0.196)	(0.236)	(0.271)
Sales growth × MP shock	0.0454	-0.114	-0.266	-0.370**
	(0.0625)	(0.136)	(0.197)	(0.168)
Sales growth × GDP growth	-0.0258	0.266	0.467	0.129
	(0.0777)	(0.238)	(0.317)	(0.312)
Size	-0.695***	-5.400***	-10.26***	-15.75***
	(0.182)	(0.906)	(1.754)	(2.385)
Size × MP shock	-0.0187	0.108	-0.265	-0.768
	(0.0857)	(0.305)	(0.419)	(0.533)
Size × GDP growth	0.0754	0.0404	0.167	0.621
	(0.168)	(0.531)	(1.039)	(1.118)
Avg. bond maturity	-0.00414	-0.234	-0.370	-0.423
	(0.0494)	(0.269)	(0.439)	(0.566)
Avg. bond maturity × MP shock	0.0271	0.000360	0.0244	0.0274
	(0.0332)	(0.196)	(0.205)	(0.121)
Avg. bond maturity × GDP growth	0.0594	0.414	0.599	0.500
	(0.0579)	(0.286)	(0.412)	(0.371)
Firm FE	Yes	Yes	Yes	Yes
Industry-quarter FE	Yes	Yes	Yes	Yes
R^2	.2	.37	.48	.56
N	13,568	13,495	13,115	12,643

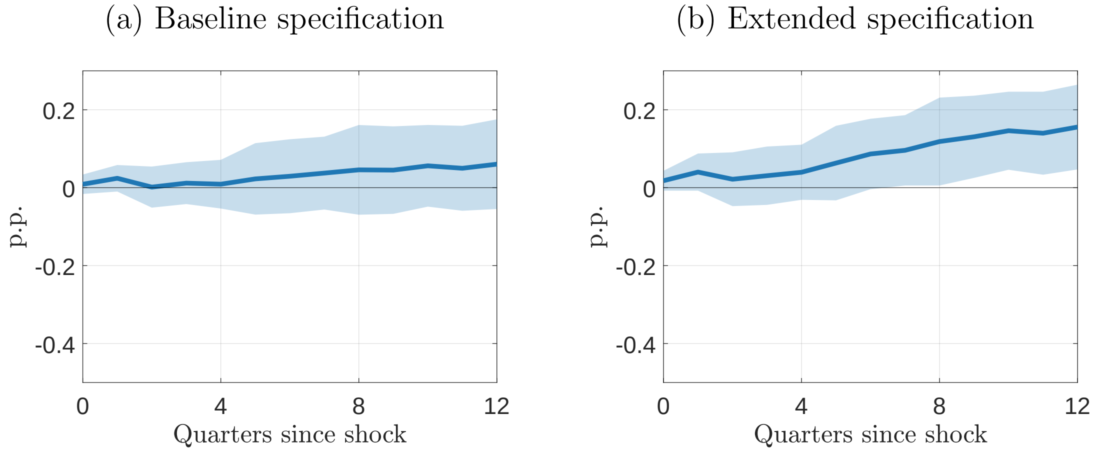
Notes: The table shows all estimated coefficients from the extended local projection (3.3). The coefficient estimates are standardized to capture the effects of a one standard deviation change in demeaned $\mathcal{M}_{it}$ and other covariates, a one standard deviation change in the monetary policy shock, and a 1 p.p. change in GDP growth. Standard errors (in parentheses) are clustered by firm and quarter.

Figure 3.B.2: Differential response of other variables associated with higher $\mathcal{M}_{it}$ using baseline local projection



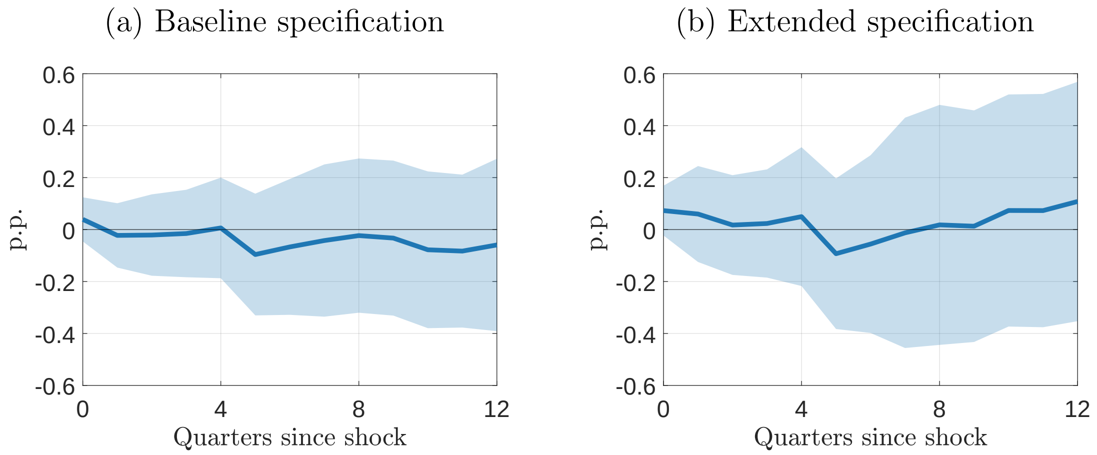
Notes: The figure shows the estimated β_1^h coefficients based on equation (3.2), but where the left-hand side is $\Delta^{h+1}\log(\text{debt})_{it+h}$ in panel (a), $\Delta^{h+1}\log(\text{sales})_{it+h}$ in panel (b), $\Delta^{h+1}\log(\text{employment})_{it+h}$ in panel (c), and $\Delta^{h+1}\log(\text{cost of goods sold})_{it+h}$ in panel (d). The β_1^h estimates are standardized to capture the differential response (approx. in p.p.) to a one standard deviation increase in $\varepsilon_t^{\text{mp}}$ associated with a one standard deviation higher $\mathcal{M}_{it}$. Shaded areas indicate 95% confidence bands two-way clustered by firms and quarters.

Figure 3.B.3: Differential investment response associated with lagged maturing bond share $\mathcal{M}_{t-1}$



Notes: Panel (a) shows the estimated β_1^h coefficients based on the baseline local projection (3.2) using $\mathcal{M}_{it-1}$ instead of $\mathcal{M}_{it}$. Panel (b) shows the estimated β_1^h coefficients based on the extended local projection (3.3), using $(\mathcal{M}_{it-1}-\overline{\mathcal{M}}_i)$ instead of $(\mathcal{M}_{it}-\overline{\mathcal{M}}_i)$. The β_1^h estimates are standardized to capture the differential response (approx. in p.p.) to a one standard deviation increase in $\varepsilon_t^{\text{mp}}$ associated with a one standard deviation higher $\mathcal{M}_{it-1}$ and $(\mathcal{M}_{it-1}-\overline{\mathcal{M}}_i)$, respectively. Shaded areas indicate 95% confidence bands two-way clustered by firms and quarters.

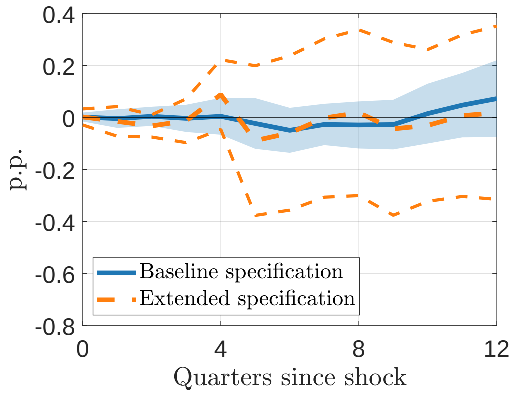
Figure 3.B.4: Differential investment response associated with higher Compustat maturing debt share



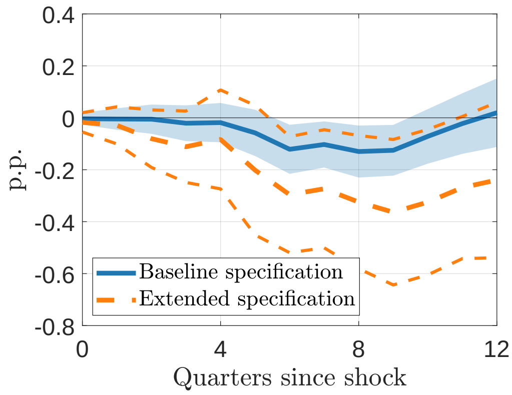
Notes: Panel (a) shows the estimated β_1^h coefficients based on the baseline local projection (3.2), using $\widetilde{\mathcal{M}}_{it}$ instead of $\mathcal{M}_{it}$. Panel (b) shows the estimated β_1^h coefficients based on the extended local projection (3.3), using $(\widetilde{\mathcal{M}}_{it}-\overline{\widetilde{\mathcal{M}}}_i)$ instead of replacing $(\mathcal{M}_{it}-\overline{\mathcal{M}}_i)$. The variable $\widetilde{\mathcal{M}}_{it} = (\text{current liabilities})_{it}/(\text{total debt})_{it-1}$ measures maturing debt based on Compustat data only. The β_1^h estimates are standardized to capture the differential response (approx. in p.p.) to a one standard deviation increase in $\varepsilon_t^{\text{mp}}$ associated with a one standard deviation higher $\widetilde{\mathcal{M}}_{it}$ and $(\widetilde{\mathcal{M}}_{it}-\overline{\widetilde{\mathcal{M}}}_i)$, respectively. Shaded areas indicate 95% confidence bands two-way clustered by firms and quarters.

Figure 3.B.5: Differential investment response associated with higher maturing bond share including callable bonds or bonds with variable coupon

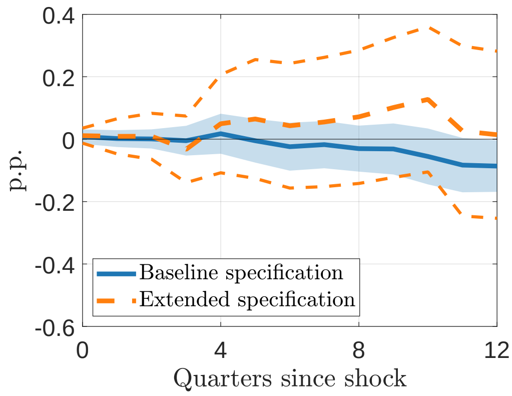
(a) $\mathcal{M}_{it}$ including only callable bonds



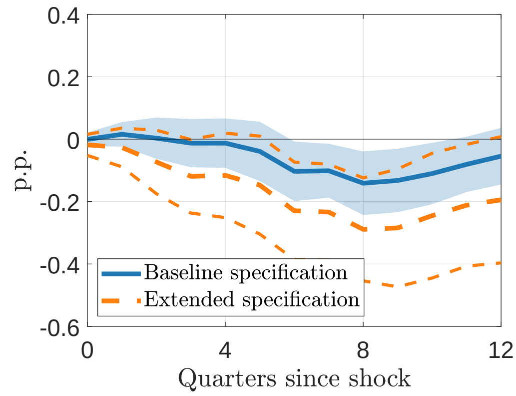
(b) $\mathcal{M}_{it}$ including callable and non-callable bonds



(c) $\mathcal{M}_{it}$ including only variable coupon bonds

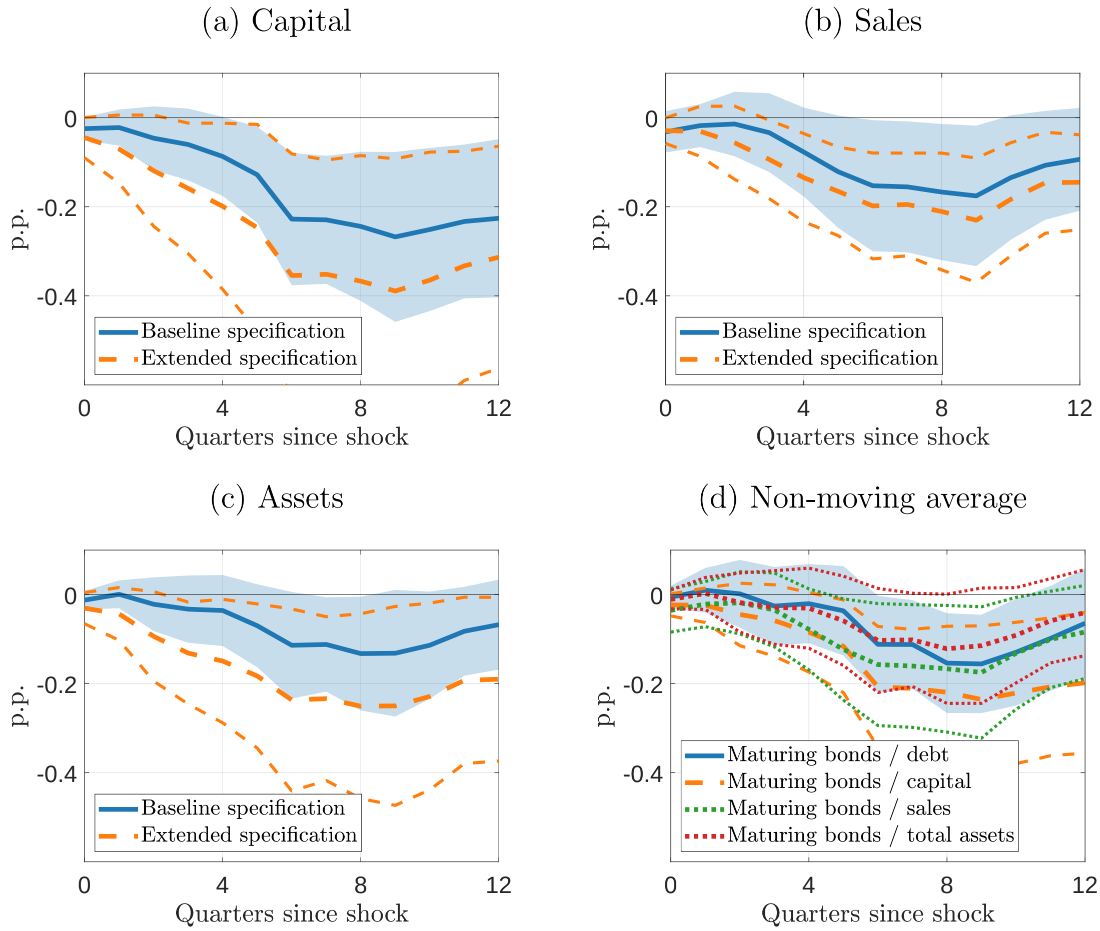


(d) $\mathcal{M}_{it}$ including variable coupon bonds and fixed coupon bonds



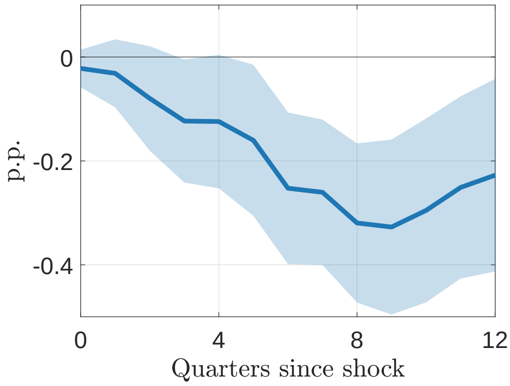
Notes: The figure shows the estimated β_1^h coefficients based on the baseline local projection (3.2) (solid lines) and extended local projection (3.3) (dashed lines), for various alternative definitions of the maturing bond share $\mathcal{M}_{it}$. In our main findings, $\mathcal{M}_{it}$ includes only non-callable fixed coupon bonds. In panel (a), we re-define $\mathcal{M}_{it}$ based on callable (fixed coupon) bonds. In panel (b), we include both callable and non-callable (fixed coupon) bonds. In panel (c), we re-define $\mathcal{M}_{it}$ based on variable coupon (non-callable) bonds. In panel (d), we include both variable coupon and fixed coupon (non-callable) bonds. The β_1^h estimates are standardized to capture the differential response (approx. in p.p.) to a one standard deviation increase in $\varepsilon_t^{\text{mp}}$ associated with a one standard deviation higher $\mathcal{M}_{it}$. Shaded areas indicate 95% confidence bands two-way clustered by firms and quarters.

Figure 3.B.6: Differential investment response associated with $\mathcal{M}_{it}$ using alternative denominators



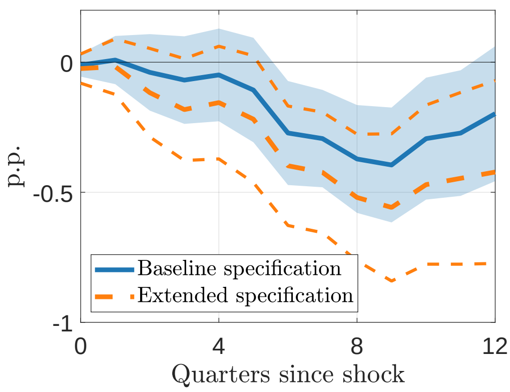
Notes: In panels (a) to (c) the figure shows the estimated β_1^h coefficients based on the baseline local projection (3.2) (solid lines) and extended local projection (3.3) (dashed lines), for various alternative definitions of $\mathcal{M}_{it}$. In panel (a), we re-define $\mathcal{M}_{it}$ as the ratio of maturing bonds over the average capital stock in the preceding four quarters, in (b) the denominator is average sales, in (c) average assets. In panel (d) the figure shows the estimated β_1^h coefficients based on the baseline local projection (3.2) using as denominator debt, capital, sales, or assets in the preceding quarter, instead of constructing a moving average. The β_1^h estimates are standardized to capture the differential response (approx. in p.p.) to a one standard deviation increase in $\varepsilon_t^{\mathrm{mp}}$ associated with a one standard deviation higher $\mathcal{M}_{it}$ for baseline specifications and and $(\mathcal{M}_{it} - \overline{\mathcal{M}}_i)$ for extended specifications. Shaded areas indicate 95% confidence bands two-way clustered by firms and quarters.

Figure 3.B.7: Differential investment response associated with $\mathcal{M}_{it}$ when including firm age as control variable



Notes: The figure shows the estimated β_1^h coefficients based on the extended local projection (3.3) where here Z_{it} additionally includes firm age. The β_1^h estimates are standardized to capture the differential response (approx. in p.p.) to a one standard deviation increase in $\varepsilon_t^{\text{mp}}$ associated with a one standard deviation higher $(\mathcal{M}_{it} - \overline{\mathcal{M}}_i)$. Shaded areas indicate 95% confidence bands two-way clustered by firms and quarters.

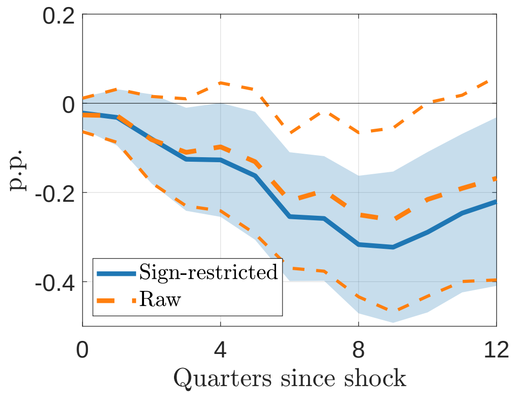
Figure 3.B.8: Differential investment response associated with $\mathcal{M}_{it}$, based on book value of capital



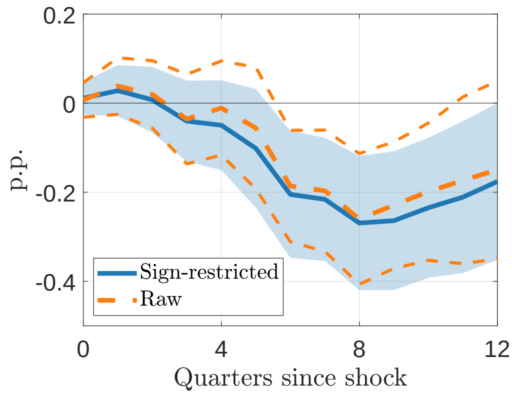
Notes: The figure shows the estimated β_1^h coefficients based on the baseline local projection (3.2) (solid lines) and extended local projection (3.3) (dashed lines), using book value of capital (deflated net fixed assets) instead of capital stocks constructed using a perpetual inventory method. The β_1^h estimates are standardized to capture the differential response (approx. in p.p.) to a one standard deviation increase in $\varepsilon_t^{\text{mp}}$ associated with a one standard deviation higher $\mathcal{M}_{it}$ for the baseline specification and $(\mathcal{M}_{it} - \overline{\mathcal{M}}_i)$ for the extended specification, respectively. Shaded areas indicate 95% confidence bands two-way clustered by firms and quarters.

Figure 3.B.9: Differential investment response associated with $\mathcal{M}_{it}$ for alternative monetary policy shocks

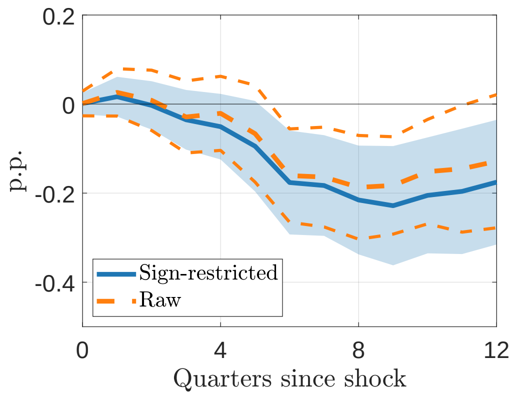
(a) 1-quarter federal funds future



(b) 2-quarter eurodollar future



(c) 3-quarter eurodollar future

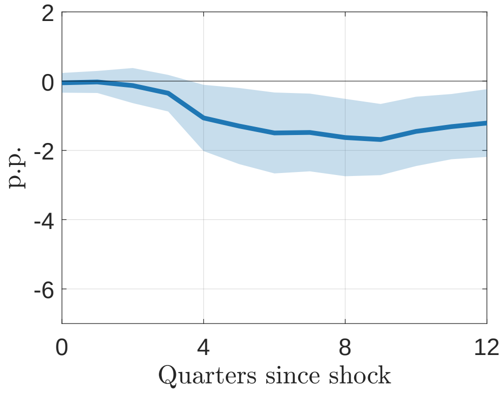


(d) 4-quarter eurodollar future

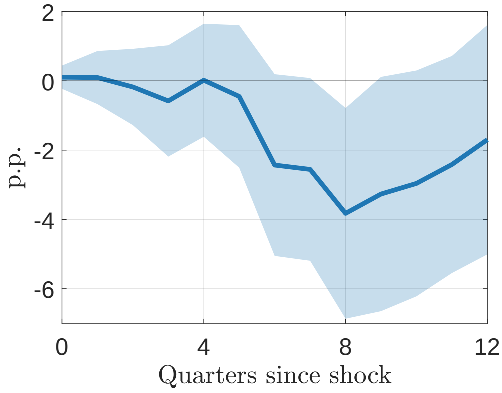
Notes: The figure shows the estimated β_1^h coefficients based on the baseline local projection (3.2) using various alternative monetary policy shocks $\varepsilon_t^{\text{mp}}$. In panel (a), $\varepsilon_t^{\text{mp}}$ is the surprise change (in a 30 minute window around regular FOMC meetings) in the one-quarter ahead federal funds future, in (b) the two-quarter ahead eurodollar future, in (c) the three-quarter ahead eurodollar future, and in (d) the four-quarter ahead eurodollar future. Solid lines show the responses based on sign-restricted shocks and dashed lines additionally show the responses based on raw surprises. The β_1^h estimates are standardized to capture the differential response (approx. in p.p.) to a one standard deviation increase in $\varepsilon_t^{\text{mp}}$ associated with a one standard deviation higher $\mathcal{M}_{it}$. Shaded areas indicate 95% confidence bands two-way clustered by firms and quarters.

Figure 3.B.10: Differential investment response associated with $\mathcal{M}_{it}$ using dummy specification of bond maturity

(a) Differential effect of $\mathcal{M}_{it} > 0$



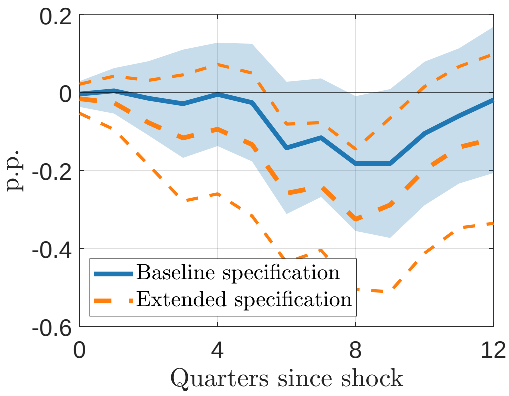
(b) Differential effect of $\mathcal{M}_{it} > 15$



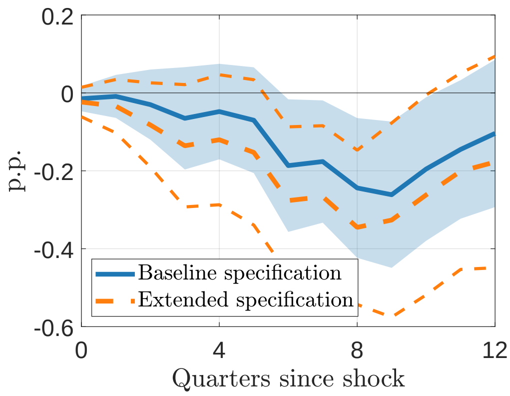
Notes: The figure shows the estimated β_1^h coefficients based on the baseline local projection (3.2), using $\mathbb{1}\{\mathcal{M}_{it} > 0\}$ instead of $\mathcal{M}_{it}$ in panel (a) and $\mathbb{1}\{\mathcal{M}_{it} > 15\}$ instead of $\mathcal{M}_{it}$ in panel (b). The β_1^h estimates are standardized to capture the differential response (approx. in p.p.) to a one standard deviation increase in $\varepsilon_t^{\text{mp}}$ associated with, respectively, $\mathcal{M}_{it} > 0$ and $\mathcal{M}_{it} > 15$ (i.e., 15 % of debt). Shaded areas indicate 95% confidence bands two-way clustered by firms and quarters.

Figure 3.B.11: Differential investment response associated with $\mathcal{M}_{it}$ based on alternative samples

(a) Pre-Great Recession



(b) Exclude Great Recession



Notes: The figure shows the estimated β_1^h coefficients based on the baseline local projection (3.2) (solid lines) and extended local projection (3.3) (dashed lines), using alternative samples. Panel (a) uses only monetary policy shocks until 2008Q2. Panel (b) excludes monetary policy shocks between 2008Q3 and 2009Q2. The β_1^h estimates are standardized to capture the differential response (approx. in p.p.) to a one standard deviation increase in $\varepsilon_t^{\text{mp}}$ associated with a one standard deviation higher $\mathcal{M}_{it}$ for the baseline specification and $(\mathcal{M}_{it} - \overline{\mathcal{M}}_i)$ for the extended specification, respectively. Shaded areas indicate 95% confidence bands two-way clustered by firms and quarters.

3.C Model

In this section we provide additional details of the model set up in Section 3.3 (Appendix 3.C.1) and derive the first-order conditions presented in Section 3.4 (Appendix 3.C.2).

3.C.1 Model: Details

Production firms' labor demand. A production firm i enters period t with productivity z_{it} and capital k_{it}. Given the price of undifferentiated output p_t and the wage rate w_t, optimal labor demand l_{it} solves a simple static maximization problem. The first-order condition with respect to l_{it} in (3.2) is:

$$l_{it} = \left(\frac{\zeta(1-\psi) p_t z_{it} k_{it}^{\psi\zeta}}{w_t} \right)^{\frac{1}{1-\zeta(1-\psi)}} \tag{3.C.1}$$

This implies that firm revenue net of labor costs is

$$\max_{l_{it}} \quad p_t z_{it} \left(k_{it}^{\psi} l_{it}^{1-\psi} \right)^{\zeta} - w_t l_{it} = A_{it} k_{it}^{\alpha}, \tag{3.C.2}$$

where

$$A_{it} \equiv (p_t z_{it})^{\frac{1}{1-\zeta(1-\psi)}} [1 - \zeta(1-\psi)] \left(\frac{\zeta(1-\psi)}{w_{it}} \right)^{\frac{\zeta(1-\psi)}{1-\zeta(1-\psi)}} \quad \text{and} \quad \alpha \equiv \frac{\zeta\psi}{1-\zeta(1-\psi)}. \tag{3.C.3}$$

This is used in equation (3.3).

Retail firms. Retailer $j \in [0,1]$ buys y_{jt} units of undifferentiated goods from production firms at price p_t and converts them into a quantity $\tilde{y}_{jt}$ of differentiated retail goods which is sold to the final goods sector at price $\tilde{p}_{jt}$. Period profits are

$$\tilde{p}_{jt}\tilde{y}_{jt} - p_t y_{jt} - \frac{\lambda}{2} \left(\frac{\tilde{p}_{jt}}{\tilde{p}_{jt-1}} - 1 \right)^2 Y_t. \tag{3.C.4}$$

Rotemberg-style costs of price adjustment are expressed as a fraction of aggregate real output Y_t. Retail goods are bought by a perfectly competitive final goods sector which produces final goods Y_t at constant returns to scale:

$$Y_t = \left[\int_0^1 \tilde{y}_{jt}^{\frac{\rho-1}{\rho}} dj \right]^{\frac{\rho}{\rho-1}}, \quad \text{where} \quad \rho > 1. \tag{3.C.5}$$

Profit maximization in the final goods sector yields a downward sloping demand curve for variety j:

$$\tilde{y}_{jt} = \left(\frac{P_t}{\tilde{p}_{jt}} \right)^{\rho} Y_t, \quad \text{with} \quad P_t = \left[\int_0^1 \tilde{p}_{jt}^{1-\rho} dj \right]^{\frac{1}{1-\rho}} \tag{3.C.6}$$

Imperfect substitutability among different varieties gives each retailer some amount of market power. Optimal dynamic price setting by retailer j gives the following first-order condition for $\tilde{p}_{jt}$:

$$\tilde{y}_{jt} - \rho\left(\frac{\tilde{p}_{jt} - p_t}{\tilde{p}_{jt}}\right)\tilde{y}_{jt} - \lambda\frac{Y_t}{\tilde{p}_{jt-1}}\left(\frac{\tilde{p}_{jt}}{\tilde{p}_{jt-1}} - 1\right) + \mathbb{E}_{S_{t+1}|S_t}\Lambda_{t,t+1}\lambda\frac{Y_{t+1}}{\tilde{p}_{jt}}\left(\frac{\tilde{p}_{jt+1}}{\tilde{p}_{jt}} - 1\right)\frac{\tilde{p}_{jt+1}}{\tilde{p}_{jt}} = 0 \tag{3.C.7}$$

From symmetry ($\tilde{p}_{jt} = P_t$ and $\tilde{y}_{jt} = Y_t$), it follows that

$$1 - \rho\left(\frac{P_t - p_t}{P_t}\right) - \lambda\frac{1}{P_{t-1}}\left(\frac{P_t}{P_{t-1}} - 1\right) + \mathbb{E}_{S_{t+1}|S_t}\Lambda_{t,t+1}\lambda\frac{Y_{t+1}}{P_t Y_t}\left(\frac{P_{t+1}}{P_t} - 1\right)\frac{P_{t+1}}{P_t} = 0. \tag{3.C.8}$$

The final good is the numéraire: $P_t = 1$. Using $\pi_t = P_t/P_{t-1}$ yields the New Keynesian Phillips Curve in (2.F.8).

Market clearing. Labor market clearing implies

$$L = \int_x l(x; S)\mu(x)dx. \tag{3.C.9}$$

The aggregate amount of final goods Y is

$$Y = \int_x y(x; S)\mu(x)dx. \tag{3.C.10}$$

Output net of fixed costs of operation and default costs is

$$Y^{net} \equiv Y - \int_x \left[f + \xi\int_\varepsilon \mathcal{D}(x,\varepsilon; S)\underline{q}(x,\varepsilon; S)\varphi(\varepsilon)d\varepsilon\right]\mu(x)dx. \tag{3.C.11}$$

Final goods market clearing implies that

$$Y^{net} = C + \mathcal{G} + \mathcal{H} + I, \tag{3.C.12}$$

where C is aggregate consumption, and $\mathcal{G}$ and $\mathcal{H}$ are aggregate equity and debt issuance costs. Aggregate equity issuance costs are

$$\begin{aligned}\mathcal{G} = &\int_x\int_\varepsilon\int_{z'} G\left(e(x,\varepsilon,z'; S)\right)\Pi(z'|z)dz'(1-\kappa)[1-\mathcal{D}(x,\varepsilon; S)]\varphi(\varepsilon)d\varepsilon\mu(x)dx \\ &+ \int_{x'} \tilde{G}(x'; S)\mathcal{E}(x'; S)dx',\end{aligned} \tag{3.C.13}$$

where $\tilde{G}(x'; S)$ is equity issuance costs of entrants starting in state x'. Aggregate debt issuance costs are

$$\begin{aligned}\mathcal{H} = &\int_x\int_\varepsilon\int_{z'} H\left(b^{S'}(x,\varepsilon,z'; S), b^{L'}(x,\varepsilon,z'; S), b^L(x)/\pi\right) \\ &\cdot \Pi(z'|z)dz'(1-\kappa)[1-\mathcal{D}(x,\varepsilon; S)]\varphi(\varepsilon)d\varepsilon\mu(x)dx + \int_{x'} \tilde{H}(x'; S)\mathcal{E}(x'; S)dx',\end{aligned} \tag{3.C.14}$$

where $\tilde{H}(x';S)$ is debt issuance costs of entrants starting in state x'. Aggregate investment I follows from (3.14):

$$I = K\left[\frac{\phi-1}{\phi}\delta^{-\frac{1}{\phi}}\left(\frac{K'}{K} - 1 + \delta\frac{\phi}{\phi-1}\right)\right]^{\frac{\phi}{\phi-1}} \tag{3.C.15}$$

Capital goods market clearing implies:

$$K = \int_x k(x)\mu(x)dx, \quad \text{and} \quad K' = \int_{x'} k'(x')\mu(x')dx' \tag{3.C.16}$$

Finally, GDP is equal to $C + I$.

3.C.2 Characterization: Details

To derive the first-order conditions in Section 3.4.1 we express the firm problem (3.20) in terms of three choice variables: the scale of production k', and the amounts of short-term debt $b^{S'}$ and long-term debt $b^{L'}$:

$$\begin{aligned} W^C(x,\varepsilon,z';S) = q(x,\varepsilon;S) - Qk' + b^{S'}p^S + \left(b^{L'} - \frac{(1-\gamma)b^L}{\pi}\right)p^L - G(e) - H\left(b^{S'}, b^{L'}, \frac{b^L}{\pi}\right) \\ + \mathbb{E}_{S'|S}\Lambda\int_{\varepsilon'} V(x',\varepsilon';S')\varphi(\varepsilon')d\varepsilon', \end{aligned} \tag{3.C.17}$$

where $x = (z,k,b^S,b^L)$ and the real market value of firm assets $q(x,\varepsilon;S)$ is specified in (3.3). The firm's short-term bond price p^S is

$$\begin{aligned} p^S =& \mathbb{E}_{S'|S}\Lambda\int_{\varepsilon'}\Bigg[[1-\mathcal{D}(x',\varepsilon';S')]\frac{1+c}{\pi'} \\ &+ \mathcal{D}(x',\varepsilon';S')\frac{1-\xi}{b^{S'}+b^{L'}}\left[Q'k' + (1-\tau)\left[A'k'^{\alpha} + (\varepsilon'-\delta)Q'k' - f\right]\right]\Bigg]\varphi(\varepsilon')d\varepsilon', \end{aligned} \tag{3.C.18}$$

where $\mathcal{D}(x',\varepsilon';S') = 1$ iff $W(x',\varepsilon';S') < 0$ in (3.20), and $x' = (z',k',b^{S'},b^{L'})$. The long-term bond price p^L is

$$\begin{aligned} p^L =& \mathbb{E}_{S'|S}\Lambda\int_{\varepsilon'}\Bigg[[1-\mathcal{D}(x',\varepsilon';S')]\frac{\gamma + c + (1-\gamma)\mathbb{E}_{z''|z'}g(x',\varepsilon',z'';S')}{\pi'} \\ &+ \mathcal{D}(x',\varepsilon';S')\frac{1-\xi}{b^{S'}+b^{L'}}\left[Q'k' + (1-\tau)\left[A'k'^{\alpha} + (\varepsilon'-\delta)Q'k' - f\right]\right]\Bigg]\varphi(\varepsilon')d\varepsilon'. \end{aligned} \tag{3.C.19}$$

It follows that both p^S and p^L are functions of k', $b^{S'}$, and $b^{L'}$. Equity issuance costs are

$$G(e) = \nu\left(\max\{e,0\}\right)^2, \quad \text{where:} \quad e = Qk' - q(x,\varepsilon;S) - b^{S'}p^S - \left(b^{L'} - \frac{(1-\gamma)b^L}{\pi}\right)p^L. \tag{3.C.20}$$

Debt issuance costs are

$$H\left(b^{S'}, b^{L'}, \frac{b^L}{\pi}\right) = \eta\left(b^{S'} + \max\left\{b^{L'} - \frac{(1-\gamma)b^L}{\pi}, 0\right\}\right)^2. \tag{3.C.21}$$

It follows that the firm objective (3.C.17) is a function of the three choice variables k', $b^{S'}$, and $b^{L'}$.

First-order condition for capital. The firm's first-order condition with respect to capital k' is:

$$\begin{aligned}&\left[1 + \frac{\partial G(e)}{\partial e}\right]\left[-Q + b^{S'}\frac{\partial p^S}{\partial k'} + \left(b^{L'} - \frac{(1-\gamma)b^L}{\pi}\right)\frac{\partial p^L}{\partial k'}\right] \\ &+ \mathbb{E}_{S'|S}\Lambda\int_{\varepsilon'}[1 - \mathcal{D}(x', \varepsilon'; S')]\frac{\partial W(x', \varepsilon'; S')}{\partial k'}\varphi(\varepsilon')d\varepsilon' = 0\,,\end{aligned} \tag{3.C.22}$$

where

$$\frac{\partial W(x', \varepsilon'; S')}{\partial k'} = \frac{\partial q'}{\partial k'}\left[(1-\kappa)\mathbb{E}_{z''|z'}\left(1 + \frac{\partial G(e')}{\partial e'}\right) + \kappa\left(1 - \frac{(1-\gamma)b^{L'}}{\pi'}\mathbb{E}_{z''|z'}\frac{\partial \tilde{g}(q', b', z''; S')}{\partial q'}\right)\right], \tag{3.C.23}$$

and $$\frac{\partial q'}{\partial k'} = \left[Q' + (1-\tau)\left[A'\alpha k'^{\alpha-1} + (\varepsilon' - \delta)Q'\right]\right]. \tag{3.C.24}$$

Equation (3.C.23) uses the fact that the future price of long-term debt $g(x', \bar{\varepsilon}', z''; S')$ can be expressed as a function of the reduced state vector $(q', b', z''; S')$ (as explained in Section 3.5.1). Written in this way, the future price of long-term debt $\tilde{g}(q', b', z''; S')$ depends on the endogenous firm states

$$q' = q(x', \varepsilon'; S') = Q'k' - \frac{b^{S'}}{\pi'} - \frac{\gamma b^{L'}}{\pi'} + (1-\tau)\left[A'k'^{\alpha} + (\varepsilon' - \delta)Q'k' - f - \frac{c(b^{S'} + b^{L'})}{\pi'}\right] \tag{3.C.25}$$

and $b' = (1-\gamma)b^{L'}$. To compute $\partial p^S/\partial k'$ and $\partial p^L/\partial k'$ in (3.C.22), we first derive how k' affects the firm's default decision. Let $\bar{\varepsilon}'$ denote the threshold realization of the capital quality shock ε' such that $W(x', \bar{\varepsilon}'; S') = 0$ in (3.20). At this threshold realization $\bar{\varepsilon}'$, the firm is just indifferent between defaulting and servicing its current debt obligations, i.e.,

$$(1-\kappa)\mathbb{E}_{z''|z'}W^C(x', \bar{\varepsilon}', z''; S') + \kappa\left(q' - \frac{(1-\gamma)b^{L'}}{\pi'}\mathbb{E}_{z''|z'}\tilde{g}(q', b', z''; S')\right) = 0. \tag{3.C.26}$$

Applying the implicit function theorem to (3.C.26), we derive

$$\frac{\partial \bar{\varepsilon}'}{\partial k'} = -\frac{\frac{\partial q'}{\partial k'}}{\frac{\partial q'}{\partial \bar{\varepsilon}'}} = -\frac{Q' + (1-\tau)\left[A'\alpha k'^{\alpha-1} + (\bar{\varepsilon}' - \delta)Q'\right]}{(1-\tau)Q'k'}. \tag{3.C.27}$$

The derivative of p^S with respect to k' is then given by

$$\begin{aligned}\frac{\partial p^S}{\partial k'} =&\mathbb{E}_{S'|S}\Lambda\Bigg[\int_{-\infty}^{\bar{\varepsilon}'} \frac{1-\xi}{b^{S'}+b^{L'}}\left[Q'+(1-\tau)\left[A'\alpha k'^{\alpha-1}+(\varepsilon'-\delta)Q'\right]\right]\varphi(\varepsilon')d\varepsilon' \\ &+\varphi(\bar{\varepsilon}')\frac{\partial\bar{\varepsilon}'}{\partial k'}\left[-\frac{1+c}{\pi'}+\frac{(1-\xi)}{b^{S'}+b^{L'}}\left[Q'k'+(1-\tau)\left[A'k'^{\alpha}+(\bar{\varepsilon}'-\delta)Q'k'-f\right]\right]\right]\Bigg].\end{aligned} \tag{3.C.28}$$

It follows for the derivative of p^L with respect to k':

$$\begin{aligned}\frac{\partial p^L}{\partial k'} =&\mathbb{E}_{S'|S}\Lambda\Bigg[\int_{\bar{\varepsilon}'}^{\infty} \frac{1-\gamma}{\pi'}\mathbb{E}_{z''|z'}\frac{\partial\tilde{g}(q',b',z'';S')}{\partial q'}\frac{\partial q'}{\partial k'}\varphi(\varepsilon')d\varepsilon' \\ &+\int_{-\infty}^{\bar{\varepsilon}'}\frac{1-\xi}{b^{S'}+b^{L'}}\left[Q'+(1-\tau)\left[A'\alpha k'^{\alpha-1}+(\varepsilon'-\delta)Q'\right]\right]\varphi(\varepsilon')d\varepsilon' \\ &+\varphi(\bar{\varepsilon}')\frac{\partial\bar{\varepsilon}'}{\partial k'}\Bigg[-\frac{\gamma+c+(1-\gamma)\mathbb{E}_{z''|z'}\tilde{g}(q',b',z'';S')}{\pi'} \\ &+\frac{1-\xi}{b^{S'}+b^{L'}}\left[Q'k'+(1-\tau)\left[A'k'^{\alpha}+(\bar{\varepsilon}'-\delta)Q'k'-f\right]\right]\Bigg]\Bigg]\end{aligned} \tag{3.C.29}$$

First-order condition for short-term debt. The firm's first-order condition with respect to $b^{S'}$ is

$$\begin{aligned}&\left[1+\frac{\partial G(e)}{\partial e}\right]\left[p^S+b^{S'}\frac{\partial p^S}{\partial b^{S'}}+\left(b^{L'}-\frac{(1-\gamma)b^L}{\pi}\right)\frac{\partial p^L}{\partial b^{S'}}\right]-\frac{\partial H(b^{S'},b^{L'},b^L/\pi)}{\partial b^{S'}} \\ &+\mathbb{E}_{S'|S}\Lambda\int_{\bar{\varepsilon}'}^{\infty}\frac{\partial W(x',\varepsilon';S')}{\partial b^{S'}}\varphi(\varepsilon')d\varepsilon'=0\,,\end{aligned} \tag{3.C.30}$$

where

$$\frac{\partial W(x',\varepsilon';S')}{\partial b^{S'}}=\frac{\partial q'}{\partial b^{S'}}\left[(1-\kappa)\mathbb{E}_{z''|z'}\left(1+\frac{\partial G(e')}{\partial e'}\right)+\kappa\left(1-\frac{(1-\gamma)b^{L'}}{\pi'}\mathbb{E}_{z''|z'}\frac{\partial\tilde{g}(q',b',z'';S')}{\partial q'}\right)\right], \tag{3.C.31}$$

and $$\frac{\partial q'}{\partial b^{S'}}=-\frac{1+(1-\tau)c}{\pi'}. \tag{3.C.32}$$

The derivative of p^S with respect to $b^{S'}$ is

$$\begin{aligned}\frac{\partial p^S}{\partial b^{S'}} =&\mathbb{E}\Lambda\Bigg[-\int_{-\infty}^{\bar{\varepsilon}'}\frac{1-\xi}{(b^{S'}+b^{L'})^2}\left[Q'k'+(1-\tau)\left[A'k'^{\alpha}+(\varepsilon'-\delta)Q'k'-f\right]\right]\varphi(\varepsilon')d\varepsilon' \\ &+\varphi(\bar{\varepsilon}')\frac{\partial\bar{\varepsilon}'}{\partial b^{S'}}\left[-\frac{1+c}{\pi'}+\frac{1-\xi}{b^{S'}+b^{L'}}\left[Q'k'+(1-\tau)\left[A'k'^{\alpha}+(\bar{\varepsilon}'-\delta)Q'k'-f\right]\right]\right]\Bigg],\end{aligned} \tag{3.C.33}$$

where $$\frac{\partial\bar{\varepsilon}'}{\partial b^{S'}}=-\frac{\frac{\partial q'}{\partial b^{S'}}}{\frac{\partial q'}{\partial\bar{\varepsilon}'}}=\frac{1+(1-\tau)c}{\pi'(1-\tau)Q'k'}. \tag{3.C.34}$$

Finally, we derive the derivative of p^L with respect to $b^{S'}$:

$$\begin{aligned}\frac{\partial p^L}{\partial b^{S'}} =&\mathbb{E}_{S'|S}\Lambda\Bigg[\int_{\bar{\varepsilon}'}^{\infty}\frac{1-\gamma}{\pi'}\mathbb{E}_{z''|z'}\frac{\partial\tilde{g}(q',b',z'';S')}{\partial q'}\frac{\partial q'}{\partial b^{S'}}\varphi(\varepsilon')d\varepsilon' \\ &-\int_{-\infty}^{\bar{\varepsilon}'}\frac{1-\xi}{(b^{S'}+b^{L'})^2}\left[Q'k'+(1-\tau)\left[A'k'^{\alpha}+(\varepsilon'-\delta)Q'k'-f\right]\right]\varphi(\varepsilon')d\varepsilon' \\ &+\varphi(\bar{\varepsilon}')\frac{\partial\bar{\varepsilon}'}{\partial b^{S'}}\Bigg[-\frac{\gamma+c+(1-\gamma)\mathbb{E}_{z''|z'}\tilde{g}(q',b',z'';S')}{\pi'} \\ &+\frac{1-\xi}{b^{S'}+b^{L'}}\left[Q'k'+(1-\tau)\left[A'k'^{\alpha}+(\bar{\varepsilon}'-\delta)Q'k'-f\right]\right]\Bigg]\Bigg] \end{aligned} \tag{3.C.35}$$

First-order condition for long-term debt. The firm's first-order condition with respect to $b^{L'}$ is

$$\begin{aligned}&\left[1+\frac{\partial G(e)}{\partial e}\right]\left[p^L+b^{S'}\frac{\partial p^S}{\partial b^{L'}}+\left(b^{L'}-\frac{(1-\gamma)b^L}{\pi}\right)\frac{\partial p^L}{\partial b^{L'}}\right]-\frac{\partial H(b^{S'},b^{L'},b^L/\pi)}{\partial b^{L'}} \\ &+\mathbb{E}_{S'|S}\Lambda\int_{\bar{\varepsilon}'}^{\infty}\frac{\partial W(x',\varepsilon';S')}{\partial b^{L'}}\varphi(\varepsilon')d\varepsilon'=0\,,\end{aligned} \tag{3.C.36}$$

where

$$\begin{aligned}\frac{\partial W(x',\varepsilon';S')}{\partial b^{L'}}=&\frac{\partial q'}{\partial b^{L'}}\left[(1-\kappa)\mathbb{E}_{z''|z'}\left(1+\frac{\partial G(e')}{\partial e'}\right)+\kappa\left(1-\frac{(1-\gamma)b^{L'}}{\pi'}\mathbb{E}_{z''|z'}\frac{\partial\tilde{g}(q',b',z'';S')}{\partial q'}\right)\right] \\ &+\frac{\partial b'}{\partial b^{L'}}\mathbb{E}_{z''|z'}\left[(1-\kappa)\frac{\partial\tilde{W}^C(q',b',z'';S')}{\partial b'}-\frac{\kappa}{\pi'}\left(\tilde{g}(q',b',z'';S')+b'\frac{\partial\tilde{g}(q',b',z'';S')}{\partial b'}\right)\right],\end{aligned} \tag{3.C.37}$$

$$\text{with}\quad \frac{\partial q'}{\partial b^{L'}}=-\frac{\gamma+(1-\tau)c}{\pi'}\quad\text{and}\quad\frac{\partial b'}{\partial b^{L'}}=1-\gamma. \tag{3.C.38}$$

Equation (3.C.37) uses the fact that the future value $W^C(x',\bar{\varepsilon}',z'';S')$ can be expressed as a function of the reduced state vector $\tilde{W}^C(q',b',z'';S')$ (as explained in Section 3.5.1). The derivative of p^S with respect to $b^{L'}$ is

$$\begin{aligned}\frac{\partial p^S}{\partial b^{L'}}=&\mathbb{E}_{S'|S}\Lambda\Bigg[-\int_{-\infty}^{\bar{\varepsilon}'}\frac{1-\xi}{(b^{S'}+b^{L'})^2}\left[Q'k'+(1-\tau)\left[A'k'^{\alpha}+(\varepsilon'-\delta)Q'k'-f\right]\right]\varphi(\varepsilon')d\varepsilon' \\ &+\varphi(\bar{\varepsilon}')\frac{\partial\bar{\varepsilon}'}{\partial b^{L'}}\left[-\frac{1+c}{\pi'}+\frac{1-\xi}{b^{S'}+b^{L'}}\left[Q'k'+(1-\tau)\left[A'k'^{\alpha}+(\bar{\varepsilon}'-\delta)Q'k'-f\right]\right]\right]\Bigg],\end{aligned} \tag{3.C.39}$$

where

$$\frac{\partial\bar{\varepsilon}'}{\partial b^{L'}}=-\frac{\frac{\partial q'}{\partial b^{L'}}}{\frac{\partial q'}{\partial\bar{\varepsilon}'}}-\frac{\frac{\partial b'}{\partial b^{L'}}}{\frac{\partial q'}{\partial\bar{\varepsilon}'}}\frac{\mathbb{E}_{z''|z'}\left[(1-\kappa)\frac{\partial\tilde{W}^C(q',b',z'';S')}{\partial b'}-\frac{\kappa}{\pi'}\left(\tilde{g}(q',b',z'';S')+b'\frac{\partial\tilde{g}(q',b',z'';S')}{\partial b'}\right)\right]}{\left[(1-\kappa)\mathbb{E}_{z''|z'}\left(1+\frac{\partial G(e')}{\partial e'}\right)+\kappa\left(1-\frac{b'}{\pi'}\mathbb{E}_{z''|z'}\frac{\partial\tilde{g}(q',b',z'';S')}{\partial q'}\right)\right]}. \tag{3.C.40}$$

Similarly, we derive the derivative of p^L with respect to $b^{L'}$:

$$\begin{aligned}\frac{\partial p^L}{\partial b^{L'}} =&\mathbb{E}\Lambda\Bigg[\int_{\bar\varepsilon'}^{\infty}\frac{1-\gamma}{\pi'}\mathbb{E}_{z''|z'}\left(\frac{\partial\tilde g(q',b',z'';S')}{\partial q'}\frac{\partial q'}{\partial b^{L'}}+\frac{\partial\tilde g(q',b',z'';S')}{\partial b'}\frac{\partial b'}{\partial b^{L'}}\right)\varphi(\varepsilon')d\varepsilon'\\ &-\int_{-\infty}^{\bar\varepsilon'}\frac{1-\xi}{(b^{S'}+b^{L'})^2}\left[Q'k'+(1-\tau)\left[A'k'^{\alpha}+(\varepsilon'-\delta)Q'k'-f\right]\right]\varphi(\varepsilon')d\varepsilon'\\ &+\varphi(\bar\varepsilon')\frac{\partial\bar\varepsilon'}{\partial b^{L'}}\Bigg[-\frac{\gamma+c+(1-\gamma)\mathbb{E}_{z''|z'}\tilde g(q',b',z'';S')}{\pi'}\\ &+\frac{1-\xi}{b^{S'}+b^{L'}}\left[Q'k'+(1-\tau)\left[A'k'^{\alpha}+(\bar\varepsilon'-\delta)Q'k'-f\right]\right]\Bigg]\Bigg]\end{aligned} \tag{3.C.41}$$

The effect of a marginal increase in b' on $\tilde W^C(q',b',z'';S')$ in (3.C.37) can be derived using (3.C.17):

$$\frac{\partial\tilde W^C(q,b,z';S)}{\partial b}=\frac{\partial W(x',\varepsilon';S')}{\partial(1-\gamma)b^L}=-\frac{1}{\pi}\frac{\partial H(b^{S'},b^{L'},b^L/\pi)}{\partial\frac{(1-\gamma)b^L}{\pi}}-\frac{p^L}{\pi}\left[1+\frac{\partial G(e)}{\partial e}\right] \tag{3.C.42}$$

Iterating forward one time period, this implies

$$\frac{\partial\tilde W^C(q',b',z'';S')}{\partial b'}=-\frac{1}{\pi'}\left(\frac{\partial H(\tilde b^{S''},\tilde b^{L''},b^{L'}/\pi')}{\partial\frac{(1-\gamma)b^{L'}}{\pi'}}+\tilde g\left(q',b',z'';S'\right)\left[1+\frac{\partial G(e')}{\partial e'}\right]\right). \tag{3.C.43}$$

3.D Quantitative results

This section of the appendix complements the quantitative analysis in Section 3.5. We define a number of moments used in the model (Appendix 3.D.1), give more details on the empirical moments used (Appendix 3.D.2), present additional steady state results (Appendix 3.D.3), provide details on the analysis of the heterogenous effects of monetary policy shocks (Appendix 3.D.4), and decribe the models used to highlight the importance of heterogeneous debt maturity (Appendix 3.D.5).

3.D.1 Model moments

The total amount of firm debt is the sum of future principal payments:

$$b^S+\gamma b^L+(1-\gamma)\gamma b^L+(1-\gamma)^2\gamma b^L+...=b^S+\gamma b^L\sum_{j=0}^{\infty}(1-\gamma)^j=b^S+b^L. \tag{3.D.1}$$

Firm leverage (total debt over total assets) is given by $(b^S+b^L)/k$.

In Table 3.2, we target the share of debt due within one year:

$$\frac{b^S_{it}+\gamma b^L_{it}+(1-\gamma)\gamma b^L_{it}+(1-\gamma)^2\gamma b^L_{it}+(1-\gamma)^3\gamma b^L_{it}}{\frac{1}{4}\sum_{j=0}^{3}\left(b^S_{it-j}+b^L_{it-j}\right)}. \tag{3.D.2}$$

As in the empirical part of the paper, we use a four-quarter moving average of debt in the denominator.[32] For firms which are younger than four quarters, the denominator is average debt over the maximum number of past quarters available.

The maturing bond share $\mathcal{M}$ from (3.4) measures the share of total debt which matures within one quarter:

$$\mathcal{M}_{it} = \frac{b_{it}^S + \gamma b_{it}^L}{b_{it}^S + b_{it}^L}. \tag{3.D.3}$$

In Figures 3.3 and 3.5, we use average total debt over the preceding four quarters (as in (3.D.2)) as denominator of $\mathcal{M}_{it}$ to be consistent with the empirical specification in Section 3.2. All model results are virtually indistinguishable when using the current level of debt as the denominator instead.

The Macaulay duration of long-term debt is the weighted average term to maturity of the cash flow from a riskless bond divided by its steady state market price:

$$\mu = \frac{1}{P_r^L} \sum_{j=1}^{\infty} j(1-\gamma)^{j-1} \frac{c+\gamma}{(1+r^*)^j} = \frac{c+\gamma}{P_r^L} \frac{1+r^*}{(\gamma+r^*)^2}, \tag{3.D.4}$$

where P_r^L is the price of a riskless nominal long-term bond:

$$P_r^L = \mathbb{E} \sum_{j=1}^{\infty} (1-\gamma)^{j-1} \frac{c+\gamma}{(1+i)^j} \tag{3.D.5}$$

In steady state ($i = r^*$), this implies that $P_r^L = (c+\gamma)/(r^*+\gamma)$ with Macaulay duration

$$\mu = \frac{1+r^*}{\gamma+r^*}. \tag{3.D.6}$$

The credit spread on short-term debt compares the annualized gross return from buying a firm's nominal short-term debt (in the absence of default) to the annualized gross return from buying riskless nominal short-term debt:

$$spr^S \equiv \left(\frac{1+c}{p^S}\right)^4 - \left(\frac{1+c}{P_r^S}\right)^4, \tag{3.D.7}$$

where P_r^S is the price of a riskless short-term bond: $P_r^S = (1+c)/(1+i)$.

The credit spread on long-term debt compares the annualized gross return from buying a firm's nominal long-term debt (in the absence of default and assuming constant p^L) to the annualized gross return from buying riskless nominal long-term debt:

$$\begin{aligned} spr^L &\equiv \left(\frac{\gamma+c+(1-\gamma)p^L}{p^L}\right)^4 - \left(\frac{\gamma+c+(1-\gamma)P_r^L}{P_r^L}\right)^4 \\ &= \left(\frac{\gamma+c}{p^L}+1-\gamma\right)^4 - \left(\frac{\gamma+c}{P_r^L}+1-\gamma\right)^4 \end{aligned} \tag{3.D.8}$$

[32]Note that b_{it}^S and b_{it}^L denote debt levels chosen at the end of period $t-1$ and outstanding at the beginning of period t.

The average credit spread used in Figure 3.1 is defined as

$$\frac{b^{S'}}{b^{S'}+b^{L'}}spr^S + \frac{b^{L'}}{b^{S'}+b^{L'}}spr^L. \tag{3.D.9}$$

Equity issuance of firm i at time t is the average of quarterly equity issuance over the preceding four quarters relative to firm assets:

$$\frac{1}{4}\cdot(\max\{0,e_{it}\}+\max\{0,e_{it-1}\}+\max\{0,e_{it-2}\}+\max\{0,e_{it-3}\})\cdot\frac{1}{k_{it}} \tag{3.D.10}$$

We use an average of quarterly equity issuance over four quarters to be consistent with the empirical moment used in Table 3.2.

Firm capital growth is $\log(k_{it})-\log(k_{it-1})$. The capital growth moments in Table 3.2 are medians across firm-level averages and standard deviations of quarterly firm-level capital growth.

The firm exit rate is total exit (endogenous through default and exogenous) per quarter:

$$\int_x\int_\varepsilon \mathcal{D}(x,\varepsilon;S)\varphi(\varepsilon)d\varepsilon\mu(x)dx+\kappa\left(1-\int_x\int_\varepsilon \mathcal{D}(x,\varepsilon;S)\varphi(\varepsilon)d\varepsilon\mu(x)dx\right) \tag{3.D.11}$$

Finally, the value of firm entry is $W^C(x,\varepsilon,z';S)$ for the firm state corresponding to $q=0$, $b=0$, and $z'=z^e$.

3.D.2 Empirical moments

In this section, we provide details on the empirical moments used in Table 3.2. As described in Section 3.2, we use quarterly firm-level balance sheet data from Compustat and FISD bond-level information. The time sample is 1995–2017. We exclude firms that are not incorporated in the U.S. and we delete firms in the public administration, finance, insurance, real estate, and utilities sectors. Negative observations of total assets (`atq`), fixed assets (`ppegtq` and `ppentq`), and short-term and long-term debt (`dlcq`, `dlttq`) are set to missing.

Firm leverage is total debt (`dlcq`+`dlttq`) divided by assets (`atq`). The share of debt due within one year is debt in current liabilities (`dlcq`) divided by the moving average of total firm debt (`dlcq`+`dlttq`) over the last four quarters. This procedure smoothes out seasonal factors and other transitory fluctuations. If less than four past quarters of total debt are available, we use average debt over the maximum number of past quarters available as denominator. The credit spread on long-term debt is constructed using firm-level credit ratings combined with rating-specific corporate bond spreads, following Arellano et al. (2019). We use quarterly Standard & Poor's credit ratings from Compustat Monthly Updates. Based on this rating, each firm-quarter is assigned the time-varying median spread of the corresponding rating class from the FISD data. Because FISD data only includes bonds with maturity above one quarter, this data is informative with respect to long-term credit spreads in our model. See Jungherr and Schott (2021) for details on the construction of time-varying rating-specific credit spreads using FISD data.

For leverage, the credit spread on long-term debt, and the share of debt due within a year we exclude observations below the 1st and above the 99th percentile. The share of debt due within a year is winsorized at 100%.

Equity issuance is defined as the average of quarterly sale of common and preferred stock over the preceding four quarters divided by assets (`atq`). Quarterly sale of common and preferred stock is constructed from the yearly cumulative variable `sstky`, where missing entries are set to zero. We use an average of quarterly equity issuance over four quarters to reduce the skewness of equity issuance caused by rare but large positive spikes.

Firm-level capital stocks are constructed using the perpetual inventory method described in Appendix 3.A.3. The capital growth moments in Table 3.2 are medians across firm-specific averages and standard deviations of quarterly firm-level capital growth. The firm exit rate is the quarterly value of the yearly exit rate of 8.7% reported in Ottonello and Winberry (2020).

3.D.3 Steady state results: Details

In our solution method described in Section 3.5.1, we exploit the fact that the idiosyncratic state $(z, k, b^S, b^L, \varepsilon, z'; S)$ in the firm problem (3.20) can be summarized by the reduced state vector $(q, b, z'; S)$ which includes firm assets $q = q(z, k, b^S, b^L, \varepsilon; S)$ and outstanding long-term debt $b = (1-\gamma)b^L$. We create grids for the endogenous firm states q and b which are specific to the exogenous firm state z'. The results presented in the paper are computed using a grid of five distinct firm productivity levels z'. Figures 3.D.1 and 3.D.2 show firm policy functions and the firm distribution over the lowest three levels of firm productivity z'.

As shown in Section 3.5.3, the model generates the fact that smaller firms borrow at shorter maturities and therefore have higher shares of maturing debt. The model generates this fact because low-productivity firms have higher default risk. This means that for them the price of long-term debt is more sensitive to the issuance of additional long-term debt. The derivative $\partial p^L / \partial b^{L'}$ in the first order condition for long-term debt (3.3) is steeper for low-productivity firms. This is illustrated in Figure 3.D.3.

Figure 3.D.4 shows additional steady state results on the co-movement of firm age with leverage, credit spreads, and debt maturity. In the data, firm size and leverage are increasing in age whereas credit spreads and the maturing debt share are falling. The model replicates these untargeted patterns.

3.D.4 Heterogenous effects of monetary policy shocks: Details

Figure 3.3 shows the estimated β_1^h coefficients from (3.2) using simulated model data. We construct these estimates as follows. Starting from the steady state of the model, we simulate two panels of a large number of firms for 50 time periods. In the first simulation firms are subject to idiosyncratic shocks in capital quality ε and productivity z', as well as exogenous exit, but there are no monetary policy shocks, i.e., the economy remains in steady state. In the second simulation, all idiosyncratic firm shocks are exactly identical to the first simulation. The only difference is a one-time innovation to $\varepsilon_t^{\mathrm{mp}}$ which on impact induces a 30bp increase in the nominal interest rate i. By regressing the difference in firm-level capital growth between the two simulations at various time horizons h on the pre-shock maturing bond share, we obtain β_1^h in (3.2) displayed in Figure 3.3. The estimates are standardized to measure the differential response associated with a one standard deviation higher $\mathcal{M}_{it}$. The estimates shown in Figure 3.5, as well as in Figures 3.D.5 and 3.D.6 using debt, sales, employment, and credit spreads as additional firm outcomes are constructed correspondingly.

3.D.5 Aggregate implications of heterogeneous debt maturity: Details

In Section 3.5.6, we compare the benchmark model to two alternative economies: an economy without long-term debt, and an economy without heterogeneity.

Economy without long-term debt. In the short-term debt model, the setup is identical to the benchmark model with endogenous debt maturity described above. The key difference is that we set $\gamma = 1$, i.e., there is no long-term debt. The remaining parameters are recalibrated to match the same empirical targets as in the benchmark model. As there is no trade-off between short-term debt and long-term debt, we set the debt issuance cost parameter η to zero and do not target the average share of debt due within one year. Because there is no long-term credit spread in the model, we use the short-term credit spread as model moment in the calibration instead. We increase τ to 60% because otherwise either leverage or credit spreads are too low in the short-term debt model. All remaining externally set parameters are left unchanged. The calibration is summarized in Table 3.D.1.

Table 3.D.1: Model without long-term debt: Internally calibrated parameters

Parameter	Value	Target	Data	Model
σ_ε	0.935	Average firm leverage *(in %)*	34.4	34.2
ξ	0.20	Average credit spread *(in %)*	3.1	3.1
ν	0.0007	Average equity issuance *(in %)*	11.4	13.5
ρ_z	0.6	Average firm capital growth *(in %)*	1.0	1.1
$\bar{z}$	0.184	Std. of firm capital growth *(in %)*	8.3	9.7
κ	0.014	Firm exit rate *(in %)*	2.2	2.3
f	0.2615	Steady state value of firm entry	-	0

Economy without heterogeneity. We solve an alternative model in which all firms are ex-ante identical every period. To do so, we make three assumptions:

(1.) All firms have the same constant productivity level $z = 1$. (2.) We set the equity issuance cost parameter ν to zero. This implies that current cash flow and existing assets q do not appear in firms' first order conditions. The variable q no longer affects firm choices. (3.) We assume that all new entrants pay an entry cost which is financed with long-term debt. The entry cost is set such that entrants always operate with the same amount of outstanding long-term debt b as incumbent firms. This makes sure that entrants do not differ from incumbents because of different histories of long-term debt issuance. The setup is otherwise identical to the benchmark model with firm heterogeneity described above. In this model, firms differ ex-post in terms of realized earnings but all firms are ex-ante identical in the sense that they choose identical policies in every period.

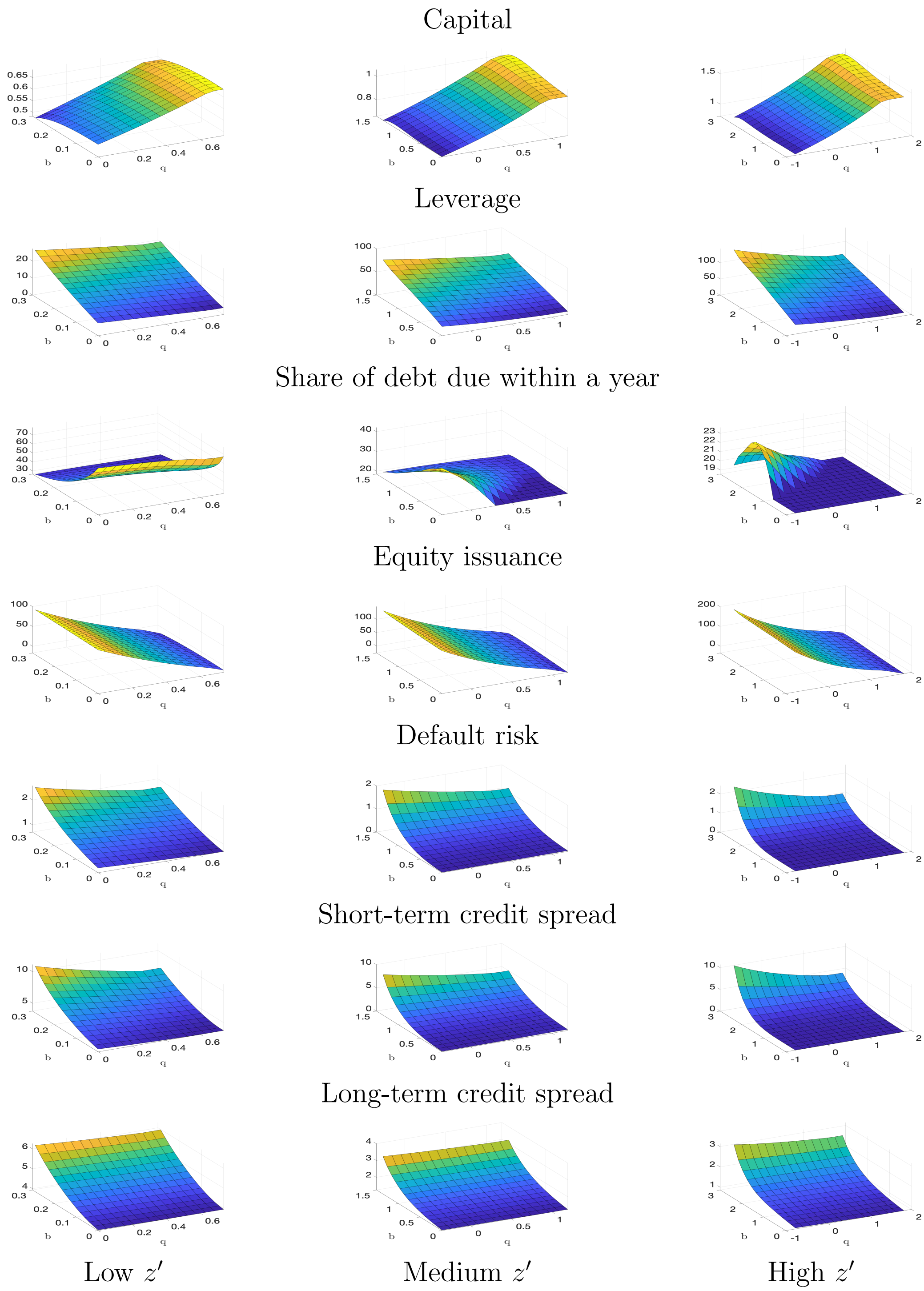
We recalibrate the parameters to match the same empirical targets as in the benchmark model. Because firm productivity is constant, there is no role for the parameters ρ_z and $\bar{z}$ and the associated empirical targets. We also remove equity issuance from the list of our empirical targets, because the equity issuance cost parameter is set to $\nu = 0$. The calibration is summarized in Table 3.D.2.

Table 3.D.2: Model without heterogeneity: Internally calibrated parameters

Parameter	Value	Target	Data	Model
σ_ε	0.75	Average firm leverage *(in %)*	34.4	31.1
ξ	0.90	Average credit spread on long-term debt *(in %)*	3.1	3.3
η	0.0045	Average share of debt due within a year *(in %)*	30.5	31.1
κ	0.0151	Firm exit rate *(in %)*	2.2	2.3
f	0.327	Steady state value of firm entry	–	0

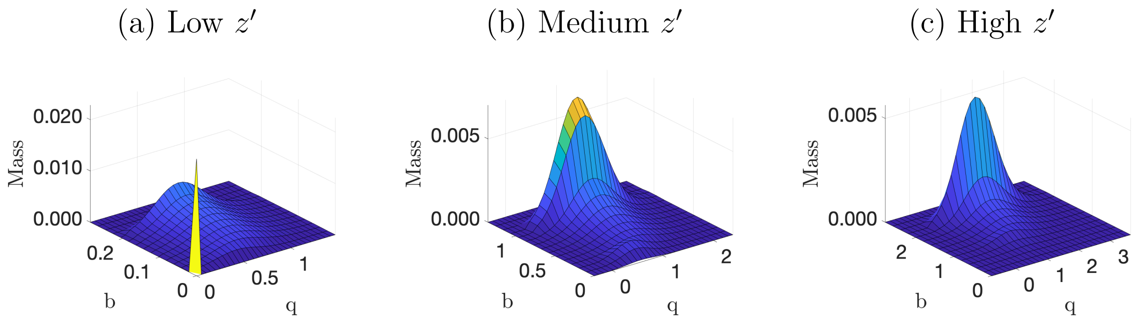
Model comparison relative to frictionless model. An alternative way to compare the different models in Figure 3.6 is to show their response as differences relative to a model with a frictionless production sector (without taxes, default costs, and equity or debt issuance costs). This is done in Figure 3.D.7.

Figure 3.D.1: Steady state policy functions



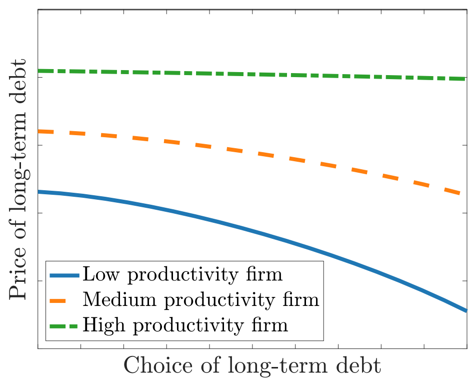
Notes: On the x-axis are firm assets $q = q(z, k, b^S, b^L, \varepsilon; S)$ normalized by average firm capital. On the y-axis is outstanding long-term debt $b = (1-\gamma)b^L$ normalized by average firm debt. Policy functions for *Capital* (k') are normalized by average firm capital. All remaining firm policies are in %. *Leverage* is total firm debt over assets ($(b^{S'} + b^{L'})/k'$); the *Share of debt due within a year* is $(b^{S'} + \gamma b^{L'}(1 + 1 - \gamma + (1-\gamma)^2 + (1-\gamma)^3)/(b^{S'} + b^{L'})$; *Equity issuance* is relative to firm assets (e/k'); *Default risk* is quarterly; *Short-term* and *Long-term credit spread* are annualized.

Figure 3.D.2: Steady state firm distribution



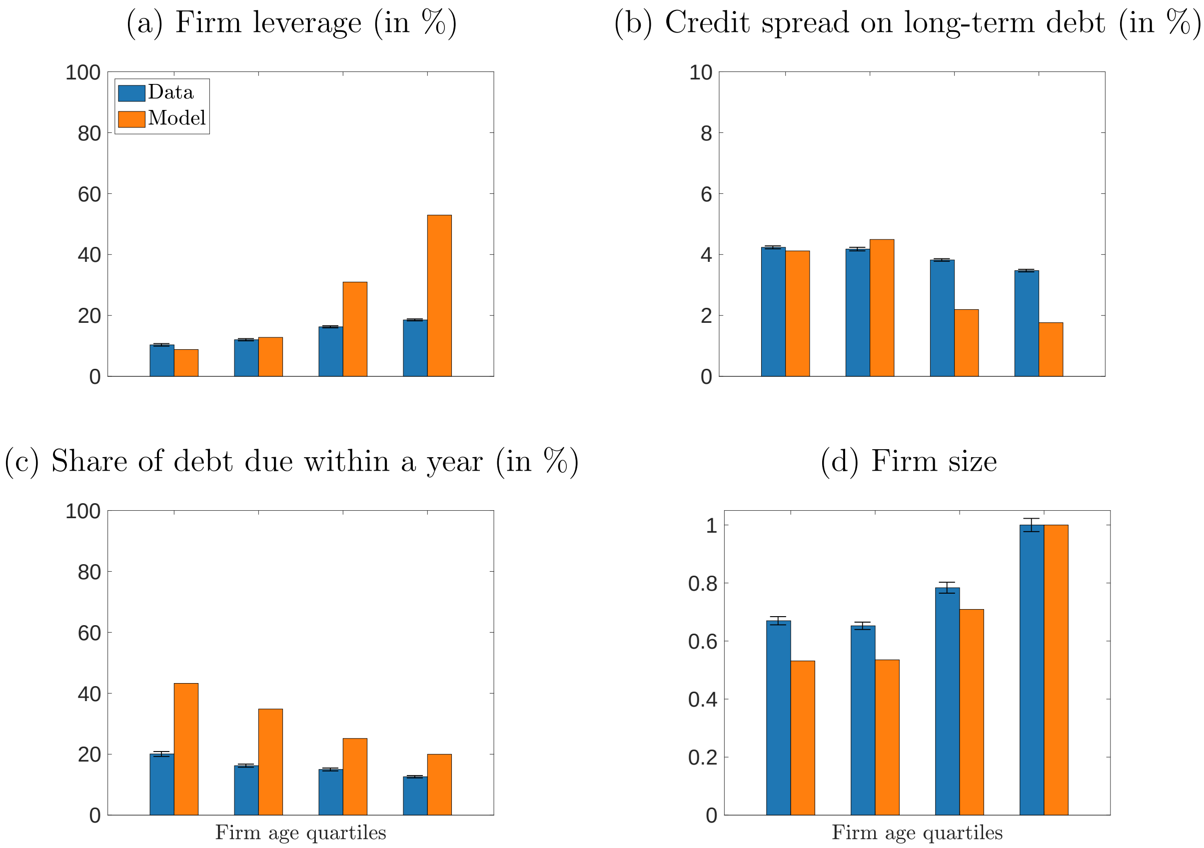
Notes: The steady state firm distribution is plotted for different levels of firm productivity z' against firm assets $q = q(z, k, b^S, b^L, \varepsilon; S)$ and outstanding long-term debt $b = (1-\gamma)b^L$. Assets q are normalized by average firm capital; outstanding long-term debt b is normalized by average firm debt. In panel (a), a large mass point is noticeable at $q = 0$ and $b = 0$ which is the state of new entrants.

Figure 3.D.3: Price of long-term debt p^L



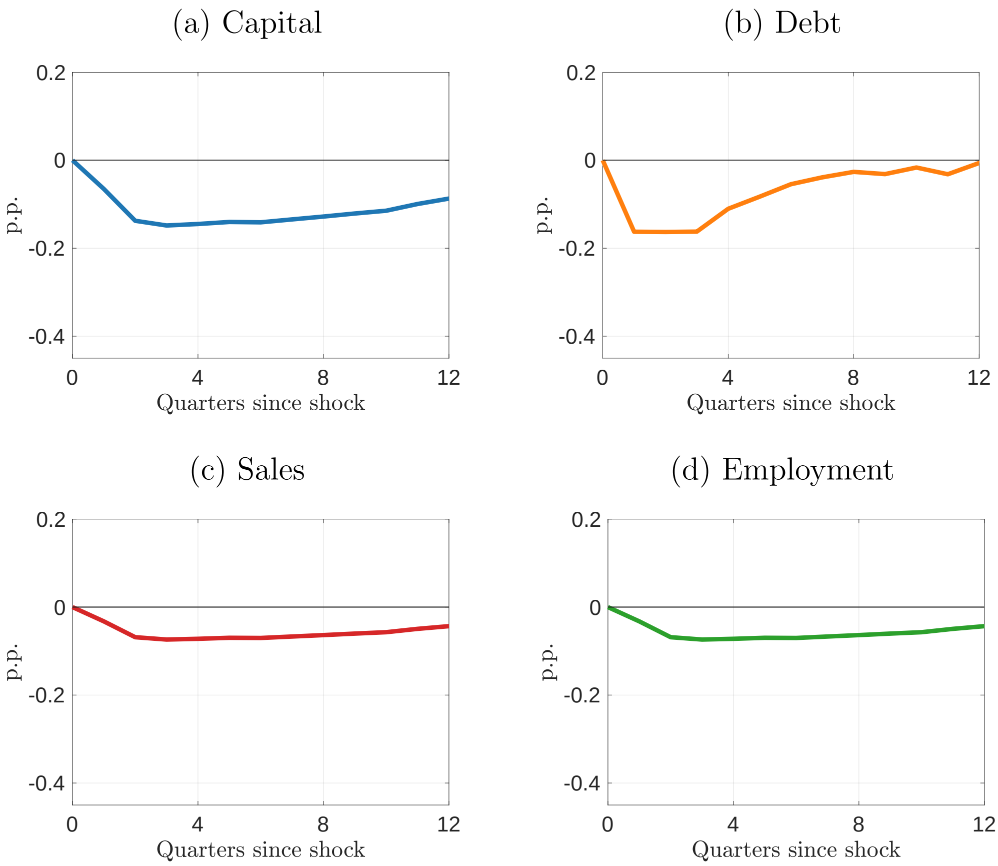
Notes: The price of long-term debt p^L in (3.12) is shown as a function of the firm's choice of long-term debt $b^{L'}$ for a given state of firm assets q and outstanding long-term debt b, and three different productivity levels z'. All firm-level choices besides $b^{L'}$ (i.e., capital k' and short-term debt $b^{S'}$) are held at their steady-state values.

Figure 3.D.4: Firm variables conditional on age



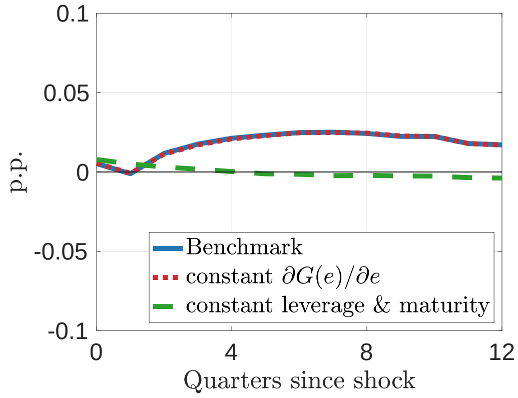
Notes: For each variable, median values are shown by age quartile. The data sample is 1995-2017. Firm-level data on age (quarters since initial public offering), leverage, the share of debt due within a year, and size (log total firm assets relative to top age quartile) is from Compustat. Firm-level credit spreads are computed using data from Compustat and FISD. Empirical median values are shown with 95% confidence intervals. Model moments are computed from the stationary distribution of the model. In the data and the model, observations with age higher than 60 quarters are excluded. See Appendix 3.D.1 and 3.D.2 for details.

Figure 3.D.5: Model: Differential firm-level responses associated with $\mathcal{M}_{it}$



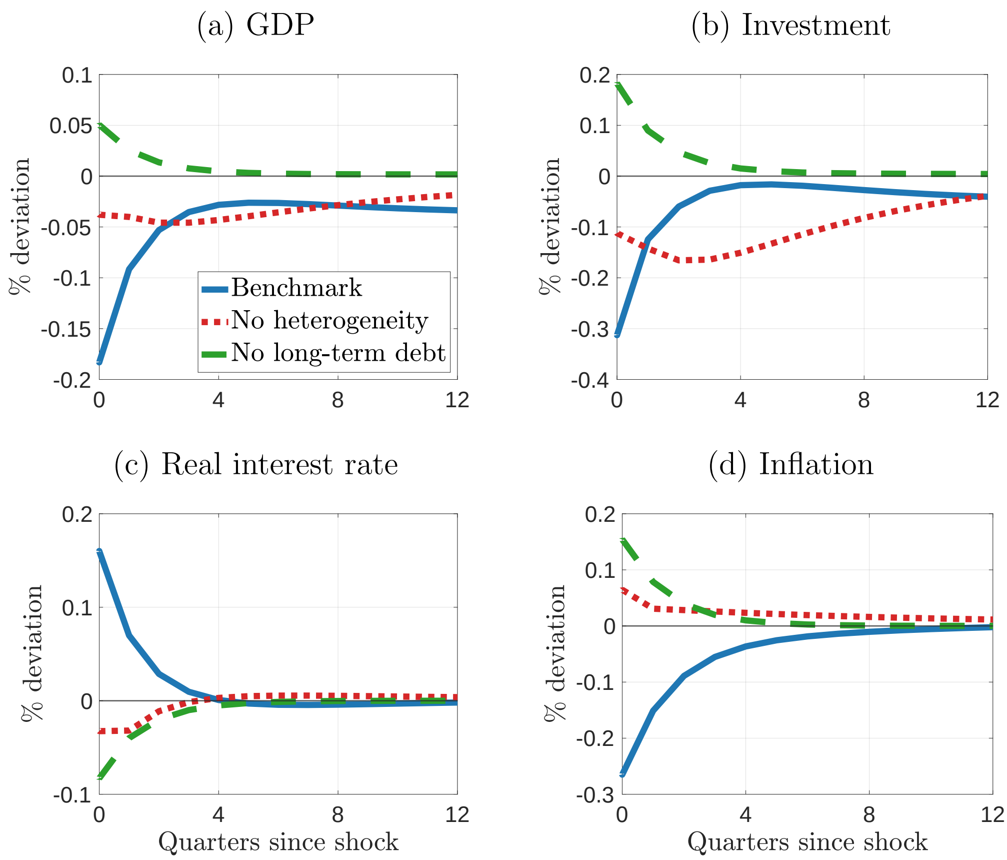
Notes: The lines show the differential response of capital growth, debt growth, sales growth, and employment growth associated with $\mathcal{M}_{it}$ in simulated model data. All values are standardized to capture the differential response (in p.p.) to a one standard deviation (30bp) increase in the nominal interest rate i associated with a one standard deviation higher $\mathcal{M}_{it}$. The differential capital response in panel (a) is identical to Figure 3.3. Debt growth in panel (b) is change in total firm debt relative to pre-shock firm capital (as a control for firm size). Sales growth in panel (c) is log changes in sales y. Employment growth in panel (d) is log changes in labor demand l.

Figure 3.D.6: Counterfactuals: Differential credit spread response associated with $\mathcal{M}_{it}$



Notes: The lines show the differential response of average credit spreads associated with $\mathcal{M}_{it}$ in simulated model data. All values are standardized to capture the differential response (in p.p.) to a one standard deviation (30bp) increase in the nominal interest rate i associated with a one standard deviation higher $\mathcal{M}_{it}$. The blue solid line shows the value from the benchmark model. The red dotted line shows the corresponding value in a counterfactual economy with fixed marginal equity issuance costs. It is barely indistinguishable from the blue solid line. The green dashed line shows the corresponding value in a counterfactual economy with fixed leverage and debt maturity.

Figure 3.D.7: Aggregate response to monetary policy shock: Model comparison relative to frictionless model



Notes: Model responses of Figure 3.6 are shown as difference relative to the response in a model with a frictionless production sector (without taxes, default costs, and equity or debt issuance costs). A value less than zero thus implies a stronger negative response than the frictionless model and vice versa. Blue solid lines correspond to the benchmark economy, green dashed ones to the economy without long-term debt and red dotted ones to the economy without heterogeneity.

Chapter 4

Tax Thy Neighbor: Corporate Tax Pass-through into Downstream Consumer Prices in a Monetary Union*

We estimate the response of product-level retail prices to changes in the corporate tax rates paid by wholesale producers (pass-through). Under perfect competition in goods and factor markets, pass-through of corporate taxes should be zero, and their incidence mainly falls on factor prices. We use variation in tax rates across time and space in Germany, where municipalities set the local business tax once a year, to provide estimates of tax pass-through into the retail prices of more than 125,000 food and personal care products sold across Germany. By leveraging 1,058 changes in the local business tax rate between 2013 and 2017, we find that a one percentage point tax increase results in a 0.4% increase in the retail prices of goods produced by taxed firms and purchased by consumers in the rest of Germany, who thus end up bearing a substantial share of the tax burden. This finding suggests that manufacturers may exploit their market power to shield profits from corporate taxes, complicating the analysis of the redistributive effects of tax reforms. We also explore various dimensions of heterogeneity in pass-through related to market power, including producer size, market shares, and retail store types. While producer heterogeneity does not seem to matter, the significant pass-through of corporate taxes to consumer prices in the low inflation period covered by our sample is mostly due to price changes in supermarkets and hypermarkets.

4.1 Introduction

Who pays for local corporate tax increases in a highly integrated monetary union? The ability to set different local tax rates is usually extolled as a virtue of fiscal federalism. But goods and capital mobility imply that the tax incidence may fall on shareholders, workers in the jurisdiction setting corporate taxes, or consumers

*Joint work with Luca Dedola and Chiara Osbat.

not only in the same jurisdiction but also in other regions in the monetary union, where the goods of taxed firms are exported to. In contrast to a closed economy where the burden of corporate taxes falls fully on shareholders, as shown in the seminal paper by Harberger (1962), full goods and capital mobility and perfectly competitive markets imply that the burden falls mainly on labour, the less mobile (even though generally tax-exempt) factor. If goods markets are not perfectly competitive and firms have market power, then the tax burden will also be borne by consumers, with additional distributive implications. Nevertheless, the effects of corporate tax policies on firms' prices are a rarely analysed issue.

In this paper, we estimate the pass-through of corporate taxes into retail prices in Germany using municipality-level variation in local business tax rates. In this respect, we consider Germany as a highly integrated currency area, comprising many small open economies with no trade frictions and a great deal of capital mobility. We build on Baker, Sun, and Yannelis (2020), who are the first to empirically estimate the pass-through of state-level corporate taxes into retail prices in the United States, using barcode-level retail prices from household scanner data. We complement the results of Baker et al. (2020) by using store-level scanner data and especially by exploiting the German institutional setup of local corporate taxes, which are set at the municipal level.[1] The ensuing much more granular variation in tax changes helps in addressing some well known identification challenges (see also Fuest, Peichl, and Siegloch, 2018).[2] In particular, because we relate local tax changes to price changes outside the production municipality and flexibly control for demand and supply factors separately, the estimated effect is not contaminated by shocks that jointly drive prices and tax rates. Moreover, the local business tax is the main fiscal tool of German municipalities that affects firms, in contrast with central governments and even less decentralised regional fiscal authorities that have multiple tools at their disposal.[3] We also analyse determinants of pass-through of corporate taxes to retail prices, especially concerning heterogeneity in producers and retailers, which plays a key role in theories of imperfect competition and firms' market power.

Specifically, we look at 1,058 tax changes between 2013 and 2017, match-

[1]Differently from Germany, where all firms pay corporate taxes, papers on the tax incidence in the US must also take into account whether a firm is incorporated or not, because corporate taxes in the US depend on the legal form of the firm. Harberger (1962) shows that the tax burden falls on all owners of capital, independently of whether they are incorporated or not. Gravelle and Kotlikoff (1988) shows that when accounting for the endogenous decision to incorporate as well as for dual production of the same good by corporates and non-corporates, the incidence does not change much but the excess tax burden increases substantially.

[2]Several other studies focus on the intra-national variation of corporate tax rates, arguing that this makes it easier to control for unobserved factors. For example, Ljungqvist and Smolyansky (2016) use a difference-in-differences approach at the US state-level and show that a one percentage point increase in the top marginal corporate income tax rate reduces employment by 0.3–0.5% and income by 0.3–0.6%.

[3]German municipalities also set two real estate taxes. One applies to arable land (*Grundsteuer A*) and one on built-up areas (*Grundsteuer B*). Similar to the local business tax, the tax rate is a federally set base level multiplied by local scaling factors. Our estimated pass-through of corporate taxes to prices is robust to controlling for changes in local scaling factors of the real estate taxes.

ing German municipalities with firms' headquarters and with the prices of their products from supermarket scanner data. Similarly to Baker et al. (2020), the identification of our empirical results uses the fact that retail product prices are observed in locations different from where producers are subject to the corporate tax. This allows controlling for local business cycles that may jointly influence prices and tax rates. Our main finding is that local corporate tax pass-through into retail prices of goods "exported" to the rest of Germany is substantial. On average, a one percentage point increase in the local corporate tax rate raises the retail prices of the exported products of taxed firms by around 0.4%.

The municipality-level variation in corporate tax rates used in our analysis was previously considered by Fuest et al. (2018), who argue that it is largely exogenous. They find that a one percentage point tax increase lowers firm-level wages by 0.3%. By using the same tax variation, we can compare their result on the corporate income tax pass-through to wages to ours on retail prices more directly, stressing the novelty of our results. An estimated pass-through of taxes to retail prices of about 40% has the following two implications. First, the fact that consumer prices are affected by tax changes implies significant adjustment in wholesale prices, which are then passed-through by retailers. Second, under the mild assumption that supermarkets do not magnify wholesale price changes, corporate taxes elicit changes in the latter prices large enough to keep net-of-tax profit margins, and thus markups, of producing firms broadly constant.

Because firms are located across different German municipalities and sell their products to many other jurisdictions across Germany through retailers, we frame our analysis in a model of a currency area, consisting of many small open economies trading under minimal frictions through a retail sector, similar to Corsetti and Dedola (2005) and Hong and Li (2017). We show in the model that the effect of corporate taxes on prices depends on the elasticity of consumer demand for each product, the share of retail costs relative to wholesale cots, the share of tax-deductible input costs and the effect of corporate tax on input costs (Fuest et al., 2018).

In line with our model, we extend our empirical analysis to allow for heterogeneity in pass-through of corporate taxes to retail prices. In particular, we estimate category-specific pass-through for 20 COICOP-level categories, but find no significant heterogeneity. We also explore the role of producers and retailers heterogeneity. We allow for heterogeneity in producers' size and market shares, but find little evidence of it. This is interesting, because models with oligopolistic competition or non-CES demand curves, would predict lower pass-through for firms with larger market shares, other things equal (Atkeson and Burstein, 2008; Kimball, 1995). Consistently with the above results, we also find no significant effect of competitors' tax changes on prices of other firms' products. Furthermore, we consider pass-through heterogeneity in terms of income (GDP per capita) in the sales region, finding larger point estimates for high-income regions but no statistically significant differences. However, we do find significant difference in pass-through across store types: While prices in drug stores and discounters are hardly affected, we find significant pass-through into retail prices in supermarkets and hypermarkets.

Related literature. We contribute to four strands of the literature. First, our finding that consumers bear some of the burden of corporate taxes could be appreciated against the backdrop of a large body of literature that has instead examined the effects of corporate taxes on factor prices in various settings, but mostly focusing on closed vs. open economies. Our findings, together with those in Baker et al. (2020), point to the need to include imperfect competition in goods markets into the analysis of the costs and benefits of corporate taxes. Auerbach (2006) discusses possible consequences of relaxing some of the assumptions in Harberger (1962), e.g., allowing for imperfect competition in goods markets and introducing risk. Gravelle (2013) focuses on relaxing the closed-economy setup and reviews the literature on corporate tax incidence in open-economy general equilibrium models, where the relatively higher mobility of capital versus labour increases the tax incidence on wages when factor substitution is low. The reduced incidence on capital arising from the open economy setting is mitigated when the elasticity of substitution of products is low.

Second, we contribute to the vast literature on the role of imperfect competition and market power in price setting, by showing that corporate taxes affect consumer prices. This constitutes a clear deviation from perfectly competitive markets. Moreover, we contribute to the strand of the literature on imperfect competition that focuses on heterogeneous pass-through and markup adjustment. By showing that products with relatively large market share and firms with relatively high total sales have no significantly different pass-through, we complement theoretical and empirical findings in the context of exchange rate pass-through, which show that pass-through decreases with market power, as proxied by market shares and firm size (Atkeson and Burstein, 2008; Auer and Schoenle, 2016; Amiti, Itskhoki, and Konings, 2019).

Third, we contribute to the literature on networks and vertical interactions (see, e.g., Hong and Li, 2017), by showing that the pass-through of shocks to wholesale producers into retail prices is substantial, in particular for supermarkets and hypermarkets relative to discounters. There could be various structural reasons for this finding. On the one hand, producers may discriminate between discounters and other stores, e.g. since they may perceive little market power for sales by the former, or they may be less able to apply price increases to them. On the other hand, retailers may transmit the shocks differently to their customers depending on their own market power, with discounters absorbing price increases into their profit margins, contrary to supermarkets and hypermarkets.

Fourth, we contribute to the literature on price adjustment in currency areas. Similarly to Fuest et al. (2018), McKenzie and Ferede (2017) use the fact that in Canada corporate taxes change across provinces to cast the problem in an open-economy setting across provinces.[4] In light of the open economy literature they predict a high pass-through on local wages. Using provincial data they estimate that in the long run a 1% tax increase lowers wages by 0.11%. Our contribution

[4] Relatedly, Becker, Egger, and Merlo (2012) show that corporate tax rates have an effect on the location decision of multinational enterprises. In particular, they find using German data that higher corporate tax rates reduces employment and fixed assets of foreign MNEs.

shows that tax shocks to producers are passed through not only into their local factor prices, but also into the retail prices of their "exports" to other German regions.

4.2 Institutional setup and data construction

Corporate taxes in Germany are set at the federal and the local level.[5] The tax base and the firms subject to the tax base are defined at the federal level, while the tax rate contains both a federal component and a component that is set in each of more than 10,000 municipalities (*Gemeinden*). Specifically, the tax base is operating profits with some adjustments, for example to account for non-deductibility of equity-based financing and only partial deductibility of debt-based financing. Unlike in the United States, where corporate income is taxed for so-called C-corporations but not for "pass-through" entities, in Germany both incorporated and not incorporated entities are subject to this tax.[6]

The municipality-specific corporate tax rate is computed by multiplying the federally-set basic rate (*Steuermesszahl*) with the local scaling factor (*Hebesatz*) set by the municipalities. Since 2008, well before the start of our sample, the basic rate has been constant at 3.5%. Each year, usually in the last quarter, municipalities decide on the local scaling factor for the next year, becoming effective on January 1. It must be set at least to 2 but is not restricted otherwise (implying that the overall corporate tax rate is at least 7%).[7]

We collect and assemble official data from the Statistical Offices of the 16 German Länder (*Statistische Landesämter*) on yearly municipality-level corporate tax rates. Figure 4.1a shows the significant geographical variation in the level of municipality-level scaling factors. The average scaling factor is 3.62, which results in a corporate tax rate of 12.7%. The largest scaling factor is observed at 9 in the town of Dierfeld in Rheinland-Pfalz (so that the overall corporate tax rate is 31.5%).

In this paper, we construct a unique dataset that links product-level retail prices to municipality-level tax rates based on the location of the producers.[8] We obtain product-level prices from the marketing company Information Resources, Inc. (IRi) (Bronnenberg, Kruger, and Mela, 2008). The German IRi data are collected by point-of-sale scanners and comprise the weekly value and quantity sold of 309,3228 products, identified by barcodes (called EANs, UPCs, or GTINs),

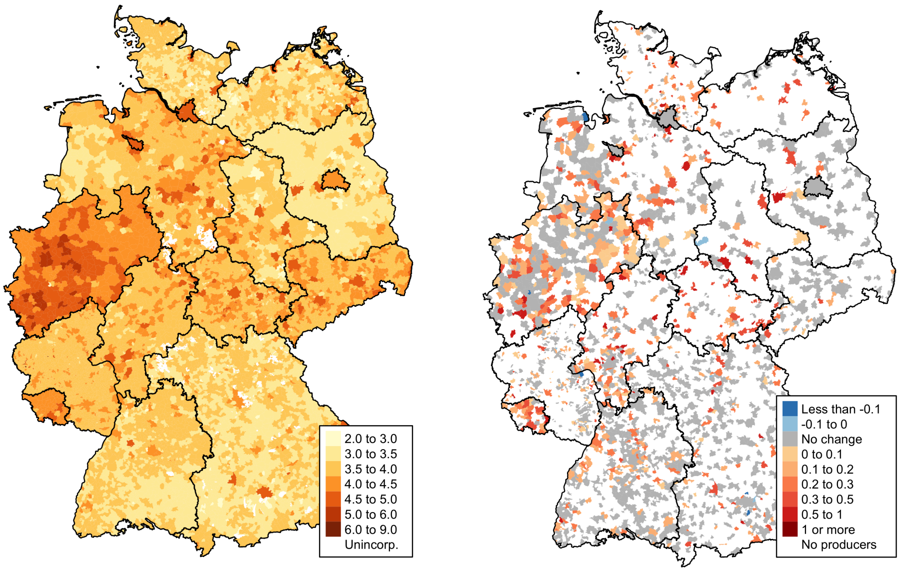
[5]In this paper, we focus on the local business tax component of the corporate tax, following Fuest et al. (2018). There is also the federal corporate income tax (*Körperschaftssteuer*) and the federal personal income tax (*Einkommenssteuer*). The local business tax (*Gewerbesteuer*) yields about 7% of total tax revenue.

[6]Specifically, the self employed as well as firms operating in agriculture and forestry are exempt, but they do not belong in our sample of products.

[7]Note that scaling factors are also commonly reported in percentage points, such that the minimum is 200. If a firm has establishments in many municipalities (or an establishment extends over more municipalities), the tax is apportioned in proportion to the wage bill in each municipality.

[8]Table 4.A.1 in Appendix 1.A.1 contains an overview of all data sources used.

Figure 4.1: Geographical variation in tax scaling factors



(a) Scaling factors in 2017

(b) In-sample cumulative changes between 2013 and 2017

Notes: Panel (a): municipality-specific corporate tax scaling factors in 2017. The effective corporate tax is computed as 3.5% times the scaling factor. Panel (b): Cumulative changes in the municipality scaling factor between 2013 and 2017, which is the sample period for consumer prices used in this paper. Grey areas indicate no change in the scaling factor. White areas indicate municipalities in which no producer location is observed in our sample.

across 10,412 distinct shops from 16 (anonymized) retail chains in 95 two-digit ZIP codes between 2013 and 2017. Thereby, product prices are recorded also in regions other than the one where producers are located.[9] The products are so-called fast-moving consumer goods, including (mostly processed) food, beverages, tobacco, toiletries, and other personal and household care items. The coverage of food, beverages and tobacco accounts for 74 of the 187 ECOICOP categories for goods comprising the Harmonised Index of Consumer Prices (HICP).

To obtain the municipality-level tax rate that applies to the producer of a given product, we match the product-specific barcodes in the IRi data with firm information from the GS1 GEPIR database, which contains producer identity and location. Since we are interested in the pass-through of taxes to prices for German firms, we restrict our attention to barcodes that are registered in Germany.[10] Because the large number of distinct products in the data set prevents us from querying information for every barcode, we focus on a subset of barcodes so as to cover every distinct producer firm. The subset of barcodes is determined as

[9] We aggregate the weekly data to annual frequency, as described below, to match the frequency of tax changes.

[10] Namely, to barcodes beginning with digits 40–44.

follows.[11] First, for most of the barcodes the first seven digits identify the firm, so we focus on barcodes with different seven-digit starting sequences. Second, because for some firms GS1 identifiers are longer than seven digits (up to eleven digits), we also add barcodes with the same starting sequence but attached to different "vendors", which is a coarse firm/brand name variable in the IRi data.

Given this set of barcodes, we obtain detailed associated producer information, including its location, from GS1, the company administrating and licensing barcodes. It is natural to assume that this official address identifies the producer's headquarter and thus where the corporate tax is paid. Note that this information reflects the most recent address of the firm; we are not able to track the historical locations of firms.[12] We are able to obtain the identity and location for 65% of barcodes representing different firms.

We merge the firm information back to the product-level data based on this firm identifier. This yields the producer location for 61% of all German products in the IRi data. Based on the reported postcode and city, we can attach firms to a municipality and thereby the applicable corporate tax rate for the firm over time.[13] Appendix 4.A.3 provides details on the matching of products to firms and municipalities.

Equipped with the concrete firm name and firm location, we are also able to search for these firms in the Orbis database provided by Bureau van Dijk (BvD). We are able to find 77% of firms in Orbis. The Orbis data then helps us to construct a proxy for the existence of establishments and branch offices and their location. This is relevant because of the apportionment rule in the corporate tax code. If a firm produces in several municipalities, the corporate tax code requires that the tax base is divided among municipalities according to the wage bill accruing there.[14] In other words, the true corporate tax rate is a wage-bill-weighted average across the tax rates of all municipalities in which the firm operates. Unfortunately, no establishment-level wage bill data is available to us, so that we cannot compute this. Instead, we address this issue in a robustness exercise by excluding firms with known establishments located outside the headquarter municipality.

Table 4.1 reports the number of products sold in Germany by the different levels of information we have on them. First, of all 311,787 products sold, 175,255 have German barcodes. Considering the products for which we find location information – and, hence, tax information – cuts the number of products to 126,527 in 2,100

[11] The information behind different barcodes registered by the same firm is mostly identical, so this approach is sufficient to determine the location of every product's producer.

[12] This is a potential source of measurement error. However, our sample covers recent years, so that the current addresses of the firms should largely be valid. Moreover, due to the short nature of the sample, re-locations are unlikely to have occurred often.

[13] Neither the city name nor the postcode uniquely identifies a municipality. Municipalities may share names or postcodes. Thus, the matching of firm location to municipalities is done in an iterative way. First, we match to municipalities with a unique name. Then, we match to municipalities for which name and one-digit postcodes are unique, then for those whose name and two-digit postcode are unique, and so on. Firms remain unmatched to municipalities if the city and postcode do not match any municipality.

[14] This should be relevant mainly for large firms. We would expect that small firms do not distribute their administration and production across cities.

Table 4.1: Summary statistics of the matched data

	Barcodes	Vendors	Producers	Municipalities	Sales (bn. €)
Universe of products	311767	11581	–	–	118659.4
Products with German barcodes	175255	6378	–	–	59374.6
... with known firm location	127527	4265	4684	2100	45672.5
... with Orbis information	117351	3368	3739	1620	44240.1

Notes: This table summarises the number of barcodes (individual products at the EAN level), *vendors* as defined by IRi, producers as defined by the GS1 company prefix, municipalities, and the total sales revenue, for the universe of products in the IRi data (row 1), for products with German barcodes (EANs starting with digits 40–44, and excluding private labels; row 2), and for the subset of German products for which we have producer location information (i.e., producer identity and a matched municipality; row 3), and for the subset of those with a match to the Orbis database (row 4). Each row is a strict subset of the previous row.

municipalities. When we also match to those that are present in Orbis the number drops further to 117,351 in 1,620 municipalities (though we will use this subsample only for some robustness checks). Total sales of the sample matched with firm locations covers a large share, roughly 75%, of German barcode sales.

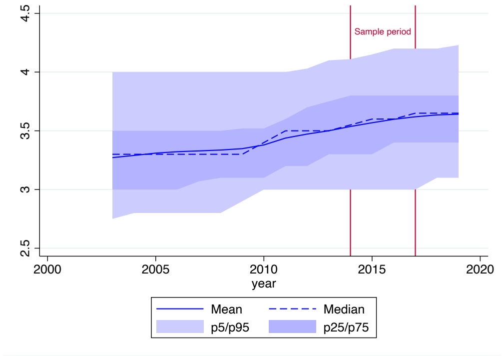
Figure 4.1 shows the variation in the changes in the scaling factors between 2013 and 2017 across Germany. Panel (b) focuses on municipalities which correspond to at least one producer location in our data (white areas indicate municipalities in which no firm was identified in the scanner data). Figure 4.2 (c) reports additional descriptive statistics on tax municipality-level changes. Our matched data set contains producers in 2,100 different municipalities, i.e., around 20% of all municipalities in Germany. Nevertheless, the municipalities in our sample account for a population of 52 million, i.e., around 60% of the German population. In these municipalities the frequency of tax change was 14.1%, close to 15.1% across all municipalities (see Figure 4.2 (b)). The distribution, including the mean, of tax changes is similar in the municipalities in our sample and in all municipalities in Germany, as can be seen from the histogram in Figure 4.2 (b). Our sample period, 2013–2017, is representative of the long-term upward trend in corporate taxes in Germany, see Figure 4.2 (a).

We aggregate the IRi price data as follows. We start with prices per unit for each product, store and week, computed as sales over quantity sold. We then compute annual quantity-weighted average prices of each product in each store and year. We then compute log changes of these store-year specific average prices. We include only prices from stores that were operative for the full current and previous year, in order to avoid possible shop composition effects. We then take the simple average of the log price changes over all stores within a two-digit ZIP code region, retail chain and year.[15] We denote these average log price changes as $\Delta \log p_{isrt}$ where i denotes a product, r a retail chain, s a two-digit ZIP code

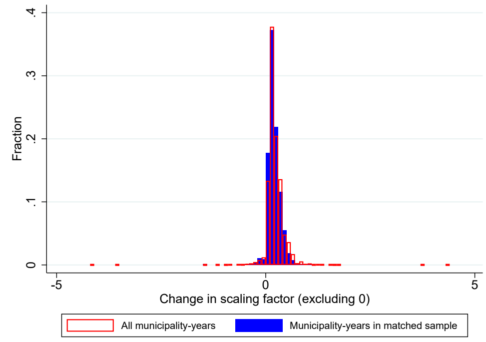
[15] Using quantity-weighted averages yields the same results.

Figure 4.2: Changes in the corporate tax scaling factor

(a) Distribution of scaling factors over time



(b) Histogram at municipality-year level



(c) Descriptives

	Municipalities	Pop.	Freq. change	Mean	Min	Max
All Germany	11,172	83m	15.1%	0.030	-4.20	4.30
Matched sample	2,120	52m	14.1%	0.025	-1.5	1.01

Notes: Panel (a): moments of municipality-specific corporate tax scaling factors over time. The effective corporate tax is computed as 3.5% times the scaling factor. Panel (b): histogram of municipality-year-specific changes in corporate tax scaling factors, for the years 2014–2017. Panel (c): Corresponding descriptives on municipality–year-specific changes in corporate tax scaling factors, for the years 2014–2017.

region, and t years. Appendix 4.A.2 provides more details.

For our panel regressions, we trim the yearly distributions of average log price changes at their 1% and 99% quantiles. We exclude in all regressions the price changes which refer to the two-digit zip code region in which the product is produced, i.e., where sold region and produced region overlap. Effectively, in our empirical analysis we look at how corporate taxes in a municipality affect the retail prices of products originating in this municipality in all other German jurisdictions.

4.3 Theoretical framework

The German institutional setup, where each production firm, located in one of many municipalities with different local tax rates, sells their products mainly outside of that municipality through retailers, and where interest rates are determined at the national level, can be thought of as a currency area comprising many small open economies with no trade frictions and a great deal of capital mobility. To analyse how corporate tax rates may influence prices, we set up a model similar to Corsetti and Dedola (2005) and Hong and Li (2017).

We consider an economy with many local markets m, where in each market a retailer sets the retail price of product i in sector j, P_{ij}^m, as a markup over marginal cost, taking marginal costs as given. The retailer's marginal cost consists of the wholesale price Q_{ij}^m, which is set by the production firm, and an additional distribution cost D_j^m. This cost, which for simplicity depends only on the sector and market, captures factors related to distribution, inventory, advertising, as well as retail inputs like land, capital and labour. Assuming that the retailer has market power and faces a CES final demand curve, the retail price of product i in sector j, sold in region m is

$$P_{ij}^m = \frac{\rho_j^m}{\rho_j^m - 1}(Q_{ij}^m + D_m^j), \tag{4.1}$$

where ρ_j^m is the price elasticity of the demand Y_{ij}^m.

The product wholesale price is set by a production firm, which is generically located in a different region than the retailer, but can sell to all regions m. The manufacturer sets the wholesale price, taking into account its own demand elasticity, which depends indirectly on the retail price. The manufacturer of product i in sector j has a Cobb–Douglas production function using labour L_{ij} and capital K_{ij}, with output elasticities α and $1-\alpha$, respectively, subject to idiosyncratic productivity Z_{ij}. The manufacturer pays the firm-specific (in practice municipality-specific) corporate tax rate τ_{ij} on its revenues, after subtracting labour costs and other deductibles.[16] Denoting the firm-specific wage as W_{ij} and the user cost of capital by R (common to all firms under the assumption of perfect capital mobility), the manufacturer's post-tax profits are given by:

$$\pi_{ij} = (1 - \tau_{ij})\left(\sum_m Q_{ij}^m Y_{ij}^m - W_{ij}L_{ij}\right) - RK_{ij}, \tag{4.2}$$

The assumed production function implies

$$\sum_m Y_{ij}^m = Z_{ij}L_{ij}^{\alpha}K_{ij}^{1-\alpha} \tag{4.3}$$

and, therefore, individual firms' marginal costs are the same for all regions where they sell. Standard static profit maximization yields the following optimal price as a markup over the firm-specific marginal costs, MC_{ij}, scaled by the corporate tax rate:

$$Q_{ij}^m = \frac{\lambda_{ij}^m}{\lambda_{ij}^m - 1}\frac{MC_{ij}}{1 - \tau_{ij}}, \tag{4.4}$$

where λ_{ij}^m represents the possibly region-specific manufacturer's perceived elasticity of demand, $\lambda_{ij}^m \equiv -\frac{\partial Q_{ij}^m}{\partial Y_{ij}^m}\frac{Y_{ij}^m}{Q_{ij}^m} = \rho_j^m \frac{\partial P_{ij}^m}{\partial Q_{ij}^m}\frac{Q_{ij}^m}{P_{ij}^m}$. Firms' before-tax markups are scaled by the corporate tax rate, as the tax reduces the marginal revenue of an additional unit sold by $(1-\tau)$, other things equal. As noted by Hong and Li (2017), producers face a lower elasticity of demand than retailers when the latter do not pass-through wholesale price changes completely: this is the case in our setting if retail costs are strictly positive. Therefore, in general it holds that

[16] In Germany, equity-financed capital is partly deductible (see Fuest et al., 2018).

$\frac{\partial P_{ij}^m}{\partial Q_{ij}^m}\frac{Q_{ij}^m}{P_{ij}^m} < 1$. Namely, under the maintained assumption of CES demand, it holds that $\frac{\partial P_{ij}^m}{\partial Q_{ij}^m}\frac{Q_{ij}^m}{P_{ij}^m} = \frac{Q_{ij}^m}{Q_{ij}^m + D_j^m}$.[17] Using the optimal pricing rules, we obtain the following expression for the equilibrium retail price:

$$P_{ij}^m = \frac{\rho_j^m}{\rho_j^m - 1}\left(\frac{\lambda_{ij}^m}{\lambda_{ij}^m - 1}\frac{MC_{ij}}{1-\tau_{ij}} + D_j^m\right) \tag{4.5}$$

Under the assumption that firms take wages as given, we have $\frac{\partial \log MC_{ij}}{\partial \log(1-\tau_{ij})} = \alpha$. That is, when the tax rate increases, (after-tax) marginal costs *fall.* In particular, the decrease in marginal costs is proportional to the share of deductible inputs in production costs, here α. This reflects the fact that after-tax costs are given by $(1-\tau_{ij})\, W_{ij}L_{ij} - RK_{ij}$. While after-tax marginal costs fall by α% after a one percentage point tax increase, post-tax revenues fall by 1%. In response to this, the production firm increases the wholesale price by $(1-\alpha)$% so as to keep the post-tax markup constant.[18]

We can compute the pass-through from corporate taxes to retail prices as follows:

$$d\log P_{ij}^m = -\left(1 - \frac{1}{\rho_j^m - 1}\frac{D_j^m}{Q_{ij}^m}\right)\frac{Q_{ij}^m}{Q_{ij}^m + D_j^m}\,(1-\alpha)\, d\log\left(1-\tau_{ij}\right) \tag{4.6}$$

This expression shows that, other things equal, tax pass-through to retail prices will be larger, the higher the price elasticity of retail demand, the lower the share of distribution costs in retail costs, and the lower the share of deductible inputs in production costs. In particular, while production firms raise wholesale prices by $(1-\alpha)$% after a one percentage point tax increase, retailers increase prices by less if distribution costs are positive. Moreover, they choose lower pass-through if the retail demand elasticity is higher.

4.4 Empirical strategy

To estimate the causal effect of tax changes on price changes, we leverage the dichotomy between the location of sales and the location of production, following Baker et al. (2020). This dichotomy allows us to control for region-time fixed effects pertaining to the sold region. In addition we also include production region time fixed effects, exploiting the more granular variation of tax rates in our data at the municipality level, whereas in the US corporate taxes are set at the state level. Specifically, we compare annual price changes of goods produced by firms located in a given two-digit ZIP code area, but being subject to different municipality-level

[17] Note that for $\lambda_{ij}^m = \rho_j^m \frac{Q_{ij}^m}{Q_{ij}^m + D_j^m}$ to be well-defined by being greater than one, it needs to hold that $\frac{Q_{ij}^m}{Q_{ij}^m + D_j^m} > \frac{1}{\rho_j^m}$.

[18] However, if manufacturers are able to influence their own wages (or the prices of other deductible inputs) and shift the incidence of tax changes on workers, then $\frac{\partial \log MC_{ij}}{\partial \log(1-\tau_{ij})}$ can differ from α. In general, a tax increase can increase retail prices as long as $\frac{\partial \log MC_{ij}}{\partial \log(1-\tau_{ij})} < 1$.

corporate tax rate changes, focusing on the response of price changes in different two-digit ZIP code areas from the production one. By focusing on *within-sold-region and within-production-region variation* in price changes and tax changes, we flexibly difference out supply- and demand-driven business cycle factors that could jointly affect prices and taxes. The remaining variation in corporate tax rate changes due to factors operating at a more disaggregated level than the two-digit ZIP code is thus arguably exogenous. This view is corroborated by the fact that, in the data, local corporate tax changes are not predicted by changes in county-level GDP or unemployment, as shown by Fuest et al. (2018).[19] Another advantage of our setup is that the corporate tax is the main fiscal lever of municipalities on firms. In other words, municipalities do not include corporate taxes in complex fiscal packages like those of state or federal governments, which could affect firms with other tools and would thus blur the signal about the actual shocks impinging on firms.[20]

We run panel regressions of price changes $\Delta \log p_{isrt}$ of product i, manufactured in production-region (two-digit ZIP code) p, and sold in retail chain r within sold-region (two-digit ZIP code) s, and year t, on net-of-tax factor changes $\Delta \log(1-\tau_{it})$. Therefore, we include fixed effects by sold-region-year, α_{st}, where the sold region is a two-digit ZIP code area, and by production-location-year, α_{pt}, where the production location is either a state or a two-digit ZIP code region.[21] The regression equation, which can be motivated by taking time differences of a suitably extended version of the structural equation (4.6) is

$$\Delta \log p_{isrt} = \alpha_i + \alpha_{st} + \alpha_{pt} + \beta(-\Delta \log(1-\tau_{it})) + \Gamma X_{it} + \varepsilon_{isrt}, \tag{4.1}$$

where α_i is a product fixed effect, α_{st} is a two-digit ZIP code sold location by year fixed effect, and α_{pt} is a two-digit ZIP code production location by year fixed effect. X_{it} can contain control variables at the municipality and county (*Kreis*) level, specifically, four lags of changes of the production municipality unemployment rate and four lags of growth rates in the production county-specific debt. The coefficient of interest β captures the elasticity of the price with respect to the negative net-of-tax factor. We choose this normalization such that an increase in the regressor corresponds to an increase in the corporate tax rate. Since $-\Delta \log(1-\tau) \approx \Delta\tau$, this elasticity is approximately equivalent to the semi-elasticity of the price with respect to the tax. That is, β indicates the average relative increase in prices in response to a one percentage point increase in the corporate tax rate. We cluster standard errors at the municipality level, which allows for arbitrary serial correlation of shocks within municipalities.

One can lend a causal interpretation to the coefficent β, to the extent that our right-hand side variables, including the location-time fixed effects, control for

[19]Counties or districts (*Kreise and kreisfreie Städte*) are the administrative level between municipalities and states in Germany. There are, on average, 25 municipalities in each county.

[20]In a robustness check below, we control for changes in the local scaling factors that apply to real estate taxes, which are two additional instruments of municipalities. Our estimates remain practically unchanged.

[21]Each two-digit ZIP can contain multiple stores and municipalities, thus this variation helps to estimate, respectively, sold-region-year fixed effects, α_{st}, and production-region-year fixed effects α_{pt}.

the co-movement between prices and taxes that arise from business cycle effects. Specifically, the sold-location-year fixed effect differences out local demand conditions. Such factors would induce endogeneity of tax changes if a fall in prices in a sold-region, leading to declining profits, would lead in turn to the municipality reducing the corporate tax rate to support local firms. The production-region fixed effect differences out local supply-related factors that could affect the cost of all goods manufactured in the production region. Such factors would induce endogeneity if municipalities were to be induced to change corporate taxes to, e.g., alleviate the impact of wage or other cost increases on firms. Additionally, the inclusion of local unemployment rates and debt in the regression proxies for changes in production costs and other determinants of prices at an even more disaggregated local level. Indeed, in line with the results in Fuest et al. (2018), the remaining variation in corporate tax changes at the municipal level can be thought of as largely exogenous.

4.5 Results

4.5.1 Average pass-through of local corporate taxes into retail prices in other jurisdictions

In this section we present the estimated pass-through of corporate taxes into consumer prices in regions outside the production location. Table 4.1 reports the estimated pass-through coefficient based on three specifications of equation (4.1). Column (1) uses sold-region-year and production-region-year fixed effects but no further controls, while columns (2) and (3) add four lags of changes of the production municipality unemployment rate and four lags of growth rates in the production county-specific debt level. All specifications include product fixed effects to account for product-specific price trends (Adam and Weber, 2022). The point estimates of the pass-through coefficient are all positive, ranging from 0.425 in column (3) to 0.525 in column (1), and highly significant.

The positive coefficients imply that prices increase in response to an increase in the corporate tax rate. Specifically, the coefficient in column (3) implies that a one percentage point increase in the local corporate tax rate leads on average to an approximately 0.425% increase in the retail price of products exported from the affected municipality, relative to the prices of all other products originating

Table 4.1: Estimated pass-through from corporate taxes to consumer prices

	(1)	(2)	(3)
	Δ log price	Δ log price	Δ log price
$-\Delta\log(1-\text{tax})$	0.525^{***}	0.538^{***}	0.425^{**}
	(0.171)	(0.182)	(0.209)
Observations	19434155	18871628	14091803
Product FE	✓	✓	✓
Sold-region × year FE	✓	✓	✓
Production-region × year FE	✓	✓	✓
Production-muni. UE controls		✓	✓
Production-district debt controls			✓

Notes: Results from estimating $\Delta \log p_{isrt} = \alpha_i + \alpha_{st} + \alpha_{pt} + \beta(-\Delta \log(1-\tau_{it})) + \Gamma X_{it} + \varepsilon_{isrt}$. Prices are observed at the product, retail chain, two-digit zip code sold location, and year level. Tax rates vary by production municipality and year. In the panel regression, α_i is a product fixed effect, α_{st} is a two-digit ZIP code ("region") sold location by year fixed effect, α_{pt} is a two-digit production location ZIP code ("region") production location by year fixed effect. Depending on the specification, X_{it} contains four lags of changes of the production municipality unemployment rate and four lags of growth rates in the production county-specific debt level. Standard errors (in parentheses) are clustered at the municipality level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

from municipalities in different regions.[22][23][24] This result is remarkable given the evidence in Fuest et al. (2018) that firm wages fall with a corporate tax increase, easing producers' costs. In contrast, the fact that retail prices instead increase suggests significant adjustment in wholesale firms' prices, which is in turn passed-through into higher retail prices by supermarkets.

As a first robustness check, we exclude products of firms with multiple establishments. While we only observe the tax rate in the headquarter municipality, the effective tax rate for a firm is a wage-bill-weighted average over all production establishment municipalities. Our results may thus be influenced by the presence of multi-establishment firms. As explained in Section 4.2, we match IRi data with

[22]In Table 4.B.2 in the Appendix shows these regressions when using directly $\Delta\tau$ as the main regressor. The findings are quantitatively very similar. The results are also robust to trimming price changes at different cutoffs, see Table 4.B.3a and robust to using sales-filtered price data, see Table 4.B.3b.

[23]Table 4.B.1 in the Appendix shows that the estimated pass-through is robust to controlling for changes in local scaling factors applying to two real estate taxes, which are two additional fiscal instruments at municipalities' disposal. The first real estate tax (*Grundsteuer A*) applies to arable land and the second (*Grundsteuer B*) on built-up areas. While the local business tax generated total tax revenues of 55 billion euro in 2019, the revenues from the real estate taxes were lower. The revenues from the tax on arable land amounted to 0.4 billion euro and the revenues from the tax on built-up areas amounted to 14 billion euro.

[24]During the period covered by our sample, tax changes have been predominantly positive. Of the 1,058 observed tax changes, only 31 were tax cuts. Standard models would predict the effects to be symmetric across otherwise similar tax increases and decreases.

Table 4.2: Robustness to excluding firms with branches

	(1)	(2)
	Δ log price	
	All Orbis firms	without branch
$-\Delta\log(1-\text{tax})$	0.413**	0.581***
	(0.209)	(0.203)
Observations	13564215	6591425
Product FE	✓	✓
Sold-region × year FE	✓	✓
Production-region × year FE	✓	✓
Production-muni. UE controls	✓	✓
Production-district debt controls	✓	✓

Notes: Results from estimating $\Delta\log p_{isrt} = \alpha_i + \alpha_{st} + \alpha_{pt} + \beta(-\Delta\log(1-\tau_{it})) + \Gamma X_{it} + \varepsilon_{isrt}$, as in Table 4.1, using three different subsamples: Column (1) includes only prices of products for which the producing firm is observed in the Orbis database. Column (2) further restricts to observations for which the producing firm does *not* have recorded branches in Orbis. In this case, we can exclude issues of tax apportionment. X_{it} contains four lags of changes of the production municipality unemployment rate and four lags of growth rates in the production county-specific debt level. Standard errors (in parentheses) are clustered at the municipality level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Orbis firm data in order to exclude multi-establishment firms. Orbis includes some information about what the data provider calls "branches". A branch is a recorded firm presence outside of the location of the headquarter. Column (1) in Table 4.2 repeats the pass-through estimation for all Orbis firms as a benchmark. Column (2) then includes only the product prices of firms without branches. It turns out that our benchmark estimate is robust to excluding multi-establishment firms, as the pass-through coefficient is again insignificantly different from 0.5 (although there is, unfortunately, no guarantee that the Orbis information perfectly captures establishments).

We further assess the robustness of our results, showing that they are not driven by retailer-specific or product category-specific effects. We do so by adding more granular fixed effects, see Table 4.3. Column (1) reproduces the baseline specification from Table 4.1. Column (2) adds retail chain-sold-region-year fixed effects, to capture demand factors that are specific to a given retail chain in a given region. Effectively, we then compare the retail prices of products exported from a municipality subject to a corporate tax change, to those of firms located outside that region's municipality, within a given retail chain. Column (3) adds category-sold-region-year fixed effects to capture factors that are specific to a given product category in a given region, thereby analogously comparing relative price changes of products in the same category sold in the same region. The results indicate that our benchmark estimates are robust to controlling for more granular

Table 4.3: Robustness to adding more granular fixed effects

	(1)	(2)	(3)
	Δ log price	Δ log price	Δ log price
$-\Delta\log(1-\text{tax})$	0.425**	0.416**	0.339**
	(0.209)	(0.204)	(0.166)
Observations	14091803	14091803	14091803
Product FE	✓	✓	✓
Production-region × year FE	✓	✓	✓
Production-muni. UE controls	✓	✓	✓
Production-district debt controls	✓	✓	✓
Sold-region × year FE	✓		
Sold-region × retailer × year FE		✓	
Sold-region × category × year FE			✓

Notes: Results from estimating $\Delta\log p_{isrt} = \alpha_{i(rs)} + \alpha_{s(r)t} + \alpha_{pt} + \beta(-\Delta\log(1-\tau_{it})) + \Gamma X_{it} + \varepsilon_{isrt}$, with different levels of fixed effects. α_i is a product fixed effect. $\alpha_{s(r)t}$ is a two-digit ZIP code ("region") sold location by year fixed effect in specification (1), a sold location by retailer by year fixed effect in specification (2) and a sold location by category by year fixed effect in specification (3). α_{pt} is a two-digit production location ZIP code ("region") production location by year fixed effect. X_{it} contains four lags of changes of the production municipality unemployment rate and four lags of growth rates in the production county-specific debt level. Standard errors (in parentheses) are clustered at the municipality level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

sources of unobserved heterogeneity, with the point estimates of the pass-through coefficient ranging from 0.339 to 0.416. However, the differences in these point estimates are not statistically significant.

Although the local fixed effects that we include in our panel regressions control flexibly for common local shocks, we carry out another test to address concerns of exogeneity of the tax rate changes. This placebo-type exercise checks if randomly re-allocating tax changes across municipalities within a narrowly defined region also results in price changes. If this was the case, prices would change either due to unobserved local shocks or due to spillovers. Specifically, for any municipality we randomly draw a tax change from the population of tax changes observed in municipalities that are located in either the same two-digit ZIP code region or the same county. We then re-run the baseline regression as in equation (4.1). Table 4.4 shows the results, which reveal small and insignificant coefficients. This corroborates our finding that prices indeed increase due to municipality-specific changes in tax rates.

Finally, we estimate the dynamic effects of municipality-level tax changes on retail prices. To this end, we extend the panel regression (4.1) to an event study

Table 4.4: Placebo exercise

	(1) Δ log price	(2) Δ log price
$-\Delta\log(1-\text{tax})$, randomised within 2-digit ZIP code	0.118	
	(0.128)	
$-\Delta\log(1-\text{tax})$, randomised within district		-0.103
		(0.135)
Observations	16795538	10835955
Product FE	✓	✓
Sold-region × year FE	✓	✓
Production-region × year FE	✓	✓

Notes: Results from estimating $\Delta\log p_{isrt} = \alpha_i + \alpha_{st} + \alpha_{pt} + \beta(-\Delta\widetilde{\log(1-\tau_{it})}) + \varepsilon_{isrt}$. The regression is set up as in Table 4.1, but uses randomised regressors $(-\Delta\widetilde{\log(1-\tau_{it})})$. In particular, column (1) randomises the value of $\Delta\log(1-\tau_{it})$ by drawing a random $\Delta\log(1-\tau_{it})$ with replacement from the population of municipalities within the two-digit production location ZIP code. The exercise in column (2) draws a random $\Delta\log(1-\tau_{it})$ from the population of municipalities within the county of the production location. This leads to fewer observations because some counties are also one municipality, in which case we do not consider them for randomisation B. Standard errors (in parentheses) are clustered at the municipality level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

regression:

$$\Delta\log\text{price}_{isrt} = \alpha_i + \alpha_{st} + \alpha_{pt} + \sum_{k=-3}^{3}\beta_k(-\Delta\log(1-\tau_{it-k})) + \Gamma X_{it} + \varepsilon_{isrt} \quad (4.1)$$

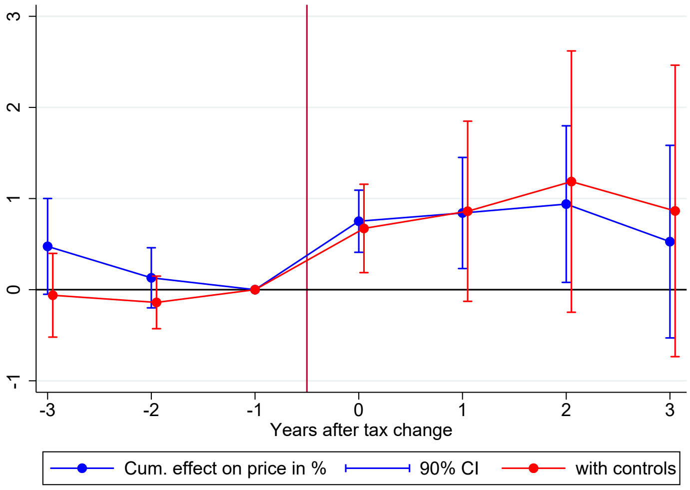
Figure 4.1 plots the estimated coefficients $\{\beta_k\}$. For $k = -3, ..., 3$, β_k indicates the effect of a tax change in period t on retail prices in period $t+k$. The coefficient β_{-1} is normalized to zero. The results show that prior to a tax change, there are no significant changes in retail prices. This flat pre-trend is consistent with the exogeneity of tax changes. In the years after the tax change, prices increase significantly and stay persistently higher.[25]

4.5.2 The role of market shares, firm size, and competitor tax changes

According to models of oligopolistic firms (Atkeson and Burstein, 2008) and models with kinked, non-CES demand curves (Kimball, 1995), pass-through should depend on market shares, and specifically it should fall the larger the latter, possibly increasing again for very large market shares. Since our dataset is uniquely

[25] This is in line with the fact that tax changes are empirically highly persistent.

Figure 4.1: Dynamic effects of a corporate tax change



Notes: In panel (a) this figure plots, for a horizon of h years after the tax change, the sum of coefficients $\sum_{k=0}^{h} \beta_k$ from the regression $\Delta \log \text{price}_{isrt} = \alpha_i + \alpha_{st} + \alpha_{pt} + \sum_{k=-3}^{3} \beta_k(-\Delta \log(1 - \tau_{it-k})) + \Gamma X_{it} + \varepsilon_{isrt}$, where β_{-1} is normalised to zero. In panel (b) price changes are replaced by quantity changes. The panel regression is otherwise set up as in Table 4.1. The whiskers show 90% confidence bands based on standard errors clustered at the municipality level.

suited to investigate this hypothesis given that it includes very granular information on sales for each "barcode" product, we estimate the pass-through as a function of different market shares. We proceed as follows. For a given definition of a market m, we compute the market share of a product sold in that market as

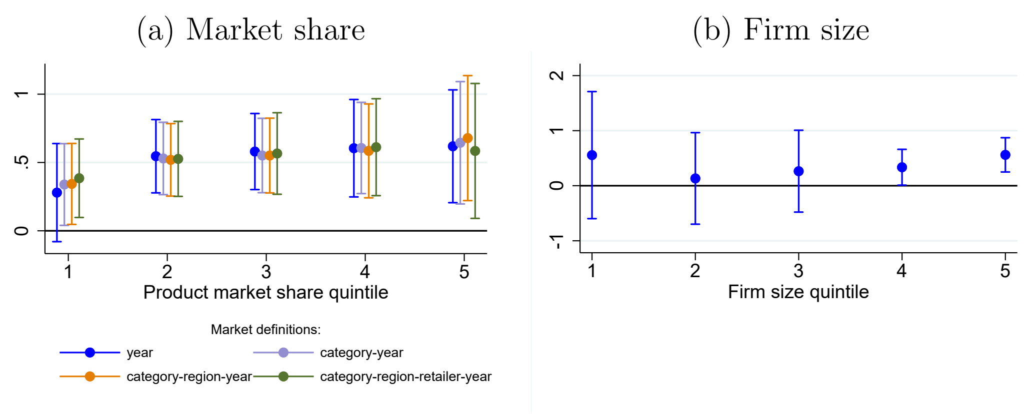
$$s_{isrt}^{(m)} = \frac{\text{sales}_{isrt}}{\sum_{i's'r' \in m} \text{sales}_{i's'r't}}. \tag{4.2}$$

where again s and r refer to a sold-region and a retailer. We use the following four market definitions: First, all products sold in a given year across all categories and regions; this is the market share of each individual product in all sales of German supermarkets.[26] Second, all products sold in a given COICOP product category and in a given year, across all regions; this is the market share of each individual product in all sales in its category.[27] Third, all products sold in a given category, in a given two-digit ZIP code region, in a given year; this is the market share of each individual product in its category at the regional level. Fourth, all products sold in a given category, in a given two-digit ZIP code region, by a given retailer,

[26] The denominator includes therefore also sales of products for which we do not observe the producer identity in our sample.

[27] We manually map the roughly 200 categories in the IRi data set into twenty COICOP level-3 categories. The Classification of Individual Consumption by Purpose (COICOP) is used, for example, in the euro area Harmonized Index of Consumer Prices (HICP).

Figure 4.2: The role of market share and firm size for pass-through



Notes: Effect of an increase of corporate tax rates on retail prices by product market share and by firm size. Observations are sorted into quintile bins according to product-level sales within the market and according to total product sales for the producing firms (in the given year). The figure then plots bin-specific coefficients β_{q_k} from the regression $\Delta \log \text{price}_{isrt} = \alpha_i + \alpha_{st} + \alpha_{pt} + \sum_{k=1}^{5} \beta_{q_k} \mathbb{1}\{i \in q_k\}(-\Delta \log(1 - \tau_{it})) + \varepsilon_{isrt}$. Standard errors are clustered at the municipality level. Confidence intervals are at the 90% level.

in a given year; this is the same market share as the previous one, but computed within each specific retailer.[28]

We sort observations into quintiles based on market share. Thereby, we assign an observation to the lowest product market share quintile if $s_{i,s,r,t}$ is in the lowest 20% of market shares in the market share distribution of market m.[29] We then estimate, as for product categories above, an extension of the panel regression (4.1) where we interact the net-of-tax factor with dummy variables representing the market share quintile. Figure 4.2 (a) plots the quintile-specific pass-through coefficients for the various market definitions. There are no statistically significant differences between the estimated effects across market share bins. However, the point estimates and the statistical significance of the individual estimates suggest a modestly stronger effect of corporate taxes on prices for products with larger market shares, across the second to the fifth quintiles. This pattern holds irrespective of the specific market definition.

Similarly, we estimate pass-through conditional on the size of the production firm. We compute firm size as the total sum of all product sales of a firm in all regions and retailers (for a given year). We assign an observation into quintile group k if the production firm's size is in the kth qunitile of the firm size distribution (for a given year), where every firm is only counted once. We then again estimate the panel regression (4.1) extended to firm size quintile-specific coefficients. Figure 4.2

[28] Our retail scanner data does not observe sales by hard discounters. Therefore, total market sales are only partly captured and thus market shares may be mismeasured. However, this caveat does not apply to the category-region-retailer-year measure, which is retailer-specific and therefore does not depend on sales in other retailers.

[29] Note that an equivalent binning would arise from sorting according to sales within the market.

Table 4.5: Estimated pass-through of own and competitor tax changes

	(1)	(2)	(3)
	Δ log price	Δ log price	Δ log price
Market definition m:	Category	Cat.-Region	Cat.-Reg.-Retailer
$-\Delta\log(1-\text{tax})$	0.519***	0.507***	0.521***
	(0.170)	(0.170)	(0.171)
$-\overline{\Delta\log(1-\text{tax})}_{-i}^{(m)}$	1.956	0.488	0.279
	(1.738)	(0.797)	0.552
Observations	19434155	19434155	19434155
Product FE	✓	✓	✓
Sold-region × year FE	✓	✓	✓
Production-region × year FE	✓	✓	✓

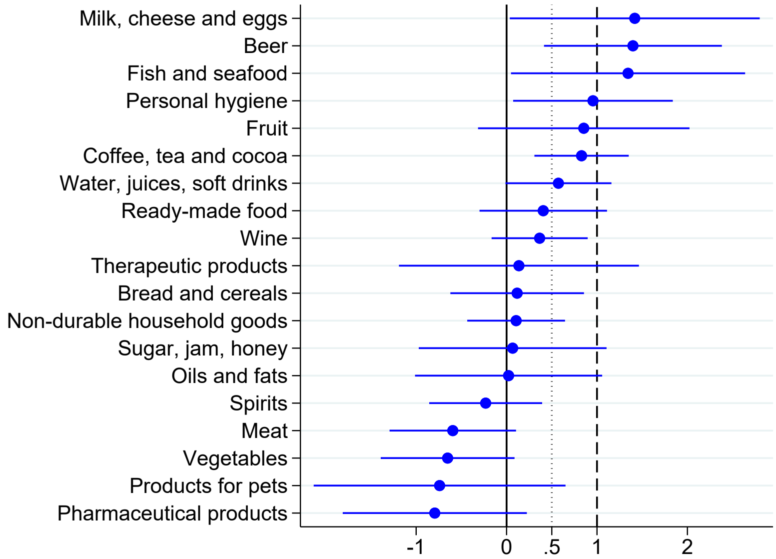
Notes: Results from estimating $\Delta\log p_{isrt} = \alpha_i + \alpha_{st} + \alpha_{pt} + \beta(-\Delta\log(1-\tau_{it})) + \delta\overline{\Delta\log(1-\tau)}_{-isrt} + \varepsilon_{isrt}$. The average competitor tax change is defined as $\overline{\Delta\log(1-\tau)}_{-isrt}^{(m)} := \sum_{i',s',r'\in m, i'\neq i} s_{i's'r't}^{(m)}\Delta\log(1-\tau_{i't})$ where s is the competitor market share in market m, based on the market definition being applied in the respective column. Prices are observed at the product, retail chain, two-digit zip code sold location, and year level. Tax rates vary by production municipality and year. In the panel regression, α_i is a product fixed effect, α_{st} is a two-digit ZIP code ("region") sold location by year fixed effect, α_{pt} is a two-digit production location ZIP code ("region") production location by year fixed effect. Standard errors (in parentheses) are clustered at the municipality level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

(b) plots the quintile-specific pass-through coefficients. While the coefficients are again not statistically different from each other, a similar pattern as for market shares emerges. The point estimates tend to increase with firm size and, notably, only the coefficients for the top 40% of firms are statistically different from zero. This weakly suggests again that if anything pass-through is stronger for larger firms.

This pattern of broad insensitivity of pass-through – both to market shares and firm size – is inconsistent with pass-through in models of oligopolistic firms (Atkeson and Burstein, 2008) and with kinked, non-CES demand curves (Kimball, 1995), and different from previous empirical results on exchange rate pass-through by Auer and Schoenle (2016) and Amiti et al. (2019). Nevertheless, our finding that larger firms are those whose retail prices react more to corporate taxes is best appreciated in light of the evidence in Fuest et al. (2018) that the same set of firms also does not pass-through tax changes to their workers' wages. This is consistent with the presumption that market power allows larger firms to shift the tax incidence to their consumers rather than their workers.

In an imperfectly competitive market environment, firms may respond not only to corporate taxes levied on their own profits, but also to (changes in) the corporate taxes in other jurisdictions that are levied on their competitors and lead the latter to change their prices. We test for this possibility by extending our

Figure 4.3: Category-specific pass-through



Notes: Effect of an increase of corporate tax rates on retail prices by product category. The figure plots category-specific coefficients $\beta_{(c)}$ from the regression $\Delta \log \text{price}_{isrt} = \alpha_i + \alpha_{st} + \alpha_{pt} + \sum_c \beta_{(c)} \mathbb{1}\{i \in c\}(-\Delta \log(1-\tau_{it})) + \varepsilon_{isrt}$. Standard errors are clustered at the municipality-level. Confidence intervals are at the 90% level.

baseline regression (4.1) to include a competitor tax change variable defined by

$$\overline{\Delta \log(1-\tau)}^{(m)}_{-isrt} := \sum_{i',s',r' \in m, i' \neq i} s_{i's'r't} \, \Delta \log(1-\tau_{i't}). \tag{4.3}$$

where s is the competitor market share as defined in equation 4.2, for which the market m is again defined at the category, category-region, or category-region-retailer level.

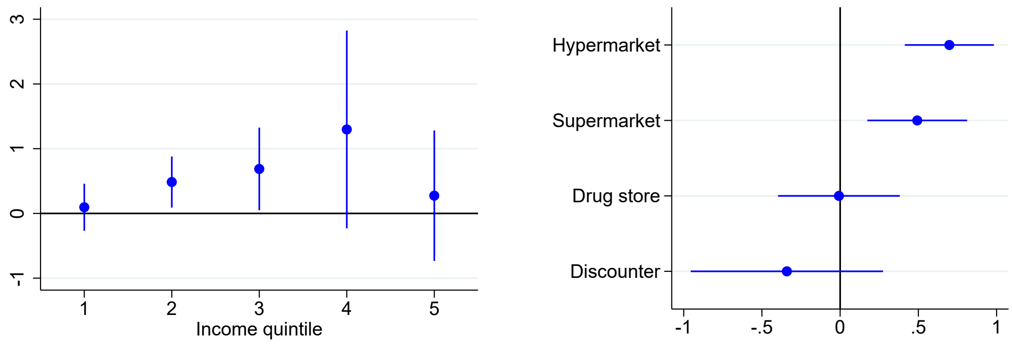
Table 4.5 shows the result of the extended regression, revealing that we do not find a significant effect of changes in competitor taxes. The estimates of the own-tax pass-through are unchanged when conditioning on competitor tax changes. This result is consistent with the finding that pass-through does not significantly vary by market shares and thus similarly suggests only weak strategic complementarities.

4.5.3 Heterogeneous pass-through: Product categories, regional income, and retailer types

In this section we explore possible heterogeneity in pass-through of corporate taxes to retail prices along several broad dimensions: product categories, regional household income, and retailer type. Looking at equation (6), pass-through may be different across these dimensions. First, product categories differ in the price elasticity of demand or distribution costs. Second, price elasticities may also differ across regions, as a function of households income levels (e.g., Anderson, Rebelo,

Figure 4.4: Pass-through by sales region income and by retail store

(a) Sales region income (b) Store type



Notes: Panel (a): Effect of an increase of corporate tax rates on retail prices (vertical axis) by sales region income (horizontal axis). The figure plots income quintile-specific coefficients β_{q_k} from the regression $\Delta \log \text{price}_{isrt} = \alpha_i + \alpha_{st} + \alpha_{pt} + \sum_{k=1}^{5} \beta_{q_k} \mathbb{1}\{i \in q_k\}(-\Delta \log(1 - \tau_{it})) + \varepsilon_{isrt}$. Panel (b): Effect of an increase of corporate tax rates on retail prices (horizontal axis) by retail store (vertical axis). The figure plots store type-specific coefficients $\beta_{(\tilde{r})}$ from the regression $\Delta \log \text{price}_{isrt} = \alpha_i + \alpha_{st} + \alpha_{pt} + \sum_{\tilde{r}} \beta_{(\tilde{r})} \mathbb{1}\{r \in \tilde{r}\}(-\Delta \log(1-\tau_{it})) + \varepsilon_{isrt}$, where $\tilde{r}$ denotes a hypermarket, supermarket, drug store, or discounter. Standard errors are clustered at the municipality-level. Confidence intervals are at the 90% level.

and Wong (2020) find that markups increase with local GDP per capita in the US). Third, retailers may differ in their distribution costs or face alternative levels of price elasticities due to their consumers having different preferences or different degrees search effort, which may in turn make it optimal for firms and retailers to implement different degrees of pass-through.

Pass-through by product categories. To estimate pass-through by product category, we estimate an extension of the panel regression (4.1) where we interact the net-of-tax factor with dummy variables representing each product category in our sample. As product categories we consider COICOP categories, as above. Figure 4.3 shows the results. We find no categories that exhibit statistically different pass-through from our baseline estimate, partly on account of large uncertainty in some categories. While there is dispersion in category-specific pass-through estimates, there are no extreme outliers. This suggests that the significant estimate of average pass-through comes from pooling all categories, while there is no strong evidence for heterogeneity across product categories.

Pass-through by income in the sales region. We also investigate heterogeneity in pass-through for sales regions with different income levels. To do so, we enrich our dataset with GDP per capita at the district level. We then interact the net-of-tax factor with dummy variables representing quintiles of the year-specific distribution of district GDP per capita across observations in the estimation sample (similar to market shares as in Section 4.5.2). Panel (a) in Figure 4.4 shows the results. We find that pass-through tends to rise with regional incomes, with the exception of the highest-income regions. However, these differences are statistically insignificant.

Pass-through by retail store type. Our dataset covers four types of retail stores: supermarkets, hypermarkets, drug stores, and discounters.[30] We estimate the heterogeneity in pass-through by again interacting the tax change with a dummy indicating the store type. Panel (b) in Figure 4.4 shows the results. We find significant differences across stores: While drug stores and discounters display both quantitatively and statistically insignificant pass-through, we find that hypermarkets and supermarkets exhibit sizable pass-through of around 50%. This reveals that the significant average pass-through of corporate taxes on retail prices is mainly driven by price adjustments in supermarkets and hypermarkets.[31]

There could be various structural reasons for this result. On the one hand, producers may be following different pricing strategies between discounters and other stores, e.g. they may perceive little market power for sales by the former, or they may be less able to apply price increases to them. On the other hand, if producers don't discriminate across store types and raise prices across the board, retailers may transmit the shocks differently to their customers depending on their own market power, with discounters absorbing price increases into their profit margins, contrary to supermarkets and hypermarkets.

4.6 Conclusion

In this paper we estimate how changes in local corporate tax rates in Germany affect retail prices of products of taxed firms that are exported to other German jurisdictions. We find that a one percentage point increase in the local corporate tax rate leads on average to an approximately 0.4% increase in the municipality's retail "export" price relative to the prices of all other products originating from municipalities in different regions. Our results suggest that upstream firms are able to increase prices to protect their markups, and retailers pass-through the wholesale price increases into higher retail prices. This is remarkable given evidence in Fuest et al. (2018) that wages fall with a corporate tax increase, putting downward pressure on producers' costs.

[30] The types of store are defined as follows. (1) Traditional stores are outlets with a range of goods consisting mainly of groceries (excluding specialty stores) with a surface area from 200 to 799 square metres. This includes supermarkets, which have a surface area larger than 400 square metres, but in the text we use the term "supermarkets" for all traditional stores independent of size. (2) Hypermarkets are self-service retail stores with large surface size (larger than 800 square metres) that are not discounters and offer groceries as well as consumer durables and consumer goods mostly for short to medium-term use. (3) Discounters are self-service stores carrying mainly groceries in a limited range with emphasis on low prices. (4) Drugstores are self-service retail outlets carrying medicines and cosmetics as their core product range.

[31] The heterogeneous effects across store types are not driven by product composition. We directly test for this by including a product by sold-region by year fixed effect as a robustness check, see Table 4.B.4. Thereby we focus on within-product(-region) variation across stores. The differences between store types remain robust.

We find that that firms and products with larger market shares do not exhibit lower pass-through, contrary to theoretical predictions and earlier findings in the context of exchange rate pass-through. We also find that competitor tax changes to not lead to significant price changes. These two findings both suggest that strategic complementarities are weak. Nevertheless, our finding that larger firms are those whose prices react more to corporate taxes is best appreciated in light of the evidence in Fuest et al. (2018) that this set of firms also does not pass-through tax changes to their workers' wages. This is consistent with the idea that market power allows larger firms to choose to shift the tax incidence to their consumers rather than their workers, as done by smaller firms. At any rate, the evidence in Fuest et al. (2018) and in our paper strongly points to the fact that shareholders may be able to shield a great deal of the incidence of corporate taxes, at least in Germany.

We also document substantial heterogeneity in pass-through across store types: While drug stores and discounters do not pass-through price increases, we find significant pass-through of tax changes for prices charged in supermarkets and hypermarkets. In contrast, pass-through heterogeneity across other dimensions, including across product categories and in terms of income in the sales region, is limited.

Appendices for Chapter 4

4.A Data

This appendix describes the data sources used in the paper and how the data is mapped and aggregated. Table 4.A.1 provides an overview of all data sources used. The following sections describe them in detail.

4.A.1 Administrative data

Municipality tax scaling factors. We obtain annual local scaling factors for each municipality (*Gemeinde*) which are provided by the *Statistische Bibliothek* as *Hebesätze der Realsteuern* separately by state and by the years 2003–2018. These files differ slightly across years with respect to their structure, which needs to be taken into account when appending them to one data set.

Municipalities are uniquely identified by *Amtlicher Gemeindeschlüssel* (`AGS`). `AGS` is an eight-digit key that contains identification of a municipality's state (digits 1–2), *Regierungsbezirk* (given state, digit 3), county (*Kreis*, given the state and *Regierungsbezirk*, digits 4–5), and municipality (given the state, county, and *Regierungsbezirk*, digits 6–8).

In the official data, some `AGS` are less than eight digits long (respecting leading zeroes). This is because the records omit the state identifier from the `AGS` which we then add. The `AGS` of Berlin is sometimes erroneously recorded as a ten-digit code; we delete the superfluous lagging zeroes. Some of the `AGS` are not correct based on the fact that they do not begin with the right state identifier. In this

Table 4.A.1: Summary of data sources

Data	Source	Granularity	Identifier	Time
Administrative data:				
Local tax scaling factors (*Hebesätze*)	*Statistische Bibliothek*	Municip.	`AGS` × year	2003–2019
– Local business tax scaling factor				
– Real estate tax A scaling factor				
– Real estate tax B scaling factor				
– Indication of territory reform				
Municipality info	Destatis	Municip.	`AGS` × year	2003–2019
– Postcode of administration	(*Gemeindeverzeichnis*)			
(*Verwaltungssitz*)				
– Population				
Municipality economic data	*Regionaldatenbank*	Municip.	`AGS` × year	2008–2017
– Number of employed				
– Number of unemployed				
County economic indicators	*Regionaldatenbank*	County	5d-`AGS` × year	2010–2019
– Total debt				
– GDP (per capita/per worker)				
Regional maps of Germany	*GeoBasis-DE / BKG*	Municip.	`AGS`	2017
– Municipalities (*VG-250*)				
– States (*NUTS-250*)				
Retail price data:				
Supermarket sales across Germany	IRi	Barcode/	`EAN` × store-ID	2013–2017
– Weekly unit sales		store/time	× 2d-ZIP × week	
– Weekly EUR sales				
– Vendor of product				
– IRi product category				
– two-digit postcode of store				
– IRi store keyaccount				
– IRi store type				
Firm information data:				
GS1 records of individual barcodes	GS1 GEPIR	Barcode	`EAN`	
– Exact firm name				
– City and postcode				
– GS1 Company Prefix				
Orbis data:				
Orbis branch information	Orbis / Bureau Van Dijk	Branch	`bvdidnumber`	
– Branch city				
– Headquarter city				
COICIOP-IRi category mapping:				
COICOP-3 category			`category` (IRi)	

Notes: Regional identifiers: `AGS` is *Amtlicher Gemeindeschlüssel* (official municipality key). *BKG* is the *Bundesamt für Kartografie und Geodäsie.*

case, we use the *GVISys* (*Gemeindeverzeichnis-Informationssystem*) variable to back out the correct `AGS`.

Moreover it contains information about potential territory changes that happened in the corresponding year. We record such indication as a binary indicator.

Municipality information. Additional information on each municipality is provided by Destatis. We obtain these for the years 2003–2018 as well; again, differing column structures have to be taken into account when appending these files. This data contains the total population of the municipality and the postcode, which helps us to map firms to municipalities. However, note that postcodes do not identify municipalities and vice versa. Postcodes are defined by the German postal service *Deutsche Post*. Single municipalities can have many postcodes (in case of a large city), but also one postcode can be attached to many municipalities (small cities). To identify the state of a postcode area, one needs to know up to four digits. The postcode that is part of Destatis data refers to the postcode where a municipality's administration centre (*Verwaltungssitz*) is located. Nevertheless, knowing approximately the postcode of a municipality will help us in matching firms to municipalities.[32]

This data also includes information on unincorporated areas (*gemeindefreie Gebiete*) which are not not governed by a local municipal corporation and hence do not have their own local business tax scaling factor. We effectively ignore these areas.

Municipality (un-)employment data. We obtain the number of employed (subject to social insurance contributions, *sozialversicherungspflichtige Beschäftigte*) and unemployed persons by municipality and year for 2008–2017, which are years relevant for our empirical exercise, from *Regionaldatenbank Deutschland*.

County debt data. We obtain total debt for each county (*Landkreis* or *Kreisfreie Stadt*) and year also from *Regionaldatenbank Deutschland*. Counties are identified by the first five digits of `AGS`. Some counties do not report their debt. In general, this data is only available from 2010 to 2019.

Municipality map of Germany. From the federal cartography office, the *Bundesamt für Kartografie und Geodäsie*, we obtain shape files that allow producing a map of all municipalities in Germany, which we use to illustrate the geographical variation in our data. We use the map as of 2017 for simplicity. Figure 4.A.1 (a) draws the municipality and state borders.

Matched data. We match the municipality scaling factor with the postcode and population data based on `AGS` and year. Table 4.A.2 shows the number of municipalities, thereof "normal" ones and ones with territory changes, across years. Unincorporated areas are ignored by only considering municipalities that are part of the local scaling factor data.

We then match the (un-)employment data based on `AGS` and year. We obtained only the years relevant for our empirical exercise. Within these years, a number of municipalities are missing, as they do not report these numbers. For the remaining municipalities, we compute an (approximative) municipality level unemployment rate as the fraction of unemployed to unemployed and employed.

[32] This data also contains the `ARS` key, which is richer than `AGS`. After digit 5 of the `AGS` a four-digit identifier of a *Gemeindeverband* (municipality union) is inserted. Leaving these digits out of the `ARS` gives the `AGS`. However, it is not necessary for our data mapping.

Table 4.A.2: Number of municipalities across years

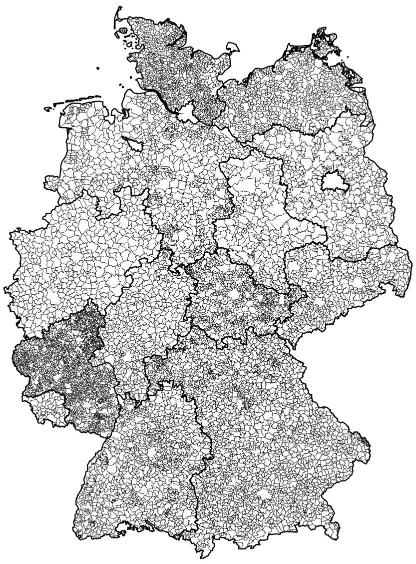
Year	Total	Normal	with UE rate	and with debt	No. of scaling fct. changes
2003	12630	12465			–
2004	12434	12321			1031
2005	12342	12249			1341
2006	12313	12227			991
2007	12266	12194			496
2008	12227	12163	9567		486
2009	11996	11917	8306		528
2010	11442	11312	8215	8209	1031
2011	11294	11179	8315	8309	2016
2012	11224	11113	9033	9027	1443
2013	11161	11058	9000	8994	1390
2014	11117	11025	9633	9627	2153
2015	11093	11037	9599	9593	1698
2016	11059	11007	9842	9836	1465
2017	11055	11011	9842	9837	1178
2018	11014	10959			932
2019	10799	10715			700

Notes: *Normal* municipalities means those without territory change.

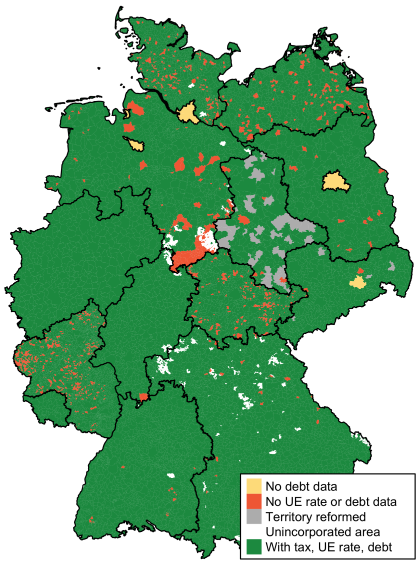
Based on the five-digit `AGS` and year we match the municipality data with the county-level data on total debt. Debt data is available for all counties except 11 (including the city states Berlin, Hamburg, and Bremen), a total of 61 municipality-years between 2010 and 2019.

Table 4.A.2 summarises the number of available municipalities according to data richness. Figure 4.A.1 (b) illustrates the data availability across municipalities for the year 2017.

Figure 4.A.1: Geography of municipalities and data availability



(a) Municipality and state borders



(b) Data availability (year 2017)

Table 4.A.3: An example observation from the raw IRi data

Variable	Example
EAN	40015340025782
store-ID	'63386112'
week-ID	'1875'
unit sales	925
value sales [EUR]	638.25
price per unit [EUR]	0.69
category	BIER
vendor	BINDING
volume	500.00ML
zip	63***
keyaccount id	'4'
store type id	'4'

4.A.2 IRi data

Structure of raw retail scanner price data. The retail scanner price data we use observes weekly sales of individual products, identified by barcodes (`EAN`), in individual stores across Germany. An individual product is, for example, a 500ml can of beer with the barcode 40015340025782. Table 4.A.3 shows one individual observation for such a product in the raw data. The data allows us to observe how often a product was sold in a particular store and a particular week. For example, in the week of August 3, 2015 one store in our data sold 925 units of the 500ml can, and thereby generated a revenue of EUR 638.25. Moreover, the data contains a product category classification (there are 217 categories defined by IRi), a coarse name of the manufacturer (vendor), and store characteristics.

Because of data protection, stores are partly anonymised in our data. That is, we do not know the exact identity of a store but only their approximate location and their type. The approximate location is given by the first two digits of their location postcode. The retailer is given by the IRi keyaccount and store type, which can be hypermarket, supermarket, discount, or drugstore.

By means of comparing the sold units to the value of sales, this implies a store-week specific price-per-unit of

$$p_{i,\text{store},w} = \frac{\text{EUR sales}_{i,\text{store},w}}{\text{unit sales}_{i,\text{store},w}}.$$

In our empirical analysis, however, we aggregate our data from the product-store-week level to the product-retailer type-year level. This has two reasons. First, reducing the number of observations improves computational tractability. Second, tax changes are at the yearly level and we are interested in the medium-run effects on prices, and because stores are identified only up to their approximate location and type, we can aggregate the prices to this level of granularity without losing identifying information. The aggregation is explained next, together with sample selection.

Sample selection and aggregation. We condition on sales data from individual stores and years for which the store was operative throughout the year. That is, we filter out stores for which we see less than 51 weeks recorded across all products. Then, we aggregate price changes to the store *type* by region by year level. Store types are defined by the combination of IRi keyaccount and IRi store type. Regions are defined as two-digit postcode areas.

First, we compute the store-level average price for product i in year t:

$$p_{i,\text{store},t} := \frac{\sum_{w \in t} \text{EUR sales}_{i,\text{store},w}}{\sum_{w \in t} \text{unit sales}_{i,\text{store},w}}$$

Note that this is equivalent to a unit-weighted average across weekly per-unit prices.

Second, we compute the store-level year-over-year price change:

$$\Delta \log p_{i,\text{store},t} = \log(p_{i,\text{store},t}) - \log(p_{i,\text{store},t-1})$$

Third, for a two-digit postcode region r, store type s, and year t, we compute the *average* year-over-year price change (with slight abuse of notation):

$$\Delta \log p_{i,s,r,t} := \frac{1}{N_{(r,s),t}} \sum_{\text{store} \in (r,s)} \Delta \log p_{i,\text{store},t}$$

where $N_{(r,s),t}$ is the number of type s stores in region r in year t.

As explained in the main text, for our diff-in-diff analysis, we only consider price changes observations that refer to a sales location outside of the producer location. Specifically, we exclude product price changes $\Delta \log p_{i,s,r,t}$ which, according to our further data work explained below, are produced by manufacturers that are located in a municipality that belongs to the two-digit postcode region r.

4.A.3 Firm information

Barcode structure and manufacturer identification. Individual products are identified by barcodes, called `EAN` in IRi data. EAN stands for European Articel Number. Barcodes around the world are administrated by the firm GS1. According to GS1, the term EAN was superseeded by the GTIN concept, which stands for Global Trade Item Number. In this paper, we call `EAN` the barcode identifier in IRi data and `GTIN` the *equivalent* barcode registered with IRi. `EAN`s can be converted into the `GTIN` form by removing digits 2–3 and adding a check digit according to a known formula. This formula is explained at `https://www.gs1.org/services/how-calculate-check-digit-manually`.

The `GTIN` contains two important pieces of information with respect to the producer of the firm, which by definition maintained throughout the paper, is the firm that registered the product with GS1. First, it identifies the country location of the producer through the first three digits of the barcode. In particular, German producers are identified by digits 400–440. The meanings of all country prefixes are listed at `https://www.gs1.org/standards/id-keys/company-prefix`.

The product barcode also identifies the producer by the company prefix. Whenever a firm becomes a member of GS1, in order to register barcodes, it obtains a company prefix with which all registered barcodes begin. This company prefix is usually seven digits long, but can also be up to eleven digits long. The length of the company prefix cannot be inferred directly. We learn the company prefix precisely in the barcode request step explained below.

Table 4.A.4: Example: IRi `EAN`, GS1 `GTIN` and country/company identification

(1)	IRi `EAN`:	40015340025782
(2)	Remove digits 3–4:	405340025782
(3)	Add check digit to get GS1 `GTIN`:	4053400257822
(4)	Identify country and company:	$\underbrace{\underbrace{405}_{\text{country}}\ 3400}_{\text{company}}\underbrace{257822}_{\text{product}}$

For illustration, Table 4.A.4 uses the example of a can of beer to illustrate the conversion of `EAN` to `GTIN`.

Selection of individual firm information obtained. We want to learn the company identification prefix and the company-related information in the GS1 database for all German products in our sample. We focus on German firms because they are all subject to the same corporate taxation. To this end, we select all barcodes that start with digits 400–440, which are the country prefixes for Germany.

We select a subsample of barcodes that is intended to cover all distinct producers in the sample. At this point we have not obtained firm information for all barcodes individually because downloading this information for more than 150,000 barcodes was infeasible. Instead, we select a subset of `GTIN` barcodes that (i) start with distinct seven-digit sequences and (ii) have distinct vendor names in the IRi data. The first property makes sure to select one `GTIN` for every producer, if all company prefixes are seven digits long. However, since some are longer, but this is not visible from the barcode directly, we impose the second property which means that if the first seven digits are the same but the vendor information differs, we sample multiple `GTIN`s, with the intention to obtain information on (at least) one barcode per actual producer.

Information request for barcodes from GS1 GEPIR. Ultimately, we request information for 11,693 individual barcodes from the commercial database GS1 GEPIR. The majority of queries, roughly 75%, is successful, yielding company prefix and company information. The remaining quarter of queries is not successful for a variety of reasons. Table 4.A.5 lists the split-up. Most importantly, some company information is not made public by GS1 (row 2). Some barcodes are outdated and cannot be obtained any more (row 3) or are invalid (row 4 and row 5). For some barcodes, the returned company prefix does not match with the

Table 4.A.5: Success of individual information requests

Return Code	No.
Query Successful	8,384
Company information witheld	1,492
Prefix no longer subscribed	949
Record not found	636
Unknown GS1 Prefix	6
Company prefix mismatch	5
Query successful but links to GS1 company information	221
Total	11,693

requested barcodes (row 6). We also drop such pathological cases. Lastly, some barcode requests are successful, but the barcode contains only the information about GS1 itself (row 7). We also ignore these.

Note that the 8,384 successful queries are for individual barcodes, which are partly produced by the same firm. Ex-post we find that we have obtained information for barcodes of 5951 different firms, based on the GS1 company prefix.

Attaching firm information to remaining barcodes. For the 8,384 barcodes for which we successfully gathered firm information, we attach the received producer information back to all barcodes in the following way. The information contains the exact company prefix, which can be seven digits or longer. Based on this, we attach this information to all products for which the `GTIN` starts with this sequence.

Using postal addresses to determine municipality. The information contains for every producer their address including the postcode and city name. However, this information does not map easily into municipalities. Complications arise because cities/municipalities can have multiple postcodes, so the postcode in the administrative data does not need to match the postcode of the firm address. Municipalities may also have "suburbs" that show up as firm locations or the cities are spelled slightly differently, e.g., by omitting parts of the official municipality name (e.g., Frankfurt instead of Frankfurt am Main).

We first prepare the administrative data as follows: We remove all parts of the municipality names that describe the city level, i.e.: ", Stadt", ", St.", ", Hansestadt", ", Landeshauptstadt", "Universitätsstadt", ", Hochschulstadt", ", Kreisstadt", ", Wissenschaftsstadt", ", Univeristäts- und Hansestadt", ", gr.kr.St". Moreover we remove all suffixes in brackets (such as "(Main)") and replace both Frankfurt am Main and Frankfurt an der Oder by "Frankfurt", and later distinguish the two based on the different postcodes. We also remove municipality-years with territory reforms.

The official data contains two instances where two `AGS` have the same municipality name and postcode, resepectively: Hamfelde (`AGS` 01053049 and 0153070) and Köthel (`AGS` 01062026 and 01062040). We delete these from the data before matching to firms.

To match firms to municipalities, we rely on municipality names and postcodes. For a match to be valid, we require that the first two digits of the firm's postcode and the municipality postcode are the same. We then match based on municipality names if the municipality name is unique. If it is not unique, we additionally use the first two digits of the postcode if the combination therewith is unique, otherwise the also the third digit, and so on. This way, we are able to match 5018 of 5951 firms.

In a second step, we use the Stata function `matchit` to match firms' city to municipalities using fuzzy string matching. This algorithm accounts for typos in the firm locations and other slight perturbations of the city names. The algorithm produces a number of candidate matches with associated similarity scores. We drop candidate matches if the first digit of the postcodes do not match. Of the remaining candidates, we directly accept matches it if turns out that the address city name is an exact match to the corresponding first part of the municipality name (e.g., Radolfzell instead of Radolfzell am Bodensee). We then focus on matches with the highest similarity score. If postcodes match exactly, we accept the match. Apart from this, we accept matches with a similarity score of more than 0.75 and screen each match manually. This increases the number of matched firms by another 412 to 5430, i.e., 91% of the ones identified in the producer-level information.

4.A.4 Orbis data

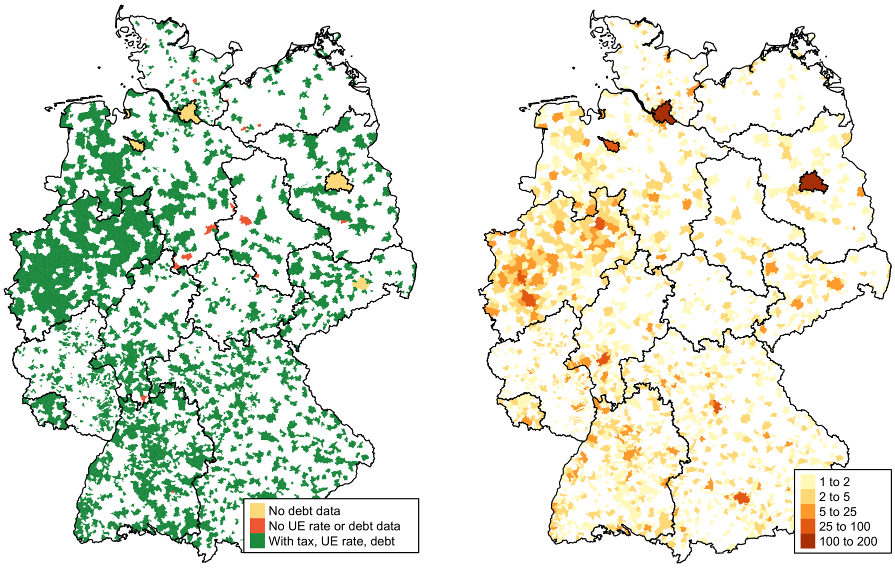
Matching to Orbis based on firm name and location. To match the firm information from the web information to Orbis data, we use the matching software on the web platform of Orbis. We supply the tool with firm name and location, which the tool matches to Orbis records, yielding the Orbis identifier `bvdidnumber`. We manually go through all matches and check them for correctness. We find 4585 matches, i.e., 77%, in the Orbis database.

Work with Orbis branch information. Orbis data contains information about branches of firms. We check if for a given `bvdidnumber` there are multiple branch cities which are different from the firm's main city. In this case we record it as a multi-branch firm. Of the firms we identify in the previous step and linked to Orbis, 74% have more than one branch.

4.A.5 Matched dataset

We ultimately enrich the IRi price data with the additional data sources described above. Table 4.1 (in the main text) summarises the sample after each step. First, we condition on German barcodes, i.e., `EAN`s starting with digits 40–44. This reduces the sample of products, as shown by row 2 in the table. Second, we attach the producer–municipality data. This step includes the matching of producer information to products and the matching of municipalities to producers, as ex-

Figure 4.A.2: Geographic coverage in matched data



(a) Data availability (2017) (b) Number of firms (2017)

plained above. This leads to the sub-population of products described by row 3. Finally, we also attach the Orbis information, which leads to row 4.

The matched data covers production in all regions of Germany with no abnormal geographic clustering, as shown by Figure 4.A.2. North Rhine-Westphalia stands out in being especially densely covered. The number of firms in individual municipalities varies between one firm for most to up to 173 in Hamburg.

4.B Additional results

Table 4.B.1: Results when controlling for changes in local real estate taxes

	(1) Δ log price	(2) Δ log price	(3) Δ log price
$-\Delta \log(1 - \text{corporate tax})$	0.488**	0.524***	0.531***
	(0.208)	(0.203)	(0.204)
Δ scaling factor real estate tax A	-0.00303		-0.00163
	(0.00201)		(0.00218)
Δ scaling factor real estate tax B		-0.00349	-0.00253
		(0.00214)	(0.00244)
Observations	14091803	14091803	14091803
Product FE	✓	✓	✓
Sold-region × year FE	✓	✓	✓
Production-region × year FE	✓	✓	✓
Production-muni. UE controls	✓	✓	✓
Production-district debt controls	✓	✓	✓

Notes: *Real estate tax A* refers to the tax on arable land. *Real estate tax B* refers to the tax on built-up land. See also Table 4.1.

Table 4.B.2: Comparing results with $\Delta\,\tau$ and $\Delta \log(1 - \tau)$

	(1) Δ log price	(2) Δ log price	(3) Δ log price	(4) Δ log price	(5) Δ log price	(6) Δ log price
$-\Delta \log(1 - \text{tax})$	0.525***		0.538***		0.425**	
	(0.171)		(0.182)		(0.209)	
Δ tax		0.606***		0.622***		0.490**
		(0.198)		(0.211)		(0.242)
Observations	19434155	19434155	18871628	18871628	14091803	14091803
Product FE	✓	✓	✓	✓	✓	✓
Sold-region × year FE	✓	✓	✓	✓	✓	✓
Production-region × year FE	✓	✓	✓	✓	✓	✓
Production-muni. UE controls			✓	✓	✓	✓
Production-district debt controls					✓	✓

Notes: See Table 4.1.

Table 4.B.3: Comparing results with different trimmings and with sales filtering

(a) Posted prices (baseline)

	(1) $(p1, p99)$	(2) $(-0.33, 0.33)$	(3) $(-0.2, 0.2)$	(4) $(-0.5, 0.5)$
$-\Delta\log(1-\text{tax})$	0.425^{**}	0.452^{**}	0.393^{**}	0.482^{**}
	(0.209)	(0.201)	(0.178)	(0.209)
Observations	14091803	13998007	13528490	14183456
Product FE	✓	✓	✓	✓
Sold-region × year FE	✓	✓	✓	✓
Production-region × year FE	✓	✓	✓	✓
Production-muni. UE controls	✓	✓	✓	✓
Production-district debt controls	✓	✓	✓	✓

(b) Sales-filtered prices

	(1) $(p1, p99)$	(2) $(-0.33, 0.33)$	(3) $(-0.2, 0.2)$	(4) $(-0.5, 0.5)$
$-\Delta\log(1-\text{tax})$	0.375^{*}	0.393^{**}	0.352^{**}	0.419^{**}
	(0.205)	(0.196)	(0.174)	(0.204)
Observations	14092680	13992737	13519954	14182186
Product FE	✓	✓	✓	✓
Sold-region × year FE	✓	✓	✓	✓
Production-region × year FE	✓	✓	✓	✓
Production-muni. UE controls	✓	✓	✓	✓
Production-district debt controls	✓	✓	✓	✓

Notes: Panel (a) uses observed, posted prices as in our baseline. Panel (b) uses price changes based on a simple V-filter at weekly frequency. Column (1) represents the baseline data treatment where price changes are trimmed at the year-specific 1% and 99% quantiles. Columns (2)-(4) represent different trimmings, where, price changes are trimmed instead at alternative absolute cut-offs. See also Table 4.1.

Table 4.B.4: Heterogeneous pass-through across retail store types: Product-region-year FE

	(1) Δ log price	(2) Δ log price
Discounter $\times -\Delta\log(1-\text{tax})$	-0.340	
	(0.373)	
Drug store $\times -\Delta\log(1-\text{tax})$	-0.00755	0.520
	(0.236)	(0.395)
Supermarket $\times -\Delta\log(1-\text{tax})$	0.493**	0.929**
	(0.194)	(0.459)
Hypermarket $\times -\Delta\log(1-\text{tax})$	0.698***	1.284***
	(0.173)	(0.452)
Observations	19434155	14677639
Product FE	yes	(red.)
Sold-region $\times$ year FE	yes	(red.)
Production-region $\times$ year FE	yes	(red.)
Product $\times$ sold-region $\times$ year FE	no	yes

Notes: Column (1) repeats the estimates shown in Figure 4.4 (b). Column (2) adds a product by sold-region by year FE. *(red.)* indicates that other fixed effects and regressors become redundant due to this. The disounter-specific coefficient is used as the base category and becomes unidentified.

Bibliography

Adam, K. and R. M. Billi (2006): "Optimal Monetary Policy under Commitment with a Zero Bound on Nominal Interest Rates," *Journal of Money, Credit and Banking*, 38(7), 1877–1905.

Adam, K. and A. Marcet (2011): "Internal Rationality, Imperfect Market Knowledge and Asset Prices," *Journal of Economic Theory*, 146, 1224–1252.

Adam, K., A. Marcet, and J. Beutel (2017): "Stock Price Booms and Expected Capital Gains," *Amercian Economic Review*, Vol. 107 No. 8, 2352–2408.

Adam, K., D. Matveev, and S. Nagel (2021): "Do Survey Expectations of Stock Returns Reflect Risk-Adjustments?" *Journal of Monetary Economics*, 117, 723–740.

Adam, K. and S. Nagel (2022): "Expectations Data in Asset Pricing," in *Handbook of Economic Expectations*, Elsevier, forthcoming.

Adam, K. and H. Weber (2022): "Estimating the optimal inflation target from trends in relative prices," *American Economic Journal: Macroeconomics*, forthcoming.

Adam, K. and M. Woodford (2012): "Robustly optimal monetary policy in a microfounded New Keynesian model," *Journal of Monetary Economics*, 59, 468–487.

——— (2021): "Robustly Optimal Monetary Policy in a New Keynesian Model with Housing," *Journal of Economic Theory*, 198, 105352.

Adler, K. (2020): "Financial covenants, firm financing, and investment," Working Paper.

Admati, A. R., P. M. Demarzo, M. F. Hellwig, and P. Pfleiderer (2018): "The leverage ratchet effect," *Journal of Finance*, 73, 145–198.

Aguiar, M. and M. Amador (2020): "Self-fulfilling debt dilution: Maturity and multiplicity in debt models," *American Economic Review*, 110, 2783–2818.

Aguiar, M., M. Amador, H. Hopenhayn, and I. Werning (2019): "Take the short route: Equilibrium default and debt maturity," *Econometrica*, 87, 423–462.

ALMEIDA, H., M. CAMPELLO, B. LARANJEIRA, AND S. WEISBENNER (2012): "Corporate debt maturity and the real effects of the 2007 credit crisis," *Critical Finance Review*, 1, 3–58.

ALTINKILIÇ, O. AND R. S. HANSEN (2000): "Are there economies of scale in underwriting fees? Evidence of rising external financing costs," *Review of Financial Studies*, 13, 191–218.

AMITI, M., O. ITSKHOKI, AND J. KONINGS (2019): "International Shocks, Variable Markups, and Domestic Prices," *Review of Economic Studies*, 86, 2356–2402.

ANDERSON, E., S. REBELO, AND A. WONG (2020): "Markups across space and time," Working Paper.

ANDERSON, G. AND A. CESA-BIANCHI (2020): "Crossing the credit channel: credit spreads and firm heterogeneity," Bank of England Staff Working Paper No. 854.

ANDRADE, P., J. GALÍ, H. LE BIHAN, AND J. MATHERON (2019): "The Optimal Inflation Target and the Natural Rate of Interest," *Brookings Papers on Economic Activity*, 2019(2), 173–255.

——— (2021): "Should the ECB Adjust its Strategy in the Face of Lower r*?" *Journal of Economic Dynamics and Control*, forthcoming.

ANGELETOS, G.-M., Z. HUO, AND K. A. SASTRY (2020): "Imperfect Macroeconomic Expectations: Evidence and Theory," in *NBER Macroeconomics Annual 2020, Vol. 35.*

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

ARELLANO, C., Y. BAI, AND L. BOCOLA (2020): "Sovereign default risk and firm heterogeneity," NBER Working Paper No. 23314.

ARELLANO, C., Y. BAI, AND P. J. KEHOE (2019): "Financial frictions and fluctuations in volatility," *Journal of Political Economy*, 127, 2049–2103.

ARELLANO, C. AND A. RAMANARAYANAN (2012): "Default and the maturity structure in sovereign bonds," *Journal of Political Economy*, 120, 187–232.

ARMONA, L., A. FUSTER, AND B. ZAFAR (2018): "Home Price Expectations and Behaviour: Evidence from a Randomized Information Experiment," *Review of Economic Studies*, 86, 1371–1410.

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.

Atkeson, A. and A. Burstein (2008): "Pricing-to-Market, Trade Costs, and International Relative Prices," *American Economic Review*, 98, 1998–2031.

Auclert, A. (2019): "Monetary policy and the redistribution channel," *American Economic Review*, 109, 2333–67.

Auer, R. A. and R. S. Schoenle (2016): "Market structure and exchange rate pass-through," *Journal of International Economics*, 98, 60–77.

Auerbach, A. J. (2006): "Who Bears the Corporate Tax? A Review of What We Know," *NBER Chapters*, 1–40.

Baker, S., S. T. Sun, and C. Yannelis (2020): "Corporate Taxes and Retail Prices," *National Bureau of Economic Research.*

Baley, I. and A. Blanco (2021): "Aggregate dynamics in lumpy economies," *Econometrica*, 89, 1235–1264.

Baqaee, D., E. Farhi, and K. Sangani (2021): "The Supply-Side Effects of Monetary Policy," NBER Working Paper No. 28345.

Baqaee, D. R. and E. Farhi (2017): "Productivity and Misallocation in General Equilibrium," NBER Working Paper No. 24007.

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

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2015): "X-CAPM: An Extrapolative Capital Asset Pricing Model," *Journal of Financial Economics*, 115, 1–24.

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

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

Barsky, R. B., C. L. House, and M. S. Kimball (2007): "Sticky-Price Models and Durable Goods," *American Economic Review*, 97, 984–998.

Basu, S. (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.

Bayer, C., R. Lütticke, L. Pham-Dao, and V. Tjaden (2019): "Precautionary savings, illiquid assets, and the aggregate consequences of shocks to household income risk," *Econometrica*, 87, 255–290.

Becker, S. O., P. H. Egger, and V. Merlo (2012): "How low business tax rates attract MNE activity: Municipality-level evidence from Germany," *Journal of Public Economics*, 96, 698–711.

Begenau, J. and J. Salomao (2018): "Firm financing over the business cycle," *Review of Financial Studies*, 32, 1235–1274.

Benigno, P. and L. Paciello (2014): "Monetary Policy, Doubts and Asset Prices," *Journal of Monetary Economics*, 64, 85–98.

Benmelech, E., C. Frydman, and D. Papanikolaou (2019): "Financial frictions and employment during the great depression," *Journal of Financial Economics*, 133, 541–563.

Benmelech, E., N. Kumar, and R. Rajan (2020): "The decline of secured debt," NBER Working Paper No. 26637.

Berger, D., K. Milbradt, F. Tourre, and J. Vavra (2021): "Mortgage prepayment and path-dependent effects of monetary policy," *American Economic Review*, 111, 2829–78.

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

Bizer, D. S. and P. M. DeMarzo (1992): "Sequential banking," *Journal of Political Economy*, 100, 41–61.

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

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. J. Terry (2018): "Really uncertain business cycles," *Econometrica*, 86, 1031–1065.

Bocola, L. and A. Dovis (2019): "Self-fulfilling debt crises: A quantitative analysis," *American Economic Review*, 109, 4343–77.

Bond, S., A. Hashemi, G. Kaplan, and P. Zoch (2021): "Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data," *Journal of Monetary Economics*, 121, 1–14.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020): "Overreaction to Macroeconomic Expectations," *American Economic Review*, 110, 2748–82.

Bronnenberg, B. J., M. W. Kruger, and C. F. Mela (2008): "Database paper—The IRI marketing data set," *Marketing Science*, 27, 745–748.

Buera, F. and S. Karmakar (2021): "Real effects of financial distress: The role of heterogeneity," *Economic Journal*, forthcoming.

Caggese, A., A. Gutierrez, and A. Pérez-Orive (2019): "Firm debt deflation, household precautionary savings, and the amplification of aggregate shocks," Working Paper.

Caines, C. and F. Winkler (2021): "Asset Price Beliefs and Optimal Monetary Policy," *Journal of Monetary Economics*, 123, 53–67.

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

Case, K. E., R. J. Shiller, and A. K. Thompson (2012): "What Have They Been Thinking? Homebuyer Behavior in Hot and Cold Markets," *Brookings Papers on Economic Activity*, 265–315.

Castro Cienfuegos, N. and F. Loria (2017): "A New Keynesian Perspective on Total Factor Productivity via Production Networks," Working Paper.

Chatterjee, S. and B. Eyigungor (2012): "Maturity, indebtedness, and default risk," *American Economic Review*, 102, 2674–99.

Chodorow-Reich, G. and A. Falato (2021): "The loan covenant channel: How bank health transmits to the real economy," *Journal of Finance*, forthcoming.

Chow, G. and A.-l. Lin (1971): "Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series," *Review of Economics and Statistics*, 53, 372–75.

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.

Cinelli, C. and C. Hazlett (2020): "Making sense of sensitivity: extending omitted variable bias," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 82, 39–67.

Clarida, R., J. Galí, and M. Gertler (1999): "The Science of Monetary Policy: Evidence and Some Theory," *Journal of Economic Literature*, 37, 1661–1707.

Cloyne, J., C. Ferreira, M. Froemel, and P. Surico (2020): "Monetary policy, corporate finance and investment," *Journal of the European Economic Association*, forthcoming.

Coibion, O. and Y. Gorodnichenko (2015): "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," *American Economic Review*, 105, 2644–78.

Coibion, O., Y. Gorodnichenko, and J. Wieland (2012): "The Optimal Inflation Rate in New Keynesian Models: Should Central Banks Raise Their Inflation Targets in Light of the Zero Lower Bound?" *Review of Economic Studies*, 79, 1371–1406.

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

Comin, D. and M. Mestieri (2018): "If Technology Has Arrived Everywhere, Why Has Income Diverged?" *American Economic Journal: Macroeconomics*, 10, 137–78.

Cooley, T. F. and V. Quadrini (2006): "Monetary policy and the financial decisions of firms," *Economic Theory*, 27, 243–270.

Corsetti, G. and L. Dedola (2005): "A macroeconomic model of international price discrimination," *Journal of International Economics*, 67, 129–155.

Covas, F. and W. J. Den Haan (2012): "The role of debt and equity finance over the business cycle," *Economic Journal*, 122, 1262–1286.

Crouzet, N. (2017): "Default, debt maturity, and investment dynamics," Working Paper.

——— (2018): "Aggregate implications of corporate debt choices," *Review of Economic Studies*, 85, 1635–1682.

Crouzet, N. and F. Tourre (2021): "Can the cure kill the patient? Corporate credit interventions and debt overhang," Working Paper.

Dangl, T. and J. Zechner (2021): "Debt maturity and the dynamics of leverage," *Review of Financial Studies*, 34, 5796–5840.

Darmouni, O., O. Giesecke, and A. Rodnyansky (2021): "The bond lending channel of monetary policy," Working paper.

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

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

Decker, R., J. Haltiwanger, R. Jarmin, and J. Miranda (2014): "The role of entrepreneurship in U.S. job creation and economic dynamism," *Journal of Economic Perspectives*, 28, 3–24.

DeMarzo, P. M. and Z. He (2021): "Leverage dynamics without commitment," *Journal of Finance*, 76, 1195–1250.

Deng, M. and M. Fang (2022): "Debt maturity heterogeneity and investment responses to monetary policy," *European Economic Review*, forthcoming.

Drechsel, T. (2019): "Earnings-based borrowing constraints and macroeconomic fluctuations," Working Paper.

Duchin, R., O. Ozbas, and B. A. Sensoy (2010): "Costly external finance, corporate investment, and the subprime mortgage credit crisis," *Journal of Financial Economics*, 97, 418–435.

Edmond, C., V. Midrigan, and D. Y. Xu (2021): "How costly are markups?" Working Paper.

Eggertsson, G. and S. Singh (2020): "Log-Linear Approximation versus an Exact solution at the ZLB in the New Keynesian Model," *Journal of Economic Dynamics and Control*, forthcoming.

Eggertsson, G. and M. Woodford (2003): "The Zero Interest-Rate Bound and Optimal Monetary Policy," *Brookings Papers on Economic Activity*, (1), 139–211.

Eichenbaum, M., S. Rebelo, and A. Wong (2022): "State-dependent effects of monetary policy: the refinancing channel," *American Economic Review*, 112, 721–61.

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.

Evans, C. L. and F. T. Santos (2002): "Monetary policy shocks and productivity measures in the G-7 countries," *Portuguese Economic Journal*, 1, 47–70.

Fabiani, A., L. Falasconi, and J. Heineken (2022): "Monetary policy and corporate debt maturity," Working Paper.

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.

Flynn, Z., J. Traina, and A. Gandhi (2019): "Measuring Markups with Production Data," Working Paper.

Fuest, C., A. Peichl, and S. Siegloch (2018): "Do Higher Corporate Taxes Reduce Wages? Micro Evidence from Germany," *American Economic Review*, 108, 393–418.

Fujiwara, S., Y. Iwasaki, I. Muto, K. Nishizaki, and N. Sudo (2016): "Developments in the Natural Rate of Interest in Japan," *Bank of Japan Review*.

Garga, V. and S. R. Singh (2021): "Output hysteresis and optimal monetary policy," *Journal of Monetary Economics*, 117, 871–886.

Gertler, M. and S. Gilchrist (1994): "Monetary policy, business cycles, and the behavior of small manufacturing firms," *Quarterly Journal of Economics*, 109, 309–340.

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

Gilchrist, S., J. W. Sim, and E. Zakrajšek (2014): "Uncertainty, financial frictions, and investment dynamics," NBER Working Paper No. 20038.

Gilchrist, S. and E. Zakrajšek (2012): "Credit spreads and business cycle fluctuations," *American Economic Review*, 102, 1692–1720.

Gomes, J., U. Jermann, and L. Schmid (2016): "Sticky leverage," *American Economic Review*, 106, 3800–3828.

Gomes, J. F. and L. Schmid (2021): "Equilibrium asset pricing with leverage and default," *Journal of Finance*, 76, 977–1018.

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.

Gornemann, N., K. Kuester, and M. Nakajima (2016): "Doves for the rich, hawks for the poor? Distributional consequences of monetary policy," Working Paper.

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

Gravelle, J. G. (2013): "Corporate Tax Incidence: Review of General Equilibrium Estimates and Analysis," *National Tax Journal*, 66, 185–214.

Gravelle, J. G. and L. J. Kotlikoff (1988): "Does the Harberger Model Greatly Understate the Excess Burden of the Corporate Tax? - Another Model Says Yes," NBER Working Paper No. 2742.

Green, D. (2018): "Corporate refinancing, covenants, and the agency cost of debt," Working Paper.

Greenwald, D. L. (2019): "Firm debt covenants and the macroeconomy: The interest coverage channel," Working Paper.

Greenwald, D. L., J. Krainer, and P. Paul (2021): "The credit line channel," Working Paper.

Greenwood, R. and A. Shleifer (2014): "Expectations of Returns and Expected Returns," *Review of Financial Studies*, 27 (3), 714–746.

Gurkaynak, R. S., G. Karasoy Can, and S. S. Lee (2021): "Stock market's assessment of monetary policy transmission: the cash flow effect," *Journal of Finance*, forthcoming.

Gürkaynak, R. S., B. Sack, and E. Swanson (2005): "The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models," *American economic review*, 95, 425–436.

Gutiérrez, G. and T. Philippon (2017): "Declining Competition and Investment in the US," NBER Working Paper No. 23583.

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

Harberger, A. C. (1962): "The Incidence of the Corporation Income Tax," *Journal of Political Economy*, 70, 215–240.

Hatchondo, J. C. and L. Martinez (2009): "Long-duration bonds and sovereign defaults," *Journal of International Economics*, 79, 117–125.

Hatchondo, J. C., L. Martinez, and C. Sosa-Padilla (2016): "Debt dilution and sovereign default risk," *Journal of Political Economy*, 124, 1383–1422.

Holston, K., T. Laubach, and J. C. Williams (2017): "Measuring the natural rate of interest: International trends and determinants," *Journal of International Economics*, 108, S59 – S75, 39th Annual NBER International Seminar on Macroeconomics.

Hong, G. H. and N. Li (2017): "Market structure and cost pass-through in retail," *Review of Economics and Statistics*, 99, 151–166.

Hovakimian, A., A. Kayhan, and S. Titman (2011): "Are corporate default probabilities consistent with the static trade-off theory?" *Review of Financial Studies*, 25, 315–340.

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

Ippolito, F., A. K. Ozdagli, and A. Perez-Orive (2018): "The transmission of monetary policy through bank lending: the floating rate channel," *Journal of Monetary Economics*, 95, 49–71.

Ivashina, V. and B. Vallee (2020): "Weak credit covenants," NBER Working Paper No. 27316.

Jarociński, M. and P. Karadi (2020): "Deconstructing monetary policy surprises - The role of information shocks," *American Economic Journal: Macroeconomics*, 12, 1–43.

Jeenas, P. (2019): "Firm balance sheet liquidity, monetary policy shocks, and investment dynamics," Working paper.

Jermann, U. J. and H. Xiang (2021): "Dynamic banking with non-maturing deposits," Working Paper.

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

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

Jungherr, J. and I. Schott (2021): "Optimal debt maturity and firm investment," *Review of Economic Dynamics*, 42, 110–132.

——— (2022): "Slow debt, deep recessions," *American Economic Journal: Macroeconomics*, 1, 224–259.

Kalemli-Ozcan, S., L. Laeven, and D. Moreno (2018): "Debt Overhang, Rollover Risk, and Corporate Investment: Evidence from the European Crisis," NBER Working Paper No. 24555.

Kaplan, G., K. Mitman, and G. L. Violante (2020): "The Housing Boom and Bust: Model Meets Evidence," *Journal of Political Economy*, 128, 3285–3345.

Kaplan, G., B. Moll, and G. L. Violante (2018): "Monetary policy according to HANK," *American Economic Review*, 108, 697–743.

Kara, E. (2015): "The reset inflation puzzle and the heterogeneity in price stickiness," *Journal of Monetary Economics*, 76, 29–37.

Karabarbounis, M. and P. Macnamara (2021): "Misallocation and financial frictions: The role of long-term financing," *Review of Economic Dynamics*, 40, 44–63.

KHAN, A., T. SENGA, AND J. K. THOMAS (2016): "Default risk and aggregate fluctuations in an economy with production heterogeneity," Working Paper.

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.

KING, R. G. AND S. T. REBELO (1999): "Resuscitating real business cycles," *Handbook of macroeconomics*, 1, 927–1007.

KLEIN, P., P. KRUSELL, AND J.-V. RÍOS-RULL (2008): "Time-consistent public policy," *Review of Economic Studies*, 75, 789–808.

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

KOBY, Y. AND C. WOLF (2020): "Aggregation in heterogeneous-firm models: theory and measurement," Working Paper.

KOHLHAS, A. N. AND A. WALTHER (2021): "Asymmetric Attention," *American Economic Review*, 119, 2879–2925.

KUCHLER, T. AND B. ZAFAR (2019): "Personal Experiences and Expectations about Aggregate Outcomes," *Journal of Finance*, 75, 2491–2542.

LELAND, H. (1994): "Bond prices, yield spreads, and optimal capital structure with default risk," Finance Working paper No. 240, Institute of Business and Economic Research, University of California, Berkeley.

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

LIAN, C. AND Y. MA (2021): "Anatomy of corporate borrowing constraints," *Quarterly Journal of Economics*, 136, 229–291.

LJUNGQVIST, A. AND M. SMOLYANSKY (2016): "To Cut or Not to Cut? On the Impact of Corporate Taxes on Employment and Income," Working Paper.

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

MA, C. (2020): "Momentum and Reversion to Fundamentals: Are They Captured by Subjective Expectations of House Prices?" *Journal of Housing Economics*, 49, 101687.

Malmendier, U. and S. Nagel (2011): "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking," *Quarterly Journal of Economics*, 126, 373–416.

——— (2015): "Learning from Inflation Experiences," *Quarterly Journal of Economics*, 131, 53–87.

Marcet, A. and R. Marimon (2019): "Recursive contracts," *Econometrica*, 87, 1589–1631.

McKenzie, K. J. and E. Ferede (2017): "The Incidence of the Corporate Income Tax on Wages: Evidence from Canadian Provinces," *SPP Research Papers*, 10.

Meier, M. (2020): "Supply Chain Disruptions, Time To Build, and the Business Cycle," Working Paper.

Meier, M. and T. Reinelt (2020): "Monetary policy, markup dispersion, and aggregate TFP," Working Paper.

Mele, A., K. Molnar, and S. Santoro (2020): "On the perils of stabilizing prices when agents are learning," *Journal of Monetary Economics*, 115, 339–353.

Meyer-Gohde, A. (2014): "Risky linear approximations," Working Paper.

Miranda-Agrippino, S. and G. Ricco (2021): "The transmission of monetary policy shocks," *American Economic Journal: Macroeconomics*, 13, 74–107.

Molnar, K. and S. Santoro (2014): "Optimal monetary policy when agents are learning," *European Economic Review*, 66, 39–62.

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

Myers, S. C. (1977): "Determinants of corporate borrowing," *Journal of Financial Economics*, 5, 147 – 175.

Myers, S. C. and N. S. Majluf (1984): "Corporate financing and investment decisions when firms have information that investors do not have," *Journal of Financial Economics*, 13, 187–221.

Nagel, S. and Z. Xu (2018): "Asset Pricing with Fading Memory," Working Paper.

Nakamura, E. and J. Steinsson (2010): "Monetary Non-neutrality in a Multisector Menu Cost Model," *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," *Quarterly Journal of Economics*, 133, 1933–1980.

Nekarda, C. J. and V. A. Ramey (2020): "The Cyclical Behavior of the Price-Cost Markup," *Journal of Money, Credit and Banking*, 52, 319–353.

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," *Econometrica*, 88, 2473–2502.

Pasten, E., R. Schoenle, and M. Weber (2020): "The propagation of monetary policy shocks in a heterogeneous production economy," *Journal of Monetary Economics*, 116, 1–22.

Perla, J., C. Pflueger, and M. Szkup (2020): "Doubling down on debt: limited liability as a financial friction," NBER Working Paper No. 27747.

Piazzesi, M. and M. Schneider (2006): *Equilibrium Yield Curves*, University of Chicago Press, vol. 21 of *NBER Macroeconomics Annual*, chap. 6, 389–442.

Plagborg-Møller, M. and C. K. Wolf (2021): "Local Projections and VARs Estimate the Same Impulse Responses," *Econometrica*, 89, 955–980.

Poeschl, J. (2020): "Corporate debt maturity and investment over the business cycle," Working Paper.

Reiter, M. (2009): "Solving heterogeneous-agent models by projection and perturbation," *Journal of Economic Dynamics and Control*, 33, 649–665.

Reiter, M. and L. Zessner-Spitzenberg (2020): "Long-term bank lending and the transfer of aggregate risk," Working Paper.

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

Rubbo, E. (2020): "Networks, Phillips Curves and Monetary Policy," Working Paper.

Schmitt-Grohé, S. and M. Uribe (2004): "Solving dynamic general equilibrium models using a second-order approximation to the policy function," *Journal of Economic Dynamics and Control*, 28, 755–775.

Stambaugh, R. F. (1999): "Predictive Regressions," *Journal of Financial Economics*, 54, 375–421.

Svensson, L. E. O. (2018): *Festschrift in Honour of Vitor Constancio*, European Central Bank, chap. The Future of Monetary Policy and Macroprudential Policy, 69–123.

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," Working Paper.

Winberry, T. (2021): "Lumpy investment, business cycles, and stimulus policy," *American Economic Review*, 111, 364–96.

Wong, A. (2019): "Refinancing and the transmission of monetary policy to consumption," Working Paper.

Woodford, M. (2003): *Interest and Prices*, Princeton University Press.

Xiang, H. (2020): "Time inconsistency and financial covenants," Working Paper.

Yun, T. (1996): "Nominal Price Rigidity, Money Supply Endogeneity, and Business Cycles," *Journal of Monetary Economics*, 27, 345–370.

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.

Curriculum Vitae

2016–2022	University of Mannheim Center for Doctoral Studies in Economics *Ph.D. in Economics (Dr. rer. pol.)*
2014–2016	University of Oxford *M.Phil. in Economics*
2012	Brown University *Visiting Student*
2010–2014	University of Tübingen *B.Sc. in Economics*