Essays in Macroeconomics

Inauguraldissertation

zur Erlangung des Grades eines Doktors
der Wirtschaftswissenschaften

durch

die Rechts- und Staatswissenschaftliche Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Matthias Gnewuch

aus Ahlen

2023

Dekan:	Prof. Dr. Jürgen von Hagen
Erstreferent:	Prof. Dr. Keith Kuester
Zweitreferent:	JProf. Donghai Zhang, Ph.D.
Tag der mündlichen Prüfung:	27. Februar 2023

Acknowledgements

This dissertation has greatly benefited from the support and advice of many individuals. First of all, I would like to thank Keith Kuester, my first supervisor, for his invaluable guidance in numerous situations. His insightful feedback and perspectives have not only improved each chapter of this dissertation but also enhanced my understanding of economics. I am also extremely grateful to Donghai Zhang, my second supervisor, for his forthright advice and countless productive discussions which have led to the second chapter of this dissertation. Moreover, I would like to thank Christian Bayer, my third supervisor, not only for excellent feedback but also for his inspiring attitude that always motivated me to work on this dissertation.

More generally, I am grateful to many current and former faculty members at the Institute of Macroeconomics and Econometrics (IME)—in particular Benjamin Born, Thomas Hintermaier, and Joachim Jungherr—for their advice and support on various occasions. I would like to thank Heide Baumung and the administrative staff of the IME for their help with all sorts of issues. I am also thankful for many fruitful discussions and enjoyable lunches with my colleagues at the IME—in particular Kai, Alina, Lucas, Gasper, and Ricardo. I gratefully acknowledge financial support from the Bonn Graduate School of Economics (BGSE), the Research Training Group 2281 “The Macroeconomics of Inequality”, and the IME.

I was fortunate to start my doctoral studies at the BGSE alongside a great group of fellow Ph.D. students. I am particularly grateful to Timo, Martin, Luca, Chui Yee, Ximeng, Tobias, and Marton for joint hiking trips, bike tours, lunches and dinners, vacations, and many other activities that have made this entire journey so much more enjoyable. I also want to thank Holger Gerhardt for organizing the weekly BGSE soccer games, which I greatly enjoyed. Special thanks go to Annika for always supporting me, especially during the tough times of working on this dissertation. Above all, I am beyond grateful to my family, especially my parents Vera and Jürgen. Without their unconditional support, encouragement, and confidence in me, I would never have been in the position to write this dissertation. Thank you!

Matthias Gnewuch
Bonn, December 2022

Contents

List of Figures

List of Tables

Introduction

A central goal of macroeconomic policy is the stabilization of business cycle fluctuations. Stabilization policies, including monetary and fiscal policy, are implemented based on the paradigm that stable economic conditions are preferable to recurrent boom-bust cycles. This thesis contributes to our understanding of the effects and transmission mechanisms of macroeconomic stabilization policies, and how these depend on the decisions and interaction of heterogeneous firms. The first chapter investigates sovereign debt-based quantitative easing in the euro area and how its effects differ across countries. The second chapter studies the investment channel of conventional monetary policy in light of lumpy firm-level investment behavior. The third chapter explores how the strategic interaction of firms shapes aggregate fluctuations and the role of competition policy in macroeconomic stabilization.

Chapter 1, titled “Spillover Effects of Sovereign Debt-Based Quantitative Easing in the Euro Area”, studies central bank purchases of sovereign debt in the setting of the euro area. In normal times, central banks use interest rate policy to steer the economy. However, when interest rates are constrained by the Effective Lower Bound (ELB), central banks must resort to unconventional monetary policy measures to provide stimulus to the economy. In the aftermath of the Great Recession as well as during the COVID-19 crisis, large-scale asset purchases, also referred to as “Quantitative Easing”, were heavily used.

Estimating the causal effects of large-scale asset purchases is an intricate task because they are employed in view of economic conditions and jointly with other policy measures. In this chapter, I develop a novel strategy to identify the effects of central bank purchases of sovereign debt which addresses these challenges. The idea is to measure sovereign yield changes around official central bank communication that are unrelated to movements in risk-free interest rates or risk premiums. These yield changes reflect the anticipation of shifts in the effective supply of government debt available to the public, caused by news about central bank asset purchases. I document that asset purchase news about government debt not only affect sovereign bond markets but also have substantial spillover effects on corporate bond and stock markets, within and beyond the euro area. Most interestingly, spillovers are unequal across euro-area countries, as stock prices rise most in low-risk countries with very large firms. In contrast, sovereign yields fall rather homogeneously.

Chapter 2, titled “Monetary Policy, Firm Heterogeneity, and the Distribution of Investment Rates”, is joint work with Donghai Zhang and studies the investment channel of conventional monetary policy while considering lumpy firm-level investment behavior and firm life-cycle dynamics. According to the investment channel, expansionary monetary policy stimulates the economy by boosting aggregate investment. At the same time, it is well-known that the investment behavior of individual firms is lumpy—instead of regularly investing small amounts, they make big but infrequent investments. This raises the question of whether lumpy investment matters for the transmission of monetary policy.

To address this question, the chapter combines empirical evidence from firm-level micro data with a quantitative heterogeneous-firm model featuring lumpy investment and firm life-cycle dynamics and presents two main findings. First, monetary policy reshapes the distribution of investment rates. Specifically, an interest rate cut leads to fewer small or zero investment rates and more large investment rates. This corroborates the relevance of the extensive margin—firms deciding whether to invest or not—for the investment channel of monetary policy. Second, young firms are more sensitive to monetary policy than old firms, in the model as in the data. This is noteworthy because the model does not feature a financial accelerator mechanism which is however typically used to explain such empirical evidence. Both findings suggest that monetary policy is more effective in stimulating the economy in an expansion than in a recession.

Chapter 3, titled “Market Power and Macroeconomic Fluctuations”, starts from the observation that crises do not affect all firms equally. A natural disaster, for example, disrupts the production only of firms that are located in a specific region. A financial crisis particularly hurts firms that rely on external financing to fund their operations. I collectively refer to supply disruptions that affect some firms more than others within many industries as asymmetric supply shocks.

In this chapter, I study the aggregate consequences of asymmetric supply shocks. Most importantly, I show that when firms interact strategically, the aggregate effects of asymmetric supply shocks depend on the intensity of competition among firms. The reason is the profit-maximizing behavior of firms. When an adverse shock, such as a natural disaster, disrupts the production of some firms, their unharmed competitors as a result face a higher demand for their goods. When these firms have high market power, they find it optimal to respond by primarily raising prices instead of expanding production. In contrast, when these firms have low market power, they primarily raise production, not prices, and thereby help to stabilize aggregate output. Therefore, a more competitive economy is more resilient to asymmetric supply shocks. This finding implies that economic policy which fosters competition among firms not only reduces market power and markups but also provides macroeconomic stabilization.

Chapter 1

Spillover Effects of Sovereign Debt-Based Quantitative Easing in the Euro Area*

1.1 Introduction

Since the Great Recession, central banks around the world have introduced large-scale asset purchase programs to provide stimulus to economies at or near the Effective Lower Bound (ELB). Initially, these programs were considered to be exceptional and temporary, hence the label *unconventional* monetary policy. Yet, central bank balance sheets have remained at elevated levels and again rose sharply in the wake of the COVID-19 crisis. It thus appears that asset purchase programs have become part of the toolkit of regular central bank policy. In light of this, it is of first-order importance to understand the effects and transmission mechanisms of asset purchases. This is an intricate task, because asset purchases are employed in view of economic conditions and oftentimes announced jointly with other policy measures. Thus, to understand the causal effects of asset purchases, one has to account for the endogeneity of asset purchases to economic conditions and for concurrent policy announcements.

* I would like to thank the editors and two anonymous referees for providing exceptionally helpful guidance. I would also like to thank Keith Kuester, Donghai Zhang, Christian Bayer, Benjamin Born, Martin Dohmen, Peter Karadi, Marek Jarociński, Roberto De Santis, Michele Lenza, Oreste Tristani, Evi Pappa, as well as seminar participants at the European Central Bank, the VfS Annual Conference 2020, the 7th Workshop on Empirical Monetary Economics, the 13th RGS Doctoral Conference in Economics, and the 3rd Bonn-Mannheim PhD Workshop for helpful comments and suggestions. I thank the DFG for supporting my research as part of the research training group RTG 2281 "The Macroeconomics of Inequality". Parts of this paper were written while the author was at the European Central Bank. The views expressed in this paper are those of the author. They do not necessarily coincide with those of the European Central Bank.

In this paper, I propose a novel strategy to identify the effects of central bank purchases of government debt, henceforth “asset purchases”, which accounts for these concerns. It is well-known that monetary policy decisions affect many financial variables, including government bond yields, risk-free interest rates and risk premiums. Movements in none of these variables, around policy communication events, are guaranteed to arise from news about asset purchases only. The fundamental idea is that long-term government bond yields, instead, include a component that can be argued to reflect only news about asset purchases. I measure this component, which I label the “scarcity premium”, as the government bond yield in excess of the risk-free interest rate and the risk premium. This component arises due to the particular structure of the market for long-term government debt. Namely, the demand for long-term government debt is rather inelastic because there exists a unique clientele for long-term, safe nominal assets (Krishnamurthy and Vissing-Jorgensen, 2011). At the same time, the supply of debt by governments is limited and relatively inelastic in the short to medium term, as well. In this market environment, central bank asset purchases reduce the effective supply of government debt available to the public and, thus, affect the scarcity premium. I, then, use movements in the scarcity premium around ECB policy communication events to measure the effects of news about asset purchases. I find that asset purchases reduce euro-area sovereign bond yields rather homogeneously. Thus, I do not find that sovereign yield spreads of other countries vis-à-vis Germany fall significantly. At the same time, asset purchases increase stock prices; the effect being largest in Germany, France, and the Netherlands, i.e. countries with relatively few concerns about sovereign solvency. These two pieces of evidence may raise doubts about the prevalent view that asset purchases mostly benefit highly indebted countries.[1]

More in detail, the identification strategy developed in this paper builds on high-frequency identification. By looking at changes in monetary policy expectations in a narrow window around official central bank communication, we can identify the unexpected component of policy decisions. This deals with the endogeneity of policy decisions to economic conditions. To measure the change in expectations about asset purchases, changes in long-term interest rates, such as government bond yields, are commonly used. However, because multiple policy decisions are typically announced jointly, changes in long-term interest rates reflect not only news about asset purchases, but also news about other monetary policy measures and information about the state of the economy. The latter, Jarociński and Karadi (2020) show how to distinguish from interest rate movements due to monetary policy. Altavilla et al. (2019), on whose work I build and whose database I use, highlight that it is important to

1. For example, *The Economist* wrote “A fear often heard in the northern countries of the currency bloc [...] is that QE, by lowering the financing costs of indebted southern governments, allows them to avoid painful reforms. It is true that loose money has benefited highly indebted countries the most.” (The Economist, October 10, 2019, “What to make of the strife at the ECB”)

separate the different monetary policy measures which move interest rates. Thus, ideally, we would want to distinguish news about asset purchases from both, news about other policy measures and information about the state of the economy. For this purpose, I decompose long-term government bond yields of several euro-area countries into the risk-free interest rate, a risk premium, and the scarcity premium. Changes in the risk-free interest rate or the risk premium can reflect various monetary policy measures as well as the revelation of information. In contrast, changes in the scarcity premium can be argued to reflect only news about asset purchases. The supply by governments being inelastic, asset purchases reduce the effective supply of government debt available to the public. Demand also being inelastic, purchases reduce sovereign yields over and above any effect on the risk-free interest rate or the risk premium: the scarcity premium moves. At the same time, other monetary policy measures and information about the state of the economy do not significantly affect supply and demand in the market for government debt: the scarcity premium does not move. In a final step, I aggregate changes in the scarcity premiums of the four largest euro-area countries to create a single series of asset purchase news.

With this identified series at hand, I estimate the effects of asset purchases on financial markets in event-study regressions. Central bank purchases of government debt reduce not only the yields of euro-area government bonds, but also the yields of corporate bonds and non-euro area government and corporate bonds. At the same time, stock prices rise in the euro area and in other advanced economies. In addition, asset purchases reduce expected risk-free interest rates in line with a signaling effect and strongly depreciate the euro against all major currencies. I further study whether the financial effects of asset purchases differ across euro-area countries. First, I find that asset purchases reduce sovereign yields rather homogeneously. Even though point estimates are slightly larger for Spain and Italy, sovereign yield spreads vis-à-vis Germany do not fall significantly. This finding stands in contrast to some of the related literature. For example, Altavilla, Carboni, and Motto (2021) find sovereign spreads to fall on days with news about the Public Sector Purchase Programme (PSPP) between September 2014 and March 2015. A potential explanation for these contrasting findings are the different samples. Compared to previous studies, my analysis uses a later and much longer sample (October 2014 - January 2020) during which sovereign risk premiums were relatively low on average. Therefore, there was little scope for asset purchases to have heterogeneous effects by reducing risk premiums in countries with concerns about sovereign solvency. Second, I find that asset purchases have heterogeneous effects on stock prices. National stock indices increase the most in Germany (DAX), France (CAC 40), and the Netherlands (AEX). The common characteristic is that these stock indices include a number of very large firms. The evidence thus suggests that asset purchases may benefit large firms more than small firms. This heterogeneous effect on national stock indices alongside the homogeneous effect on sovereign bond yields may raise doubts about the prevalent view that asset purchases mostly benefit highly indebted countries.

Institutional Background. The analysis focuses on asset purchases under the ECB's PSPP. Choosing the euro area as a setting strengthens the identification strategy that I propose, for two reasons. First, the identification strategy builds on bond scarcity, which allegedly was more severe in Europe than elsewhere, as Coeuré (2018) explains. Second, I can exploit that the same monetary policy applies to a number of countries.

I focus on the PSPP because it is by far the largest of the ECB's asset purchase programs put in place before the COVID-19 crisis. The left panel of Figure 1.1 illustrates quarterly holdings under the four asset purchase programs, which comprise the ECB's Asset Purchase Programme (APP). By December 2019, the ECB held around €2100bn worth of euro-area government and agency debt under the PSPP, which amounts to roughly 20% of euro-area annual Gross Domestic Product (GDP). In comparison, the other programs were relatively small. The ECB held assets worth around €264bn under the Covered Bond Purchase Programme (CBPP), €184bn under the Corporate Sector Purchase Programme (CSPP), and €28bn under the Asset-Backed Securities Purchase Programme (ABSPP).

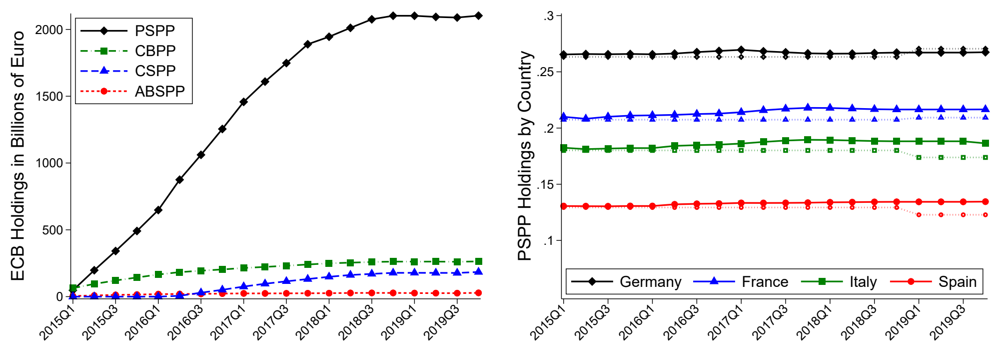


Figure 1.1. The European Central Bank's Asset Purchase Programme

Notes: The left panel shows ECB holdings under the Public Sector Purchase Programme (PSPP), the Covered Bond Purchase Programme (CBPP), the Corporate Sector Purchase Programme (CSPP), and the Asset-Backed Securities Purchase Programme (ABSPP) by quarter in billions of euro. In the right panel, solid lines depict ECB holdings under the Public Sector Purchase Programme (PSPP) by country as a share of total PSPP holdings (excluding supranationals), by quarter. Dashed lines depict the share according to the ECB's capital key. Source: ECB.

Throughout, the distribution of PSPP purchases across eligible countries was guided by the ECB's capital key. Therefore, the distribution was known in advance. The right panel of Figure 1.1 shows that country shares of holdings under the PSPP fluctuated little over time and that these shares closely align with the respective shares according to the ECB's capital key.[2] In contrast, both the duration of the

2. A new capital key entered into force on January 1, 2019, causing small changes in the prescribed country shares. To not disrupt market conditions, the ECB decided to adjust its portfolio al-

program and monthly purchase amounts were adjusted several times.[3] Therefore, the future amount of ECB purchases was uncertain. Hence, any ECB communication possibly changed expectations about the future amount of ECB purchases, but not about its distribution across countries.

Related Literature. This paper relates to three strands of the literature. First and foremost, it relates to the literature identifying monetary policy shocks from high-frequency monetary surprises. A seminal contribution is Gürkaynak, Sack, and Swanson (2005a), who show that more than one factor is required to explain U.S. high-frequency monetary surprises. Brand, Buncic, and Turunen (2010) apply their methodology to the euro area. A number of recent papers, including Nakamura and Steinsson (2018), Zhang (2019), Kerssenfischer (2019), Cieslak and Schrimpf (2019), and Jarociński and Karadi (2020), emphasize that non-monetary news, such as the dissemination of central bank information, are an important aspect of central bank communication. In a similar spirit, Kroencke, Schmeling, and Schrimpf (2021) identify changes in how investors evaluate and price risks during FOMC meetings. Andrade and Ferroni (2021) identify Delphic and Odyssean forward guidance shocks in the euro area. Swanson (2021) separates conventional monetary policy, forward guidance and asset purchase shocks for the U.S. by means of a factor rotation. Lewis (2019) identifies asset purchase shocks for the U.S. alongside other monetary policy shocks based on intraday time-varying volatility.

In the literature focusing on euro-area monetary policy shocks, two highly influential and closely related papers are Jarociński and Karadi (2020) and Altavilla et al. (2019). Jarociński and Karadi (2020) decompose interest rate surprises into policy and information shocks and show that these two shocks have very different macroeconomic effects. However, they do not further decompose policy shocks into shocks due to particular policy measures. Altavilla et al. (2019) decompose interest rate surprises into target, timing, forward guidance, and QE shocks using the factor rotation methodology of Swanson (2021) and the split of the ECB communication into a press release and a press conference. They document differences in the financial effects of the different policy measures and make another significant contribution by publishing and maintaining the Euro Area Monetary Policy Event-Study Database (EA-MPD), which I utilize. A limitation is that their methodology does not separate policy from information shocks. The contribution of this paper is to fill the gap between these two papers and provide a measure of news about asset purchases, which is not subject to either of the limitations mentioned above. In contrast to the policy shock of Jarociński and Karadi (2020), the measure of news about asset purchases that I develop is separated from news about other policy measures.

location very gradually. Therefore, an increase in the total amount of purchases would still imply an increase in the amount purchased from each country.

3. A detailed timeline of major events can be found in Hammermann et al. (2019).

In contrast to the QE factor of Altavilla et al. (2019), the measure does not correlate with information shocks.

Second, I add to the literature on the heterogeneous effects of asset purchases across regions. Wieladek and Pascual (2016), Burriel and Galesi (2018), and Hachula, Piffer, and Rieth (2020) provide VAR-based assessments of heterogeneous real effects across countries of the ECB's unconventional monetary policies. Altavilla, Carboni, and Motto (2021), Altavilla et al. (2019), and De Santis (2020) examine the heterogeneous effects of the PSPP on euro-area sovereign bond yields. Georgiadis and Gräb (2016) and Bubeck, Habib, and Manganelli (2018) evaluate the financial spillover effects of the PSPP beyond the euro area. To the best of my knowledge, my paper is the first to document heterogeneous effects of the PSPP on euro-area national stock indices.

Third, I relate to the much broader literature studying the financial market impact of large-scale asset purchases. Krishnamurthy and Vissing-Jorgensen (2011) provide an early assessment of the channels of the Federal Reserve's QE programs. Droste, Gorodnichenko, and Ray (2021) draw conclusions about the impact of asset purchases using evidence from Treasury auctions. D'Amico and King (2013) distinguish between stock and flow effects of asset purchases.[4] Krishnamurthy, Nagel, and Vissing-Jorgensen (2018) consider the effects of the ECB's Securities Markets Programme (SMP) and Outright Monetary Transactions (OMT). Koijen et al. (2021), Bergant, Fidora, and Schmitz (2018) and Albertazzi, Becker, and Boucinha (2021) analyze portfolio flows before and during the PSPP period using quarterly transaction-level data.

The remainder of this paper is structured as follows. Section 2 explains the identification of asset purchase news, presents the resulting series and compares the identification strategy to existing methodologies. Section 3 estimates the effects of asset purchases on financial markets and interprets the findings. Section 4 performs a number of robustness checks and Section 5 concludes.

1.2 Identification Strategy

A large literature strives to identify exogenous variation, so called shocks, in monetary policy in order to study its causal effects and mechanisms. This is an intricate task for several reasons. First, monetary policy decisions are taken in view of current and future economic conditions, making them endogenous to the state of the economy. Second, economic agents form expectations about future monetary policy

4. By construction, asset purchase news only measure stock effects of asset purchases, i.e. the effect of the ECB holding (or announcing to hold) a certain stock of assets. Potential flow effects of asset purchases, i.e. the effect of the actual purchases of these assets, are not measured. D'Amico and King (2013) find stock effects to be quantitatively more important. See Schlepper et al. (2018) for an analysis of the flow effects of purchases under the PSPP in the market for German sovereign bonds.

decisions, such that many decisions are anticipated. A popular strategy to deal with both issues and identify plausibly exogenous variation in monetary policy is high-frequency identification. The idea is to measure monetary policy shocks using the change in monetary policy expectations in a narrow window around official policy communication events.[5]

High-frequency identification requires a high-frequency measure which reflects monetary policy expectations. For expectations about unconventional monetary policy, long-term interest rates, such as government bond yields, are commonly used in the literature.[6] However, long-term interest rates reflect expectations of multiple policy measures. Since several policy decisions are typically announced jointly, the change in long-term interest rates in an event window reflects potentially a combination of a conventional monetary policy shock, a forward guidance shock, and news about asset purchases[7]. To deal with this multidimensionality of central bank communication, Altavilla et al. (2019) use a factor rotation methodology to identify target, timing, forward guidance, and QE shocks from high-frequency changes in interest rates of various maturities.

While the method used by Altavilla et al. (2019) has greatly enhanced our understanding of monetary policy, it also relies on the strong assumption that all interest rate movements in event windows reflect news about monetary policy. However, Jarociński and Karadi (2020) among others find that interest rate movements in event windows reflect not only policy decisions, but also the revelation of information about the state of the economy, so called central bank information shocks. Therefore, this paper proposes an alternative way to measure news about asset purchases, which does not rely on ruling out information shocks.

While the objective is similar, the approach in this paper differs from the two aforementioned papers. In particular, to measure monetary policy expectations, I do not use high-frequency changes in risk-free interest rates, precisely because these reflect multiple monetary policy and also information shocks. Instead, I construct a component of government bond yields, labeled the "scarcity premium", which around central bank communication arguably reflects only a single monetary policy shock, namely, news about asset purchases. The fundamental idea is that asset

5. Since not the policy decision itself, but the change in policy expectations is measured, the shock is unexpected. Moreover, since initial monetary policy expectations take into account the state of the economy and there are by assumption no other news during during the narrow window, the shock is not endogenous to economic conditions. See Ramey (2016) for a more detailed discussion.

6. For example, Andrade et al. (2016) use the German 5-year sovereign yield, Hachula, Piffer, and Rieth (2020) use a number of euro-area 2-year sovereign yields excluding Germany, and Gambetti and Musso (2020) use a GDP-weighted euro-area 10-year yield. Altavilla et al. (2019) and Andrade and Ferroni (2021) use interest rate swap rates of various maturities.

7. Throughout, I refer to this shock as news about asset purchases, or, *asset purchase news*. Other papers use the terms *asset purchase shocks*, *QE shocks*, or *LSAP shocks* to refer to the same or a very similarly defined shock.

purchases, as opposed to other monetary policy or information shocks, reduce the effective supply of government debt and thereby affect the scarcity premium. I explain the measurement of the scarcity premium and the construction of asset purchase news in detail in Section 1.2.1. I confirm that the resulting series reflects news about asset purchases in Section 1.2.2, and demonstrate that it is indeed unrelated to other monetary policy and information shocks in Section 1.2.3. In Section 1.2.4, I compare the identification strategy to the approach of Altavilla et al. (2019).

1.2.1 Measuring Asset Purchase News

The basic idea of my identification strategy is to not use high-frequency changes in the entire government bond yield, but to decompose the yield and isolate a component, which around ECB communication reflects only news about asset purchases. I draw on Krishnamurthy, Nagel, and Vissing-Jorgensen (2018)[8] and consider the following decomposition of the nominal yield of a sovereign bond of country c with remaining term to maturity T at time t:

$$\begin{aligned} yield_t^{c,T} &= \underbrace{i_t^T}_{\text{Overnight Index Swap (OIS) rate with maturity } T} \\ &+ \underbrace{CountryRiskPremium_t^{c,T}}_{\text{Credit Default Swap (CDS) rate of country } c \text{ with maturity } T} \\ &+ \underbrace{ScarcityPremium_t^{c,T}}_{\text{Component of interest}} \end{aligned}$$

The first component, i^T, is the risk-free nominal interest rate associated with the remaining maturity. It is straightforward to measure this component from maturity-matched interest rate swaps using the Euro OverNight Index Average (EONIA) as the underlying floating rate.[9] The second component, the country-specific risk premium,

8. Krishnamurthy, Nagel, and Vissing-Jorgensen (2018) use a similar decomposition to understand sovereign yield changes in response to specific ECB policy communication events, associated with news about the Securities Markets Programme (SMP), the Outright Monetary Transactions (OMT), and the Long-Term Refinancing Operations (LTROs). I draw on their decomposition, but reverse the logic. Instead of assuming the type of news and assessing the effect on yield components, I assume that one particular yield component reflects only one type of news and thereby back out a quantitative measure of this type of news.

9. These interest rate swap contracts are used to hedge interest rate exposure. The buyer periodically pays a fixed rate (the swap rate) to the seller and receives the current (floating) rate in return, thereby trading interest rate risk. Without uncertainty, the swap rate would equal the average expected interest rate. With uncertainty, the swap rate will also include a risk premium which compensates the seller for bearing the interest rate risk. Therefore, the overnight index swap (OIS) rate using the Euro OverNight Index Average (EONIA) as the underlying floating rate is a convenient measure of the Euro risk-free interest rate along with the interest rate risk premium. Lloyd (2021) explains that counterparty risk in OIS contracts is minor. EONIA OIS contracts are fairly liquid and available for a wide range of maturities ranging from two weeks to thirty years.

compensates the investor for bearing the risk of and loss in case of sovereign default and currency redenomination. For the time period of interest, it is straightforward to measure this premium from maturity-matched credit default swap (CDS) rates traded under the 2014 ISDA Credit Derivatives Protocol.[10] CDS contracts based on the 2014 ISDA Credit Derivatives Protocol, which are traded since September 22, 2014, insure against sovereign default and currency redenomination. CDS contracts under the previous 2003 ISDA Credit Derivatives Protocol do not insure against currency redenomination for G-7 countries.[11] Therefore, I restrict my sample to the period after the introduction of the 2014 CDS protocol to be able to account for redenomination risk.

The third component, which I label the scarcity premium[12], is measured residually. This premium should be negligible under no-arbitrage considerations (Duffie, 1999). However, in the data it is non-negligible and negative due to the particularities of sovereign bonds.[13] On the one hand, sovereign bonds are commonly used to collateralize transactions, serve as a safe storage facility, and count as high-quality liquid assets towards banks' *liquidity coverage ratio*.[14] Therefore, holding them provides utility to banks and financial institutions. This gives rise to a rather inelastic, or, in the words Krishnamurthy and Vissing-Jorgensen (2011), *unique* demand for sovereign bonds. On the other hand, the supply of sovereign bonds by euro-area governments is limited and rather inelastic due to constraints imposed by the European Union's fiscal rules. First and foremost, the Stability and Growth Pact limits government deficits and debt. The combination of this unique demand for and limited supply of sovereign debt gives rise to the scarcity premium.[15] Since this component is computed as a residual, the measured scarcity premium will also reflect any additional drivers of the sovereign yield, which are not captured by the other two components.[16]

10. Available maturities range from six months to thirty years. I use CDS contracts denominated in euro.

11. Three countries, for which I compute this decomposition, are G-7 countries, namely Germany, France, and Italy.

12. I refer to this component as a premium, because, while it *decreases* the yield, it *increases* the price of the bond. This nomenclature follows the definition of the liquidity premium in Nagel (2016).

13. Over the sample period used in this paper (10/2014 - 01/2020), the scarcity premium is on average negative for all four countries (Germany, France, Italy, Spain), for which the decomposition is done. It is largest (in absolute terms) for Germany with on average 40 basis points.

14. Under the Basel III regulatory framework, banks are required to have a liquidity coverage ratio (LCR) of 100% or higher. The LCR is defined as high-quality liquid assets divided by total net liquidity outflows over 30 days. Sovereign bonds count as high-quality liquid assets without haircut.

15. I choose this nomenclature to emphasize that the premium is a result of particularities on the demand and the supply side. Closely related concepts which solely emphasize particularities on the demand side are the *safety premium* (Krishnamurthy and Vissing-Jorgensen, 2011), the *convenience yield* (Krishnamurthy and Vissing-Jorgensen, 2012) or the *utility premium*.

16. As discussed in the following paragraph, this is not an issue as long as such additional drivers are not affected by central bank communication besides news about asset purchases.

I propose to use the change in this scarcity premium around ECB policy communication as a measure of asset purchase news. Thereby, I make two key assumptions, which are visualized in Figure 1.2. First, I assume that news about asset purchases affect the scarcity premium (Relevance Assumption). The supply of sovereign debt by governments and the demand by investors being rather inelastic, this must be the case, because asset purchases reduce the effective supply of sovereign bonds available to the public. Since actual purchases do not take place until well after their announcement, however, this requires that financial markets are sufficiently forward-looking. Second, I assume that other elements of central bank communication, including conventional monetary policy shocks, forward guidance shocks and central bank information shocks, do not affect the scarcity premium (Exogeneity Assumption). The rationale is that these shocks affect sovereign yields only via the risk-free interest rate and the country-specific risk premium. None of these shocks affects the scarcity premium, because they do not affect demand and supply for sovereign bonds in a relevant magnitude. I provide empirical support for this assumption in Section 1.2.3. Nevertheless, two concerns remain. First, the ECB could affect the scarcity premium directly by announcing major changes to its collateral framework. I confirm that no major changes were announced on Governing Council Meeting days in the sample.[17] Second, ECB communication could potentially induce safe haven flows by changing investors' perception of risk. Fontana and Scheicher (2016) discuss in the context of the euro crisis that a flight to safety may have shifted bond demand from peripheral to core euro-area countries. To address this concern, I extract the common component of changes in the scarcity premium of several countries to not capture shifts in bond demand between euro-area countries. I explain this final step in the following.

In principle, the decomposition explained above can be applied to sovereign bonds of any maturity and any euro-area country, whose bonds were bought under the PSPP. I focus on 10-year sovereign yields, because bond scarcity is a bigger issue among long-term bonds.[18] Moreover, these bonds did not trade below the Deposit Facility Rate (DFR) before December 2016, which would have made them ineligible for ECB purchases at that point in time. This is particularly relevant for Germany and France, whose short-term government bonds did periodically trade below the DFR. Furthermore, I focus on euro-area countries with a large and highly liquid market for government debt, which is necessary to compute yield changes in a narrow window

17. There are two potentially interfering regulatory changes. On December 8, 2016, the Eurosystem introduced cash collateral for PSPP securities lending facilities. On December 14, 2017, there were changes to collateral eligibility criteria for unsecured bank bonds. However, neither of these regulatory changes seems quantitatively important.

18. For example, over the sample period used in this paper (10/2014 - 01/2020), the average scarcity premium for German 5-year bonds (-32 basis points) is smaller in absolute terms than the average premium for 10-year bonds (-40 basis points).

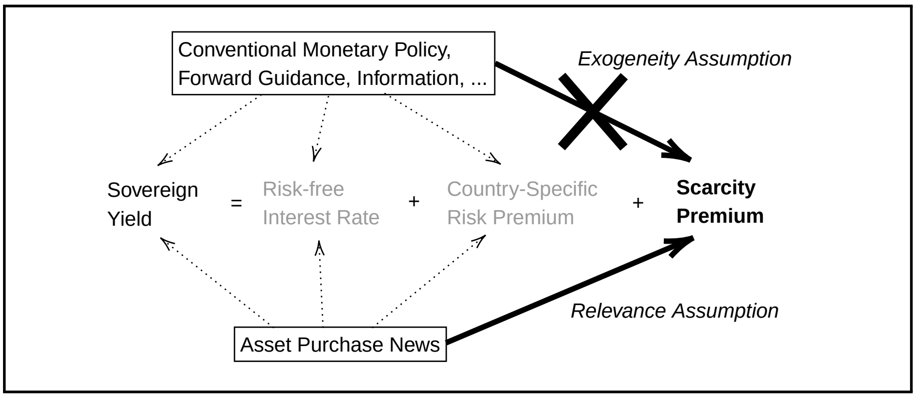


Figure 1.2. Visualization of Identifying Assumptions

around policy communication. I select Germany (DE), France (FR), Italy (IT), and Spain (ES).

I obtain changes in 10-year sovereign yields and 10-year OIS rates in a narrow window around ECB communication on Governing Council Meeting days from the Euro Area Monetary Policy Event Study Database (EA-MPD) made available by Altavilla et al. (2019). I use the full *monetary event window* to capture news about asset purchases from the press conference and the press release.[19] To the best of my knowledge, intraday data on CDS rates is not available. Therefore, I use daily changes in 10-year CDS rates from Thomson Reuters Eikon.[20]

In order to create a single measure of asset purchase news and minimize the influence of country-specific noise, I aggregate information from all four countries. Therefore, I extract the first principal component of the changes in the scarcity premiums around ECB policy communication on Governing Council Meeting days. Beforehand, the country-specific series are standardized to unit variance to avoid mechanically giving higher weights to more volatile series. This approach resembles Gürkaynak, Sack, and Swanson (2005a) and Nakamura and Steinsson (2018), who condense information from interest rates of various maturities using principal

19. The monetary event window (13:30 CET - 15:45 CET) brackets both, the press release published at 13:45 Central European Time (CET), and the press conference starting at 14:30 CET and lasting for around an hour. Before March 2016, the ECB communicated news about its unconventional policies in the press conference only. Since then, some information about unconventional policies is already included the press release. Thus, to capture all news about asset purchases, it is necessary to use the full monetary event window.

20. The lack of data on intraday CDS rate changes makes it necessary to trade off noise in the scarcity premium outside the monetary event window against changes in the CDS rate outside the window. Since in the relevant time period, CDS rates were relatively low and there was little concern about debt sustainability, I opt for this mixed-frequency approach. In a previous draft, I opted for daily data throughout and found broadly similar results. There are three missing observations in the series of German CDS rate changes which are set to 0.

component analysis. I discuss the aggregation of individual series in more detail in Appendix 1.D and show that the final series is not driven by any single country.

1.2.2 Series of Asset Purchase News

The principal component decomposition identifies the series of asset purchase news only up to sign and scale. I define the sign such that positive asset purchase news[21] reduce scarcity premiums and therefore constitute expansionary realizations. Moreover, I follow Altavilla et al. (2019) and normalize the scale such that asset purchase news reduce the 10-year OIS rate by one basis point. Figure 1.3 shows the resulting series of asset purchase news. The sample contains a total of 44 ECB Governing Council Meeting days between October 2014 and January 2020.

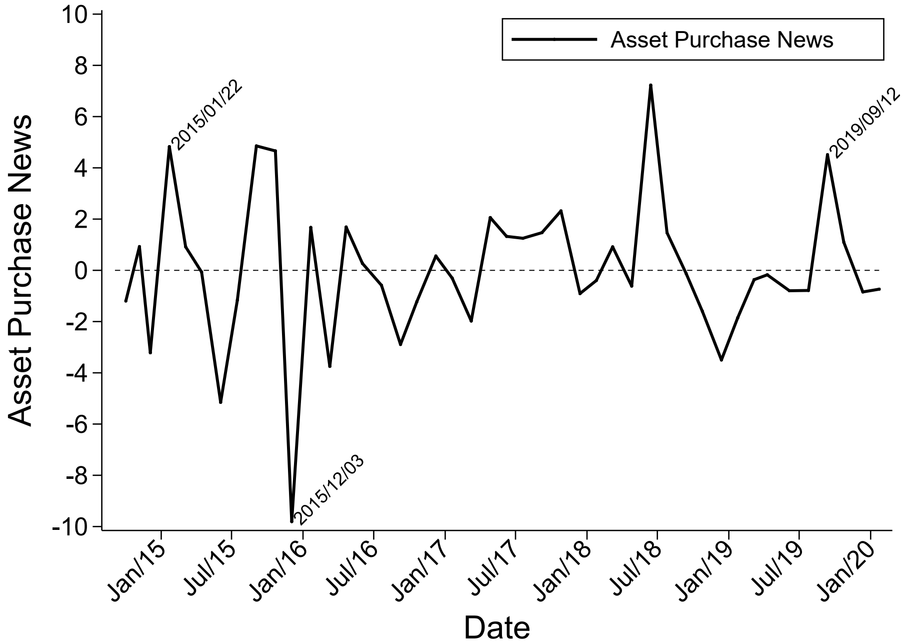


Figure 1.3. Series of Asset Purchase News

Notes: Observations refer to ECB Governing Council Meetings. Positive (negative) realizations denote expansionary (contractionary) asset purchase news. Sample: October 2014 - January 2020.

The realizations of asset purchase news align well with the interpretation in the financial press of the respective ECB communication. The first large expansionary realization occurs on January 22, 2015, the day on which the PSPP was officially announced. Although market participants were expecting the ECB to introduce a

21. In the nomenclature of the literature, asset purchase news could also be called asset purchase news *shocks*. I remain by *asset purchase news* for the sake of brevity. In addition, I frequently omit the supplement *expansionary* when referring to (positive) asset purchase news.

large-scale asset purchase program, they did not anticipate its size.[22] The largest contractionary realization is December 3, 2015. On this day, the ECB decided to extend its asset purchases for only 6 months, while markets had expected a longer extension or even an increase in the monthly amount of purchases.[23] This event was preceded by a large expansionary realization on October 22, 2015, when President Draghi surprisingly hinted at an expansion of asset purchases at the next Governing Council Meeting.[24] The last large realization is September 12, 2019, the day on which the ECB announced to restart net asset purchases. This decision was only partly expected and therefore amounted to expansionary news.[25]

1.2.3 The Exogeneity Assumption

Figure 1.3 and the discussion in Section 1.2.2 confirm that the series of asset purchase news captures key announcements regarding the PSPP, thereby supporting the relevance assumption. In this subsection, I provide empirical support for the exogeneity assumption which states that the scarcity premium is not systematically affected by other elements of central bank communication, including conventional monetary policy shocks, forward guidance shocks, and central bank information shocks.

To confirm that the series of asset purchase news does not correlate with information shocks, I compute central bank information and monetary policy shocks using a rotational sign restriction approach in the spirit of Jarociński and Karadi (2020) and as implemented in Jarociński (2021). The idea is that interest rate changes accompanied by stock price changes of the opposite sign reflect monetary policy shocks. Interest rate changes accompanied by stock price changes of the same sign reflect information shocks. With multiple underlying monetary policy shocks, the measured monetary policy shock will reflect a combination of these. However, the information shock is still the only underlying shock, which induces a positive co-movement of interest rates and stocks. Therefore, it can still be identified. As explained in Jarociński (2021), sign restrictions only provide set identification and there are three

22. The *Financial Times* wrote "Mario Draghi's bond-buying plan outstripped expectations". (Financial Times, January 22, 2015, "Mario Draghi's bond-buying plan outstrips expectations" by Claire Jones)

23. *The Guardian* wrote "Mario Draghi dashes expectations that the European Central Bank would pump more new money into the eurozone economy each month". (The Guardian, December 3, 2015, "ECB Day: markets tumble as Draghi disappoints investors - as it happened" by Graeme Wearden)

24. *The Guardian* wrote "Mario Draghi, the president of the European Central Bank, has stunned markets by signalling that he is prepared to cut interest rates and step up quantitative easing to stave off the risk of a renewed economic slump in the eurozone". (The Guardian, October 22, 2015, "Mario Draghi: ECB prepared to cut interest rates and expand QE" by Heather Stewart)

25. *Market Watch* wrote "Economists had been less certain whether the ECB would also move to relaunch its quantitative easing program at its September meeting, but policy makers did so." (Market Watch, September 12, 2019, 'ECB cuts key rate, relaunches QE to shore up eurozone economy")

Table 1.1. Correlation of Asset Purchase News with Policy and Information Shocks

	Information Shock			Policy Shock		
	Baseline	Median	Poor Man's	Baseline	Median	Poor Man's
Asset Purchase News	-0.07 (0.6546)	-0.11 (0.4863)	-0.07 (0.6533)	0.74*** (0.0000)	0.74*** (0.0000)	0.75*** (0.0000)

Notes: Table reports correlation coefficients, while *p*-values are in parentheses. Signs of all shocks are normalized to decrease interest rates. Policy and information shocks are identified from 1Y, 2Y, 5Y, and 10Y OIS rate surprises in the monetary event window. "Baseline" shocks use the rotation which matches the variance shares of poor man's shocks as in Jarociński (2021). "Median" shocks use the median rotation among all admissible rotations. "Poor Man's" shocks are computed as in Jarociński and Karadi (2020). The sample is 10/2014 - 01/2020 (N=44). See Appendix 1.A for more details on the construction of the shocks.

options to uniquely identify the series of information shocks. I report results using all three series. More details of the construction of information and policy shocks are explained in Appendix 1.A. Table 1.1 shows that the correlation of asset purchase news and information shocks is very small and insignificant, regardless of the series of information shocks. This is important because I investigate the effect of asset purchases on stock prices in the following section. The presence of information shocks would bias the estimates towards zero. Moreover, the correlation of asset purchase news with the policy shock is large and significant. This is as expected and reflects that asset purchases were the most important policy tool during this time period.

To support the assumption that the scarcity premium is not affected by other monetary policy shocks, I turn to the period before the PSPP. Before 2014, there were asset purchase news for Italy and Spain, since their government bonds were purchased under the Securities Markets Programme (2010-2012), but there weren't any asset purchase news for Germany and France. Thus, the series of scarcity premium changes for Germany and France should only pick up noise and under the exogeneity assumption be uncorrelated with monetary policy shocks.[26] As an out-of-sample test, I therefore compute the correlation of changes in these two scarcity premiums with a number of identified monetary policy shocks between July 2011 to December 2013.[27] I use the policy shocks identified in Jarociński and Karadi (2020) and Kerssenfischer (2019) using the co-movement of interest rates and stock prices. Moreover, I use the target and Delphic and Odyssean forward guidance shocks identified in Andrade and Ferroni (2021) using the co-movement of 1-year OIS and

26. To compute the scarcity premiums before October 2014, I need to use CDS rates from contracts traded under the 2003 protocol, which do not insure against currency redenomination. Therefore, the scarcity premiums potentially include a redenomination risk premium, which could respond to monetary policy actions. Bayer, Kim, and Kriwoluzky (2018) discuss redenomination risk in the euro area between January 2010 and October 2014 and argue that it was sizable even for Germany and France.

27. Before July 2011, high-frequency data on OIS rates is unavailable, as discussed in Altavilla et al. (2019). Throughout 2014, there was already discussion about large-scale asset purchases in the euro area, meaning that there may have been asset purchase news.

5-year inflation-linked swap (ILS) rates.[28] Finally, I use the press release surprise in short-term OIS rates, which is a common measure of conventional monetary policy surprises. Table 1.2 displays the correlation coefficients of these identified monetary shocks with the German and French scarcity premium changes. All correlations are small and not significantly different from zero, thereby lending support to the exogeneity assumption.

Table 1.2. Correlation of Monetary Policy Shocks with DE & FR Scarcity Premium Changes

	Δ Scarcity Premium Germany	**Δ Scarcity Premium France**
Jarocinski & Karadi (2020)		
→ Policy Shocks	0.04 *(.8362)*	0.07 *(.7293)*
Kerssenfischer (2019)		
→ Policy Shocks	-0.03 *(.8747)*	-0.01 *(.9768)*
Andrade & Ferroni (2019)		
→ Delphic Forward Guidance Shocks	0.00 *(.9996)*	0.07 *(.6973)*
→ Odyssean Forward Guidance Shocks	-0.01 *(.9614)*	0.14 *(.4500)*
→ Target Shocks	0.25 *(.1763)*	0.27 *(.1463)*
Press Release Surprises		
→ 1-M OIS Rate	0.04 *(.8235)*	0.10 *(.6015)*
→ 3-M OIS Rate	0.10 *(.6050)*	0.14 *(.4505)*
→ 1-Y OIS Rate	0.00 *(.9978)*	0.08 *(.6689)*

Notes: Sample includes all ECB GCMs from July 2011 to December 2013 (N=30). *p*-values are reported in parentheses.

1.2.4 A Comparison to Altavilla et al. (2019)

The paper most closely related to mine is Altavilla et al. (2019), who also develop a measure of news about asset purchases contained in ECB communication. They exploit that the ECB communication on Governing Council Meeting days is split into two parts. At 13:45 Central European Time (CET), a press release is published. At 14:30 CET, a press conference begins, which lasts around one hour. For identifying the effects of unconventional monetary policy measures, including asset purchases, they use a narrow window around the press conference only. Altavilla et al. (2019) extract factors from changes in interest rates (1-month to 10-years) in this narrow window and rotate them using the methodology of Swanson (2021). Thus, they extract three factors and rotate them to make the factors interpretable as a timing,

28. I thank Filippo Ferroni for sharing the series of Delphic and Odyssean forward guidance shocks.

a forward guidance (FG), and a QE factor.[29] The QE factor provides a measure of news about asset purchases, similar in interpretation to the measure developed in this paper. Figure 1.G.4 provides a visual comparison of the two measures.

There are several appealing features of the methodology employed by Altavilla et al. (2019). Most importantly, their approach aims to identify *all* shocks which move interest rates around ECB communication events. Moreover, they look at a very long sample, starting in January 2002. This paper takes a much narrower approach and aims to identify merely a single shock and for a shorter period of time, starting in October 2014. Moreover, publishing and maintaining the Euro Area Monetary Policy Event-Study Database (EA-MPD) by Altavilla et al. (2019) constitutes a significant contribution in itself.

There are two potential drawbacks of the state-of-the-art methodology, as applied by Altavilla et al. (2019) among others. This paper improves on these. First and foremost, the identification strategy does not consider that interest rate movements around central bank communication reflect not only policy decisions, but also the revelation of information about the state of the economy. Jarociński and Karadi (2020) show that it is essential to distinguish between these central bank information shocks and true monetary policy shocks. Indeed, also Altavilla et al. (2019) explain in their Section 5 that the presence of information shocks may explain the surprisingly small and insignificant estimates of the effect of some policy factors on stock markets.[30] I formalize this point in Table 1.3, which shows that the QE factor identified in Altavilla et al. (2019) appears to correlate significantly with the series of information shocks identified in the spirit of Jarociński and Karadi (2020). This table uses exactly the same series of information and policy shocks, that were used in table 1.1.[31] The FG factor identified in Altavilla et al. (2019) also correlates significantly with the series of information shocks. Moreover, both factors correlate with the policy shocks, as expected. Thus, table 1.3 shows that the QE and FG factors not only reflect monetary policy, but also the revelation of information about the state of the economy. As I have shown in table 1.1, my proposed measure does not correlate with information shocks. The identification strategy proposed in this paper circumvents the issue that movements in risk-free interest rates also reflect the revelation of information by not directly using movements in risk-free interest rates to identify asset purchase news.

A seemingly straightforward approach to get rid of information shocks in the factor rotation methodology works as follows. One could simply orthogonalize the

29. The identifying assumptions are (i) that forward guidance and QE do not load onto the 1-month OIS rate surprise, and (ii) that the QE factor has a minimal variance before August 2008. By definition, the three factors are required to be orthogonal.

30. “Therefore, the presence of these two types of policy [monetary policy; information revelation] can make the response of the stock market, on average, insignificant and can produce the results reported in Table 7.” (Altavilla et al. (2019), p.174)

31. The construction of these shocks is explained in Appendix 1.A.

Table 1.3. Correlation of QE and FG Factors with Policy and Information Shocks

	Information Shock			Policy Shock		
	Baseline	Median	Poor Man's	Baseline	Median	Poor Man's
QE Factor	0.39***	0.36***	0.37***	0.58***	0.60***	0.59***
	(0.0038)	(0.0078)	(0.0057)	(0.0000)	(0.0000)	(0.0000)
FG Factor	0.28**	0.27*	0.33**	0.23*	0.24*	0.2
	(0.0447)	(0.0550)	(0.0167)	(0.0987)	(0.0799)	(0.1436)

Notes: Table reports correlation coefficients, while p-values are in parentheses. Signs of all shocks are normalized to decrease interest rates. Policy and information shocks are identified from 1Y, 2Y, 5Y, and 10Y OIS rate surprises in the monetary event window. "Baseline" shocks use the rotation which matches the variance shares of poor man's shocks as in Jarociński (2021). "Median" shocks use the median rotation among all admissible rotations. "Poor Man's" shocks are computed as in Jarociński and Karadi (2020). The sample is 01/2014 - 01/2020 (N=53). See Appendix 1.A for more details on the construction of the shocks.

three factors with respect to the identified information shock. This would work under the assumption that the decomposition of interest rate surprises into three policy factors in the first place is not affected by the presence of information shocks.[32] This, however, may be problematic in light of the evidence above that the information shock correlates with two of the estimated factors (QE, FG). The information shock therefore induces a correlation among the factors. Yet, factors must be uncorrelated by definition. Therefore, the decomposition of interest rate surprises into three factors in the presence of information shocks must be different from the decomposition without concurrent information shocks.

Another approach to get rid of information shocks in the factor rotation methodology would be to combine the approaches of Jarociński and Karadi (2020) and Altavilla et al. (2019). One could simply augment the data matrix used in Altavilla et al. (2019) with the stock market surprise and rotate the factors using a combination of the restrictions of both papers. Using this augmented data matrix, one would expect to find 3 factors (timing, forward guidance, information) in the pre-QE sample (01/2002 - 01/2014, as in Altavilla et al. (2019)) and one additional factor (QE) in the full sample (01/2002 - 01/2020). Surprisingly, using the Cragg and Donald (1997) test, I find three factors in both samples.[33] Finding the same amount of factors in both samples is at odds with the prior that a new factor (QE) emerged after 2014. It also impedes using the identifying assumption, to separate QE and FG, that one factor (QE) was only active in the later part of the sample and should therefore have a minimal variance beforehand. For sure there are ways to identify a QE factor

32. That is, each factor picks up only the policy measure it is supposed to pick up plus the information shock. E.g. $QEFactor_t = \epsilon_t^{QE} + \gamma_1 \epsilon_t^{Info}$, $FGFactor_t = \epsilon_t^{FG} + \gamma_2 \epsilon_t^{Info}$, etc.

33. The Cragg and Donald (1997) test is a bottom-up test, which tests the hypothesis of k factors against the alternative hypothesis of there being more than k factors. In the pre-QE sample, the hypothesis of two factors can be rejected ($p = 0.0420$), while the hypothesis of three factors cannot ($p = 0.4009$). Similarly, in the full sample, the hypothesis of two factors can be rejected ($p = 0.0054$), while the hypothesis of three factors cannot be rejected ($p = 0.4917$).

which is unrelated to the revelation of information in a factor rotation framework. However, this exploration shows that doing so is not straightforward and might require different variables, a different sample, and different identifying assumptions. While this approach certainly constitutes an interesting avenue for future research, I pursue an entirely different strategy in this paper. Instead of choosing surprises of variables which are affected by many shocks and seeking to disentangle them, I construct a variable, the scarcity premium, which arguably reflects only a single shock.

There is another concern with the way the existing literature proceeds. Namely, it identifies the QE factor from the press conference window only. This is fine before March 2016, because then, the press release only contained information about the ECB's interest rate decisions. Thus, for measuring news about asset purchases, it was sufficient to look at the press conference window. Since March 2016, however, the press release also includes key decisions regarding the ECB's asset purchase programs. Therefore, to measure all news about asset purchases, one needs to look at both, the press release window and the press conference window, as this paper does when measuring asset purchase news. A limitation, however, of the identification strategy developed in this paper is that it cannot exploit the split of the communication into two windows. This is because changes in CDS rates are only available at a daily frequency. Therefore, it is not possible to investigate whether measured asset purchase news stem from information released with the press release or during the press conference.

1.3 Results

With the series of asset purchase news at hand, I explore the effects of central bank purchases of government debt on financial markets. Particular attention is paid to heterogeneous effects across euro-area countries, as well as spillover effects beyond the market for euro-area sovereign debt.

To estimate the effects of asset purchase news, I use the following regression specification commonly used in the event-study literature:

$$y_t = \alpha + \beta s_t + \epsilon_t \tag{1.1}$$

where y_t is the one-day change in some financial variable of interest, s_t is the series of asset purchase news, and ϵ_t is the error term. The parameter of interest is β, which captures the effect of asset purchase news on the dependent variable. Throughout, I use standard errors robust to heteroskedasticity and autocorrelation.[34] Any noise

34. Since the regressor, i.e. the asset purchase news, is generated, standard errors should reflect the additional uncertainty arising from its construction. However, confidence intervals constructed using a wild bootstrap in the spirit of Swanson (2021) are barely distinguishable from those constructed

remaining in the measure of asset purchase news introduces an attenuation bias, distorting the estimate $\hat{\beta}$ towards zero. In that case, $\hat{\beta}$ provides a lower bound on the true effect of asset purchases. I focus on the impact effects of asset purchase news and document their persistence over the following days in Appendix 1.B.

1.3.1 Risk-Free Interest Rates

To understand how asset purchases affect financial markets, it is crucial to know how these affect risk-free interest rates. Usually, and also in the case of the PSPP, large-scale asset purchase programs are employed to provide additional monetary stimulus to an economy at or near the ELB. Therefore, short-term risk-free interest rates cannot be reduced much more. Nevertheless, asset purchases can still reduce long-term risk-free interest rates by reducing expected future short-term interest rates. This is referred to as the signaling channel of asset purchases.[35] On the one hand, asset purchases can be seen as a commitment to keep rates low, because the central bank would make losses on its purchased assets if it raised interest rates. On the other hand, asset purchases may signal that the central bank is willing to maintain an accommodative policy stance in the future.

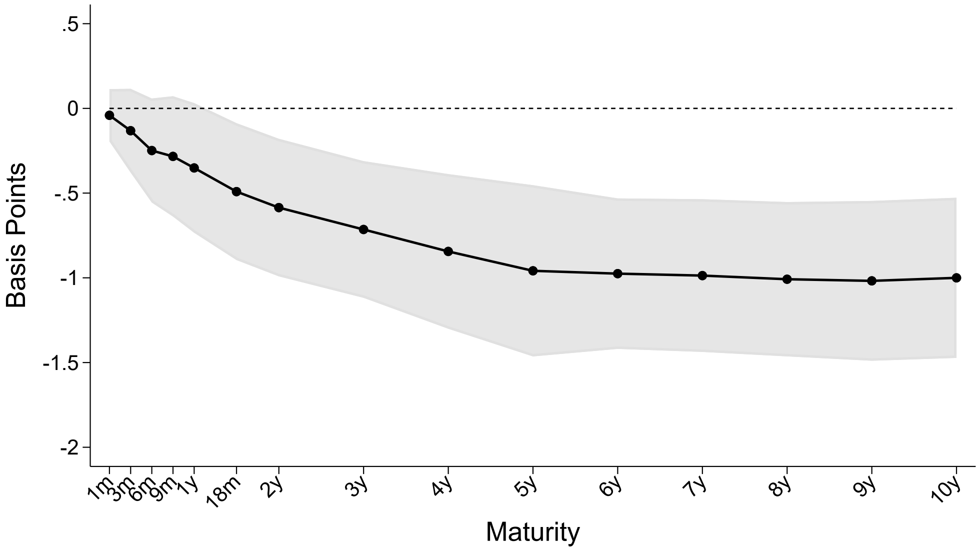
Figure 1.4 illustrates the impact effect of asset purchase news on risk-free nominal interest rates of various maturities (upper panel) and implied forward rates at various horizons (lower panel). As in Section 1.2, I use interest rate swaps to measure risk-free interest rates and implied forward rates.[36] Asset purchases do significantly reduce long-term interest rates as can be seen from the upper panel. Recall that the magnitude of the effect on the 10-year OIS rate is normalized to one basis point as in Altavilla et al. (2019). The lower panel provides strong support for the signaling channel, as asset purchase news reduce implied forward rates, while the effect is largest and most significant between 1- and 5-years ahead. The peak effect at the 5-year horizon is later than observed for conventional monetary policy and forward guidance shocks.[37]

with asymptotic standard errors robust to heteroskedasticity and autocorrelation. Therefore, I remain with asymptotic standard errors. Gürkaynak, Sack, and Swanson (2005a) similarly observe that bootstrapping standard errors leads to almost identical results.

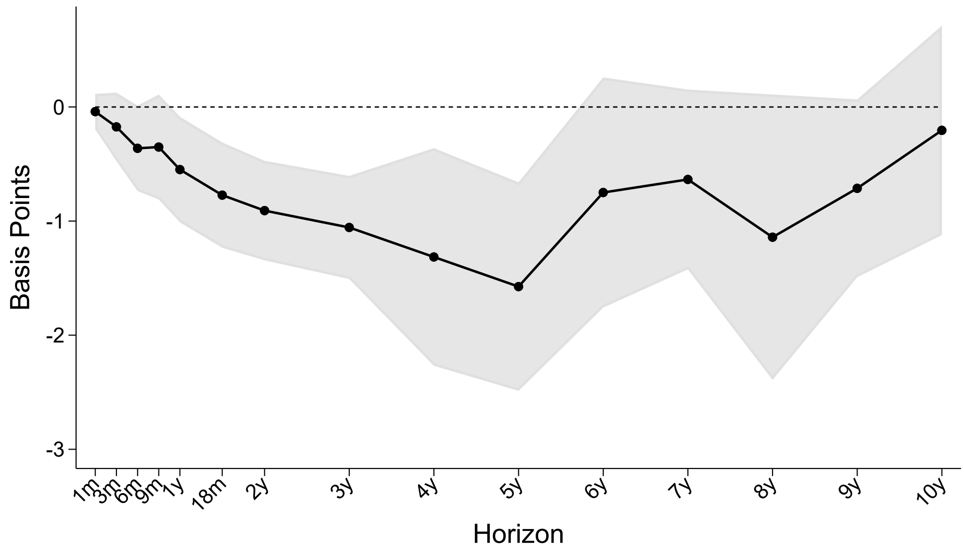
35. Bauer and Rudebusch (2014) discuss the importance of the signaling channel for the Fed's QE programs and argue that it contributed 40-50% to the decline of long-term Treasury yields.

36. As explained in footnote 9, interest rate swap rates include an interest rate risk premium. Therefore, the results in Figure 1.4 may not only reflect falling interest rate expectations, but also reduced interest rate risk.

37. For example, Brand, Buncic, and Turunen (2010) estimate a downward-sloping maturity response pattern to ECB policy decisions. Altavilla et al. (2019) estimate a very similar downward-sloping response pattern to their target factor. Moreover, the effects of forward guidance in Andrade and Ferroni (2021) and Altavilla et al. (2019) peak earlier than 5-years ahead.



(a) OIS Rates



(b) OIS-implied Forward Rates

Figure 1.4. Response of Risk-Free Interest Rates to Asset Purchase News

Notes: The dots represent the estimated $\hat{\beta}$ from separate regressions: $y_t = \alpha + \beta s_t + \varepsilon_t$, where y_t is the daily change, measured in basis points. Shocks are scaled to reduce the 10-year OIS rate by one basis point. Shaded areas depict the 90% confidence interval using robust standard errors.

1.3.2 Bond Markets

It is largely undisputed that the ECB was successful in reducing sovereign yields in the euro area with the PSPP. In this subsection, I investigate the drivers of this effect, heterogeneous effects across euro-area countries, and spillover effects to corporate and non-euro-area bond markets.

1.3.2.1 Euro-Area Sovereign Bonds

To understand how news about asset purchases affect euro-area sovereign bond markets, I reuse the bond yield decomposition introduced in Section 1.2 for the identification of asset purchase news. Thus, I consider sovereign yields to be the sum of the risk-free interest rate, a risk premium, and a scarcity premium. The previous subsection presented evidence that news about asset purchases reduce risk-free interest rates. Thus, yields should fall accordingly. Moreover, asset purchases may influence bond yields through the risk premium by affecting solvency considerations. Finally, asset purchase news must affect the scarcity premium in sovereign yields, since they are identified via changes in this scarcity premium.

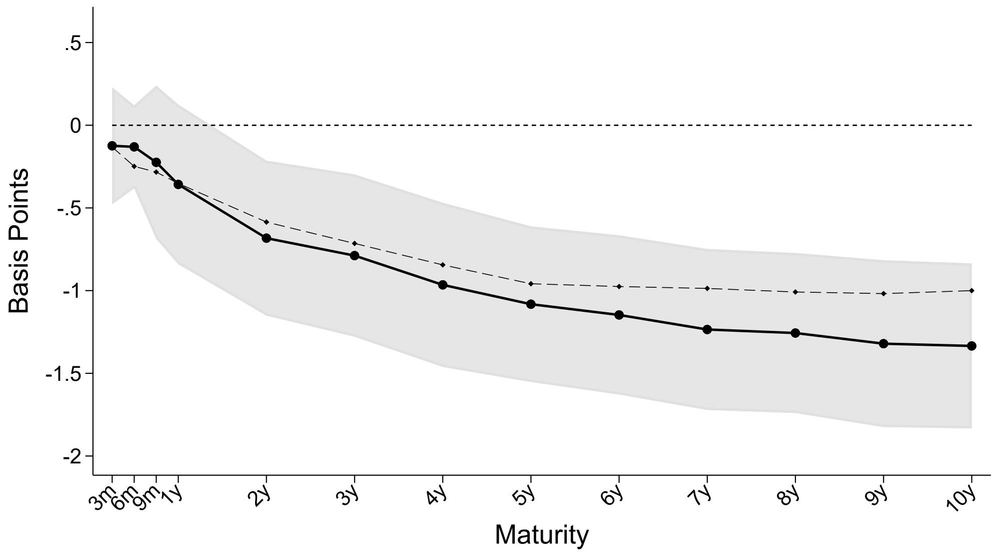


Figure 1.5. Response of German Sovereign Yields to Asset Purchase News

Notes: The dots represent the estimated $\hat{\beta}$ from separate regressions: $y_t = \alpha + \beta s_t + \varepsilon_t$, where y_t is the daily change, measured in basis points. Shocks are scaled to reduce the 10-year OIS rate by one basis point. Shaded areas depict the 90% confidence interval using standard errors robust to heteroskedasticity.

Figure 1.5 confirms that asset purchase news reduce sovereign yields across maturities, using the example of German bonds.[38] The effect is significant at the 10% significance level for maturities of two years and longer. As can be seen from the same figure, the magnitude of the effect exceeds the effect on risk-free interest rates at longer horizons, implying that either the risk premium or the scarcity premium, or both, were affected as well. This comes as no surprise, since I use changes in scarcity premiums to identify asset purchase news. A concern might therefore be that the measured statistical relationship is to some extent mechanical, because I use the German 10-year yield as a dependent variable and also for the construction

38. I use German bonds here, because these are often considered risk-free. The effects on French, Italian, and Spanish sovereign yields are displayed in Figure 1.G.1.

of asset purchase news. The robustness exercise in Section 1.4.1 shows, however, that estimates barely change if Germany is left out of the construction of the series of asset purchase news.

A recurring question with respect to asset purchases in the euro area is whether some countries are affected more than others. A common narrative is that sovereign yields fall most in countries with initially high sovereign yields. On the one hand, this could be due to asset purchases reducing risk premia and particularly so in countries with initially high risk premia. On the other hand, this could be due to a portfolio rebalancing towards riskier bonds in search for yield. Altavilla, Carboni, and Motto (2021) find evidence for the former mechanism using a number of events in late 2014 and early 2015.

Figure 1.6 shows the effect on 10-year sovereign yields of the ten largest euro-area countries. While I do find the effects to be strongest on Spanish and Italian yields, the point estimates are very similar across countries. Figure 1.G.2 confirms that sovereign spreads vis-à-vis Germany do not fall significantly. Figure 1.6 also decomposes the effect on 10-year yields into the effect on the three yield components. This reveals that country-specific risk premia (light gray bars) react very little and fall only for Portugal. Thus, I find little evidence for the "credit risk channel" of asset purchases, which holds that asset purchases reduce bond yields by reducing sovereign risk premia. The slight heterogeneity in responses across countries is mostly driven by the scarcity premium (medium gray bars). This heterogeneity is consistent with a portfolio rebalancing in search for yield, but could also reflect other reasons. The robustness exercise in Section 1.4.1 shows that the lack of heterogeneity is unchanged if the country on the left hand side of the regression is left out of the construction of the series of asset purchase news. Another interesting observation from Figure 1.6, which is in line with the evidence in Altavilla et al. (2019), is that the fall in the risk-free rate explains the majority of the reduction in sovereign yields.

In sum, the evidence shows that asset purchases reduce sovereign yields across countries and maturities. In contrast to some of the related literature, I do not find the effects to be very heterogeneous across countries. For example, Altavilla, Carboni, and Motto (2021) find asset purchases to narrow euro-area sovereign spreads. These contrasting findings can potentially be explained by the different samples. Altavilla, Carboni, and Motto (2021) use a number of events between September 2014 and March 2015. The current paper uses a later and much longer sample (October 2014 - January 2020), during which sovereign risk premia were on average lower.[39] Thus, there was less scope for asset purchases to have heterogeneous ef-

39. Credit default swap rates under the 2003 protocol, which measure the default risk premium, were on average 148 (Italy) and 106 (Spain) basis points over the period September 2014 - March 2015. The averages fall to 130 (Italy) and 79 (Spain) basis points over the period October 2014 - January 2020.

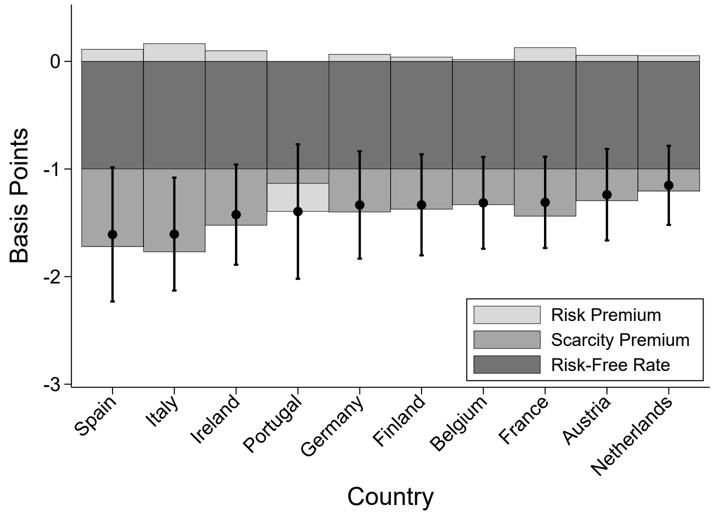


Figure 1.6. Response of 10-Year Sovereign Yields to Asset Purchase News

Notes: The dots represent the estimated $\hat{\beta}$ from separate regressions: $y_t = \alpha + \beta s_t + \varepsilon_t$, where y_t is the daily change, measured in basis points. Shocks are scaled to reduce the 10-year OIS rate by one basis point. Whiskers depict the 90% confidence interval using standard errors robust to heteroskedasticity. Dark gray bars depict the effect on risk-free interest rates, medium gray bars the effect on scarcity premiums, and light gray bars the effect on country-risk premiums.

fects by reducing sovereign risk premia. Figure 1.G.3 shows that using the QE factor of Altavilla et al. (2019) in the sample used throughout this paper, one finds similarly homogeneous effects across countries. This corroborates that the finding of no heterogeneous effects is a feature of the sample period and not of the series of shocks. I conclude that the effect of asset purchases on sovereign yields differs across countries only under certain circumstances, such as elevated sovereign risk premia.

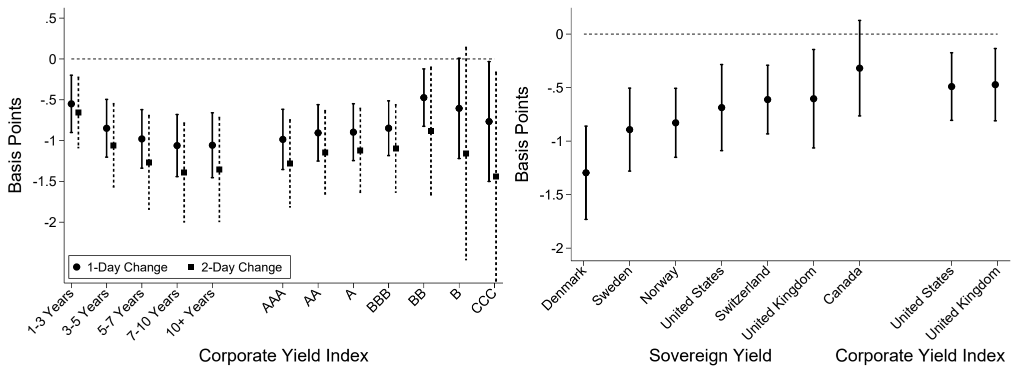
1.3.2.2 Corporate Bonds

In light of the previous results, one would expect there to be spillover effects of asset purchase news to the market for corporate bonds. On the one hand, the fall in risk-free interest rates should reduce corporate yields. On the other hand, risk premia might fall or there might be a portfolio rebalancing towards corporate bonds. Importantly, there is no direct effect on the corporate bond market, as under the PSPP, the ECB only purchased government and supranational bonds. I do not measure asset purchase news with respect to the Corporate Sector Purchase Programme (CSPP), under which the ECB directly purchased corporate bonds.[40] This section focuses on the effect of asset purchase news on a number of euro-area corporate bond

40. In Section 1.4.2, I discuss the CSPP in more detail and verify that it does not drive the results regarding the corporate sector.

indices.[41] Due to a lack of country-specific corporate bond indices, I do not study heterogeneities across euro-area countries in the effect of asset purchase news on corporate yields. Nonetheless, this constitutes a highly interesting avenue for future research.

Figure 1.7a displays the effect of asset purchase news on euro-area corporate bond yield indices of various maturities and credit ratings. Evidently, asset purchases reduce corporate yields across the board. In line with the effect on risk-free interest rates and sovereign yields, the effect increases with the remaining maturity. However, speculative grade corporate bond yields fall less strongly than investment grade yields in regressions using the 1-day yield change. This suggests that the market for speculative grade corporate debt is relatively illiquid, which hampers the transmission of lower interest rates. Indeed, using 2-day changes, the effect is more homogeneous across rating buckets. Moreover, using 2-day yield changes, the effect of asset purchase news is larger on corporate yield indices in general, which is suggestive of illiquidities in all segments of the market for corporate debt. Recalling that the effect of asset purchase news on 10-year risk-free interest rates is normalized to 1 basis point, we find corporate spreads over the risk-free rate to fall only when using 2-day changes.[42] Thus, there seems to be an effect on corporate bond yields beyond the risk-free rate reduction, so via risk premia or a portfolio rebalancing.



(a) Euro-Area Corporate Bond Yields

(b) Non-Euro Area Bond Yields

Figure 1.7. Spillovers to Euro-Area Corporate and Non-Euro Area Bonds

Notes: The dots represent the estimated $\hat{\beta}$ from separate regressions: $y_t = \alpha + \beta s_t + \varepsilon_t$, where y_t is the daily (or 2-day) change, measured in basis points. Shocks are scaled to reduce the 10-year OIS rate by one basis point. Whiskers depict the 90% confidence interval using standard errors robust to heteroskedasticity.

41. I use the Bank of America Merrill Lynch EMU Corporates Non-Financial AAA, AA, A, and BBB indices, the Merrill Lynch Euro High Yield BB, B, and CCC and Lower indices, as well as the Bank of America Merrill Lynch EMU Corporates Non-Financial 1-3Yr, 3-5Yr, 5-7Yr, 7-10Yr, and 10+Yr indices.

42. For example, the spread between the 7-10-year corporate yield index and the 10-year OIS rate falls by a mere 6 basis points ($t = -0.56$) in a 1-day regression, but by almost 36 basis points ($t = -1.76$) in a 2-day regression.

1.3.2.3 Non-Euro Area Bonds

Figure 1.7b shows that asset purchases by the ECB not only affect euro-area sovereign and corporate bond markets, but also bond markets beyond the euro area. 10-year sovereign yields of several advanced economies with tight financial linkages to the euro area (Denmark, Sweden, Norway, United States, Switzerland, United Kingdom) fall significantly in response to asset purchase news. Canadian sovereign yields also fall but insignificantly. Corporate yields in the U.S. and U.K. also fall significantly. On the one hand, these effects could be driven by reduced risk-free interest rates also in these other economies. On the other hand, they could reflect a portfolio rebalancing towards non-euro area bonds. The latter mechanism aligns well with Bergant, Fidora, and Schmitz (2018), who find evidence for a portfolio rebalancing towards debt instruments issued in non-euro area advanced economies during the PSPP period using quarterly portfolio holdings data.

1.3.3 Exchange Rates

Before turning to stock markets, it is useful to estimate how asset purchase news affect exchange rates. Figures 1.6 and 1.7b have shown that euro-area sovereign bond yields fall to a larger extent than sovereign bond yields outside the euro area. According to the uncovered interest rate parity, this should go hand in hand with a depreciation of the euro. Indeed, Figure 1.8 shows that asset purchase news significantly depreciate the euro against all major currencies.

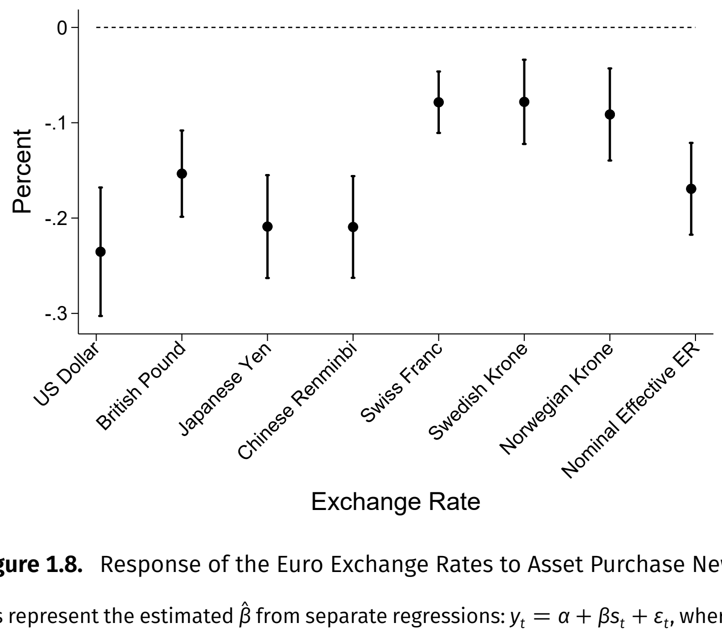


Figure 1.8. Response of the Euro Exchange Rates to Asset Purchase News

Notes: The dots represent the estimated $\hat{\beta}$ from separate regressions: $y_t = \alpha + \beta s_t + \varepsilon_t$, where y_t is the daily change, measured in percent. Shocks are scaled to reduce the 10-year OIS rate by one basis point. Whiskers depict the 90% confidence interval using standard errors robust to heteroskedasticity. Exchange rates are denoted in foreign currency per euro.

The magnitude of the depreciation is quite large, as a shock which reduces the 10-year OIS rate by one basis point depreciates the euro vis-à-vis the U.S. dollar by almost 0.25%. This magnitude exceeds most estimates of the effect of ECB conventional monetary policy on the exchange rate.[43] These findings echo Glick and Leduc (2018), who find that the Fed's unconventional monetary policy announcements had a much larger effect on the dollar exchange rates than previous conventional monetary policy announcements.

1.3.4 Stock Markets

There are plenty of reasons to expect asset purchases to increase stock prices. First, the results in Section 1.3.1 show that asset purchases reduce risk-free interest rates, which all else equal implies higher stock prices through a discounting effect. Moreover, asset purchases improve financing conditions for firms, as shown in Section 1.3.2.2, which may increase stock prices via higher expected dividends. Finally, depreciated exchange rates, as shown in Section 1.3.3, or generally higher growth expectations, may increase stock prices.

1.3.4.1 Euro-Area Stocks

Figure 1.9 shows the effect of asset purchase news on euro-area national stock indices. As expected, stock prices rise significantly in the ten largest euro-area countries. The European STOXX 50 index similarly increases by almost 0.3% in response to asset purchase news, which reduce the 10-year OIS rate by one basis point. The QE factor of Altavilla et al. (2019) has a much smaller effect on the STOXX 50, in line with the conjecture that this measure of asset purchases also reflects central bank information shocks, which bias the estimate towards zero.

There is an interesting heterogeneity across euro-area countries in the magnitude of the effect of asset purchases on stock indices. Stock prices rise most in Germany (DAX), France (CAC 40), the Netherlands (AEX), and Italy (FTSE MIB). What these four countries have in common is that their national stock indices include a number of very large firms. By total market capitalization and market capitalization per constituent, the German, French and Dutch indices are a lot larger than the other indices. The Italian index is the ranks fifth. There are several mechanisms which could explain why asset purchases potentially benefit large firms more than small firms.[44] On the one hand, larger firms might have better bond market access and therefore be able to make better use of the improved financing conditions. On the other hand, larger firms might rely more on exports and therefore benefit

43. For example, Altavilla et al. (2019) find their target factor, scaled to reduce the 1-month OIS rate by one basis point, to depreciate the euro by 0.06%.

44. Of course, none of the firms listed on these stock indices are "small" firms by usual definitions. Nevertheless, there are still large differences in size among these public firms.

more from the depreciated exchange rate. Moreover, large firms potentially benefit more from low interest rates in general, as in the model of Liu, Mian, and Sufi (2022). Further investigating the reasons for this heterogeneity is high on my research agenda. I confirm that this finding is robust to using 2-day changes in stock prices (Section 1.4.3), excluding key CSPP dates (Section 1.4.2), and controlling for macroeconomic data releases (Appendix 1.C).

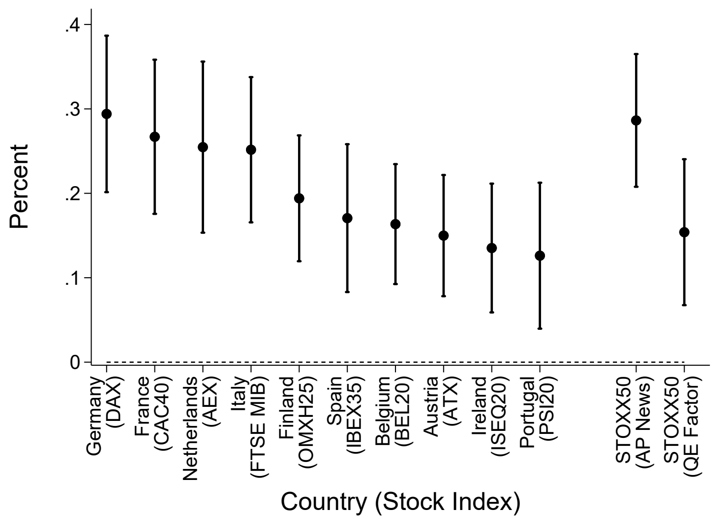


Figure 1.9. Response of National Stock Indices to Asset Purchase News

Notes: The dots represent the estimated $\hat{\beta}$ from separate regressions: $y_t = \alpha + \beta s_t + \varepsilon_t$, where y_t is the daily change, measured in percent. Shocks are scaled to reduce the 10-year OIS rate by one basis point. Whiskers depict the 90% confidence interval using standard errors robust to heteroskedasticity.

1.3.4.2 Non-Euro Area Stocks

Figure 1.10 shows that asset purchases not only increase stock prices in the euro area, but also beyond it. Stock indices in a number of advanced economies with tight financial linkages to the euro area (Sweden, United Kingdom, Denmark, Switzerland, Norway, Canada, United States) rise significantly. Again, there are several potential mechanisms for this, including lower discount rates and higher growth expectations. However, finding large effects on non-euro-area stock prices speaks against a central role of the exchange rate for the effect of asset purchase news on stock markets. Firms outside the euro area lose competitiveness due to the appreciation of their own currencies against the euro, but nevertheless see rising stock prices. In particular Norway, Sweden, and Switzerland should be hit hard by this, since a large share of their exports go to the euro area.

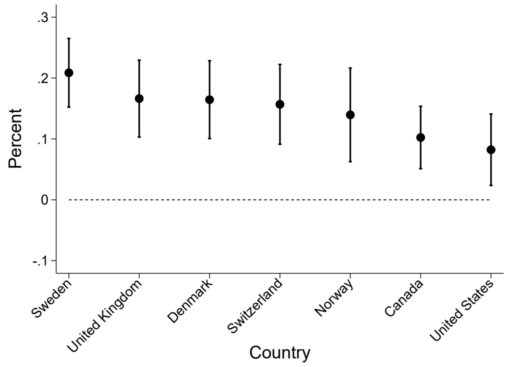


Figure 1.10. Response of Stock Indices Beyond the Euro Area

Notes: The dots represent the estimated $\hat{\beta}$ from separate regressions: $y_t = \alpha + \beta s_t + \varepsilon_t$, where y_t is the daily change, measured in percent. Shocks are scaled to reduce the 10-year OIS rate by one basis point. Whiskers depict the 90% confidence interval using standard errors robust to heteroskedasticity.

1.3.5 Comparison to the Literature

Since there already exists a substantial amount of research on the financial effects of the ECB's unconventional monetary policy measures, I compare the results found in this paper with those documented in the related literature. I focus on studies explicitly analyzing the effects of the PSPP. For an overview of the effects of the ECB's unconventional monetary policy measures more generally, see Fratzscher, Lo Duca, and Straub (2016), Rogers, Scotti, and Wright (2014), Dell'Ariccia, Rabanal, and Sandri (2018) and references therein.

I first of all document that the asset purchases under the PSPP reduce euro-area sovereign yields. This is a common finding and it is largely undisputed that the ECB was successful in reducing euro-area sovereign yields with the PSPP. A more interesting question is whether the fall in yields is heterogeneous across countries. I document yields to fall in a relatively homogeneous manner in response to asset purchase news. Thus, I do not find asset purchases to significantly narrow sovereign spreads. This is a surprising finding, since several papers, including Altavilla, Carboni, and Motto (2021) and De Santis (2020) find asset purchases to narrow sovereign spreads. To reconcile these disparate findings, it is important to note that these papers focus on the first months of the PSPP and the period leading up to it, during which sovereign risk premia were still somewhat elevated.[45] I instead document the lack of significant heterogeneity in a later and much longer sample from October 2014 until January 2020. During this period of time, sovereign risk premia were on average lower, such that there was less scope for asset purchases to have heterogeneous effects by reducing sovereign risk premia. I show that using

45. Altavilla, Carboni, and Motto (2021) consider events between September 2014 and March 2015. De Santis (2020) uses a sample from September 2014 to October 2015.

the shocks of Altavilla et al. (2019) in this longer sample, one finds a similar lack of heterogeneous effects. I conclude that the effect of asset purchases on sovereign yields differs across countries only under certain circumstances, such as elevated sovereign risk premia.

Second, I document spillover effects of asset purchases to euro-area corporate bond yields and stock prices. This is in line with Altavilla, Carboni, and Motto (2021), who similarly find spillovers to corporate yields, which increase in size when 2-day yield changes are used, and euro-area stock prices. Georgiadis and Gräb (2016) and Bubeck, Habib, and Manganelli (2018) also document significant and sizeable effects of the PSPP on euro-area equities. As discussed before, Altavilla et al. (2019) find surprisingly small and insignificant effects of QE on stock prices, potentially due to the presence of information shocks. Moreover, I document that asset purchases increase national stock indices more strongly in countries with very large firms (Germany, France, Netherlands, Italy). To the best of my knowledge, this finding has not been documented previously. De Santis (2020) points out that stock prices rose most in Germany and Italy on average across three QE dates in 2015, but does not further analyze this observation.

Finally, I show that asset purchases reduce bond yields and increase stock prices in advanced economies beyond the euro area, and depreciate the euro. These three effects are also documented in Georgiadis and Gräb (2016) and Bubeck, Habib, and Manganelli (2018). A small difference is that Georgiadis and Gräb (2016) do not find yields in the U.S. to fall significantly, whereas Bubeck, Habib, and Manganelli (2018) and I do find them to fall significantly. The finding that asset purchases depreciate the euro is also documented in Altavilla, Carboni, and Motto (2021), Altavilla et al. (2019), and Dedola et al. (2021).

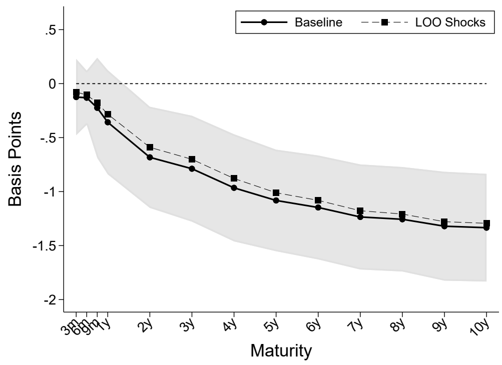
1.4 Robustness Checks

I now discuss a number of robustness exercises. In the interest of space, I focus on Figures 1.6 and 1.9 to show that the most interesting findings of heterogeneous effects on euro-area national stock indices and a lack thereof on sovereign bond yields are robust.

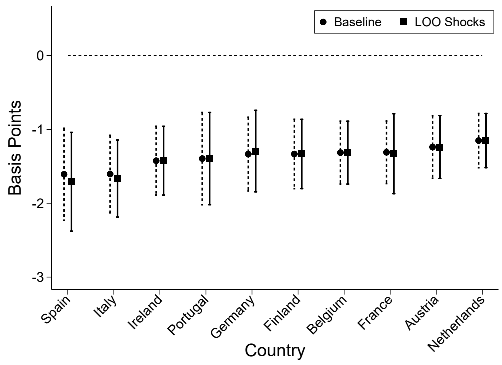
1.4.1 Leave-One-Out Asset Purchase News

I use some financial variables, such as 10-year sovereign yields, for the identification of asset purchase news and as outcome variables. Therefore, a concern might be that the measured statistical relationships between the series of asset purchase news and those variables is to some extent mechanical. To investigate this issue, I construct leave-one-out asset purchase news series. These follow the identification strategy explained in Section 1.2, but leave out one country in the aggregation of country-specific series of scarcity premium changes. Figure 1.11 replicates Figures 1.5, 1.6

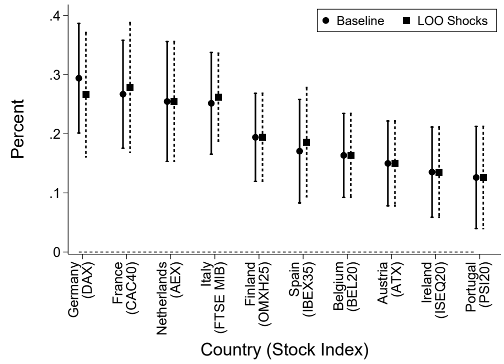
and 1.9 using the respective leave-one-out series for each country. As an example, panel (a) of Figure 1.11 uses a series of asset purchase news constructed from French, Spanish, and Italian scarcity premium changes, because German yields are used as outcome variables. Evidently, estimates using leave-one-out shocks (squares) barely differ from the baseline estimates (circles). There are no meaningful differences in panels (b) and (c), either. This rules out the concern that the results regarding sovereign yields merely reflect a mechanical correlation. Note that the leave-one-out shock series for countries not used in the construction of baseline asset purchase news, such as Belgium, equal the baseline series.



(a) Sovereign Yields (Germany)



(b) Sovereign Yields (Euro Area)



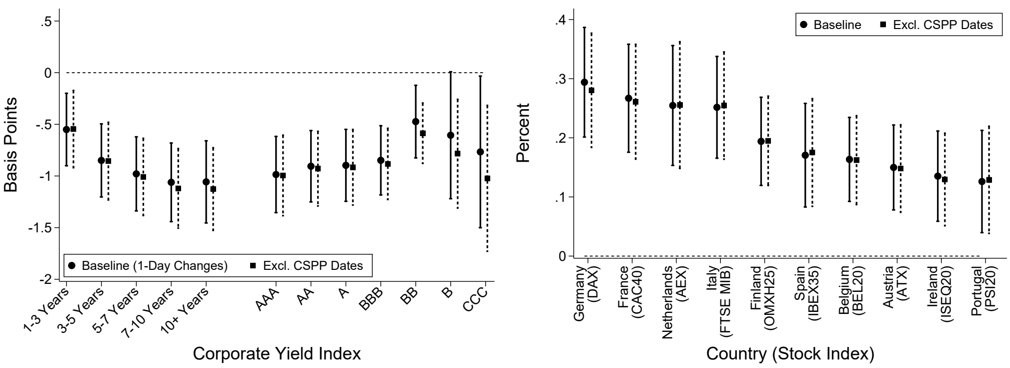
(c) National Stock Indices

Figure 1.11. Robustness Exercise - Leave-One-Out (LOO) Asset Purchase News

Notes: The dots represent the estimated $\hat{\beta}$ from separate regressions: $y_t = \alpha + \beta s_t + \varepsilon_t$, where y_t is the daily change, measured in basis points or percent (panel c). The squares represent the estimated $\hat{\beta}$ from separate regressions which use leave-one-out shocks. Shocks are scaled to reduce the 10-year OIS rate by 1 basis point. The shaded area / whiskers depicts the 90% confidence interval using standard errors robust to heteroskedasticity.

1.4.2 Corporate Sector Purchase Programme

As discussed and demonstrated in Section 1.2, the series of asset purchase news successfully captures key announcements regarding the PSPP and is unrelated to conventional monetary policy shocks, forward guidance shocks, and central bank



(a) Corporate Yields **(b)** National Stock Indices

Figure 1.12. Robustness Exercise - Corporate Sector Purchase Programme

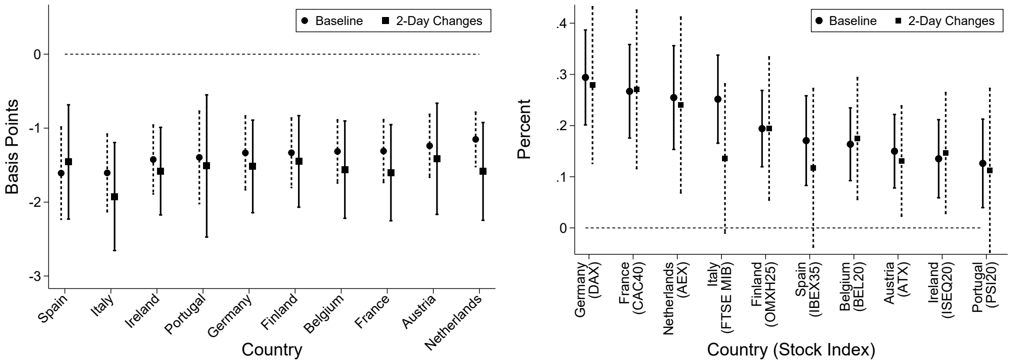
Notes: The dots represent the estimated $\hat{\beta}$ from separate regressions: $y_t = \alpha + \beta s_t + \varepsilon_t$, where y_t is the daily change, measured in basis points. The squares represent the estimated $\hat{\beta}$ from separate regressions where the three CSPP dates have been excluded. Shocks are scaled to reduce the 10-year OIS rate by 1 basis point (before excluding dates). Whiskers depict the 90% confidence interval using standard errors robust to heteroskedasticity.

information shocks. However, the series of asset purchase news might be correlated with news about *other* asset purchase programs, in particular the Corporate Sector Purchase Programme (CSPP), for two reasons. First, announcements of corporate bond purchases might have spillover effects to the sovereign bond market, therefore affecting the scarcity premium in sovereign bonds and, thereby, my measure of asset purchase news. Second, in the later part of my sample, the ECB oftentimes made announcements about total purchase amounts under the APP without specifying amounts under each single program. On the contrary, the left panel of Figure 1.1 showed that the PSPP sizewise clearly dominates the other asset purchase programs. By December 2019, the ECB had spent more than 10 times as much under the PSPP as compared to the CSPP (€2100bn vs. €184bn). Relative to the amount of eligible bonds, the amount purchased under the PSPP also clearly exceeds the amount purchased under the CSPP.[46]

To verify empirically that my results regarding the corporate sector are not driven by the CSPP, I replicate Figures 1.7a and 1.9 while leaving out major CSPP announcements. According to Dedola et al. (2021), such major announcements were made on March 10, April 21, and June 2 in 2016. Figure 1.12 shows that excluding these three dates leaves the effect of asset purchase news on corporate bonds and stock indices almost unchanged. I conclude that the results regarding the corporate sector are robust to excluding dates with major CSPP announcements.

46. Relative to the respective eligible bond universe, the size of the PSPP exceeds the size of the CSPP almost by a factor of 2.

1.4.3 2-Day Changes



(a) Sovereign Yields (Euro-Area) **(b)** National Stock Indices

Figure 1.13. Robustness Exercise - 2-Day Changes

Notes: The dots represent the estimated $\hat{\beta}$ from separate regressions: $y_t = \alpha + \beta s_t + \varepsilon_t$, where y_t is the daily change, measured in basis points (panel a) or percent (panel b). The squares represent the estimated $\hat{\beta}$ from separate regressions where y_t is the two-day change. Shocks are scaled to reduce the 10-year OIS rate by 1 basis point. Whiskers depict the 90% confidence interval using standard errors robust to heteroskedasticity.

Section 1.3.2.2 showed that using 2-day changes in outcome variables makes a difference for the estimated effects of asset purchase news on corporate bond yields. To investigate whether this is also the case for other variables, Figure 1.13 replicates Figures 1.6 and 1.9 using 2-day changes in outcome variables. The left panel shows that the estimated effects on sovereign yields tend to be a bit larger. However, the lack of meaningful heterogeneity remains. Figure 1.G.2 confirms that spreads do not change significantly using 2-day changes. The right panel shows that the estimated effects on national stock indices remain largely unchanged, except for Italy. The Italian stock index rises a lot less strongly using 2-day changes. Nevertheless, the main result that stock prices rise more strongly in Germany, France, and the Netherlands, the countries with particularly large firms, is unchanged.

1.5 Conclusion

In this paper, I propose a novel strategy to identify the effects of central bank purchases of government debt. I build on high-frequency identification and propose to use changes in the "scarcity premium", i.e. the component of government bond yields in excess of the risk-free interest rate and the risk premium, to measure news about asset purchases. The idea is that central bank asset purchases reduce the supply of government debt available to the public. The supply by governments and demand by investors being relatively inelastic, asset purchases thus reduce government bond yields, even in the absence of movements in risk-free interest rates or the risk premium. Hence, asset purchases affect the scarcity premium, in contrast to other monetary policy measures, which affect government bond yields only via the risk-free interest rate and the risk premium.

Employing the identified series, I estimate the effects of asset purchases on financial markets. I find that central bank purchases of government debt reduce not only the yields of euro-area government bonds, but also the yields of corporate bonds and non-euro area government and corporate bonds. At the same time, stock prices rise in the euro area and in other advanced economies. In addition, asset purchases reduce risk-free interest rates and strongly depreciate the euro against all major currencies. Investigating differences across euro-area countries, I find that asset purchases reduce sovereign bond yields rather homogeneously. That is, sovereign yield spreads vis-à-vis Germany do not fall significantly. In contrast, I find that asset purchases have heterogeneous effects on stock prices. National stock indices increase the most in Germany (DAX), France (CAC 40), and the Netherlands (AEX), i.e. countries with relatively few concerns about sovereign solvency. These two pieces of evidence may raise doubts about the prevalent view that asset purchases mostly benefit highly indebted countries.

Directly investigating the real effects of asset purchases, for example by means of a structural VAR with an external instrument, is challenging due to data limitations, but constitutes an interesting avenue for future research. The identification strategy developed in this paper and the series of asset purchase news may provide a helpful starting point, in particular in contexts where concurrent information shocks are a concern.

Appendix 1.A Factor Shocks

Throughout the paper, I use data up until January 31, 2020. This sample selection aims to exclude the COVID-19 crisis from the main analysis. An extended sample including the COVID-19 period is considered in Appendix 1.F.

Updated Factors. To be able to use the entire sample, I re-estimate the factors of Altavilla et al. (2019) in this longer sample and use these updated series throughout the paper. I confirm that the original and re-estimated factors align very well in the original sample (01/2002 - 09/2018). The correlation coefficients exceed 0.995 for all three conference factors.

Policy & Information Shocks. I identify policy and information shocks using the rotational sign restriction approach used in Jarociński and Karadi (2020) and as implemented in Jarociński (2021). I use the monetary event window surprises in the period during which the QE factor is active (01/2014 - 01/2020). I choose as interest rate measure the first principal component of the standardized 1Y, 2Y, 5Y, and 10Y OIS rate surprises, since shorter-horizon OIS rates were severely constrained by the ELB. I use all three options to compute information shocks explained in Jarociński (2021). The easiest approach are the “poor man’s sign restrictions”. Here, interest rate surprises are classified as policy (information) shocks if the stock market surprise has the opposite (same) sign. The other two options explain the interest rate and stock price surprise with two factors. The sign restriction provides a set of admissible rotations of the two factors. Then, there are two options to choose the rotation among all admissible rotations. The first option is use the rotation which ensures that the variance share of interest rate surprises explained by information shocks is the same as when using poor man's shocks. I follow Jarociński (2021) and use this as a baseline. The other option is to use the median rotation among all admissible rotations, as used in Andrade and Ferroni (2021). More details can be found in the Appendix of Jarociński (2021).

Appendix 1.B Persistence of Impact Effects

The estimated impact effects on financial variables might have a reduced relevance for the real economy, if they are not persistent. For example, financial markets could overreact to asset purchase news initially. Several studies, including Wright (2012), Rogers, Scotti, and Wright (2014), and Greenlaw et al. (2018), argue that this was the case for the Federal Reserve's Quantitative Easing programs.

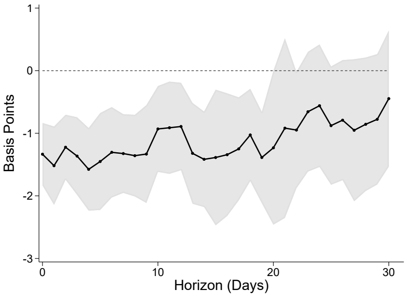
To evaluate the persistence of the impact effects of asset purchase news, I use the same event-study regression as before, but replace the one-day change as the outcome variable with the h-day change for horizons h between 0 and 30. These are essentially Jordà (2005) local projections:

$$y_{t+h} - y_{t-1} = \alpha_h + \beta_h s_t + \epsilon_{h,t}$$

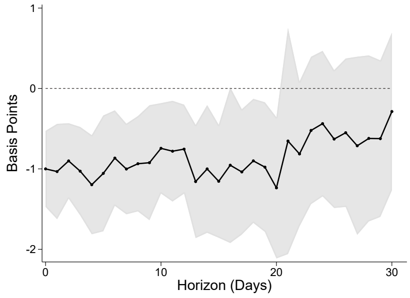
The parameter of interest is β_h, which captures the cumulative effect of asset purchase news over h trading days[47].

Figure 1.B.1 shows the effect of asset purchase news over a horizon of 30 days on the German 10-year sovereign yield, the 10-year OIS rate, the AAA corporate yield index, the Euro/Dollar exchange rate, the DAX stock index and the STOXX50 stock index. Overall, the impact effects are fairly persistent and remain significant for quite some time. Confidence intervals naturally widen due to the amount of noise accumulating over 30 trading days. This is in contrast to the U.S. evidence and more in line with Altavilla et al. (2019), who also find the announcement effects of QE in the euro area to be rather persistent. These disparate findings in euro area and U.S. can be explained by market participants learning about the effects of asset purchases over time, or the fact that the Federal Reserve's programs were implemented in times of higher financial distress.

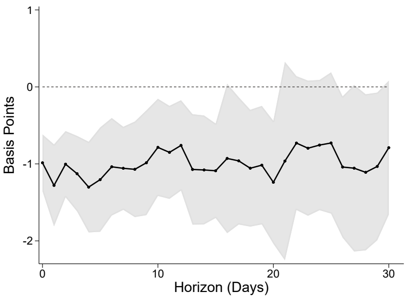
47. Days without trading, such as weekends and public holidays are excluded.



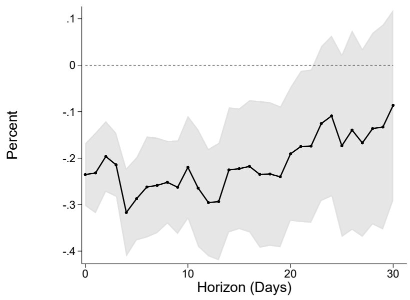
(a) German 10-Year Sovereign Yield



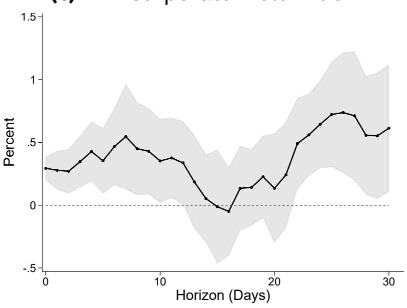
(b) 10-Year OIS Rate



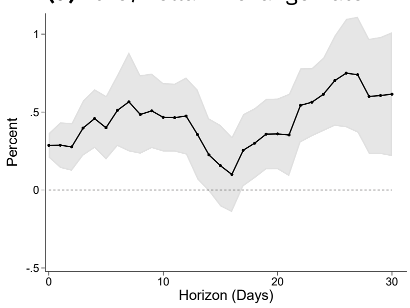
(c) AAA Corporate Yield Index



(d) Euro/Dollar Exchange Rate



(e) German Stock Index (DAX)



(f) STOXX50 Index

Figure 1.B.1. Persistence of Impact Effects

Notes: The dots represent the estimated $\hat{\beta}_h$ from separate regressions: $y_{t+h} - y_{t-1} = \alpha_h + \beta_h s_t + \varepsilon_{h,t}$, where the left-hand side is the change over h days, measured in basis points or percent, respectively. Shocks are scaled to reduce the 10-year OIS rate by one basis point. Shaded areas depict the 90% confidence interval using standard errors robust to heteroskedasticity and autocorrelation.

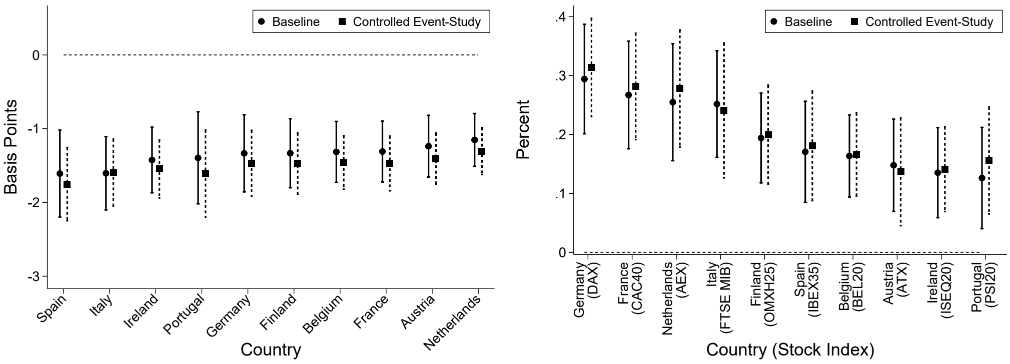
Appendix 1.C Macroeconomic Data Releases

A number of studies, including Gürkaynak, Sack, and Swanson (2005b), Swanson and Williams (2014a) and Swanson and Williams (2014b), have shown that the surprise component of macroeconomic data releases has the potential to move long-term bond yields. These surprises, also referred to as "macroeconomic news", provide information about the state of the economy. Positive news typically lift interest rate expectations, in line with the notion that the central bank will eventually tighten its policy. In view of the sovereign bond yield decomposition of Section 1.2, bond yields should rise accordingly. Beyond their effect on risk-free interest rates, it is possible that some macroeconomic news also affect the (expected) demand and supply for sovereign debt, thereby affecting the scarcity premium. In this section, I discuss the concern that macroeconomic releases coinciding with ECB Governing Council Meeting days might affect either the measure of asset purchase news or any results.

For this purpose, I retrieve data on macroeconomic releases from Bloomberg. I include all country-specific data releases for the ten largest euro-area countries as well as euro-area-wide data releases with a nonzero relevance for financial markets. Moreover, I include the 40 U.S. data releases with the highest relevance for financial markets. I drop observations where there were less than 8 participants in the prerelease survey. Then, I drop all series for which less than 10 observations remain in total. This leaves a total of 102 series from the euro area and 37 from the U.S.. I compute the surprise as the difference between the actual value and the survey median.

Macroeconomic releases occurring during the high-frequency window on ECB GCM days could in principle affect the scarcity premium and thus the series of asset purchase news. There is one important data release, which frequently coincides with the ECB communication on GCM days, namely the U.S. initial jobless claims (IJC). However, the correlation of IJC surprises and asset purchase news is small and insignificant (ρ=0.15, p-value=0.3444, N=41). Macroeconomic releases occurring outside the high-frequency window, but on ECB GCM days, cannot affect the measured yield and OIS rate surprises, but could still affect the (daily) CDS surprises used for the construction of the asset purchase news. There are no data releases (except for the IJC) which regularly coincide with ECB GCM days, however, there are some releases which coincide a few (i.e. up to 7) times. To test whether this is a quantitatively relevant concern, I reconstruct the series of asset purchase news after orthogonalizing the CDS rate changes with respect to euro-area and U.S. macroeconomic surprises. The correlation of this series with the baseline series of asset purchase news is extremely high ($\rho = 0.975$). I conclude that there is no evidence that macroeconomic data releases affect the series of asset purchase news in a quantitatively relevant way.

However, given the limited sample size, even if macroeconomic news do not affect the series of asset purchase news, they might still affect the outcome variables and therefore the reported results. To rule out this concern, I perform a “controlled” event study in the spirit of Altavilla, Giannone, and Lenza (2016). That is, I reestimate the effects of asset purchase news on sovereign yields and national stock indices controlling for all euro-area and U.S. surprises. The sample is October 1, 2014 until January 31, 2020. Figure 1.C.1 displays the results. Evidently, the estimates of the effect of asset purchase news on sovereign yields and national stock indices do not change much when controlling for macroeconomic news.



(a) Sovereign Yields

(b) National Stock Indices

Figure 1.C.1. Robustness Exercise - Controlling for Macroeconomic News

Notes: The dots represent the baseline estimates as in Figures 1.6 and 1.9. The squares represent the estimates controlling for macroeconomic surprises. Whiskers depict the 90% confidence interval using standard errors robust to heteroskedasticity.

Appendix 1.D Aggregation of Asset Purchase News

The baseline series of asset purchase news is computed as the first principal component of the (standardized) series of scarcity premium changes around ECB communication on Governing Council Meeting days. This aggregated series is highly correlated with the country-specific series. The correlations are 0.85 (FR), 0.79 (IT), 0.77 (ES), and 0.70 (DE). This reassures that the final series indeed measures a common shock, which is reflected in each country-specific series.

Vice versa, no single country is driving the baseline series. This can be assessed by computing four series of "leave-one-out" asset purchase news, by leaving out one country at a time in the aggregation procedure. The correlation of these four series with the baseline series never falls below 0.96, reassuring that no single country is disproportionately affecting the aggregated series.

An alternative aggregation procedure would be to use a GDP-weighted average of unstandardized scarcity premium changes. This produces a remarkably similar series of asset purchase news, the correlation with the baseline series is 0.99. Practically, GDP-weighting and standardizing to unit variance do (almost) the same thing. Both reduce the influence of the Italian and the Spanish series, because these countries have the lowest GDP and also the highest variance of scarcity premium changes.

Appendix 1.E Liquidity in CDS Markets

A potential concern with the identification strategy is that the market for CDS contracts is not as liquid as the markets for sovereign debt and OIS contracts. This potentially makes the scarcity premium reflect changes in the country-specific risk premium. As an example, if central bank communication changes the risk premium and this is correctly priced in sovereign yields, while CDS rates do not change because of illiquidity, the scarcity premium will reflect the change in the risk premium.

However, using a longer window (daily) for CDS surprises than for sovereign yield and OIS rate surprises (high-frequency), as discussed in Section 1.2, mitigates this concern to some extent. In addition, judging by the number of missing observations over the entire sample period, the CDS contracts for Italy and Spain (the countries with higher and more volatile CDS rates and generally more concern about sovereign risk) are relatively more liquid.

As a robustness check, I re-construct the series of asset purchase news using CDS contracts denominated in U.S. Dollar instead of those denominated in Euro, since USD CDS are typically more liquid than Euro CDS. Of course, USD CDS rates depend on the exchange rate, which is itself affected by monetary policy. For this reason, the baseline series uses Euro CDS. Nevertheless, the series of USD-CDS asset purchase news is highly correlated ($\rho = 0.94$) with the baseline series of asset purchase news, suggesting that CDS illiquidity is not a major concern for the identification strategy.

Appendix 1.F Extended Sample

The baseline sample used throughout this paper includes all ECB Governing Council Meetings between October 2014 and January 2020 (N=44). In this section, I show how the results are affected if this sample is extended.

The baseline sample begins in October 2014, because CDS contracts based on the 2014 ISDA Credit Derivatives Protocol are traded since September 22, 2014. These contracts insure against sovereign default and currency redenomination. CDS contracts under the previous 2003 ISDA Credit Derivatives Protocol do not insure against currency redenomination for G-7 countries (including Germany, France and Italy). Thus, before October 2014, it is not possible to separate the scarcity premium from a *redenomination risk premium* for these three countries. In consequence, a stronger identifying (exogeneity) assumption is needed. That is, one has to assume that all elements of ECB communication, except for asset purchase news, do not affect the redenomination risk premium. To avoid having to make this assumption, the baseline sample begins in October 2014, when CDS contracts based on the 2014 ISDA protocol become available. However, bearing the caveat in mind, it is of course possible to extend the series of asset purchase news backwards using CDS contracts under the 2003 ISDA protocol. This could be interesting because, even though the PSPP was officially announced in January 2015, expectations about large-scale asset purchases in the euro area started to form before October 2014. For example, Mario Draghi's speech at the Jackson Hole Symposium in August 2014 is often considered as having signaled that large-scale asset purchases are possible in the future.

The baseline sample ends in January 2020 with the last Governing Council Meeting before the COVID-19 crisis. The current release of the EA-MPD includes 3 additional Governing Council Meetings, so it is possible to extend the sample until June 2020. This could be interesting in order to see whether the identification strategy also works for the Pandemic Emergency Purchase Programme (PEPP). The reason not to include these three observations in the main analysis is that the PSPP and the PEPP are programs with very different intentions. The PSPP was intended to stimulate inflation and growth in the euro area as a whole. Thus, it did not intend to have heterogeneous effects across countries. The PEPP instead was primarily intended to mitigate concerns about sovereign debt sustainability in some euro-area countries. Therefore, even though purchases remained proportional, there was a relatively clear heterogeneous intention, as in some countries, there was a lot more concern about sovereign debt (e.g. Italy, Spain) than in others (e.g. Germany). Thus, focusing on the PSPP in the baseline sample provides a nice setting to analyze how asset purchases intended to stimulate economic activity work in times of low financial distress.

Figure 1.F.1 plots the series of asset purchase news for the baseline sample and an "extended" sample from January 2014 to June 2020. Evidently, shocks identified between October 2014 and January 2020 barely change. Before October 2014, there is one relatively large contractionary realization in April 2014, which is indeed one of the Governing Council Meetings where QE was discussed.[48] After January 2020, there is one relatively large expansionary realization in June 2020, which is also in line with the financial press.[49]

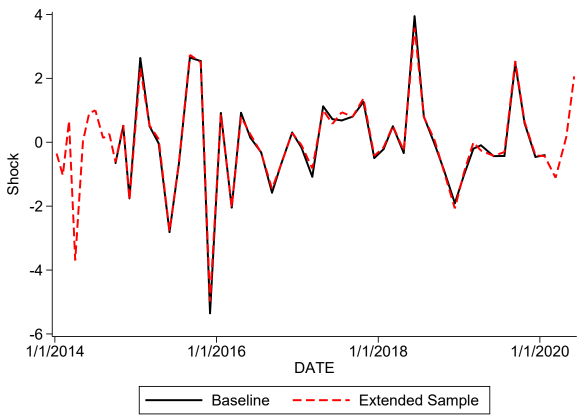
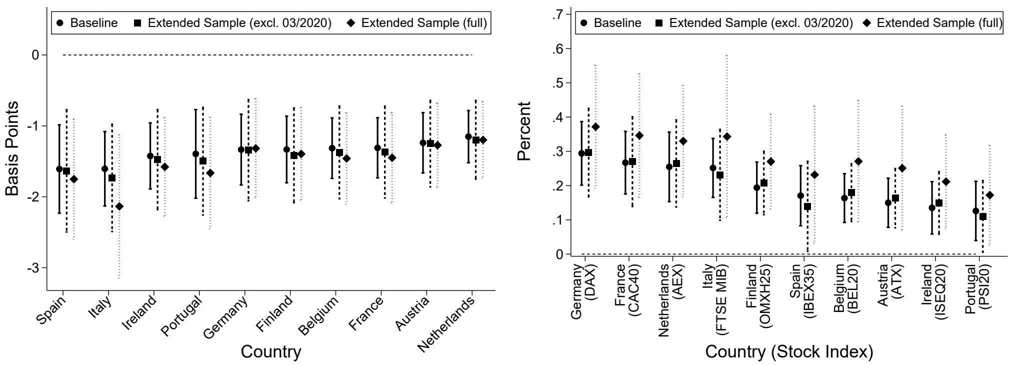


Figure 1.F.1. Series of Asset Purchase News (Baseline & Extended Sample)

Figure 1.F.2 shows that the effect of asset purchases on euro-area sovereign yields and national stock indices remains largely unchanged in the extended sample, as long as the March 2020 Governing Council Meeting remains excluded. Including

48. e.g. https://www.reuters.com/article/markets-forex-idUSL1N0MV1A720140403
49. e.g. https://www.cnbc.com/2020/06/04/european-central-bank-ramps-up-its-pandemic-bond-buying-to-1point35-trillion-euros.html

this date changes the magnitude of the coefficients for Italian sovereign yields as well as for stock prices, even though the measure of asset purchase news on this date is small (and negative). The reason is that outcome variables, in particular the Italian 10-year sovereign yield and all stock indices, moved dramatically in consequence of the comment made by President Lagarde that the ECB is "not here to close spreads". This event fits the interpretation of a "risk shock" (Kroencke, Schmeling, and Schrimpf, 2021), since sovereign yields moved a lot, driven by sovereign risk premiums. Scarcity premiums increased a bit, in line with the interpretation that this event constituted rather contractionary asset purchase news. Either way, the main qualitative results of heterogeneous effects on stock indices and a lack of a significant effects on sovereign yield spreads vis-à-vis Germany, remain unchanged also in the full extended sample.



(a) Sovereign Yields **(b)** National Stock Indices

Figure 1.F.2. Robustness Exercise - Extended Sample

Notes: The dots represent the baseline estimates as in Figures 1.6 and 1.9. The squares represent the estimates in the extended sample, excluding the March 2020 GCM. The diamonds represent the estimates from the full extended sample. Whiskers depict the 90% confidence intervals using standard errors robust to heteroskedasticity.

Appendix 1.G Additional Figures

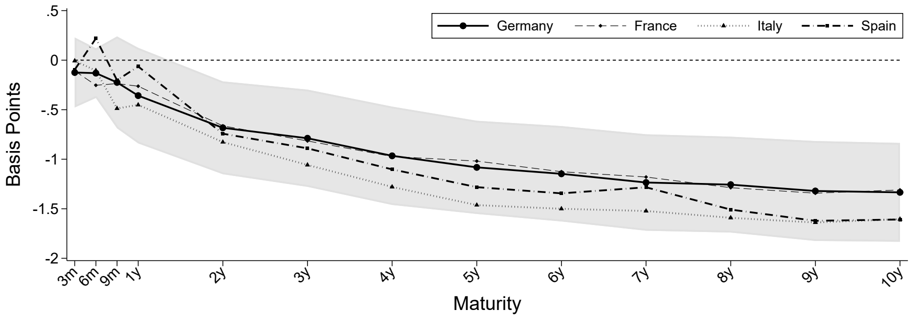


Figure 1.G.1. Response of Euro-Area Sovereign Yields to Asset Purchase News

Notes: The dots represent the estimated $\hat{\beta}$ from separate regressions: $y_t = \alpha + \beta s_t + \varepsilon_t$, where y_t is the daily change, measured in basis points. Shaded areas depict the 90% confidence interval using standard errors robust to heteroskedasticity for German sovereign yields.

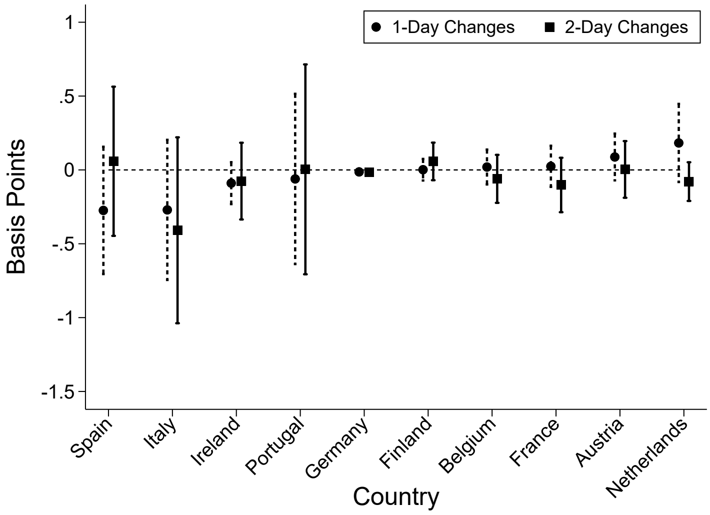


Figure 1.G.2. 10-Year Sovereign Spreads (vis-a-vis Germany) to Asset Purchase News

Notes: The dots represent the estimated $\hat{\beta}$ from separate regressions: $y_t = \alpha + \beta s_t + \varepsilon_t$, where y_t is the daily change, measured in basis points. Shocks are scaled to reduce the 10-year OIS rate by one basis point. Whiskers depict the 90% confidence interval using standard errors robust to heteroskedasticity.

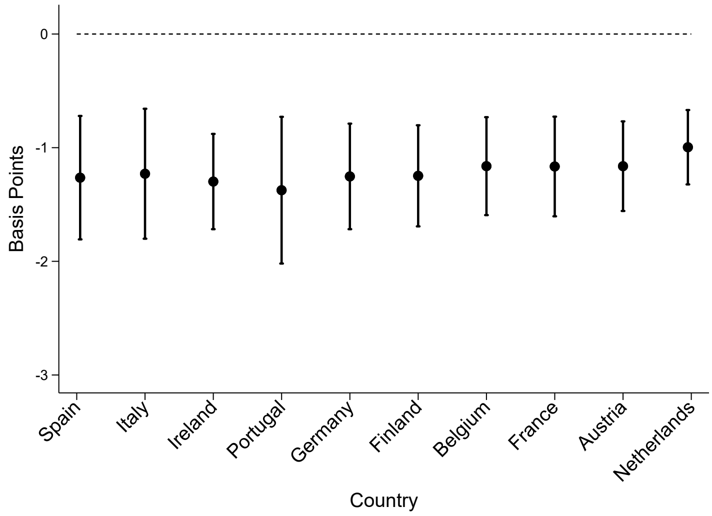


Figure 1.G.3. Response of 10-Year Euro-Area Sovereign Yields to QE Factor

Notes: The dots represent the estimated $\hat{\beta}$ from separate regressions: $y_t = \alpha + \beta s_t + \varepsilon_t$, where y_t is the daily change, measured in basis points. Shocks are scaled to reduce the 10-year OIS rate by one basis point. Whiskers depict the 90% confidence interval using standard errors robust to heteroskedasticity.

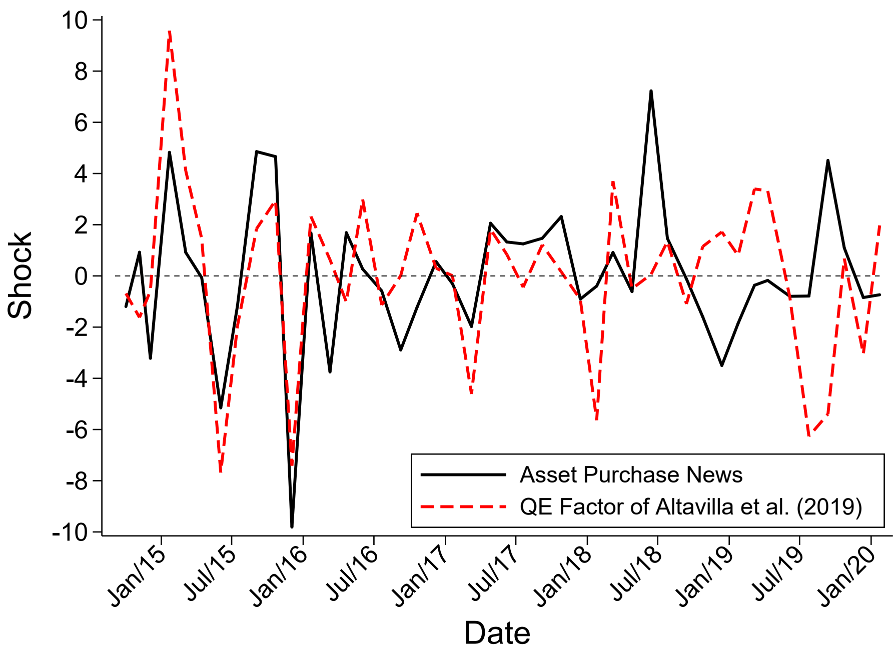


Figure 1.G.4. Series of Asset Purchase News & QE Factor of Altavilla et al. (2019)

Notes: Both series are scaled to reduce the 10-year OIS rate by one basis point. Sample: October 2014 - January 2020 (N=44).

References

Albertazzi, Ugo, Bo Becker, and Miguel Boucinha. 2021. "Portfolio rebalancing and the transmission of large-scale asset purchase programs: Evidence from the Euro area." *Journal of Financial Intermediation* 48: 100896. [8]

Altavilla, Carlo, Luca Brugnolini, Refet Gürkaynak, Roberto Motto, and Giuseppe Ragusa. 2019. "Measuring euro area monetary policy." *Journal of Monetary Economics* 108: 162–79. [4, 7–10, 13, 14, 16–19, 21, 24, 25, 28, 31, 35, 36]

Altavilla, Carlo, Giacomo Carboni, and Roberto Motto. 2021. "Asset purchase programmes and financial markets: lessons from the euro area." *International Journal of Central Banking* 17 (4). [5, 8, 24, 30, 31]

Altavilla, Carlo, Domenico Giannone, and Michele Lenza. 2016. "The Financial and Macroeconomic Effects of the OMT Announcements." *International Journal of Central Banking* 12 (3): 29–57. [38]

Andrade, Philippe, Johannes Breckenfelder, Fiorella De Fiore, Peter Karadi, and Oreste Tristani. 2016. "The Reanchoring Channel of QE - The ECB's Asset Purchase Programme and Long-Term Inflation Expectations." *ECB Working Paper No. 1956, European Central Bank.* [9]

Andrade, Philippe, and Filippo Ferroni. 2021. "Delphic and Odyssean monetary policy shocks: Evidence from the euro area." *Journal of Monetary Economics* 117: 816–32. [7, 9, 16, 21, 36]

Bauer, Michael, and Glenn Rudebusch. 2014. "The signaling channel for Federal Reserve bond purchases." *International Journal of Central Banking* 10 (3): 233–89. [21]

Bayer, Christian, Chi Kim, and Alexander Kriwoluzky. 2018. "The Term Structure of Redenomination Risk." *DIW Discussion Paper No. 1740, Deutsches Institut für Wirtschaftsforschung.* [16]

Bergant, Katharina, Michael Fidora, and Martin Schmitz. 2018. "International capital flows at the security level-evidence from the ECB's asset purchase programme." *ECMI Working Paper No. 7 October 2018, European Capital Markets Institute.* [8, 27]

Brand, Claus, Daniel Buncic, and Jarkko Turunen. 2010. "The impact of ECB monetary policy decisons and communication on the yield curve." *Journal of the European Economic Association* 8 (6): 1266–98. [7, 21]

Bubeck, Johannes, Maurizio Michael Habib, and Simone Manganelli. 2018. "The portfolio of euro area fund investors and ECB monetary policy announcements." *Journal of International Money and Finance* 89: 103–26. [8, 31]

Burriel, Pablo, and Alessandro Galesi. 2018. "Uncovering the heterogeneous effects of ECB unconventional monetary policies across euro area countries." *European Economic Review* 101: 210–29. [8]

Cieslak, Anna, and Andreas Schrimpf. 2019. "Non-monetary news in central bank communication." *Journal of International Economics* 118: 293–315. [7]

Coeuré, Benoît. 2018. "The persistence and signalling power of central bank asset purchase programmes." *Speech at the 2018 US Monetary Policy Forum, New York City.* [6]

Cragg, John, and Stephen Donald. 1997. "Inferring the rank of a matrix." *Journal of Econometrics* 76 (1-2): 223–50. [19]

D'Amico, Stefania, and Thomas King. 2013. "Flow and stock effects of large-scale treasury purchases: Evidence on the importance of local supply." *Journal of Financial Economics* 108 (2): 425–48. [8]

De Santis, Roberto. 2020. "Impact of the asset purchase programme on euro area government bond yields using market news." *Economic Modelling* 86: 192–209. [8, 30, 31]

Dedola, Luca, Georgios Georgiadis, Johannes Gräb, and Arnaud Mehl. 2021. "Does a big bazooka matter? Quantitative easing policies and exchange rates." *Journal of Monetary Economics* 117: 489–506. [31, 33]

Dell'Ariccia, Giovanni, Pau Rabanal, and Damiano Sandri. 2018. "Unconventional monetary policies in the Euro Area, Japan, and the United Kingdom." *Journal of Economic Perspectives* 32 (4): 147–72. [30]

Droste, Michael, Yuriy Gorodnichenko, and Walker Ray. 2021. "Unbundling Quantitative Easing: Taking a Cue from Treasury Auctions." *Working Paper, University of California Berkeley.* [8]

Duffie, Darrell. 1999. "Credit Swap Valuation." *Financial Analysts Journal* 55 (1): 73–87. [11]

Fontana, Alessandro, and Martin Scheicher. 2016. "An analysis of euro area sovereign CDS and their relation with government bonds." *Journal of Banking and Finance* 62: 126–40. [12]

Fratzscher, Marcel, Marco Lo Duca, and Roland Straub. 2016. "ECB Unconventional Monetary Policy Actions: Market Impact, international Spillovers and Transmission Channels." *IMF Economic Review* 64 (1): 36–74. [30]

Gambetti, Luca, and Alberto Musso. 2020. "The effects of the ECB's expanded asset purchase programme." *European Economic Review* 130. [9]

Georgiadis, Georgios, and Johannes Gräb. 2016. "Global financial market impact of the announcement of the ECB's asset purchase programme." *Journal of Financial Stability* 26: 257–65. [8, 31]

Glick, Reuven, and Sylvain Leduc. 2018. "Unconventional monetary policy and the dollar: conventional signs, unconventional magnitudes." *International Journal of Central Banking* 14: 103–52. [28]

Greenlaw, David, James D Hamilton, Ethan Harris, and Kenneth West. 2018. "A Skeptical View of the Impact of the Fed's Balance Sheet." *NBER Working Paper No. 24687, National Bureau of Economic Research.* [36]

Gürkaynak, Refet S, Brian Sack, and Eric T. Swanson. 2005a. "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements." *International Journal of Central Banking* 1: 55–93. [7, 13, 21]

Gürkaynak, Refet S, Brian Sack, and Eric T. Swanson. 2005b. "The Sensitivity of Long-Term Interest Rates to Economic News : Evidence and Implications for Macroeconomic Models." *American Economic Review* 95 (1): 425–36. [37]

Hachula, Michael, Michele Piffer, and Malte Rieth. 2020. "Unconventional monetary policy, fiscal side effects, and euro area (im) balances." *Journal of the European Economic Association* 18 (1): 202–31. [8, 9]

Hammermann, Felix, Kieran Leonard, Stefano Nardelli, and Julian von Landesberger. 2019. "Taking stock of the Eurosystem's asset purchase programme after the end of net asset purchases." *ECB Economic Bulletin Articles, European Central Bank.* [7]

Jarociński, Marek. 2021. "Central Bank Information Effects and Transatlantic Spillovers." *ECB Working Paper No. 2482, European Central Bank.* [15, 16, 19, 35, 36]

Jarociński, Marek, and Peter Karadi. 2020. "Deconstructing monetary policy surprises—the role of information shocks." *American Economic Journal: Macroeconomics* 12 (2): 1–43. [4, 7, 9, 15, 16, 18, 19, 35]

Jordà, Òscar. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95 (1): 161–82. [36]

Kerssenfischer, Mark. 2019. "Euro Area Monetary Policy Effects : New Evidence from High-Frequency Futures Data." *Deutsche Bundesbank Discussion Paper No 07/2019, Bundesbank.* [7, 16]

Koijen, Ralph, François Koulischer, Benoit Nguyen, and Motohiro Yogo. 2021. "Inspecting the mechanism of quantitative easing in the euro area." *Journal of Financial Economics* 140 (1): 1–20. [8]

Krishnamurthy, Arvind, Stefan Nagel, and Annette Vissing-Jorgensen. 2018. "ECB Policies Involving Government Bond Purchases: Impact and Channels." *Review of Finance* 22 (1): 1–44. [8, 10]

Krishnamurthy, Arvind, and Annette Vissing-Jorgensen. 2011. "The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy." *Brookings Papers on Economic Activity* 2011 (2): 215–87. [4, 8, 11]

Krishnamurthy, Arvind, and Annette Vissing-Jorgensen. 2012. "The Aggregate Demand for Treasury Debt." *Journal of Political Economy* 120 (2): 233–67. [11]

Kroencke, Tim A, Maik Schmeling, and Andreas Schrimpf. 2021. "The FOMC risk shift." *Journal of Monetary Economics* 120: 21–39. [7, 42]

Lewis, Daniel. 2019. "Announcement-Specific Decompositions of Unconventional Monetary Policy Shocks and Their Macroeconomic Effects." *Working Paper, Federal Reserve Bank of New York.* [7]

Liu, Ernest, Atif Mian, and Amir Sufi. 2022. "Low Interest Rates, Market Power, and Productivity Growth." *Econometrica* 90 (1): 193–221. [29]

Lloyd, Simon P. 2021. "Overnight indexed swap-implied interest rate expectations." *Finance Research Letters* 38: 101430. [10]

Nagel, Stefan. 2016. "The liquidity premium of near-money assets." *Quarterly Journal of Economics* 131 (4): 1927–71. [11]

Nakamura, Emi, and Jon Steinsson. 2018. "High Frequency Identification of Monetary Non-Neutrality: The Information Effect." *Quarterly Journal of Economics* 133 (3): 1283–330. [7, 13]

Ramey, Valerie. 2016. "Macroeconomic Shocks and Their Propagation." *Handbook of Macroeconomics* 2: 71–162. [9]

Rogers, John H, Chiara Scotti, and Jonathan H Wright. 2014. "Evaluating asset-market effects of unconventional monetary policy: a multi-country review." *Economic Policy* 29 (80): 749–99. [30, 36]

Schlepper, Kathi, Heiko Hofer, Ryan Riordan, and Andreas Schrimpf. 2018. "The Market Microstructure of Central Bank Bond Purchases." *Journal of Financial and Quantitative Analysis,* 1–65. [8]

Swanson, Eric T. 2021. "Measuring the effects of federal reserve forward guidance and asset purchases on financial markets." *Journal of Monetary Economics* 118: 32–53. [7, 17, 20]

Swanson, Eric T., and John C. Williams. 2014a. "Measuring the Effect of the Zero Lower Bound on Medium-and Longer-Term Interest Rates." *American Economic Review* 104 (10): 3154–85. [37]

Swanson, Eric T., and John C. Williams. 2014b. "Measuring the effect of the zero lower bound on yields and exchange rates in the U.K. and Germany." *Journal of International Economics* 92. [37]

Wieladek, Tomasz, and Antonio Garcia Pascual. 2016. "The European Central Bank's QE: A New Hope." *CESifo Working Paper No. 5946, Center for Economic Studies and Ifo Institute (CESifo) Munich.* [8]

Wright, Jonathan H. 2012. "What does monetary policy do to long-term interest rates at the zero lower bound?" *Economic Journal* 122 (564): 447–66. [36]

Zhang, Donghai. 2019. "Term Structure, Forecast Revision and the Information Channel of Monetary Policy." *Working Paper, University of Bonn.* [7]

Chapter 2

Monetary Policy, Firm Heterogeneity, and the Distribution of Investment Rates*

Joint with Donghai Zhang

2.1 Introduction

Understanding the investment channel of monetary policy is important for policymakers because investment is a sizable and the most volatile component of aggregate GDP. To this end, the literature has extensively studied the effect of monetary policy on the *average* investment rate.[1] However, this estimated effect on the average investment rate can reflect a *parallel shifting* of the entire distribution or a change in the *shape* of the distribution. How does monetary policy affect the distribution of investment rates? Which part of the distribution is responsible for the change in the average investment rate? Moreover, a growing academic literature studies the heterogeneous effects of monetary policy on the investment behavior of different groups of firms, see, e.g., Gertler and Gilchrist (1994), Ottonello and Winberry (2020), Jeenas (2019), and Cloyne et al. (2020).[2] Which part of the distribution drives these heterogeneous effects on *average* investment rates? The answers

* We thank Isaac Baley, Christian Bayer, Yongsung Chang, Russell Cooper, Maren Froemel, Joachim Jungherr, Keith Kuester, Matthias Meier, Emi Nakamura, Haozhou Tang, Petr Sedláček, Vincent Sterk as well as seminar/workshop participants at the EEA and Peking University for insightful discussions. We thank Lixing Wang for his research assistance in constructing Section 3 of the paper.

1. See, for example, Ottonello and Winberry (2020), Jeenas (2019), or Cloyne et al. (2020).

2. Cloyne et al. (2020) document that investment rates of young firms are on average more sensitive to monetary policy than those of old firms. Gertler and Gilchrist (1994) show a similar result for small and large firms. Clearly, these findings are connected, as age and size are strongly correlated in the data. In this paper, we focus on age but emphasize and show that our results are similar when comparing small and large firms.

to these questions are important to understand the transmission of monetary policy. In particular, they are informative about the frictions that matter for the (heterogeneous) effects of monetary policy on firm investment decisions.

We provide three pieces of evidence that address the raised questions. First, monetary policy affects the *shape* of the distribution of investment rates. Specifically, an expansionary monetary policy shock leads to fewer small or zero investment rates and more large investment rates. Second, the change in the shape of the investment rate distribution is more pronounced among young (small) firms than among old (large) firms. This paper emphasizes the relevance of the extensive margin—firms deciding whether to invest or not—in explaining these findings. Third, a decomposition exercise indicates that the extensive margin accounts for around 50% of the effect of monetary policy on the *average* investment rate and for more than 50% of the *heterogeneous effect* on firms of different age groups.

Theoretically, we provide a heterogeneous-firm model that combines capital adjustment costs, firm entry and exit, and nominal rigidities. The presence of fixed adjustment costs gives rise to an investment channel of monetary policy along the extensive margin. That is, an interest rate cut induces some firms to switch from not investing to making a large investment. Therefore, monetary policy reshapes the distribution of investment rates. In addition, the calibrated quantitative model generates sizable heterogeneous effects of monetary policy on the average investment rates of firms of different age groups. The reason is that young firms are more easily induced to make an investment than old firms. The model attributes more than 50% of the heterogeneous effect across age groups to the extensive margin, as in the data.

In more detail, we address the empirical questions using quarterly Compustat data combined with identified monetary policy shocks as in Ottonello and Winberry (2020) and Cloyne et al. (2020). In contrast to the existing studies, we estimate the effects of monetary policy on different quantiles of the investment rate distribution rather than solely focusing on the first moment of the distribution—the average investment rate. We uncover that the upper quantiles respond substantially more to a monetary policy shock than the lower quantiles do. This finding suggests that monetary policy reshapes the distribution of investment rates by reducing the mass of firms making a small or no investment and increasing the mass of firms making a sizeable investment.

To visualize the change in the distribution, we approximate the shape of the entire distribution of investment rates before and after a monetary shock. We do so by fitting a flexible skewed t-distribution, using quantiles of the distribution and their estimated impulse response functions as inputs.[3] Comparing the approximated skewed t-distributions before and after an expansionary monetary policy shock, we

3. This approach has recently been applied by Adrian, Boyarchenko, and Giannone (2019) to transform conditional quantiles into the conditional distribution of GDP growth.

illustrate that fewer firms make a small or no investment and more firms make a large investment—**Fact 1**. This novel evidence suggests the presence of a quantitatively relevant investment channel of monetary policy along the extensive margin.

Conducting the same exercise for young and old firms separately, we uncover that the effect of monetary policy on the shape of the distribution of investment rates is more pronounced among young firms than among old firms—**Fact 2**. This result implies that the extensive margin investment channel is particularly important for young firms. The following exercises further support this conclusion. We estimate the effects of monetary policy on the *spike rate*, defined as the fraction of firms whose quarterly investment rate exceeds 10%, and on the *inaction rate*, defined as the fraction of firms whose quarterly investment rate is smaller than 0.5% in absolute value.[4] The spike rate rises and the inaction rate drops more strongly for young firms than for old firms. Both differences are statistically significant.

The empirical literature has documented that young firms' *average* investment rates are more sensitive to monetary policy than old firms', see, e.g., Cloyne et al. (2020). Gertler and Gilchrist (1994) show a similar result for small and large firms. Conventional wisdom views these findings as supporting the financial accelerator mechanism, based on the narrative that young firms are financially constrained[5] and monetary policy affects financial conditions. The novel empirical evidence presented in this paper suggests that besides financial acceleration, the extensive margin investment decision, arising from fixed adjustment costs, is important for the heterogeneous sensitivity of young and old firms. The final empirical exercise quantifies the relative importance of the extensive margin.

We decompose the effect of monetary policy on the average investment rate into contributions arising from the extensive and intensive margin, respectively. We use the change in the spike rate to proxy for the extensive margin. Our decomposition suggests that the extensive margin accounts for 50% of the total effect of a monetary policy shock on the average investment rate. In addition, more than 50% of the heterogeneous sensitivity of young and old firms' average investment rates to monetary policy is due to the extensive margin—**Fact 3**.

4. In annual data, an investment spike is typically defined as an investment rate above 20%, so about twice the average investment rate, which, in most representative datasets, ranges between 10% and 12% (Zwick and Mahon, 2017). Since we do not use annual, but quarterly data and Compustat features higher average investment rates, as discussed in Appendix 2.B.3, we define an investment spike to be a quarterly investment rate exceeding 10%. This too is an investment rate roughly twice the average investment rate. Inaction is typically defined as an annual investment rate less than 1% in absolute value. For the same reasons as above, we define inaction as a quarterly investment rate smaller than 0.5% in absolute value.

5. Rauh (2006), Fee, Hadlock, and Pierce (2009), Hadlock and Pierce (2010), and more recently Cloyne et al. (2020) argue that young firms are more likely financially constrained than old firms. Gertler and Gilchrist (1994) rely on the narrative that "*...the costs of external finance apply mainly to younger firms, firms with a high degree of idiosyncratic risk, and firms that are not well collateralized. These are, on average, smaller firms...*" to motivate the use of firm size as a proxy for financial frictions.

The second part of the paper interprets the empirical findings through the lens of a quantitative model. Fixed capital adjustment costs give rise to an investment channel of monetary policy along the extensive margin, consistent with the empirical findings. As a result, monetary policy affects the distribution of investment rates.

Moreover, the model highlights that the presence of an *extensive margin* investment decision creates *heterogeneous* effects on *average* investment rates of young and old firms. Entry and exit give rise to endogenous firm life cycles and an age distribution. The age-group-specific average investment rate is the fraction of investing firms (hazard rate) times the investment rate conditional on investing. The heterogeneous effect on different age groups along the extensive margin arises from two channels. First, monetary policy has a heterogeneous effect on hazard rates. More specifically, an interest rate cut induces more young than old firms to switch from inaction to making an investment. The reason is that young firms are on average further away from their optimal level of capital and therefore more easily induced to make an investment. Second, even without a heterogeneous effect on hazard rates, we would observe a higher *average* sensitivity of young firms. This is because young firms have a higher investment rate conditional on adjusting, again, because they are on average further away from their optimal level of capital. Overall, the model predicts that monetary policy affects the distribution of investment rates due to the extensive margin investment decision and that these effects are more pronounced among young firms, in line with the empirical evidence.

After illustrating the mechanisms in a simple model, we quantify them in a general equilibrium heterogeneous-firm model calibrated to match moments of the cross-sectional investment rate distribution and firm life-cycle patterns. According to the quantitative model, young firms are almost twice as sensitive to a monetary policy shock as old firms, explaining a large portion of the observed heterogeneous sensitivity in the data. A decomposition exercise demonstrates that the extensive margin is quantitatively dominant.

Our findings have important implications for both academic research and the conduct of monetary policy. We present a mechanism that makes firms typically classified as financially constrained more sensitive to monetary policy *even in the absence* of financial acceleration.[6] Thus, there is an issue of observational equivalence: The observed heterogeneous sensitivity of young (small) firms can arise not only due to a financial accelerator mechanism but also due to the presence of fixed adjustment costs as explained above. In addition, to the extent that age is correlated with popular proxies of financial frictions, as documented in Cloyne et al. (2020), the issue

6. Even though the capital adjustment costs that we impose can in principle be interpreted as stand-ins for financial frictions, our model does not feature a financial accelerator mechanism. The idea of the financial accelerator mechanism is that monetary policy changes the tightness of financial constraints. By construction, the capital adjustment costs are not affected by aggregate shocks, including monetary policy shocks, and therefore, there is no financial accelerator mechanism.

of observational equivalence extends beyond the comparison of firms by age or size. However, one should not interpret our results as rejecting the financial accelerator mechanism. Likely, both financial frictions and the non-financial mechanisms that we emphasize in this paper are relevant for the heterogeneous sensitivity observed in the data. Our findings highlight that there remain unresolved challenges when it comes to identifying the financial accelerator mechanism in the data.

Understanding the frictions underlying firms' (heterogeneous) investment decisions is important for guiding macroeconomic policy in recessions. Financial conditions are typically tighter in recessions, which is associated with a stronger financial accelerator mechanism. Therefore, if financial frictions are more important for firms' decisions, one would expect monetary and fiscal policy to be more effective in recessions. On the contrary, if lumpy investment behavior is more important, macroeconomic policies are less effective in times of economic slack. This is because recessions are typically associated with fewer firms at the margin of adjusting (Winberry, 2021). Our paper supports the quantitative relevance of lumpy investment behavior and argues that the heterogeneous sensitivity of young firms is not sufficient evidence for financial acceleration. The relevance of the lumpy investment is consistent with separate evidence uncovered in the empirical literature: monetary and fiscal policy interventions are less potent in recessions (Tenreyro and Thwaites, 2016; Ramey and Zubairy, 2018).

Literature Review. The evidence presented in this paper contributes to the empirical literature that studies the investment channel of monetary policy: see, e.g., Christiano, Eichenbaum, and Evans (2005) using aggregate data and Gertler and Gilchrist (1994), Ottonello and Winberry (2020), Jeenas (2019), and Cloyne et al. (2020) using firm-level data. So far, this literature has focused on the effects on average investment rates or on aggregate investment. Our contribution is to show that monetary policy reshapes the distribution of investment rates and that this effect is more pronounced among young and small firms.

Our paper also contributes to the literature emphasizing the extensive margin of firm investment, i.e. the relevance of fixed adjustment costs. A long debate has focused on whether lumpy firm-level investment behavior matters for aggregate investment and in particular for its responsiveness to shocks over the business cycle. Important contributions include Caballero, Engel, and Haltiwanger (1995), Caballero and Engel (1999), Thomas (2002), Khan and Thomas (2003), Khan and Thomas (2008), Bachmann, Caballero, and Engel (2013), House (2014), Koby and Wolf (2020), Winberry (2021), and Baley and Blanco (2021). Monetary policy shocks in models with fixed adjustment costs have been analyzed in Reiter, Sveen, and Weinke (2013), Reiter, Sveen, and Weinke (2020), and Fang (2022). We develop a heterogeneous-firm model with three features: fixed adjustment costs, firm life-cycle dynamics, and a New Keynesian sticky-price setup. We contribute to the theoretical strand of this literature by demonstrating the importance of the extensive margin for the heterogeneous sensitivity of firm-level investment rates across firm groups.

The empirical strand of the literature on lumpy investment has mainly produced two types of evidence. First, the *unconditional* distribution of firm-level investment rates is in line with the presence of fixed adjustment costs, see, e.g., Caballero, Engel, and Haltiwanger (1995), Cooper, Haltiwanger, and Power (1999), and Cooper and Haltiwanger (2006). Second, the behavior of *aggregate* investment in response to macroeconomic shocks is in line with the presence of fixed adjustment costs, see, e.g. Caballero and Engel (1999), Bachmann, Caballero, and Engel (2013) and Fang (2022). We contribute to the empirical strand by documenting the response of the *entire distribution* of investment rates to monetary policy shocks. The evidence supports the quantitative relevance of the investment channel of monetary policy along the extensive margin.[7] Furthermore, we emphasize that the heterogeneous effects along the extensive margin across age groups are consistent with a model with fixed adjustment costs and endogenous firm life cycles.

The issue of observational equivalence that we raise in this paper contributes to the literature aiming to document the financial accelerator mechanism in firm-level data. Several recent papers compare the investment behavior of groups of "likely financially constrained" and "likely financially unconstrained" firms after monetary policy shocks. To group firms, some observable proxy variable, which plausibly correlates with the severity of financial constraints, is used. For example, Gertler and Gilchrist (1994) use size to group firms, Ottonello and Winberry (2020) use leverage and distance to default, and Jeenas (2019) uses liquidity. Cloyne et al. (2020) advocate the use of firm age as a proxy for financial constraints, because it correlates with most other proxy variables while being exogenous to firm decisions. A higher sensitivity to monetary policy shocks of "likely financially constrained" firms is taken as evidence supporting the financial accelerator mechanism, based on the logic that there is a heterogeneous effect on the marginal *cost* of investing. We show that two common proxies for financial constraints—firm age and firm size—predict a higher sensitivity to monetary policy shocks even in the absence of financial frictions. This is because fixed adjustment costs create a heterogeneous effect on the (marginal) *benefit* of investing. This finding does not speak against age and size being correlated with financial constraints. However, it illustrates that age and size, and therefore all proxy variables which correlate with them, also correlate with the

7. Gourio and Kashyap (2007) emphasize the cyclicality of the spike rate of firms' investments—another statistic in addition to the average investment rate. In contemporaneous work, Lee (2022) estimates the effect of monetary policy shocks on the spike rates of small and large firms. We investigate the effect on the entire distribution of investment rates in addition to spike rates. Lee (2022) uses the estimates by firm size to *calibrate* a real business cycle model with *size-dependent* fixed adjustment costs and aggregate TFP shocks. We *rationalize* our findings in a New Keynesian sticky-price model. Importantly, firm entry and exit give rise to endogenous firm life cycles in our setting that explain the heterogeneous effects along the extensive margin by firm age.

relevance of non-financial constraints which make firms sensitive to monetary policy. In this sense, our paper relates to Crouzet and Mehrotra (2020), who argue that large firms are less cyclical than small firms because they are better diversified across industries, but not because of financial frictions.[8]

The remainder of this paper is organized as follows. Section 2.2 presents our empirical findings. Section 2.3 outlines the simple model and explains its key mechanisms. Section 2.4 presents the full New Keynesian heterogeneous-firm model. Section 2.5 calibrates the model and analyzes the effects of a monetary policy shock. Section 2.6 concludes.

2.2 Empirical Evidence

We present three pieces of evidence that are important to understand the investment channel of monetary policy. Section 2.2.1 introduces the data used throughout this paper. Section 2.2.2 describes the local projection model used to estimate impulse response functions (IRFs). Section 2.2.3 documents the effects of monetary policy on the distribution of investment rates. Section 2.2.4 presents the heterogeneous effects of monetary policy by firm age. Section 2.2.5 decomposes the (heterogeneous) effects of monetary policy into contributions arising from the extensive and intensive margin, respectively.

2.2.1 Firm-Level Data

We use quarterly firm-level data from Compustat. Our sample begins with 1986Q1 and ends with 2018Q4. We exclude firms with incomplete or questionable information (e.g. negative reported sales) and those not suitable for our analysis (e.g. financial firms) from the sample. Details on the sample selection are relegated to Appendix 2.B.1. Since information on firm age in Compustat is scarce, we merge age information from WorldScope and Jay Ritter's database, as explained in Appendix 2.B.2.

Capital stocks reported in Compustat are *accounting* capital stocks and do not perfectly reflect *economic* capital stocks.[9] To address this issue, we use a Perpetual Inventory Method (PIM) to compute real economic capital stocks, building on Bach-

8. In addition, our argument aligns well with Farre-Mensa and Ljungqvist (2016) who show that supposedly financially constrained firms do not behave as if they were constrained and also differ systematically from supposedly less constrained firms along other important dimensions. Along the same lines, Dinlersoz et al. (2018) argue that only private firms, but not public ones—i.e. Compustat firms which are the ones examined in the above-mentioned papers—appear financially constrained.

9. On the one hand, accounting depreciation is driven by tax incentives and usually exceeds economic depreciation. On the other hand, accounting capital stocks are reported at historical prices, not current prices. With positive inflation, both issues make economic capital stocks exceed accounting capital stocks.

mann and Bayer (2014). Details of this procedure are explained in Appendix 2.B.3. Our baseline measure of the investment rate is $i_{jt} = \frac{CAPX_{jt} - SPPE_{jt}}{INVDEF_t \times k_{jt-1}}$, thus, real capital expenditures (CAPX) net of sales of capital (SPPE) divided by the lagged real economic capital stock (k). More details on the construction of variables are given in Appendix 2.B.4.

For parts of the subsequent analysis, we aggregate the firm-level data to quarterly investment rate distributions and moments thereof.[10] The distribution of investment rates, shown in Figure 2.1, depicts some well-known features of investment rate distributions. That is, the distribution has a positive skewness, a long right tail, substantial mass at 0, and very few negative observations.

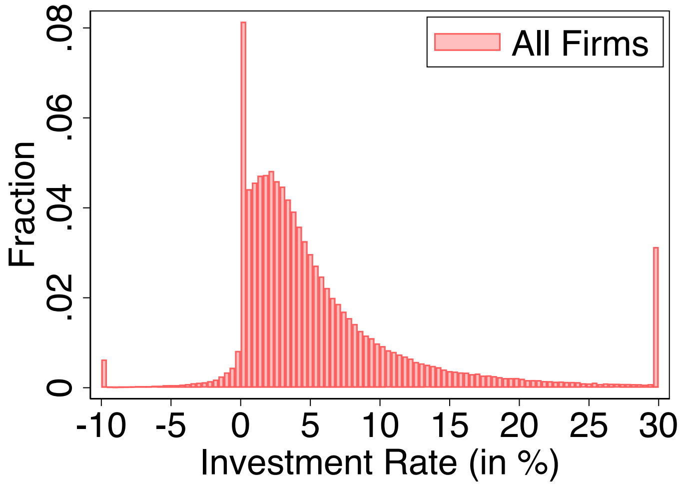


Figure 2.1. Distribution of Investment Rates

Notes: This figure plots the distribution of quarterly investment rates of firms in Compustat. The investment rate is real capital expenditures (CAPX) net of sales of capital (SPPE) divided by the lagged real economic capital stock. Sample: 1986Q1 - 2018Q4.

2.2.2 Local Projection: Method to Construct the IRFs

To estimate the effects of monetary policy shocks, we estimate the following simple local projection (LP) model:

$$y_{t+h} - y_{t-1} = \alpha^h + \beta^h \epsilon_t^{MP} + \sum_{j=2}^{4} \gamma^j \mathbb{1}\{q_{t+h} = j\} + e_{t+h} \qquad (2.1)$$

where y_t indicates the outcome variable, ϵ_t^{MP} is the monetary policy shock, q_t is the calendar quarter, γ^j are quarter dummies that are included to address seasonality,

10. Moments that are sensitive to outliers, such as the mean, are winsorized. Quite importantly, winsorizing is done by group and quarter. This ensures that the process does not systematically bias our sample.

and h denotes the horizon. We use the monetary policy shocks implied by the Proxy SVAR in Gertler and Karadi (2015). These are extracted after updating the time series data used in the VAR as well as the high-frequency instruments. Details are relegated to Appendix 2.B.5. The shocks are scaled to reduce the 1-year Treasury rate on impact by 25 basis points. Throughout, we use Newey-West standard errors to account for heteroskedasticity and autocorrelation.

Before turning to our novel findings, we verify that the monetary policy shocks have plausible effects on aggregate variables. We show in Figure 2.B.1 that an expansionary shock leads to hump-shaped increases in both investment and real GDP. The peak effects are 1.4% (investment) and 0.35% (real GDP), respectively.

2.2.3 Fact 1: Shape of the Distribution of Investment Rates

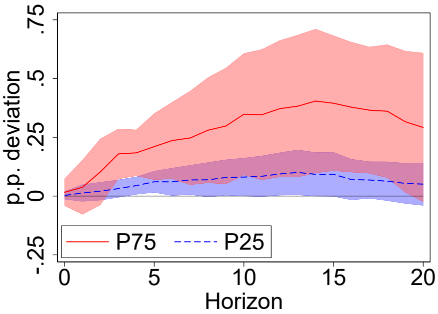
The literature has extensively studied the effect of monetary policy on the *average* investment rate.[11] However, this estimated effect on the average investment rate can reflect a *parallel shifting* of the entire distribution or a change in the *shape* of the distribution. We now investigate how monetary policy affects the distribution of firm-level investment rates by estimating the effects on different quantiles of the investment rate distribution. This is done by using the time series of the respective quantiles of the distribution as outcome variables in the empirical model (2.1). If the increase in the average investment rate reflects a parallel shifting of the distribution, the effect on all quantiles must be identical.

Figure 2.2 shows the effect of monetary policy shocks on selected quantiles of the investment rate distribution. Panel (a) plots the responses of the 25th (in blue) and the 75th (in red) percentiles. It is evident the right tail (the 75th percentile) responds more strongly than the left tail (the 25th percentile) of the investment rate distribution. At the peak, the 75th percentile of the investment rate distribution rises by 40 basis points, while the 25th percentile rises by only 10 basis points. This difference is statistically significant, as illustrated by the IRF of the corresponding interquantile range (Panel b). These findings are robust to alternative choices of quantiles: see Panels (c) to (f).[12] The disproportionate change in the right tail compared to the left tail indicates that monetary policy affects the shape of the investment rate distribution. The following exercise formalizes the mapping between the heterogeneous effect on different quantiles and changes in the shape of the investment rate distribution.

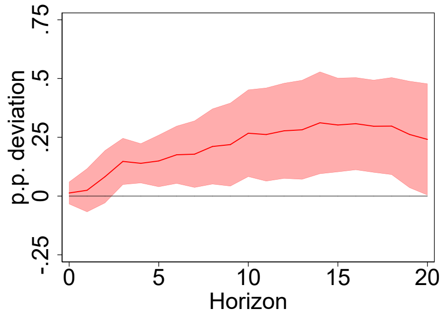
From Quantiles to Distribution. To visualize the effect of monetary policy on the distribution of investment rates, we use parametric approximations of the investment rate distribution. The approximations use quantiles of the distribution and

11. We show the effect of monetary policy on the average investment rate in Panel (a) of Figure 2.7.

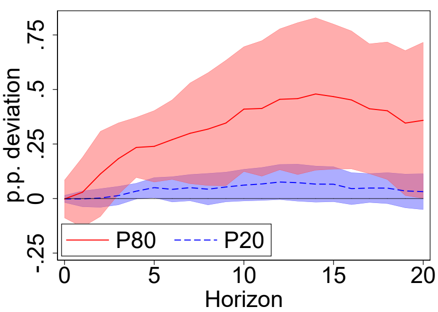
12. The impulse response functions of additional quantiles are reported in Figure 2.A.1.



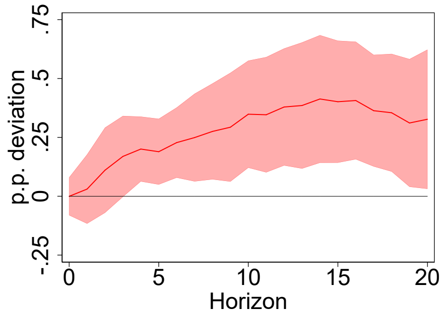
(a) Quantiles (75th & 25th)



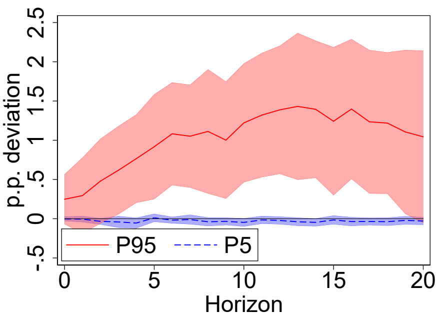
(b) Interquantile Range: P75 - P25



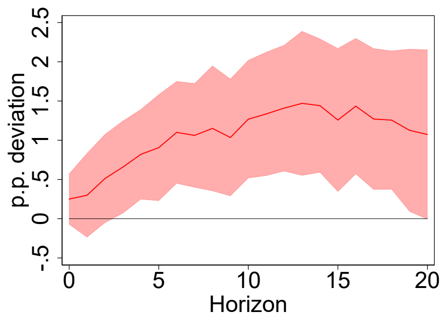
(c) Quantiles (80th & 20th)



(d) Interquantile Range: P80 - P20



(e) Quantiles (95th & 5th)



(f) Interquantile Range: P95 - P5

Figure 2.2. Effect of Monetary Policy Shock on Different Quantiles of Investment Rates

Notes: This figure plots the effect of a monetary policy shock on statistics of the investment rate distribution. The lines represent the estimated $\hat{\beta}^h$ from separate regressions: $y_{t+h} - y_{t-1} = \alpha^h + \beta^h \varepsilon_t^{MP} + \sum_{j=2}^{4} \gamma^j \mathbb{1}\{q_{t+h} = j\} + e_{t+h}$, using monetary policy shocks ε_t^{MP} implied by the Proxy SVAR in Gertler and Karadi (2015). The shocks are scaled to reduce the 1-year Treasury rate by 25 basis points. The shaded areas indicate the 90% confidence intervals constructed using standard errors that are robust to heteroskedasticity and autocorrelation. Sample: 1986Q1 - 2018Q4.

their impulse response functions as inputs. Since the investment rate distribution is skewed with a fat right tail, we choose a flexible skewed-*t* distribution (Azzalini and Capitanio, 2003) to match it.[13] This approach has recently been applied by

13. The skewed-*t* distribution for a variable *y* features the following density function:

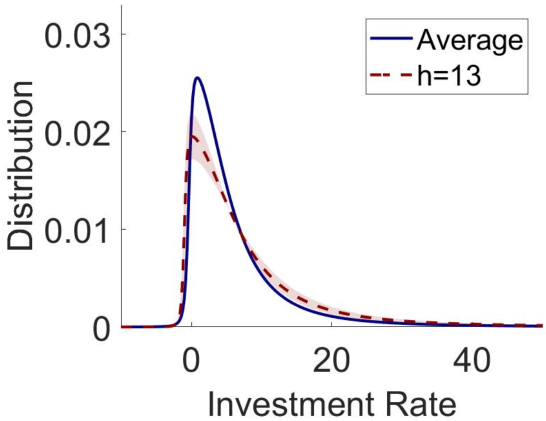


Figure 2.3. Effect of Monetary Policy Shock on Distribution of Investment Rates

Notes: This figure plots the approximated average distribution of investment rates (blue line) as well as the approximated distribution at horizon 13 (peak effect) after a monetary policy shock (red dashed line). The red shaded areas depict the 90% confidence bands constructed using the corresponding confidence bands of the responses of quantiles. The monetary policy shock is scaled by a factor of 10 to make differences in the distribution better visible.

Adrian, Boyarchenko, and Giannone (2019) to transform conditional quantiles into the conditional distribution of GDP growth.

To transform quantiles into a distribution, we estimate the parameters of the skewed-t distribution to match nine quantiles of the investment rate distribution.[14] The blue line in Figure 2.3 plots the average distribution using quantiles of the time-averaged distribution of firm-level investment rates. To fit the distribution of investment rates after a monetary policy shock, we repeat the estimation with the same quantiles but add the impulse responses for a given horizon h.

The red line in Figure 2.3 plots the distribution at the horizon at which the effect of the monetary policy shock peaks.[15] There is a clearly visible change in the distribution of investment rates after an expansionary monetary policy shock. In particular, there are fewer small investment rates and more large investment rates. This suggests that the *average* effect of monetary policy on firm investment rates is driven to a sizeable degree by the extensive margin, i.e., firms switch from making a small

$$f(y|\mu,\sigma,\alpha,\nu) = \frac{2}{\sigma} t\left(\frac{y-\mu}{\sigma};\nu\right) T\left(\alpha \frac{y-\mu}{\sigma}\sqrt{\frac{\nu+1}{\nu+\left(\frac{y-\mu}{\sigma}\right)^2}};\nu+1\right), \tag{2.2}$$

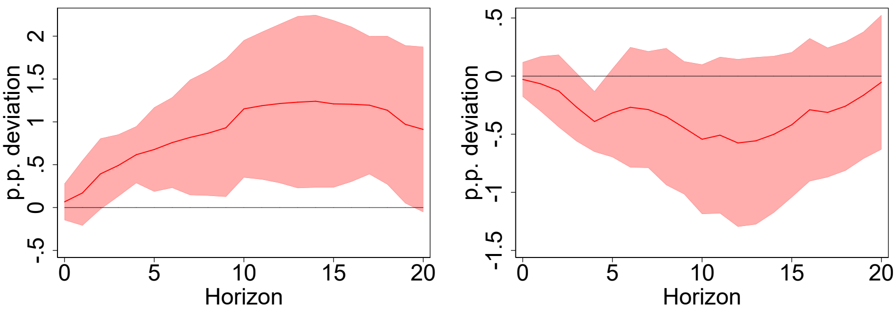
where t and T denote the density and cumulative distribution function of the t-distribution, respectively. μ determines the location of the distribution, σ is a scale parameter, ν controls the fatness of the tails, and α governs the skewness as it controls how much the standard t-distribution is skewed.

14. Specifically, we match the 5th, 10th, 15th, 20th, 25th, 35th, 50th, 75th, and 95th percentiles. Our findings are robust to alternative choices of quantiles.

15. Horizon 13 is when the peak effect on the average investment rate is reached.

or no investment to making a large investment. This aligns well with the evidence that *unconditional* fluctuations in aggregate investment are driven primarily by the extensive margin (Gourio and Kashyap, 2007).

Effect on the Spike and the Inaction Rate. To further investigate the hypothesis that the extensive margin is important for the effect of monetary policy on firm investment behavior, we look at two additional statistics of the investment rate distribution, namely, the *spike rate*, defined as the fraction of firms whose quarterly investment rate exceeds 10%, and the *inaction rate*, defined as the fraction of firms whose quarterly investment rate is smaller than 0.5% in absolute value.[16] Indeed, we find that following an expansionary monetary policy shock, the inaction rate falls and the spike rate rises substantially, as shown in Figure 2.4.



(a) Spike Rate

(b) Inaction Rate

Figure 2.4. Effect of Monetary Policy Shock on Spike and Inaction Rates

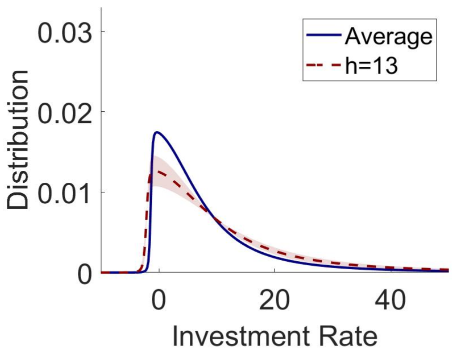
Notes: This figure plots the effect of a monetary policy shock on the spike rate and the inaction rate of all firms. A spike is an investment rate exceeding 10%, inaction is an investment rate less than 0.5% in absolute value. The lines represent the estimated $\hat{\beta}^h$ from separate regressions: $y_{t+h} - y_{t-1} = \alpha^h + \beta^h \varepsilon_t^{MP} + \sum_{j=2}^{4} \gamma^j \mathbb{1}\{q_{t+h} = j\} + e_{t+h}$, using monetary policy shocks ε_t^{MP} implied by the Proxy SVAR in Gertler and Karadi (2015). The shocks are scaled to reduce the 1-year Treasury rate by 25 basis points. The shaded areas indicate the 90% confidence intervals constructed using standard errors that are robust to heteroskedasticity and autocorrelation. Sample: 1986Q1 - 2018Q4.

2.2.4 Fact 2: Heterogeneous Effects across Age Groups

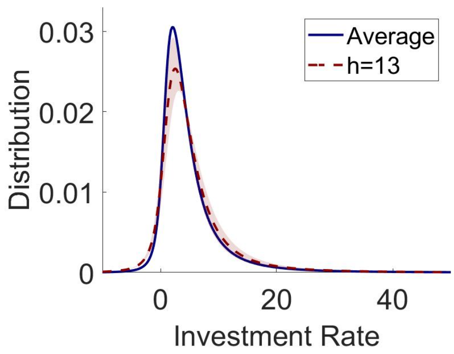
In addition to documenting the effect of monetary policy on the overall distribution of investment rates, we investigate the effect on *group-specific* investment rate distributions. In particular, we look at age-specific distributions. Cloyne et al. (2020) have documented that after an expansionary monetary shock, young firms increase their investment rates *on average* by much more than old firms. We replicate this finding in Figure 2.A.2. Yet, this difference in averages is only to a limited extent informative about the effect of monetary policy on the age-specific distributions.

16. The choice of cutoffs reflects that the investment data is quarterly and features relatively high average investment rates. See footnote 4 for more details.

Heterogeneous Effects on Quantiles of the Investment Rate Distributions. Figures 2.A.3 and 2.A.4 show that the disproportionate effects of monetary policy on the right tail of the investment rate distribution, documented for all firms in Figure 2.2, are present among both the group of young firms and the group of old firms. Quantitatively, these effects are much more pronounced among young firms, however. Using the IRFs of the quantiles, we visualize the effect of monetary policy on the age-specific investment rate distributions using parametric approximations.



(a) Young Firms



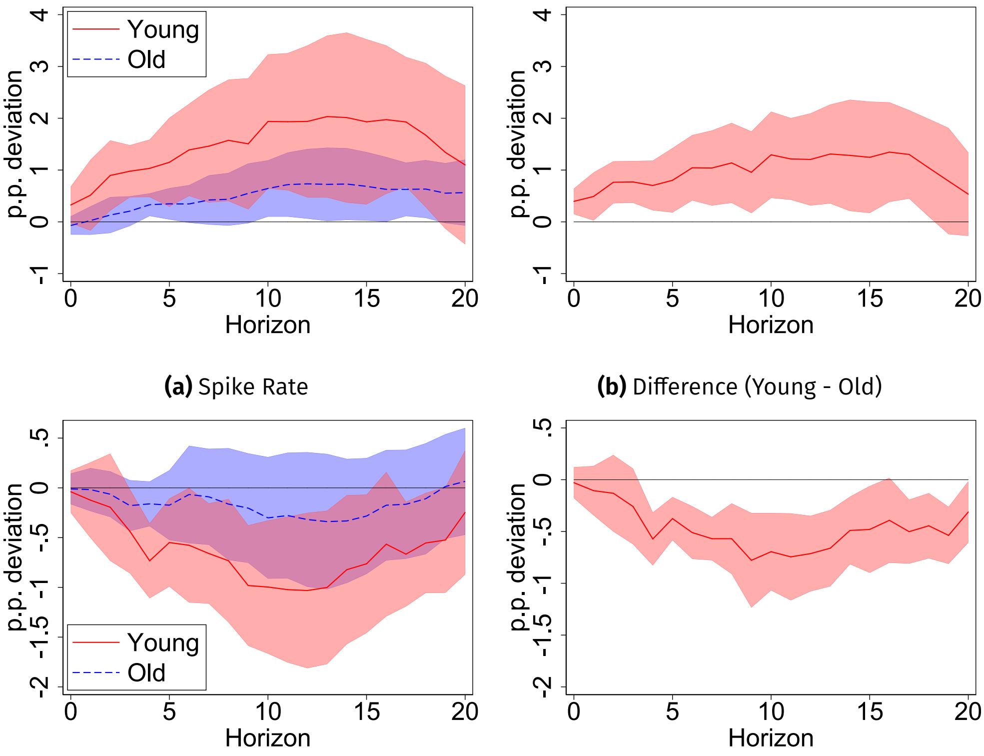
(b) Old Firms

Figure 2.5. Effect of Monetary Policy on Age-Specific Distributions of Investment Rates

Notes: This figure plots the approximated average distributions of investment rates (blue lines) as well as the approximated distributions at horizon 13 (peak effect) after a monetary policy shock (red dashed line) for young (panel a) and old (panel b) firms. Young (old) firms are less (more) than 15 years old. The red shaded areas depict the 90% confidence bands constructed using the corresponding confidence bands of the responses of quantiles. The monetary policy shock is scaled by a factor of 10 to make differences in the distribution better visible.

Heterogeneous Effects on Distributions of Investment Rates. Figure 2.5 compares the average distribution of investment rates of young (left panel) and old (right panel) firms with the distribution after a monetary policy shock. We find that the shape of the distribution changes more visibly for young firms. In particular, the decrease in small investment rates and increase in large investment rates is more pronounced. This suggests that the extensive margin is not only important for the *average* effect of monetary policy on investment rates, but also for the *heterogeneous* effect across age groups.

Heterogeneous Effects on Spike and Inaction Rates. To lend further support to the hypothesis that the extensive margin is important for the heterogeneous sensitivity of young and old firms, we look at two additional statistics of the investment rate distribution, namely, the *spike rate* and the *inaction rate*. Figure 2.6 shows that after a monetary policy shock the spike rate rises and the inaction rate drops more strongly for young firms. Both differences are statistically significant.



(a) Spike Rate

(b) Difference (Young - Old)

(c) Inaction Rate

(d) Difference (Young - Old)

Figure 2.6. Effect of Monetary Policy Shock on Age-Specific Spike & Inaction Rates

Notes: This figure plots the effect of a monetary policy shock on the spike rate and the inaction rate of young and old firms. Young (old) firms are less (more) than 15 years old. A spike is an investment rate exceeding 10%, inaction is an investment rate less than 0.5% in absolute value. The lines represent the estimated $\hat{\beta}^h$ from separate regressions: $y_{t+h} - y_{t-1} = \alpha^h + \beta^h \varepsilon_t^{MP} + \sum_{j=2}^{4} \gamma^j \mathbb{1}\{q_{t+h} = j\} + e_{t+h}$, using monetary policy shocks ε_t^{MP} implied by the Proxy SVAR in Gertler and Karadi (2015). The shocks are scaled to reduce the 1-year Treasury rate by 25 basis points. The shaded areas are the 90% confidence intervals constructed using standard errors that are robust to heteroskedasticity and autocorrelation. Sample: 1986Q1 - 2018Q4.

2.2.5 Fact 3: The Relative Importance of the Extensive Margin

Finally, we perform a simple decomposition exercise to gain some insights about the relative importance of the intensive and extensive margin. For this purpose, we classify investment rate observations into "spikes" ($i_{j,t} > 10\%$, as before) and "normal" investments ($i_{j,t} \leq 10\%$). It follows that the average (potentially group-specific) investment rate in period t is

$$\bar{i}_t = \psi_t^s i_t^s + (1 - \psi_t^s) i_t^n \tag{2.3}$$

where ψ_t^s is the fraction of firms undertaking a "spike" in period t and i_t^s and i_t^n are the average investment rates conditional on "spike" and "normal", respectively. Then, the effect of a monetary policy shock on the average investment rate can be decomposed as follows:[17]

$$\frac{\partial \mathbb{E}(\overline{i_t})}{\partial \epsilon^{MP}} \approx \underbrace{\frac{\partial \mathbb{E}(\psi_t^s)}{\partial \epsilon^{MP}} \left(\mathbb{E}(i_t^s) - \mathbb{E}(i_t^n)\right)}_{\text{Extensive Margin}} + \underbrace{\mathbb{E}(\psi_t^s) \frac{\partial \mathbb{E}(i_t^s)}{\partial \epsilon^{MP}} + \left(1 - \mathbb{E}(\psi_t^s)\right) \frac{\partial \mathbb{E}(i_t^n)}{\partial \epsilon^{MP}}}_{\text{Intensive Margin}} \tag{2.4}$$

Intuitively, the extensive margin reflects the change in the average investment rate that results *only* from changes in the spike rate, while the conditional investment rates are held fixed. Vice versa, the intensive margin reflects the change in the average investment rate that results *only* from changes in the conditional investment rates, while the spike rate is held fixed.

To implement this decomposition, we construct hypothetical average investment rates that would prevail if there were no changes in the extensive margin ($\overline{i_t}^{int}$) or the intensive margin ($\overline{i_t}^{ext}$):

$$\overline{i_t}^{int} = \overline{\psi^s} i_t^s + (1 - \overline{\psi^s}) i_t^n, \tag{2.5}$$

$$\overline{i_t}^{ext} = \psi_t^s \overline{i^s} + (1 - \psi_t^s) \overline{i^n}. \tag{2.6}$$

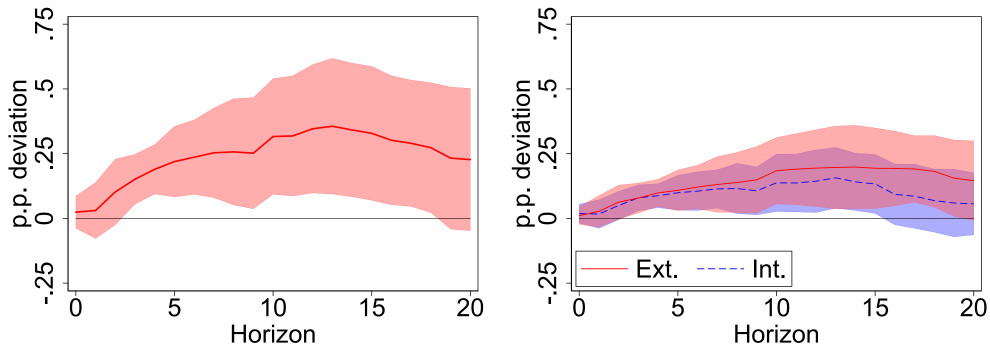
$\overline{i_t}^{int}$ captures fluctuations in the average investment rate arising only from the intensive margin, because the spike rate, $\overline{\psi^s}$, equals its average over time. Vice versa, $\overline{i_t}^{ext}$ captures fluctuations in the average investment rate arising only from the extensive margin, because the conditional investment rates, $\overline{i^n}$ and $\overline{i^s}$, equal their respective averages over time.

Decomposition of the Effect of a Monetary Policy Shock. According to Equation (2.4), the IRF of the average investment rate ($\overline{i_t}$) is approximately equal to the sum of the IRFs of the two hypothetical investment rates ($\overline{i_t}^{int}$ and $\overline{i_t}^{ext}$). Figure 2.7a plots the total effect of a monetary policy shock on the average investment rate and Figure 2.7b presents the decomposition. It is evident that both margins contribute about 50% to the effect of monetary policy on the *average* investment rate.

Decomposition of the Heterogeneous Effect of a Monetary Policy Shock. Figure 2.8a plots the estimated impulse response function of the difference between the average investment rates of young and old firms to an expansionary monetary policy shock ($\frac{\partial \mathbb{E}(\overline{i_{Y,t+h}} - \overline{i_{O,t+h}})}{\partial \epsilon_t^{MP}}$). The average investment rate of young firms responds more strongly to a monetary policy shock than that of old firms. This confirms the findings of Cloyne et al. (2020).

Figure 2.8b decomposes the heterogeneous effect into the contributions arising from the extensive margin ($\frac{\partial \mathbb{E}(\overline{i_{Y,t+h}}^{ext} - \overline{i_{O,t+h}}^{ext})}{\partial \epsilon_t^{MP}}$) and the intensive margin

17. This decomposition ignores two covariance terms ($Cov(\psi_t^s, i_t^s)$, $Cov(\psi_t^s, i_t^n)$), which can also be affected by the monetary shock. In the data, their contribution to the total effect on the average investment rate is very small, however.



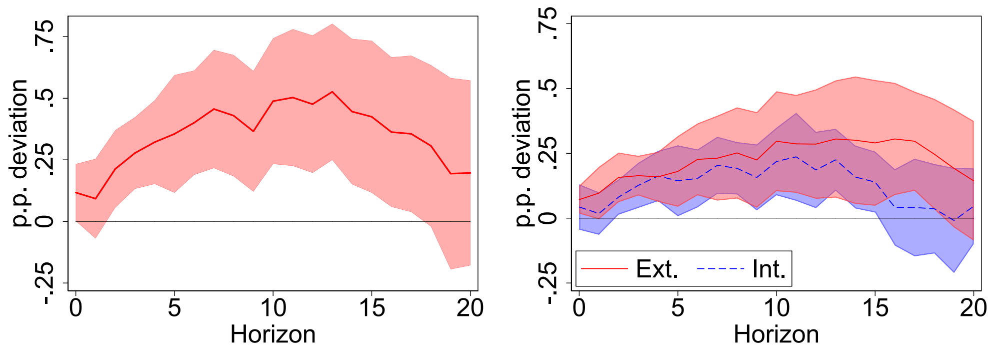
(a) Effect on Average Inv. Rate

(b) Decomposition: Ext. vs. Int. Margin

Figure 2.7. Decomposition of the Effect of Monetary Policy on the Average Inv. Rate

Notes: Panel (a) of this figure shows the effect of a monetary policy shock on the average investment rate ($\overline{i_t}$). Panel (b) decomposes this effect into an intensive ($\overline{i_t}^{int}$) and an extensive margin ($\overline{i_t}^{ext}$) contribution, using equation (2.4). The lines represent the estimated $\hat{\beta}^h$ from separate regressions: $y_{t+h} - y_{t-1} = \alpha^h + \beta^h \varepsilon_t^{MP} + \sum_{j=2}^{4} \gamma^j \mathbb{1}\{q_{t+h} = j\} + e_{t+h}$, using monetary policy shocks ε_t^{MP} implied by the Proxy SVAR in Gertler and Karadi (2015). The shocks are scaled to reduce the 1-year Treasury rate by 25 basis points. The shaded areas are the 90% confidence intervals constructed using standard errors that are robust to heteroskedasticity and autocorrelation. Sample: 1986Q1 - 2018Q4.

$(\frac{\partial \mathbb{E}(\overline{i_{Y,t+h}}^{int} - \overline{i_{O,t+h}}^{int})}{\partial \epsilon_t^{MP}})$. It shows that the extensive margin explains more than 50% of the heterogeneous sensitivity of young and old firms.



(a) Heterogeneous Effect (Young - Old)

(b) Decomposition: Ext. vs. Int. Margin

Figure 2.8. Decomposition of the Heterogeneous Effect of a Monetary Policy Shock

Notes: Panel (a) of this figure shows the heterogeneous effect of a monetary policy shock on the average investment rate of young firms as opposed to old firms. Panel (b) decomposes this heterogeneous effect into an intensive and an extensive margin contribution, using equation (2.4). Young (old) firms are less (more) than 15 years old. The lines represent the estimated $\hat{\beta}^h$ from separate regressions: $y_{t+h} - y_{t-1} = \alpha^h + \beta^h \varepsilon_t^{MP} + \sum_{j=2}^{4} \gamma^j \mathbb{1}\{q_{t+h} = j\} + e_{t+h}$, using monetary policy shocks ε_t^{MP} implied by the Proxy SVAR in Gertler and Karadi (2015). The shocks are scaled to reduce the 1-year Treasury rate by 25 basis points. The shaded areas are the 90% confidence intervals constructed using standard errors that are robust to heteroskedasticity and autocorrelation. Sample: 1986Q1 - 2018Q4.

Summary of the Empirical Evidence. We have documented three empirical findings. First, monetary policy reshapes the distribution of investment rates. Specifically, an interest rate cut leads to fewer small or zero investment rates and more large investment rates. Second, the change in the distribution is more pronounced among young firms than among old firms. Third, the extensive margin accounts for around 50% of the effect of monetary policy on the *average* investment rate and for more than 50% of the heterogeneous effect on young and old firms.

Appendix 2.D shows that similar findings emerge when we compare small and large firms, instead of young and old ones.

The second part of the paper presents a theoretical model to interpret these empirical findings.

2.3 A Simple Model

In Section 2.4, we build a heterogeneous-firm life-cycle model with capital adjustment costs and nominal rigidities. The purpose is to explain the observed effects of interest rate changes on the distribution of investment rates and why these effects are stronger among young firms. In the current section, we illustrate the mechanisms at work through the lens of a simple two-period model. Most importantly, the model features fixed capital adjustment costs which create an extensive margin investment decision.

In this simple model, we compare small and large firms. Since age and size are strongly correlated both in the data and in the quantitative model, all intuitions we provide in the simple model hold true when comparing young and old firms in the quantitative model. In Appendix 2.D, we compare the heterogeneous sensitivity by age and by size in the data and in the quantitative model.

The simple model consists of two periods. In period 0, firms are endowed with k_0 units of capital and choose the next period's capital k_1. The price of one unit of capital relative to the price of the consumption good is q. In period 1, firms transform capital into the consumption good (y) using the decreasing returns to scale production technology $y = k_1^\theta$ with $\theta < 1$. Sales are discounted at the real interest rate r, and capital depreciates fully during production.

In the absence of adjustment costs, the firms' profit-maximization problem is

$$\max_{k_1} \frac{1}{1+r} k_1^\theta - q(k_1 - k_0). \tag{2.7}$$

From the first-order condition for k_1, we obtain the optimal amount of capital that the firm chooses for period 1

$$k_1^* = \left(\frac{\theta}{(1+r)q} \right)^{\frac{1}{1-\theta}} \tag{2.8}$$

and the optimal (gross) investment rate as a function of firm size $i^*(k_0) = \frac{k_1^*}{k_0}$.

We now introduce some features from the quantitative model. First, there is a unit mass of firms within each size category k_0 and firms are indexed by j. Second, adjusting the stock of capital is subject to a fixed adjustment cost $\xi_j \in [0, \overline{\xi}]$, which is drawn from a uniform distribution. Moreover, we assume that the economy is populated by firms whose initial capital stocks are below the desired level, i.e., $k_{j,0} < k_1^*, \ \forall k_0$.[18]

The optimization problem of a firm j with an initial stock of capital k_0 has changed to:

$$\max_{k_{1,j}} \frac{1}{1+r} k_{1,j}^{\theta} - q(k_{1,j} - k_0) - \xi_j \mathbb{1}\{k_{1,j} \neq k_0\}, \tag{2.9}$$

where $\mathbb{1}\{k_{1,j} \neq k_0\}$ is an indicator variable that equals 1 if $k_{1,j} \neq k_0$ and 0 otherwise. To solve this problem, let $VA(k_0)$ denote the value added of adjusting capital while ignoring the fixed adjustment cost:

$$VA(k_0) = \frac{1}{1+r} k_1^{*\theta} - q(k_1^* - k_0) - \frac{1}{1+r} k_0^{\theta}, \tag{2.10}$$

where k_1^* is the optimal amount of capital that firms will acquire conditional on adjusting as defined by equation (2.8).

Considering the adjustment cost, a firm j adjusts capital if and only if the value added exceeds the costs, i.e., $VA(k_0) > \xi_j$. The threshold value of ξ_j, which makes a firm indifferent between adjusting or not, is defined by $\xi^T(k_0) \equiv VA(k_0)$. This implies a cutoff rule, i.e., a firm j will adjust its capital stock if and only if $\xi_j < \xi^T(k_0)$. From equation (2.10), it is evident that this cutoff value not only depends on the initial size of the firm but also on the interest rate r and the other parameters of the model.

The average investment rate among firms of size k_0 is:

$$\bar{i}(k_0) = \lambda(k_0) \times i^*(k_0) \tag{2.11}$$

where $\lambda(k_0) = \frac{\xi^T(k_0)}{\overline{\xi}} \in [0, 1]$ denotes the share of firms of size k_0 that choose to invest, i.e. the hazard rate. Conditional on investing, firms choose the optimal investment rate $i^*(k_0)$ as defined above.

The group-specific interest rate sensitivity of the investment rate is:

$$\frac{\partial \bar{i}(k_0)}{\partial r} = \underbrace{\frac{\partial \lambda(k_0)}{\partial r} i^*(k_0)}_{\text{Extensive Margin}} + \underbrace{\lambda(k_0) \frac{i^*(k_0)}{\partial r}}_{\text{Intensive Margin}}, \tag{2.12}$$

which features two components. There is an intensive margin effect, $\lambda(k_0)\frac{i^*(k_0)}{\partial r}$, because firms that would be adjusting anyways choose a different investment rate.

18. In the steady state of the quantitative model, there are also some firms with capital stocks above their desired level. However, quantitatively, these firms play a minor role.

Moreover, there is an extensive margin effect, $\frac{\partial\lambda(k_0)}{\partial r}i^*(k_0)$, because more or fewer firms choose to invest at all. Motivated by our empirical findings, this paper emphasizes the extensive margin effect.

Proposition 2.1 provides the main theoretical findings of this paper, which regard the effect of interest rate changes on the hazard rate ($\frac{\partial\lambda(k_0)}{\partial r}$) as well as how the sensitivity of the average investment rate due to the extensive margin changes with firm size.

Proposition 2.1. *In an economy populated by heterogeneous firms that face fixed adjustment costs as described above, it holds that*

(1) An interest rate cut increases the hazard rate: $\frac{\partial\lambda(k_0)}{\partial r} < 0$

(2) The sensitivity of the average investment rate to interest rate changes via the extensive margin is decreasing (in absolute terms) in firm size: $\frac{\partial\left(\frac{\partial\lambda(k_0)}{\partial r}i^*(k_0)\right)}{\partial k_0} > 0$

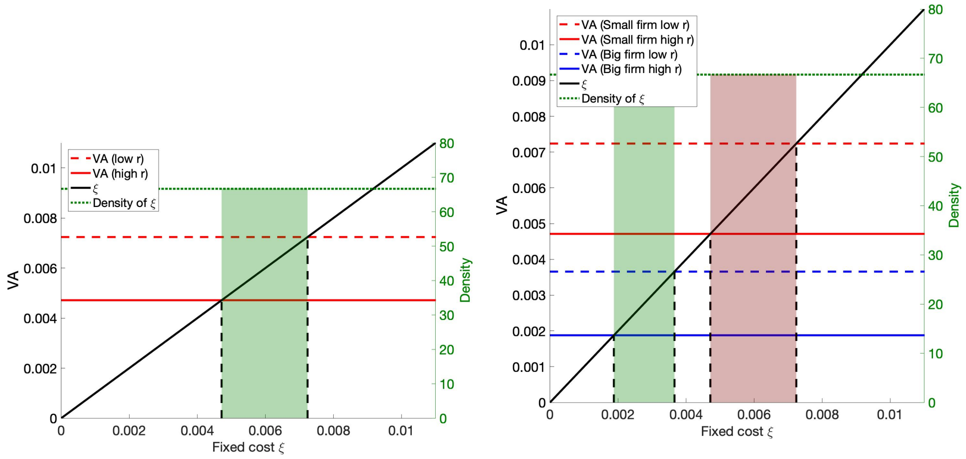
Proof. See Appendix 2.C. □

The first part of Proposition 2.1 establishes that an interest rate cut increases the hazard rate in line with the empirical evidence shown in Figure 2.6. The costs of investing (cost of additional capital, adjustment cost) are paid in period 0, whereas the benefits materialize in period 1. When the interest rate falls, the discounted benefit of investing rises. Hence, the value added of adjusting and thus the hazard rate rise.[19]

Figure 2.9a provides visual intuition by plotting the value added for a given k_0, $VA(k_0)$, against the random fixed cost ξ. The black upward-sloping line is the 45° line indicating the points where VA equals ξ. The intercept of the two curves pins down the cutoff value ξ^T. The green dotted line plots the density function of ξ (uniform distribution). The area under the density function to the left of the cutoff value ξ^T is the mass of adjusting firms. An interest rate cut shifts the VA curve upwards. As a result, the cutoff value ξ^T increases and so does the mass of adjusting firms as indicated by the green shaded area.

The second part of Proposition 2.1 establishes that the effect of an interest rate cut on the group-specific average investment rate via the extensive margin is larger among small firms. To understand this result, it is useful to compare the extensive margin effect for groups of small (S) and large (L) firms:

19. In the quantitative model, there are of course additional effects, but the main intuition – an interest rate cut raising the value added of investing – remains the same.



(a) r-Sensitivity of Hazard Rate **(b) r-Sensitivity by Firm Size**

Figure 2.9. Intuition for Proposition 2.1

Notes: This figure plots the value added of investing (VA) of a firm against the random fixed cost ξ. The black upward-sloping line is the 45° line indicating the points where VA equals ξ. The intercept of the two curves pins down the threshold value of ξ^T. The green dotted line plots the density function of ξ (uniform distribution). The area under the density function to the left of the threshold value ξ^T is the hazard rate. The shaded area in Panel (a) plots the difference in the hazard rate after an interest rate change. Panel (b) plots the difference in the hazard rate for a small and a big firm.

$$
\begin{aligned}
HetExt_{S-L} &= \underbrace{\frac{\partial \lambda(k_{0,S})}{\partial r} i^*(k_{0,S})}_{\text{Small Firms}} - \underbrace{\frac{\partial \lambda(k_{0,L})}{\partial r} i^*(k_{0,L})}_{\text{Large Firms}} \\
&= \underbrace{\frac{\partial \lambda(k_{0,S})}{\partial r}\left(i^*(k_{0,S}) - i^*(k_{0,L})\right)}_{\text{Heterogeneous Size Effect}} + \underbrace{\left(\frac{\partial \lambda(k_{0,S})}{\partial r} - \frac{\partial \lambda(k_{0,L})}{\partial r}\right) i^*(k_{0,L})}_{\text{Heterogeneous Hazard Rate Increase}}
\end{aligned}
\tag{2.13}
$$

This decomposition shows that there are two mechanisms. First, there is the *heterogeneous size effect*, due to which even if an interest rate cut had the same effect on hazard rates of small and large firms, there would be a differential effect on average investment rates. This is because among the *new* adjusters, small firms have higher investment rates conditional on adjusting ($i^*(k_{0,S}) - i^*(k_{0,L}) > 0$). This follows from the observation that in this simple model, conditional on investing, all firms choose k_1^* and the investment rate is defined by $i^* = \frac{k_1^*}{k_0}$. In the absence of an extensive margin investment decision, this effect would disappear because $\frac{\partial \lambda(k_0)}{\partial r} = 0$.

Second, interestingly, an interest rate cut increases the hazard rate of small firms by more than the hazard rate of large firms. This result aligns well with the empirical evidence that the spike rate of small (young) firms reacts more strongly to a monetary shock than the spike rate of large (old) firms (see Figure 2.D.3 for size and Figure 2.6 for age). As discussed above, the hazard rate rises, because the value added of investing rises, which happens because the discounted benefit of investing rises. This increase in the discounted benefit of investing is larger for small firms. The reason for this is that small firms have a higher marginal product of capital because of decreasing returns to scale. Hence, the interest rate cut has a larger effect on the hazard rate of small firms.

Figure 2.9b provides visual intuition for the heterogeneous effect of an interest rate cut on hazard rates. The cut in the interest rate shifts the VA of small firms (red lines) up by more than the VA of big firms (blue lines). As a result, the change in the hazard rate is more pronounced for small firms (red-shaded area) than for big firms (green-shaded area).

To sum up, we have highlighted two effects in this simple model. First, an interest rate cut increases the hazard rate, i.e. the fraction of firms deciding to make an investment. Therefore, a change in the interest rate changes the distribution of investment rates. Second, the average investment rate of small firms responds more strongly along the extensive margin to interest rate changes than the average investment rate of large firms.

Regarding the second effect, it is worth pointing out that small firms are more sensitive to interest rate changes *in the absence* of a financial accelerator mechanism. The basic idea of the financial accelerator mechanism is that interest rate changes affect financing conditions and small firms are more exposed to financing conditions than large firms. Then, interest rate changes have a heterogeneous effect on investment because there is a heterogeneous effect on the *cost* of investing, as e.g. in Ottonello and Winberry (2020). In contrast, in this model, there is a heterogeneous effect of interest rate changes on investment because of a heterogeneous effect on the *benefit* of investing.[20] This is because small firms have a higher marginal product of capital.

In the next section, we quantify the mechanisms highlighted in this section in a general equilibrium model.

2.4 A Quantitative Model

We build a New Keynesian model with heterogeneous firms subject to convex and fixed capital adjustment costs and entry and exit. These features have been studied separately; see, e.g., Khan and Thomas (2008), Clementi and Palazzo (2016), Ottonello and Winberry (2020), and Koby and Wolf (2020). The novelty of our model is to combine all these ingredients that are relevant for the understanding of the effect of monetary policy on age-specific investment rate distributions.

20. Even though the capital adjustment costs that we impose can in principle be interpreted as stand-ins for financial frictions, the model does not feature a financial accelerator mechanism. This is because, by construction, the capital adjustment costs are themselves not affected by aggregate shocks, including monetary policy shocks.

2.4.1 Investment Block

There exists a continuum of production firms[21] in the economy. Each firm j produces a quantity y_{jt} of the intermediate good using the production function

$$y_{jt} = z_{jt}k_{jt}^{\theta}n_{jt}^{\nu} \qquad \text{with } \theta, \nu > 0 \text{ and } \theta + \nu < 1 \tag{2.14}$$

where z_{jt} is total factor productivity (TFP), k_{jt} is the capital stock, and n_{jt} is the labor input. Productivity z_{jt} is subject to idiosyncratic shocks and follows an AR(1) process in logs

$$\log z_{jt} = \rho_z \log z_{jt-1} + \sigma_z \epsilon_{jt}^z \qquad \text{with } \epsilon_{jt}^z \sim \mathcal{N}(0,1) \tag{2.15}$$

Labor n_{jt} can be adjusted frictionlessly in every period. Capital k_{jt} is accumulated according to

$$k_{jt+1} = (1-\delta)k_{jt} + i_{jt} \tag{2.16}$$

where i_{jt} is investment and δ the depreciation rate. The relative price of capital (in terms of the final good) is q_t.

Following Bachmann, Caballero, and Engel (2013), we include *maintenance investment*. That is, a fraction χ of the depreciation δk_{jt} that occurs during the production process needs to be replaced immediately. At the end of the period, firms have $(1-\delta(1-\chi))k_{jt}$ units of capital and decide how much to invest voluntarily. To this voluntary investment, i_{jt}^{ν}, there are capital adjustment costs, which need to be paid if $i_{jt}^{\nu} \neq 0$.[22] Total adjustment costs consist of a random fixed adjustment cost $w_t\xi_{jt}$, where ξ_{jt} is distributed uniformly between 0 and $\bar{\xi}$, and a convex adjustment cost $\frac{\phi}{2}\frac{(i_{jt}^{\nu})^2}{k_{jt}}$:

$$AC(k_{jt}, k_{jt+1}, \xi_{jt}) = w_t\xi_{jt}\mathbb{1}\{k_{jt+1} \neq (1-\delta(1-\chi))k_{jt}\} + \frac{\phi}{2}\frac{(k_{jt+1}-(1-\delta(1-\chi))k_{jt})^2}{k_{jt}} \tag{2.17}$$

where w_t is the real wage. Total investment is the sum of voluntary investment and maintenance investment.

Entry & Exit. Firms face independent and identically distributed (i.i.d.) exit shocks ϵ_{jt}^{exit} and are forced to exit the economy at the end of the period with probability π^{exit}. Each period, a fixed mass of newborn firms enters the economy. These entrants are endowed with k_0 units of capital and draw their initial productivity level from the distribution $\mu^{ent} \sim \mathcal{N}(0, \frac{\sigma_z^2}{1-\rho_z^2})$, which is the ergodic distribution of (2.15).

21. We normalize the mass of firms to 1. Since entry and exit are exogenous, the mass of firms does not vary in response to aggregate shocks. While our model also features retailers, a final good producer, and a capital good producer, we only refer to intermediate good producers as firms.

22. Matching the empirical distribution of investment rates requires a rich adjustment cost specification, as discussed in Cooper and Haltiwanger (2006).

Timing. Within any period, the timing is as follows. At stage one, idiosyncratic TFP shocks to incumbent firms realize. At stage two, a fixed mass of firms enters the economy. Entrants draw their initial productivity from μ^{ent} and are endowed with k_0 units of capital from the household. Henceforth, they are indistinguishable from incumbent firms. At stage three, firms hire labor, production takes place, and firms conduct maintenance investment. At stage four, exit shocks realize and random fixed adjustment costs are drawn. Exiting firms sell their capital stock and leave the economy. Continuing firms decide whether to adjust their capital stock or remain inactive.

Value Functions. We characterize the firm's optimization problem recursively. The individual state variables are total factor productivity z and capital k. Subscripts for individual variables are henceforth dropped for readability and primes denote next period's values. The beginning-of-period real firm value is

$$V_t(z,k) \;=\; \max_n p_t z k^{\theta} n^{\nu} - w_t n + \pi^{exit} CV_t^{exit}(z,k) + (1-\pi^{exit}) \int_0^{\bar{\xi}} CV_t(z,k,\xi) d\xi \tag{2.18}$$

where CV_t^{exit} and CV_t denote the continuation values of exiting and surviving firms, respectively. With probability π^{exit}, a firm is forced to exit after the production stage. Exiting firms have the liquidation value

$$CV_t^{exit}(z,k) \;=\; (1-\delta) q_t k. \tag{2.19}$$

Note that exiting firms do not need to pay capital adjustment costs. Therefore, maintenance investment does not affect the liquidation value.

The continuation value of a surviving firm is

$$CV_t(z,k,\xi) \;=\; \max\,\{CV_t^a(z,k,\xi), CV_t^n(z,k)\}, \tag{2.20}$$

which reflects that surviving firms can decide whether to adjust their capital stock (CV_t^a) or not (CV_t^n). The continuation value of not adjusting is:

$$CV_t^n(z,k) \;=\; \mathbb{E}_t\big[\Lambda_{t+1} V_{t+1}(z', (1-\delta(1-\chi))k)\big] - q_t \chi \delta k, \tag{2.21}$$

while the continuation value of a firm that adjusts its capital stock is:

$$CV_t^a(z,k,\xi) \;=\; \max_{k'} \mathbb{E}_t\big[\Lambda_{t+1} V_{t+1}(z',k')\big] - q_t\big(k' - (1-\delta)k\big) - AC(k,k',\xi). \tag{2.22}$$

Policy Functions. The labor decision in equation (2.18) is static and independent of the capital decision

$$n_t^*(z,k) \;=\; \left(\frac{p_t \nu z k^\theta}{w_t}\right)^{\frac{1}{1-\nu}}. \tag{2.23}$$

Thus, earnings net of labor costs are

$$\pi_t(z,k) \;\equiv\; p_t z k^\theta (n_t^*)^\nu - w_t n^*. \tag{2.24}$$

The optimal capital decision is computed as follows. First of all, the solution to the maximization problem in equation (2.22) is the policy function $k_t^a(z,k)$, which is independent of ξ. This policy function allows us to compute $CV_t^a(z,k,\xi)$. Since, $CV_t^a(z,k,\xi)$ depends on ξ linearly, we can formulate a cutoff rule for the maximization problem in equation (2.20). Firms choose to adjust capital if and only if their fixed adjustment cost draw ξ is smaller or equal $\xi_t^T(z,k)$:

$$k_t^*(z,k,\xi) \;=\; \begin{cases} k_t^a(z,k) & \text{if } \xi \;\leq\; \xi_t^T(z,k) \\ (1-\delta(1-\chi))k & \text{if } \xi \;>\; \xi_t^T(z,k) \end{cases} \tag{2.25}$$

where $\xi_t^T(z,k) = \frac{CV_t^a(z,k,\xi=0) - CV_t^n(z,k)}{w_t}$.

As in the simple model, we define the hazard rate $\lambda_t(z,k)$ as:

$$\lambda_t(z,k) \;=\; \begin{cases} 0 & \text{if } \xi_t^T(z,k) \;\leq\; 0 \\ \frac{\xi_t^T(z,k)}{\bar{\xi}} & \text{if } 0 \;<\; \xi_t^T(z,k) \;\leq\; \bar{\xi}\,. \\ 1 & \text{if } \bar{\xi} \;<\; \xi_t^T(z,k) \end{cases} \tag{2.26}$$

2.4.2 New Keynesian Block

We separate nominal rigidities from the investment block of the model. A fixed mass of retailers $i \in [0,1]$ produces differentiated varieties $\widetilde{y}_{it}$ from the undifferentiated intermediate goods produced by the production firms. There is a one-to-one production technology $\widetilde{y}_{it} = y_{it}$, where y_{it} is the amount of the intermediate good retailer i purchases. Retailers face Rotemberg quadratic price adjustment costs $\frac{\varphi}{2}\left(\frac{\widetilde{p}_{it}}{\widetilde{p}_{it-1}} - 1\right)^2 Y_t$, where $\widetilde{p}_{it}$ is the relative price of variety i.

A representative final good producer aggregates the differentiated varieties optimally into the final good according to

$$Y_t \;=\; \left(\int \widetilde{y}_{it}^{\frac{\gamma-1}{\gamma}}\, di\right)^{\frac{\gamma}{\gamma-1}} \tag{2.27}$$

The resulting demand function for retail good $\widetilde{y}_{it}$ is:

$$\widetilde{y}_{it} \;=\; \left(\frac{\widetilde{p}_{it}}{P_t}\right)^{-\gamma} Y_t, \tag{2.28}$$

where $P_t = \left(\int_0^1 \widetilde{p}_{it}^{1-\gamma} di\right)^{\frac{1}{1-\gamma}}$ is the price of the final good.

The optimization problem of a monopolistically competitive retailer i is:

$$\max_{\{\widetilde{p}_{it}\}} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \Lambda_t \left\{ (\widetilde{p}_{it} - p_t)\widetilde{y}_{it} - \frac{\varphi}{2}\left(\frac{\widetilde{p}_{it}}{\widetilde{p}_{it-1}} - 1\right)^2 Y_t \right\} \right] \tag{2.29}$$

subject to the demand curve (2.28). We log-linearize the optimality condition of the retailer's problem to obtain the familiar New Keynesian Phillips Curve (NKPC):

$$\log(1+\pi_t) \;=\; \frac{\gamma-1}{\varphi} \log \frac{p_t}{p^*} + \beta\, \mathbb{E}_t \log(1+\pi_{t+1}) \tag{2.30}$$

where $\pi_t \equiv P_t/P_{t-1} - 1$ is the inflation rate, $p^* = \frac{\gamma-1}{\gamma}$ is the relative price (in terms of the final good) of the intermediate good in steady state.

2.4.3 Capital Good Producer

There is a representative capital good producer operating in a perfectly competitive market. It transforms units of the final good into new capital subject to external capital adjustment costs:

$$I_t \;=\; \left[\frac{\delta^{1/\kappa}}{1-1/\kappa} \left(\frac{I_t^Q}{K_t}\right)^{1-1/\kappa} - \frac{\delta}{\kappa - 1} \right] K_t, \tag{2.31}$$

where I_t^Q represents the amount of the final good used, I_t is the amount of new capital produced, and K_t is the total stock of capital in the beginning of period t. The parameter κ determines the strength of external capital adjustment costs. The static optimization problem is:

$$\max_{I_t} q_t I_t - I_t^Q. \tag{2.32}$$

Optimal behavior implies that the relative price of capital (q_t) has to satisfy the following condition

$$q_t \;=\; \left(\frac{I_t^Q/K_t}{\delta} \right)^{1/\kappa}. \tag{2.33}$$

2.4.4 The Central Bank

The central bank sets the nominal interest rate r_t^n according to a Taylor rule

$$\log(1+r_t^n) \;=\; \rho_r \log(1+r_{t-1}^n) + (1-\rho_r)\left[\log\frac{1}{\beta} + \varphi_\pi \log(1+\pi_t)\right] + \epsilon_t^m, \tag{2.34}$$

where ϵ_t^m is a monetary policy shock, ρ_r is the interest rate smoothing parameter, and φ_π is the reaction coefficient to inflation.

2.4.5 Household

There is a representative household, which consumes C_t^h, supplies labor N_t^h, and saves or borrows in one-period non-contingent bonds B_t^h.

The household's objective is to maximize expected lifetime utility

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\log(C_t^h) - \psi N_t^h\right), \tag{2.35}$$

subject to the flow budget constraint:

$$P_t C_t^h + Q_t^B B_t^h \ \leq\ B_{t-1}^h + W_t N_t^h + \Pi_t, \tag{2.36}$$

where Q_t^B is the nominal one-period risk-free bond price (one unit of B_t pays one unit of currency at $t+1$), W_t is the nominal wage, and Π_t subsumes additional transfers to and from the household.[23]

Solving the household's optimization problem leads to the following optimality conditions

$$\Lambda_{t+1} \ \equiv\ \beta \mathbb{E}_t \left[\frac{C_t^h}{C_{t+1}^h}\right], \tag{2.37}$$

$$w_t \ =\ \psi C_t^h, \tag{2.38}$$

where Λ_{t+1} is the household's stochastic discount factor between period t and $t+1$, and w_t is the real wage.

Appendix 2.E.1 defines an equilibrium in this economy.

2.5 Quantitative Results

We calibrate the model to the U.S. economy. Wherever possible, we rely on data sources that refer to the entire economy.

We fix a subset of parameters to conventional values. These parameters are listed in Table 2.1. Given these fixed parameters, we fit the remaining parameters to match the moments listed in Table 2.3. The fitted parameters are listed in Table 2.2.

Since a model period corresponds to a quarter, the discount factor is set to $\beta = 0.99$. The labor disutility parameter is set to $\psi = 0.45$.[24] Capital and labor coefficients are set to standard values, that is, $\theta = 0.21$ and $\nu = 0.64$ (Ottonello and Winberry, 2020). The depreciation rate δ generates an annual aggregate investment

23. Π_t includes dividends from intermediate good producers, retailers, and the final good producer, as well as the initial capital endowment k_0, which entering firms receive from the household. We follow Winberry (2021) and do not rebate back adjustment costs to the household in a lump-sum manner. Therefore, convex adjustment costs do exhaust the aggregate resource constraint.

24. This value follows from normalizing the steady-state real wage w to 1.

rate of 7.7% as reported in Zwick and Mahon (2017). We target the standard deviation of idiosyncratic TFP shocks σ_z, but fix their persistence ρ_z due to the identification problem discussed in Clementi and Palazzo (2015). We set ρ_z to 0.95 (Khan and Thomas, 2008; Bloom et al., 2018). The exogenous exit probability π^{exit} is set to 1.625% as in Koby and Wolf (2020).[25]

We choose standard values for the parameters of the New Keynesian block, i.e. $\varphi = 90$ and $\gamma = 10$ (Ottonello and Winberry, 2020). The coefficient on inflation in the Taylor rule φ_π is set to 1.5 and the interest rate smoothing parameter ρ_r is set to 0.75. External capital adjustment costs κ are set to 8 to roughly match the peak effect of a monetary policy shock on investment relative to the peak effect on output documented in Section 2.2.

Table 2.1. Fixed Parameters

Parameter	Description	Value
Household		
β	Discount factor	0.99
ψ	Labor Disutility	0.45
Investment Block		
θ	Capital Coefficient	0.21
ν	Labor Coefficient	0.64
δ	Depreciation Rate	0.01925
ρ_z	Persistence of TFP Shock	0.95
π^{exit}	Exogenous Exit Probability	0.01625
New Keynesian Block		
ϕ	Price Adjustment Cost	90
γ	Elasticity of Substitution over Intermediate Goods	10
ϕ_π	Taylor Rule Coefficient on Inflation	1.5
ρ_r	Interest Rate Smoothing	0.75
κ	External Capital Adjustment Costs	8

Table 2.2. Fitted Parameters

Parameter	Description	Value
σ_z	Volatility of TFP Shock	0.060
k_0	Initial Capital of Entrants	4.025
$\bar{\xi}$	Upper Bound on Fixed Adjustment Cost	0.350
ϕ	Convex Adjustment Cost	0.750
χ	Maintenance Investment Parameter	0.375

25. This exit probability brings the age distribution as close to the data as possible without using age-specific exit probabilities.

Table 2.3. Empirical & Simulated Moments

Moment	Data	Model
Standard Deviation of Investment Rates	0.200	0.203
Average Investment Rate	0.119	0.136
Autocorrelation of Investment Rates	0.380	0.377
Relative Size of Entrants	0.285	0.297
Relative Spike Rate of Old Firms	0.400	0.380

Notes: Data moments related to investment rates are taken from Zwick and Mahon (2017) (Appendix, Table B.1, Unbalanced Sample). The relative spike rate of old firms is computed from Compustat data. Corresponding model moments are computed from a simulation of a large panel of firms. The relative size of entrants is taken from Business Dynamics Statistics (BDS). In the model, this moment can be computed from the steady-state distribution.

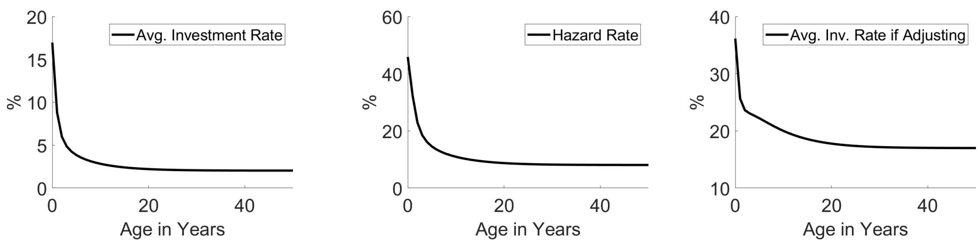
The five parameters listed in Table 2.2 are chosen to match the five targeted moments listed in Table 2.3. Even though all parameters are calibrated jointly, we briefly explain which moments are particularly informative about which parameters.

First, we target the standard deviation of investment rates, because it is informative about the volatility of idiosyncratic TFP shocks. Second, we target the average investment rate as it is informative about both adjustment cost parameters. Increasing either adjustment cost dampens investment rates in particular of young firms and therefore the average investment rate. Third, we target the autocorrelation of investment rates, because it is informative about the relative importance of fixed and convex adjustment costs. Convex adjustment costs generate a positive autocorrelation, whereas fixed adjustment costs generate a negative or zero autocorrelation. For these three moments, we use the statistics reported in Zwick and Mahon (2017). Fourth, we target the relative size of entrants, which is informative about the initial capital of entrants. This moment is computed from Business Dynamics Statistics (BDS) data. Fifth, we target the spike rate of old firms relative to the spike rate of young firms, which is informative about the maintenance investment parameter. The more depreciation is undone by maintenance investment, the less frequently do old firms need to make an extensive margin investment. Thus, a higher maintenance parameter leads to a lower spike rate among old firms. This moment needs to be computed from Compustat data since it is the only data source that includes both investment rates and firm age.

2.5.1 Firm Life-Cycles and the Aggregate Effects of Monetary Policy

Before moving to the key findings of the paper, we show that the model is capable of reproducing well-known facts regarding (i) firms' life cycles and (ii) the aggregate effects of monetary policy shocks.

Firm Life-Cycle Profiles. Figure 2.10 shows that the model matches several untargeted investment life-cycle profiles. The empirical counterparts are shown in Figure 2.A.5. Panel (a) shows that the average investment rate is higher for young firms and falls monotonically in age. Panels (b) and (c) decompose this average investment rate into the average probability to invest ("hazard rate") and the average investment rate conditional on investing. Evidently, the observation that young firms have higher average investment rates is driven in part by higher hazard rates and in part by a higher investment rate conditional on investing.



(a) Avg. Investment Rate **(b)** Hazard Rate **(c)** Avg. Cond. Inv. Rate

Figure 2.10. Life-Cycle Profiles

Notes: Investment rates and the hazard rate refer to a quarter. Averages are computed from the steady state distribution.

The Aggregate Effects of Monetary Policy Shocks. We study the effects of an unexpected expansionary monetary policy shock followed by a perfect foresight transition back to steady state.[26] Figure 2.A.6 plots the impulse response functions of aggregates and prices, which confirm that our model produces the typical New Keynesian effects of a monetary policy shock.[27]

2.5.2 Monetary Policy and the Distribution of Investment Rates

Turning to the main focus of this paper, Figure 2.11 plots the effect of a monetary policy shock on the distribution of investment rates. Specifically, it plots the distribution of investment rates in steady state (blue bars) and in the period when an expansionary monetary policy shock has hit the economy (red bars). It is apparent that monetary policy affects some firms' extensive margin investment decision and therefore the distribution of investment rates: after an interest rate cut, there are fewer inactive firms and more firms choosing to make an investment. This observation corresponds to **Fact 1** documented in Section 2.2.3. Figure 2.12 shows that

26. The size of the monetary shock is chosen to roughly match the peak effects on output and investment seen in the data.

27. Impulse response functions of aggregates are not hump-shaped as in Christiano, Eichenbaum, and Evans (2005), because our model does not feature habit formation.

also the impulse response functions of spike and inaction rates as well as the relative movements of the quantiles of the investment rate distribution match the empirical evidence (see Figures 2.2 and 2.4).

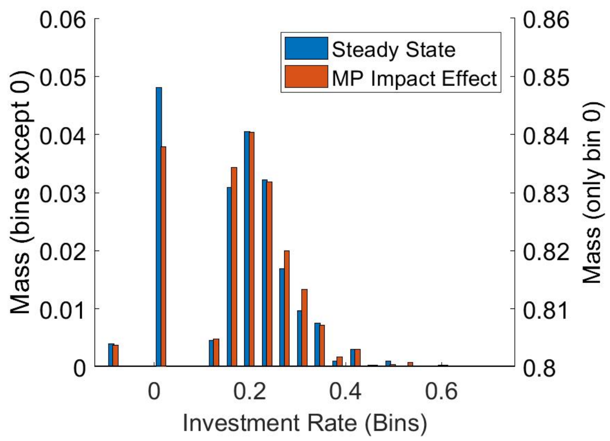
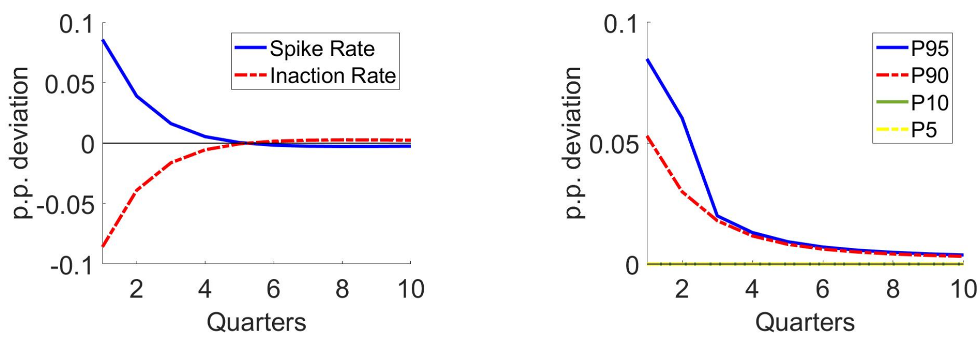


Figure 2.11. Effect of Monetary Policy on the Distribution of Investment Rates

Notes: This figure plots the distribution of investment rates in steady state (blue bars) and after a monetary policy shock (red bars). The monetary policy shock is scaled by a factor of 10 to make differences in the distribution better visible.

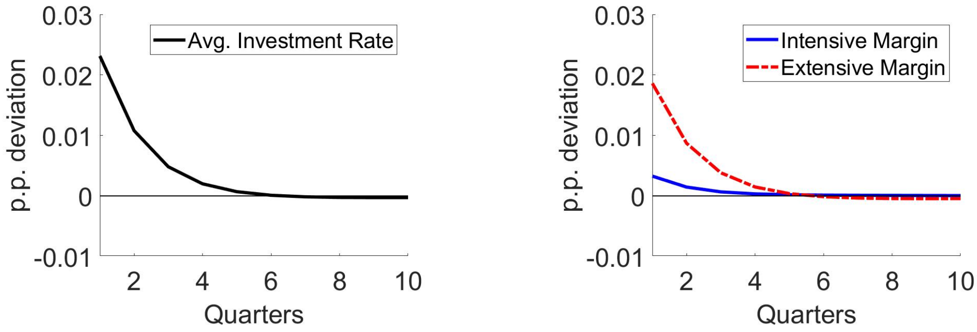


(a) Spike & Inaction Rate **(b)** Quantiles of Distribution

Figure 2.12. Effect of Monetary Policy on Spike Rate, Inaction Rate, Quantiles

Notes: Panel (a) of this figure plots the effect of a monetary policy shock on the spike and inaction rate of all firms. Panel (b) plots the IRFs of certain quantiles of the investment rate distribution.

As in the data, monetary policy affects the average investment rate not only via the extensive margin but also via the intensive margin. To assess the relative importance of both margins, we decompose the effect on the average investment rate into contributions of the extensive and intensive margin, similar to the empirical exercise



(a) Avg. Investment Rate **(b)** Decomposition

Figure 2.13. Effects of Monetary Policy: Extensive & Intensive Margin

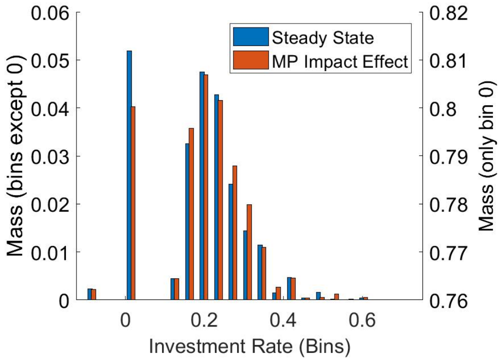
Notes: Panel (a) of this figure plots the effect of a monetary policy shock on the average investment rate of all firms. Panel (b) decomposes the IRF in panel (a) into an extensive margin contribution and an intensive margin contribution.

presented in Figure 2.7.[28] Figure 2.13 presents the findings. Panel (a) plots the effect of an expansionary monetary policy shock on the average investment rate. Panel (b) plots the decomposition. Evidently, the model attributes a significant portion of the change in the average investment rate to the extensive margin (**Fact 3**).

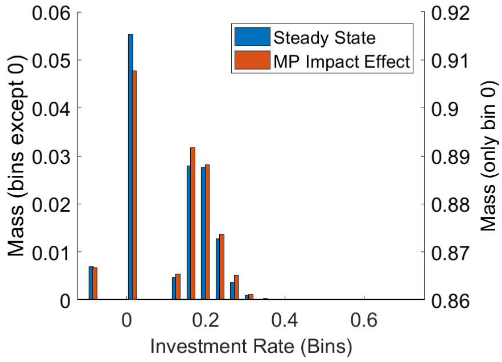
Heterogeneous Sensitivity of Young Firms. In addition, the model reproduces the empirical finding that the effect of monetary policy on the distribution of investment rates is heterogeneous across different age groups, as shown in Figure 2.14. This corresponds to **Fact 2**. Panels (a) and (b) plot the distribution of investment rates before and after an expansionary monetary policy shock of young and old firms, respectively. The bottom panels plot the changes in the distribution. This shows that after an interest rate cut, there are more young firms than old firms switching from being inactive to making a large investment. That is, the effect of monetary policy along the extensive margin is more pronounced among young firms.

Due to the heterogeneous effect along the extensive margin, monetary policy affects *average* investment rates differently across age groups. Figure 2.15 plots the effect of an expansionary monetary policy shock on average investment rates by age group. Panel (a) shows that young firms increase their investment rates on average more strongly than old firms. Panel (b) decomposes this heterogeneous sensitivity into extensive and intensive margin contributions, similar to the empirical exercise shown in Figure 2.8b, and demonstrates that the total difference is driven by the extensive margin (**Fact 3**). Panel (c) further decomposes the extensive margin into the two mechanisms identified in the simple model (Equation 2.13). On the one

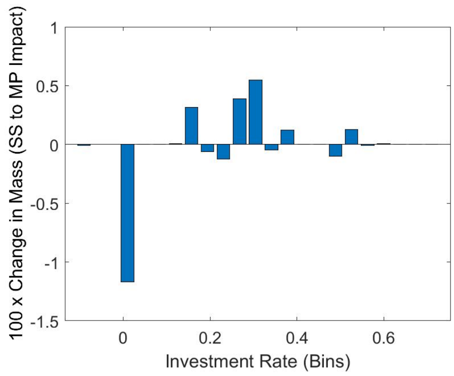
28. This decomposition is computed by holding either hazard rates at steady-state levels (intensive margin contribution) or investment rates conditional on investing at steady-state levels (extensive margin contribution), see also Equation (2.4).



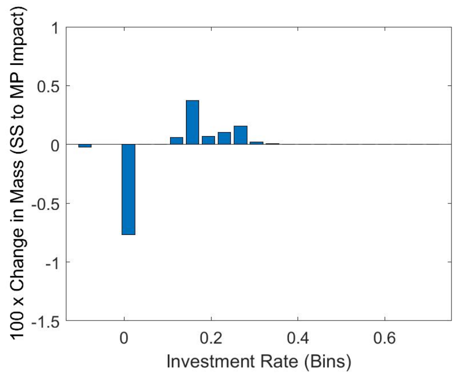
(a) Young Firms



(b) Old Firms



(c) Young Firms (Difference)



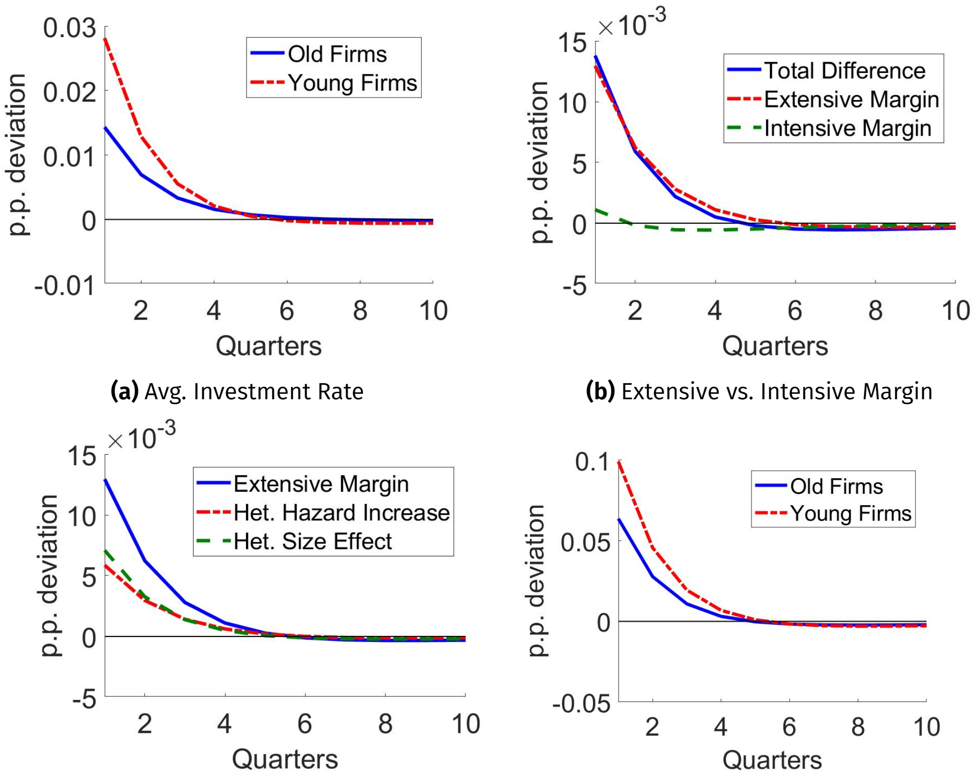
(d) Old Firms (Difference)

Figure 2.14. Effect of Monetary Policy on the Distribution of Inv. Rates (by Age Group)

Notes: Panels (a) and (b) of this figure plot the distribution of investment rates of young (old) firms in steady state and after a monetary policy shock. Panels (c) and (d) plot the difference between the two distributions for young (old) firms. The monetary policy shock is scaled by a factor of 10 to make differences in the distribution better visible.

hand, the hazard rate rises more strongly among young firms (heterogeneous hazard rate increase), which is separately shown in panel (d). On the other hand, new young adjusters on average have a higher investment rate than new old adjusters (heterogeneous size effect). Quantitatively, the heterogeneous size effect is slightly more important.

To summarize, these results confirm that the two effects identified in the simple model in Section 2.3 hold and are quantitatively relevant in a calibrated general equilibrium heterogeneous-firm model. First, there is an important investment channel of monetary policy along the extensive margin. Second, this effect does not affect all firms homogeneously: young (small) firms' average investment rates are more sensitive to monetary policy even in the absence of a financial accelerator mechanism.



(a) Avg. Investment Rate

(b) Extensive vs. Intensive Margin

(c) Extensive Margin Decomposition

(d) Spike Rate

Figure 2.15. Heterogeneous Effect of Monetary Policy (by Age Group)

Notes: Panel (a) of this figure plots the effect of a monetary policy shock on the average investment rates of young and old firms. Panel (b) decomposes the difference between the two IRFs in panel (a) into an extensive margin contribution and an intensive margin contribution. Panel (c) further decomposes the IRF of the extensive margin contribution in panel (b) into the heterogeneous hazard rate increase and the heterogeneous size effect. Panel (d) plots the effect of a monetary policy shock on the spike rates of young and old firms.

2.6 Conclusion

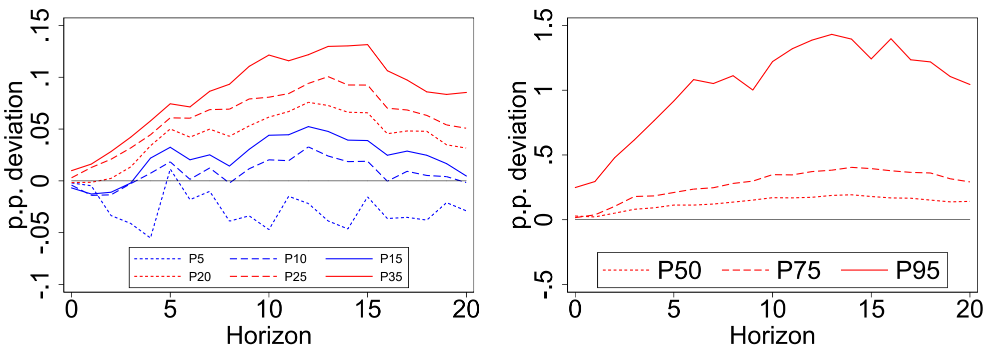
In this paper, we highlight two features of the investment channel of monetary policy. First, there is a quantitatively relevant investment channel along the extensive margin. That is, an interest rate cut induces some firms to switch from making a small or no investment to making a sizeable one. Second, along the extensive margin, young firms respond more strongly to interest rate changes than old firms. Therefore, young firms are more sensitive to monetary policy even in the absence of a financial accelerator mechanism.

We present three pieces of evidence in line with these effects. First, monetary policy affects the shape of the distribution of investment rates. Specifically, an interest rate cut leads to fewer small or zero investment rates and more large investment rates. Second, this change in the distribution is more pronounced among young firms than among old firms. Third, a decomposition exercise indicates that the extensive margin accounts for around 50% of the effect of monetary policy on the average investment rate and for more than 50% of the heterogeneous effect on firms of different age groups.

We build a heterogeneous-firm model that combines fixed adjustment costs, firm life-cycle dynamics, and a New Keynesian sticky-price setup to interpret these novel empirical findings. In the model, monetary policy affects firms' investment decisions along the intensive and, importantly, along the extensive margin. The extensive margin investment channel arises due to fixed capital adjustment costs. Quantitatively, the extensive margin explains a large chunk of the effect of monetary policy on the average investment rate as well as of the heterogeneous sensitivity of young firms—as in the data.

Our findings have important implications for both academic research and the conduct of monetary policy. First, the paper raises the issue of observational equivalence: firms typically classified as financially constrained (young/small) are more sensitive to monetary policy even in the absence of a financial accelerator mechanism. Second, understanding the frictions underlying firms' (heterogeneous) investment decisions is important for guiding macroeconomic policies in recessions. The financial accelerator mechanism suggests that macroeconomic policies are more effective in downturns. In contrast, the presence of an extensive margin investment decision—which we highlight in this paper—makes monetary and fiscal policy interventions less potent in recessions.

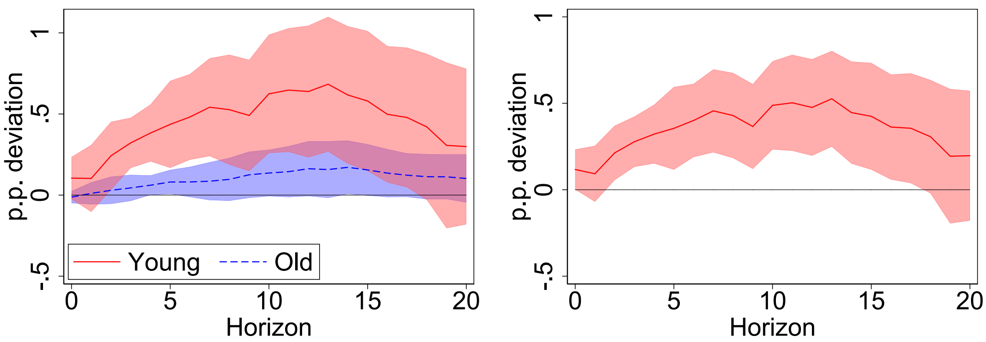
Appendix 2.A Additional Figures



(a) P5 - P35 **(b)** P50, P75 & P90

Figure 2.A.1. Effect of Monetary Policy on Quantiles of the Inv. Rate Distribution

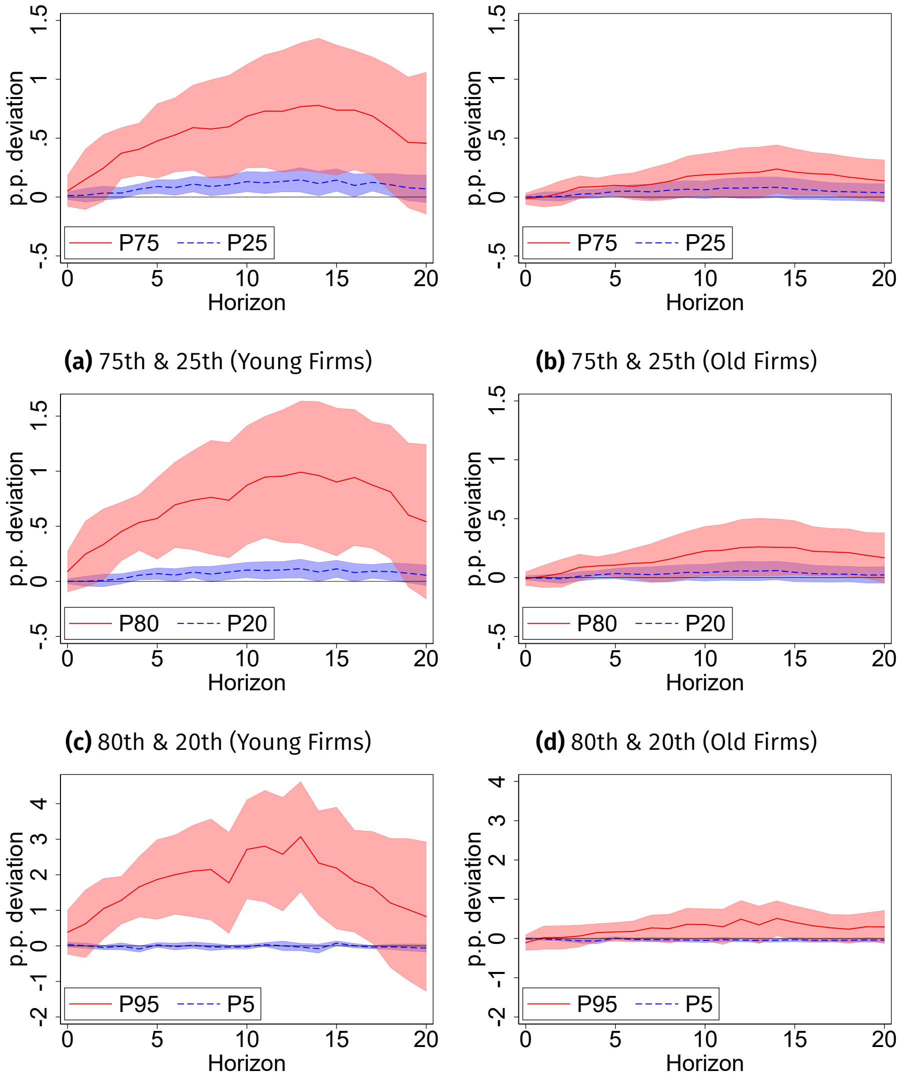
Notes: The lines represent the estimated $\hat{\beta}^h$ from separate regressions: $y_{t+h} - y_{t-1} = \alpha^h + \beta^h \varepsilon_t^{MP} + \sum_{j=2}^{4} \gamma^j \mathbb{1}\{q_{t+h} = j\} + e_{t+h}$, using monetary policy shocks ε_t^{MP} implied by the Proxy SVAR in Gertler and Karadi (2015). The shocks are scaled to reduce the 1-year Treasury rate by 25 basis points. Sample: 1986Q1 - 2018Q4.



(a) Average Investment Rates **(b)** Difference (Young - Old)

Figure 2.A.2. Effect of Monetary Policy on Age-Group-Specific Avg. Inv. Rates

Notes: Young (old) firms are less (more) than 15 years old. The lines represent the estimated $\hat{\beta}^h$ from separate regressions: $y_{t+h} - y_{t-1} = \alpha^h + \beta^h \varepsilon_t^{MP} + \sum_{j=2}^{4} \gamma^j \mathbb{1}\{q_{t+h} = j\} + e_{t+h}$, using monetary policy shocks ε_t^{MP} implied by the Proxy SVAR in Gertler and Karadi (2015). The shocks are scaled to reduce the 1-year Treasury rate by 25 basis points. The shaded areas are the 90% confidence intervals constructed using standard errors that are robust to heteroskedasticity and autocorrelation. Sample: 1986Q1 - 2018Q4.



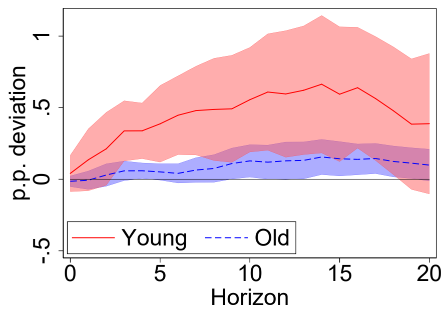
(a) 75th & 25th (Young Firms) **(b)** 75th & 25th (Old Firms)

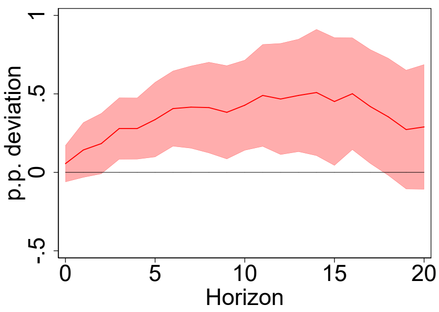
(c) 80th & 20th (Young Firms) **(d)** 80th & 20th (Old Firms)

(e) 95th & 5th (Young Firms) **(f)** 95th & 5th (Old Firms)

Figure 2.A.3. Effect on Quantiles of Age-Group-Specific Inv. Rate Distributions

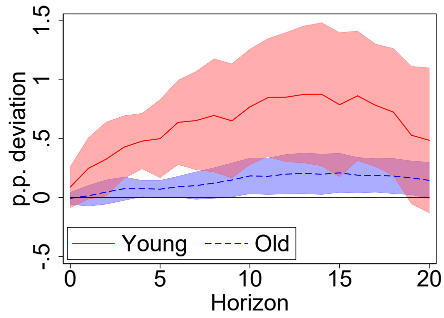
Notes: This figure plots the effect of a monetary policy shock on quantiles of the age-specific investment rate distributions. Young (old) firms are firms less (more) than 15 years old. The lines represent the estimated $\hat{\beta}^h$ from separate regressions: $y_{t+h} - y_{t-1} = \alpha^h + \beta^h \varepsilon_t^{MP} + \sum_{j=2}^{4} \gamma^j \mathbb{1}\{q_{t+h} = j\} + e_{t+h}$, using monetary policy shocks ε_t^{MP} implied by the Proxy SVAR in Gertler and Karadi (2015). The shocks are scaled to reduce the 1-year Treasury rate by 25 basis points. The shaded areas indicate the 90% confidence intervals constructed using standard errors that are robust to heteroskedasticity and autocorrelation. Sample: 1986Q1 - 2018Q4.

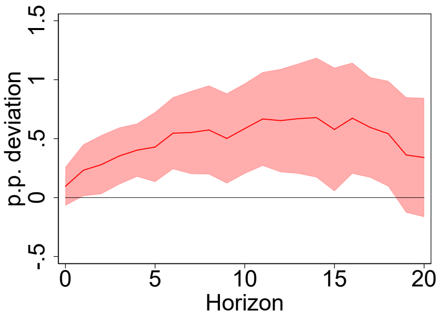




(a) Interquantile Range: P75 - P25

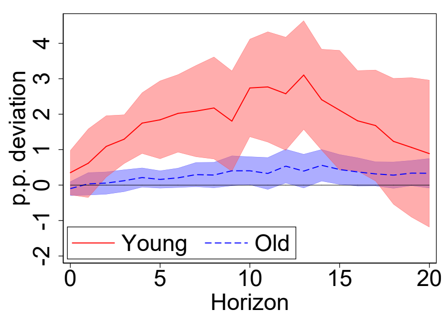
(b) Difference (Young - Old)

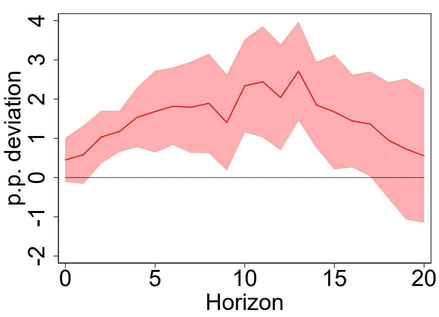




(c) Interquantile Range: P80 - P20

(d) Difference (Young - Old)



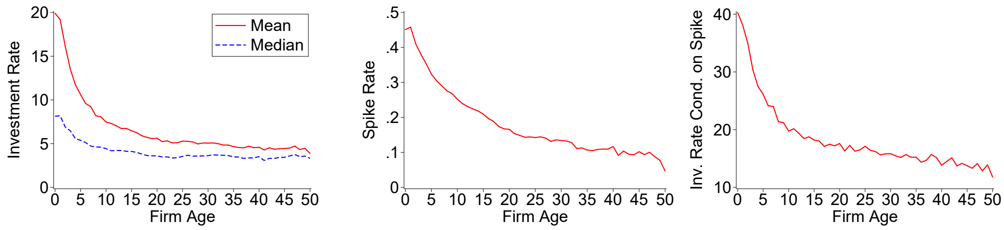


(e) Interquantile Range: P95 - P5

(f) Difference (Young - Old)

Figure 2.A.4. Effect on Interquantile Ranges of Age-Group-Specific Inv. Rate Distributions

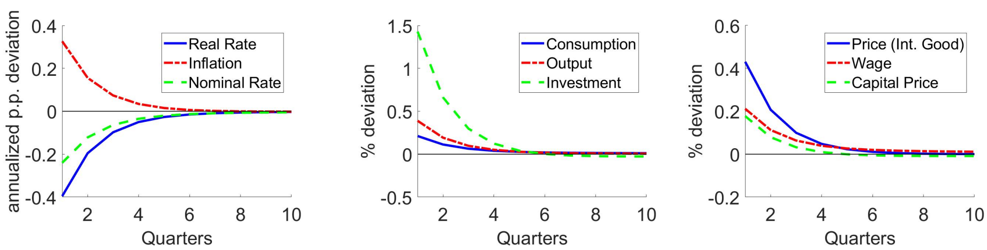
Notes: This figure plots the effect of a monetary policy shock on statistics of the age-specific investment rate distributions. Young (old) firms are firms less (more) than 15 years old. The lines represent the estimated $\hat{\beta}^h$ from separate regressions: $y_{t+h} - y_{t-1} = \alpha^h + \beta^h \varepsilon_t^{MP} + \sum_{j=2}^{4} \gamma^j \mathbb{1}\{q_{t+h} = j\} + e_{t+h}$, using monetary policy shocks ε_t^{MP} implied by the Proxy SVAR in Gertler and Karadi (2015). The shocks are scaled to reduce the 1-year Treasury rate by 25 basis points. The shaded areas indicate the 90% confidence intervals constructed using standard errors that are robust to heteroskedasticity and autocorrelation. Sample: 1986Q1 - 2018Q4.



(a) Investment Rate **(b)** Spike Rate **(c)** Avg. Cond. Inv. Rate

Figure 2.A.5. Empirical Life-Cycle Profiles

Notes: Investment rates and the spike rate refer to a quarter. A spike is defined as an investment rate ≥ 10%. The average conditional investment rate (panel c) is the average investment rate among all firms with an investment rate ≥ 10%.



(a) Interest Rates & Inflation **(b)** Cons., Output & Investm. **(c)** Prices

Figure 2.A.6. Aggregate Effects of an Expansionary Monetary Policy Shock

Notes: This figure plots the effects of a monetary policy shock on interest rates, inflation, aggregates, and prices.

Appendix 2.B Empirical Appendix

2.B.1 Sample Selection

We use the Compustat North America Fundamentals Quarterly database. Observations are uniquely identified by GVKEY & DATADATE. In line with the literature, we exclude observations that fall under the following criteria

(1) not incorporated in the United States (based on FIC)
(2) native currency not U.S. Dollar (based on CURNCDQ)
(3) fiscal quarter does not match calendar quarter (based on FYR)
(4) specific sectors
- Utilities (SIC 4900-4999)
- Financial Industry (SIC 6000-6999)
- Non-operating Establishments (SIC 9995)
- Industrial Conglomerates (SIC 9997)
- Non-classifiable (NAICS > 999900)

(5) missing industry information (SIC or NAICS code)
(6) missing capital expenditures (based on CAPX)
(7) missing or non-positive total assets (AT) or net capital (PPENT)
(8) negative sales (SALEQ)
(9) acquisitions (based on AQCY) exceed 5% of total assets (in absolute terms)
(10) missing or implausible age information (see Appendix 2.B.2)
(11) outlier in the Perpetual Inventory Method (see Appendix 2.B.3)

Our sample begins with 1986Q1 and ends with 2018Q4. In a final step, we exclude firms which we observe for less than 20 quarters unless they are still in the sample in the final period. This ensures that we do not mechanically exclude all firms incorporated in the last five years of our sample.

2.B.2 Firm Age

We use data on firm age from WorldScope and Jay Ritter's database[29]. WorldScope provides the date of incorporation (Variable: INCORPDATE), while Jay Ritter's database provides the founding date. Both are merged with Compustat based on CUSIP. We define as the firm entry quarter the minimum of both dates if both are available. We do not use information on the initial public offering (IPO) of a firm

29. https://site.warrington.ufl.edu/ritter/

to determine its age, since the time between incorporation and IPO can vary substantially. However, we use the IPO date to detect implausible age information. We exclude firms for which the IPO date reported in Compustat (IPODATE) precedes the firm entry quarter by more than four quarters. Similarly, we exclude firms that appear in Compustat more than four quarters before the firm entry quarter.[30] Finally, we merge information on the beginning of trading from CRSP (Variable: BEGDAT) based on CUSIP and likewise exclude firms with trading more than four quarters before the firm entry quarter.

2.B.3 Perpetual Inventory Method

Accounting capital stocks $k^a_{j,t}$ as reported in Compustat deviate from *economic* capital stocks for at least two reasons. First, accounting depreciation is driven by tax incentives and usually exceeds economic depreciation. Second, accounting capital stocks are reported at historical prices, not current prices. With positive inflation, both issues make the economic capital stock exceed the accounting capital stock. Therefore, we use a Perpetual Inventory Method (PIM) to compute real economic capital stocks, building on Bachmann and Bayer (2014).

Investment. In principle, there are two options to measure net nominal quarterly investment. First, investment can be measured directly ($I^{dir}_{j,t}$) from the Statement of Cash Flows as capital expenditures (CAPX) less the sale of PPE (SPPE)[31]. Second, investment can be backed out ($I^{indir}_{j,t}$) from the change in PPE (D.PPENT) plus depreciation (DPQ), using Balance Sheet and Income Statement information. Either measure needs to be deflated to obtain real investment. We use INVDEF from FRED, which has the advantage of being quality-adjusted. We prefer the direct investment measure, since the indirect measure basically captures any change to PPE, including changes due to acquisitions. Nevertheless, we want to exclude observations where both investment measures differ strongly. To this end, we compute investment rates using lagged net accounting capital (L.PPENT), compute the absolute difference between both and discard the top 1% of that distribution.

Depreciation Rates. We obtain economic depreciation rates from the Bureau of Economic Analysis' (BEA) Fixed Asset Accounts. Specifically, we retrieve current-cost net stock and depreciation of private fixed assets by year and industry.[32] We

30. We do not construct firm age from the first appearance in Compustat. An inspection of the data reveals that this would result in wrongly classifying a number of old and established firms as young. Cloyne et al. (2020) do exactly this. However, they show in an earlier working paper version that results are unchanged if only age information from WorldScope is used.

31. We follow Belo, Lin, and Bazdresch (2014) and set missing values of SPPE to zero.

32. The Fixed Asset Accounts also provide depreciation rates by asset type (Equipment, Structures, Intellectual Property Products), which we do not use since the firm-level data does not include information on capital stocks or capital expenditure by asset type.

calculate annual depreciation rates by industry and assume a constant depreciation rate within the calendar year to calculate quarterly depreciation rates.

Real Economic Capital Stocks. We initialize a firm's capital stock with the net (real) accounting capital stock $k^a_{j,1}$ (PPENT / INVDEF) whenever this variable is first observed. We iterate forward using deflated investment and the economic depreciation rate.

$$k^{(1)}_{j,1} = k^a_{j,1} \tag{2.B.1}$$

$$k^{(1)}_{j,t+1} = (1-\delta^e_t)k^{(1)}_{j,t} + \frac{p^I_t}{p_{2009,t}} I^{dir}_{j,t} \tag{2.B.2}$$

Comparing $k^{(1)}_{j,t}$ and $k^a_{j,t}$ shows non-negligible discrepancies. On average, the economic capital stock is larger, confirming the hypothesis that accounting capital stocks are understated. This makes it problematic to use the accounting capital stock as a starting value in the PIM. As a remedy, we again follow Bachmann and Bayer (2014) and use an iterative procedure to re-scale the starting value. We compute a time-invariant scaling factor ϕ at the sector level and use it to re-scale the starting value as follows. We iterate until ϕ converges. The procedure is initialized with $k^{(0)}_{j,t} = k^a_{j,t}$ and $\phi^{(0)} = 1$.

$$\phi^{(n)} = \frac{1}{NT} \sum_{j,t} \frac{k^{(n)}_{j,t}}{k^{(n-1)}_{j,t}} \quad \text{[and not in top or bottom 1\%]} \tag{2.B.3}$$

$$k^{(n+1)}_{j,1} = \phi^{(n)} k^{(n)}_{j,1} \tag{2.B.4}$$

Outliers. We exclude firms for which the economic capital stock becomes negative at any point in time. This can arise if there is a sale of capital, which exceeds current economic capital. Further, we compute the deviation between (real) accounting and economic capital stocks and discard the top 1% of that distribution. Finally, we discard firms for which we have less than 20 observations unless they are still in the sample in the final quarter.

Evaluation. Our estimated real economic capital stock is still highly correlated with the real accounting capital stock. A simple regression has an R^2 of above 0.96 and shows that the economic capital stock is on average slightly higher (by about 4%), as expected. The investment rate (net real investment over lagged real economic capital) is highly correlated ($\rho > 0.98$) with the accounting investment rate used in Cloyne et al. (2020). A simple regression shows that on average, the economic investment rate is lower (by about 13%) than the accounting investment rate, also as expected due to the underreporting of accounting capital stocks.

2.B.4 Variable Construction

Most of our variables follow the definitions in the literature. Our baseline measure of the investment rate is $i_{jt} = \frac{CAPX_{jt} - SPPE_{jt}}{INVDEF_t \times k_{jt-1}}$, thus, real capital expenditures (CAPX) net of sales of capital (SPPE) divided by the lagged real economic capital stock, computed as described previously. To measure size, we use the log of total assets (AT).

2.B.5 Identification of Monetary Policy Shocks

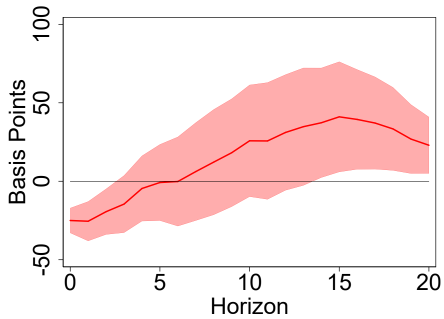
We use the monetary policy shocks implied by the proxy SVAR used in Gertler and Karadi (2015). We calculate them according to the following procedure. First, we update the data used in the Gertler and Karadi (2015) baseline SVAR. They use monthly data from 1979M7 to 2012M6. We update all time series to 2019M12. The SVAR includes (the log of) industrial production (FRED: INDPRO), (the log of) the consumer price index (FRED: CPIAUCSL), the one-year government bond rate (FRED: GS1), and the excess bond premium (Source: https://www.federalreserve.gov/econresdata/notes/feds-notes/2016/files/ebp_csv.csv, retrieved in February 2020). Moreover, we update the instrument (cumulative high-frequency FF4 surprises) to 2015M10. Then, we run the SVAR and compute the implied structural monetary policy shocks. See the appendix of Mertens and Ravn (2013) for details. Importantly, even though the instrument is only available until 2015M10, we can compute the structural monetary policy shock until 2019M12.

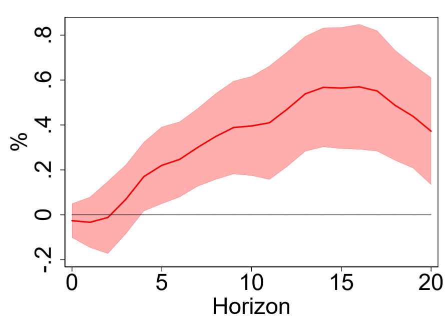
2.B.6 Effects of Monetary Policy using Aggregate Data

Using time series data from FRED, we document the aggregate effects of the monetary policy shocks we utilize. Qualitatively, these are quite similar to Gertler and Karadi (2015). Panel (a) of Figure 2.B.1 shows that a monetary policy shock decreases the 1-year Treasury rate (FRED: GS1) for roughly 4 quarters. Thereafter, it overshoots, as observed in Gertler and Karadi (2015). Panels (b) and (c) show that (real) investment (FRED: PNFI) and the relative price of capital goods (FRED: PIRIC) increase strongly. The peak effect on investment is roughly 1.4%. As we will show in the model, the endogenous response of the relative price of capital generates a heterogeneous effect on young and old firms. Panel (d) shows that real GDP (FRED: GDPC1) also increases following an expansionary shock. The peak effect is about 0.35%.

2.B.7 Aggregate vs. Compustat Data

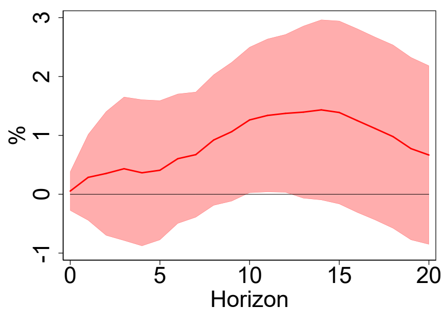
Following other studies in the literature, we use Compustat data because it offers quarterly firm-level data including information on investment rates and firm age. However, Compustat firms, being public firms, are by no means a random or representative sample of the universe of firms in the economy. In the following, we show

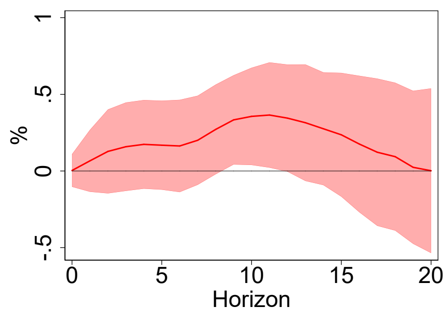




(a) 1-Year Treasury Rate

(b) Relative Price of Investment Goods





(c) Investment

(d) GDP

Figure 2.B.1. Aggregate Effects of a Monetary Policy Shock

Notes: The lines represent the estimated $\hat{\beta}^h$ from separate regressions: $y_{t+h} - y_{t-1} = \alpha^h + \beta^h \varepsilon_t^{MP} + \sum_{j=2}^{4} \gamma^j \mathbb{1}\{q_{t+h} = j\} + e_{t+h}$, using monetary policy shocks ε_t^{MP} implied by the Proxy SVAR in Gertler and Karadi (2015). The shocks are scaled to reduce the 1-year Treasury rate by 25 basis points. The shaded areas are the 90% confidence intervals constructed using standard errors that are robust to heteroskedasticity and autocorrelation. Sample: 1986Q1 - 2018Q4. All variables except for the 1-year Treasury rate are in logs.

that data on investment from Compustat *qualitatively* aligns well with aggregate data from national accounts. At the same time, there are substantial *quantitative* differences.

Figure 2.B.2 compares three quarterly *aggregate* investment rates. The first one is computed from national account data, following the procedure described in Bachmann, Caballero, and Engel (2013). The other two are constructed from our sample of Compustat firms reflecting two alternative ways of constructing capital. The first one uses investment and capital as computed with the perpetual inventory method ("PIM"). The second one uses investment and capital as reported in Compustat ("Accounting"). As expected, aggregate investment rates from Compustat have a substantially higher level, which is at least in part due to the capital measurement issues described in Appendix 2.B.3. The PIM addresses these issues to some extent, but the level of the investment rate remains substantially above the national-account investment rate. Moreover, the Compustat aggregate investment rates are substantially

more volatile. Despite these differences in the level and volatility, the investment rates are highly correlated.[33]

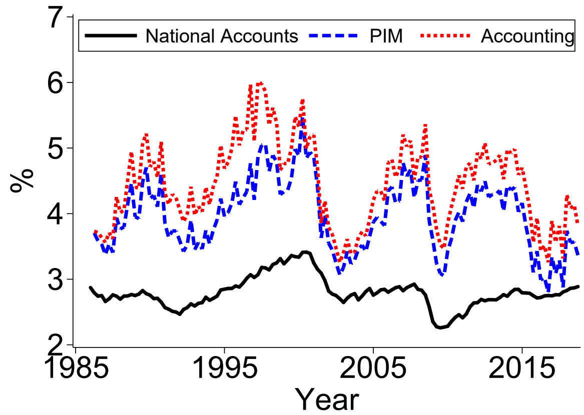
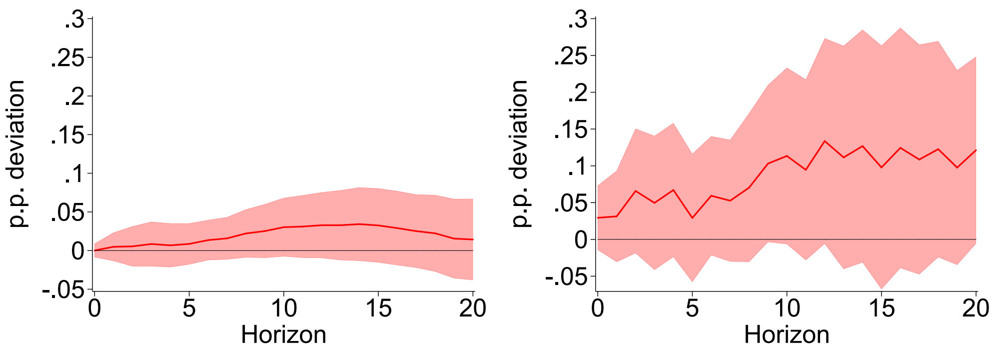


Figure 2.B.2. Aggregate Investment Rates

Notes: Both Compustat investment rates are seasonally adjusted using quarterly dummy variables. This deals with the observation that reported investment rates are typically higher in the fourth quarter (Xu and Zwick, 2021).



(a) National Accounts Data **(b)** Compustat Data (PIM)

Figure 2.B.3. Effect of Monetary Policy Shock on Aggregate Investment Rates

Notes: The lines represent the estimated $\hat{\beta}^h$ from separate regressions: $y_{t+h} - y_{t-1} = \alpha^h + \beta^h \varepsilon_t^{MP} + \sum_{j=2}^{4} \gamma^j \mathbb{1}\{q_{t+h} = j\} + e_{t+h}$, using monetary policy shocks ε_t^{MP} implied by the Proxy SVAR in Gertler and Karadi (2015). The shocks are scaled to reduce the 1-year Treasury rate by 25 basis points. The shaded areas are the 90% confidence intervals constructed using standard errors that are robust to heteroskedasticity and autocorrelation. Sample: 1986Q1 - 2018Q4.

It is not surprising to find that the differing volatilities lead to different estimates regarding the effect of monetary policy. Figure 2.B.3 plots the impulse response func-

33. The aggregate investment rate from national accounts has a correlation of $\rho = 0.6$ with the "PIM" investment rate and of $\rho = 0.54$ with the "Accounting" investment rate. Both Compustat investment rates are highly correlated ($\rho = 0.95$).

tion of the aggregate investment rate from national accounts and from Compustat (PIM). While the shape is very similar, the magnitude differs substantially. The peak effects are about 0.03 percentage points (national accounts) and 0.13 percentage points (Compustat), respectively.

In Section 2.5, we calibrate our model to the entire U.S. economy, not only to Compustat firms. Therefore, the model quantitatively reflects the former number, not the latter one.

Appendix 2.C Proofs

Proposition 2.1. *In an economy populated by heterogeneous firms that face fixed adjustment costs as described above, it holds that*

(1) An interest rate cut increases the hazard rate: $\frac{\partial \lambda(k_0)}{\partial r} < 0$

(2) The sensitivity of the average investment rate to interest rate changes via the extensive margin is decreasing (in absolute terms) in firm size: $\frac{\partial\left(\frac{\partial \lambda(k_0)}{\partial r} i^*(k_0)\right)}{\partial k_0} > 0$

Proof. Rearranging equation (2.10), the value added of adjusting capital while ignoring the fixed adjustment cost is:

$$VA(k_0) = \frac{1}{1+r}\left(k_1^{*\theta} - k_0^{\theta}\right) - q(k_1^* - k_0) \tag{2.C.1}$$

where k_1^* was defined in equation (2.8). Using the definition of the cutoff $\xi^T(k_0)$ and the hazard rate $\lambda(k_0)$ from the main text, we have

$$\lambda(k_0) = \frac{1}{\bar{\xi}} VA(k_0). \tag{2.C.2}$$

Taking the derivative w.r.t. the real interest rate, we get

$$\frac{\partial \lambda(k_0)}{\partial r} = -\frac{1}{\bar{\xi}} \frac{1}{(1+r)^2}\left(k_1^{*\theta} - k_0^{\theta}\right) < 0, \tag{2.C.3}$$

which proves the first part of the proposition. Note that $k_0 < k_1^*$ by assumption.

The second part of the proposition requires

$$\frac{\partial\left(\frac{\partial \lambda(k_0)}{\partial r} i^*(k_0)\right)}{\partial k_0} = \frac{\partial^2 \lambda(k_0)}{\partial r \partial k_0} i^*(k_0) + \frac{\partial \lambda(k_0)}{\partial r}\frac{\partial i^*(k_0)}{\partial k_0} > 0. \tag{2.C.4}$$

The first term is positive, because

$$\frac{\partial^2 \lambda(k_0)}{\partial r \partial k_0} = \frac{1}{\bar{\xi}} \frac{1}{(1+r)^2} \theta k_0^{\theta-1} > 0 \tag{2.C.5}$$

and $i^*(k_0) > 0$ because $k_0, k_1 > 0$. The second term is positive because

$$\frac{\partial i^*(k_0)}{\partial k_0} = -k_1^* k_0^{-2} < 0 \tag{2.C.6}$$

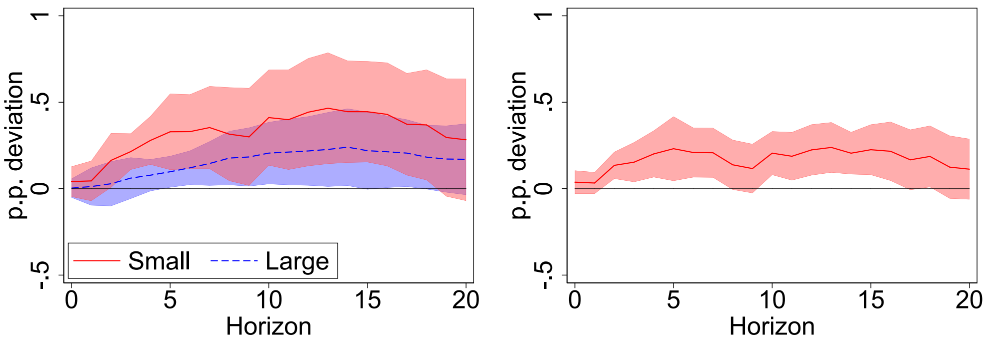
and $\frac{\partial \lambda(k_0)}{\partial r} < 0$ as shown in equation (2.C.3). Thus, the inequality in equation (2.C.4) holds which completes the proof.

□

Appendix 2.D Heterogeneous Sensitivity by Firm Size

Cloyne et al. (2020) have shown that being young is a better predictor of a firm's sensitivity to monetary policy shocks than being small. We replicate this finding in Figure 2.D.1. Firms that are smaller than the median are at the peak on average 24 basis points more sensitive than firms that are larger than the median. In comparison, young firms are at the peak on average 53 basis points more sensitive than old firms, as shown in Figure 2.A.2. This weaker heterogeneous sensitivity goes along with a weaker heterogeneous sensitivity of the extensive margin, as shown in Figure 2.D.3, which replicates Figure 2.6 while grouping firms by size instead of age. In addition, the change in the distribution differs somewhat less across size groups than across age groups, as can be seen from Figures 2.D.4 and 2.A.3.

Our model is able to replicate the finding that young age is a better predictor of firms' sensitivity to monetary policy shocks than small size. This is evident from Figure 2.D.2, which replicates Figure 2.15, panel (a), while grouping firms by size instead of age. Firms that are smaller than the median are on impact more sensitive than firms larger than the median, but the difference is by about 50% smaller than the gap between young and old firms. Intuitively, age is the better predictor of sensitivity, because young firms are more likely to be "close to making a large investment". This is because young firms are born small and will almost certainly grow in the future. In contrast, small firms may or may not be "close to making a large investment". This is because some firms are small because they are very unproductive, such that the low level of capital is their desired level of capital. In a nutshell, size correlates positively with productivity, while age is uncorrelated with productivity.

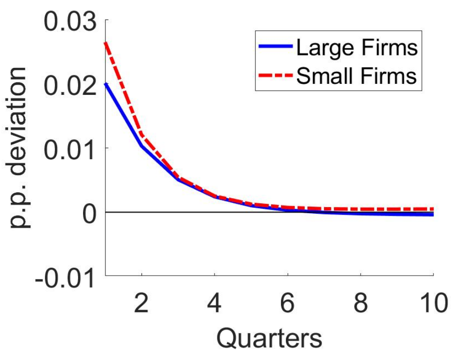


(a) Average Investment Rates

(b) Difference (Small - Large)

Figure 2.D.1. Effect of Monetary Policy Shock on Average Investment Rates by Size

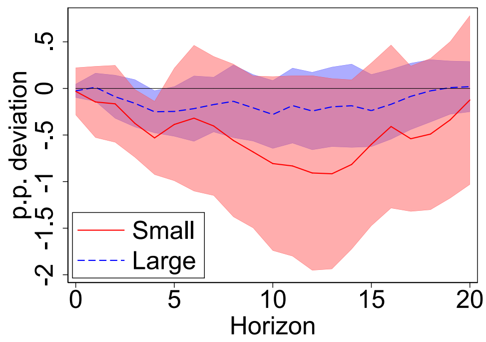
Notes: Small (large) firms are firms smaller (larger) than the median in a given quarter. The lines represent the estimated $\hat{\beta}^h$ from separate regressions: $y_{t+h} - y_{t-1} = \alpha^h + \beta^h \varepsilon_t^{MP} + \sum_{j=2}^{4} \gamma^j \mathbb{1}\{q_{t+h} = j\} + e_{t+h}$, using monetary policy shocks ε_t^{MP} implied by the Proxy SVAR in Gertler and Karadi (2015). The shocks are scaled to reduce the 1-year Treasury rate by 25 basis points. The shaded areas are the 90% confidence intervals constructed using standard errors that are robust to heteroskedasticity and autocorrelation. Sample: 1986Q1 - 2018Q4.



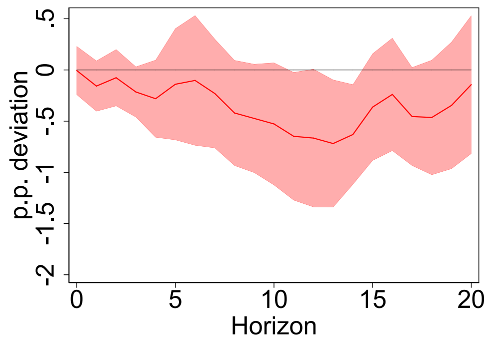
(a) Avg. Investment Rate

Figure 2.D.2. Effect of Monetary Policy Shock on Average Investment Rates by Size Group (Model)

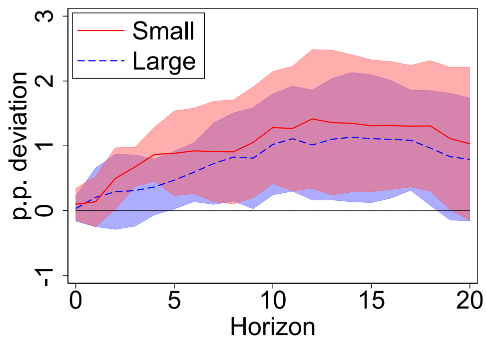
Notes: This figure plots the effect of an expansionary monetary policy shock on the average investment rates of small and large firms. Small (large) firms are firms smaller (larger) than the median in a given quarter.



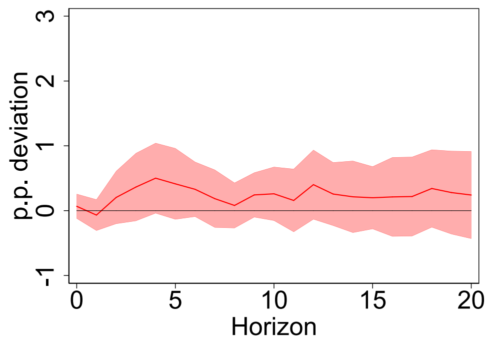
(a) Spike Rate (by Size)



(b) Spike Rate (Difference)



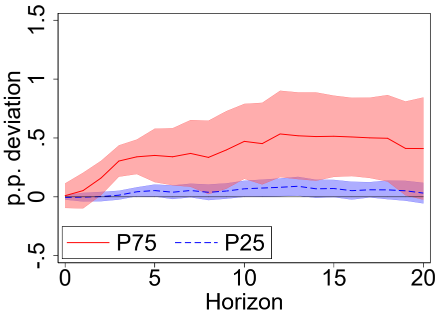
(c) Inaction Rate (by Size)



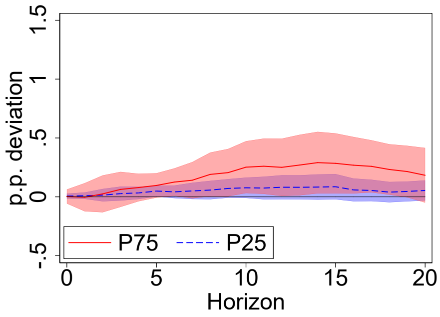
(d) Inaction Rate (Difference)

Figure 2.D.3. Effect on Group-Specific Spike & Inaction Rates (by Size)

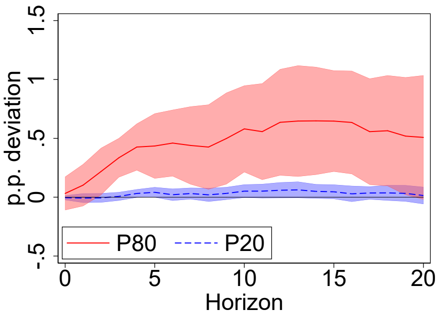
Notes: This figure plots the effect of a monetary policy shock on the spike rate and the inaction rate of small and large firms. Small (large) firms are firms smaller (larger) than the median in a given quarter. A spike rate is an investment rate exceeding 10%, an inaction rate is an investment rate less than 0.5% in absolute value. The lines represent the estimated $\hat{\beta}^h$ from separate regressions: $y_{t+h} - y_{t-1} = \alpha^h + \beta^h \varepsilon_t^{MP} + \sum_{j=2}^{4} \gamma^j \mathbb{1}\{q_{t+h} = j\} + e_{t+h}$, using monetary policy shocks ε_t^{MP} implied by the Proxy SVAR in Gertler and Karadi (2015). The shocks are scaled to reduce the 1-year Treasury rate by 25 basis points. The shaded areas are the 90% confidence intervals constructed using standard errors that are robust to heteroskedasticity and autocorrelation. Sample: 1986Q1 - 2018Q4.



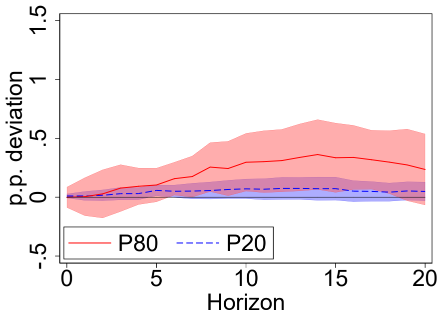
(a) 75th & 25th (Small Firms)



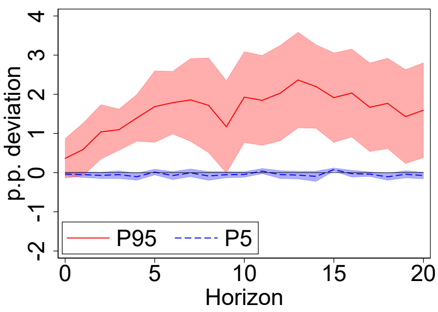
(b) 75th & 25th (Large Firms)



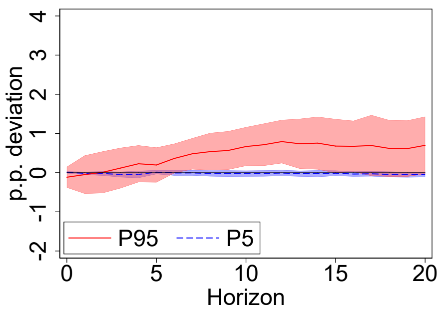
(c) 80th & 20th (Small Firms)



(d) 80th & 20th (Large Firms)



(e) 95th & 5th (Small Firms)



(f) 95th & 5th (Large Firms)

Figure 2.D.4. Effect on Quantiles of Size-Group-Specific Inv. Rate Distributions

Notes: This figure plots the effect of a monetary policy shock on quantiles of the size-specific investment rate distributions. Small (large) firms are firms smaller (larger) than the median in a given quarter. The lines represent the estimated $\hat{\beta}^h$ from separate regressions: $y_{t+h} - y_{t-1} = \alpha^h + \beta^h \varepsilon_t^{MP} + \sum_{j=2}^{4} \gamma^j \mathbb{1}\{q_{t+h} = j\} + e_{t+h}$, using monetary policy shocks ε_t^{MP} implied by the Proxy SVAR in Gertler and Karadi (2015). The shocks are scaled to reduce the 1-year Treasury rate by 25 basis points. The shaded areas indicate the 90% confidence intervals constructed using standard errors that are robust to heteroskedasticity and autocorrelation. Sample: 1986Q1 - 2018Q4.

Appendix 2.E Quantitative Model Appendix

2.E.1 Equilibrium Definition

A recursive competitive equilibrium in this model is a set of value functions $\{V_t(z,k)$, $CV_t^{exit}(z,k)$, $CV_t^a(z,k,\xi)$, $CV_t^n(z,k)\}$, policy functions $\{n_t^*(z,k)$, $k_t^*(z,k,\xi)$, $\xi_t^T(z,k)\}$, quantities $\{C_t, Y_t, I_t^Q, K_t, N_t\}$, prices $\{p_t, w_t, \pi_t, \Lambda_{t+1}, q_t\}$, and distributions $\{\mu_t(z,k)\}$ such that all agents in the economy behave optimally, the distribution of firms is consistent with decision rules, and all markets clear:

(1) Investment Block: Taking all prices as given, $V_t(z,k)$, $CV_t^{exit}(z,k)$, $CV_t^a(z,k,\xi)$, and $CV_t^n(z,k)$ solve the Bellman equation with associated decision rules $n_t^*(z,k)$, $k_t^*(z,k,\xi)$, and $\xi_t^T(z,k)$.

(2) Household Block: Taking prices as given, C_t and C_{t+1} satisfy the household's optimality conditions (2.37) and (2.38).

(3) New Keynesian Block: The New Keynesian Phillips Curve holds. The Taylor rule holds. Taking prices a given, I_t^Q satisfies (2.33).

(4) All markets (final good, capital, labor) clear.

(5) The distribution of firms, $\mu_t(z,k)$, evolves as implied by the decision rules $k^*(z,k,\xi)$ and $\xi_t^T(z,k)$, the exogenous process for firm-level productivity, and considering exogenous exits and entrants with capital k_0 and productivity from μ^{ent}.

References

Adrian, Tobias, Nina Boyarchenko, and Domenico Giannone. 2019. "Vulnerable Growth." *American Economic Review* 109 (4): 1263–89. [50, 59]

Azzalini, A., and A. Capitanio. 2003. "Distributions Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew *t*-distribution." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65: 367–89. [58]

Bachmann, Rüdiger, and Christian Bayer. 2014. "Investment Dispersion and the Business Cycle." *American Economic Review* 104 (4): 1392–416. [55, 88, 89]

Bachmann, Rüdiger, Ricardo J Caballero, and Eduardo M R A Engel. 2013. "Aggregate Implications of Lumpy Investment: New Evidence and a DSGE Model." *American Economic Journal: Macroeconomics* 5 (4): 29–67. [53, 54, 70, 91]

Baley, Isaac, and Andrés Blanco. 2021. "Aggregate dynamics in lumpy economies." *Econometrica* 89 (3): 1235–64. [53]

Belo, Frederico, Xiaoji Lin, and Santiago Bazdresch. 2014. "Labor hiring, investment, and stock return predictability in the cross section." *Journal of Political Economy* 122 (1): 129–77. [88]

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry. 2018. "Really Uncertain Business Cycles." *Econometrica* 86 (3): 1031–65. [75]

Caballero, Ricardo J, and Eduardo MRA Engel. 1999. "Explaining investment dynamics in US manufacturing: a generalized (S, s) approach." *Econometrica* 67 (4): 783–826. [53, 54]

Caballero, Ricardo J, Eduardo MRA Engel, and John C Haltiwanger. 1995. "Plant-level adjustment and aggregate investment dynamics." *Brookings Papers on Economic Activity,* 1–54. [53, 54]

Christiano, Lawrence J, Martin Eichenbaum, and Charles L Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113 (1). [53, 77]

Clementi, Gian Luca, and Berardino Palazzo. 2015. "On The Calibration of Competitive Industry Dynamics Models." *Mimeo, New York University.* [75]

Clementi, Gian Luca, and Berardino Palazzo. 2016. "Entry, exit, firm dynamics, and aggregate fluctuations." *American Economic Journal: Macroeconomics* 8 (3): 1–41. [69]

Cloyne, James, Clodomiro Ferreira, Maren Froemel, and Paolo Surico. 2020. "Monetary Policy, Corporate Finance and Investment." *NBER Working Paper No. 25366, National Bureau of Economic Research.* [49–54, 60, 63, 88, 89, 94]

Cooper, Russell, John Haltiwanger, and Laura Power. 1999. "Machine Replacement and the Business Cycle: Lumps and Bumps." *American Economic Review* 89 (4): 921–46. [54]

Cooper, Russell W, and John Haltiwanger. 2006. "On the Nature of Capital Adjustment Costs." *Review of Economic Studies* 73 (3): 611–33. [54, 70]

Crouzet, Nicolas, and Neil R Mehrotra. 2020. "Small and large firms over the business cycle." *American Economic Review* 110 (11): 3549–601. [55]

Dinlersoz, Emin, Sebnem Kalemli-Ozcan, Henry Hyatt, and Veronika Penciakova. 2018. "Leverage over the Life Cycle and Implications for Firm Growth and Shock Responsiveness." *National Bureau of Economic Research Working Paper Series* No. 25226. [55]

Fang, Min. 2022. "Lumpy Investment, Uncertainty, and Monetary Policy." *Mimeo, University of Toronto,* 1–57. [53, 54]

Farre-Mensa, Joan, and Alexander Ljungqvist. 2016. "Do measures of financial constraints measure financial constraints?" *Review of Financial Studies* 29 (2): 271–308. [55]

Fee, C. Edward, Charles J. Hadlock, and Joshua R. Pierce. 2009. “Investment, Financing Constraints, and Internal Capital Markets: Evidence from the Advertising Expenditures of Multinational Firms.” *Review of Financial Studies* 22 (6): 2361–92. [51]

Gertler, Mark, and Simon Gilchrist. 1994. “MONETARY POLICY, BUSINESS CYCLES, AND THE BEHAVIOR OF SMALL MANUFACTURING FIRMS.” *Quarterly Journal of Economics* 109 (2). [49, 51, 53, 54]

Gertler, Mark, and Peter Karadi. 2015. “Monetary Policy Surprises, Credit Costs and Economic Activity.” *American Economic Journal: Macroeconomics* 7 (1): 44–76. [57, 58, 60, 62, 64, 83–85, 90–92, 95–97]

Gourio, Francois, and Anil K Kashyap. 2007. “Investment spikes: New facts and a general equilibrium exploration.” *Journal of Monetary Economics* 54: 1–22. [54, 60]

Hadlock, Charles J., and Joshua R. Pierce. 2010. “New Evidence on Measuring Financial Constraints: Moving Beyond the KZ Index.” *Review of Financial Studies* 23 (5): 1909–40. [51]

House, Christopher L. 2014. “Fixed costs and long-lived investments.” *Journal of Monetary Economics* 68 (1): 86–100. [53]

Jeenas, Priit. 2019. “Firm Balance Sheet Liquidity, Monetary Policy Shocks, and Investment Dynamics.” *Mimeo, Universitat Pompeu Fabra.* [49, 53, 54]

Khan, Aubhik, and Julia Thomas. 2003. “Nonconvex factor adjustments in equilibrium business cycle models: do nonlinearities matter?” *Journal of Monetary Economics* 50 (2): 331–60. [53]

Khan, Aubhik, and Julia Thomas. 2008. “IDIOSYNCRATIC SHOCKS AND THE ROLE OF NONCONVEXITIES.” *Econometrica* 76 (2): 395–436. [53, 69, 75]

Koby, Yann, and Christian Wolf. 2020. “Aggregation in Heterogeneous-Firm Models: Theory and Measurement.” *Mimeo, Princeton University.* [53, 69, 75]

Lee, Hanbaek. 2022. “Striking While the Iron Is Cold: Fragility after a Surge of Lumpy Investments.” *Mimeo, University of Tokyo.* [54]

Mertens, Karel, and Morten O Ravn. 2013. “The dynamic effects of personal and corporate income tax changes in the United States.” *American Economic Review* 103 (4): 1212–47. [90]

Ottonello, Pablo, and Thomas Winberry. 2020. “Financial heterogeneity and the investment channel of monetary policy.” *Econometrica* 88 (6): 2473–502. [49, 50, 53, 54, 69, 74, 75]

Ramey, Valerie A, and Sarah Zubairy. 2018. “Government spending multipliers in good times and in bad: evidence from US historical data.” *Journal of Political Economy* 126 (2): 850–901. [53]

Rauh, Joshua. 2006. “Investment and Financing Constraints: Evidence from the Funding of Corporate Pension Plans.” *Journal of Finance* 61 (1): 33–71. [51]

Reiter, Michael, Tommy Sveen, and Lutz Weinke. 2013. “Lumpy investment and the monetary transmission mechanism.” *Journal of Monetary Economics* 60 (7): 821–34. [53]

Reiter, Michael, Tommy Sveen, and Lutz Weinke. 2020. “Agency costs and the monetary transmission mechanism.” *B.E. Journal of Macroeconomics* 20 (1): 1–11. [53]

Tenreyro, Silvana, and Gregory Thwaites. 2016. “Pushing on a string: US monetary policy is less powerful in recessions.” *American Economic Journal: Macroeconomics* 8 (4): 43–74. [53]

Thomas, Julia. 2002. “Is lumpy investment relevant for the business cycle?” *Journal of Political Economy* 110 (3): 508–34. [53]

Winberry, Thomas. 2021. “Lumpy investment, business cycles, and stimulus policy.” *American Economic Review* 111 (1): 364–96. [53, 74]

Xu, Qiping, and Eric Zwick. 2021. “Tax policy and abnormal investment behavior.” *Journal of Finance.* [92]

Zwick, Eric, and James Mahon. 2017. “Tax Policy and Heterogeneous Investment Behavior.” *American Economic Review* 59 (3): 379–88. [51, 75, 76]

Chapter 3

Market Power and Macroeconomic Fluctuations*

3.1 Introduction

Crises affect firms unequally. Typically, their direct effects concentrate on some subset of firms, as the following examples illustrate. Natural disasters, such as floods or earthquakes, disrupt the production only of firms that are located in a specific region. Shortages of natural gas concern only firms that rely on this particular source of energy in their production process. Financial crises directly affect only firms that rely on external financing to fund their operations. All of these supply disruptions are neither aggregate nor industry-specific. Instead, they affect some firms more than others *within many industries*. I collectively refer to such disruptions as *asymmetric supply shocks*. The current paper investigates the aggregate effects of asymmetric supply shocks. Most importantly, the aggregate consequences are shown to depend on the intensity of competition among firms. A less competitive economy is less resilient to asymmetric supply shocks.

The current paper builds a model with oligopolistic competition and firm heterogeneity in order to study the aggregate effects of asymmetric supply shocks. I show analytically that a lower intensity of competition among firms makes an economy more vulnerable to asymmetric supply shocks. The mechanism relies on the profit-maximizing behavior of firms. When an adverse shock, such as a natural disaster, disrupts the production of a subset of firms, their unharmed competitors as a result

* I would like to thank Christian Bayer, Keith Kuester and Donghai Zhang for their invaluable support and guidance. Moreover, I would like to thank Rüdiger Bachmann, Florin Bilbiie, Jeff Campbell, Dean Corbae, Alessandro Ferrari, Basile Grassi, Thomas Hintermaier, Matthias Hoelzlein, Joachim Jungherr, Julian Schärer as well as seminar participants at the University of Notre Dame, the RTG-2281 Summer School 2022, the University of Bonn and the 1st PhD Workshop on Imperfect Competition in Macroeconomics for helpful comments and suggestions. I thank the DFG for supporting my research as part of the research training group RTG 2281 "The Macroeconomics of Inequality".

face a higher demand for their goods. When these firms have high market power, they find it optimal to primarily raise prices instead of expanding production. In contrast, when these firms have low market power, they primarily raise production, not prices, and thereby help to stabilize aggregate output. I calibrate the model to the U.S. economy and find that the welfare costs of asymmetric supply shocks increase substantially when the intensity of competition falls. This finding is particularly concerning in view of the rise in market power documented by De Loecker, Eeckhout, and Unger (2020). I also derive implications for competition policy. The main mechanism implies that fostering competition among firms not only reduces markups but also stabilizes the economy. Finally, I test the mechanism in firm-level as well as time-series data. The evidence supports the main mechanism: I find that higher markups are associated with higher volatility.

More in detail, to investigate the aggregate effects of asymmetric supply shocks, I study model environments with two essential features. First, firms have market power and compete strategically within narrow industries. Second, competing firms are heterogeneous and there are asymmetric supply shocks. These shocks affect one or more firms differently than one or more other firms within many industries and thereby change the distribution of sales across firms. To introduce the first essential model feature—strategic competition among firms—I build on the oligopoly framework of Atkeson and Burstein (2008). There is a large number of industries and, in each of them, a small number of firms that produce differentiated goods. Due to the limited number of firms in each industry, firms have market power and interact strategically. Their degree of market power, and thus their profit-maximizing production decision, depends on the intensity of competition. For example, when there are few competitors in an industry, each firm produces little and sells at a high markup, because consumers have few alternative products to choose from.

The second essential model feature—firm heterogeneity—is introduced in a simple and tractable form for the main analysis. That is, there are only two types of firms, active and inactive ones. The total number of firms in each industry is constant, but the share of active firms fluctuates over time. Exogenous changes in this share of active firms constitute asymmetric supply shocks. Since all active firms are identical, a decrease in the number of active firms constitutes a change in the distribution of sales because fewer firms produce larger amounts each. Intuitively, changes in the share of active firms can be interpreted as the result of regional shocks, such as natural disasters or strikes. Each region is home to some firms of each industry. Therefore, when some region is hit by a natural disaster, firms located there shut down, and the share of active firms in each industry falls.

In this framework, I derive analytically the aggregate effects of asymmetric supply shocks, i.e. changes in the share of active firms, in partial equilibrium. I show that a given change in the share of active firms has larger effects on aggregate output and the aggregate markup when the intensity of competition is low, i.e. the number of firms is low. The reason is the profit-maximizing behavior of the remaining active firms, which suddenly face higher demand and have more market power. A reduction in the number of active firms from 4 to 3 (i.e. by 25%) gives the remaining firms substantially more market power, because their market shares rise significantly (0.25 to 0.33). Therefore, they raise prices substantially and expand production relatively little. As a result, the aggregate markup rises and aggregate output falls substantially. In contrast, a reduction in the number of active firms from 40 to 30 (i.e. also by 25%), gives the remaining firms only a small increase in market power, because their market shares rise only slightly (0.025 to 0.033). Thus, they barely increase prices and primarily expand production. Thereby, they help to stabilize aggregate output. This result is reminiscent of Gabaix (2011), even though the mechanism is distinct. Gabaix (2011) shows that the aggregate effects of firm-specific shocks dissipate when the number of firms becomes very large and individual firms become very small. This is intuitive as a shock to 1 firm out of 4 firms can be expected to have a larger aggregate effect than a shock to 1 firm out of 40 firms. In contrast, I consider shocks that affect a given share of firms (e.g. 1 out of 4 or 10 out of 40) and therefore have the same direct effect irrespective of the number of firms. The aggregate effects nonetheless depend on the number of firms because the change in the market structure does.

A corollary to the main result is that when the number of firms becomes very large (high intensity of competition), asymmetric supply shocks become irrelevant for aggregate outcomes. This finding connects to the irrelevance of firm heterogeneity in models without market power but decreasing returns to scale, as discussed in Khan and Thomas (2008), Winberry (2021), and Koby and Wolf (2020). In these frameworks, firm heterogeneity is irrelevant when the technological returns to scale—governed by an exogenous parameter—are close to constant. In both cases, the irrelevance follows from firm profit functions that are close to linear. The important difference is that the number of firms (intensity of competition), which determines the relevance of asymmetric supply shocks in the framework used in this paper, is not policy-invariant and can in principle be affected by competition policy.

Next, I turn to the general equilibrium effects of asymmetric supply shocks. I estimate their welfare costs to a representative household, which I find to increase exponentially when the intensity of competition falls, in line with the partial equilibrium results. Moreover, I decompose the total welfare costs into two components. First, asymmetric supply shocks cause fluctuations in consumption and labor and thereby reduce welfare for the risk-averse household. Second, asymmetric supply shocks further reduce welfare by bringing average consumption below steady-state consumption. This happens because output is a concave function of the number of active firms. As both cost components ultimately result from the market power of firms, both become larger when the intensity of competition decreases.

Thereafter, I investigate optimal competition policy in the face of asymmetric supply shocks, assuming that a government authority chooses the intensity of competition. While I have extensively discussed the benefits of a higher number of firms, so far the model did not include a cost to a higher number of firms. Therefore, I now assume that each firm—active or not—incurs a per-period operating cost, similar to Jaimovich and Floetotto (2008). It is straightforward to see that optimal competition policy depends on the volatility of asymmetric supply shocks. When competition policy takes their presence into account, it optimally prescribes a higher number of firms and thereby makes consumption both higher on average and more stable.

The simple form of firm heterogeneity with active and inactive firms is useful to derive and illustrate the aggregate effects of asymmetric supply shocks. However, I emphasize that the main mechanism applies to a much broader class of models with some form of firm heterogeneity and supply disruptions that can be considered asymmetric supply shocks, because they change the distribution of sales across firms within industries. In particular, there are many firm heterogeneity frameworks that make some firms within an industry more exposed to certain fluctuations than other firms. First, when firms in an industry are spread across several regions, region-specific shocks, such as natural disasters, strikes, or country-specific productivity shocks as in Atkeson and Burstein (2008), reallocate sales across regions and thus firms. Second, when firms in an industry use different inputs or have different production functions, they are differently exposed to changes in the price or the availability of inputs. Shortages of natural gas immediately only affect firms that use natural gas, instead of oil, as a source of energy. Moreover, firms with a relatively labor-intensive production process are more exposed to wage changes than relatively capital-intensive competitors. Third, in models with financial frictions, such as Khan and Thomas (2013) or Ottonello and Winberry (2020), financially constrained firms are more exposed to aggregate shocks than financially unconstrained firms, because aggregate shocks affect the tightness of financial constraints. Therefore, a financial tightening reallocates sales from constrained to unconstrained firms. In addition, frameworks with endogenous entry and exit, such as Bilbiie, Ghironi, and Melitz (2012), or with changes in the distribution of idiosyncratic shocks, e.g. due to volatility shocks as in Bloom (2009) or skewness shocks as in Salgado, Guvenen, and Bloom (2019), give rise to changes in the distribution of sales and thus asymmetric supply shocks.

Finally, I provide empirical evidence which supports the main mechanism. The key insight from the model analysis is that when the intensity of competition among firms is low, markups are not only *high*, but also *volatile*. According to the model, this positive relationship between the level and the volatility of markups holds at the firm-level, at the industry-level, and at the aggregate level. I test this prediction in firm-level micro data from Compustat as well as in aggregate time-series data. I investigate the volatility of prices (markups) instead of the volatility of quantities (output) for two reasons. First, while a low intensity of competition unambiguously predicts a higher volatility of industry-level and aggregate output, there is no clear prediction regarding the volatility of firm-level output. Second, output is also affected by symmetric supply and demand shocks, while markups are not.

Building on the work of De Loecker and Warzynski (2012), De Loecker, Eeckhout, and Unger (2020) and Baqaee and Farhi (2020), I estimate annual firm-level markups for firms included in Compustat. Using these estimates, I show that there is a positive correlation between the firm-level average markup and the firm-level markup volatility, as predicted by the model. However, according to the model, this relationship is not linear, but convex. To test this relationship by means of OLS, I therefore derive a linear relationship. In particular, I show that the model predicts a *linear* relationship between the level and the volatility of the *inverse* markup. Across a range of empirical specifications, I find support for this relationship in the micro data.

Moreover, I assess the relationship between the level of the markup and its volatility in aggregate time-series data. I employ the widely-used model of Smets and Wouters (2007) as a "measurement device" in order to obtain a quarterly time series of the aggregate (target) markup. In the model of Smets and Wouters (2007), this variable evolves over time due to exogenous price-markup shocks for which asymmetric supply shocks can provide a micro-foundation. These shocks explain around 7.4% of fluctuations in consumption and therefore provide a quantitatively relevant source of aggregate volatility. The estimated time series of the aggregate markup captures the recent increase in markups documented in firm-level data (De Loecker, Eeckhout, and Unger, 2020). However, the series shows that the aggregate markup has not only been high in recent years, but was also high during the 1970s. Both periods of a high aggregate markup coincide with a high volatility of the aggregate markup, supporting the main mechanism of the model.

Related Literature. This paper relates to several strands of the literature. The shocks that I introduce—asymmetric supply shocks—relate to the literature on the supply-side origins of aggregate fluctuations. Early contributions have proposed aggregate (e.g. Kydland and Prescott, 1982) and sector-specific (e.g. Long and Plosser, 1983) supply shocks, typically to productivity, as drivers of business cycles. More recently, firm-specific shocks have been proposed as a source of aggregate fluctuations (e.g. Gabaix, 2011). Motivated by the COVID-19 crisis, Guerrieri et al. (2022) study shocks to a subset of sectors of the economy. I add to this literature by introducing and studying *shocks to a subset of firms within many industries*, referred to as asymmetric supply shocks. In contrast to aggregate shocks, sector-specific shocks, and shocks to a subset of sectors, asymmetric supply shocks affect the market structure within industries (sectors[1]), because they do not affect all firms in an industry

1. The terms "sector" and "industry" both describe a group of firms that operate in the same segment of the economy. A "sector" typically describes a large segment of the economy, e.g. the manufacturing sector, while an "industry" typically refers to a smaller, more specific group of firms. Since

symmetrically. In contrast to firm-specific shocks, asymmetric supply shocks affect firms in many industries. Moreover, they affect a given share of firms, not a given number of firms (e.g. one firm), within an industry. Therefore, asymmetric supply shocks do not vanish by a law of large numbers when the number of firms becomes very large and individual firms become very small.

The key feature of asymmetric supply shocks is that they change the market structure in many industries. Therefore, this paper relates to a growing literature on the implications of imperfect competition among firms and endogenous market structure for macroeconomic outcomes. One strand of this literature has focused on long-run trends. De Loecker, Eeckhout, and Unger (2020) and Covarrubias, Gutiérrez, and Philippon (2020) document substantial increases in markups, industry concentration, and profit rates in the United States over the past decades and discuss their macroeconomic implications. Another strand of this literature has focused on the interaction of aggregate shocks with the market structure. Jaimovich and Floetotto (2008) and Bilbiie, Ghironi, and Melitz (2012) show how firm entry and exit amplifies the aggregate effects of productivity shocks in frameworks with endogenous markups. Mongey (2021) finds a larger degree of monetary non-neutrality under oligopolistic competition than under monopolistic competition in a dynamic setting with price rigidities. Burstein, Carvalho, and Grassi (2020) extend the granular macroeconomic model of Gabaix (2011) to oligopolistic competition and show that variable markups dampen the aggregate effects of idiosyncratic shocks.

More recently, attention has shifted to the intersection of the two strands, i.e. the implications of the current market structure and the level of markups and concentration for the amplification of aggregate shocks. Wang and Werning (2022) find that higher industry concentration leads to a larger degree of monetary non-neutrality in a framework with price rigidities. Ferrari and Queirós (2022) show that the amplification of aggregate productivity shocks via firm entry and exit is stronger when idiosyncratic productivity is more dispersed, because more firms are close to the entry (exit) threshold. Jaimovich and Floetotto (2008) and Corhay, Kung, and Schmid (2020) also study aggregate productivity shocks and their amplification via endogenous entry and exit. They find that higher markups are associated with higher aggregate volatility, because of the convex relationship between markups and the number of homogeneous firms. I add to this literature by showing that higher markups lead to larger aggregate effects of asymmetric supply shocks, which arise in a broad class of models with firm heterogeneity. Thus, higher markups are associated with higher aggregate volatility even in the absence of changes in the number of firms.

Any shock that changes the distribution of sales among firms in an industry is also asymmetric and can thus be considered an asymmetric supply shock. Therefore, this paper has implications for a large body of work investigating the transmission

I am interested in the strategic interaction of small groups of firms, I discuss “industries” throughout this paper.

of aggregate shocks in models with some form of firm heterogeneity. For example, the presence of financial frictions in the models of Khan and Thomas (2013), Khan, Senga, and Thomas (2016), and Ottonello and Winberry (2020) implies that any aggregate shock propagates asymmetrically. As a result, the mechanism highlighted in this paper becomes relevant as soon as the assumption of a continuum of firms that do not interact strategically—a common simplification in models with firm heterogeneity—is dropped.

Finally, this paper relates to the literature on the welfare costs of markups and optimal competition policy in macroeconomic models. Bilbiie, Ghironi, and Melitz (2019) discuss the optimal number of varieties in a model with endogenous product creation and monopolistic competition, as well as how to implement the optimal allocation using taxes on consumption and dividends. Edmond, Midrigan, and Xu (2018) quantify the costs of markups and investigate the welfare consequences of a variety of subsidies in a model with firm heterogeneity and endogenous entry. Boar and Midrigan (2019) characterize optimal product market policy in an economy in which firms with market power are owned by a subset of heterogeneous households. I add to this literature by showing that competition increases welfare by reducing aggregate volatility and quantifying the implications for optimal competition policy.

Organization. The remainder of this paper is organized as follows. Section 3.2 presents the main model with oligopolistic competition, firm heterogeneity, and asymmetric supply shocks. Section 3.3 demonstrates analytically and quantitatively that the aggregate effects of asymmetric supply shocks are larger when the intensity of competition among firms is lower. Moreover, optimal competition policy is discussed. In Section 3.4, I provide empirical evidence in support of the main mechanism using firm-level data as well as aggregate time-series data. Section 3.5 concludes.

3.2 Model

In this section, I build a general equilibrium model with oligopolistic competition and firm heterogeneity. The purpose of the model is to study the aggregate effects of asymmetric supply shocks.

The core of the model is a supply side with two main features. First, firms have market power and compete strategically within narrow industries. Building on the framework of Atkeson and Burstein (2008), there is a large number of industries and, in each of them, a small number of firms. Second, firms are heterogeneous and there are shocks that change the distribution of sales across firms within industries. These shocks are referred to as asymmetric supply shocks because they affect one or more firms differently than one or more other firms.

In Section 3.2.1, I describe a simple and tractable industry setup that introduces firm heterogeneity and asymmetric supply shocks in a parsimonious manner. Despite its simplicity, this setup suffices to illustrate the main results in Section 3.3. In Section 3.2.2, I explain the broader class of firm heterogeneity setups to which the main results apply. Thereafter, I integrate the simple industry setup into the larger supply-side structure in Section 3.2.3 and explain how firms interact strategically. Finally, the representative household, which constitutes the intentionally simplistic demand side of the model, is presented in Section 3.2.4.

3.2.1 Simple Industry Setup

There exists a large number of industries j and within each industry, there are $\widetilde{N}_j$ firms, which are indexed by $i \in \{1, ..., \widetilde{N}_j\}$. Each firm ij produces an intermediate good y_{ij} according to a constant-returns-to-scale production technology

$$y_{ijt} = z_{ijt} l_{ijt} \tag{3.1}$$

where z_{ijt} is a firm-specific component and l_{ijt} is the labor input. Firms are heterogeneous due to the firm-specific component, which is a binary variable, i.e. $z_{ijt} \in \{0, 1\}$. Thus, there are only two types of firms. Firms with $z_{ijt} = 0$ have a labor productivity of 0 and therefore optimally shut down in period t. Hence, I will refer to firms as *active* ($z_{ijt} = 1$) and *inactive* ($z_{ijt} = 0$). The share of active firms is λ_t, such that the number of active firms in industry j in period t is

$$N_{jt} = \lambda_t \widetilde{N}_j \tag{3.2}$$

The share of active firms, λ_t, fluctuates over time. These fluctuations in λ_t constitute the *asymmetric supply shocks* in this simple setup. Regardless of the remaining features of the economy, the equilibrium distribution of sales across firms within industry j changes when λ_t changes. As an example, consider an industry with $\widetilde{N}_j = 4$ firms. When three firms are active, the equilibrium distribution of sales shares must be $\{⅓, ⅓, ⅓, 0\}$, because all active firms are identical. With only two active firms, the equilibrium distribution of sales shares is $\{½, ½, 0, 0\}$.

Many supply disruptions may force some firms in each industry to temporarily shut down and thus serve as a micro-foundation for changes in λ_t. Regional shocks, such as natural disasters or strikes, provide an intuitive example. If the $\widetilde{N}_j$ firms are distributed equally across a number of regions and regions are hit by regional shocks from time to time, λ_t reflects the share of undisrupted and active regions. A low value of λ_t means that more regions, and thus firms, than usual are inactive.

3.2.2 Overview of Asymmetric Supply Shocks

Asymmetric supply shocks are defined as shocks which—within many industries—affect one or more firms differently than one or more other firms and thereby change the distribution of sales across firms. In the previous subsection, I have presented regional shocks with active and inactive firms as a simple and tractable example.

However, there exists a fairly broad class of firm heterogeneity frameworks that give rise to supply disruptions that can be considered asymmetric supply shocks. In this subsection, I organize and discuss some of these frameworks. In contrast to the simple setup, many of these examples feature heterogeneity among active firms. In Appendix 3.A.1, I therefore present a generalized industry setup that does not impose restrictions on z_{ijt} and thus allows for heterogeneity among active firms. The main results, derived using the simple setup in Section 3.3, are shown to hold in the generalized setup in Appendix 3.A.2.

Heterogeneous Exposure. Many firm heterogeneity frameworks make some firms within an industry more exposed to certain disruptions than other firms. Three groups of examples appear particularly relevant. First, when firms in an industry are distributed across several regions, region-specific shocks, such as natural disasters, strikes, or regional lockdowns, reallocate sales across regions and thus firms. In a framework with multiple countries, such as Atkeson and Burstein (2008), country-specific productivity shocks also belong to this category. Second, when firms in an industry use different inputs or have different production functions, they are differently exposed to changes in the price or the availability of inputs. Shortages of natural gas immediately only affect firms that use natural gas, instead of oil, as a source of energy. Lockdowns in some part of the world, such as China, only affect firms getting their inputs from this particular region. Moreover, firms with a relatively labor-intensive production process are more exposed to wage changes than relatively capital-intensive competitors. Third, firm-level frictions, in particular financial frictions as in Khan and Thomas (2013), Khan, Senga, and Thomas (2016), and Ottonello and Winberry (2020), make firms differently exposed to aggregate shocks. Financial shocks to the tightness of borrowing constraints as in Khan and Thomas (2013) or Khan, Senga, and Thomas (2016) immediately affect only firms which are “financially constrained”, in contrast to “financially unconstrained” firms. A financial tightening would thus reallocate production and sales from constrained to unconstrained firms. Other aggregate shocks, such as monetary policy shocks in Ottonello and Winberry (2020), endogenously change the tightness of borrowing constraints and therefore set in motion the same mechanism.

Idiosyncratic Shocks. Idiosyncratic shocks to productivity, demand, capital quality, or some other firm-level state variable can be considered a special case of asymmetric supply shocks. Idiosyncratic shocks reallocate market shares between the firm facing the idiosyncratic shock and all other firms in the industry which are not directly affected. However, idiosyncratic shocks only matter for *aggregate* outcomes when firms are not atomistic, e.g. as in the setup of Burstein, Carvalho, and Grassi (2020) with a finite number of industries. In contrast, when there is a continuum of industries, as in Atkeson and Burstein (2008), idiosyncratic shocks “wash out” and do not have aggregate effects. Yet, shocks to the distribution of these idiosyncratic shocks still do have aggregate effects, because they change the distribution of sales in all industries. Examples of these asymmetric supply shocks include shocks to the dispersion (e.g. Bloom, 2009; Bachmann and Bayer, 2014; Ferrari and Queirós, 2022) or skewness (e.g. Salgado, Guvenen, and Bloom, 2019) of idiosyncratic shocks.

Extensive Margin. Closely related to the simple setup is a class of models with homogeneous active firms and endogenous fluctuations in the number of active firms due to endogenous firm entry and exit (e.g. Jaimovich and Floetotto, 2008; Bilbiie, Ghironi, and Melitz, 2012; Corhay, Kung, and Schmid, 2020). In these frameworks, aggregate shocks, e.g. to aggregate productivity, affect firm entry and exit decisions and therefore the number of active firms. Thus, an otherwise perfectly symmetric aggregate shock becomes an asymmetric supply shock due to its propagation via endogenous entry and exit. A similar mechanism is at work in models which feature firms that endogenously choose the number of industries to enter or markets to serve (e.g. Sedláček and Sterk, 2017). Symmetric aggregate shocks now affect how many markets any firm serves, and therefore the number of active firms in any market. Again, there is an asymmetric propagation of otherwise symmetric shock.

In sum, what all of these examples have in common is that there is some form of firm heterogeneity and a shock that changes the distribution of sales within industries. Of course, many of the aforementioned disruptions, such as financial shocks, not only have asymmetric effects on firms, but also symmetric effects (e.g. lower aggregate demand). However, I focus on the analysis of the asymmetric effects and therefore investigate the aggregate effects of asymmetric supply shocks in Section 3.3. Beforehand, I integrate the simple industry setup into the larger supply-side structure.

3.2.3 Supply Side

To study the aggregate effects of asymmetric supply shocks, I need a framework that not only features firm heterogeneity and asymmetric supply shocks but also firms that have market power and interact strategically. Therefore, I integrate the simple industry setup of Section 3.2.1 into the oligopolistic competition framework of Atkeson and Burstein (2008). The simple setup suffices to illustrate the main results regarding the aggregate effects of asymmetric supply shocks. An analysis of other asymmetric supply shocks is relegated to Appendix 3.A.

The production side of the economy consists of three layers. There is a competitive final consumption good producer, a continuum of industries, and in each industry a small number of firms producing differentiated intermediate goods.

Consumption Good Production. A competitive final consumption good producer aggregates the industry goods Y_{jt} of a continuum of industries $j \in [0, 1]$ according to

$$Y_t^C = \left[\int_0^1 Y_{jt}^{\frac{\eta-1}{\eta}} \, dj \right]^{\frac{\eta}{\eta-1}} \quad \text{with} \quad \eta > 1 \tag{3.3}$$

where Y_t^C is the quantity of the final consumption good.[2] The parameter η captures the elasticity of substitution *across* industries.

Industry Good Production. The industry good Y_{jt} is an aggregate of the intermediate goods y_{ijt} produced by the N_{jt} active firms in industry j

$$Y_{jt} = N_{jt}^{\frac{1}{1-\rho}} \left[\sum_{i=1}^{N_{jt}} y_{ijt}^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}} \quad \text{with} \quad \rho > 1 \tag{3.4}$$

where the term $N_{jt}^{\frac{1}{1-\rho}}$ neutralizes love of variety effects[3], as in De Loecker, Eeckhout, and Mongey (2021).[4] The parameter ρ captures the elasticity of substitution *within* industries.

Intermediate Good Production. Industries are modeled as introduced in Section 3.2.1. That is, intermediate good firms operate a constant-returns-to-scale production technology $y_{ijt} = z_{ijt} l_{ijt}$, where the firm-specific component $z_{ijt} \in \{0, 1\}$ creates active and inactive firms. The number of active firms, N_{jt}, is the product of the share of active firms, λ_t, and the number of firms, $\widetilde{N}_j$. Since active firms are homogeneous, *in equilibrium* they produce the same amount, i.e. $y_{ijt} = y_{jt} \; \forall \; i \in \{1, ..., N_{jt}\}$, where y_{jt} is the output of any firm in industry j. Combining this insight with equation (3.4), it follows that in equilibrium industry output is $Y_{jt} = N_{jt} y_{jt}$.[5]

Firm Optimization. The objective of intermediate good firms is to maximize profits, d_{ijt}, which are defined by

$$d_{ijt} = \left(\frac{p_{ijt}}{P_t^C} \right) y_{ijt} - w_t l_{ijt} \tag{3.5}$$

2. The price index for the final consumption good is given by $P_t^C = \left[\int_0^1 P_{jt}^{1-\eta} dj \right]^{\frac{1}{1-\eta}}$ where P_{jt} is the price index for industry j.

3. Neutralizing love of variety effects is not necessary for Proposition 3.1, but simplifies the exposition. Without a love of variety, a change in the number of active firms affects market power and markups while leaving measured productivity unchanged. Changes in measured productivity are discussed in Appendix 3.A.

4. The price index for the industry good is given by $P_{jt} = N_{jt}^{\frac{1}{\rho-1}} \left[\sum_{i=1}^{N_j} p_{ijt}^{1-\rho} \right]^{\frac{1}{1-\rho}}$ where p_{ijt} is the price of the intermediate good produced by firm i in industry j.

5. With homogeneous active firms, in equilibrium it must be the case that $p_{ijt} = p_{jt} \; \forall \; i \in \{1, ..., N_{jt}\}$, where p_{jt} is the price of any firm in industry j. Combining this insight with the price index for the industry good yields $P_{jt} = p_{jt}$.

where p_{ijt} is the price charged by firm i in industry j, P_t^C is the price index for the final consumption good, and w_t is the real wage. Firms compete by choosing quantities (Cournot competition[6]) and face the demand curve

$$\frac{p_{ijt}}{P_t^C} = \left(\frac{y_{ijt}}{Y_{jt}}\right)^{-1/\rho} \left(\frac{Y_{jt}}{Y_t^C}\right)^{-1/\eta} N_{jt}^{-1/\rho} \tag{3.6}$$

which results from optimizing behavior of industry and consumption good producers.

Under optimal behavior, firms set a (gross) markup over marginal costs, μ_{ijt}, which depends on the number of active firms in the industry

$$\frac{p_{ijt}}{P_t^C} = \mu_{ijt}(N_{jt})\,\frac{w_t}{z_{ijt}} \tag{3.7}$$

The optimal markup is a function of the number of active firms, because the demand elasticity faced by firm i in industry j, $\epsilon_{ijt}(N_{jt})$, is a function of the number of active firms

$$\mu_{ijt}(N_{jt}) = \frac{\epsilon_{ijt}(N_{jt})}{\epsilon_{ijt}(N_{jt}) - 1} \qquad \text{where} \qquad \epsilon_{ijt}(N_{jt}) = \left[\frac{1}{\eta}\frac{1}{N_{jt}} + \frac{1}{\rho}\left(1 - \frac{1}{N_{jt}}\right)\right]^{-1} \tag{3.8}$$

The demand elasticity is a weighted harmonic average of the elasticity of substitution across industries, η, and the elasticity of substitution within industries, ρ. This reflects that firms compete both *within* industries, where the relevant elasticity of substitution is ρ, and *across* industries, where the relevant elasticity of substitution is η. Firms internalize that their actions affect not only their own demand, but also the demand for the industry good. The weight given to the elasticity of substitution across industries, η, is $\frac{1}{N_{jt}}$, which equals the market share of a single firm in the industry. This reflects that when there are fewer firms, any one firm becomes larger and has a larger influence on industry demand. Therefore, the demand elasticity depends on the number of active firms.

Combining the optimal markup (3.8) with the demand curve (3.6) yields equilibrium firm output

$$y_{ijt} = \mu_{ijt}(N_{jt})^{-\eta} w_t^{-\eta} \frac{Y_t^C}{N_{jt}} \tag{3.9}$$

6. The main result, Proposition 3.1, also holds under Bertrand competition. See Burstein, Carvalho, and Grassi (2020) for a discussion and comparison of Cournot and Bertrand competition in a similar framework.

Aggregation. In equilibrium, all active firms in an industry choose the same output quantity, y_{ijt}, and the same markup, μ_{ijt}. Combining the equation for industry output with the equation for optimal firm-level output (3.9) yields

$$Y_{jt} \;=\; N_{jt} y_{ijt} \;=\; N_{jt}\mu_{ijt}(N_{jt})^{-\eta}\left(\frac{w_t}{Z_t}\right)^{-\eta}\frac{Y_t^C}{N_{jt}} \tag{3.10}$$

Moreover, the industry markup, μ_{jt}, which is defined as the ratio of industry sales to labor payments[7], is equal to the markup of the any active firm

$$\mu_{jt} \;=\; \mu_{ijt}(N_{jt}) \tag{3.11}$$

To keep the model as parsimonious as possible, I assume that all industries $j \in [0,1]$ are homogeneous. That is, the number of firms in each industry j is $\widetilde{N}_j = \widetilde{N}$ and the number of active firms in each industry is $N_{jt} = N_t = \lambda_t \widetilde{N}$. Therefore, in equilibrium, industry output and the industry markup are identical for all j, i.e. $Y_{jt} = Y_t$ and $\mu_{jt} = \mu_t \;\forall\, j \in [0,1]$, where Y_t is the industry output and μ_t the industry markup of any industry.[8] Combining this insight with equation (3.3), it follows that

$$Y_t^C \;=\; Y_t \tag{3.12}$$

Moreover, the aggregate markup, μ_t^C, defined as the ratio of aggregate sales and labor payments, is equal to the industry markup

$$\mu_t^C \;=\; \mu_t \tag{3.13}$$

Finally, it is important to point out that aggregate productivity, *TFP*, in this economy is constant and thus not affected by asymmetric supply shocks:

$$TFP \;=\; \frac{Y_t^C}{L_t} \;=\; 1 \tag{3.14}$$

where $L_t = N_t l_{ijt}$. As discussed in Appendix 3.A, there are of course examples of asymmetric supply shocks that do affect aggregate productivity.

7. Formally, as shown in Burstein, Carvalho, and Grassi (2020), the industry markup, defined as $\mu_{jt} = \frac{(P_{jt}/P_t^C)Y_{jt}}{w_t L_{jt}}$ can be rewritten as a sales-weighted harmonic average of firm markups.

8. With homogeneous industries, in equilibrium $P_{jt} = P_t \;\forall\, j \in [0,1]$, where P_t is the price index of any industry. Combining this insight with the price index for the consumption good yields $P_t^C = P_t$.

3.2.4 Household

There is a representative household that consumes the final consumption good, C_t, supplies labor, L_t, and owns all firms in the economy. The household has Epstein-Zin preferences and maximizes

$$W_t = u(C_t, L_t) + \beta \left(\mathbb{E}_t W_{t+1}^{1-\alpha}\right)^{1/(1-\alpha)} \tag{3.15}$$

where the risk aversion parameter α allows specifying a coefficient of relative risk aversion which differs from the inverse of the intertemporal elasticity of substitution.[9] The period utility function is standard,

$$u(C_t, L_t) = \frac{C_t^{1-\sigma}}{1-\sigma} + \psi \frac{(1-L_t)^{1-\chi}}{1-\chi} \tag{3.16}$$

The household maximizes (3.15) subject to a sequence of budget constraints

$$C_t = w_t L_t + D_t \tag{3.17}$$

where D_t subsumes dividends of all firms. Optimization gives rise to a standard wage-Euler equation

$$C_t^{\sigma} \psi (1-L_t)^{-\chi} = w_t \tag{3.18}$$

3.2.5 Stochastic Process

The only source of aggregate uncertainty are changes in the share of active firms, λ_t. The variable λ_t must remain on the interval $(0, 1]$, such that the number of active firms, N_t, remains above 0 and below $\widetilde{N}$.[10] To implement this, λ_t is the logistic transformation of an otherwise standard AR(1) process

$$\epsilon_t^{\lambda} = (1-\rho_\lambda)\overline{\lambda} + \rho_\lambda \epsilon_{t-1}^{\lambda} + \sigma_\lambda \nu_t \qquad \text{with } \nu_t \sim \mathcal{N}(0,1) \tag{3.19}$$

$$\lambda_t = \frac{1}{1 + e^{-(\epsilon_t^{\lambda} - \overline{\lambda})}} \tag{3.20}$$

where $\overline{\lambda}$ is the steady-state value of λ_t, σ_λ determines the volatility of shocks to λ_t, and ρ_λ their persistence.

9. When $\alpha = 0$, the coefficient of relative risk aversion coincides with the inverse of the intertemporal elasticity of substitution, and Epstein-Zin preferences coincide with standard expected utility preferences.

10. To avoid integer constraints, I assume that the number of active firms, N_t, is a continuous variable. Therefore, λ_t can take any value on the interval $(0, 1]$ and not only values from the set $[1/\widetilde{N}, 2/\widetilde{N}, ..., 1]$. Integer constraints are an issue only in the simple setup. In the generalized setup, outlined in Appendix 3.A.1, the *effective* number of firms is anyways a continuous variable, because active firms can be heterogeneous.

3.3 Model Analysis

I now study the aggregate implications of asymmetric supply shocks. Throughout, I pay particular attention to how the intensity of competition among firms matters for the aggregate effects of these shocks. First, I derive some analytical results in partial equilibrium. Thereafter, I calibrate the model in order to obtain quantitative results in general equilibrium. Finally, I investigate optimal competition policy in the face of asymmetric supply shocks.

3.3.1 Analytical Results

I begin by characterizing analytically the aggregate effects of an asymmetric supply shock, i.e. a change in the share of active firms, λ_t. I do so in partial equilibrium, meaning that the real wage, w_t, and demand for the final consumption good, $Y_t^{C,D}$, are held constant.

First, it is important to notice that the partial equilibrium effects on aggregate output and the aggregate markup are inextricably linked. To see this, consider the following decomposition of the elasticity of aggregate output (supply), $Y_t^{C,S}$, with respect to λ_t:

$$\begin{aligned}
\frac{\mathrm{dlog}(Y_t^{C,S})}{\mathrm{dlog}(\lambda_t)} &= \underbrace{\frac{\mathrm{dlog}(N_t)}{\mathrm{dlog}(\lambda_t)}}_{\text{Direct Effect}} + \left(\underbrace{\frac{\mathrm{dlog}(y_{it})}{\mathrm{dlog}(\mu_{it})} \frac{\mathrm{dlog}(\mu_{it})}{\mathrm{dlog}(N_t)} \frac{\mathrm{dlog}(N_t)}{\mathrm{dlog}(\lambda_t)} - \frac{\mathrm{dlog}(N_t)}{\mathrm{dlog}(\lambda_t)}}_{\text{Spillover Effect}} \right) \\
&= \underbrace{\frac{\mathrm{dlog}(y_{it})}{\mathrm{dlog}(\mu_{it})}}_{-\eta} \frac{\mathrm{dlog}(\mu_{it})}{\mathrm{dlog}(N_t)} \underbrace{\frac{\mathrm{dlog}(N_t)}{\mathrm{dlog}(\lambda_t)}}_{1} \\
&= -\eta \frac{\mathrm{dlog}(\mu_{it})}{\mathrm{dlog}(N_t)}
\end{aligned} \tag{3.21}$$

The (positive) direct effect reflects that aggregate output increases because more firms are active. The (negative) spillover effect reflects that all active firms produce less. The total effect boils down to the elasticity of the firm-level markup with respect to the number of active firms and a constant. The same elasticity governs the effect of the asymmetric supply shock on the aggregate markup:

$$\frac{\mathrm{dlog}(\mu_t^C)}{\mathrm{dlog}(\lambda_t)} = \frac{\mathrm{dlog}(\mu_{it})}{\mathrm{dlog}(N_t)} \underbrace{\frac{\mathrm{dlog}(N_t)}{\mathrm{dlog}(\lambda_t)}}_{1} \tag{3.22}$$

Thus, the effects on the aggregate markup and aggregate output are closely linked, which comes as no surprise, given that both effects are the result of the firm-level price-quantity trade-off. The central elasticity of the firm-level markup with respect to the number of active firms is

$$\frac{\mathrm{dlog}(\mu_{it})}{\mathrm{dlog}(N_t)} = -\frac{\mu_{it}}{N_t}\left(\frac{1}{\eta} - \frac{1}{\rho}\right) \tag{3.23}$$

Under the standard parameter restriction $\rho > \eta$, an increase in the number of active firms decreases the markup.[11] From this it follows that a positive asymmetric supply shock, i.e. an increase in λ_t, increases aggregate output and reduces the aggregate markup:

$$\frac{\mathrm{dlog}(Y_t^{C,S})}{\mathrm{dlog}(\lambda_t)} > 0 \tag{3.24}$$

$$\frac{\mathrm{dlog}(\mu_t^C)}{\mathrm{dlog}(\lambda_t)} < 0 \tag{3.25}$$

Importantly, however, these aggregate effects of asymmetric supply shocks depend on the time-invariant intensity of competition among firms, captured by $\widetilde{N}$. This is because the central elasticity of the markup with respect to the number of active firms (equation 3.23) depends on $\widetilde{N}$, since $N_t = \lambda_t \widetilde{N}$. This observation leads to the main analytical result, summarized in Proposition 3.1.

Proposition 3.1. *In a more competitive economy (higher number of firms $\widetilde{N}$), an asymmetric supply shock has a smaller absolute effect on aggregate output and the aggregate markup:*

Proof.

$$\frac{d\left(\frac{\mathrm{dlog}(\mu_{it})}{\mathrm{dlog}(N_t)}\right)}{d\widetilde{N}} = \frac{\mu_{it}}{N_t}\left(\frac{1}{\eta} - \frac{1}{\rho}\right)\left[1 + \frac{\mu_{it}}{N_t^2}\left(\frac{1}{\eta} - \frac{1}{\rho}\right)\right] > 0$$

From this, it follows that

$$\frac{d\left(\frac{\mathrm{dlog}(Y_t^{C,S})}{\mathrm{dlog}(\lambda_t)}\right)}{d\widetilde{N}} = -\eta\frac{d\left(\frac{\mathrm{dlog}(\mu_{it})}{\mathrm{dlog}(N_t)}\right)}{d\widetilde{N}} < 0$$

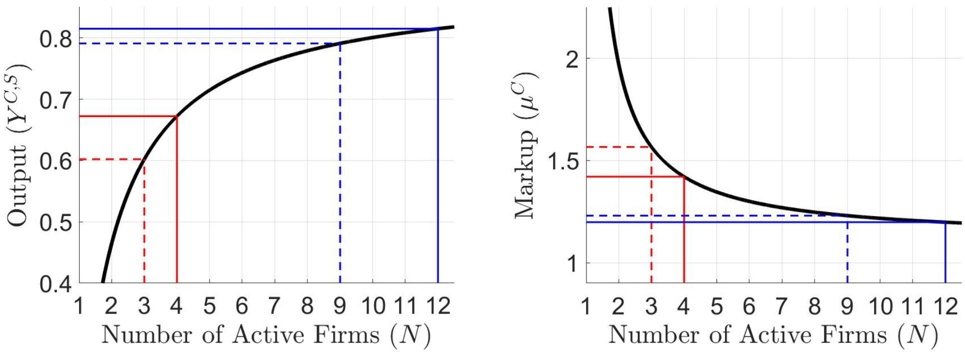
$$\frac{d\left(\frac{\mathrm{dlog}(\mu_t^C)}{\mathrm{dlog}(\lambda_t)}\right)}{d\widetilde{N}} = \frac{d\left(\frac{\mathrm{dlog}(\mu_{it})}{\mathrm{dlog}(N_t)}\right)}{d\widetilde{N}} > 0$$

□

11. This parameter restriction states that the elasticity of substitution is higher within industries than across industries. Intuitively, the consumer is more willing to substitute a Coke and a Pepsi, than a soda and a t-shirt.

Intuition. The intuition for Proposition 3.1 goes as follows. Suppose there is a decrease in the number of active firms by 25% (λ falls from 1 to 0.75). The direct effect is that industry output falls by 25%. However, the remaining 75% of firms now face more demand—both in absolute terms and relative to total industry demand. They respond to this by increasing output and by increasing prices (i.e. markups). The combination of the two depends on the increase in market power that the remaining firms experience, which again depends on the increase in their market share. In an economy with 40 firms per industry, the market share of remaining firms grows from 2.5% to 3.33%, which implies a fairly small increase in market power. In an economy with 4 firms per industry, the market share of the remaining firms grows from 25% to 33.3%, which constitutes a sizeable increase in market power. Therefore, the remaining firms raise their markups by more and output by less than the remaining firms in the economy with 40 firms. In consequence, the total effects on aggregate output and the aggregate markup are larger in the economy with only 4 firms to start with. Figure 3.1 illustrates a very similar example. Panel (a) plots aggregate output ($Y^{C,S}$) as a function of the number of active firms (N). The solid red line depicts an economy with a low intensity of competition, the number of firms being 4. The dashed red line illustrates the fall in output when the number of active firms falls by 25% (λ falls from 1 to 0.75). The solid blue line depicts an economy with a higher intensity of competition, the number of firms being 12. When the number of active firms falls by 25%, as illustrated with the blue dashed line, output falls, but much less than in the economy with a low intensity of competition. This illustrates the main result that a given shock (decrease in λ from 1 to 0.75) has smaller aggregate effects when the intensity of competition is high. Panel (b) plots the aggregate markup as a function of the number of active firms (N). Analogously to panel (a), the same shock has a smaller aggregate effect when the intensity of competition is high.

Irrelevance of Asymmetric Supply Shocks. A straightforward implication of Proposition 3.1 is that when the number of firms, $\widetilde{N}$, gets very large, asymmetric supply shocks become irrelevant for aggregate outcomes. This limit case is the familiar monopolistic competition setup in which firms charge a constant markup of $\mu_{ijt} = \frac{\rho}{\rho-1}$. This result, which is summarized in Corollary 3.1, connects to the literature on aggregation in heterogeneous-firm models, which has shown that when firms are atomistic—a common simplifying assumption—and profit functions become linear, firm-level frictions become irrelevant for aggregate outcomes (Koby and Wolf, 2020; Winberry, 2021). Starting with Khan and Thomas (2008), the focus has been on firm-level capital adjustment costs generating "lumpy investment behavior", but Koby and Wolf (2020) have shown more recently that the irrelevance result also applies to firm-level financial frictions. In these frameworks, curvature in the profit function is governed by the exogenous parameter determining the degree of (decreasing) returns-to-scale. Thus, the curvature of profit functions is exogenous.



(a) Output **(b)** Markup

Figure 3.1. Intuition for Proposition 3.1

Notes: This figure illustrates the aggregate effects of a drop in the number of active firms by 25%. The black lines depict aggregate output (left panel) and the aggregate markup (right panel) as a function of the number of active firms, using $\rho = 10$ and $\eta = 1.13$, which are the parameters used in the quantitative model below. The red lines refer to a low-competition economy with 4 firms, the blue lines refer to a high-competition economy with 12 firms. Solid lines depict the initial state ($\lambda = 1$), and dashed lines depict the state after the drop in the number of active firms ($\lambda = 0.75$).

Corollary 3.1 states that in the framework discussed in this paper, asymmetric supply shocks equally become irrelevant when profit functions become linear. However, as the curvature of profit functions stems from market power, it is not exogenous, but depends on the intensity of competition among firms. This implies that the aggregate relevance of asymmetric supply shocks and firm heterogeneity more generally is not policy-invariant. In Section 3.3.4, I therefore discuss optimal competition policy in the face of asymmetric supply shocks.

Corollary 3.1. *When the number of firms becomes very large ($\widetilde{N} \to \infty$), asymmetric supply shocks become irrelevant for aggregate output and the aggregate markup.*

$$\lim_{\widetilde{N}\to\infty} \frac{dlog(Y_t^{C,S})}{dlog(\lambda_t)} = 0 \quad and \quad \frac{dlog(\mu_t^C)}{dlog(\lambda_t)} = 0 \tag{3.26}$$

Alternative Asymmetric Supply Shocks. As I show in Appendix 3.A, it is straightforward to extend Proposition 3.1 to a generalized industry setup which allows for other sources of firm heterogeneity and different asymmetric supply shocks. All that is necessary is to replace the number of active firms with the *effective number of firms*. Intuitively, the effective number of firms is the number of homogeneous firms that delivers the same industry concentration as a given distribution of heterogeneous firms.[12] Asymmetric supply shocks change the effective number of firms and Proposition 3.2 shows that a given change has larger aggregate effects when the intensity of competition is low.

3.3.2 Calibration

To study the quantitative implications of asymmetric supply shocks in general equilibrium, I calibrate the model to the U.S. economy. The calibrated parameters are summarized in Table 3.1. The parameterization of the household follows Rudebusch and Swanson (2012). The discount factor is set to $\beta = 0.99$, which generates an annual real interest rate close to 4%. The labor disutility parameter, ψ, is chosen such that in steady state the household spends a third of its time endowment working. The curvature of the utility function with respect to consumption is set to $\sigma = 2$, which implies an intertemporal elasticity of substitution (IES) of 0.5. The curvature of the utility function with respect to labor is set to $\chi = 3$, which implies a Frisch elasticity of labor supply of 2/3. The risk aversion parameter is set to $\alpha = -148.3$. The resulting coefficient of relative risk aversion (CRRA) is 75 as in Rudebusch and Swanson (2012).

The steady-state share of active firms is normalized to $\overline{\lambda} = 0.5$. Then, the number of firms per industry is calibrated to $\widetilde{N} = 7.46$, such that the steady-state number of active firms in an industry matches the median effective number of firms in a market calculated in Mongey (2021).[13] The elasticity of substitution within industries is set to $\rho = 10$ as in Atkeson and Burstein (2008), Mongey (2021), and Wang and Werning (2022). The elasticity of substitution across industries is calibrated to $\eta = 1.13$ in order to generate a steady-state (sales-weighted) average markup of 1.45. This roughly corresponds to the average value between 2000 and 2010 according to De Loecker, Eeckhout, and Unger (2020). While being a fairly high markup value, it is still substantially below the latest value of 1.62 reported for the year 2016.[14] Finally, the volatility and persistence of asymmetric supply shocks are calibrated to match the observed fluctuations in the detrended (log) labor share, because in the model the labor share is equal to the inverse of the gross markup. The persistence parameter is set to $\rho_\lambda = 0.95$, which generates an auto-correlation of 0.71. The volatility parameter is set to $\sigma_\lambda = 0.052$, which implies a standard deviation of 1.04%.[15]

12. Formally, the effective number of firms is the inverse of the Herfindahl–Hirschman concentration index.

13. To calculate this number, a market is defined as an IRI product category within a state.

14. There exists a wide range of estimates of the aggregate markup for the U.S. economy. This is due to difficulties in both measuring and aggregating firm-level markups; see Edmond, Midrigan, and Xu (2018), Basu (2019), and De Ridder, Grassi, and Morzenti (2021).

15. To calculate these targets, I use the series “Nonfarm Business Sector: Labor Share for All Employed Persons” (PRS85006173) from FRED. Both the data and the model-generated data are detrended using an HP-filter with smoothing parameter $\lambda = 1600$.

Table 3.1. Calibration

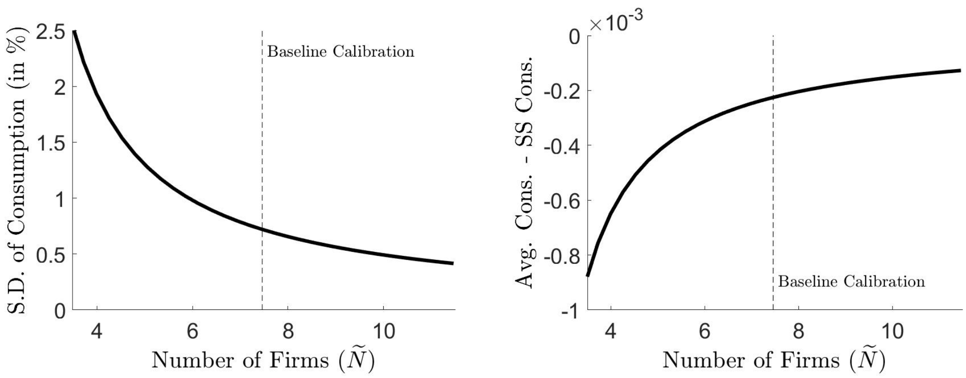
Param.	Description	Value	Target / Source
Household			
β	Discount factor	0.99	$r^{ann} \approx 4\%$
ψ	Labor disutility	1.64	$L_{SS} = 1/3$
σ	Curvature of util. w.r.t. C	2	IES = 0.5
χ	Curvature of util. w.r.t. L	3	Frisch elasticity = 2/3
α	Risk aversion parameter	-148.3	CRRA = 75 (Rudebusch and Swanson, 2012)
Firms			
$\bar{\lambda}$	Share of active firms in SS	0.5	Normalization
$\widetilde{N}$	Number of firms per ind.	7.46	$N_{SS} = 3.73$ (Mongey, 2021)
ρ	Elast. of subst. within ind.	10	Atkeson and Burstein (2008)
η	Elast. of subst. across ind.	1.13	Avg. $\mu = 1.45$ (De Loecker et al., 2020)
ρ_λ	Persist. of fluct. in λ	0.95	ρ(log(Labor Share) $= 0.71$ (detrended)
σ_λ	SD of innovations to λ	0.052	σ(log(Labor Share)) $= 1.04\%$ (detrended)

3.3.3 Welfare Costs of Asymmetric Supply Shocks

It is instructive to decompose the welfare costs of asymmetric supply shocks into two components. First, as shown previously in partial equilibrium, asymmetric supply shocks cause fluctuations in aggregate output and the aggregate markup. In general equilibrium, fluctuations in the aggregate markup cause fluctuations in the real wage (equation 3.7), which again cause fluctuations in labor and consumption (equation 3.18). Since the household is risk-averse, these fluctuations in consumption and labor reduce welfare. I refer to this cost component as the “volatility effect”.

However, asymmetric supply shocks do not only cause aggregate fluctuations, but also affect the *average* state of the economy. To see this, note that aggregate output is a concave function of the number of active firms, as shown in panel (a) of Figure 3.1. Due to this concavity, fluctuations in the number of active firms bring *average* output below *steady-state* output. By the same logic, the average markup exceeds the steady-state markup, because the markup is a convex function of the number of active firms (see panel (b) of Figure 3.1). Since markups are distortionary, steady-state output and consumption are already suboptimally low, and an even lower average consumption level reduces welfare. This cost component is referred to as the “mean effect”.

The Role of Competition. The main result of this paper is that a higher intensity of competition among firms, which in this model is equivalent to a higher number of firms, is welfare-improving because it reduces the costs of asymmetric supply shocks. To support this argument, I first of all inspect both components of the welfare costs in isolation. The left panel of Figure 3.2 shows how the volatility of consumption, constituting the volatility effect, depends on the intensity of competition. Along the x-axis, the intensity of competition, i.e. the number of firms $\widetilde{N}$, changes while all other parameters are held constant. Evidently, when there is less competition (low $\widetilde{N}$), the volatility of consumption resulting from asymmetric supply shocks increases in a convex manner. For example, when the number of firms is reduced by 50%, the volatility of consumption roughly triples. This shows that the main insight from Proposition 3.1—competition makes the economy more resilient—holds in general equilibrium. The right panel of Figure 3.2 shows how the difference between average consumption and steady-state consumption, constituting the mean effect, depends on the intensity of competition. Very similar to the volatility effect, this difference increases in a convex manner when the intensity of competition falls. The similarity is not surprising in light of the fact that both effects are caused by the non-linear relationship between the firm-level markup and the number of firms (equation 3.23).



(a) Volatility Effect

(b) Mean Effect

Figure 3.2. Volatility & Mean Effects by Intensity of Competition

Notes: This figure illustrates how the cost of asymmetric supply shocks depends on the number of firms. The left panel plots the standard deviation of the log of aggregate consumption by the number of firms ("volatility effect"). The right panel plots the difference between average consumption and steady-state consumption by the number of firms ("mean effect"). The dashed vertical lines depict the baseline calibration ($\widetilde{N} = 7.46$).

Finally, Figure 3.3 presents estimates of the welfare costs of asymmetric supply shocks for a range of competition intensities. At the baseline calibration, the household would be willing to give up around 0.035% of steady-state consumption to erase asymmetric supply shocks. Of the total costs, around 69% are due to the mean effect and 31% are due to the volatility effect. These modest numbers both for the total costs as well as for the volatility effect reflect the small costs of business cycles in models with a simple representative household, an observation dating back to Lucas (1987). Adding features such as countercyclical income risk as in Storesletten, Telmer, and Yaron (2001) or high and persistent individual consumption risk as in De Santis (2007) to the household would greatly amplify the costs of business cycles and therefore of asymmetric supply shocks. However, since these features would also complicate the analysis, I refrain from doing so and instead emphasize that the estimated welfare costs most likely present a fairly low lower bound.

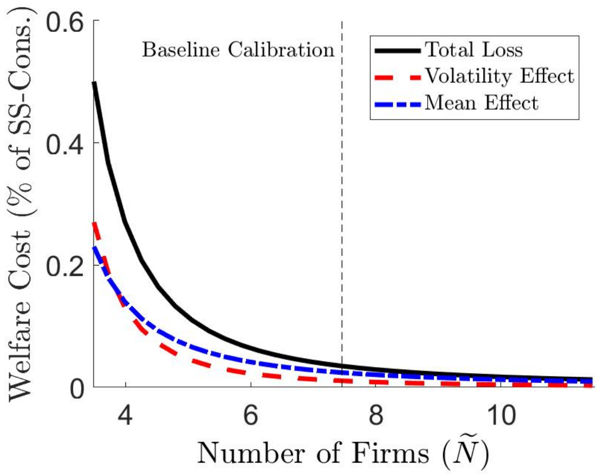


Figure 3.3. Welfare Cost of Asymmetric Supply Shocks by Intensity of Competition

Notes: This figure illustrates how the welfare cost of asymmetric supply shocks depends on the number of firms. The welfare cost is calculated as the share of steady-state consumption that the household would be willing to pay for the removal of asymmetric supply shocks. The black line depicts the total loss, the red line the volatility effect, and the blue line the mean effect. The dashed vertical line depicts the baseline calibration ($\widetilde{N} = 7.46$).

Either way, the main point of this paper is not about the level of the welfare costs of asymmetric supply shocks, but about how the costs vary with the intensity of competition among firms. In line with Proposition 3.1 and the results shown in Figure 3.2, Figure 3.3 confirms that the welfare costs are monotonically decreasing in the number of firms. Note that this calculation does not include welfare gains from a higher or lower intensity of competition in steady state, but only from reducing the costs of asymmetric supply shocks. Two additional features of Figure 3.3 stand out. First, the welfare costs of asymmetric supply shocks are a convex function of the number of firms, reflecting the convexity of the markup itself, as evident from Figure 3.1. That is, increasing the number of firms by 1 reduces the welfare costs by 0.01 percentage points, whereas reducing the number of firms by 1 increases them by 0.017 percentage points. Second, the smaller the number of firms gets, the more important the volatility effect becomes. When the number of firms falls by 50% compared to the baseline calibration, the volatility effect even quantitatively dominates the mean effect.

3.3.4 Optimal Competition Policy

To study optimal competition policy in the face of asymmetric supply shocks, I now add to the model a government competition authority. I assume that the government can choose a time-invariant number of firms in an industry, $\widetilde{N}$, and thereby the intensity of competition among firms. Thus far, there are benefits to competition, but no "cost" of having a high number of firms in the economy. Therefore, the optimal competition policy would simply be $\widetilde{N} = \infty$. To make optimal competition policy less trivial, I henceforth assume that there is a cost to having a high number of firms. Such a cost could arise because firms incur overhead operating costs as in Jaimovich and Floetotto (2008) or because there is firm churn and a sunk entry cost as in Bilbiie, Ghironi, and Melitz (2012).

Costly Firms. I assume that each firm (active or not) incurs a per-period operating cost δ, akin to the framework of Jaimovich and Floetotto (2008), but with two key differences. First, the operating cost is paid for and the number of firms is chosen in a welfare-maximizing manner by the government. Thus, the number of firms is optimal, which is not necessarily the case if entry is a private decision as in Jaimovich and Floetotto (2008). Second, the number of firms is time-invariant and therefore does not respond to shocks. The cost function takes the following functional form:

$$F_N = \delta\left(\widetilde{N} - \widetilde{N}_{oSS}\right) \tag{3.27}$$

The parameter determining the marginal cost of an additional firm is calibrated to $\delta = 0.0015$, such that the baseline steady state is socially optimal *in the absence* of asymmetric supply shocks. Moreover, the subtraction of the number of firms in the baseline steady state, $\widetilde{N}_{oSS}$, sets the total cost to zero in the baseline steady state. Therefore, the steady state in the model with costly firms is exactly the same as the baseline steady state without operating costs. The government runs a balanced budget in each period and finances the operating cost with lump-sum taxes raised from the household

$$T_t = F_N \tag{3.28}$$

I assume that there are no other policy options available, such as a labor subsidy, which could solve the markup distortion entirely (Bilbiie, Ghironi, and Melitz, 2019).

Quantitative Analysis. Absent asymmetric supply shocks, the government trades off the cost of a high number of firms with the static benefit of a high number of firms. Absent shocks, a higher number of firms increases welfare because it reduces markups and thereby the distortion in the household's consumption-labor decision. As explained in Bilbiie, Ghironi, and Melitz (2019) among others, with positive markups, leisure is too cheap and therefore, labor supply and consumption are too low. The presence of asymmetric supply shocks adds two benefits of a high number of firms, as explained in the previous subsection, while leaving the cost of firms unchanged. Therefore, adding shocks must induce the government to choose a higher intensity of competition. Quantitatively, I find that the government increases the number of firms in the economy by 1.1% when asymmetric supply shocks, calibrated as before, are introduced. Thereby, the planner reduces the standard deviation of (log) consumption by 1.45% (volatility effect) and decreases the gap between average and steady-state consumption by 1.53% (mean effect). In addition, steady-state consumption rises by 0.08% which reflects the static benefit of competition. Steady-state output rises by 0.11%. The gap between the changes in consumption and output is due to the higher total operating cost. As discussed above, these numbers should be interpreted as a lower bound given that the model features very low costs of business cycles.

3.4 Empirical Evidence

The main insight from the preceding analysis is that when the intensity of competition among firms is low, markups are not only *high*, but also *volatile*. According to the model, this positive relationship between the level and the volatility of markups holds at the firm level, at the industry level, and at the aggregate level. In this section, I test this prediction in firm-level micro data (Section 3.4.1) as well as aggregate time-series data (Section 3.4.2).

3.4.1 Evidence from Firm-Level Data

The fundamental source of the main insight is the convex relationship at the firm level between the markup and the market share.[16] Therefore, when facing the same shocks, firms with a higher market share not only have a higher markup, but also a more volatile markup. I now turn to firm-level micro data from Compustat to investigate this model prediction.

Throughout, I focus on the volatility of firm-level markups instead of the volatility of firm-level output. The reason is that the model predicts a monotone relationship between the intensity of competition and the volatility of output at the industry level and at the aggregate level, but not at the firm-level.[17] Likewise, there is no monotone relationship between the intensity of competition and the volatility of firm-level sales.[18]

3.4.1.1 Data & Markup Estimation

I use annual firm-level data from Compustat North America. The data treatment is described in detail in Appendix 3.B.1. Markups are estimated according to the

16. In the simple industry setup studied in Section 3.3, the market share only depends on the number of active firms ($s_{ijt} = \frac{1}{N_{jt}}$), which is why the markup is a function of the number of active firms.

17. To see this, note that the elasticity of firm-level output with respect to the number of active firms is $\frac{\mathrm{dlog}(y_{ijt})}{\mathrm{dlog}(N_{jt})} = -\eta \frac{\mathrm{dlog}(\mu_{ijt})}{\mathrm{dlog}(N_{jt})} - 1$, which can be positive or negative. Thus, a change in the intensity of competition, which changes $\frac{\mathrm{dlog}(\mu_{ijt})}{\mathrm{dlog}(N_{jt})}$, can increase or decrease the volatility of firm-level output.

18. The elasticity of firm-level sales with respect to the number of active firms is $\frac{\mathrm{dlog}(p_{ijt}y_{ijt})}{\mathrm{dlog}(N_{jt})} = (1-\eta)\frac{\mathrm{dlog}(\mu_{ijt})}{\mathrm{dlog}(N_{jt})} - 1$, which can be positive or negative.

production approach due to De Loecker and Warzynski (2012), which I summarize in Appendix 3.B.2. Details of the implementation of this estimation procedure are relegated to Appendix 3.B.3. All steps broadly follow De Loecker, Eeckhout, and Unger (2020) and Baqaee and Farhi (2020).

3.4.1.2 Basic Correlations

First of all, I document basic correlations between the level and the volatility of firm-level markups. In the face of the caveats that come with the dataset and the estimation of markups, I primarily use measures that are not sensitive to outliers. As a measure of the level of the markup, I use the median markup of firm i over time. As a measure of the volatility of the markup, I use the interquartile range of markups of firm i.

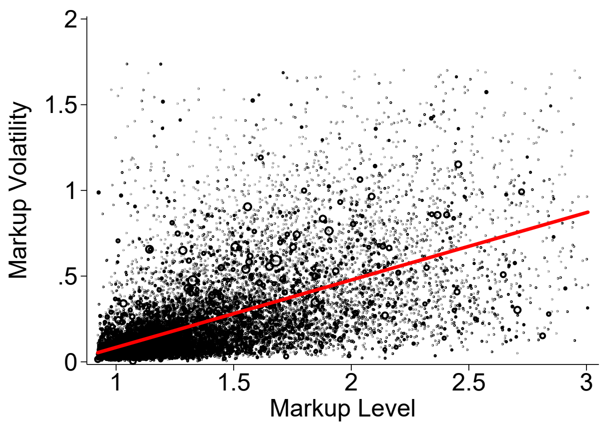


Figure 3.4. Firm-Level Markups & Markup Volatility

Notes: Each circle depicts one firm. The markup level is the median markup. The markup volatility is the interquartile range of the markup. Both variables are trimmed (1%). Firms are weighted by their average sales share. The red line shows the linear fit.

Figure 3.4 displays the relationship between the level and the volatility of markups at the firm level. There is a clear positive relationship, as predicted by the model. Table 3.2 confirms this positive relationship by regressing the interquartile range of markups on the median markup and a constant. There is an economically and statistically significant positive relationship, irrespective of whether firms are weighted by their average economy-wide sales share (column 1) or not (column 2), or whether industry fixed effects are included (column 3). A potential concern might be that firm-specific markup trends induce a correlation between the level and the volatility of the markup. To address this, I compute the volatility after taking out a firm-specific linear trend (column 4) and from changes in markups (column 5). In addition, Figure 3.B.1 and Table 3.B.1 show that the results are robust to using log markups.

Table 3.2. Firm-Level Markups & Markup Volatility

	(1) IQR (μ)	(2) IQR (μ)	(3) IQR (μ)	(4) IQR (μ)	(5) IQR ($d\mu$)
Median (μ)	0.393***	0.392***	0.375***	0.282***	0.173***
	(0.025)	(0.007)	(0.025)	(0.024)	(0.015)
Constant	-0.309***	-0.278***	-0.318***	-0.219***	-0.133***
	(0.030)	(0.009)	(0.058)	(0.029)	(0.018)
Observations	12282	12282	12282	12257	12259
R^2	0.386	0.307	0.456	0.344	0.357
Weights	Yes	No	Yes	Yes	Yes
Industry FE	No	No	Yes	No	No
Linear Trend	No	No	No	Yes	No

Notes: Each column displays coefficients from a separate regression: $Volatility_i(\mu_{it}) = \beta_0 + \beta_1 * Level_i(\mu_{it}) + \varepsilon_i$. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All variables are trimmed (1%).

3.4.1.3 Testing the Model

The model not only predicts that the volatility of the firm-level markup increases with the level of markup, but also that this relationship has a particular non-linear (convex) shape. Therefore, to properly test this relationship in the data, I derive from the model a linear relationship, which can then be estimated by OLS, between a measure of the volatility and a measure of the level of the markup.

While the markup itself is a non-linear function of the market share (s_{ijt})[19], the inverse of the markup is a linear function thereof

$$\mu_{ijt}^{-1} = \frac{\rho - 1}{\rho} - \frac{\frac{\rho}{\eta} - 1}{\rho} s_{ijt} \tag{3.29}$$

This model equation cannot be estimated in Compustat data, because no information on the sales share in the relevant market (s_{ijt}) is available.[20] Burstein, Carvalho, and Grassi (2020) present evidence for this relationship in French administrative data, however.

The market share fluctuates around some long-run average ($\overline{s_{ij}}$)

$$s_{ijt} = \overline{s_{ij}} \epsilon_{ijt} \tag{3.30}$$

19. For the derivation of this equation in the generalized setup, see Appendix 3.A, in particular equations (3.A.2) and (3.A.3).

20. The relevant market share would be the market share within a narrowly defined industry. Of course, it is possible to assign Compustat firms to broad sectors, as done in Baqaee and Farhi (2020), but a sector is too broad to constitute a market, in which firms interact strategically.

due to asymmetric supply shocks (ϵ_{ijt}).[21] It follows from these two equations that the standard deviation of the inverse markup, $\sigma(\mu_{ijt}^{-1})$, can be written as

$$\sigma(\mu_{ijt}^{-1}) = \frac{\frac{\rho}{\eta} - 1}{\rho} \overline{s_{ij}} \sigma(\epsilon_{ijt}) \tag{3.31}$$

where $\sigma(\epsilon_{ijt})$ is the standard deviation of ϵ_{ijt}. Since the market share, $\overline{s_{ij}}$, is unobserved, I replace it using equation (3.29) to get

$$\sigma(\mu_{ijt}^{-1}) = \frac{\rho - 1}{\rho} \sigma(\epsilon_{ijt}) - \overline{\mu_{ij}^{-1}} \sigma(\epsilon_{ijt}) \tag{3.32}$$

Hence, the model predicts a negative linear relationship between the average level, $\overline{\mu_{ij}^{-1}}$, and the volatility, $\sigma(\mu_{ijt}^{-1})$, of the *inverse* markup.

I estimate equation (3.32) in order to test the model prediction and obtain an estimate of $\sigma(\epsilon_{ijt})$. To do so, I now use the standard deviation and the mean as measures of the volatility and the level of the markup. Table 3.3 displays the results. Across specifications, there is a significant negative relationship between the level and the volatility of inverse markups, as predicted by the model. The estimate for $\sigma(\epsilon_{ijt})$ ranges from 0.039 (column 4) to 0.054 (column 1). I include the same set of specifications as in Table 3.2, i.e. weighted by the average economy-wide sales share (column 1), unweighted (column 2), with industry fixed effects (column 3), taking out a linear firm-level trend (column 4), and in changes (column 5).[22] Column 6 estimates the interquartile range of ϵ_{ijt} which should be less sensitive to outliers.

In terms of magnitudes, the estimated $\sigma(\epsilon_{ijt})$ is somewhat smaller but certainly "in the same ballpark" as its model counterpart, $\sigma(\overline{\lambda}/\lambda_t) = 0.082$.[23] The discrepancy might have several reasons. On the one hand, the model might overstate the volatility of asymmetric supply shocks by assigning to them *all* fluctuations in the labor share. On the other hand, the relationship might be understated in Compustat data, which is a very particular sample of firms.[24]

21. In the simple model, the average market share is $\overline{s_{ij}} = \frac{1}{\bar{N}_j \bar{\lambda}}$ and the shock is $\epsilon_{ijt} = \frac{\bar{\lambda}}{\lambda_t}$.

22. Note that column 5 does not estimate $\sigma(\epsilon_{ijt})$, but $\sigma(d\epsilon_{ijt})$. These two statistics coincide only if there is no persistence in ϵ_{ijt}. The relatively smaller estimate for $\sigma(d\epsilon_{ijt})$ in column 5 suggests that there is persistence in ϵ_{ijt}, as calibrated in the quantitative model in Section 3.3.

23. Since the Compustat data is annual, but the model is quarterly, I aggregate the model-generated data to annual frequency.

24. To give an example of why this relationship might be understated in Compustat, note that Compustat includes primarily large firms which most likely operate in more than one market. Ideally, we would then estimate the relationship of interest using market-specific markups instead of one firm-level markup. Of course, this is not possible with the data at hand. The firm-level markup, which can be estimated, will be an average of the market-specific markups and therefore have a lower volatility than the market-specific markups if shocks are not perfectly correlated across markets.

Table 3.3. Firm-Level Inv. Markups & Inv. Markup Volatility

	(1) SD (μ^{-1})	(2) SD (μ^{-1})	(3) SD (μ^{-1})	(4) SD (μ^{-1})	(5) SD ($d\mu^{-1}$)	(6) IQR (μ^{-1})
Mean (μ^{-1})	-0.054***	-0.054***	-0.053***	-0.039***	-0.021***	
	(0.008)	(0.003)	(0.008)	(0.007)	(0.006)	
Median (μ^{-1})						-0.099***
						(0.014)
Constant	0.117***	0.130***	0.110***	0.086***	0.065***	0.183***
	(0.007)	(0.002)	(0.010)	(0.006)	(0.005)	(0.012)
Observations	12250	12250	12250	12250	12249	12250
R^2	0.037	0.025	0.231	0.030	0.010	0.049
Weights	Yes	No	Yes	Yes	Yes	Yes
Industry FE	No	No	Yes	No	No	No
Linear Trend	No	No	No	Yes	No	No

Notes: Each column displays coefficients from a separate regression: $Volatility_i(\mu_{it}^{-1}) = \beta_0 + \beta_1 * Level_i(\mu_{it}^{-1}) + \varepsilon_i$. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All variables are trimmed (1%).

3.4.2 Evidence from Aggregate Data

According to the model, the positive relationship between the level and the volatility of markups holds not only at the firm level, but also at the aggregate level. To test this relationship at the aggregate level, I now turn to time-series data.

However, measuring a time series of the aggregate markup for the U.S. economy is not straightforward. One approach would be to aggregate the firm-level markups estimated for firms in Compustat, as done in De Loecker, Eeckhout, and Unger (2020). There are, however, three issues with this method. First, an estimated firm-level markup is only available for a small and non-random subsample of firms, i.e. firms in Compustat. Second, quarterly data is scarce in Compustat before the 1980s. Therefore, time series over long horizons can only be computed at an annual frequency. Third, there are different ways to aggregate firm-level markups that lead to very different aggregate patterns, as discussed in Edmond, Midrigan, and Xu (2018). Therefore, I choose a different, indirect approach to measuring the aggregate markup which avoids these issues. That is, I employ the widely-used medium-scale DSGE model of Smets and Wouters (2007) as a "measurement device" to obtain a quarterly time series for the aggregate markup from 1957Q1 to 2019Q4. An additional benefit of this approach is that it provides an estimate of the share of consumption fluctuations that can be attributed to asymmetric supply shocks. This is because asymmetric supply shocks can be interpreted as a micro-foundation for *price-markup shocks*, which are a common element of medium-scale DSGE models, as I explain next.

3.4.2.1 Asymmetric Supply Shocks and Price-Markup Shocks

The model presented in Section 3.2 does not feature any price rigidities, which constitutes an important difference to New Keynesian models such as the one used in Smets and Wouters (2007). Therefore, firms always sell their products at the optimal (i.e., profit-maximizing) markup over the current marginal cost (see equation 3.7), thus, at the optimal price. When some friction or cost to price adjustments is introduced, as in Smets and Wouters (2007), this is not the case anymore, and a gap between the actual markup, μ_{ijt}, and the optimal ("target") markup, μ^*_{ijt}, can arise. While the actual markup—via inflation—is affected by all sorts of shocks, the target markup fluctuates around its steady-state value only due to specific exogenous shocks, referred to as price-markup shocks.[25]

In the model of Section 3.2, the target markup, $\mu^*_{ijt} = \frac{\epsilon_{ijt}}{\epsilon_{ijt}-1}$ also fluctuates over time. However, its changes are not exogenous, but arise endogenously in frameworks with oligopolistic competition and firm heterogeneity because asymmetric supply shocks affect the demand elasticity, ϵ_{ijt}.[26] Therefore, I henceforth interpret asymmetric supply shocks as a micro-foundation for price-markup shocks.[27] The similarity between endogenous markup fluctuations and "cost-push" shocks is also pointed out in Bilbiie, Fujiwara, and Ghironi (2014).[28] I abstract from *interactions* between oligopolistic competition and price stickiness, as investigated in Mongey (2021) and Wang and Werning (2022).

3.4.2.2 Estimation

Assuming that price-markup shocks reflect asymmetric supply shocks, I take the estimation of a linearized medium-scale DSGE model in Smets and Wouters (2007) "off-the-shelf" in order to estimate a time series of the target markup and to get a sense of the quantitative relevance of price-markup shocks. As the linearization of

25. For example, in Smets and Wouters (2003), the target markup fluctuates around its steady-state value ($\mu^*_{SS} = 1 + \lambda_p$) according to: $\mu^*_t = 1 + \lambda_p + \nu^p_t$ where ν^p_t is i.i.d.-normal.

26. In the simple model, the demand elasticity is a function of the number of active firms, N_{jt}, see equation (3.8). In the general model, it is a function of the market share, s_{ijt}, see equation (3.A.2).

27. Of course, ideally one would establish this link by calculating a correctly-weighted aggregate concentration measure and comparing its fluctuations with the estimated price-markup shocks. Gutiérrez, Jones, and Philippon (2021) compute an aggregate concentration measure from Compustat firms for the period 1989 - 2015 (see their Figure A.5, Panel F), which aligns quite well with the time series of the price-markup shock estimated below (see Figure 3.5). However, to the best of my knowledge, there is no longer time series available or one which covers all sectors.

28. Bilbiie, Fujiwara, and Ghironi (2014) make this point in a model in which the markup depends on the mass of firms due to a preference specification that features an "exponential love-of-variety".

the model eliminates any non-linearities, it is irrelevant that the model of Smets and Wouters (2007) does not feature the non-linearity highlighted in this paper, i.e. the aggregate markup becoming more volatile when its level is higher. In principle, this non-linearity calls for a non-linear estimation, which however is outside the scope of this paper.

I estimate the linearized model of Smets and Wouters (2007) using an updated sample from 1957Q1 to 2019Q4 as in Bayer, Born, and Luetticke (2020). The only adjustment of the model, also following Bayer, Born, and Luetticke (2020), is the removal of the moving-average components of the wage- and price-markup shock processes. That is, I specify AR(1) processes for the wage- and price-markup shocks, instead of ARMA(1,1) processes.[29]

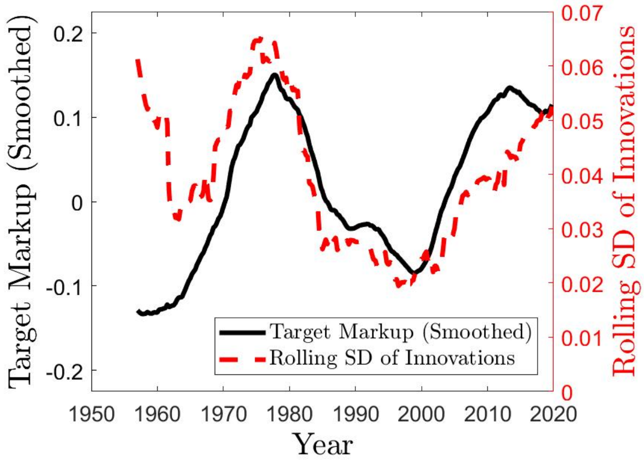


Figure 3.5. Markups & Markup Volatility in Estimated DSGE Model

Notes: The time series for the target markup results from the estimation of the model of Smets and Wouters (2007). Sample: 1957Q1 - 2019Q4. The smoothed target markup and the rolling-window standard deviation of innovations are both calculated over a symmetric 7-year window.

I find price-markup shocks (i.e., asymmetric supply shocks) to explain up to 7.4% of the fluctuations in consumption. Thus, these disturbances are a quantitatively relevant source of aggregate fluctuations. Due to the high persistence of the shocks, the share of fluctuations explained at short horizons is substantially lower than at longer horizons. These numbers are in line with the literature. For example, Smets and Wouters (2007) find price-markup shocks to explain up to 12% of fluctuations

29. The price-markup shock process is estimated to be quite persistent, with a posterior mode of the autoregressive parameter of $\rho_\mu = 0.8665$, in line with Smets and Wouters (2007) and Bayer, Born, and Luetticke (2020), who both find ρ_μ to be around 0.9. The posterior mode of the standard deviation of innovations to the price-markup shock is 0.0431, which is lower than the 0.14 estimated in Smets and Wouters (2007). Still, I estimate a higher volatility of the price-markup shock as the process does not include a moving-average term as in Smets and Wouters (2007).

in output. In addition, Bayer, Born, and Luetticke (2020) find that the relevance of price-markup shocks increases once the representative household is replaced by heterogeneous households. Interestingly, they find price-markup shocks to also be a key driver of fluctuations in income and wealth inequality.

Figure 3.5 plots the smoothed time series for the target markup (black line) alongside the rolling-window standard deviation of innovations to the target markup (red line). Two observations stand out. First, the markup has not only been high in recent years but was also high during the 1970s. Second, both periods of a high markup coincide with a high standard deviation of innovations to the markup. The correlation of the two series is 0.49. This finding supports the prediction of the model that the volatility of the aggregate markup is high when its level is high.

3.4.3 Discussion: Conditional Evidence

Both subsections, using firm-level data as well as time-series data, present *unconditional* correlations. The empirical analysis in Ferrari and Queirós (2022) provides complementary evidence *conditional* on a particular asymmetric supply shock. They show that after the financial crisis, which can be interpreted as an asymmetric supply shock (see Section 3.2.2), labor shares fell more strongly in industries that were more concentrated at the onset of the crisis. This suggests that markups increased more strongly in industries with a lower intensity of competition, which is in line with the main mechanism discussed in this paper.

3.5 Conclusion

In this paper, I study the aggregate effects of asymmetric supply shocks. Most importantly, I show that a high intensity of competition among firms makes an economy resilient to asymmetric supply shocks. The main mechanism relies on the profit-maximizing behavior of firms that compete strategically within narrow industries. In response to negative shocks to their competitors, firms with high market power find it optimal to not stabilize total output by expanding production, but to raise the prices of their goods instead. In contrast, firms with low market power find it optimal to primarily expand their production, thereby stabilizing total output.

This mechanism provides an additional reason why the secular increases in market power, markups, and industry concentration, documented by De Loecker, Eeckhout, and Unger (2020) and Covarrubias, Gutiérrez, and Philippon (2020), are troubling. They not only reduce consumer welfare in static economies (Edmond, Midrigan, and Xu, 2018) but also increase aggregate volatility, which further reduces welfare of risk-averse consumers. Competition policy must take into account that by leaning against these trends, it can not only reduce markups but also provide macroeconomic stabilization.

I emphasize that the main mechanism is relevant for all supply disruptions that change the distribution of sales shares among firms within industries—and thereby change industry concentration and the effective number of firms. A broad class of models with firm heterogeneity gives rise to such disruptions, referred to as asymmetric supply shocks. Some well-known examples include models with firm financial heterogeneity (Khan and Thomas, 2013; Ottonello and Winberry, 2020), regional shocks (Atkeson and Burstein, 2008), or time-varying distributions of idiosyncratic shocks (Bloom et al., 2018; Salgado, Guvenen, and Bloom, 2019).

In light of this wide range of supply disruptions that can cause fluctuations in market shares and industry concentration, two interesting questions for future research emerge. First, what are quantitatively the main drivers of fluctuations in concentration at the industry level and at the aggregate level? Second, do the main drivers and the extent of fluctuations in concentration depend on the intensity of competition among firms, as has been observed for banks (Corbae and D'Erasmo, 2021)? Addressing both questions requires linking product market data (prices, quantities) with firm-level information (productivity, financial resources) as done in Gilchrist et al. (2017) or Suveg (2021).

Appendix 3.A Model Appendix

3.A.1 A Generalized Industry Setup

In this section, I describe a generalized industry setup that allows for various forms of firm heterogeneity and asymmetric supply shocks. The behavior of the industry good producers and the final consumption good producer remains unchanged. I continue to assume that within each industry j, there are $\widetilde{N}_j$ firms, which are indexed by $i \in \{1, ..., \widetilde{N}_j\}$. Each firm ij produces the intermediate good y_{ij} according to the constant-returns-to-scale production technology $y_{ijt} = z_{ijt} l_{ijt}$.

Firm heterogeneity still originates from the firm-specific component z_{ijt}. The crucial difference with respect to the simple industry setup in Section 3.2.1 is that no restrictions are imposed on z_{ijt}. In particular, z_{ijt} is not restricted to being a binary variable anymore. Many types of firm heterogeneity and asymmetric supply shocks map into this setup with appropriate choices for the firm-specific component. Some important examples, including financial shocks, are discussed below in Section 3.A.3. Beforehand, I explain firm behavior and aggregate outcomes in this generalized industry setup.

Firm Optimization. Firms continue to maximize profits (equation 3.5) under Cournot competition subject to the demand curve (equation 3.6). Under optimal behavior, firms still set a markup over marginal costs which depends on the intensity of competition in their industry (see equation 3.7). However, with firm heterogeneity among active firms, the number of active firms, N_{jt}, is not sufficient to characterize the intensity of competition and market power anymore. Instead, a firm's market power and thus its optimal markup now depends on its market share defined by

$$s_{ijt} = \frac{p_{ijt}y_{ijt}}{P_{jt}Y_{jt}} \tag{3.A.1}$$

In particular, the optimal markup (see equation 3.8) becomes

$$\mu_{ijt}(s_{ijt}) = \frac{\epsilon_{ijt}(s_{ijt})}{\epsilon_{ijt}(s_{ijt}) - 1} \qquad \text{where} \qquad \epsilon_{ijt}(s_{ijt}) = \left[\frac{1}{\eta}s_{ijt} + \frac{1}{\rho}\left(1 - s_{ijt}\right)\right]^{-1} \tag{3.A.2}$$

As discussed in Section 3.3.1, in partial equilibrium, firms' price-setting and production decisions are inextricably linked. In general equilibrium, production is demand-determined and firms supply any quantity at the markup μ_{ijt}. Therefore, I focus on the markup (price) decision in this section and do not separately discuss the output (quantity) decision. It is helpful to rewrite the optimal markup (equation 3.A.2) as

$$\mu_{ijt} = \frac{\rho}{\rho - 1}\left[1 - \frac{\frac{\rho}{\eta} - 1}{\rho - 1}s_{ijt}\right]^{-1} \tag{3.A.3}$$

In addition, combining the definition of the market share (equation 3.A.1) with the price equation (equation 3.7) and the demand curve (equation 3.6), yields the following expression for the market share

$$s_{ijt} = \frac{z_{ijt}^{\rho-1}\mu_{ijt}^{1-\rho}}{\sum_{k=1}^{\widetilde{N}_j} z_{kjt}^{\rho-1}\mu_{kjt}^{1-\rho}} \tag{3.A.4}$$

Given the firm-specific component z_{ijt} for all firms i, equations (3.A.3) and (3.A.4) can be used to solve for all firms' market shares, s_{ijt}, and markups, μ_{ijt}, in period t. Importantly, this is possible regardless of the specific distribution of firm-specific components, z_{ijt}.

Industry Aggregates. The industry markup, defined by $\mu_{jt} = \frac{(P_{jt}/P_t^C)Y_{jt}}{w_t L_{jt}}$, can be rewritten, using equation (3.A.3), as a function of the Herfindahl–Hirschman index (*HHI*), a measure of industry concentration

$$\mu_{jt} = \frac{\rho}{\rho - 1}\left[1 - \frac{\frac{\rho}{\eta} - 1}{\rho - 1}HHI_{jt}\right]^{-1} \tag{3.A.5}$$

where the *HHI* is calculated as the sum of squared market shares, $HHI_{jt} = \sum_{i=1}^{\widetilde{N}_{jt}} s_{ijt}^2$.

With heterogeneity among active firms, asymmetric supply shocks do not only have an effect on the industry markup μ_{jt}, but also on industry productivity Z_{jt}. Industry productivity is defined by

$$Z_{jt} = \frac{Y_{jt}}{L_{jt}} = \frac{N_{jt}^{\frac{1}{1-\rho}} \left[\sum_{i=1}^{\widetilde{N_{jt}}} \mu_{ijt}^{1-\rho} \right]^{\frac{\rho}{\rho-1}}}{\sum_{i=1}^{\widetilde{N_{jt}}} \mu_{ijt}^{-\rho}} \tag{3.A.6}$$

where again $N_{jt}^{\frac{1}{1-\rho}}$ is the term arising from the cancellation of love of variety effects.[30] It is easy to see that with symmetric firms, $Z_{jt} = 1$ and changes in N_{jt}, as in the main text, do not affect Z_{jt}. However, as soon as there are love of variety effects or active firms are heterogeneous, industry productivity is affected by asymmetric supply shocks. The analytical results below focus on the effect on markups, however. An in-depth analysis of the effects of asymmetric supply shocks on productivity is left for future research.

Aggregation of Industries. In the following, I discuss the effect of asymmetric supply shocks on the industry markup, μ_{jt}. Under the assumption that all industries are identical, as made in the main text, the aggregate markup equals the industry markup: $\mu_t^C = \mu_{jt}$.[31] Thus, the effect of an asymmetric supply shock on the industry markup equals the effect on the aggregate markup. However, in principle, the framework allows for industry heterogeneity and thus shocks that affect only a subset of industries.

3.A.2 Analytical Results

To see that the analytical results derived in the main paper generalize to this more general framework, it is helpful to define the effective number of firms as

$$N_{jt}^{eff} = HHI_{jt}^{-1} = \left(\sum_{i=1}^{\widetilde{N_{jt}}} s_{ijt}^2 \right)^{-1} \tag{3.A.7}$$

Intuitively, the effective number of firms is the number of homogeneous firms which results in the same intensity of competition (and thus the same industry concentration) as a given distribution of heterogeneous firms. Using equation (3.A.5), the

30. As before, the number of active firms is defined as the number of firms with a positive market share. All firms with $z_{ijt} > 0$ have a positive market share.

31. Moreover, assuming that all industries are identical, aggregate productivity equals industry productivity: $Z_t = Z_{jt}$. See Burstein, Carvalho, and Grassi (2020) for a more detailed discussion of industry heterogeneity in a similar framework.

industry markup is

$$\mu_{jt} = \frac{\rho}{\rho-1}\left[1 - \frac{\frac{\rho}{\eta}-1}{\rho-1}\left(N_{jt}^{eff}\right)^{-1}\right]^{-1} \tag{3.A.8}$$

whereas in the baseline model, the industry markup can be written as

$$\mu_{jt} = \frac{\rho}{\rho-1}\left[1 - \frac{\frac{\rho}{\eta}-1}{\rho-1}N_{jt}^{-1}\right]^{-1} \tag{3.A.9}$$

Comparing these two equations, it is obvious that a change in the effective number of firms, N_{jt}^{eff}, in the generalized setup has exactly the same effect on the industry markup as a change in the number of active firms, N_{jt}, in the baseline model. Proposition 3.2 therefore extends Proposition 3.1 to the generalized setup.

Proposition 3.2. *In a more competitive industry (higher steady-state effective number of firms $\widetilde{N_j^{eff}}$), an asymmetric supply shock (log-change in effective number of firms) has a smaller absolute effect on the industry markup:*

$$\frac{d\left(\frac{dlog(\mu_{jt})}{dlog(N_{jt}^{eff})}\right)}{d\widetilde{N_j^{eff}}} > 0$$

Proof. This follows immediately from replacing N with N^{eff} in the first line of the proof of Proposition 3.1. □

Numerical Example. Intuitively, Proposition 3.2 shows that a given percentage change in the effective number of firms (industry concentration) has a larger effect on the industry markup when the steady-state effective number of firms (industry concentration) is low (high) to begin with. To provide a numerical example, consider an industry with two firms in which market shares are reallocated from $s = \{0.5, 0.5\}$ to $s = \{0.4, 0.6\}$. The HHI increases from 0.5 to 0.52, i.e. by 4%. In consequence, the industry markup increases by 3.2%, using the parameter values from Table 3.1. In contrast, if the industry was populated by four firms, i.e. splitting each firm in two, the same reallocation of market shares would be from $s = \{0.25, 0.25, 0.25, 0.25\}$ to $s = \{0.2, 0.2, 0.3, 0.3\}$. The HHI increases from 0.25 to 0.26, which again is an increase of 4%. However, the resulting change in the industry markup would only be 1.13%, so a lot less than in the industry with only two firms, in line with Proposition 3.2.

Discussion. Many supply disruptions can cause changes in industry concentration (effective number of firms) as in the numerical example above. Some of these are discussed in more detail below in Section 3.A.3. However, depending on the example, the effect of an asymmetric supply shock on industry concentration may also depend on the intensity of competition. To see this, consider the following decomposition of the effect of an arbitrary asymmetric supply shock ϵ_t^A on the industry markup:

$$\frac{\mathrm{dlog}(\mu_{jt})}{\mathrm{dlog}(\epsilon_{jt}^A)} = \underbrace{\frac{\mathrm{dlog}(\mu_{jt})}{\mathrm{dlog}(HHI_{jt})}}_{A}\underbrace{\frac{\mathrm{dlog}(HHI_{jt})}{\mathrm{dlog}(\epsilon_t^A)}}_{B} \tag{3.A.10}$$

Proposition 3.2 has established that the former elasticity ("A") is decreasing (in absolute terms) in the intensity of competition.[32] However, whether the latter elasticity ("B") also depends on the intensity of competition depends on the precise example. In the baseline model, this was not the case, because $\frac{\mathrm{dlog}(N_t)}{\mathrm{dlog}(\lambda_t)} = 1$ (see equation 3.22). In contrast, in the financial frictions example below, this is the case. However, in sum, the elasticity of the industry markup with respect to the asymmetric supply shock continues to be decreasing (in absolute terms) in the intensity of competition.

3.A.3 Examples of Asymmetric Supply Shocks

The previous subsection has shown that a given percentage change in industry concentration (the effective number of firms) has larger aggregate effects when industry concentration is high (the effective number of firms is low) to begin with. Now, I show that many types of firm heterogeneity map into the generalized setup explained in Section 3.A.1 and give rise to changes in industry concentration, such that they can be considered asymmetric supply shocks. I first discuss in detail a framework with firm heterogeneity due to financial frictions. Thereafter, I briefly describe additional examples.

Financial Frictions. In heterogeneous-firm models with financial frictions, such as Khan and Thomas (2013), Khan, Senga, and Thomas (2016) or Ottonello and Winberry (2020), aggregate shocks affect concentration via their effect on the tightness of financial constraints. To illustrate this, I consider a model in which firms produce using capital and labor according to the production function $y_{ijt} = k_{ijt}^{\theta} l_{ijt}$ with $\theta < 1$. Capital is purchased at the end of the previous period, such that at time t, only labor can be adjusted. Defining $z_{ijt} = k_{ijt}^{\theta}$, it is evident, that this example maps into the generalized industry setup outlined above. In addition, I assume that within each (symmetric) industry there are two types of firms: financially constrained firms and financially unconstrained ones. Financially unconstrained firms ("u") choose the optimal level of capital k_{jt}^*. I normalize $z_{ujt} = (k_{jt}^*)^{\theta} = 1$. Financially

32. Note that $\mathrm{dlog}(HHI_{jt}) = -\mathrm{dlog}(N_{jt}^{eff})$.

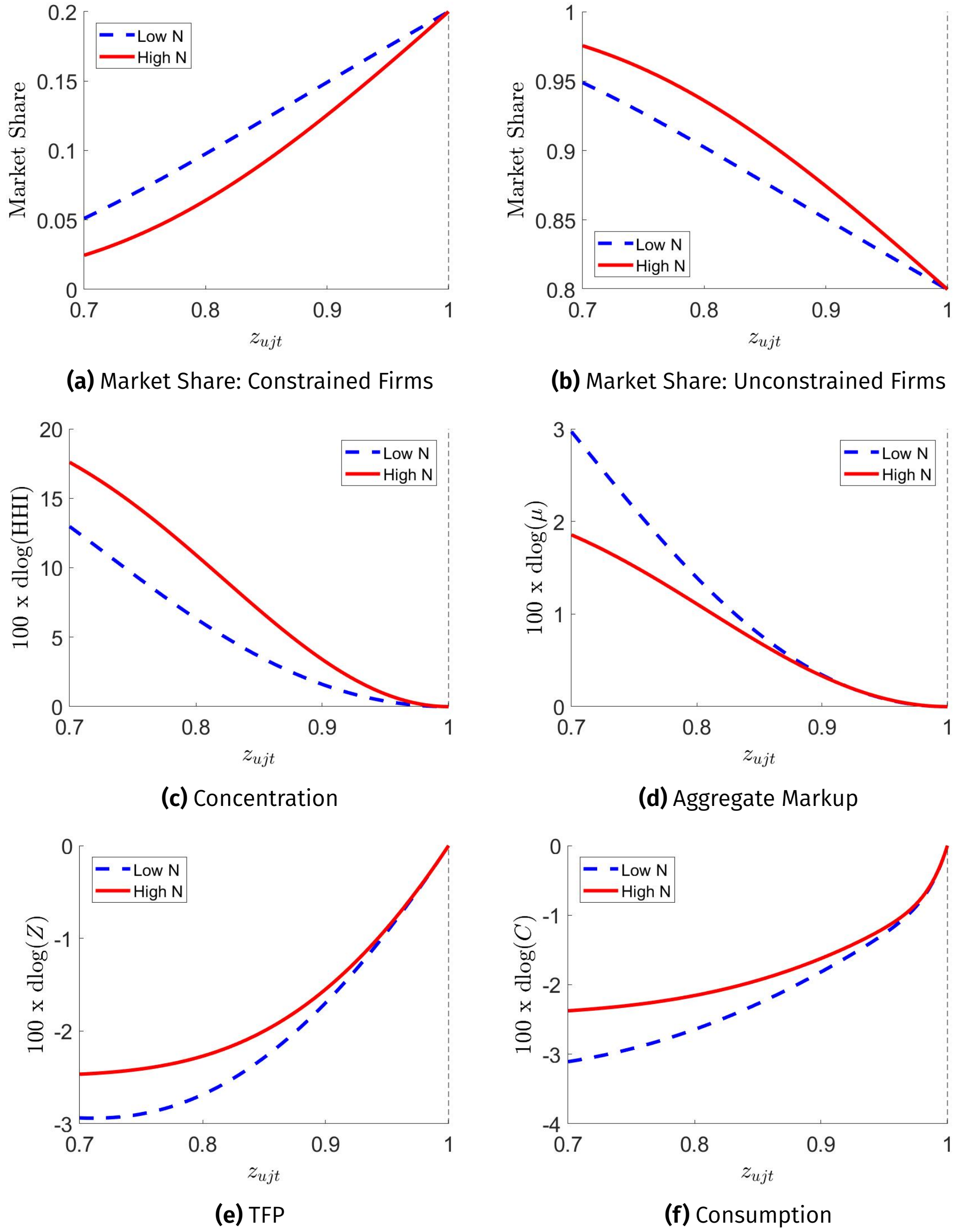
constrained firms ("c") can only purchase capital up to an (always binding) limit, γ_t, such that $k_{cjt} \leq \gamma_t \leq k^*_{jt}$. Hence, $z_{cjt} = \gamma_t^\theta \leq 1$.

In this setup, financial shocks, i.e. changes in the tightness of the financial constraint γ_t, directly affect financially constrained firms (z_{cjt}), but not financially unconstrained firms (z_{ujt}). To illustrate the effects, I provide a numerical example: All firms are initially of the same size ($z_{cjt-1} = 1$), and a negative financial shock hits, such that $z_{cjt} < 1$. I compare two economies, one with a high intensity of competition ($N = 10$) and one with a low intensity of competition ($N = 5$). In both economies, 20% of firms are financially constrained. Figure 3.A.1 plots the effects for a range of values for z_{cjt}, i.e. a range of shock sizes. Panel (a) shows that the financial tightening (decrease in γ_t, so decrease in z_{cjt}) leads to a fall in the market share of these 20% of constrained firms. Vice versa, the market share of the remaining 80% of unconstrained firms increases, as shown in panel (b). This leads to an increase in concentration, i.e. the *HHI*, as shown in panel (c). Notably, the increase in concentration is larger (in percent) when there are more firms (red line). This is because, with more firms, it is easier to substitute away from the 20% of constrained and hence less productive firms.

In sum, it is therefore theoretically ambiguous whether a higher intensity of competition stabilizes the economy, or not. On the one hand, higher competition *increases* the effect of a financial tightening on industry concentration (panel (c)). On the other hand, a given percentage change in concentration leads to a larger percentage change in the aggregate markup according to Proposition 3.2. Panel (d) shows that in this example, the stabilizing effect of competition dominates and the financial tightening has a smaller effect on the aggregate markup when there are more firms. In addition, the decrease in TFP is smaller when competition is intense, as shown in panel (e). The reason is the same as before: the industry as a whole is better able to substitute away from the constrained firms when there are many firms. It follows that the effect of the financial tightening on aggregate consumption is smaller when there is a high intensity of competition, as shown in panel (f).

In models with endogenous financial constraints, e.g. Ottonello and Winberry (2020), all aggregate shocks affect concentration and thus the industry markup. For the particular case of monetary policy shocks, supporting empirical evidence can be found in Meier and Reinelt (2022). Ferrando et al. (2021) and Furceri et al. (2021) also empirically investigate the transmission of monetary policy shocks via firms considering both financial frictions and market power.

Idiosyncratic Shocks. Idiosyncratic shocks to productivity, demand, capital quality, or some other firm-level state variable can be considered a special case of asymmetric supply shocks. For the case of idiosyncratic productivity shocks, one would simply define z_{ijt} as firm productivity, which follows some exogenous process. For the case of capital quality shocks, one would define $z_{ijt} = \epsilon^{CQ}_{ijt} k_{ijt}$, where k_{ijt} is capital and ϵ^{CQ}_{ijt} is a capital quality shock, drawn from some exogenous distribution. Either



(a) Market Share: Constrained Firms

(b) Market Share: Unconstrained Firms

(c) Concentration

(d) Aggregate Markup

(e) TFP

(f) Consumption

Figure 3.A.1. Financial Tightening as an Asymmetric Supply Shock

Notes: On the x-axis, various sizes of financial tightening shocks are shown. The black vertical dashed lines indicate the state before any financial tightening ($z_{ujt} = 1$). Red solid (blue dashed) lines represent an economy with a high (low) intensity of competition, i.e. a high (low) number of firms.

way, the idiosyncratic shock affects the firm-specific component of one firm, while not directly affecting any other firms in the industry. Therefore, market shares are reallocated and—expect for knife-edge cases—industry concentration changes and thus the industry markup.[33] However, idiosyncratic shocks only matter for *aggregate* outcomes, i.e. the aggregate markup, when firms are not atomistic, e.g. as in the setup of Burstein, Carvalho, and Grassi (2020) with a large but finite number of industries. In contrast, when there is a continuum of industries, as in Atkeson and Burstein (2008), idiosyncratic shocks "wash out" and do not have aggregate effects. Yet, shocks to the *distribution* of these idiosyncratic shocks still do have aggregate effects, because they change the distribution of firm-specific components and thus reallocate sales shares in *all* industries. Examples of these asymmetric supply shocks include shocks to the dispersion (e.g. Bloom, 2009; Bachmann and Bayer, 2014; Ferrari and Queirós, 2022) or skewness (e.g. Salgado, Guvenen, and Bloom, 2019) of idiosyncratic shocks.

Multi-Country Setups. In the two-country model of Atkeson and Burstein (2008), country-specific TFP shocks are asymmetric supply shocks, because they only affect firms from one country. This can be represented in the generalized setup by setting $z_{ijt} \in \{z_t^A, z_t^B\}$, where z_t^A and z_t^B are the country-specific productivity levels for countries A and B, respectively. A change in country-specific TFP reallocates market shares between the firms of the two countries and therefore changes concentration and the industry markup.

Appendix 3.B Data Appendix

3.B.1 Data Treatment

For the empirical analysis in Section 3.4.1, I use annual firm-level data from Compustat North America. The data treatment described here broadly follows De Loecker, Eeckhout, and Unger (2020) and Baqaee and Farhi (2020). From the beginning, I exclude

(1) firms not incorporated in the United States (based on FIC)
(2) observations with missing or non-classifiable industry (NAICS)
(3) observations with missing or non-positive sales (SALE), cost of goods sold (COGS), or total assets (AT)

Deflators. Sales and cost of goods sold are deflated using the price index for gross output from KLEMS. Capital expenditures and capital are deflated using the price index for gross fixed capital formation from KLEMS. These deflators are available from 1970 to 2015 which limits the analysis to these years.

33. See Burstein, Carvalho, and Grassi (2020), Proposition 1 for a result regarding the *sign* of the change in the industry markup.

3.B.2 Markup Estimation

The fundamental issue is that markups, defined as the price of a good divided by its marginal cost, are not observable. While the price of a good is typically observable, the marginal cost of producing it is not. Basu (2019) provides a summary and discussion of the various approaches developed in the literature to deal with this issue. I estimate firm-level markups using the popular "production function approach" due to De Loecker and Warzynski (2012). This approach is also used in De Loecker, Eeckhout, and Unger (2020) and Baqaee and Farhi (2020).

I briefly describe the "production function approach" to estimating firm-level markups, largely following the exposition of De Loecker and Warzynski (2012). A firm i at time t produces output y_{it} using the production technology $y_{it} = F_{it}(z_{it}, k_{it}, v_{it})$, where z_{it} is productivity, k_{it} is capital and v_{it} is a variable input factor. The firm minimizes costs, subject to producing the quantity $\overline{y_{it}}$. The Lagrangian function is

$$L_{it} = r_{it}k_{it} + p^v_{it}v_{it} + \lambda_{it}\left(\overline{y_{it}} - y_{it}(\cdot)\right) \tag{3.B.1}$$

where the factor prices r_{it} and p^v_{it} are taken as given by the firm and λ_{it} is the Lagrange multiplier. Note that λ_{it} reflects the marginal cost of production, because $\frac{\partial L_{it}}{\partial Q_{it}} = \lambda_{it}$. The first-order condition for the variable input is

$$\frac{\partial L_{it}}{\partial v_{it}} = p^v_{it} - \lambda_{it}\frac{\partial y_{it}(\cdot)}{\partial v_{it}} = 0 \tag{3.B.2}$$

Rearranging and expanding this condition yields

$$\underbrace{\frac{p_{it}}{\lambda_{it}}}_{\mu_{it}} = \underbrace{\frac{\partial y_{it}(\cdot)}{\partial v_{it}}\frac{v_{it}}{y_{it}}}_{\theta^v_{it}} \underbrace{\frac{p_{it}y_{it}}{p^v_{it}v_{it}}}_{(s^v_{it})^{-1}} \tag{3.B.3}$$

Thus, the markup, μ_{it}, is simply the product of the output elasticity on the variable input, θ^v_{it}, and the inverse of the share of the variable input's expenditure in total sales, $(s^v_{it})^{-1}$.

Estimating Output Elasticities. While the share of inputs in total sales can easily be calculated in most datasets, output elasticities need to be estimated. There exists an extensive literature on the estimation of production functions. I follow the implementation in Baqaee and Farhi (2020), who use the methodology of Olley and Pakes (1996) with the correction proposed by Ackerberg, Caves, and Frazer (2015). I briefly outline the main idea, building on the aforementioned papers, which provide a much more complete description. Assuming a Cobb-Douglas functional form for $F_{it}(z_{it}, k_{it}, v_{it})$ and taking logs, the production function can be written as

$$\log(y_{it}) = \beta_0 + \theta^k \log(k_{it}) + \theta^v \log(v_{it}) + z_{it} + \epsilon_{it} \tag{3.B.4}$$

where z_{it} is a productivity shock, which is observed by the firm before choosing the variable input. ϵ_{it} is measurement error observed only after choosing inputs. The presence of productivity shocks would bias an OLS estimate of θ^v. The idea of Olley and Pakes (1996) is to control for z_{it} using investment i_{it} as a "proxy" variable, because it is observable and under mild assumptions monotonically increasing in z_{it}. This enables estimating the output elasticity on the variable input, θ^v, using GMM.

3.B.3 Implementation

I now describe how the markup estimation described in Appendix 3.B.2 is applied to the Compustat data treated as described in Appendix 3.B.1. As before, the steps broadly follow De Loecker, Eeckhout, and Unger (2020) and Baqaee and Farhi (2020).

Output Elasticities. The output elasticity of a variable input is estimated using the Stata package *prodest*, which implements Olley and Pakes (1996) and the correction of Ackerberg, Caves, and Frazer (2015). I use (log) sales as the output variable, (log) cost of goods sold (COGS) as the variable input, (log) capital (PPEGT) as the state variable, and (log) investment (CAPX) as the proxy variable. Additional controls are SIC-3-digit and SIC-4-digit sales shares to deal with the issue that sales do not measure quantities, but revenue.

To deal with outliers, the top and bottom 5% of the year-specific COGS-to-sales and XSGA-to-sales ratios are excluded for the estimation of the production function. I estimate time-varying elasticities by using 11-year rolling windows. Choosing relatively long windows ensures fairly stable parameter estimates. Moreover, I estimate industry-specific elasticities, grouping industries based on 2-digit NAICS codes.

Markups. Having estimated output elasticities of cost of goods sold, the markup is simply computed as the product of the elasticity with the inverse expenditure share on COGS, i.e. $\frac{SALE}{COGS}$.

Firm-Level Dataset. Inspecting the distribution of firm markups shows that the dataset includes many outliers. Therefore, I trim the top and bottom 7.5% of the year-specific markup distribution. Even after this, estimated markups range from 0.52 to 7.45, which are in an economic sense extreme values. In addition, I exclude firms for which fewer than 6 observations are available. This ensures that I can sensibly calculate measures of markup volatility at the firm level, such as the interquartile range and the standard deviation.

3.B.4 Additional Results

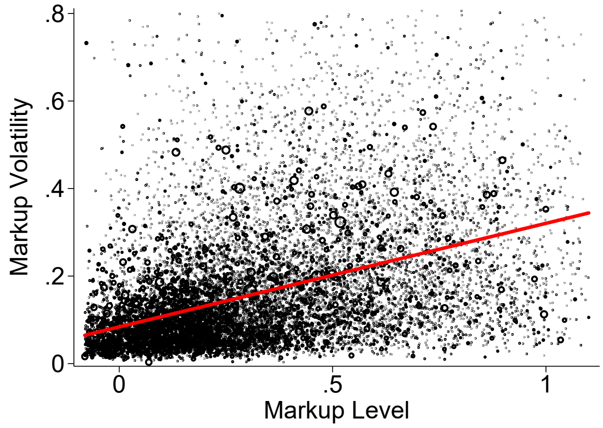


Figure 3.B.1. Firm-Level (Log) Markups & (Log) Markup Volatility

Notes: Each circle depicts one firm. The markup level is the median log markup. The markup volatility is the interquartile range of the log markup. Both variables are trimmed (1%). Firms are weighted by their average sales share. The red line shows the linear fit.

Table 3.B.1. Firm-Level (Log) Markups & (Log) Markup Volatility

	(1) IQR (log μ)	(2) IQR (log μ)	(3) IQR (log μ)	(4) IQR (log μ)	(5) IQR (dlog μ)
Median (log μ)	0.237***	0.218***	0.226***	0.164***	0.089***
	(0.017)	(0.005)	(0.017)	(0.016)	(0.008)
Constant	0.084***	0.107***	0.069*	0.062***	0.044***
	(0.005)	(0.002)	(0.039)	(0.004)	(0.002)
Observations	12257	12257	12257	12250	12251
R^2	0.211	0.146	0.316	0.188	0.161
Weights	Yes	No	Yes	Yes	Yes
Industry FE	No	No	Yes	No	No
Linear Trend	No	No	No	Yes	No

Notes: Each column displays coefficients from a separate regression: $Volatility_i(\log(\mu_{it})) = \beta_0 + \beta_1 * Level_i(\log(\mu_{it})) + \varepsilon_i$. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All variables are trimmed (1%).

References

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer. 2015. "Identification properties of recent production function estimators." *Econometrica* 83 (6): 2411–51. [140, 141]

Atkeson, Andrew, and Ariel Burstein. 2008. "Pricing-to-market, trade costs, and international relative prices." *American Economic Review* 98 (5): 1998–2031. [102, 104, 107, 109, 110, 119, 120, 132, 139]

Bachmann, Rüdiger, and Christian Bayer. 2014. "Investment dispersion and the business cycle." *American Economic Review* 104 (4): 1392–416. [110, 139]

Baqaee, David Rezza, and Emmanuel Farhi. 2020. "Productivity and misallocation in general equilibrium." *Quarterly Journal of Economics* 135 (1): 105–63. [105, 125, 126, 139–141]

Basu, Susanto. 2019. "Are price-cost markups rising in the United States? A discussion of the evidence." *Journal of Economic Perspectives* 33 (3): 3–22. [119, 140]

Bayer, Christian, Benjamin Born, and Ralph Luetticke. 2020. "Shocks, frictions, and inequality in US business cycles." *CEPR Discussion Paper No. DP14364.* [130, 131]

Bilbiie, Florin O, Ippei Fujiwara, and Fabio Ghironi. 2014. "Optimal monetary policy with endogenous entry and product variety." *Journal of Monetary Economics* 64: 1–20. [129]

Bilbiie, Florin O, Fabio Ghironi, and Marc J Melitz. 2012. "Endogenous entry, product variety, and business cycles." *Journal of Political Economy* 120 (2): 304–45. [104, 106, 110, 123]

Bilbiie, Florin O, Fabio Ghironi, and Marc J Melitz. 2019. "Monopoly power and endogenous product variety: Distortions and remedies." *American Economic Journal: Macroeconomics* 11 (4): 140–74. [107, 123]

Bloom, Nicholas. 2009. "The impact of uncertainty shocks." *Econometrica* 77 (3): 623–85. [104, 110, 139]

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry. 2018. "Really Uncertain Business Cycles." *Econometrica* 86 (3): 1031–65. [132]

Boar, Corina, and Virgiliu Midrigan. 2019. "Markups and inequality." *NBER Working Paper No. 25952, National Bureau of Economic Research.* [107]

Burstein, Ariel, Vasco M Carvalho, and Basile Grassi. 2020. "Bottom-up markup fluctuations." *NBER Working Paper No. 27958, National Bureau of Economic Research.* [106, 109, 112, 113, 126, 134, 139]

Corbae, Dean, and Pablo D'Erasmo. 2021. "Capital buffers in a quantitative model of banking industry dynamics." *Econometrica* 89 (6): 2975–3023. [132]

Corhay, Alexandre, Howard Kung, and Lukas Schmid. 2020. "Competition, markups, and predictable returns." *Review of Financial Studies* 33 (12): 5906–39. [106, 110]

Covarrubias, Matias, Germán Gutiérrez, and Thomas Philippon. 2020. "From Good to Bad Concentration? US Industries over the Past 30 Years." *NBER Macroeconomics Annual* 34: 1–46. [106, 131]

De Loecker, Jan, Jan Eeckhout, and Simon Mongey. 2021. "Quantifying market power and business dynamism in the macroeconomy." *NBER Working Paper No. 28761, National Bureau of Economic Research.* [111]

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2020. "The rise of market power and the macroeconomic implications." *Quarterly Journal of Economics* 135 (2): 561–644. [102, 105, 106, 119, 120, 125, 128, 131, 139–141]

De Loecker, Jan, and Frederic Warzynski. 2012. "Markups and firm-level export status." *American Economic Review* 102 (6): 2437–71. [105, 125, 140]

De Ridder, Maarten, Basile Grassi, and Giovanni Morzenti. 2021. "The hitchhiker's guide to markup estimation." *Mimeo, London School of Economics.* [119]

De Santis, Massimiliano. 2007. "Individual Consumption Risk and the Welfare Cost of Business Cycles." *American Economic Review* 97 (4): 1488–506. [121]

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. 2018. "How costly are markups?" *NBER Working Paper No. 24800, National Bureau of Economic Research.* [107, 119, 128, 131]

Ferrando, Annalisa, Peter McAdam, Filippos Petroulakis, and Xavier Vives. 2021. "Product market structure and monetary policy: evidence from the Euro Area." *ECB Working Paper No. 2632, European Central Bank.* [137]

Ferrari, Alessandro, and Francisco Queirós. 2022. "Firm Heterogeneity, Market Power and Macroeconomic Fragility." *Mimeo, University of Zurich.* [106, 110, 131, 139]

Furceri, Davide, Romain Duval, Marina M Tavares, and Raphael Lee. 2021. "Market Power and Monetary Policy Transmission." *IMF Working Papers* 2021 (184). [137]

Gabaix, Xavier. 2011. "The granular origins of aggregate fluctuations." *Econometrica* 79 (3): 733–72. [103, 105, 106]

Gilchrist, Simon, Raphael Schoenle, Jae Sim, and Egon Zakrajšek. 2017. "Inflation dynamics during the financial crisis." *American Economic Review* 107 (3): 785–823. [132]

Guerrieri, Veronica, Guido Lorenzoni, Ludwig Straub, and Iván Werning. 2022. "Macroeconomic Implications of COVID-19: Can Negative Supply Shocks Cause Demand Shortages?" *American Economic Review* 112 (5): 1437–74. [105]

Gutiérrez, Germán, Callum Jones, and Thomas Philippon. 2021. "Entry costs and aggregate dynamics." *Journal of Monetary Economics* 124: S77–S91. [129]

Jaimovich, Nir, and Max Floetotto. 2008. "Firm dynamics, markup variations, and the business cycle." *Journal of Monetary Economics* 55 (7): 1238–52. [104, 106, 110, 123]

Khan, Aubhik, Tatsuro Senga, and Julia Thomas. 2016. "Default Risk and Aggregate Fluctuations in an Economy with Production Heterogeneity." *Mimeo, The Ohio State University.* [107, 109, 136]

Khan, Aubhik, and Julia Thomas. 2008. "IDIOSYNCRATIC SHOCKS AND THE ROLE OF NONCONVEXITIES." *Econometrica* 76 (2): 395–436. [103, 117]

Khan, Aubhik, and Julia Thomas. 2013. "Credit shocks and aggregate fluctuations in an economy with production heterogeneity." *Journal of Political Economy* 121 (6): 1055–107. [104, 107, 109, 132, 136]

Koby, Yann, and Christian Wolf. 2020. "Aggregation in Heterogeneous-Firm Models: Theory and Measurement." *Mimeo, Princeton University.* [103, 117]

Kydland, Finn E, and Edward C Prescott. 1982. "Time to build and aggregate fluctuations." *Econometrica,* 1345–70. [105]

Long, John B, and Charles I Plosser. 1983. "Real business cycles." *Journal of Political Economy* 91 (1): 39–69. [105]

Lucas, Robert E. 1987. *Models of Business Cycles.* Blackwell. [121]

Meier, Matthias, and Timo Reinelt. 2022. "Monetary Policy, Markup Dispersion, and Aggregate TFP." *Mimeo, University of Mannheim.* [137]

Mongey, Simon. 2021. "Market structure and monetary non-neutrality." *NBER Working Paper No. 29233, National Bureau of Economic Research.* [106, 119, 120, 129]

Olley, G Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64 (6): 1263–97. [140, 141]

Ottonello, Pablo, and Thomas Winberry. 2020. "Financial Heterogeneity and the Investment Channel of Monetary Policy." *Econometrica* 88 (6): 2473–502. [104, 107, 109, 132, 136, 137]

Rudebusch, Glenn, and Eric Swanson. 2012. "The bond premium in a DSGE model with long-run real and nominal risks." *American Economic Journal: Macroeconomics* 4 (1): 105–43. [119, 120]

Salgado, Sergio, Fatih Guvenen, and Nicholas Bloom. 2019. "Skewed business cycles." *NBER Working Paper No. 26565, National Bureau of Economic Research.* [104, 110, 132, 139]

Sedláček, Petr, and Vincent Sterk. 2017. "The growth potential of startups over the business cycle." *American Economic Review* 107 (10): 3182–210. [110]

Smets, Frank, and Rafael Wouters. 2003. "An estimated dynamic stochastic general equilibrium model of the euro area." *Journal of the European Economic Association* 1 (5): 1123–75. [129]

Smets, Frank, and Rafael Wouters. 2007. "Shocks and frictions in US business cycles: A Bayesian DSGE approach." *American Economic Review* 97 (3): 586–606. [105, 128–130]

Storesletten, Kjetil, Chris I. Telmer, and Amir Yaron. 2001. "The welfare cost of business cycles revisited: Finite lives and cyclical variation in idiosyncratic risk." *European Economic Review* 45 (7): 1311–39. [121]

Suveg, Melinda. 2021. "Does Firm Exit Raise Prices?" *IFN Working Paper No. 1414, Research Institute of Industrial Economics.* [132]

Wang, Olivier, and Iván Werning. 2022. "Dynamic oligopoly and price stickiness." *American Economic Review* 112 (8): 2815–49. [106, 119, 129]

Winberry, Thomas. 2021. "Lumpy investment, business cycles, and stimulus policy." *American Economic Review* 111 (1): 364–96. [103, 117]

Publikationsverzeichnis

Gnewuch, Matthias. 2022. „Spillover effects of sovereign debt-based quantitative easing in the euro area“. *European Economic Review* 145. https://doi.org/10.1016/j.euroecorev.2022.104133.