



Uncertainty, Investment and Firm Dynamics

Siavash Mohades

PhD Dissertation to be defended at LUISS Guido Carli and Maastricht University

Supervisors:

Prof. Maria Savona

Prof. Tania Treibich

Prof. Giulia Piccillo

September 2024

Acknowledgments

This thesis is the final result of four years of hard work, support, and guidance by so many individuals at LUISS Guido Carli and Maastricht University, and it will always be something for which I am infinitely grateful.

First and foremost, I would like to express my gratitude to my supervisors, Professor Maria Savona, Professor Tania Treibich, and Professor Giulia Piccillo. Your insightful feedback, patience, and encouragement have been determinative in this chain of work that brings forth this thesis. Your craft in the various fields guided me through the intricacies of my research, and I learned something new from each one of you. I learnt from the three of you to be patient and always foresee younger colleagues' potential. You believed in me and my capabilities when many did not. I will take you as examples in teaching and coaching, and I am confident I have chosen good examples.

I would also like to express my gratitude to LUISS Guido Carli and Maastricht University for providing the academic setting and means to put together this work. Your institutions have been instrumental in the completion of this thesis. I am deeply grateful for the opportunities and resources you have provided. Particularly, I would like to thank the faculty members of both universities who pave the way for those stepping into their footprints in the future daily. Furthermore, I would like to thank the administration offices at both institutes, who helped shape a joint degree contract with time-varying conditions.

I wish to express my profound gratitude to my colleagues and friends who comprise the Department of Economics and Finance at LUISS and MILE at Maastricht University. Your repeated instances of comradeship, as well as the intellectual discussions we've had, have significantly lightened my work and made it more enjoyable. I am especially grateful for the support given to me by the PhD community at both institutes. Your critical impetus for important parts of my work has been invaluable. Most importantly, I felt at home at both institutes, and both communities' presence was sorely missed when I was attending the office at the other.

Above all, I will be eternally grateful to my family. Without Neira's unconditional support, my PhD would have never been anywhere close to a conclusion. Her love, patience, and unwavering belief in me provided the strength I needed during the most challenging moments of this journey. She stood by my side at my lowest and did not allow a 1,000-mile distance between Maastricht and Rome to even put a shadow of doubt on our warm, eternal love. Neira's encouragement and sacrifices have been my constant source of inspiration, and I owe much of this achievement to her steadfast support. My efforts in the following years and decades will focus on repaying what she has done for me over the past six years. You remained my closest friend and soulmate all this time, and we held each other's hands, watching the world turn upside down and back. Thank you, Neira, for all you have done. Let's enjoy *our life* together!

To Mom, Dad, Sahand and Mahshid:

بعد از چهار سال و نیم، بالاخره دوره دکترای من هم تموم شد. مطمئنم که به سرانجام رسیدن این مسیر بدون کمک های شما از روز اول امکان پذیر نبود. از روز اول کلاس اول دبستان که هر روز و شب به فکر من بودید و در کنارم بودید تا همین روزهای آخر دکترای من. ممنونم بابت همه کنارم بودن ها، حتی وقتی تصور یک دهم موفقیت های امروز غیر ممکن بود. این مدرک بیشتر به خاطر زحمات و صبر شماست و نه من. پژوهش که کار سختی نیست، کار مهم و اصلی مثل شما بودن و رفتار کردن تو زندگیه.

ممنون که هر روز و شب در کنار من و در قلب من بودید، حتی وقتی برای تلاش برای آینده ای بهتر مجبور به ترک اصفهان و بعد ایران شدم. امیدوارم بتونم توی این دوران جدید زندگی پسر و برادر بهتری باشم و مهم تر از همه، وقت بیشتری با شما بگذرونم.

To Jasmina, Zekerijah and Elma, heel erg bedankt dat jullie door de jaren heen mijn familie zijn geworden. Jullie onwrikbare steun door de jaren heen is hartverwarmend geweest en daar ben ik dankbaar voor.

Lastly, I would like to thank Amir, who has been the best comrade I could have imagined. We have been side by side from our bachelor days in Tehran to Rome and the Netherlands. Thank you, Amir, for remaining my longest and dearest friend through thick and thin.

I thank you all for being indispensable parts of this journey. I am grateful that I had the opportunity to share it with you as my friends, family, and colleagues.

Contents

Acknowledgments	3
Introduction	9
References	13
1 Unpacking Economic Uncertainty	15
1.1. Introduction	16
1.2. Literature Review	18
1.3. Data and descriptive statistics	20
1.3.1. Data and variables	20
1.3.2. Descriptive statistics	22
1.4. Methods: measuring economic uncertainty	22
1.4.1. Conceptual approach	22
1.4.2. Firm-Level Uncertainty (FLU)	24
1.4.3. Sector-Level Uncertainty (SLU)	24
1.4.4. Aggregate-Level Uncertainty (ALU)	26
1.4.5. Estimation Method: Dynamic Panel using a GMM Estimator	27
1.4.6. Descriptive Statistics of Uncertainty	27
1.5. Validation	28
1.5.1. Evolution of Uncertainty Over Time	28
1.5.2. Uncertainty and the Business Cycle	29
1.5.3. Comparison with existing measures	31
1.6. Heterogeneity of uncertainty across firms and sectors	32
1.6.1. Heterogeneity in types of uncertainty across firms	32
1.6.2. The role of size and age in explaining uncertainty heterogeneity	33
1.6.3. Size and uncertainty heterogeneity	33
1.6.4. Age and uncertainty heterogeneity	35
1.6.5. Heterogeneity of uncertainty across sectors	37
1.7. Robustness Checks	38
1.8. Conclusion	39
References	40
Appendix	45
A. Additional Graphs	45
B. Regressions of Uncertainty Derivation	46
2 Firms' Investment and Capacity Utilisation	47
2.1. Introduction	48
2.2. Conceptual Framework	51

2.2.1. Investment & Capacity Utilisation	51
2.2.2. The Role of Uncertainty and Financial Constraints	53
2.3. Empirical Approach	53
2.3.1. Estimation Methods	55
2.4. Data and Variables	57
2.4.1. Dataset	57
2.4.2. Variables	57
2.4.3. Descriptive Statistics	62
2.4.4. Investment, Capacity Utilisation and Business Cycle	63
2.5. Results	67
2.6. Robustness Checks	72
2.6.1. Timing of the Survey: From Forward to Backward Looking	72
2.6.2. Alternative Investment Rate Specifications	72
2.6.3. Firm-Specific Capacity Utilisation Target	73
2.6.4. Expectation-Based Uncertainty Measures	73
2.6.5. Frontier Firms Effect	74
2.6.6. Can Entry and Exit drive our results?	74
2.7. Discussion & Conclusion	74
References	76
Appendix	84
A. Additional Graphs and Tables	84
A.1. Capacity Utilisation Heterogeneity at the Sector Level	84
A.2. Market Dynamics in Italy	84
A.3. Sectoral Division of INVIND Data	85
A.4. Autocorrelation of Uncertainty Measures	85
A.5. Summary Statistics for Growth Window and Non-Growth Window	86
A.6. First-stage Regression	87
B. Further Analysis - Regression Tables	87
C. Robustness Checks - Regression Tables	91
D. Questions Used from the INVIND Survey	110
3 Data, Uncertainty and R&D Investments	113
3.1. Introduction	114
3.2. Theoretical Framework	116
3.2.1. From Data to R&D	116
3.2.2. Complementarity of Data and R&D	117
3.3. The Scope of Our Study: Choice of Italy, Data & Variables	118
3.3.1. Benchmarking Italy: Intangible Investments and Productivity	118
3.3.2. Data and Variables	119
3.4. Empirical Strategy	121
3.4.1. GDPR as an Instrument for Data	122
3.4.2. More data, more R&D?	125
3.4.3. Estimating Complementarity of Data and R&D Investments for Productivity	126
3.5. Results	127

3.5.1. First-stage Regression: GDPR and Investments in Data	127
3.5.2. Second stage regression - Data and R&D	128
3.5.3. Complementarity Effects	130
3.6. Robustness Checks	132
3.6.1. Inclusion of Fixed Effects	133
3.6.2. Manufacturing vs Services	133
3.6.3. GDPR 2016 instead of 2018	133
3.6.4. Data and R&D Per Labour	133
3.6.5. Estimating Complementarity via Instrumental Variable	134
3.6.6. Pre-Post GDPR Subsample Study	134
3.7. Conclusion	134
References	136
Appendix	143
A. Additional Tables	143
B. Lagged Productivity Gains	147
C. Robustness Checks - Regression Tables	149
Summary	157

Introduction

"As a rule, panics do not destroy capital; they merely reveal the extent to which it has been previously destroyed by its betrayal into hopelessly unproductive works."

— John Stuart Mill

Economic growth is widely understood to be the backbone of improving business conditions and living standards. As a result, over the past century, much of the attention in the economics literature has focused on drawing models that identify and isolate the contribution of various inputs to growth. These studies have focused on the role of (in)tangible capital accumulation, human capital enrichment and technology progress. While there is conflicting evidence on the cross-country role of capital accumulation via investments, the slowdown of corporate investments in the 21st century has coincided with a period of persistent low productivity growth (Acemoglu et al., 2014; Gutiérrez and Philippon, 2016). Plausibly, scholars' efforts have focused increasingly on reformulating the theoretical and empirical models that study the determinants of investments, the role of capital accumulation in growth and the role of (lack of) investments in the productivity growth slowdown.

Firms' low investments in new products and manufacturing lines can be due to multiple factors. First, the financialisation of the advanced economies as a result of globalisation and the increasing competitiveness of the emerging markets has moved the inputs towards the services and finance sectors (Rodrik, 2016)- which are less capital-intensive than manufacturing. Second, environmental and redistributive regulations and policies have lowered manufacturing firms' investments (Falk and Shelton, 2018). Third, and most importantly, demand uncertainty coupled with lower aggregate demand growth compared to the 20th century and lower market power has curbed investments (Bontempi, Golinelli and Parigi, 2010). In this thesis, we argue that the perception of uncertainty about future demand and investment outcomes is sizeably heterogeneous across firms and drives the low expansion investments of firms. Furthermore, realising a part of firms' overall uncertainty via higher investments in data encourages firms' R&D investments.

Interestingly, the literature in the decades prior to this slowdown had already focused on the concerning role of uncertainty in driving investments down (Abel and Eberly, 1994; Dixit, Dixit and Pindyck, 1994). Economic uncertainty acts as a significant determinant of firm behaviour and outcomes, such as growth, investment, and innovation. Firms' decisions regarding inputs and production are deeply intertwined with their uncertain expectations of future sales. These decisions are often irreversible or subject to high adjustment costs. During periods of heightened uncertainty, firms tend to delay or scale back their investments and hiring to avoid these potential costs. Uncertainty can also affect business owners' confidence about the future and affect their perspectives regarding growth (Bachmann, Elstner and Sims, 2013). At the macroeconomic level, these effects can significantly contribute to business cycle downturns (Bloom et al., 2018; Kumar, Gorodnichenko and Coibion, 2023; Bianchi, Kung and Tirskikh, 2023).

However, unlike other economic indicators such as GDP or inflation, uncertainty is not directly measurable or observable. It manifests through various channels, such as changes in demand or supply chains, policy shifts, or unexpected global events, making it a complex concept to quantify. Therefore, not only does tracking its cycles become challenging, but assessing its various impacts proves complex as well. To this end, we introduce a new method to measure uncertainty at the firm, sector, and aggregate levels using firm-level data and reveal firm and sector-level heterogeneities in levels and the decomposition of uncertainty. By applying this method to Italian firms' data, we find that higher uncertainty collides with an increase in low-growth potential firms' investments, but it does not affect high-growth potential firms' investments in plants and machinery. Furthermore, investments in data are found to increase R&D investments, potentially via lowering uncertainty since data acts as information and allows the firm to manage its various sources of uncertainty.

The first chapter introduces a novel method to measure and decompose economic uncertainty using widely available firm-level data sources. The availability of this data for the majority of the economies implies that this method is applicable to most countries. We apply this uncertainty measure in the second chapter to explore its interplay with capacity utilisation, financial constraints, and firm growth to reveal the role of uncertainty as a barrier during growth periods. The third chapter builds on the insights from the first two, demonstrating how firms' investments in data can reduce uncertainty and enhance their willingness to invest in innovative activities such as R&D.

Chapter 1: Unpacking Economic Uncertainty: Measuring the Firm, Sector and Aggregate Components

Existing measures of uncertainty are typically estimated in a preliminary step using market data (Bloom, Bond and Van Reenen, 2007), media sources (Baker, Bloom and Davis, 2016), or macroeconomic indicators (Jurado, Ludvigson and Ng, 2015), and then treated as if they were observable data series (Carriero, Clark and Marcellino, 2018). These measures often fail to capture the complexity and heterogeneity of firms' salient uncertainty. While efforts have been made to create disaggregated uncertainty measures, much of the economic literature on the impact of uncertainty on firms' decisions has focused on aggregated uncertainty, assuming a uniform uncertainty process across firms. To address this issue, this chapter raises two questions: "How is firm uncertainty decomposed into firm, sector, and aggregate sources? To what extent does heterogeneity of uncertainty exist across firms?"

This chapter introduces a novel method to measure firms' uncertainty, proxied by their sales volatility, and decomposes a multi-layered system of indices: commonalities among all firms (potentially driven by macro-level developments such as economic policy and aggregate demand shocks), sectoral commonalities (such as supply chain disruptions, technological shocks, and input price variations), and firm-specific factors. The approach used to measure uncertainty involves disaggregated firm-level data to construct different measures of uncertainty at aggregate, sectoral, and firm levels by decomposing the volatility of firms' sales, termed "Overall Uncertainty" (OU). The chapter also contributes to the literature by introducing the concept of uncertainty heterogeneity and emphasises the differences in firms' perceptions of uncertainty and its origins, as well as the varying impacts of uncertainty on different firms and sectors.

The chapter validates this approach by demonstrating that the computed aggregate-level uncertainty for the U.S. economy aligns with existing measures of macroeconomic uncertainty, such as those by Baker, Bloom and Davis (2016) and Jurado, Ludvigson and Ng (2015). Using data from the Quarterly Compustat database, our results reveal significant heterogeneity in uncertainty across

firms based on their characteristics and sectors. For example, larger firms tend to experience lower uncertainty than smaller firms. Moreover, high uncertainty episodes are more pronounced in the manufacturing sector compared to the services and mining sectors, corroborating findings on the sectoral heterogeneity of uncertainty by [Born and Pfeifer \(2021\)](#) and [Parast and Subramanian \(2021\)](#).

These findings have significant implications for policymakers, particularly in understanding the vulnerability of smaller firms to macroeconomic conditions. The results underscore the need for fiscal and monetary policies that consider the differential impacts on firms' uncertainty, which in turn affects their performance, investment decisions, and hiring choices. This understanding can empower policymakers to make informed decisions, aligning with empirical evidence on the stagnation of U.S. investments and industrial production over recent decades, especially in the manufacturing sector.

Chapter 2: Firms' Investment & Capacity Utilisation: The Role of Uncertainty and Financial Constraints

The second chapter in this thesis examines the role of capacity utilisation in firms' investments. Previous research has identified a trigger point in capacity utilisation at which firms typically invest and expand ([Brown and Mawson, 2013](#); [Coad et al., 2021](#)). In essence, once firms reach a "capacity target", some of them invest in new production capacity to prepare for periods of high demand. However, the previous efforts in the literature studying the relationship between capacity utilisation and investments of growth have not studied the role of uncertainty and financial constraints.

Our empirical analysis uses the INVIND survey data from the Bank of Italy for the period 2002-2023 and documents a structural break in the capacity utilisation of Italian manufacturing firms following the Great Recession and the European debt crisis. This shift is accompanied by a simultaneous decline in investments as documented previously by [Bond, Rodano and Serrano-Velarde \(2015\)](#); [Buseti, Giordano and Zevi \(2015\)](#), alongside heightened uncertainty and tighter financial constraints. Consequently, low capacity utilisation, coupled with high uncertainty and constrained financing, may have contributed to the infrequent expansion of Italian firms during the last two decades. We pose the following research questions: "Does the probability of investing in production capacity depend on firms' capacity targets and the extent to which they exceed their capacity utilisation target?" and "How do financial constraints and demand uncertainty influence the relationship between capacity utilisation and investment?"

The empirical analysis in this chapter includes applying an instrument variable analysis to solve for the endogeneity between financial constraints and investments. Additionally, by constructing investment spikes following [Grazzi, Jacoby and Treibich \(2016\)](#); [Arata \(2019\)](#), we isolate the expansion investments, which represent the firm growth from other types of investments, such as replacement for wear and tear. The impact of capacity utilisation on investment is measured via the difference between the observed capacity and the target, which is measured as the sectoral average conditional on the business cycle. Furthermore, we construct a dummy for observations with capacity utilisation higher than the target and call these observations interchangeably at the growth window and operating overcapacity. By using this dummy, we document the non-linearity in the relationship between capacity utilisation and investment. Briefly, the extent of operations above the target is less important than reaching the target.

Our findings confirm that firms with high capacity utilisation are more inclined to invest compared to those with lower capacity utilisation. This chapter sheds light on the heterogeneous impact of uncertainty on firms. The results suggest that firms operating at high capacity are not affected by

uncertainty, whereas those at lower capacity levels exhibit increased investments during uncertain times. This chapter contributes to the literature on firm-level heterogeneity by introducing a new dimension based on capacity utilisation and uncertainty. The results have significant implications for industrial policy, suggesting that identifying and understanding these trigger points can aid in designing cost-effective strategies to encourage investment during critical growth periods.

Chapter 3: Data, Uncertainty and R&D Investments

The third chapter explores the effect of higher investments in data on firms' Research and Development (R&D) activities. We hypothesise that investing in data results in lower uncertainty, thereby motivating firms to engage more in innovative activities like R&D. It then moves on to studying the multiplier effect of these investments on improving firms' productivity to establish the complementarity among these investments. Therefore, this chapter hypothesises that investing more in data leads to more innovation via a reduction in uncertainty of the firm. Furthermore, the chapter examines whether the combination of these two investments increases firms' labour productivity through more data-driven decision-making that leads to more efficiency. Ultimately, the chapter poses the questions: "Does investing in data affect R&D investments?" and "Is there a cumulative effect on productivity from sole investments in data, sole investments in R&D, and joint investments in data and R&D? Is there a multiplier effect?"

Investing in data results in lower uncertainty, which reduces marginal production costs. Lower costs yield larger market concentration and firm size, which allows the firm to boost its R&D investments (Eeckhout and Veldkamp, 2022; Brynjolfsson, Jin and Wang, 2023). Using an instrumental variable (IV) approach with data from the INVIND survey of the Bank of Italy between 2002-2023, this chapter identifies a causal impact of data investments on R&D. The econometric framework in this chapter uses the introduction of the General Data Protection Regulation (GDPR) as an instrument for data to address the endogeneity between data and R&D. By treating data and analytics as information, the study demonstrates that reducing uncertainty is a primary benefit of these investments, which can facilitate firms' innovative activities. The results show a causal relationship between investments in data and software and the tendency to invest in innovative activities, highlighting the potential for a virtuous cycle between investments in intangible assets and innovation.

Importantly and to measure firms have gained efficiency-related gains, we test whether the combination of these investments contributes to higher labour productivity in both short and long terms, building on the previous literature on the role of intangibles in improving productivity, such as Corrado et al. (2021); Añón Higón, Gómez and Vargas (2017); Goodridge and Haskel (2023). The results in this chapter indeed confirm the conjecture that investing in these intangible assets and their combination positively impacts productivity. Therefore, not only does investing more in data lead to more innovation, but their usage and combination positively affects labour productivity. These findings have important implications for academic research and policy and suggest that data and analytics investments can reduce uncertainty and enhance firm performance, ultimately encouraging a more innovative and productive business environment.

References

- Abel, Andrew, and Janice Eberly.** 1994. “A Unified Model of Investment under Uncertainty.” American Economic Review, 84(5): 1369–84.
- Acemoglu, Daron, David Autor, David Dorn, Gordon H. Hanson, and Brendan Price.** 2014. “Return of the Solow Paradox? IT, Productivity, and Employment in US Manufacturing.” American Economic Review, 104(5): 394–99.
- Añón Higón, Dolores, Jaime Gómez, and Pilar Vargas.** 2017. “Complementarities in innovation strategy: do intangibles play a role in enhancing firm performance?” Industrial and Corporate Change, 26(5): 865–886.
- Arata, Yoshiyuki.** 2019. “Firm growth and Laplace distribution: The importance of large jumps.” Journal of Economic Dynamics and Control, 103: 63–82.
- Bachmann, Rüdiger, Steffen Elstner, and Eric R. Sims.** 2013. “Uncertainty and Economic Activity: Evidence from Business Survey Data.” American Economic Journal: Macroeconomics, 5(2): 217–49.
- Baker, Scott R., Nicholas Bloom, and Steven J. Davis.** 2016. “Measuring Economic Policy Uncertainty*.” The Quarterly Journal of Economics, 131(4): 1593–1636.
- Bianchi, Francesco, Howard Kung, and Mikhail Tirsikh.** 2023. “The origins and effects of macroeconomic uncertainty.” Quantitative Economics, 14(3): 855–896.
- Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry.** 2018. “Really Uncertain Business Cycles.” Econometrica, 86(3): 1031–1065.
- Bloom, Nick, Stephen Bond, and John Van Reenen.** 2007. “Uncertainty and Investment Dynamics.” The Review of Economic Studies, 74(2): 391–415.
- Bond, Stephen R, Giacomo Rodano, and Nicolas Andre Benigno Serrano-Velarde.** 2015. “Investment dynamics in Italy: Financing constraints, demand and uncertainty.” Bank of Italy Occasional Paper, , (283).
- Bontempi, Maria Elena, Roberto Golinelli, and Giuseppe Parigi.** 2010. “Why demand uncertainty curbs investment: Evidence from a panel of Italian manufacturing firms.” Journal of Macroeconomics, 32(1): 218–238.
- Born, Benjamin, and Johannes Pfeifer.** 2021. “Uncertainty-driven business cycles: Assessing the markup channel.” Quantitative economics, 12(2): 587–623.
- Brown, Ross, and Suzanne Mawson.** 2013. “Trigger points and high-growth firms: A conceptualisation and review of public policy implications.” Journal of Small Business and Enterprise Development, 20(2): 279–295.
- Brynjolfsson, Erik, Wang Jin, and Xiupeng Wang.** 2023. “Information technology, firm size, and industrial concentration.” National Bureau of Economic Research.
- Busetti, Fabio, Claire Giordano, and Giordano Zevi.** 2015. “Main Drivers of the Recent Decline in Italy’s Non-Construction Investment.” Bank of Italy Occasional Paper, , (276).

- Carriero, Andrea, Todd E Clark, and Massimiliano Marcellino.** 2018. “Measuring uncertainty and its impact on the economy.” Review of Economics and Statistics, 100(5): 799–815.
- Coad, Alex, Clemens Domnick, Florian Flachenecker, Peter Harasztosi, Mario Lorenzo Janiri, Rozalia Pal, and Mercedes Teruel.** 2021. “Capacity constraints as a trigger for high growth.” Small Business Economics, 1–31.
- Corrado, Carol, Chiara Criscuolo, Jonathan Haskel, Alexander Himbert, and Cecilia Jona-Lasinio.** 2021. “New evidence on intangibles, diffusion and productivity.”
- Dixit, R.K., A.K. Dixit, and R.S. Pindyck.** 1994. Investment Under Uncertainty. Princeton University Press.
- Eeckhout, Jan, and Laura Veldkamp.** 2022. “Data and market power.” National Bureau of Economic Research.
- Falk, Nathan, and Cameron A Shelton.** 2018. “Fleeing a lame duck: Policy uncertainty and manufacturing investment in US states.” American Economic Journal: Economic Policy, 10(4): 135–152.
- Goodridge, Peter, and Jonathan Haskel.** 2023. “Accounting for the slowdown in UK innovation and productivity.” Economica, 90(359): 780–812.
- Grazzi, Marco, Nadia Jacoby, and Tania Treibich.** 2016. “Dynamics of investment and firm performance: Comparative evidence from manufacturing industries.” Empirical Economics, 51: 125–179.
- Gutiérrez, Germán, and Thomas Philippon.** 2016. “Investment-less Growth: An Empirical Investigation.” National Bureau of Economic Research Working Paper 22897.
- Jurado, Kyle, Sydney C. Ludvigson, and Serena Ng.** 2015. “Measuring Uncertainty.” American Economic Review, 105(3): 1177–1216.
- Kumar, Saten, Yuriy Gorodnichenko, and Olivier Coibion.** 2023. “The effect of macroeconomic uncertainty on firm decisions.” Econometrica, 91(4): 1297–1332.
- Parast, Mahour Mellat, and Nachiappan Subramanian.** 2021. “An Examination of the Effect of Supply Chain Disruption Risk Drivers on Organizational Performance: Evidence From Chinese Supply Chains.” Supply Chain Management: An International Journal, 26(4): 548–562.
- Rodrik, Dani.** 2016. “Premature deindustrialization.” Journal of economic growth, 21: 1–33.

Chapter 1

Unpacking Economic Uncertainty¹

Measuring the Firm, Sector and Aggregate Components

Abstract

We introduce a novel method for measuring economic uncertainty at the firm, sector, and aggregate levels using sales volatility and validate it by comparison with existing macroeconomic uncertainty measures. We use firm-level data from Compustat from 2000-2022 to construct our uncertainty measures for the U.S. economy. Our findings highlight that i) macroeconomic conditions are the predominant source of firms' uncertainty, ii) diverse firm traits yield notable heterogeneity, and iii) the manufacturing sector exhibits the highest level of uncertainty. Our findings shed light on the importance of firm and sectoral heterogeneity in studying uncertainty and its effects on economic activity.

Keywords: Measuring Uncertainty, Firm Heterogeneity, Balance Sheet Data, Business Fluctuations

¹This chapter is co-authored with Giulia Piccillo and Tania Treibich

1.1 Introduction

"No credible theory of investment can be built on the assumption of perfect foresight and arbitrage over time."
— Robert M. Solow

Economic uncertainty is a pervasive phenomenon that has significant implications for macroeconomic theory and firms' decisions. Such importance has been evidenced by its increased salience following major economic events, such as the Great Recession and the COVID-19 pandemic. In these recent major events, uncertainty has been shown to be both a symptom and a potential driving force of economic turmoil (Bloom, 2009; Alfaro, Bloom and Lin, 2023). From a microeconomic point of view, economic uncertainty affects firms' decisions through several channels with substantial macroeconomic effects (Bloom et al., 2018; Kumar, Gorodnichenko and Coibion, 2023; Bianchi, Kung and Tirsikh, 2023).

Firms' decisions on inputs and production rely on their expectations of future sales, which are informed by past sales and inventory dynamics (Gennaioli, Ma and Shleifer, 2016). These decisions are often irreversible or subject to high adjustment costs, which can lead to significant deadweight losses if made incorrectly. However, forming accurate expectations is difficult - and the resulting uncertainty can be just as costly as adjustment costs or irreversibility losses. In uncertain times, firms struggle to form precise expectations of future sales, causing delays or reductions in investment and hiring (Dixit, Dixit and Pindyck, 1994; Abel and Eberly, 1994; Abel et al., 1996). This inefficiency can lead to significant macroeconomic slumps².

Economic uncertainty is multifaceted, as various factors, including changes in government policies, global economic events, natural disasters, and technological progress, can cause it. From a firm's perspective, uncertainty rooted in aggregate, sectoral and firm conditions requires fundamentally different decisions to buffer against. Furthermore, economic uncertainty is a complex phenomenon because it can arise from "risk", "ambiguity" or "model misspecification" (Cagetti et al., 2015; Hansen and Sargent, 2021; Aydogan et al., 2023)³. For example, media coverage and public sentiment regarding a significant event can influence economic uncertainty, impacting consumer behaviour and investment decisions even further. Finally, a firm might be impacted by an economic (or policy) event in two different ways. On the one hand, the exposure to the event may be heterogeneous (different impacts on the firm's demand and or/and production). On the other hand, the perception of the event might also differ (i.e. the firm's perception of the extent of uncertainty itself – the uncertainty of uncertainty). Additionally, uncertainty levels and dynamics need not be identical across firms and sectors. For instance, the closure of mines and industries in China during the COVID outbreak impacted the manufacturing and mining sectors globally, while services were less affected (Ozili and Arun, 2023). This heterogeneity underscores the significance of complexity and diverse pathways through which uncertainty can escalate and influence the business environment.

Existing measures of uncertainty are often based on aggregate data, such as media coverage (Baker, Bloom and Davis, 2016) and macroeconomic variables (Jurado, Ludvigson and Ng, 2015) and often fail to capture the complexity and heterogeneity of firms' salient uncertainty. Furthermore, these measures require large datasets that might not be available in all countries, and therefore, they are not fully replicable. Even though there have been efforts to provide disaggregated uncertainty measures, the

²For more details, please refer to Bloom (2009); Dixit, Dixit and Pindyck (1994)

³This concerns the nature of production and consumption activities - for instance, supply chain developments, change of habits and consumers' taste, among others.

economic literature on the impact of uncertainty on firms’ decisions has mainly focused on aggregated uncertainty by assuming an identical uncertainty process for all firms, for instance, in [Bloom, Bond and van Reenen \(2007\)](#). Therefore, we tackle this gap in the literature and present our research question: “How is firm uncertainty decomposed into firm, sector, and aggregate sources? To what extent does heterogeneity of uncertainty exist across firms?”

We uncover and introduce uncertainty heterogeneity across sectors and firms by decomposing firms’ sales volatility into a multi-layered system of indices: a component that roots in commonalities of all firms (potentially macro-level developments such as economic policy and aggregated demand shocks), a component that originates in sectoral commonalities (for instance through supply chain disruptions, technological shocks and input prices) and finally a component that is specific to each firm through its decisions and activities, which ideally should be the idiosyncratic tranche of uncertainty. We use each layer to reveal differences across firms in the contribution of macroeconomic, sectoral and firm-level sources of uncertainty and exploit time and cross-sectional variations in the different measures.

Our methodology consists of first calculating the volatility of firms’ sales or “Overall Uncertainty” (OU).⁴ We then identify the proportion of uncertainty which is specific to firms by controlling for their characteristics. The sector-level uncertainty is then calculated by controlling for a common factor in each sector, computed using firms’ balance sheet data. The weighted remainder of overall uncertainty determines the aggregate uncertainty after the firm and sector-level uncertainty components have been identified and deducted. We validate our approach by showing how our computed aggregate level uncertainty is consistent with existing measures of macroeconomic uncertainty in capturing major economic events such as the Great Recession and the COVID outbreak for the U.S. economy.

To our knowledge, we are the first to use a single disaggregated firm-level data source to jointly construct different measures of uncertainty at aggregate, sectoral, and firm levels. Our approach proposes a measure that is as good as the widely used and accepted measures of uncertainty in highlighting the macro-level uncertainty among firms ([Baker, Bloom and Davis, 2016](#); [Jurado, Ludvigson and Ng, 2015](#)) while relying on an accessible dataset. The resulting measures can describe the commonalities of uncertainty among firms and sectors and uncover sources of uncertainty heterogeneity by relaxing the assumption of a common uncertainty process for all. This approach is robust to the issues associated with measuring uncertainty from the unpredictable component of the sales process - such as heteroscedasticity caused by the high persistence of firms’ sales. While we apply our method to the firms present in the quarterly Compustat database in the United States, it can be easily applied to any other firm-level database. Furthermore, it is distinguished from past efforts such as [Dew-Becker and Giglio \(2023\)](#), as it requires only real activity data (instead of financial and options data) to construct aggregate and disaggregate uncertainty measures.

Our aggregate and sectoral uncertainty measures are counter-cyclical and reveal that the majority of firms’ uncertainty originates in macroeconomic conditions. However, substantial heterogeneity in uncertainty and business cycle exists among firms based on their characteristics, sectors and whether uncertainty stems from a right or left tail event, as noted by [Born and Pfeifer \(2021\)](#); [Forni, Gambetti and Sala \(2024\)](#). Larger firms experience lower uncertainty than their smaller counterparts, while firms’ age exhibits differences only in uncertainty decomposition and not in levels. Notably, the manufacturing sector faces higher uncertainty than the services and mining sectors, as documented by [Parast and Subramanian \(2021\)](#).

⁴The literature has extensively discussed the importance of firms’ volatility as an indicator of economic uncertainty ([Bloom et al., 2018](#); [Comin and Philippon, 2005](#)). We use firms’ sales volatility and Overall Uncertainty interchangeably in this paper. Section 1.2 provides a more detailed discussion.

Our paper is structured as follows: Section 1.2 briefly reviews existing methods for measuring economic uncertainty. We then describe our dataset in Section 3.3, and the application of our approach to this data is explored in detail in Section 1.4. Section 1.5 demonstrates the evolution of uncertainty over time and documents our validation approach. Section 1.6 continues by studying the decomposition of uncertainty and exploring the heterogeneity of uncertainty. We provide a series of robustness checks in Section 3.6. Section 1.8 concludes our paper.

1.2 Literature Review

This section discusses several measures of economic uncertainty and points out how our approach offers novel insight by tackling the complexity of uncertainty. These approaches mainly differ by the information source, i.e., aggregate, survey, text-based or firm-level data.⁵

Surveys have provided a rich source for constructing ex-ante uncertainty measures. D’Amico and Orphanides (2008); Sill et al. (2012) used the probabilistic responses from the Survey of Professional Forecasters to construct their uncertainty measure for the U.S. economy. Similarly, Kozeniauskas, Orlik and Veldkamp (2018) used a macro uncertainty measure as the conditional variance of GDP growth forecasts for all agents in their model. Combining household and professional surveys forecast errors, Sheen and Wang (2017) construct ex-ante macroeconomic uncertainty estimates of the one-year ahead expected state of the economy. Similarly, one can measure ex-ante firms’ uncertainty using managers’ future expectations like aggregated data. For instance, Bachmann, Elstner and Sims (2013); Bachmann et al. (2021) used survey expectations data of manufacturing firms’ managers to construct empirical proxies for time-varying business-level uncertainty. Awano et al. (2018) used firms’ expectations of their own future turnover growth, their expectations of UK GDP growth for 2018 and the uncertainty around these expectations to measure firm-level uncertainty for a sample of U.K. businesses. Altig et al. (2019) elicited subjective probability distributions from business executives about their firm outcomes at a one-year look-ahead horizon.⁶ Lamorgese et al. (2024) measured uncertainty as the forecast error of firms at each point in time by comparing the expected value of sales in $t+1$ and the realised value. The primary issue with these methods is that surveys might reflect differences in opinions and not uncertainty per se or be impacted strongly by managers’ pessimism or optimism of the cycle. Furthermore, using survey data while market and priced data are available requires rigorous elaborations, which are mostly overlooked (Thomas, 2018).

A different approach relies on text-analysis methods to identify highly uncertain times by measuring the frequency of relevant terms to economic uncertainty in either media or other text sources. Baker, Bloom and Davis (2016) provided a text-based economic policy uncertainty index (EPU), which counts the frequency of articles containing the words “uncertain or uncertainty” and “economy or economics” in the leading media of more than 25 countries.⁷ Uncertainty at the political level can also affect firms’ investments and performance, as shown by Azzimonti (2018). Using the quarterly Economist Intelligence Unit country reports, Ahir, Bloom and Furceri (2022) constructed a quarterly World

⁵A more comprehensive review of the standard measurements of uncertainty and their empirical considerations is provided by Cascaldi-Garcia et al. (2023).

⁶In terms of question design, their crucial innovation is to let survey respondents freely select support points and probabilities in five-point distributions over future sales growth, employment, and investment. Their result - called Survey Business Uncertainty - is calculated monthly for a sample of the U.S. firms and is available online [here](#)

⁷The EPU index has been used excessively as a policy uncertainty measure. For instance, Gulen and Ion (2016) discovered a robust negative relationship between firm-level capital investment and the aggregate level of uncertainty associated with future policy and regulatory outcomes. A complete library of EPU indices for multiple countries is available at: <https://www.policyuncertainty.com/>.

Uncertainty Index (WUI) for an unbalanced panel of 143 individual countries from 1952 based on the frequency of the word “uncertainty”.

The economic literature has repeatedly used aggregate data to compute uncertainty. [Baum et al. \(2006\)](#) constructed four simple proxies for macroeconomic uncertainty using the conditional variances of real GDP, industrial production index, CPI inflation rate and returns on the S&P 500 stock-market index. In the same line, [Bloom, Bond and van Reenen \(2007\)](#); [Bloom \(2009\)](#) used share price volatility as an empirical proxy for uncertainty, justified by its positive correlation with the underlying (theoretical) standard deviation of demand shocks, real sales growth volatility and the cross-sectional distribution of financial analysts’ forecasts.

Differently, [Jurado, Ludvigson and Ng \(2015\)](#) (henceforth JLN) argued that the volatility in one or few macro indicators or policy uncertainty does not fully represent macroeconomic uncertainty, or “what matters for economic decision-making is not whether particular economic indicators have become more or less variable or dispersed per se, but rather whether the economy has become more or less predictable, that is, less or more uncertain” ([Jurado, Ludvigson and Ng, 2015](#), p. 1178). Their method consists of constructing a large macroeconomic and financial indicators dataset that extracts as much information as possible to build conditional expectations of each variable in $t+h$.⁸ A weighted average of these N macroeconomic indicators forecast errors constitutes the JLN uncertainty index. Although their method and generated series became a widely used standard for the U.S. economy, the difficulty of the method, its data hunger, and inability to provide disintegrated or heterogeneous uncertainty prevented a wide replication across other economies. [Rossi and Sekhposyan \(2015\)](#) built on the same idea as JLN but relied on the unconditional likelihood of the observed outcome. Their proposed index is the percentile in the historical distribution of forecast errors associated with the realised forecast error at each date. This approach compares the realised forecast error of a macroeconomic variable of interest with its historical forecast error distribution. A realisation at the right (left) tail of the distribution hints at high “good” (“bad”) uncertainty since it was challenging to predict conditionally on the information given to the model.

Volatility has been used as a reliable uncertainty proxy applied not just to aggregate data but also to firms’ output dispersion. Firms struggle to foresee future sales during high uncertainty periods due to unforeseen economic shifts or unpredictable idiosyncratic shocks. [De Veirman and Levin \(2018\)](#) derived a measure of firm-specific volatility in sales from firms’ balance sheet data. [Baum et al. \(2006\)](#) utilised the cross-sectional dispersion of firms’ cash-to-asset ratios as a proxy for firms’ uncertainty. [Panousi and Papanikolaou \(2012\)](#) used the shocks to firms’ production function as firm-level uncertainty. Similarly, [Bloom et al. \(2018\)](#) measured microeconomic uncertainty using the firm-level TFP dispersion. [Comin and Philippon \(2005\)](#) used the standard deviation of the annual growth rate of sales and employment of firms as a measure of firm volatility. [Davis et al. \(2006\)](#) improved this method by introducing the firm size as a normalising weight for this measure. [Kozeniauskas, Orlik and Veldkamp \(2018\)](#) followed this approach by comparing this firm-dispersion measure versus macro and survey-based uncertainty measures. The standard shortcoming of these measures is their inability to identify the source of uncertainty - whether it is rooted in aggregate, sector or firm-level components. Our work builds on this method and uses firms’ sales dispersion as a proxy for the source of uncertainty, but contributes to the literature by disentangling different sources of uncertainty firms face through different levels by decomposing the sales volatility or overall uncertainty.

⁸These expectations for any given macroeconomic variable are constructed using a FAVAR model for the h -step ahead forecast-error estimation.

Existing uncertainty measurements rely on ex-ante or ex-post differences between realised values and past conditional expectations, often using macro or aggregated data or forecasts from professional surveys. These methods offer limited insight into disaggregated sources driving uncertainty and mainly address microeconomic uncertainty. Our paper addresses this gap by quantifying and dissecting firms' volatility into three uncertainty tiers, unveiling heterogeneity among firms. Moreover, many existing measures are case-specific, with limited applicability to developing economies and potential bias from media sources. In contrast, balance-sheet firm-level data is widely available in most economies, including developing regions, making it a robust data source for measuring uncertainty in any economy.

1.3 Data and descriptive statistics

1.3.1 Data and variables

Our methodology relies on firm-level information about sales and inputs to disentangle firms' volatility into different tiers of uncertainty. We focus on implementing an analysis that constructs an uncertainty measure for the U.S. economy driven solely by firms' data. We use a sample of 24,762 publicly listed American companies registered within the U.S. borders from Compustat's quarterly dataset between 2000Q1 and 2022Q4. Below, we describe the variables used in our analysis.

Sales

We use firms' sales data since its first and second moments correlate strongly with the business cycle, and its dispersion can capture the uncertain dynamics of the firms. The cyclical nature of uncertainty arises since we assume uncertainty to be time-varying following [Hassler \(1996\)](#). Also, [Bloom \(2009, 2014\)](#) found uncertainty to be primarily counter-cyclical to economic activity, indicating the importance of any uncertainty measure to be time-varying. Figure 1.1 shows the distribution of sales growth in non-recessionary times and during recessions. During recessions, we observe a leftward shift of sales growth and an increase in its second moment compared to non-recessionary times. The standard deviation in recessionary times is 8% higher than in non-recessionary times, translating into more uncertainty about the firm's sales process. The leftward shift during recessions also results in a negative sales growth average.

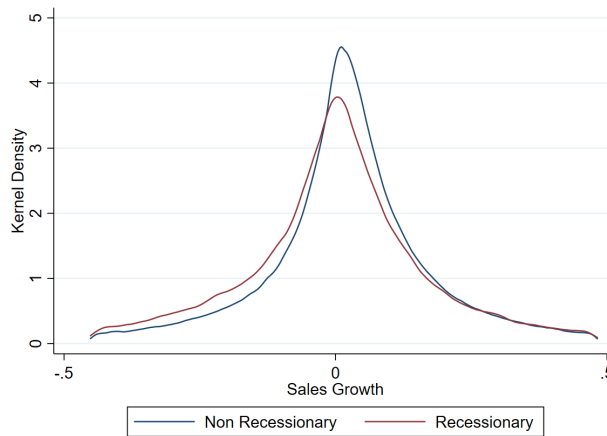


Figure 1.1: The distribution of firms' sales in recessionary and normal (non-recessionary) times. Source: Compustat database in quarterly frequency for 2000Q1-2022Q4.

Control variables and heterogeneity tiers

We use size, measured as total assets, age⁹ and capital intensity to capture the most significant variation of sales without multicollinearity (Herskovic et al., 2016, 2020; Uzeda, Tuzcuoglu and Castellnuovo, 2022). Firms’ performance, decisions and sensitivity to business cycles depend on their size and age (Clymo and Rozsypal, 2022). Smaller firms usually hold fewer assets, exposing them to low diversification and lower buffer (Aivazian, Rahaman and Zhou, 2019; Liang and Rhoades, 1991). Young firms have shorter track records, smaller customer bases, and less established reputations, making them more vulnerable to business cycles, shifts in consumer demand, and regulation changes (Comin and Mulani, 2006). They also have higher growth potential and may be more aggressive in pursuing new business opportunities, which can result in larger sales volatility. Moreover, young and small firms typically have less diversified product lines and revenue streams, which makes them more exposed to external shocks (Comin and Philippon, 2005). We also control for capital intensity, calculated as the ratio between total assets and total revenues. Additional variables in our analysis include cash flow, market value, total liabilities, and total revenue, from which we compute the commonality among sectors.

Our analyses then focus on the size and age to provide more insights regarding firms’ heterogeneity in uncertainty. Based on the distribution of firms’ total asset holdings, we divide firms into three groups: large, small and those in between. Small firms are those under the 25th percentile, and larger firms are those larger than the 75th percentile. Other firms’ total assets fall in the P25-P75 range. Our second heterogeneity tier of interest is the firms’ age. At each date, firms are grouped into their relevant subgroup based on the quartile they fall into, similar to our approach with firms’ size. We divide the distribution at each point into young and old firms and then average the values regardless of the firms entering or exiting. In this manner, we lose less information on the market dynamics than if we were to divide the sample into young and old from a whole distribution point of view.

Definition of sectors

There are multiple ways to define sectors within our data, from large partitions of many firms in each cluster to those which divide into smaller partitions. We use the 3-digit NAICS classification, which divides our sample into 111 subgroups. We call each of these groups a sector. From now on, when we use the term “sector-level uncertainty”, we refer to the uncertainty that arises in each of these subgroups. We choose to work with smaller clusters since our sample size is large enough, and the spillover effects are assumed to be larger and transformed more quickly across 3-digit clusters than broader groups. To obtain the three main “large sectors”, we re-sample our data into large sectors using 1 and 2-digit classifications of NAICS. Identifier 21 groups 1,051 firms into the mining group, 31-33 identifies the 6,844 manufacturing firms, and all firms with a first-digit identifier of either 5,6,7 or 8 are clustered under services, totalling 13,022 businesses.

Other macroeconomic indicators

In Section 1.5, we use two widely used economic uncertainty indices as benchmarks for evaluating our measure. Monthly EPU index from Baker, Bloom and Davis (2016) is averaged in each quarter to generate a quarterly index. The JLN by Jurado, Ludvigson and Ng (2015) is readily available in the

⁹We can only construct this variable for the listed firms based on their Initial Public Offering (IPO) date and the observation date.

quarterly format. In addition and to answer questions regarding the potential drivers of uncertainty in the manufacturing sector, we use the [ISM Supplier Delivery Index](#), which is available in a monthly format and average these values in each quarter to have a comparable index with respect to the rest of our analysis. Finally, recessions are defined via NBER recession bands for quarterly time series.

1.3.2 Descriptive statistics

Below, we provide the descriptive statistics of the main variables of interest that entered our analysis. It also becomes evident that the coverage of the quarterly Compustat dataset is much broader among listed firms.

Variable	Content	Observations	Mean	SE	Min	Max
S	Sales	675,919	832.37	4,120.43	-30,287.60	207,307.30
CH	Cash Flow	374,747	531.50	3,502.53	-9.44	288,606
TA	Total Assets	655,251	11,616.43	99,794.46	-12.67	4,289,453
MV	Market Value	381,924	4,494.23	29,507	0	2,901,645
LT	Total Liabilities	654,115	9,592.60	93,445.35	-0.62	4,233,271
TR	Total Revenue	612,567	859.05	4,183.90	-30,287.60	207,307.30
LI	Long-Term Investments	462,095	1,565.82	19,025.19	-5.09	2,033,489
Computed Variables						
Age (log)	Quarter - IPO	350,933	10.88	8.79	0	73.47
Capital Intensity	TA / TR	538,536	78.96	4,654.22	-280,715	2,469,984
Cash F / Tot Assets	CH / TA	372,743	0.17	0.23	-1.37	1.10

Table 1.3.1: Firm-Level Variables Descriptive Statistics

Source: Compustat database in quarterly frequency for 2000Q1-2022Q4. Geographical coverage includes firms registered in the United States and filed in the Compustat database.

1.4 Methods: measuring economic uncertainty

1.4.1 Conceptual approach

We begin by improving the approach applied to the firm-level data in [Comin and Philippon \(2005\)](#) and [Bloom et al. \(2018\)](#). Our method calculates firms' volatility as the difference between the realised anomaly of sales and the average deviation from the average. Hence, it does not differentiate the right and left-tail events and counts both as high-uncertainty events since it calculates the size of uncertainty as this distance.

These tail events are captured by large residuals of an empirical model that estimates the firms' sales. Our measure of uncertainty is a variation of the cross-sectional dispersion of firms' sales defined as the cross-sectional dispersion of the unpredictable component embedded in data as in [Uzeda, Tuzcuoglu and Castelnovo \(2022\)](#). It uses residual volatility as a proxy for uncertainty, as the cross-sectional dispersion corresponds to the part of the data which is left unexplained by the empirical model. We then disentangle different tiers of uncertainty and identify the fraction of sales uncertainty specific to each firm, sector or aggregate-level source.

We start with a measure of Overall Uncertainty (OU). This measure reflects sales volatility without imposing any firm or sector-characteristic controls. It captures the most and least unpredictable

uncertainty components of the sales process. Next, we will add a series of controls to manipulate the variations in sources of uncertainty.

OU is measured as the cross-sectional dispersion of shocks (i.e. residuals) $\varepsilon_{i,t}$ to the firm's net sales at time t compared to the firm's average $\bar{\varepsilon}_i$:

$$OU_{i,t} = \frac{(\varepsilon_{i,t} - \bar{\varepsilon}_i)^2}{T - 1} \quad (1.1)$$

Where T is the total number of quarters and the residual $\varepsilon_{i,t}$ is obtained from an AR(1) of firms' sales with fixed effects:

$$\log(S_{i,t}) = \rho \log(S_{i,t-1}) + \lambda_t + \mu_i + \varepsilon_{i,t} \quad (1.2)$$

Where $S_{i,t}$ denotes the firm i 's sales, μ_i is the firm-fixed effect to control for unobservable firm characteristics, and λ_t is the time-fixed effect to control for cyclical shocks. The OU measure introduced above is a weighted sum of uncertainty in three different levels: 1) aggregate Level Uncertainty (ALU), 2) weighted within-sector sectoral uncertainty (SLU), and 3) weighted across all firms firm-level Uncertainty (FLU).

$$OU_{i,t} = ALU_{i,t} + \omega_{i,t}^1 SLU_{i,t} + \omega_{i,t}^2 FLU_{i,t} \quad (1.3)$$

where the weights $\omega_{i,t}^1$ and $\omega_{i,t}^2$ represent the firm-level and sector-level shares of sales in the total sample, respectively:

$$\omega_{i,t}^1 = \left(\frac{\sum_{i=1 \in N_j}^{N_j} S_{i,t}}{\sum_{i=1 \in N}^N S_{i,t}} \right) \quad \& \quad \omega_{i,t}^2 = \left(\frac{S_{i,t}}{\sum_{i=1 \in N_j}^{N_j} S_{i,t}} \right) \quad (1.4)$$

In each step of our construction, we add a set of appropriate controls from firms' data so that the residuals can be used for measuring uncertainty components. We estimate two of these components (FLU and SLU) and compute ALU as the remainder of OU after FLU and SLU have been deducted.

The equation below shows an application of Equation 1.2 with a set of controls $X_{i,t}$:

$$y_{i,t} \equiv \log(S_{i,t}) = \rho \log(S_{i,t-1}) + \mu_i + \lambda_t + X_{i,t}\beta + \varepsilon_{i,t} \quad (1.5)$$

Where the conditional expectation $\mathbb{E}_t(y_{i,t}) = \rho \log(S_{i,t-1}) + X_{i,t}\beta$ provides an optimal prediction based on the Mean Squared Error (MSE) for the $M \times 1$ vector $y_{i,t}$ containing a set of firm-balanced sheet information with observations going from $t = 1$ to T , where M identifies the number of independent variables of interest, and T the time-series dimension or the number of quarters in our case. Therefore, the response variable's unpredictability (uncertainty) is associated with the error term $\varepsilon_{i,t}$.

To identify the economic uncertainty from the residuals of regression Equation 1.5 above, we need to choose a set of variables in $X_{i,t}$ that explain a considerable variation in the dependent variable $y_{i,t}$. To ensure that the error term $\varepsilon_{i,t}$ in Equation 1.5 is roughly unpredictable in each uncertainty measure, we define $X_{i,t}$ - the vector of independent variables and controls - in a way that extracts as much predictable component of firms' volatility as possible from the dependent variable - sales in our case (see also [Comin and Philippon, 2005](#)). We use combinations of fixed effects and control variables depending on the desired tier of uncertainty in each step. Table 1.4.1 below displays the set of control variables added to the $X_{i,t}$ matrix for estimating each level of uncertainty.

Measure	Time FE	Firm FE	Controls
OU	✓	✓	-
FLU	✓	✓	Age, Size & Capital Intensity
SLU	✓	✓	Sectors' Common Factors
ALU	The Residual as $ALU = OU - FLU - SLU$		

Table 1.4.1: Control variables and fixed effects in estimating each uncertainty level.

1.4.2 Firm-Level Uncertainty (FLU)

To identify the firm-level sources of uncertainty hidden and embodied in OU, we add firm-characteristic controls to our regression above and construct our preliminary FLU index similarly to OU.

$$\log(S_{i,t}) = \rho \log(S_{i,t-1}) + \lambda_t + \mu_i + \sum_{c=1}^C \beta_c X_{i,t} + \varepsilon_{i,t} \quad (1.6)$$

$$FLU_{i,t} = \frac{(\varepsilon_{i,t} - \bar{\varepsilon}_i)^2}{T - 1} \quad (1.7)$$

The set of controls in X contains capital intensity, firm age, and firm size, as guided by the previous literature. [Herskovic et al. \(2020\)](#) found that large firms are less volatile than their smaller counterparts due to more customer connections, improving the diversification of their client portfolio.¹⁰ We also consider firm age as the younger firms are more prone to experiencing firm relevant-business-conditions uncertainty such as down-sizing, CEO change and being bought out or listed. High capital intensity helps us isolate the effect of asset accumulation and the potential buffer for diversification purposes.

1.4.3 Sector-Level Uncertainty (SLU)

To measure sector-level uncertainty, one needs to estimate uncertainty at different layers of economic data and then model a large dataset to pin down all the implications that various economic system dimensions may have on uncertainty ([Uzeda, Tuzcuoglu and Castelnovo, 2022](#)). Therefore, sectoral uncertainty is viewed as a dynamic factor common to the time-varying volatility of a subset of variables corresponding to a particular sector. We repeat our previous process to measure SLU, but instead of controlling for firms' characteristics, we add a sectoral common factor ($CFact_{i,t}$) as a control. We propose to use a common factor among firms in each sector, similar to [Barigozzi et al. \(2014\)](#); [Herskovic et al. \(2020\)](#) as well as [Jurado, Ludvigson and Ng \(2015\)](#), to capture the sector-relevant variation. With its summarised but rich information, this constructed variable allows us to include a sector-related control variable that comes from the heart of our initial dataset and not external or aggregated vintages of firms' data.

We apply the dynamic factors approach ([Stock and Watson, 2011](#)) to summarise firms' data into one component that explains the co-movement and co-variation in that data-rich environment. With this variable combined with fixed effects, the isolated variation in uncertainty can be associated with unpredictable sector-specific developments that affect firms' sales. The common factor has the structure:

¹⁰It also implies that not only do small firms have higher uncertainty, but they also face higher uncertainty of uncertainty (larger second moment of uncertainty).

$$\text{Common Factor} \equiv CFact_{i,t} = \Lambda_j^{F'} \mathbf{F}_{i,t} \quad (1.8)$$

where $\mathbf{F}_t$ is an $r_F \times 1$ vector of latent common factors and Λ_j^F is a corresponding $r_F \times 1$ vector of latent factor loadings specific to each sector, defined as NAICS three-digit firm clusters. The $CFact_{i,t}$ variable is the common factor determining the commonality among firms in each sector j at time t . The factor loadings Λ_j^F are calculated through a principal component analysis - following [McCracken and Ng \(2015\)](#) and [Ma and Samaniego \(2019\)](#). This method determines the number of factors representing the firms' commonality in each sector based on the eigenvalue criterion. The values are then normalised so that $CFact_{i,t}$ follows a distribution as $\sim (\mu = 0, \sigma = 1)$. The values around 0 show that the firms at those points in time behave and experience close-to-average attributes in their sector. Values on the right and left tails - where there is a higher difference to zero - display conditions in which the firm is experiencing extreme circumstances with respect to the expected value in its sector.

In Figure 1.2, we show the construction of each sector's common factor and how the data on firms in these subgroups can be used to construct one single variable summarising the most crucial information of firms in that sector.

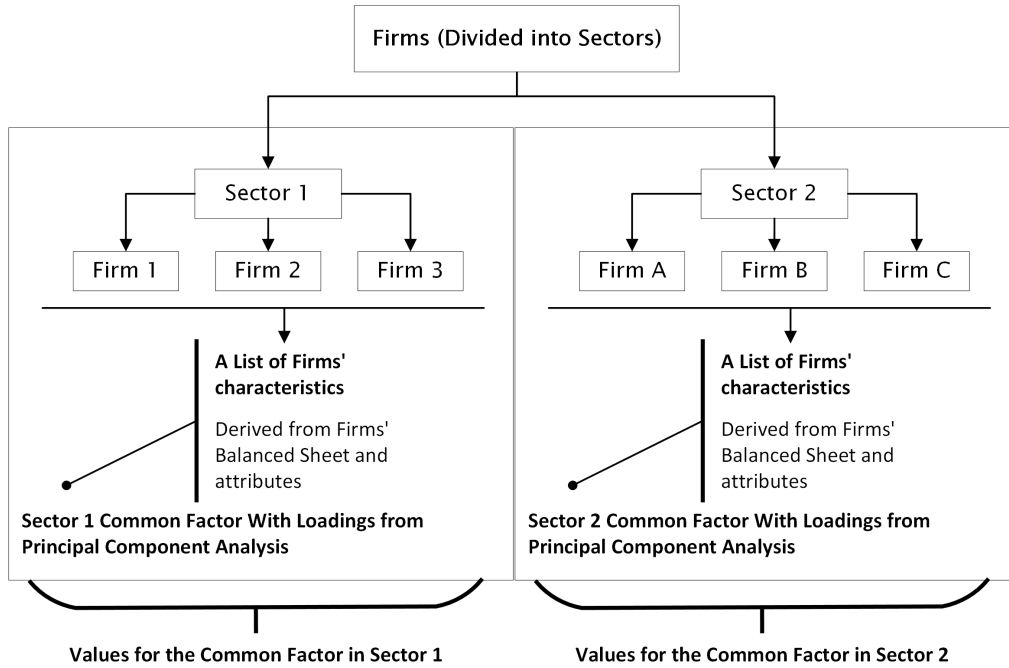


Figure 1.2: The commonality of firms' characteristics into one common factor variable for each sector.

A natural question at this point may arise regarding the choice of variables in constructing the sectoral common factor. Based on a simple Bayesian-like penalising algorithm, we find that a short list of variables containing the most information is favoured. These variables must include information on the most critical aspects of the businesses and what identifies them from other companies in other sectors. Our algorithm first constructs a common factor using all the time series for all the firms in the sector, namely $CFact^{Large}$. It then creates a list of common factors starting from an arbitrarily ordered list and calculates the correlation between each stage's resulting factor and $CFact^{Large}$. We then choose the list of variables that maximises the following fraction:

$$\text{Correlation Criterion} = \max_N \frac{\text{Corr} [CFact^N, CFact^{\text{Large}}]}{N}$$

In other words, we maximise the correlation with the large initial factor and penalise it in each step of adding variables by the total number of series that entered the factor construction. We find that the optimal vector of variables to be included is:

$$\mathbf{F}_{i,t} = \Omega \left(\text{Cash/Total Assets, Net Revenue, Long Term Investments, Total Liabilities, Age} \right)$$

After constructing the firm-specific and time-varying common factor for each sector, we add this variable to our regression on estimating sector-level uncertainty as:

$$\log(S_{i,t}) = \rho \log(S_{i,t-1}) + \lambda_t + \mu_i + CFact_{i,t} + \varepsilon_{i,t} \quad (1.9)$$

Again, $S_{i,t}$ denotes the firm i 's sales, μ_i is the firm fixed effect, and λ_t is the time fixed effect. We then define the sector-level uncertainty as the cross-sectional dispersion of the residuals of the regression above:

$$\text{SLU}_{i,t} = \frac{(\varepsilon_{i,t} - \bar{\varepsilon}_i)^2}{T - 1} \quad (1.10)$$

Our specification is robust to the choice of sector size (as defined in section 3.3). In Section 3.6, we examine our method by using the 1-digit sectors in the same analysis as above, and we show that our method is robust to a choice of larger sectors.

1.4.4 Aggregate-Level Uncertainty (ALU)

So far, we have identified the firm and sector-specific components of uncertainty in our analysis. The remainder of OU, after deducting firm and sector uncertainties, is the Aggregate-Level Uncertainty (ALU):

$$\text{ALU}_{i,t} = \text{OU}_{i,t} - \left(\frac{S_{i,t}}{\sum_{i=1 \in N_j}^{N_j} S_{i,t}} \right) \text{FLU}_{i,t} - \left(\frac{\sum_{i=1 \in N_j}^{N_j} S_{i,t}}{\sum_{i=1 \in N}^N S_{i,t}} \right) \text{SLU}_{i,t} \quad (1.11)$$

We define the weights for each tier as the relative sales of the firm over the sector and the sector over the whole economy tier to solve the granularity issue of uncertainty. Large firms might drive sector uncertainty, and large sectors' uncertainty might result in the uncertainty of the whole economy.

Ultimately, the aggregate-level uncertainty and sector-level uncertainty reflect different sources of uncertainty. Aggregate-level uncertainty reflects the overall risk of the economy, which is beyond the control of individual firms. These risks are known as systematic risks and can simultaneously affect multiple firms and sectors, such as changes in interest rates, government policies, or global economic conditions. Therefore, aggregate levels of uncertainty could remain high due to the persistent presence of these risks. On the other hand, the market structure of some industries can contribute to higher sectoral levels of uncertainty. In competitive markets, firms may have little control over prices, and uncertainty about future demand and supply conditions could persist. In contrast, firms with market power or monopoly positions may face lower uncertainty at the firm level since they have larger control over pricing and other strategic decisions. Therefore, sectors dominated by competitive markets may experience higher levels of uncertainty.

1.4.5 Estimation Method: Dynamic Panel using a GMM Estimator

Note that we are estimating a specific case of a dynamic panel of the form:

$$\log(S_{i,t}) = \sum_{j=1}^p \alpha_j \log(S_{i,t-j}) + \mathbf{X}_{i,t} \boldsymbol{\beta} + \mu_i + \lambda_t + \varepsilon_{i,t} \quad i = 1, \dots, N \quad t = 1, \dots, T \quad (1.12)$$

where in both Equations 1.6 and 1.9 the lagged component consists of one lag of sales, and the matrix of controls $\mathbf{X}$ consists of observations on either firm controls (in Equation 1.6) or sectoral common factor (in Equation 1.9). The components μ_i and λ_t capture the firm and time-fixed effects, which may be correlated with the covariates. In this form, the lagged dependent variables are correlated with the unobserved panel-level effects, making standard estimators inconsistent.¹¹

We use the methodology developed by [Arellano and Bover \(1995\)](#); [Blundell and Bond \(1998\)](#) to solve this issue and use a Generalised Method of Moments (GMM) estimator to avoid heteroscedasticity. Our methodology uses the first lag of the log of sales as an instrument for the level equation in addition to the moment conditions of lagged levels as instruments for the difference equation. The estimator in these models is robust to heteroscedasticity and provides us with residuals that are uncorrelated to the past-observed residuals. In this way, we can capture the truly unpredictable component of the sales process, independent of the uncertainty persistence from previous quarters.

1.4.6 Descriptive Statistics of Uncertainty

Table 1.4.2 presents the summary statistics for the four uncertainty measures. As shown above, these measures differ in their data requirements, leading to more observations for overall uncertainty (OU, 17,080 firms) and the lowest for aggregate uncertainty (ALU, 4,811 firms). ALU imposes the highest level of requirements as both the controls needed to construct FLU and SLU need to be available.

Negative values of ALU might result as the nature of ALU could imply the presence of confounding factors, namely factors that we either cannot observe or are out of the reach of our method. It can also be caused by a firm having a very high share of sales in the sector or operating in a sector very large compared to the economy. In this case, the weights in Equation 1.11 become too large and cause negative uncertainty. However, ALU is negative only in 7% of the sample (9,561 observations). In what follows, we replace the negative values with NA for our measure to remain as a conservative (lower bound) of aggregate uncertainty.

Our method generates uncertainty values derived from the residuals of regressions with estimations available in the Appendix B. Hence, we expect the absolute values of uncertainty measures to be very small. Therefore, we rescale these values by multiplying them by 100 to understand the descriptive statistics better. In Table 1.4.2, we report the descriptive statistics of uncertainty tiers multiplied by 100, for a sample between the 1st to the 99th percentiles to disregard outliers.

¹¹We perform the Wald-test proposed by [Wooldridge \(2010\)](#) and find significant heteroscedasticity of residuals in estimating OU, FLU and SLU with a GLS estimator.

Measure	Mean	Std.Dev	25 th	50 th	75 th	Min-Max	#Firms	#Obs
$OU \times 100$	0.220	0.506	0.009	0.046	0.183	0.000 - 4.770	16,589	578,886
$FLU \times 100$	0.145	0.269	0.008	0.039	0.149	0.000 - 2.075	5,584	162,683
$SLU \times 100$	0.191	0.511	0.005	0.030	0.130	0.000 - 6.163	5,220	146,721
$ALU \times 100$	0.258	0.572	0.016	0.068	0.228	0.000 - 5.516	4,772	122,761

Table 1.4.2: The Descriptive Statistics of Uncertainty (1st to 99th Percentiles) - Unbalanced Sample of Uncertainty Components.

We can see that the FLU lies on a smaller range and has a lower mean than SLU and ALU. Figure 12 in Appendix A reports the distribution of the four uncertainty components.

1.5 Validation

In this section, we validate our methodology by analysing whether our measures display expected behaviour over time and with respect to large economic events and the business cycle. We then demonstrate how closely our aggregate uncertainty measure co-moves with the existing measures of uncertainty in the literature.

1.5.1 Evolution of Uncertainty Over Time

We now focus on studying the evolution of different tiers of uncertainty over time. For each measure, we have obtained a value for each firm-quarter observation. For each firm for which we observe the three components (i.e. a balanced sample), we compute the relative share of FLU, SLU and ALU as part of total uncertainty (OU). In Figure 1.3 below, we report the cross-sectional average for each period and each measure.

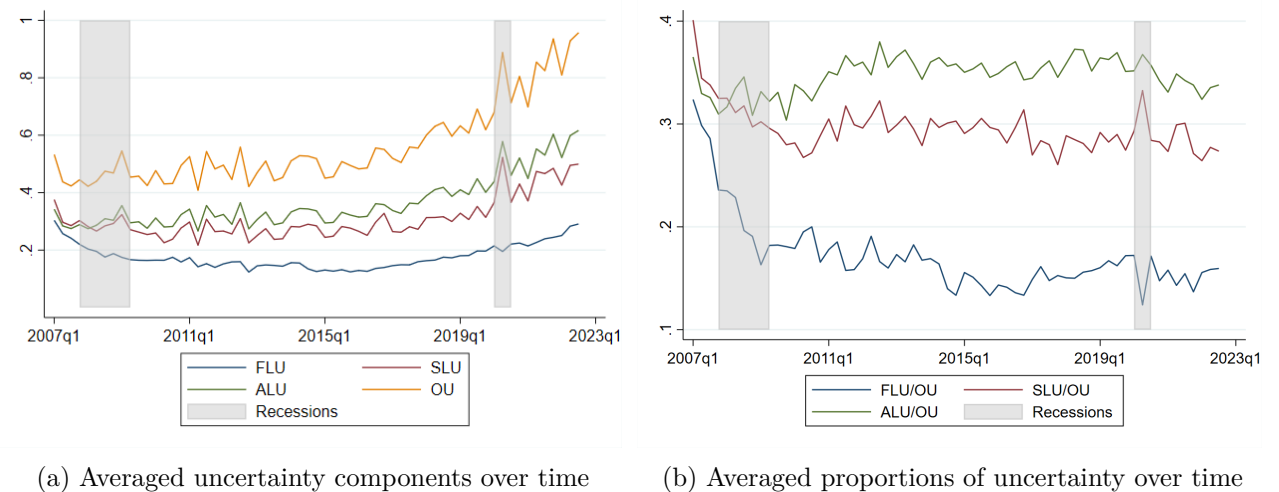


Figure 1.3: Evolution of uncertainty levels and proportions.

We can see that both the Great Recession and the COVID-19 outbreak were reflected primarily in uncertainty's sectoral and aggregate components. Furthermore, since early 2020, all uncertainty components have remained at a sharply increasing dynamic. Strikingly, uncertainty's aggregate and sectoral components have held a high proportion of uncertainty since our sample began. In contrast,

firm-level uncertainty has continued to occupy a progressively smaller share of overall uncertainty over time.

1.5.2 Uncertainty and the Business Cycle

Several studies have shown a significant increase in uncertainty measures during recessions (Bloom, Bond and van Reenen, 2007; Jurado, Ludvigson and Ng, 2015; Baker, Bloom and Davis, 2016; Bloom et al., 2022). However, whether uncertainty rises due to economic recessions or dampens growth remains open due to the challenge of reverse causality. In this section, we demonstrate that our uncertainty measures behave as expected with respect to episodes of booms and recessions. We begin by examining the correlation between uncertainty, aggregated sales, and firm-level sales growth. We then delve into how the variance of our uncertainty measures evolves over the business cycle.

Table 1.5.1 shows that our aggregate uncertainty measure is often counter-cyclical. Sector and firm-level uncertainties are strongly counter-cyclical in the whole sample and normal times, while in recessionary periods, they demonstrate both counter and pro-cyclical behaviours:

Aggregate-Level		FLU	SLU	ALU
	All Sales	-0.18***	-0.48***	-0.19***
	All Sales	0.56***	0.03***	-0.03***
	All Sales	-0.20***	-0.49***	-0.21***
Firm-Level	Sales Growth	0.00	-0.09***	-0.11***
	Sales Growth	0.00	-0.14***	-0.20***
	Sales Growth	0.01*	-0.08***	-0.10***

***: $p < 0.01$, **: $p < 0.05$ & *: $p < 0.10$

Table 1.5.1: Correlations of Uncertainty Components with the Sum of Sales (Aggregated) and Firm-Level Sales Growth in each Quarter, with Non-Recessionary vs Recessionary Periods.

The occasional pro-cyclicality of uncertainty is neither fully counter-intuitive nor unconventional. For example, Jurado, Ludvigson and Ng (2015)’s uncertainty measure is also pro-cyclical in non-recessionary times and counter-cyclical in recessionary times. Not only has the direction of the causality between business cycles and uncertainty remained an open question, but it is also challenging to assess business cycle features of either macro or financial uncertainty without the other. Furthermore, not all uncertainty is bad (Rossi and Sekhposyan, 2015; Forni, Gambetti and Sala, 2024), as our uncertainty measures also capture right-tail events. Right-tail events also contribute to uncertainty measures, which are not necessarily “bad” economic events. In our case, we combine right- and left-tale sales change observations as unpredictable (at least according to our econometric specification) events. For instance, a change in the CEO of a company might bring about creative marketing solutions, leading to a short-term increase in sales. In this case, there is a simultaneous increase in both company’s profitability and ex-post uncertainty.

Similar to Bloom et al. (2018), we argue that uncertain periods are not only periods of higher uncertainty as reflected in the first moment of the uncertainty measure but also represent higher second moments of uncertainty. Ideally, we would observe a higher variance in the recessionary periods. Not only do sales become uncertain, but the process of uncertainty becomes uncertain, translating into more sizeable heterogeneity. Initially, we find a small positive but significant correlation at the 99% confidence interval between the first and second moments of uncertainty at each tier and among each other. Forecasting sales for the next period in turbulent periods and understanding how tumultuous the economy is becomes difficult. This correlation slightly increases to 5% in NBER recessionary periods. In other words, uncertainty and its extent become less unpredictable when the economy performs poorly.

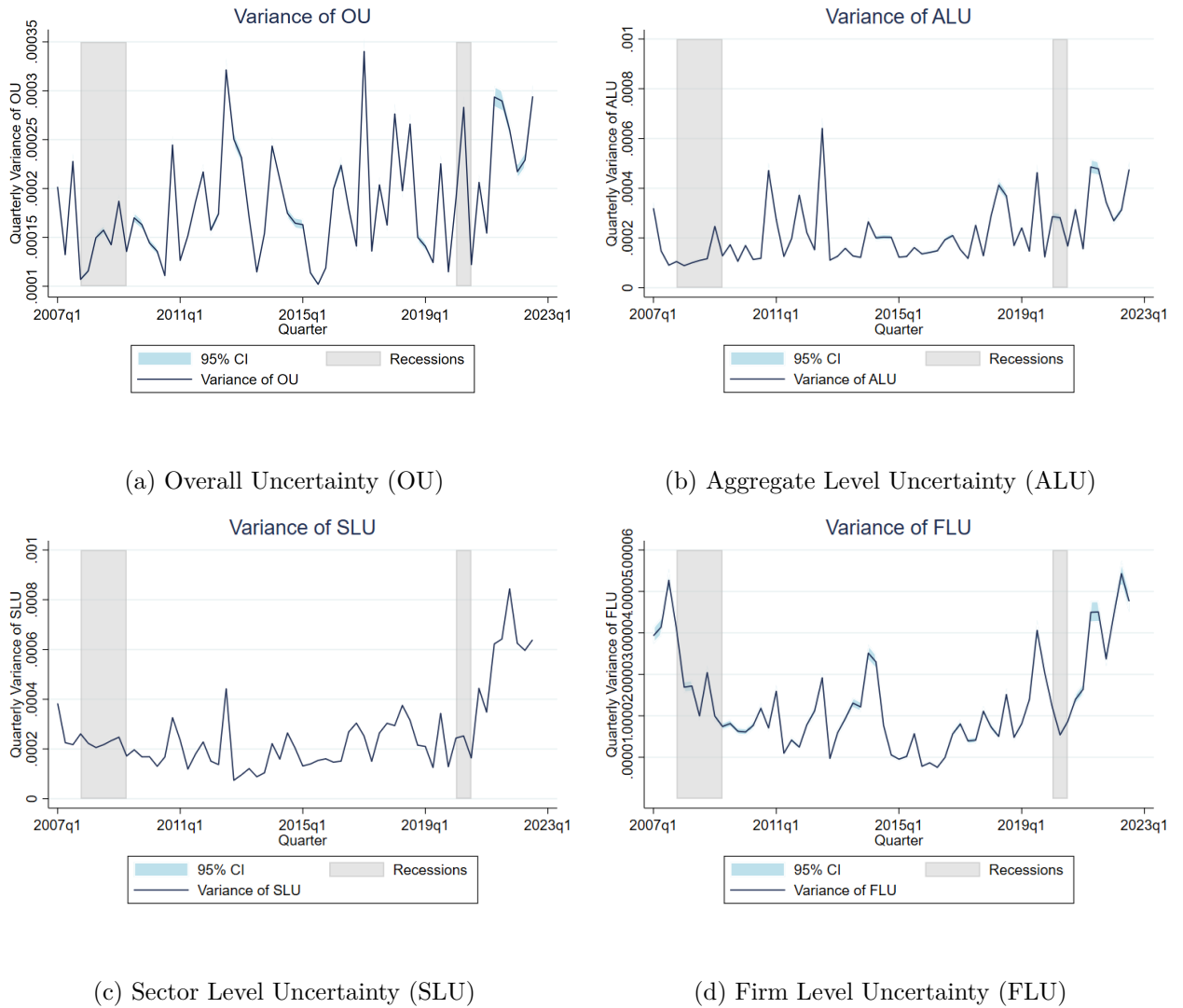


Figure 1.4: Variance of the uncertainty measures over time

Figure 1.4 above reports more systematically the variance evolution in our uncertainty measures. It displays a visually persistent lagged relationship between the first and second moments of uncertainty. After the recessionary period of 2008-09, the uncertainty about uncertainty increased consistently until around 2013. The COVID outbreak reveals instead a low variance of uncertainty, as almost every firm was hit severely. The quarters after the outbreak, on the other hand, have shown a higher uncertainty about the extent of uncertainty for firms - which could be explained by the politically

and economically troubled period after the COVID outbreak, hinting at how heterogeneously firms experienced uncertainty.

1.5.3 Comparison with existing measures

To validate whether our aggregate level uncertainty (ALU) index corresponds to other measures of economic uncertainty, we compare it with the index from [Jurado, Ludvigson and Ng \(2015\)](#) (JLN) and the Economic Policy Uncertainty measure (EPU) from [Baker, Bloom and Davis \(2016\)](#).¹² The rationale behind our focus on validating solely ALU is the lack of a corresponding measure for firm and sector-level uncertainty available for comparison.

This comparison exercise is reported in Figure 1.5 and Table 1.5.2. We first observe that similar to the JLN and EPU indices, our measure captures the sharp recession of 2007-08 and does well in pointing out the sharp peak of the uncertainty in 2020. While JLN does not capture the trade war of 2018, ALU displays significantly higher uncertainty during this period.

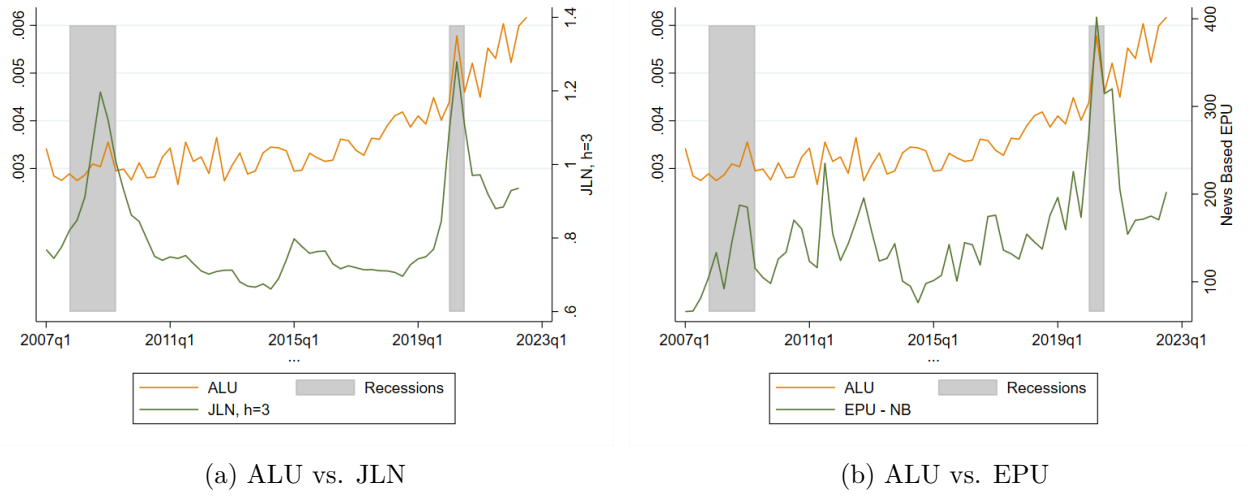


Figure 1.5: The evolution of ALU over time in comparison with existing measures of uncertainty.

Our uncertainty measure is strongly correlated to the existing indices as reported in Table 1.5.2. One may also raise whether such strong correlations are driven by recessionary or non-recessionary times. We differentiate those periods defined by NBER recession bands, divide our sample into non-recessionary and recessionary times, and compute the correlation values again. The second and third panels of Table 1.5.2 show that our measure is more strongly correlated to the commonly used uncertainty indices in the literature during the recessionary period. Interestingly, the correlation between JLN and EPU also increases during the recessionary periods and reduces vastly during the moderate years. The weaker correlation in non-recessionary periods is often justified by the nature of uncertainty that each measure is capturing - while in recessionary periods, all measures agree on a rise in uncertainty.

¹²In particular we focus on the most common versions of these indices, namely the JLN measure with 3 lags and the news-based EPU measure.

		ALU	JLN
Overall	JLN	0.41***	
sample	EPU	0.58***	0.55***
Non-recessionary	JLN	0.58***	
periods	EPU	0.51***	0.29***
Recessionary	JLN	0.72***	
periods	EPU	0.97***	0.77***

***: $p < 0.01$, **: $p < 0.05$ & *: $p < 0.10$

Table 1.5.2: Correlation Table between ALU, JLN from [Jurado, Ludvigson and Ng \(2015\)](#) & EPU from [Baker, Bloom and Davis \(2016\)](#).

1.6 Heterogeneity of uncertainty across firms and sectors

In this section, we explore the potential heterogeneities of our uncertainty measures to study the sources of differences in uncertainty across firms. Take the example of a chip-producing establishment that was immensely hit by the COVID-19 outbreak, which is reflected in aggregate uncertainty. This firm’s uncertainty - both in aggregate and sectoral components - differs significantly from that of a bike saddle producer. Even though both were severely hit by a shock, their “perception” and “experience” of uncertainty were quite different. Potentially, this heterogeneity is rooted in firms’ natural exposure to external shocks. We start this section by studying differences in the weight of the components in firms’ overall uncertainty. Then, we study the role of two critical firm characteristics, size and age, in explaining firm differences in perceiving uncertainty. Finally, we examine the presence and the extent of sectoral heterogeneity.

1.6.1 Heterogeneity in types of uncertainty across firms

So far, we have observed the evolution of uncertainty over time and some of the coral properties of aggregate uncertainty. We also discussed that the macroeconomic uncertainty lies quantitatively above firm-level uncertainty in our sample time frame. Alternatively, and to confirm the idea that most of the firms’ uncertainty roots in the less predictable component of uncertainty - e.g. macro conditions - we rank the firms based on their proportion of ALU as a share of OU (the green line in Figure 1.6). The figure also shows the ratios for the sum of aggregate and sectoral uncertainty and the sum of all three tiers as a proportion of OU, which is the horizontal line equal to 1 by definition. The distance between the blue and red lines in the figure represents the proportion of FLU out of OU. We see that most of the uncertainty firms face stems from macro sources, namely the less predictable components of volatility to firms. Furthermore, the nature of our sample - containing primarily larger firms in the economy - implies that this group must have done well in resolving their internal issues.

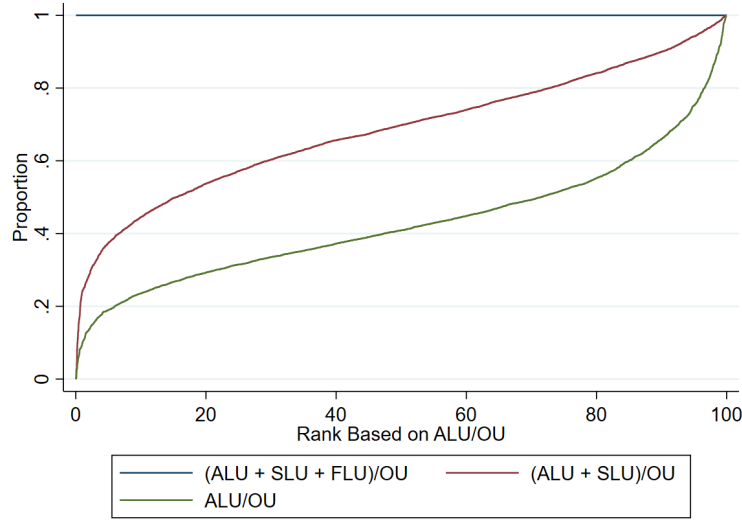


Figure 1.6: The ranked distribution of firms' proportion of macroeconomic uncertainty from their OU

We can see that from around 30th percentile based on ALU/OU, around 60% of the uncertainty stems from macro (aggregate and sectoral) sources. As the smoothness in Figure 1.6 suggests, firms are strongly heterogeneous concerning the proportionality and the nature of the uncertainty they face.

1.6.2 The role of size and age in explaining uncertainty heterogeneity

Firms face different levels of uncertainty based on their size (Ballantine, Cleveland and Koeller, 1993) and their age and experience (Drnevich and West, 2021). Business cycles affect young and small firms by a large extent (Clymo and Rozsypal, 2022). We regroup firms in the groups of small (young) and large (old) firms based on the percentile of total assets (age) in each quarter (see the details in Section 3.3). This approach allows us to capture firm dynamics better as firms grow, shrink, enter, or leave the markets. Such dynamics affect uncertainty in all of its components.

1.6.3 Size and uncertainty heterogeneity

To study the size heterogeneity, we average the overall, firm, sectoral and aggregate uncertainties within each firm size group at each quarter in our sample. Figure 1.7 shows that the average overall uncertainty is much lower for large firms with respect to their smaller counterparts across different business cycle phases.

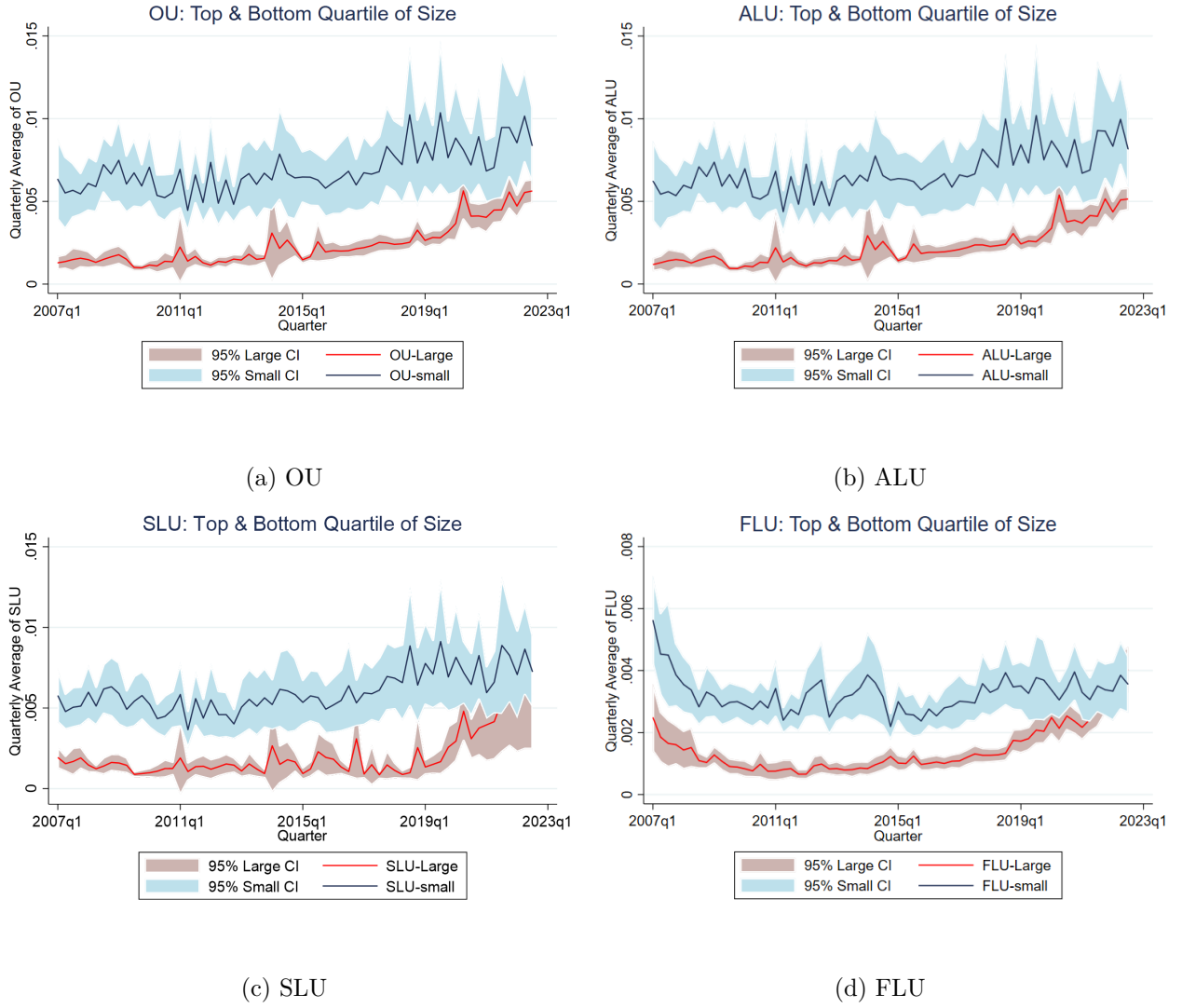


Figure 1.7: Uncertainty measures for different size groups based on total assets: large(top quartile) and small (bottom quartile) firms

The results presented above align with [Dunne, Roberts and Samuelson \(1989\)](#): firms' size is a central determinant of growth at the firm level. With small firms facing more uncertainty, there is less chance for hiring, investments and growth. The figure also shows that this result also holds for the components of uncertainty, with a slight difference that there seems to be a catch-up pattern among firms in their SLU and FLU after the COVID outbreak.

So far, we have discussed how the level of uncertainty in all components is higher among smaller firms. Our follow-up analysis concerns the decomposition of uncertainty in each group. Put simply, we study the proportions of each component in each group at each point in time. The results of these proportions are provided in Figure 13 in Appendix A. While for both groups of small and large firms, the aggregate uncertainty constitutes a large proportion of uncertainty, the extent is larger among larger firms. This result is intuitive since larger firms seem to have resolved their company's internal conditions and potentially higher market concentration in the sector, which lowers their degrees of firm and sectoral uncertainty.

1.6.4 Age and uncertainty heterogeneity

We continue by studying the role of a firm's age in determining its uncertainty. As displayed in Figure 1.8, we find no significant evidence of difference among these groups in their volatility (Overall Uncertainty) or its components. However, older firms seem to be on the rise in their firm-specific uncertainty component, while on average, the sectoral uncertainty was proportionally larger among younger firms.

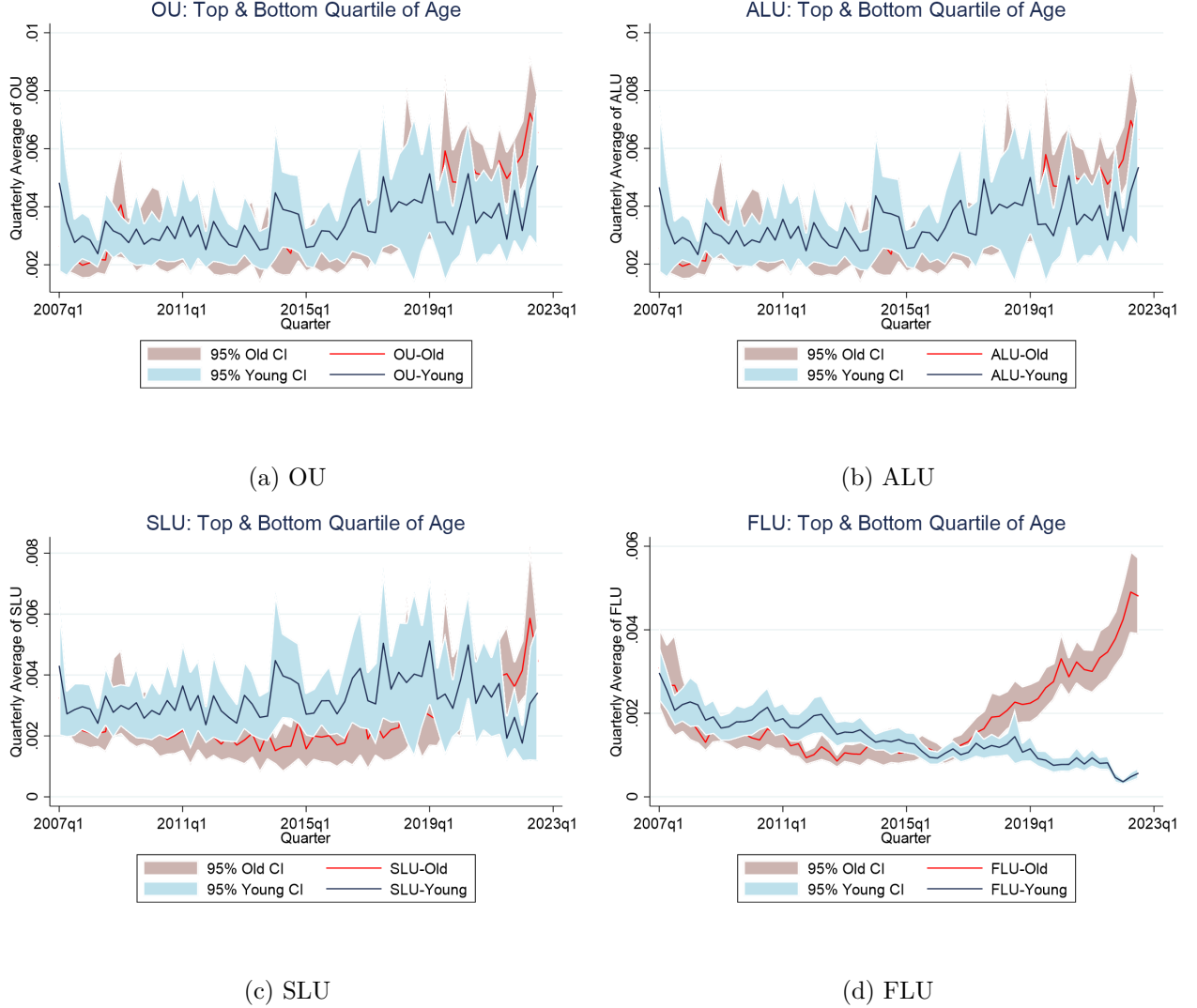


Figure 1.8: Old and young firms' uncertainty components over time.

Some notes of caution: in our sample, young means “recently listed” and not recently established. One may expect recently established firms to have more uncertainty at each point. We do not have access to the establishment date of the entities in our sample, and our attention is limited to the concept of IPO's impact on firms. Once a firm is successfully IPOed, it receives large liquidity, a better reputation, more sales, and larger investments. These dynamics could cause a short-term moderating impact on uncertainty as they provide a buffer against idiosyncratic, aggregate, and sectoral shocks.¹³

¹³To find out if younger listed firms do not necessarily have lower uncertainty, we follow an alternative approach to classify our [listed] firms as old and young. We use the benchmark of older than 15 and younger than 5, as also similarly used by Ferrando et al. (2020). Any firm in between is classified as mature. 29% of the firms will be tagged as young, 27% will be known as old and 44% of the firms will be mature. The new results are almost identical to our analysis above, indicating no significant difference in levels of uncertainty among age groups.

Next, we ask whether the proportions of types of uncertainty among young and old firms differ. We find substantially different patterns based on age (See the upper half of Figure 13 in Appendix A). Until 2020, all three uncertainty components remained roughly in the region of 0.3-0.4 for young firms. Since the COVID outbreak, however, the aggregate uncertainty proportion has increased sharply. Furthermore, we observe a substantial decrease in the firm-level component. Among larger firms, however, there has been a persistent increase in the aggregate component until 2020. Amid the supply-chain-related shocks of COVID-19 in 2020, the large firms' sectoral uncertainty (as a proportion) has been on a negative-trend path over time.

Lastly, our analysis raises the question of whether the size of the recently-listed firm matters. We analyse our framework by dividing our sample into subsamples that interact with age and size. Size and age have substantially different implications for firm growth and cannot fully replace one another [Clymo and Rozsypal \(2022\)](#). Therefore, we can also question whether large firms are more likely to be old. Accumulating assets and employees takes time. Nevertheless, with the convention of dividing firms into quartiles, we find that our sample includes many large but young companies (3.2%, as opposed to 7.6% of large-old firms). A large and young firm might be exposed to higher volatility as it faces high financial uncertainty, for example, due to its unknown performance after the IPO date ([Brown and Wiles, 2015](#)) and lack of information on past performances. However, our findings reject such a link in almost all uncertainty components.



Figure 1.9: Tiers of uncertainty for the interaction of age and size

The firms' size drives the heterogeneity across most of our sample in different uncertainty components. The striking negative trend in the aggregate uncertainty for younger firms might have policy implications: the newer firms seem to be richer in information - through their data accumulation. This sort of entry with high information can reduce the uncertainty for a firm, for interchangeably important demand and production sources of uncertainty (Eeckhout and Veldkamp, 2022).

1.6.5 Heterogeneity of uncertainty across sectors

This section studies uncertainty characteristics across large sectors (manufacturing, mining and services). We average our tiers of uncertainty for each of these sectors and find that they evolve over time, as shown in Figure 1.10:

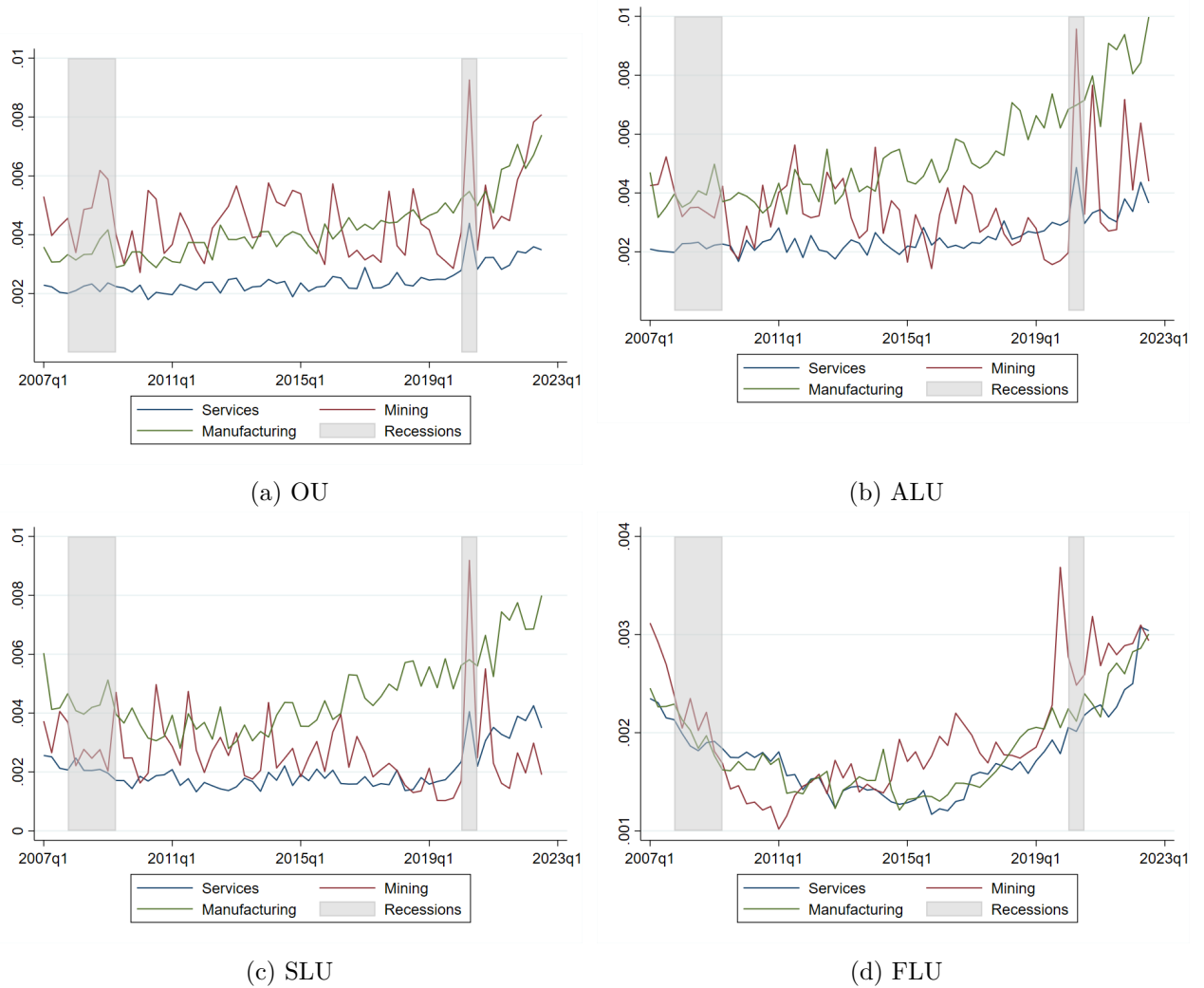


Figure 1.10: Averaged aggregate and sector level uncertainties across large sectors

The moderate levels of uncertainty in the service sector, even in the recessions of 2008 and 2020, might be captured by the idea that these activities might be less dependent on the global supply chains. The sharp peak of mining uncertainty during the 2020 supply chain crisis also stands out, while the positive trend of uncertainty in the manufacturing sector reflects the difficulties this sector has faced in conducting business in recent decades. Generally, our results confirm an aspect of the “decline in the U.S. manufacturing sector” as pointed out in detail by Charles, Hurst and Schwartz (2019): the manufacturing sector has experienced a drop in the hours worked in the last decades, as well as a

sharp increase in capital intensity and stagnated output levels. Importantly, our uncertainty measure captures the trade war between China and the U.S. (IMF, 2016; Amiti, Redding and Weinstein, 2019), also reflected in the uncertainty index of Ahir, Bloom and Furceri (2022).

Ultimately, we suspect the rise in manufacturing uncertainty can be related to supply chain disruptions. Intuition from Parast and Subramanian (2021) suggests that input supply disruptions can significantly affect firm performance and increase uncertainty in manufacturing firms through production bottlenecks. We obtain the series of the U.S. ISM manufacturing Supplier Deliveries Index from YCharts, and find a 68% correlation between the supplier deliveries index and uncertainty. This correlation drops to 33% for the SLU in services. Figure 1.11 below demonstrates how closely manufacturing uncertainty and supply delivery index correlate and how sector-level uncertainty in the services sector is affected less by supply-chain-related developments.

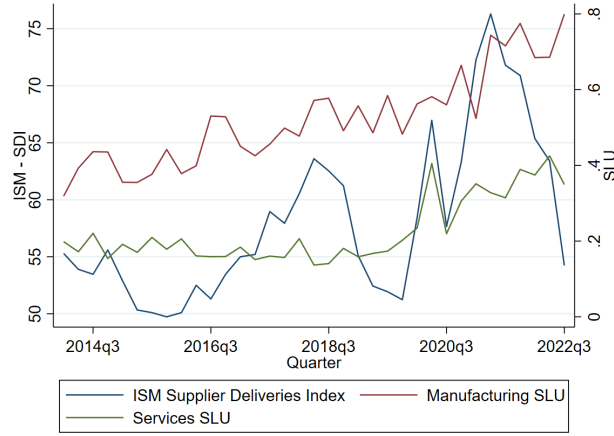


Figure 1.11: Manufacturing and services uncertainty and the Supply Delivery Index (SDI).

1.7 Robustness Checks

Our novel strategy for decomposing uncertainty into different components should be robust to alternative specifications. An uncertainty decomposition exercise should perform well in identifying the idiosyncratic component - in our case, the firm-level uncertainty, as well as a fairly uncorrelated pair of macroeconomic components, namely sectoral and aggregate levels. Our measure can identify a firm-level uncertainty measure that is correlated with an extent of only 20% with respect to aggregate and sectoral levels. In this section, we check the robustness of our strategy through multiple specifications that have the potential to perform better than our model. Our methodology is robust to these specifications and can decompose uncertainty into desired tiers appropriately.

To fully ensure that our powerhouse effectively disentangles the different components of overall uncertainty, we add the firm-level uncertainty measured in the first step of our analysis as a control in estimating the sector-level uncertainty in the second step. The regression results are in Table B.2 in Appendix B. Our baseline measure is robust to this specification: the resulted SLU and ALU are 99% correlated to their counterparts from the baseline specification.

One may argue that our approach divides the sample into too small subsamples to create a common factor among firms in each industry. To demonstrate that our specification is robust to such claims, we divide the sample into 1-digit NAICS sectors, which divides our sample into nine groups. We then include the new common factor in our estimation of SLU and show that the resulting SLU is almost

98% correlated to the baseline SLU, and the ALU is at the 90% correlation level with the baseline ALU.

Ultimately, we suspect whether the seasonality issue might arise in our estimations. To test our methodology against this argument, we perform an identical analysis with respect to our baseline estimations but with four lags in sales instead of one. We find a range of $[0.75-0.86]$ pairwise correlation between the new values for each tier and the baseline specification values. However, this new specification correlates less strongly with other economic uncertainty indices.

1.8 Conclusion

We introduce a novel uncertainty measure computed using firms' balance sheet data that generates three different tiers of uncertainty from firms' overall uncertainty. We show that our uncertainty index driven from micro-data correlates with those driven from aggregated data like macro variable dispersion or media discussion on uncertainty. Our measure of aggregate uncertainty also peaks during recessions and economic and political turbulence. We also show the presence of substantial heterogeneity among firms and sectors: large firms experience much lower uncertainty than small firms. Higher uncertainty for smaller firms means lower growth of these firms over time as they might become more and more cautious and delay their investments and hiring decisions. This is a new channel through which uncertainty potentially dampens sectoral and aggregate growth. Furthermore, we show that manufacturing has exhibited higher levels of uncertainty across our sample, except for the COVID shock, primarily represented in the mining sector. This result leaves an interesting puzzle for future research to disentangle uncertainty formally across sectors and business cycles.

These results hold significant lessons for policymakers and the academic literature on uncertainty and firm performance. Policymakers should know smaller and younger firms' vulnerability to macroeconomic conditions, established even within a predominantly large firm sample in our paper. This underscores the impact of fiscal and monetary policies on firms' uncertainty, performance, investment decisions, and hiring choices. Moreover, our findings align with the empirical findings on the stagnation of the manufacturing sector in the U.S. economy over recent decades. Lastly, our method and results contribute to the academic literature by introducing the concept of uncertainty heterogeneity, which sheds light on the varying impacts of uncertainty on firms and sectors.

All in all, our approach fills the gap regarding uncertainty heterogeneity in the literature. It provides evidence for the former and sheds light on the similarity between identified aggregate uncertainty from microdata and uncertainty measured using aggregated data. Lastly, future research can apply our measure to multiple aspects of firms' dynamics, such as growth, investments and management quality. This measure can be aggregated at different levels and be used as an uncertainty measure without endogeneity issues, while the heterogeneity established in this paper can motivate studies related to the heterogeneous effects of uncertainty on firms and their outcomes.

References

- Abel, Andrew, and Janice Eberly.** 1994. "A Unified Model of Investment under Uncertainty." American Economic Review, 84(5): 1369–84.
- Abel, Andrew B., Avinash K. Dixit, Janice C. Eberly, and Robert S. Pindyck.** 1996. "Options, the Value of Capital, and Investment*." The Quarterly Journal of Economics, 111(3): 753–777.
- Ahir, Hites, Nicholas Bloom, and Davide Furceri.** 2022. "The World Uncertainty Index." National Bureau of Economic Research Working Paper 29763.
- Aivazian, Varouj A., Mohammad M. Rahaman, and Simiao Zhou.** 2019. "Does corporate diversification provide insurance against economic disruptions?" Journal of Business Research, 100: 218–233.
- Alfaro, Iván, Nick Bloom, and Xiaoji Lin.** 2023. "The Finance Uncertainty Multiplier." Journal of Political Economy.
- Altig, David, Jose Maria Barrero, Nicholas Bloom, Steven J Davis, Brent H Meyer, and Nicholas Parker.** 2019. "Surveying Business Uncertainty." "National Bureau of Economic Research" Working Paper 25956.
- Amiti, Mary, Stephen J. Redding, and David E. Weinstein.** 2019. "The Impact of the 2018 Tariffs on Prices and Welfare." Journal of Economic Perspectives, 33(4): 187–210.
- Arellano, Manuel, and Olympia Bover.** 1995. "Another Look at the Instrumental Variable Estimation of Error-Components Models." Journal of econometrics, 68(1): 29–51.
- Awano, Gaganan, Nicholas Bloom, Ted Dolby, Paul Mizen, Rebecca Riley, Tatsuro Senga, John Van Reenen, Jenny Vyas, and Philip Wales.** 2018. "A firm-level perspective on micro- and macro-level uncertainty; An analysis of business expectations and uncertainty from the UK Management and Expectations Survey." Economic Statistics Centre of Excellence (ESCoE).
- Aydogan, Ilke, Loïc Berger, Valentina Bosetti, and Ning Liu.** 2023. "Three Layers of Uncertainty." Journal of the European Economic Association, 21(5): 2209–2236.
- Azzimonti, Marina.** 2018. "Partisan Conflict and Private Investment." Journal of Monetary Economics, 93: 114–131.
- Bachmann, Rüdiger, Kai Carstensen, Stefan Lautenbacher, and Martin Schneider.** 2021. "Uncertainty and change: survey evidence of firms' subjective beliefs." National Bureau of Economic Research.
- Bachmann, Rüdiger, Steffen Elstner, and Eric R. Sims.** 2013. "Uncertainty and Economic Activity: Evidence from Business Survey Data." American Economic Journal: Macroeconomics, 5(2): 217–49.
- Baker, Scott R., Nicholas Bloom, and Steven J. Davis.** 2016. "Measuring Economic Policy Uncertainty*." The Quarterly Journal of Economics, 131(4): 1593–1636.

- Ballantine, John W, Frederick W Cleveland, and C Timothy Koeller.** 1993. “Profitability, Uncertainty, and Firm Size.” Small Business Economics, 5: 87–100.
- Barigozzi, Matteo, Christian Brownlees, Giampiero M. Gallo, and David Veredas.** 2014. “Disentangling Systematic and Idiosyncratic Dynamics in Panels of Volatility Measures.” Journal of Econometrics, 182(2): 364–384.
- Baum, Christopher F., Mustafa Caglayan, Neslihan Ozkan, and Oleksandr Talavera.** 2006. “The Impact of Macroeconomic Uncertainty on Non-Financial Firms’ Demand for Liquidity.” Review of Financial Economics, 15(4): 289–304.
- Bianchi, Francesco, Howard Kung, and Mikhail Tirsikh.** 2023. “The origins and effects of macroeconomic uncertainty.” Quantitative Economics, 14(3): 855–896.
- Bloom, Nicholas.** 2009. “The Impact of Uncertainty Shocks.” Econometrica, 77(3): 623–685.
- Bloom, Nicholas.** 2014. “Fluctuations in Uncertainty.” Journal of Economic Perspectives, 28(2): 153–76.
- Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry.** 2018. “Really Uncertain Business Cycles.” Econometrica, 86(3): 1031–1065.
- Bloom, Nicholas, Stephen Bond, and John van Reenen.** 2007. “Uncertainty and Investment Dynamics.” The Review of Economic Studies, 74(2): 391–415.
- Bloom, Nicholas, Steven J Davis, Lucia S Foster, Scott W Ohlmacher, and Itay Saporta-Eksten.** 2022. “Investment and Subjective Uncertainty.” National Bureau of Economic Research Working Paper 30654.
- Blundell, Richard, and Stephen Bond.** 1998. “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models.” Journal of econometrics, 87(1): 115–143.
- Born, Benjamin, and Johannes Pfeifer.** 2021. “Uncertainty-driven business cycles: Assessing the markup channel.” Quantitative economics, 12(2): 587–623.
- Brown, Keith C, and Kenneth W Wiles.** 2015. “In Search of Unicorns: Private IPOs and the Changing Markets for Private Equity Investments and Corporate Control.” Journal of applied corporate finance, 27(3): 34–48.
- Cagetti, Marco, Lars Peter Hansen, Thomas Sargent, and Noah Williams.** 2015. “Robustness and Pricing with Uncertain Growth.” The Review of Financial Studies, 15(2): 363–404.
- Cascaldi-Garcia, Danilo, Cislil Sarisoy, Juan M Londono, Bo Sun, Deepa D Datta, Thiago Ferreira, Olesya Grishchenko, Mohammad R Jahan-Parvar, Francesca Loria, Sai Ma, et al.** 2023. “What is certain about uncertainty?” Journal of Economic Literature, 61(2): 624–654.
- Charles, Kerwin Kofi, Erik Hurst, and Mariel Schwartz.** 2019. “The Transformation of Manufacturing and the Decline in US Employment.” NBER Macroeconomics Annual, 33(1): 307–372.
- Clymo, Alex, and Filip Rozsypal.** 2022. “Firm Cyclicalities and Financial Frictions.”

- Comin, Diego, and Sunil Mulani.** 2006. “Diverging trends in aggregate and firm volatility.” The Review of Economics and Statistics, 88(2): 374–383.
- Comin, Diego, and Thomas Philippon.** 2005. “The Rise in Firm-Level Volatility: Causes and Consequences.” NBER Macroeconomics Annual, 20.
- D’Amico, Stefania, and Athanasios Orphanides.** 2008. “Uncertainty and Disagreement in Economic Forecasting.” Board of Governors of the Federal Reserve System (U.S.) Finance and Economics Discussion Series 2008-56.
- Davis, Steven J, John Haltiwanger, Ron Jarmin, Javier Miranda, Christopher Foote, and Eva Nagypal.** 2006. “Volatility and dispersion in business growth rates: Publicly traded versus privately held firms [with comments and discussion].” NBER macroeconomics annual, 21: 107–179.
- De Veirman, Emmanuel, and Andrew Levin.** 2018. “Cyclical Changes in Firm Volatility.” Journal of Money, Credit and Banking, 50(2-3): 317–349.
- Dew-Becker, Ian, and Stefano Giglio.** 2023. “Cross-sectional uncertainty and the business cycle: evidence from 40 years of options data.” American Economic Journal: Macroeconomics, 15(2): 65–96.
- Dixit, R.K., A.K. Dixit, and R.S. Pindyck.** 1994. Investment Under Uncertainty. Princeton University Press.
- Drnevich, Paul Louis, and Joel West.** 2021. “Performance Implications of Technological Uncertainty, Age, and Size for Small Businesses.” Journal of Small Business Management, 1–36.
- Dunne, Timothy, Mark J. Roberts, and Larry Samuelson.** 1989. “The Growth and Failure of U. S. Manufacturing Plants*.” The Quarterly Journal of Economics, 104(4): 671–698.
- Eeckhout, Jan, and Laura Veldkamp.** 2022. “Data and Market Power.” National Bureau of Economic Research.
- Ferrando, Annalisa, Klaas Mulier, Laurens Cherchye, Bram De Rock, and Marijn Verschelde.** 2020. “Identifying Financial Constraints.”
- Forni, Mario, Luca Gambetti, and Luca Sala.** 2024. “Downside and Upside Uncertainty Shocks.” Journal of the European Economic Association, jvae024.
- Gennaioli, Nicola, Yueran Ma, and Andrei Shleifer.** 2016. “Expectations and investment.” NBER Macroeconomics Annual, 30(1): 379–431.
- Gulen, Huseyin, and Mihai Ion.** 2016. “Policy Uncertainty and Corporate Investment.” The Review of Financial Studies, 29(3): 523–564.
- Hansen, Lars Peter, and Thomas J Sargent.** 2021. “Macroeconomic Uncertainty Prices When Beliefs Are Tenuous.” Journal of Econometrics, 223(1): 222–250.
- Hassler, John A. A.** 1996. “Variations in Risk and Fluctuations in Demand: A Theoretical Model.” Journal of Economic Dynamics and Control, 20(6-7): 1115–1143.

- Herskovic, Bernard, Bryan Kelly, Hanno Lustig, and Stijn Van Nieuwerburgh.** 2016. “The Common Factor in Idiosyncratic Volatility: Quantitative Asset Pricing Implications.” Journal of Financial Economics, 119(2): 249–283.
- Herskovic, Bernard, Bryan Kelly, Hanno Lustig, and Stijn Van Nieuwerburgh.** 2020. “Firm Volatility in Granular Networks.” Journal of Political Economy, 128(11): 4097–4162.
- IMF, (International Monetary Fund).** 2016. “IMF World Economic Outlook: Too Slow for Too Long.” Washington, DC: International Monetary Fund.
- Jurado, Kyle, Sydney C. Ludvigson, and Serena Ng.** 2015. “Measuring Uncertainty.” American Economic Review, 105(3): 1177–1216.
- Kozeniauskas, Nicholas, Anna Orlik, and Laura Veldkamp.** 2018. “What Are Uncertainty Shocks?” Journal of Monetary Economics, 100: 1–15.
- Kumar, Saten, Yuriy Gorodnichenko, and Olivier Coibion.** 2023. “The effect of macroeconomic uncertainty on firm decisions.” Econometrica, 91(4): 1297–1332.
- Lamorgese, Andrea, Megha Patnaik, Andrea Linarello, and Fabiano Schivardi.** 2024. “Management practices and resilience to shocks: Evidence from COVID-19.” Management Science.
- Liang, J Nellie, and Stephen A Rhoades.** 1991. “Asset diversification, firm risk, and risk-based capital requirements in banking.” Review of Industrial Organization, 6: 49–59.
- Ma, Xiaohan, and Roberto Samaniego.** 2019. “Deconstructing Uncertainty.” European Economic Review, 119: 22–41.
- McCracken, Michael W., and Serena Ng.** 2015. “FRED-MD: A Monthly Database for Macroeconomic Research.” Journal of Business & Economic Statistics, 34: 574 – 589.
- Ozili, Peterson K, and Thankom Arun.** 2023. “Spillover of COVID-19: Impact on the Global Economy.” In Managing Inflation and Supply Chain Disruptions in the Global Economy. 41–61. IGI Global.
- Panousi, Vasia, and Dimitris Papanikolaou.** 2012. “Investment, Idiosyncratic Risk, and Ownership.” The Journal of Finance, 67(3): 1113–1148.
- Parast, Mahour Mellat, and Nachiappan Subramanian.** 2021. “An Examination of the Effect of Supply Chain Disruption Risk Drivers on Organizational Performance: Evidence From Chinese Supply Chains.” Supply Chain Management: An International Journal, 26(4): 548–562.
- Rossi, Barbara, and Tatevik Sekhposyan.** 2015. “Macroeconomic Uncertainty Indices Based on Nowcast and Forecast Error Distributions.” American Economic Review, 105(5): 650–55.
- Sheen, Jeffrey, and Ben Zhe Wang.** 2017. “Estimating Macroeconomic Uncertainty from Surveys- A Mixed Frequency Approach.” Available at SSRN 3020697.
- Sill, KeiTh, et al.** 2012. “Measuring economic uncertainty using the survey of professional forecasters.” Federal Reserve Bank of Philadelphia Business Review, 92: 16–27.
- Stock, James H, and Mark Watson.** 2011. “Dynamic Factor Models.” Oxford Handbooks Online.

Thomas, P.J. 2018. "Pitfalls in survey measurements of economic parameters." Vol. 1065, 072009, IOP Publishing.

Uzeda, Luis, Kerem Tuzcuoglu, and Efrem Castelnuovo. 2022. "Sectoral Uncertainty." SSRN Electronic Journal.

Wooldridge, Jeffrey M. 2010. Econometric Analysis of Cross Section and Panel Data. MIT press.

A Additional Graphs

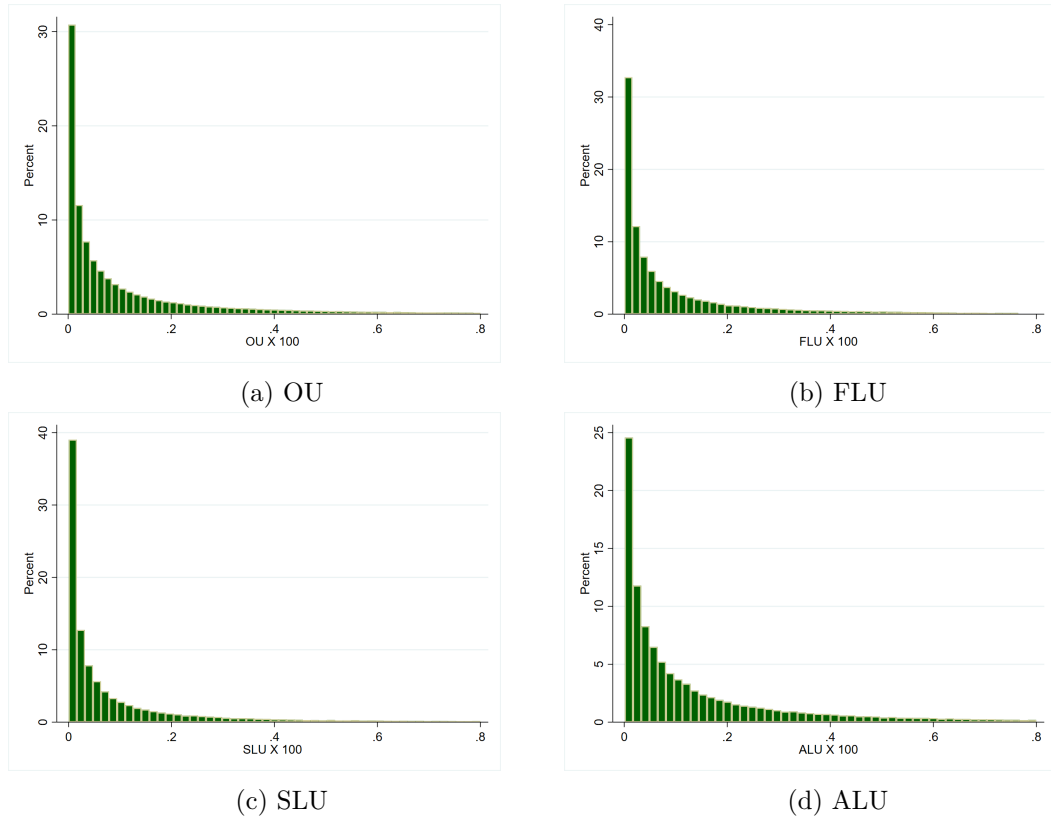


Figure 12: The distribution of uncertainty components.

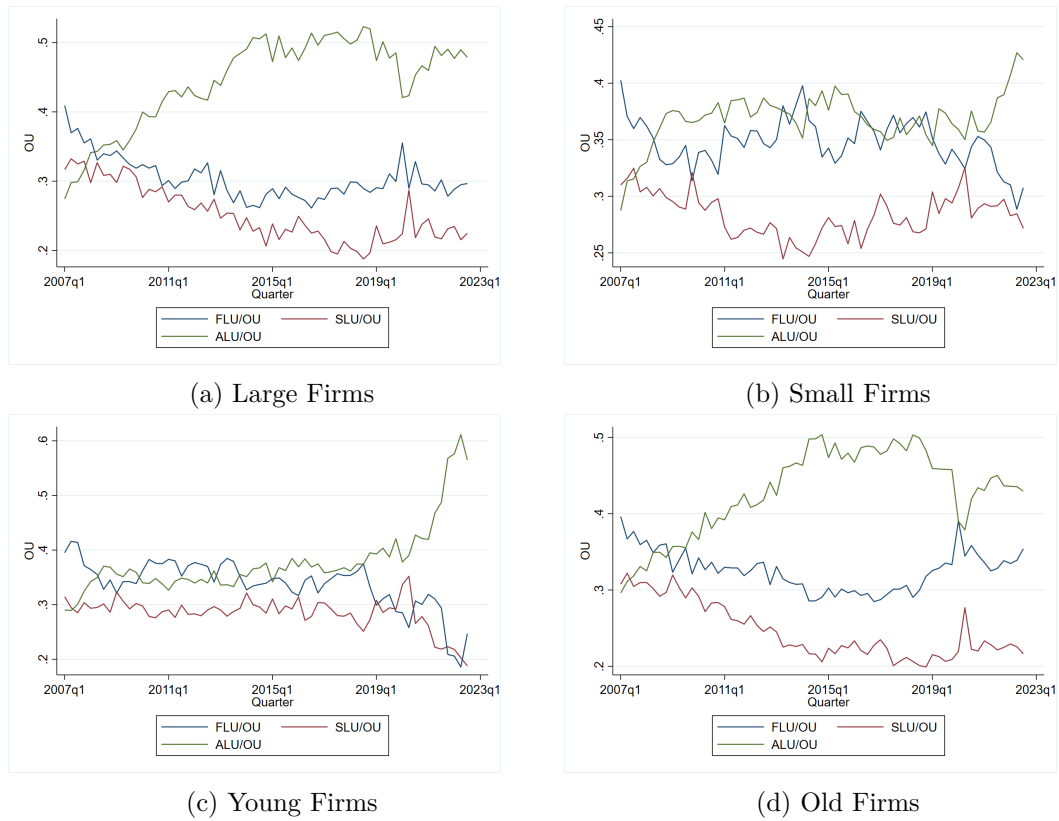


Figure 13: Large and small firms' proportions of uncertainty tiers over time

B Regressions of Uncertainty Derivation

Dependent Variable: Log of Sales _{<i>i,t</i>}			
Uncertainty Component →	(OU)	(FLU)	(SLU)
Log of Sales _{<i>i,t-1</i>}	0.3825*** (0.0017)	0.6223*** (.001264)	0.3791*** (0.0030)
Log of Age _{<i>i,t</i>}	-	0.0249*** (.0013)	-
Log of Market Value _{<i>i,t</i>}	-	0.2423*** (.0012)	-
Log of Capital Intensity _{<i>i,t</i>}	-	-0.4727*** (.0015)	-
<i>Cfact</i> _{<i>i,t</i>}	-	-	0.6486*** (0.0086)
Firm FE	✓	✓	✓
Time FE	✓	✓	✓
Observations	679,977	166,023	163,498
Firms	19,896	5,695	5,830

Table B.1: GMM Estimators for the baseline specification

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval,

* Significant with 90% Confidence Interval

Regression Result for Robustness Checks

Dependent Variable: Log of Sales _{<i>i,t</i>}	
Log of Sales _{<i>i,t-1</i>}	0.3475*** (0.0033)
<i>CFact</i> _{<i>i,t</i>}	0.6600*** (0.0104)
<i>FLU</i> _{<i>i,t</i>}	-9.0410*** (0.5230)
Firm FE	✓
Sector FE	✓
Observations	134,826
Firms	4,816

Table B.2: Regression Results for *SLU* with Controlling for FLU

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Chapter 2

Firms' Investment and Capacity Utilisation¹

The Role of Financial Constraints and Uncertainty

Abstract

Information on capacity utilisation is central to studying firms' investment behaviour. We draw a framework where firms at high capacity utilisation need to invest in order to maintain their capacity buffer to cope with future shocks. We then test the relationship between firms' deviation from this "target" and their investments by considering the role of uncertainty and financial constraints by applying an instrumental variable model to the firms present in the INVIND survey of the Bank of Italy for 2002-2023. Our findings show that the Italian firms with high growth potential or at the growth window invest in their production capacity. Importantly, we find that higher uncertainty does not affect the investments of firms operating at high capacity utilisation but increases the investment of those at low capacity utilisation. This result from our work hints at a novel inefficiency role for uncertainty through its heterogeneous impact on firms based on their capacity utilisation.

Keywords: Capacity Utilisation, Investment, Uncertainty, Financial Constraints, Firm Growth

¹This chapter is co-authored with Giulia Piccillo, Maria Savona and Tania Treibich

2.1 Introduction

"The growth of a nation's productive potential is the central factor in determining its growth in real wages and living standards.... high rates of investment and saving usually have a big payoff in promoting economic growth." — Paul A. Samuelson

Over the past few decades, the Italian economy has grappled with frequent and substantial shocks, the consequences of which have been strongly persistent. Italy not only suffered during the Great Recession of 2008-2009 (Zamagni, 2018) but also struggled to manage the fallout from the European debt crisis (Bull, 2020) and experienced frequent changes in government between 2011 and 2021 (Balduzzi et al., 2020) followed by the pandemic shock. These shocks have affected various aspects of the economy such as investment (Bond, Rodano and Serrano-Velarde, 2015; Buseti, Giordano and Zevi, 2015), productivity (Bugamelli et al., 2018), (mis)allocation of resources and low market dynamism (Calligaris et al., 2018; Cucculelli and Peruzzi, 2020), labour share (Bloise, Brunetti and Cirillo, 2021) and fiscal policy uncertainty (Anzuini, Rossi and Tommasino, 2020) among other aspects and have resulted in a lack of productivity growth overall (Dosi et al., 2012; Zeli, Bini and Nascia, 2022; Fernald and Inklaar, 2020). Even though the capacity utilisation of Italian firms has remained one of the lowest among European economies, the economic literature has overlooked the pivotal role of a structural and persistent drop in Italian firms' capacity utilisation in explaining the episodes of poor economic performance. To address the gap in the literature, we study the potential role of lower capacity utilisation on lower investments alongside other determinants of investments, such as uncertainty and financial constraints.

According to Nelson (1989), capacity utilisation is *"the ratio of actual to the maximum potential output consistent with a given capital stock"*. This measure has been repeatedly used as an indicator of macroeconomic cycles and productivity in the economic literature (Krugman, 1994; Pozzi and Schivardi, 2016; Butters, 2020), especially in the manufacturing sector (Corrado and Matthey, 1997). Once a firm (or an economy) operates at a high capacity utilisation rate (uses its inputs more effectively), it will require more investments to maintain its efficacy and achieve growth (Axsäter and Olhager, 1985; Artica, 2023). On the other hand, lower capacity utilisation is associated with higher average costs and lower productivity growth (Hulten, 1986; Butters, 2020; Ray, Walden and Chen, 2021). At the country level, we document a tight association between average capacity utilisation and aggregate investments (represented by Gross Fixed Capital Formation or GFCF) of selected European economies in Figure 2.1 below:

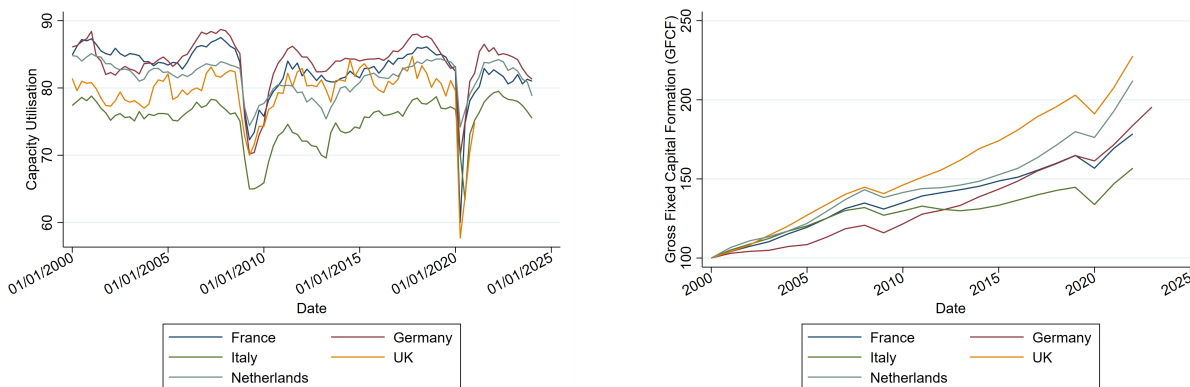


Figure 2.1: Positive relationship between aggregate annual investments and quarterly capacity utilisation at the country level. Data source: OECD (2024)

As shown in this graph, the Italian economy has experienced the lowest capacity utilisation and aggregate investments compared to other European economies. Figure 2.2 below provides a more detailed context for Italy. Within Italy, the quarterly correlation between aggregate investments and capacity utilisation is positive, with a magnitude of 0.38. Capacity utilisation is more strongly counter-cyclical with respect to the business cycle than aggregate investments, measured by their correlation with the recession dummy defined as OECD recession bands (-0.32 vs -0.10).

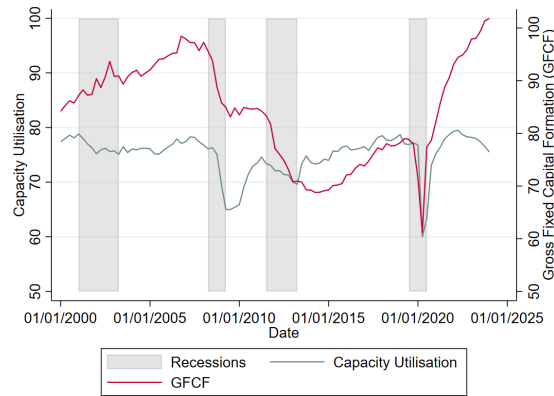


Figure 2.2: Capacity utilisation vs aggregate investments in Italy. Data source: [OECD \(2024\)](#)

Since the Great Recession and the European debt crisis, Italian firms have displayed a structurally lower capacity utilisation and low investments and have faced tight financial constraints ([Minetti and Zhu, 2011](#)) and high uncertainty ([Bontempi, Golinelli and Parigi, 2010](#); [Gufler et al., 2020](#)). In this environment, [Busetti, Giordano and Zevi \(2015\)](#) argued that low capacity utilisation may be a driver of manufacturing firms' poor outcomes. The years after the Great Recession have demonstrated low levels of GFCF, accompanied by a drop of 40 per cent between early 2007 and 2020.

Firms' investment decisions are determined primarily depending on uncertainty and business cycle ([Gourio and Kashyap, 2007](#); [Doms and Dunne, 1998](#)) and Tobin's q ([Blundell et al., 1992](#)). Real options theory has become a central explanation of why investments dampen in highly uncertain episodes of the business cycle ([Dixit and Pindyck, 1994](#)). This theory argues that uncertainty over future demand reduces firms' ongoing investment due to a higher real option and capital irreversibility, increasing the value of waiting and delaying investments.² Enriching the real options model by considering the role of capacity utilisation, [Abel et al. \(1996\)](#) set up a model where capital investment decisions are limited by firms' ability to sell later or expand the capacity: the option to expand reduces the incentive to invest, contrary to the option to disinvest. In addition, firms' investments rely on external financing via borrowing, ([Fazzari and Petersen, 1988](#); [Amiti and Weinstein, 2018](#)), which can be scarce during turbulent times for reasons such as tight credit supply, lower credit ratings and asymmetric information on firms' performance. Moreover, the firm growth and expansion process is not the result of smooth investments over time but a few extensive and costly investment events for firms, making investment a lumpy process featuring bumps and jumps ([Doms and Dunne, 1998](#); [Arata, 2019](#)).³ Periods of inaction followed by bursts of investments is a pervasive manifestation of capital adjustment frictions ([Baley and Blanco, 2021](#)). Therefore, this spikey nature of investments further

²However, some studies find that such results do not hold in competitive markets even in the presence of irreversibility ([Abel and Eberly, 1994](#)). Furthermore, [Pozzi and Schivardi \(2016\)](#) show that firms only partially respond to shocks due to adjustment frictions.

³In other words, investment is not buying a squared meter of land per day, but buying a large area at one time ([Penrose and Pitelis, 2009](#)).

extends the dependence of decisions upon investments on external financing conditions (Im, Mayer and Sussman, 2020) and increases capital adjustment costs since firms' internal funds are insufficient to finance these large investment spikes. Unlike expansion investments that display a spiky nature, smooth investment rises, and falls are usually due to the maintenance and replacement of depreciated capital. Throughout our analysis, we will take the spikiness of investments into account by isolating the large investments that may better represent expansion investments than capital replacements represented by smooth investment rates.

Previous studies have demonstrated that investment predictions of firms based on the business cycle and Tobin's q are subject to bias if the firm's capacity utilisation is not accounted for (Grullon and Ikenberry, 2021). Over the span of capacity utilisation, a trigger point (or capacity target) exists that stimulates firms' decisions to invest (Brown and Mawson, 2013), especially when faced with unanticipated demand shocks (Abel, 1981). To minimise their costs, firms choose their technology-dependent long-run capacity utilisation targets endogenously - which are lower during recessions (Nikiforos, 2011). A firm that operates above its desired capacity utilisation is known to be at the *growth window*⁴ and are more likely to invest, expand, and achieve growth (Coad and Planck, 2012). Nevertheless, some firms at the growth window will not invest and instead stall or shrink due to low growth desires or external factors (Loderer, Stulz and Waelchli, 2017; Boot and Vladimirov, 2019). Thus, firms at the growth window can follow different paths, or in other words, they have reached a "fork in the road" (Coad et al., 2021). This theory and the associated empirical results suggest that a trigger point of capacity utilisation exists through which firms decide upon their future and either invest, pause, delay, or shrink (Coad and Srhoj, 2020). We extend this framework by considering how the execution of the planned investment depends upon factors such as the prevailing uncertain macroeconomic landscape, similar to Saltari and Travaglini (2001); Bolton, Wang and Yang (2019) and availability of financial resources (potential financial constraints).

In summary, our research introduces a framework to study the interplay between firms' investments, capacity utilisation, uncertainty, and financial constraints. We aim to answer two main research questions: "Does a firm's investment likelihood depend on capacity utilisation and whether they have entered their growth window? How do financial constraints and demand uncertainty influence the relationship between capacity utilisation and investment? Is there heterogeneity between firms at and outside of the growth window?"

This paper's findings align with the previous literature and state that firms with high capacity utilisation and at the growth window have more investment incentives compared to firms with lower capacity utilisation. Even though we find that firms operating at high capacity are not sensitive to uncertainty regarding their investment decisions, the investments of those operating at low capacity are positively affected by uncertainty. Therefore, not only do our results highlight the significance of uncertainty and financial constraints for Italian firms at the growth window, but they also underscore the heterogeneity of the impact of uncertainty on firms' investments, similar to Howes (2023). In our sample, firms with low growth potential increase investments under high uncertainty regimes, while firms at the growth window remain unaffected. This insight helps us expand the literature on Italian

⁴Firms that observe a crucial state of proximity to the capacity target will see themselves at the window of opportunity for high growth in the next period. The importance of such windows in innovation, technology adaption, policy and industrial organisation has been studied previously; for instance, see Lee and Malerba (2017); Giachetti and Marchi (2017). We focus on the behaviour of these firms since they are the main drivers of aggregate investments and economic growth (Asturias et al., 2023)

firms’ inefficiencies building on [Secchi, Tamagni and Tomasi \(2016\)](#); [Schivardi, Sette and Tabellini \(2022\)](#); [Calligaris et al. \(2018\)](#).

The results of our analyses carry significant implications for industrial policy and stress that it can aim at supporting firms at the growth window. Identifying and understanding these trigger points can inform the development of cost-effective strategies to stimulate investment precisely when hesitant firms are at the critical growth window. Policymakers should recognise that trigger points are fleeting moments in a firm’s trajectory. Therefore, expediting the evaluation of decisions when supporting firms at the growth window is crucial. Our findings also validate previous research on the role of financial constraints in contracting firms’ investments. During periods of heightened borrowing constraints, such as recessions, allocating funds more effectively is needed by identifying firms with greater growth potential. Lastly, our discovery of the heterogeneous impact of uncertainty should catalyse monitoring and mitigating high uncertainty levels in the economy to prevent the inefficient allocation of resources to firms outside the growth window.

This paper is structured as follows. In Section 2.2, we outline our conceptual framework. The empirical strategy is detailed in Section 2.3. Section 3.3 presents the INVIND database and discusses the properties of the variables employed in our empirical analysis. Following that, Section 2.5 presents the results of our empirical analysis, subject to various robustness checks outlined in Section 3.6. Finally, Section 3.7 concludes our paper.

2.2 Conceptual Framework

This section presents a conceptual framework illustrating how changes in a firm’s realised capacity utilisation impact its investment decisions, building on the “fork in the road” framework introduced by [Coad et al. \(2021\)](#). Using this framework, we highlight the role of uncertainty and financial constraints in shaping these decisions.

2.2.1 Investment & Capacity Utilisation

Information on capacity utilisation can be beneficial for firms’ investment decisions. Given a fixed production capacity and a long-term utilisation rate target ([Di Domenico, 2023](#)), firms adjust their input utilisation rate to avoid the adjustment costs when they deviate from the target. However, suppose the capacity used is too high or low, and thus, the deviation from the target value is large. In that case, firms must revise their production capacity, as deviation from the capacity target and lacking a sufficient buffer threatens the firm with the opportunity cost of lost demand ([Abel, 1981](#); [Sarkar, 2009](#)). This response also enables the firm to freely adjust utilisation rates in subsequent periods as a precautionary measure, which is optimal both at the firm and social level ([Greenwood, Hercowitz and Huffman, 1988](#)). However, not all firms at high capacity invest and grow. Some firms operating at capacity stall or delay their investments due to internal conditions.⁵ The “fork in the road” framework, which resembles firms’ choice at a conjuncture of high input utilisation rate, was explained and empirically tested for firms in the “overcapacity” state by [Coad et al. \(2021\)](#). We contribute to the literature by extending this framework to explore not only the role of capacity utilisation in firms’ investment choices but also by incorporating uncertainty and financial constraints and their role in firms’ decisions to expand their production capacity by investing in new machinery.

⁵These conditions, such as the technology in use and capital intensity, matter to firms’ investment decisions, as discussed in detail by [deB. Harris \(1988\)](#).

On the other hand, if the capacity utilisation falls significantly below the target, firms perceive their situation as an excess of capital (or other inputs), lowering the probability of hiring and investment.

We now set up a conceptual framework to understand how information on capacity utilisation can determine firms' investment decisions. At the beginning of each period, firms decide about their inputs. Using all the input within their reach generates their (potential) production capacity (Y^{PC}). Next, they observe their sales (Y^S), corresponding to only a fraction of their production capacity, the sum demanded in the market throughout the period. Since firms decide to allocate part of inputs for high demand times (Fagnart, Licandro and Portier, 1999) or as a quick replacement when machinery is torn (Dagdeviren, 2016), we can assume a targeted capacity utilisation (CU^ψ), which is less than 100% (Morrison, 1985; Ray, Walden and Chen, 2021). The remaining production capacity to maintain a buffer ($100 - CU^\psi$) will be unused, stored as inventories or wasted (non-storable final product). In this case, the capacity utilisation (CU) is calculated as $CU = Y^S/Y^{PC}$. The timing and the decision plan of firms are presented in the following timeline:

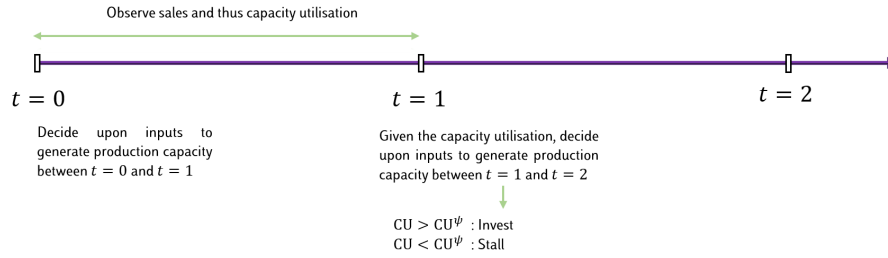


Figure 2.3 provides a snapshot of firms' decision and the corresponding tree when at the growth window ($CU > CU^\psi$). The larger rectangle represents a firm's production capacity (Y^{PC}). The blue area displays the firm's sales during the period (Y^S). The proportion of sales out of production capacity defines its capacity utilisation (CU). The white area represents the production capacity that is not sold - whether stored, unused or wasted. Using information on its utilised capacity, the firm decides upon the inputs for production in the next period. Suppose its observed capacity is larger or equal to its target CU^ψ . This firm is in the "growth window" since it can invest in production capacity and growth.⁶

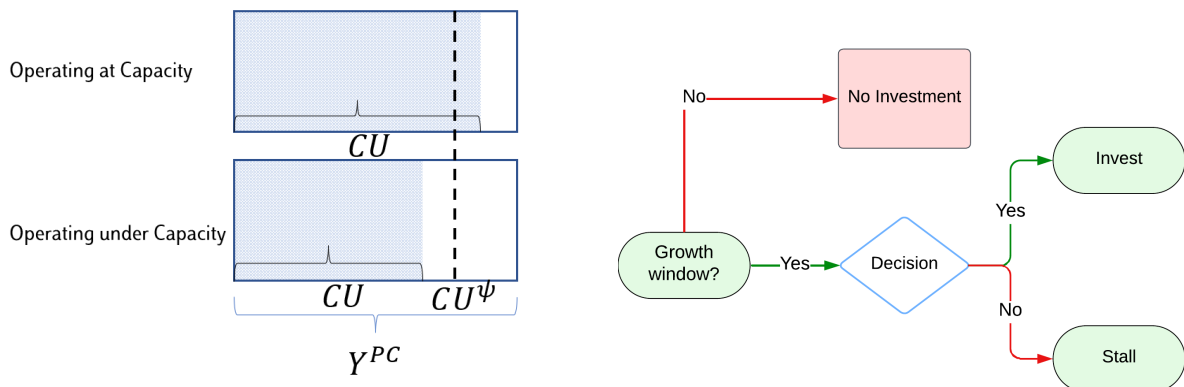


Figure 2.3: Firms decision tree at the growth window

Therefore, firms' investments depend both on whether their capacity utilisation has surpassed the target ($CU > CU^\psi$) and how distant the firm is from its target, namely the gap ($CU - CU^\psi$). The

⁶This framework can abstract from the presence of inventories (I) since their total sum is usually less than both production and sales. Therefore, whether the inventories from t are taken to $t + 1$ does not affect the mechanism.

farther the firm is from its target, positively or negatively, will increase or decrease the probability of investments. If the firm decides to invest, its production capacity will increase. On the other hand, if it does not invest, it will, at best, stall or shrink due to capital depreciation.

2.2.2 The Role of Uncertainty and Financial Constraints

So far, our framework assumes that firms' desired investment can be realised because there are no financial constraints, and the role of uncertainty was ignored.⁷ In practice, this might not be the case due to high capital intensity (large capital reserves relative to sales or labour) or higher marginal cost due to high uncertainty, higher borrowing rates and old technology. To relax a part of these assumptions, we will now move to a setting where uncertainty and financial constraints enter as external factors and affect firms' investments.

Once a firm decides to expand its capacity via investments and faces high uncertainty, its real option increases, and it becomes more cautious in investing (Bloom, Bond and Van Reenen, 2007). Furthermore, if the firm decides to fund these investments through external financing, it might not observe a fully available financing possibility in the capital market and, therefore, face financial constraints that do not allow the firm to receive the required funds. In summary, uncertainty and financial constraints are the crucial obstructions in firm decisions upon investing. In Figure 2.4 below, a firm's decision at the growth window depends on its uncertainty and borrowing constraints in two stages:

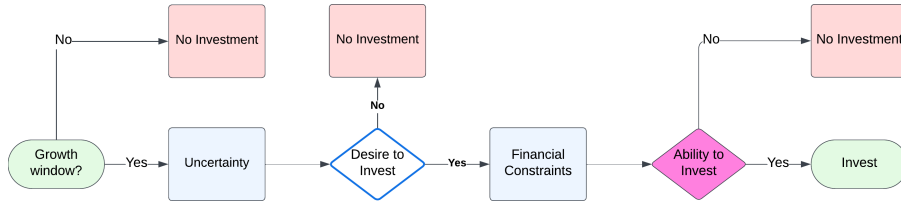


Figure 2.4: Firms decision tree at the growth window considering uncertainty and financial constraints

In summary, our conceptual framework states that when a firm's uncertainty is low, it might invest and expand its production capacity (Y^{PC}) - subject to whether its investment decision can be realised due to financial constraints in the second step. Suppose it fails to invest in production capacity due to high uncertainty or tight financial constraints. In that case, its production capacity will shrink (due to depreciation) or stall if it replaces the depreciated capital with new units.

2.3 Empirical Approach

In this section, we explain our empirical approach for testing the implications of the conceptual framework presented in Section 2.2. We test whether the positive distance of capacity utilisation to target ($CU - CU^\psi$) and being at the growth window ($CU > CU^\psi$) affect firms' investments, represented by investment rates ($I_{i,t}/\bar{Y}_{i,t-1}^S$), which normalises the investment values for firms' size and investment spikes ($\mathcal{S}_{i,t}$). In what follows, we will focus on the capacity utilisation gap ($CU - CU_{s,\tau}^\psi$) and whether firms have surpassed their target and namely, are at the growth window. We also assume this target is constant within each sector (s) and depends on the economy's state (τ).

⁷When firms operate in a market economy with differentiated goods, it is optimal to invest and grow as there are motives behind investment as long as Marginal Return to Investment is larger than Marginal Cost of Investment - $MR^I \geq MC = r^{RF}$ and there are no hurdles in internal or external financing.

Our econometric framework captures the dynamics between firms' investment and the gap between firms' capacity utilisation and a sector-specific target as follows:

$$y_{i,t} = \begin{cases} I_{i,t}/\overline{Y^S}_{i,t-1} \\ Pr[\mathcal{S}_{i,t} = 1] \end{cases} = \rho(CU_{i,t} - CU_{s,\tau}^\psi) + X_{i,t}\beta + \mu_i + \gamma_t + \varepsilon_t \quad (2.1)$$

The dependent variable $y_{i,t}$ represents firms' investment rates, measured in two ways. First, we consider the investment rate $I_{i,t}/\overline{Y^S}_{i,t-1}$ (investment normalised by sales) and in a different specification we focus on large investment events, the investment spikes, defined as the probability of spikes of the investment rate ($Pr[\mathcal{S}_{i,t} = 1]$).⁸ The vector $X_{i,t}$ contains firm-level controls, and μ_i and γ_t represent firm and time-fixed effects, respectively, to control for the impact of time in-varying firm characteristics and business cycles.

As discussed in Section 2.2, firms that operate above the target are at the growth window. Therefore, their decisions differ from those of all the other firms as they highly desire investment. To capture this non-linearity in the gap to capacity target ($CU_{i,t} - CU_{s,\tau}^\psi$), we test our framework with the overcapacity (growth window) dummy, defined as $\mathcal{O}_{i,t} = 1$ if $CU_{i,t} > CU_{s,\tau}^\psi$.

As explained in more detail in Section 2.2, uncertainty and financial constraints hinder firms' external financing of new product development and other investment opportunities. Our econometric framework, considering their role, takes the form of:

$$y_{i,t} = \begin{cases} I_{i,t}/\overline{Y^S}_{i,t-1} \\ Pr[\mathcal{S}_{i,t} = 1] \end{cases} = \alpha_1(CU_{i,t} - CU_{s,\tau}^\psi) + \alpha_2 FC_{i,t} + \alpha_3 U_{i,t} \\ + \alpha_4 FC_{i,t} \times U_{i,t} + \alpha_5 FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi) \\ + \alpha_6 U_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi) \\ + X_{i,t}\beta + \mu_i + \gamma_t + \varepsilon_{i,t} \quad (2.2)$$

In Equation 2.2, $U_{i,t}$ represents uncertainty faced by firm i at time t , and $FC_{i,t}$ is a dummy for financially constrained firm-year observations. This equation, therefore, estimates not only the impact of the gap to capacity target in determining investment rates (spikes) but also measures the role of financial constraints, uncertainty and their interactions with each other, and the capacity utilisation gap to target. Similar to Equation 2.1, we also test this framework for $\mathcal{O}_{i,t}$.

We are aware of the endogeneity concerns that might arise due to simultaneity (Almeida and Campello, 2001), omitted variables bias (Alti and Tetlock, 2014) and measurement errors since we do not have comprehensive information on firms' financial data. The sample selection bias can also drive the relationship between financial constraints and investments, as large firms primarily constitute our sample. Reverse causality can also occur when financial constraints and investments affect each other, complicating causality determination. For instance, firms that invest more intensively over time may be less financially constrained due to higher credit ratings or better bank relationships.

To address these issues, we will employ an instrumental variable approach to study the role of financial constraints without endogeneity (ex-ante financially constrained firms) and simultaneity. We follow Alfaro, Bloom and Lin (2022), who used a 5-year lagged dummy for capturing ex-ante financially constrained firms. Given our access to firms' explicit answers to whether being financially

⁸Investment spikes attempt to isolate large (expansion) investments from the normal wear and tear capital replacements. For a comprehensive review, please refer to Arata (2019); Bachmann, Elstner and Hristov (2017).

constrained, we modify their method by reducing the number of lags to one to determine the status of being financially constrained and as an instrument. Their method uses a proxy based on financial data and Moody's aggregate index. In our work, we do not rely on indirect proxies for financial constraints and use firms' direct responses to the questions regarding their financial constraints. Therefore, one lag suffices and provides us with a larger number of observations in estimations. Our instrument is well-relevant as many firms remain financially constrained for consecutive years, and the constraints in the past cannot theoretically explain the probability of delay or large investments in the present or future.

In our two-stage control function regression following Angrist and Pischke (2009), the first stage predicts the probability of remaining financially constrained for the previously constrained firms.

$$Pr(FC_{i,t} = 1) = \lambda FC_{i,t-1} + \mu_i + \gamma_t + u_{i,t} \quad (2.3)$$

The results of this regression are provided in Appendix A.6. This equation captures the probability that a firm remains consecutively financially constrained (for two periods), considering the business cycle and firms' unobserved characteristics captured by time and firm fixed effects, respectively.

We then plug the predicted probability from the first stage Equation 2.2 as $(\widehat{FC}_{i,t} \equiv E_t[Pr(FC_{i,t} = 1)] = \lambda FC_{i,t-1})$ into Equation 2.2 and estimate the new equation below:

$$\begin{aligned} y_{i,t} = \begin{cases} I_{i,t}/\overline{Y}_{i,t-1} \\ Pr[\mathcal{S}_{i,t} = 1] \end{cases} &= \alpha_1(CU_{i,t} - CU_{s,\tau}^\psi) + \alpha_2\widehat{FC}_{i,t} + \alpha_3U_{i,t} \\ &+ \alpha_4\widehat{FC}_{i,t} \times U_{i,t} + \alpha_5\widehat{FC}_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi) \\ &+ \alpha_6U_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi) \\ &+ X_{i,t}\beta + \mu_i + \gamma_t + \varepsilon_{i,t} \end{aligned} \quad (2.4)$$

This equation isolates the role of capacity utilisation, financial constraints, and uncertainty and their interaction with our outcome variables of interest without the endogeneity issues that we discussed previously. The predicted values of financial constraints ($\widehat{FC}_{i,t}$) probabilities have now replaced the dummy $FC_{i,t}$ in our attempt to solve for the endogeneity of financial constraints to uncertainty and investments.

We now proceed by demonstrating the estimation methods which will be applied to regression equations shown so far.

2.3.1 Estimation Methods

Throughout this paper, we employ two econometric methods for estimating Equations 2.2, 2.3, and 2.4, which depend on the dependent variable. We apply a standard panel OLS approach when the dependent variable is the investment rate. For regressions with binary outcomes, those involving investment spikes ($\mathcal{S}_{i,t}$) and financial constraints ($FC_{i,t}$)—we use the conditional logit (CLogit) method. This method is particularly suitable for modelling binary outcomes in matched or highly stratified datasets, as discussed by Hosmer Jr, Lemeshow and Sturdivant (2013) and for panel data as it allows for the inclusion of fixed effects, thereby controlling for unobserved characteristics.

In conditional logit models, parameter estimation is based on variation within each group. If a panel unit, such as an individual or firm, consistently exhibits the same outcome across all observations (i.e.,

all outcomes are either 0 or 1), the model cannot establish comparisons within that group, rendering it uninformative for parameter estimation. Therefore, the model will drop the firms without any investment spike from the regression.

Following [Greene \(2018\)](#), let $i = 1, 2, \dots, n$ denote firms and $t = 1, 2, \dots, T_i$ represent the observations for the i th firm. The dependent variable, y_{it} , takes a binary value of 0 or 1, corresponding to investment spikes or financial constraints. The vector $\mathbf{y}_i = (y_{i1}, \dots, y_{iT_i})$ represents the outcomes for the i th firm, while $\mathbf{x}_{it}$ is a row vector of covariates. Let k_{1i} denote the observed number of ones for the dependent variable within the i th firm, expressed as:

$$k_{1i} = \sum_{t=1}^{T_i} y_{it}$$

Thus, there are k_{1i} cases matched to $k_{2i} = T_i - k_{1i}$ controls within the i th group. According to [Hosmer Jr, Lemeshow and Sturdivant \(2013\)](#), the probability of observing $\mathbf{y}_i$ conditional on $\sum_{t=1}^{T_i} y_{it} = k_{1i}$ is:⁹

$$\Pr \left(\mathbf{y}_i \mid \sum_{t=1}^{T_i} y_{it} = k_{1i} \right) = \frac{\exp \left(\sum_{t=1}^{T_i} y_{it} \mathbf{x}_{it} \boldsymbol{\beta} \right)}{\sum_{\mathbf{d}_i \in S_i} \exp \left(\sum_{t=1}^{T_i} d_{it} \mathbf{x}_{it} \boldsymbol{\beta} \right)}$$

Here, d_{it} takes values of 0 or 1, with $\sum_{t=1}^{T_i} d_{it} = k_{1i}$, and S_i represents the set of all possible combinations of k_{1i} ones and k_{2i} zeros. Although there are $\binom{T_i}{k_{1i}}$ such combinations, the denominator of the above equation can be computed recursively without enumerating all combinations.

Let the denominator be denoted by:

$$f_i(T_i, k_{1i}) = \sum_{\mathbf{d}_i \in S_i} \exp \left(\sum_{t=1}^{T_i} d_{it} \mathbf{x}_{it} \boldsymbol{\beta} \right)$$

Computationally, the recursive formula for f_i as the number of observations increases is given by:

$$f_i(T, k) = f_i(T-1, k) + f_i(T-1, k-1) \exp(\mathbf{x}_{iT} \boldsymbol{\beta})$$

with the initial conditions $f_i(T, k) = 0$ if $T < k$ and $f_i(T, 0) = 1$.

The conditional log-likelihood function is then expressed as:

$$\ln L = \sum_{i=1}^n \left\{ \sum_{t=1}^{T_i} y_{it} \mathbf{x}_{it} \boldsymbol{\beta} - \log f_i(T_i, k_{1i}) \right\}$$

The derivatives of the conditional log-likelihood function can also be obtained recursively by differentiating the recursive formula for f_i . A maximum likelihood approach estimates the coefficients that maximise the log-likelihood function above.

This section introduced the empirical approach used in this paper as the econometric counterpart of the conceptual framework discussed in Section 2.2. We now explore the variables that enter this empirical framework and the data used for our study.

⁹The extensive derivation of the probability function is outside of our scope. For a detailed derivation, please refer to [Hosmer Jr, Lemeshow and Sturdivant \(2013\)](#) or [Greene \(2018\)](#).

2.4 Data and Variables

2.4.1 Dataset

In this work, we use the 2002-2023 panel of the “INVIND” ([Bank of Italy, 2024](#))¹⁰ firm-level dataset. This survey is administered to approximately 4,000 Italian firms with at least 20 employees in the manufacturing and services sectors. Bank of Italy has continuously collected information on firms’ key characteristics such as employment, investment (actual and planned), turnover, debt and trade receivables. The survey also collects information on annually-changing topics of particular interest for economic research (e.g. corporate strategies and governance, physical, human and organisational capital, and electric power). The INVIND surveys include questions on firms’ actual and expected capacity utilisation. Additionally, since the Great Recession of 2007-08, INVIND includes questions on firms’ financing and relevant conditions and constraints. However, this dataset does not contain information on all of the firms’ balance sheet items and plant-level production processes.

INVIND’s coverage increased after 2002, as it was introduced to more manufacturing and non-financial private service firms with 20 or more employees, excluding credit institutions, insurance companies, public services, and other social and personal services. As a result, participating firms have increased to more than 4,000 (3,000 in the manufacturing sector and 1,000 in services). In what follows, our analysis will be limited to post-2002 survey waves to avoid potential mismeasurements. Furthermore, we focus on the manufacturing sector, which constitutes the majority (72%) of all observations and is shown to be more sensitive to capacity utilisation dynamics ([Berndt and Fuss, 1982](#); [Koenig, 1994](#)). Out of 90,347 observations in 2002-2023 from 11,960 unique firms, more than 64,934 observations correspond to 8,410 unique firms in the manufacturing sector. Lastly, the variable that defines the most granular sectors in the economy available to us is based on the aggregation of [ISTAT ATECO 2002](#) and [ISTAT ATECO 2007](#) and divides the manufacturing firms into 7 sectors. Table A.1 in Appendix A provides a detailed list of this aggregation and the resulting sectors for use.

We now demonstrate the variables we use from INVIND to construct measures corresponding to the empirical analysis.

2.4.2 Variables

This section presents a detailed discussion of our variables and then discusses how these steps help frame our econometric model as described in Section 2.3. We develop various metrics for investment, uncertainty, employee and sales growth and reliable controls. The questions from the INVIND survey used in our analysis are provided in Appendix D.

Measuring Investment The INVIND survey includes questions on both realised investments each year (t) and the expected investments in the next year ($t+1$) in nominal euros, which are then converted

¹⁰This survey has been collected under different names since 1984. The current name is INVIND or “Indagine sugli investimenti delle imprese manifatturiere” (Inquiry into investments of manufacturing firms). The high-quality data collection and supervision has attracted many researchers to use this data to analyse various economic topics such as investment ([Bond, Rodano and Serrano-Velarde, 2015](#)), capacity utilisation ([Locatelli, Monteforte and Zevi, 2016](#)), productivity and firm growth ([Pozzi and Schivardi, 2016](#)) and labour outcomes ([Daruich, Di Addario and Saggio, 2023](#)). The dataset is accessible through an online platform ([RemoteExecution - REX](#)) of code execution (R and STATA). The output can only include text (.txt format). This shortcoming means that we cannot investigate the distributional properties of variables, such as their histograms or Kernel densities. Furthermore, the direct access to view the data and single observations is limited for confidentiality reasons. Hence, we cannot observe the percentiles of the variables or minimum and maximum. The use of the dataset is discussed in more detail by [Bruno, D’Aurizio and Tartaglia-Polcini \(2014\)](#).

to real values using deflators. As a firm's size typically influences investment values, the variable should be normalised using an appropriate proxy for this factor. The literature has used various forms of investment rate to tackle this issue (Belo, Lin and Bazdresch, 2014; Grazzi, Jacoby and Treibich, 2016; Alfaro, Bloom and Lin, 2022)¹¹, which commonly take the shape of $\frac{I_t}{K_{t-1}}$ or $\frac{I_t}{0.5(K_t + K_{t-1})}$, where I_t is the flow and K_t is the stock of plant and equipment. Our dataset does not offer insights on firms' capital stock. Therefore, we rely on intuitions from Baum, Caglayan and Talavera (2008) and use the average recent sales ($\overline{Y^S}_{i,t-1} = 0.5 [Y^S_{i,t} + Y^S_{i,t-1}]$) instead of capital (K) to normalise investments. In our case, the use of average recent values is motivated by the volatility of annual sales. Ultimately, to ease our notation, we show our investment rate as $I_{i,t}/\overline{Y^S}_{i,t-1}$ and calculate it as:

$$I_{i,t}/\overline{Y^S}_{i,t-1} = \frac{I_{i,t}}{0.5(Y^S_{i,t} + Y^S_{i,t-1})} \quad (2.5)$$

Using the same methodology, we also calculate the expected investment rates by replacing the investment values at t with expected values for $t + 1$ reported by the firms. This additional step will help us in the later stages of our analysis, where we attempt to find the spikes in the investment process.¹²

The abovementioned investment rate does not distinguish between expansion and replacement (or maintenance) investments. The literature has previously discussed that investment is discrete, and firms decide to invest in some years and only do little maintenance in others (Arata, 2019). In other words, most of the observed investments in plants and machinery are for replacement, maintenance or depreciation (Mobley, 2011). Thus, in order to analyse firms' investment decisions, we have to isolate the investment spikes from the other less meaningful periods and consider the lumpiness of investments. Ideally, a spike definition must satisfy two properties: 1) it should display a sudden jump in investment rates (Grazzi, Jacoby and Treibich, 2016) and 2) it must be unexpected to the firm in the previous period (Bachmann, Elstner and Hristov, 2017).

To understand at what point in the investment distribution the spikes occur, we need to observe the ordered distribution (based on percentiles) of the realised and expected investment rates. In the distribution of realised investment rates, the kink point will determine the first property for an investment spike by displaying the sudden jump. Differently, a post-kink large positive distance between the ordered distribution of realised and expected investment rates displays the second property of an investment spike, namely being unexpected to the firm.

However, we cannot observe this ordered distribution and relevant percentiles for confidentiality reasons as they grant access to single observations. Therefore, we construct a proxy for these percentiles as follows. We start by identifying the investment rate percentiles in each year and storing them in the background. We then calculate the average investment rate at percentiles, namely calculating the average value for each percentile observed over the years. Using formal notations, we can write:

$$\text{Average Annual Realised Investment Rate at Percentile } i \equiv \bar{P}^I(i) = \frac{1}{T} \sum_{t=1}^T P_t^I(i)$$

¹¹Furthermore, normalisation is carried out to reduce the skewness and smooth out any other non-normal characteristics.

¹²The availability of expected values allows us to consider the role of a spike in planned investment rates and check whether it only occurs for planned investments or happens to the expected investments. Furthermore, it allows us to understand the change in the dynamics of the ranked distribution pre and post-spike and how observed investments differ from the expected values after the spike.

$$\text{Average Annual Expected Investment Rate at Percentile } i \equiv \bar{P}^{I,e}(i) = \frac{1}{T} \sum_{t=1}^T P_t^{I,e}(i)$$

Where T is the total number of years, $\bar{P}$ is the resulted value that we can observe, and P represents the percentiles that result directly from inquiry into data, and we cannot observe. Concerned about the possibility that single observations drive these averages at the tail of the percentiles, we also calculate the range average among observations between the percentiles of interest. These results are presented below:

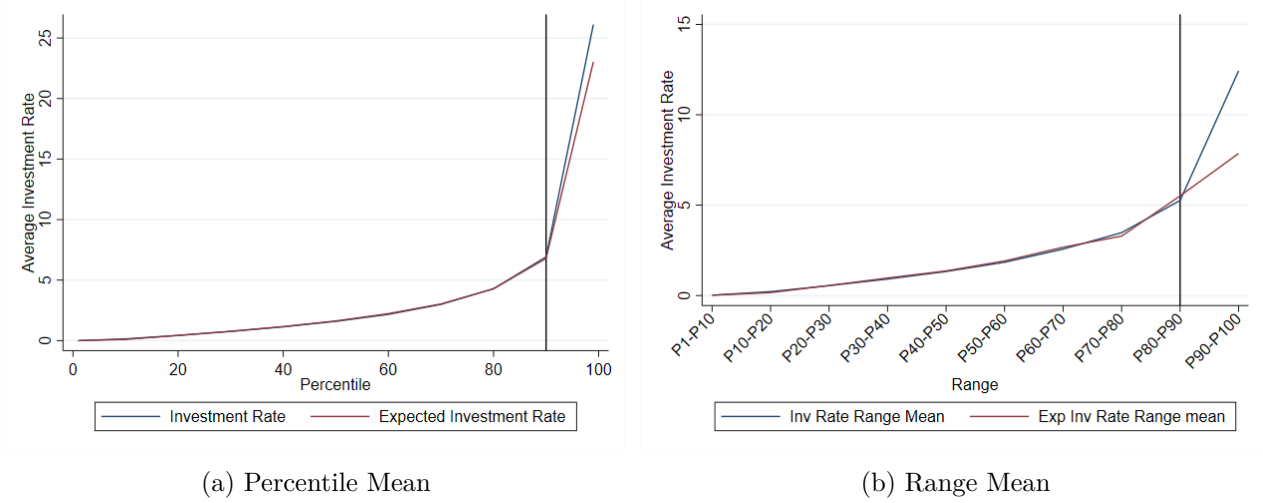


Figure 2.5: Average (expected) investment rates

We observe two phenomena. Primarily, and in Figure 2.5(a), we observe a sudden jump in investment rates above the 90th percentile, before which the growth is very smooth. This result is confirmed in Figure 2.5(b), where the spike is still present even though the jump is smoother by the intra-decile mean's construction. Hence, we can confirm the first criterion: a sudden jump at around the 90th percentile exists. Furthermore, we document another development after the 90th percentile: the distance between the blue (realised investment rates) and red (expected investment rates) expands significantly. This feature clearly shows the second important characteristic of a true spike, being unexpected to the firm in the previous period. The spike observations are assigned by setting the dummy $\mathcal{S} = 1$ for the investments at and beyond the 90th percentile of the investment rates across the whole sample. Our identification clarified that spikes are rare, and only a few firms observe spikes in each period.

Capacity utilisation gap and overcapacity measure We are interested in studying the role of the deviation of the observed capacity utilisation from the capacity target ($CU_{i,t} - CU_{s,\tau}^\psi$) as a determinant of firms' investments. First, we need to define the target ($CU_{s,\tau}^\psi$) within each sector s ¹³ as also discussed in Nikiforos (2013) and motivated by minimising costs due to returns to scale. Furthermore, these targets may differ depending on the state of the economy (τ) (Shapiro, Gordon and Summers, 1989; Dotsey and Stark, 2005) as shown by the strong cyclicity of capacity utilisation in Figures 2.1 & 2.2. During expansions - especially at the beginning - the average capacity utilisation across firms surges as firms use their resources to cope with the demand pressures. Differently, firms

¹³We assume that the one and two-digit sectors in our dataset have different technology and production processes from each other. Furthermore, a sector-wide approach allows firms with few observations, who have not responded to the survey frequently, to have a reasonable target.

adjust their utilisation rate towards lower values during recessions to avoid the adjustment costs associated with inputs. Therefore, there is a structural difference between expansion and recession periods. Aware of this issue and in our work, we use a sector-wide average of capacity utilisation in recessionary and non-recessionary periods and assign it to the firms in that sector as a capacity utilisation target. Ultimately, those firm/year observations with a higher capacity utilisation than the target will be assigned with the dummy $O_{i,t} = 1$ and interchangeably called “operating at capacity” or “at growth window”. This dummy helps us isolate the non-linear effect of capacity utilisation on firms’ investments.

Financial constraint measure Next, we categorise observations into those with and those without financial constraints. We use a set of questions in the INVIND survey implemented after 2008 to identify the firms that failed to receive loans (and funds) or chose not to inquire since they were sure they would not obtain them. We first use the question of whether firms desire to increase their debts. They were not financially constrained if they had no desire to increase their debt ($FC = 0$). Among those who desired to increase their debt, some have obtained their desired funds. We assign these observations as non-constrained ($FC = 0$). Those firms that either did not (partially) receive their funds or did not apply because they were sure they would not obtain the funds are assigned as constrained ($FC = 1$). A summary of our construction can be found below:

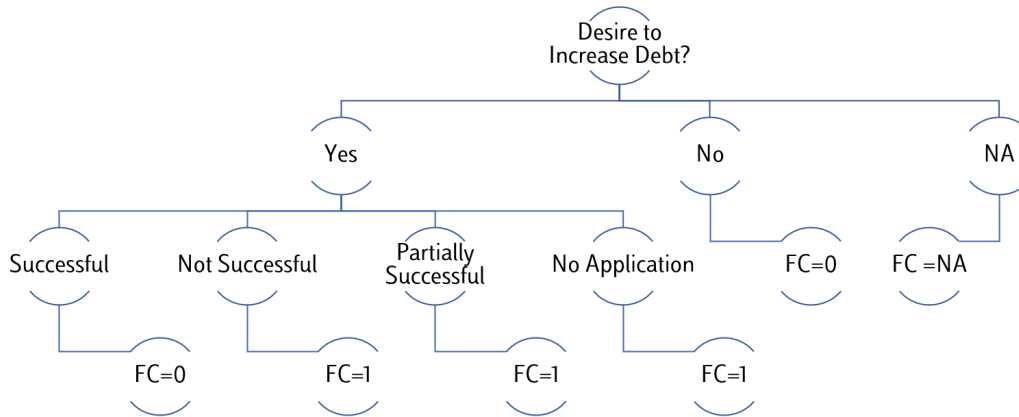


Figure 2.6: Division of firms into financially constrained, unconstrained and unassigned

Uncertainty measure To compute uncertainty, we use overall uncertainty developed by [Mohades, Piccillo and Treibich \(2024\)](#), computed as the cross-sectional dispersion of the residuals of an AR(1) process on sales. This method calculates firms’ overall uncertainty first by running the following AR(1) of sales to isolate the unpredictable component in the residual:

$$\log(S_{i,t}) = \omega \log(S_{i,t-1}) + \varepsilon_{i,t} \quad (2.6)$$

In the second step, it calculates the residual average at each firm ($\bar{\varepsilon}_i$) to measure the cross-sectional dispersion of these anomalies in T periods as:

$$U_{i,t} = \frac{(\varepsilon_{i,t} - \bar{\varepsilon}_i)^2}{T - 1} \quad (2.7)$$

This uncertainty measure captures the unpredictable component of a simple AR(1) process as the volatility. However, the observed volatility of returns is endogenous to diversification opportunities of

firms, as entrepreneurs can endogenously reduce risk by choosing safer, more conservative investments and products (Michelacci and Schivardi, 2013). To address endogeneity, we modify our measure of uncertainty by computing a sector-averaged uncertainty, or formally within each sector $s \in S$ as:

$$U_{s,t} = \frac{\sum_{i=1}^{N_s} U_{i,t}}{N_{s,t}} \quad \text{For } s = 1, \dots, S \quad (2.8)$$

The sectoral uncertainty is then normalised to a value ranging from 0 (no uncertainty) to 100 (extreme uncertainty).

Control variables The literature on the determinants of investment shows that firms' size and their growth, cash flows and profits are relevant factors determining such decisions (Nguyen, 2013; Bokpin and Onumah, 2009). Furthermore, firms choose strategic timing in their investments and hiring as these two complement each other in production. The past period(s) profits can be used as the liquid resources for ongoing investments via internal funding. Therefore, we use firms' growth (growth in the number of employees, lagged one year denoted as $L_{i,t-1}^G$) and profitability (lagged one year, denoted as $\pi_{i,t-1}$) as a measure of past performance as control variables. The information about firms' profitability in the INVIND survey consists of a five-category variable on the degree of profitability (or losses). This categorical variable consists of whether the firm is experiencing large losses, losses, no losses or profits, profits and large profits. From this variable, we construct a profitability dummy (π) that equals 1 when the firm is profitable and 0 otherwise.

Lastly, uncertainty is a pervasive phenomenon. It persists within firms over time and impacts the decision for more than one period. Our measures of uncertainty also conform to this behaviour, as shown by the high autocorrelation coefficients in Appendix A.4. Hence, our analysis also controls for lagged uncertainty ($U_{s,t-1}$).

2.4.3 Descriptive Statistics

Below, we provide the descriptive statistics of the variables used in our analysis:

Variable		Mean	Std Dev	Observations
Raw Variables				
$Y_{i,t}^S$	Sales	248,541	3,438,181	64,934
$I_{i,t}$	Investments in Plant and Machinery	8,277.66	80,087.66	38,925
$L_{i,t}$	Employees	261.21	1,118.08	64,934
Constructed Variables				
$I_{i,t}/\overline{Y}_{i,t-1}^S$	Investment Rate	3.64	26.80	38,922
$\mathcal{S}$	Spike	0.10	0.30	38,922
$CU_{i,t}$	Capacity Utilisation	76.65	15.01	52,274
$CU_{i,t} - CU_{s,\tau}^\psi$	Gap to target capacity	0.00	14.96	52,274
$\mathcal{O}_{i,t}$	Overcapacity	0.55	0.50	52,274
$U_{s,t}$	Sectoral Uncertainty	4.74	6.10	61,879
$FC_{i,t}$	Financially Constrained	0.09	0.29	43,380
$\pi_{i,t-1}$	Past Profitability	0.68	0.47	60,124
$L_{i,t-1}^G$	Past Employees Growth	0.01	0.11	51,115

Table 2.4.1: Data Description Table

Table 2.4.1 provides information on our variables' distribution's first and second moments. Our raw variables show that although small and medium firms are not included in this survey, there is significant variability across our sample. Furthermore, we see enough variation for the gap to target ($CU_{i,t} - CU_{s,\tau}^\psi$) to affect the investment rates and spikes. Additionally, more than half of the observations in our sample operate in the growth window. More than two-thirds of the observations in our sample showed profitability in the last period, while the annual employee growth rate averaged 1%. We also provide the descriptive statistics for firms at and outside the growth window in Table A.3 of Appendix A.5. Even though the firms at the growth window, on average, display less sales and investments, they are larger, have higher capacity utilisation by construction, and face less financial constraints and demand uncertainty. Furthermore, they are more profitable and more likely to have displayed growth.

We also provide a correlation matrix of these variables in Table 2.4.2. We see that firms operating at higher capacity display higher investment spikes but lower investment rates, meaning they have lower maintenance investments (displayed by lower investment rates) and larger expansion investments (displayed by more spikes). Furthermore, we can see that financial constraints occur more often for firms outside of the growth window, but they contract investments across the whole sample. Uncertainty for firms at the growth window is lower and is positively correlated to financial constraints, as demonstrated in more detail by Alfaro, Bloom and Lin (2022); Huynh (2024). Firms at the growth window have higher chances of profitability, which is positively correlated to expansion investments (spikes). Lastly, employee growth positively correlates to capacity utilisation, investments and profits.

	$I_{i,t}/\sqrt{Y_{i,t-1}^S}$	$S_{i,t}$	$I_{i,t}$	$CU_{i,t}$	$CU_{i,t} - CU_{s,\tau}^\psi$	$O_{i,t}$	$FC_{i,t}$	$\pi_{i,t}$	$U_{s,t}$	$U_{i,t}$	$L_{i,t}^G$	$Y_{i,t}^S$	$L_{i,t}$
$S_{i,t}$	0.20*												
$I_{i,t}$	0.03*	0.08*											
$CU_{i,t}$	-0.02*	0.01*	-0.00										
$CU_{i,t} - CU_{s,\tau}^\psi$	-0.02*	0.01*	-0.00	0.99*									
$O_{i,t}$	-0.00	0.01	-0.00	0.70*	0.70*								
$FC_{i,t}$	-0.01*	-0.03*	-0.02*	-0.11*	-0.10*	-0.08*							
$\pi_{i,t}$	-0.01	0.02*	0.03*	0.20*	0.19*	0.16*	-0.13*						
$U_{s,t}$	0.00	0.04*	0.04*	-0.03*	-0.02*	-0.05*	0.02*	-0.01					
$U_{i,t}$	0.04*	0.01*	0.00	-0.07*	-0.07*	-0.03*	0.02*	-0.04*	0.13*				
$L_{i,t}^G$	0.02*	0.03*	-0.00	0.10*	0.10*	0.09*	-0.05*	0.12*	-0.00	-0.02*			
$Y_{i,t}^S$	-0.00	-0.01*	0.72*	-0.02*	-0.02*	-0.00	-0.01*	0.02*	0.03*	0.00	-0.01		
$L_{i,t}$	0.00	0.01*	0.67*	-0.01*	-0.02*	0.01*	-0.02*	0.02*	0.03*	-0.01	0.00	0.56*	
Rec	0.01	-0.01	0.01	-0.05*	0.00	-0.02*	0.04*	-0.09*	0.19*	0.03*	-0.01	0.00	0.00

Table 2.4.2: Correlation Matrix. * denotes the 95% significance level

2.4.4 Investment, Capacity Utilisation and Business Cycle

We now investigate the behaviour of investments and capacity utilisation across the business cycle. Previous research has shown that firm-level investment and its spikes depend upon the business cycle (Grazzi, Jacoby and Treibich, 2016; Bachmann, Elstner and Hristov, 2017). In expanding episodes of the economy, strong motivations of growth increase firms' desire to invest in production capacity. However, consecutive periods of tightness reduce such motivations and decrease the available credit for the firms. Figure 2.7(a) demonstrates the annual share of firms with investment spikes featuring a firm investment decline up to 2014, coinciding with the European debt crisis. Since 2014 and similar to findings in the literature (OECD, 2021; Ciapanna et al., 2020; Mohades and Savona, Forthcoming), a recovery pattern towards pre-2007 values can be observed. A large part of this recovery can be attributed to the manufacturing sector and through policies that incentivised new technology and innovative investments (ISTAT, 2018).¹⁴ After all, spikes in expansionary investments can be due to purchasing plants and machines that complement innovative-related activities. In addition to firms' investments, capacity utilisation depends on the business cycle (Fagnart, Licandro and Portier, 1999). In the graph below, we demonstrate the annual average of capacity utilisation and show how it strongly co-moves with the business cycle, shown by large drops during recessionary periods. A persistent and sizeable drop after the great recession is also evident here, whereby the capacity utilisation remains lower compared to pre-2008 between 2014 and 2018.

¹⁴These policies primarily included banner Impresa 4.0 and Transizione 4.0, which attempted to offset the high regulatory burdens and levels of uncertainty faced by firms (OECD, 2021). Investment in assets supported and favoured by these policies, such as R&D and software and analytics, has proliferated after introducing these policies after the European debt crisis (Mohades and Savona, Forthcoming).

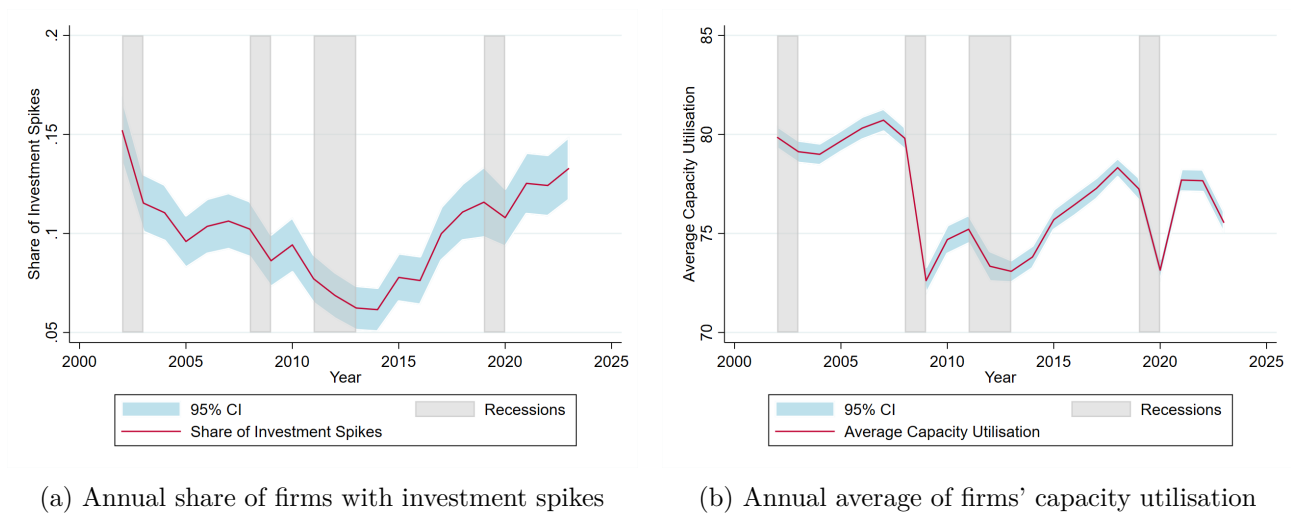


Figure 2.7: Investment spikes and capacity utilisation of Italian firms vs business cycle from 2002-2023. Data source: INVIND, [Bank of Italy](#) (2024).

Therefore, we can see a pro-cyclical pattern in investments and capacity utilisation at the aggregate level. At the beginning of the non-recessionary episodes, firms overuse their capacity to cope with the demand and avoid adjustment costs. After the initial phase of the boom, firms start investing and hiring to retain the capacity buffer to cover future unforeseen events. During recessions, firms abstain from adjusting their capital and hence adjustment costs by reducing the capacity utilisation rate. As the economy recovers, more and more inputs are used until the capacity utilisation target is reached again.

While it is expected that sectors with costly adjustment¹⁵ of capacity utilisation would behave differently from the norm, we cannot document evidence of a sizeable sector-level heterogeneity of the capacity utilisation target. The difference between sectors' targets barely reaches 3% in each business cycle phase. Table 2.4.3 presents these targets for each sector and each state of the economy:

¹⁵Some forms of production require the plant to operate non-stop as the cost of simply turning the machinery on and off is too high. Production of cement is a well-known example of such industries.

Sectors	Capacity Utilisation Target		
	Non-Recessionary	Recessionary	All Sample Mean
SS1 - Food industries, beverages and tobacco products	76.04	75.94	76.01
SS2 - Textiles, clothing, and hide, leather and footwear products	77.5	75.83	76.87
SS3 - Coke manufacturing, chemical industry, rubber and plastics	77.28	76.51	77.00
SS4 - Processing of non-metallic minerals	73.52	72.42	73.11
SS5 - Metal engineering industry	77.97	75.30	76.98
SS6 - Other manufacturing industries	77.56	76.34	77.12
SS7 - Other industries excluding construction (energy and extraction)	77.13	77.71	77.34
All Sectors	77.22	75.66	76.65

Table 2.4.3: Sector-level average (target) for capacity utilisation taken from INVIND database in annual frequency for 2002-2023 - divided by recessionary and non-recessionary periods.

The first observation is the slight difference between recessionary and non-recessionary periods across and within all sectors. The main reason for this result is the persistence of low capacity usage in non-recessionary periods after the Great Recession. The structural break that moved the average from almost 80% to 75% seems to have been quite persistent, as well as the macroeconomic indicators of the Italian economy, as also presented previously in Figures 2.2 & 2.7(b). Despite this low difference, we detect a pattern in which the business cycle impacts the sectoral target of capacity utilisation. We observe a higher capacity utilisation for 6 out of 7 manufacturing industries during the non-recessionary periods.

Figure 9 in Appendix A provides more context on the sector-level heterogeneity of capacity utilisation in different business cycle phases. While most sectors' capacity utilisation co-move across the business cycle, sectors SS4 (Processing of non-metallic minerals) and SS7 (Energy and extraction) were hit more severely during the European debt crisis. Sector SS2 (Textiles, clothing, and hide leather and footwear products) was hit the most during the Covid crisis as it suffered from multiple lockdowns, decreased consumption, and excess stock (Leal Filho et al., 2023; Arania, Putri and Saifuddin, 2022).

Lastly, we ask whether our uncertainty measure behaves as expected with the business cycle. We average the values of our uncertainty measure over the years and find that our uncertainty measure peaks during the recessions, similar to the original overall uncertainty index in Mohades, Piccillo and Treibich (2024):

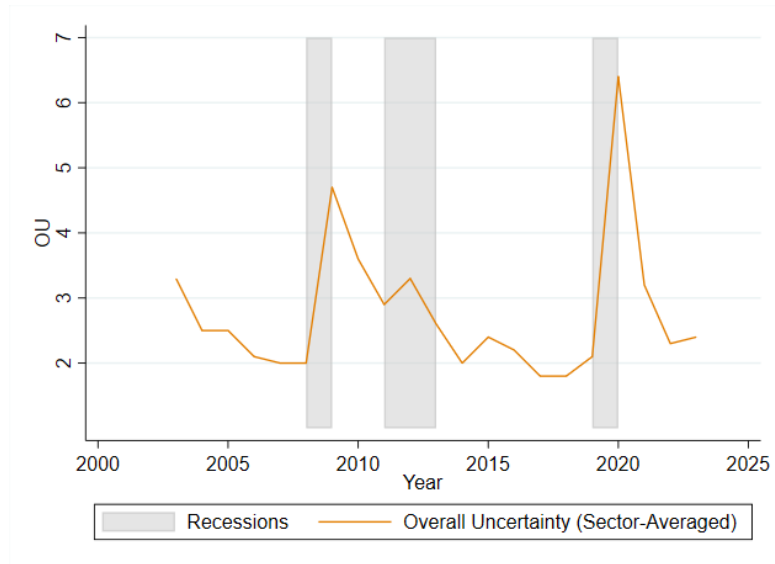


Figure 2.8: Annual Average of Italian Manufacturing Firms' Overall Uncertainty from 2002-2023. Data source: INVIND, [Bank of Italy \(2024\)](#)

The local peaks during the Great Recession, European debt crisis and the COVID outbreak reflect how this uncertainty measure behaves countercyclically with respect to the business cycle. Furthermore, this measure correlates with a magnitude of 0.52 to EPU by [Baker, Bloom and Davis \(2016\)](#).

We are now ready to discuss the results and implement the variables introduced and explored in this section in the empirical model presented in Section 2.3.

2.5 Results

This section presents the results of our estimations based on the equations presented in Section 2.3 and the implementation of the variables introduced in Section 3.3. Initially, our framework tests whether firms' investments and probability of experiencing a spike are sensitive to being at the growth window and the distance to the target. Based on the predictions of our conceptual framework in Section 2.2, we expect both the overcapacity dummy and capacity utilisation gap to display positive coefficients to investment rates and spikes. Table 2.5.1 displays our results from both regressions with the investment rate and the probability of spike as dependent variables:

	Dependent Variable			
	Inv Rate ($I/\bar{s}$)		Spike ($\mathcal{S}$)	
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.003 (0.003)		0.005* (0.002)	
$\mathcal{O}_{i,t}$		0.241** (0.093)		0.121 (0.069)
$\pi_{i,t-1}$	0.311** (0.103)	0.296** (0.103)	0.269*** (0.079)	0.275*** (0.079)
$L_{i,t-1}^G$	0.634 (0.383)	0.609 (0.383)	0.511 (0.266)	0.510 (0.266)
Constant	2.960*** (0.082)	2.832*** (0.094)		
Estimation Method	OLS	OLS	Clogit	Clogit
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Firms	3,575	3,575	896	896
Observations	23,954	23,954	8,815	8,815

Table 2.5.1: Regression Results For Equation 2.1

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

We document that observing a higher capacity utilisation distance from the target and presence at the growth window increases the probability of observing a spike or higher investment rates, in line with the “fork in the road” theory (Brown and Mawson, 2013; Coad et al., 2021). Therefore, firms at higher capacity are more likely to experience an investment spike and expand. Past profitability also positively contributes to both investment rates and spikes, which shows that Italian firms have been using internal funds to finance their investments, similar to the finding of Ughetto (2008).

In a second step corresponding to the theoretical framework explained in Section 2.2, we are interested in studying the impact of financial constraints and uncertainty on investment decisions, the probability of spike, and their interaction with the firms' capacity utilisation. While there is less discrepancy on the impact of financial constraints on firms' investments, there is substantial disagreement regarding the impact of uncertainty on investments, especially large investments (spikes). Previously, the literature has shown that the presence of irreversibility and higher uncertainty reduces the responsiveness of investment to demand shocks (Bloom, Bond and Van Reenen, 2007; Bloom,

2009; Bond and Lombardi, 2006). Hence, firms become more cautious when investing or adjusting inputs to minimise dead-weight losses in uncertain times. On the other hand, low uncertainty can trigger firms' risk-taking actions to benefit from the investing premium (Bo and Lensin, 2005), and encourage firms to take replacement investments (Mauer and Ott, 1995). The results corresponding to Equation 2.4 are provided below:

Dependent Variable: Investment Rate ($I/\bar{s}$)				
	(1)	(2)	(3)	(4)
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.005 (0.004)	0.006 (0.004)		
$\mathcal{O}_{i,t}$			0.415*** (0.116)	0.413*** (0.115)
$FC_{i,t}$	-0.583** (0.211)	-0.524* (0.205)	-0.606* (0.262)	-0.491 (0.257)
$U_{s,t}$	0.017* (0.007)	0.017* (0.007)	0.033*** (0.010)	0.030** (0.009)
$\mathcal{O}_{i,t} \times FC_{i,t}$			0.082 (0.327)	-0.088 (0.320)
$FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi)$	0.010 (0.010)	0.004 (0.010)		
$\mathcal{O}_{i,t} \times U_{s,t}$			-0.031* (0.013)	-0.026* (0.013)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times U_{s,t}$	-0.001 (0.001)	-0.001 (0.001)		
$U_{s,t} \times FC_{i,t}$	-0.022 (0.024)	-0.027 (0.023)	-0.029 (0.024)	-0.024 (0.023)
$\pi_{i,t-1}$	0.303** (0.103)	0.311** (0.103)	0.291** (0.103)	0.300** (0.102)
$L_{i,t-1}^G$	0.643 (0.380)	0.664 (0.380)	0.622 (0.379)	0.642 (0.379)
$U_{s,t-1}$	0.034*** (0.007)	0.033*** (0.007)	0.034*** (0.007)	0.033*** (0.007)
Constant	2.764*** (0.097)	2.755*** (0.096)	2.535*** (0.116)	2.530*** (0.114)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Hausman Test	$\checkmark$		$\checkmark$	
Estimation Method	OLS	OLS	OLS	OLS
Time FE	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Firm FE	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Firms	3,575	3,575	3,575	3,575
Observations	23,951	23,951	23,951	23,951

Table 2.5.2: Regression Results For Equation 2.2

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval,
* Significant with 90% Confidence Interval

Dependent Variable: Probability of Spike ($\mathcal{S}$)				
	(1)	(2)	(3)	(4)
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.007* (0.003)	0.007* (0.003)		
$\mathcal{O}_{i,t}$			0.212* (0.086)	0.223** (0.084)
$FC_{i,t}$	-0.478** (0.165)	-0.357* (0.165)	-0.410 (0.218)	-0.183 (0.201)
$U_{s,t}$	0.012* (0.005)	0.012** (0.005)	0.021** (0.008)	0.019** (0.007)
$\mathcal{O}_{i,t} \times FC_{i,t}$			-0.060 (0.276)	-0.385 (0.263)
$FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi)$	-0.003 (0.009)	-0.004 (0.008)		
$\mathcal{O}_{i,t} \times U_{s,t}$			-0.016 (0.009)	-0.014 (0.009)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times U_{s,t}$	-0.001 (0.001)	-0.001 (0.001)		
$U_{s,t} \times FC_{i,t}$	0.003 (0.014)	0.001 (0.016)	-0.005 (0.015)	0.006 (0.016)
$\pi_{i,t-1}$	0.260** (0.079)	0.268*** (0.079)	0.265*** (0.079)	0.272*** (0.079)
$L_{i,t-1}^G$	0.515 (0.266)	0.529* (0.266)	0.501 (0.266)	0.525* (0.266)
$OU_{s,t-1}$	0.011** (0.004)	0.011** (0.004)	0.011** (0.004)	0.011** (0.004)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Hausman Test		$\checkmark$		$\checkmark$
Estimation Method	CLogit	CLogit	CLogit	CLogit
Time FE	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Firm FE	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Firms	896	896	896	896
Observations	8,813	8,813	8,813	8,813

Table 2.5.3: Regression Results For Equation 2.2

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval,

* Significant with 90% Confidence Interval

In Tables 2.5.2 & 2.5.3, we show that being present at the growth window (shown by $\mathcal{O}$) increases both investment rates and the probability of investment spikes. The distance to target positively impacts the probability of a spike in investments, even after adding financial constraints and uncertainty and controlling for past uncertainty, profitability and employee growth. Importantly, these results display a different pattern for the Italian firms within our sample compared to 1998-2003, studied by Bottazzi, Secchi and Tamagni (2008), by displaying that higher incentives for investment among Italian firms with higher motives for growth: presence at the growth window increases both investment rates and the probability of observing a spike, and Italian firms are not an exception despite the difficulties of investment.

Interestingly, unlike other empirical exercises in the literature, we find that uncertainty, on average, increases the investment rates across firms, revealed by the positive coefficient of $U_{s,t}$. However, being present at the growth window cancels this positive effect out, as shown by the sum of the coefficient of $U_{s,t}$ and the coefficient of $\mathcal{O}_{i,t} \times U_{s,t}$. Our results on the positive impact of uncertainty on firms' investments are of particular interest, given our sample. In environments where firms, on average, operate below capacity, the negative impact of uncertainty on investments can be significantly smaller (Dangl, 1999). Furthermore, we interpret this heterogeneous impact of uncertainty on firms at and outside the growth window as inefficiency in allocating inputs since firms outside the growth window keep investing when uncertainty is high, similar to Bansal et al. (2019). These effects are potentially led by different values of the elasticity of substitution in production vs in consumption of high and low potential growth firms' products (Dellas and Fernandes, 2006). Furthermore, demand uncertainty is previously found to have a lower impact on firms' investments than other sources of uncertainty (Fuss and Vermeulen, 2008).

Furthermore, we demonstrate a robust negative impact of financial constraints on both investment rates and the likelihood of spikes, similar to Kamber, Smith and Thoenissen (2015), among others. Alfaro, Bloom and Lin (2022) demonstrate how more financially constrained firms cut their investments more than less constrained ones following a shock. Furthermore, they demonstrate that financial constraints can explain a large proportion of the drop in output after a shock. Lastly, our results also show that only 896 out of 3,575 firms (25%) in our regressions have experienced an investment spike at least once, which is in line with intuitions from Grazzi, Jacoby and Treibich (2016) that only a few firms observe spikes, especially in economies with tight borrowing conditions.

To better understand the mechanism behind our model, we limit the sample to observations with below-median uncertainty and no financial constraints. The rationale is based on the idea that firms without tight external conditions represent an important group for whom the role of being present at the growth window is robustly positive. We find evidence in this small subsample that presence at the growth window increases the probability of observing a spike. While we expected a stronger coefficient of both gap and overcapacity dummy for both investment variables, a small sample limits the insights we can take from these analyses as the standard deviations grow. These results are provided in Table B.1.

If both gap and overcapacity were to affect investments strongly, there would be a strong interaction between the two so that firms at the growth window closer to full capacity and hence farther from the target would be more likely to invest. We call this the intensity of the effect of presence at the growth window and the distance to the target. To study this intensity effect, we add the interaction term between gap ($CU_{i,t} - CU_{s,\tau}^\psi$) and overcapacity dummy ($\mathcal{O}_{i,t}$) to our regressions, with results presented in Tables B.2 to B.4. Our regressions in this section cannot reject the hypothesis by incorporating

both the gap to capacity utilisation target and growth window dummy: being farther ahead of the capacity utilisation target, on average, increases the probability of a spike occurrence among firms. We do not find evidence of a significant intensity difference between these groups in their reaction to the gap to capacity target. However, our baseline results remain relatively robust in the presence of this interaction term. Albeit the results of our analyses regarding no heterogeneity may seem against the odds, we believe that in a sample including small and medium enterprises (SMEs) with high potential growth, there might be more layers of heterogeneity regarding the intensity of being present at the growth window as new firms usually operate at very high capacity and invest and grow.

2.6 Robustness Checks

This section reviews multiple checks we perform to ensure the robustness of our baseline specification explored and discussed in Sections 2.3 to 2.5. We perform checks to address issues related to the survey timing (within-year expectation vs realised). We introduce alternative specifications for investment rate, capacity utilisation gap and uncertainty measurements to test the validity of our approach. We then study whether our results are driven by the frontier firms as discussed by [Añón Higón et al. \(2022\)](#). Lastly, we will explain why we believe the results presented in our analyses are not driven by the entries and exits of firms.

2.6.1 Timing of the Survey: From Forward to Backward Looking

Bank of Italy distributes and collects the INVIND survey annually between January and May. Hence, our observations' values are the expected sales and investments within the year of the survey. Due to such timing, one may be concerned that the within-year expectations might be incorrect or subject to the business cycle beliefs. We assess this issue by incorporating the realised values that firms report in the year after. Our main results remain robust to this new specification. Given the coefficients reported in Tables C.3 to C.2 and the high pairwise correlation between the baseline variable and the lagged realised value from the following year, we can conclude that our results are robust to within-year business cycle developments that might affect managers' expectations. The robustness of these findings can be attributed to firms' awareness of their annual orders and investment decisions, particularly given that the sample is predominantly composed of medium to large firms.

2.6.2 Alternative Investment Rate Specifications

Throughout our analyses, we normalise the raw investment values by the average of current and 1-year lags of sale values. To avoid the issue of persistent sales cycles impacting these values, we also run our regressions with the investment rate normalised by the current average and the past two years, hence three years in total ($t, t + 1, t + 2$). The regression results are presented in Tables C.6 to C.5. In Tables C.9 to C.8, we provide the results of our regressions with the investments normalised by labour (in year t) instead of sales. In both of these alternative analyses, we can find evidence of the robustness of our proposed framework in Sections 2.2 & 2.3. Therefore, we conclude that the unique method of normalisation of investments is not solely driving the previously presented results, and our approach is robust in terms of the definition of investment rate.

2.6.3 Firm-Specific Capacity Utilisation Target

Our baseline specification considers three dimensions of variation in the gap variable ($CU_{i,t} - CU_{s,\tau}^\psi$): variation of the realised capacity utilisation across firms, time and across sectors and regimes (recessionary and non-recessionary). However, the low variation of the target (11 industries and two regimes) might weaken our findings. We test our model using a firm-specific target to determine how robust our baseline specification performs against this potential bias. In Tables C.12 to C.11, we provide the results using a firm-specific target instead of a sector-specific one. This construction can be formally stated as a firm-specific capacity target ($CU_{i,t} - CU_{i,\tau}^\psi$) and its corresponding new overcapacity dummy $\mathcal{O}_{i,t}^{\text{alt}}$ accordingly, where $\mathcal{O}_{i,t}^{\text{alt}} = 1$ if $CU_{i,t} > CU_{i,\tau}^\psi$. Our analysis confirms the robustness of our baseline specification: regardless of the definition of the capacity target, deviation from the target or presence at the growth window motivates investments in plants and machinery. Therefore, we conclude that low variation due to aggregation at the sector level does not sabotage analysis.

2.6.4 Expectation-Based Uncertainty Measures

In this section, we consider multiple measures of uncertainty driven by firms' sales records and expectations. In Section 2.5, we provided the results with uncertainty proxied by "sectoral volatility" ($U_{s,t}$) driven by sales volatility. However, our access to survey data allows us to use firms' expectations to calculate subjective uncertainty. We are not the first to stress the importance of firms' expectations in their investment behaviour via uncertainty. Guiso and Parigi (1999) found that the cross-sectional subjective uncertainty of managers reduces investment.¹⁶ Similarly, through measuring subjective uncertainty from survey data of the U.S manufacturing firms, Altig et al. (2022); Bloom et al. (2022) found a robust negative relationship between uncertainty and investment, as well as uncertainty and firms' sales growth and hirings¹⁷. We test our results based on the firms' subjective uncertainty and forecast error. To compute subjective uncertainty, we use intuition from the method used in Lamorgese et al. (2021); Bloom et al. (2022); Bachmann et al. (2021), and use the size of the spectrum of their expected sales growth in $t+1$ as a proxy for their uncertainty at time t . Namely, we use the difference between the maximum and minimum expected sales growth for the next period as:

$$\text{Subjective Uncertainty}_{i,t} \equiv SU_{i,t} = E_t^{\text{max}} \text{Sales Growth}_{i,t+1} - E_t^{\text{min}} \text{Sales Growth}_{i,t+1}$$

We also use the forecast error as a proxy for uncertainty similar to Arslan et al. (2015). We have data regarding point expectations of firms from their future sales growth. We then calculate the forecast error as follows:

$$\text{Forecast Error}_{i,t} \equiv FCE_{i,t} = \frac{E_t \text{Sales}_{i,t+1} - \text{Sales}_{i,t+1}}{\text{Sales}_{i,t+1}}$$

The results of these analyses can be found in Tables C.13 to C.16. We show that our main takeaways are robust to uncertainty measurement. However, our model performs relatively poorly

¹⁶Their results show substantial heterogeneity among firms. Also, less access to reversible investments strengthens the impact of higher uncertainty on firms' investments.

¹⁷They also find a strong positive correlation between firms subjective uncertainty and their sales forecast error. Interestingly, flexible inputs show a positive relationship to uncertainty, demonstrating that businesses switch from less flexible to more flexible factor inputs at higher levels of uncertainty.

when expectation-based uncertainty measures are used since the response rates to the expected values are low compared to within-year values, and hence, the sample is smaller.

2.6.5 Frontier Firms Effect

We also suspect the possibility that our results are driven by firms that are very large in our sample. The literature on firm dynamics in Italy has previously found considerable heterogeneity in growth and investment among frontier (leader) and other firms (Pozzi and Schivardi, 2012; Carpenter and Rondi, 2000). We construct a dummy based on the median number of employees to distinguish the very large firms that we believe coincide with the leaders, where $\text{Large} = 1 \text{ if } EMP > \text{Percentile}(50)$. These results are provided in Tables C.17 to C.19 and demonstrate various crucial findings. First, large firms tend to invest more after controlling for firm and time-fixed effects. Second, larger firms that are financially constrained tend to invest more. However, our finding that firms at the growth window are vulnerable to uncertainty persists: even though we rule out the variation of investment rate and probability of spike due to being very large, the inefficiency due to uncertainty remains predominant. Firms with low potential growth invest more, and those with high potential and higher efficiency do not, even after controlling for frontier firms' effect.

2.6.6 Can Entry and Exit drive our results?

One might raise the concern that the macroeconomic dynamics displayed in Figure 2.2 and Figure 2.7(b) may be accompanied and be resulted by firms' exit and entries. On average, capacity utilisation surges once there are many entrants in the market. Firms that exit usually display lower utilisation rates before leaving due to lower activity in the last years of business. Therefore, the persistently low values of capacity utilisation after the Great Recession might have resulted from many firms exiting and few entering. While the Great Recession has impacted the market dynamism in Italy to a large extent, entries and exits have been impacted similarly downwards by a similar magnitude. Meanwhile, the change in market dynamism that could have driven the capacity utilisation downwards would be an increasing difference between entries and exits, not lower market dynamism via both lower entries and exits.

Figure 10 in Appendix A shows that the post-crisis market environment in the Italian economy has failed to show dynamism, as both entry and exit rates have dropped excessively compared to pre-recession (2007). If the firms' entry index was on the same cycle as capacity utilisation and its deviation from the target presented in our paper, one could claim that the dynamics between 2014 and 2020 were driven by young firms that operate at high capacity. However, we showed in this figure that this is not the case, and the entries and exits have remained irrelevant in explaining the capacity utilisation changes after the European debt crisis.

2.7 Discussion & Conclusion

Throughout the paper, we identified a new mechanism to explain the relationship between capacity utilisation and investments, considering the role of uncertainty and financial constraints. While firms at the growth window are willing to invest, financial constraints and uncertainty have lowered high-growth potential firms' investments through low borrowing possibilities. Notably, firms outside the growth window have increased their investments with higher uncertainty, hinting at the inefficiency-causing

role of uncertainty. Firms with high potential growth do not invest more during high uncertainty episodes, while firms with low potential growth invest and expand.

Our findings carry high micro and macroeconomic importance. Firstly, we explained low firm-level investments in the Italian economy in recent decades. While we observe a declining pattern in firms' capacity use and investments until 2014, a slight recovery towards the values of the pre-Great Recession can be observed in both variables after 2014. This result is in line with the previous findings of the literature, which provided multiple explanations for this recovery. Not only did the crisis episodes become less frequent between 2014 and 2020, but the policy to reduce bureaucratic and economic uncertainty has also worked toward higher confidence in the business environment.

Future research must try to understand the mechanism associated with expansion, uncertainty and capacity utilisation under different borrowing constraint regimes. This may be achieved by setting up a heterogeneous agent model where firms' reaction to observations of capacity utilisation can depend on various features of the firm and the state of the economy regarding investments and borrowing. Furthermore, it would be important to study our questions within samples, including SMEs, to understand the impact of capacity utilisation trigger points on firms at the growth window more comprehensively.

References

- Abel, Andrew, and Janice Eberly.** 1994. "A Unified Model of Investment under Uncertainty." American Economic Review, 84(5): 1369–84.
- Abel, Andrew B.** 1981. "A dynamic model of investment and capacity utilization." The Quarterly Journal of Economics, 96(3): 379–403.
- Abel, Andrew B., Avinash K. Dixit, Janice C. Eberly, and Robert S. Pindyck.** 1996. "Options, the Value of Capital, and Investment*." The Quarterly Journal of Economics, 111(3): 753–777.
- Alfaro, Iván, Nicholas Bloom, and Xiaoji Lin.** 2022. "The Finance Uncertainty Multiplier." Working Paper.
- Almeida, Heitor, and Murillo Campello.** 2001. "Financial constraints and investment-cash flow sensitivities: New research directions."
- Altı, Aydoğan, and Paul C Tetlock.** 2014. "Biased beliefs, asset prices, and investment: A structural approach." The Journal of Finance, 69(1): 325–361.
- Altig, David, Jose Maria Barrero, Nicholas Bloom, Steven J Davis, Brent Meyer, and Nicholas Parker.** 2022. "Surveying business uncertainty." Journal of Econometrics, 231(1): 282–303.
- Amiti, Mary, and David E Weinstein.** 2018. "How much do idiosyncratic bank shocks affect investment? Evidence from matched bank-firm loan data." Journal of Political Economy, 126(2): 525–587.
- Angrist, Joshua D, and Jörn-Steffen Pischke.** 2009. Mostly harmless econometrics: An empiricist's companion. Princeton university press.
- Añón Higón, Dolores, Juan A Máñez, María E Rochina-Barrachina, Amparo Sanchis, and Juan A Sanchis.** 2022. "Firms' distance to the European productivity frontier." Eurasian Business Review, 12(2): 197–228.
- Anzuini, Alessio, Luca Rossi, and Pietro Tommasino.** 2020. "Fiscal policy uncertainty and the business cycle: Time series evidence from Italy." Journal of Macroeconomics, 65: 103238.
- Arania, Febristi, Irley Maura Putri, and Muchammad Saifuddin.** 2022. "The impact of Covid-19 on textile and fashion industries: the economic perspective." Journal of Marketing Innovation (JMI), 2(1).
- Arata, Yoshiyuki.** 2019. "Firm growth and Laplace distribution: The importance of large jumps." Journal of Economic Dynamics and Control, 103: 63–82.
- Arslan, Yavuz, Aslıhan Atabek, Timur Hulagu, and Saygın Şahinöz.** 2015. "Expectation errors, uncertainty, and economic activity." Oxford Economic Papers, 67(3): 634–660.
- Artica, Rodrigo Pérez.** 2023. "The impact of excess capacity on the investment falloff." Panoeconomicus, 70(3): 403–428.

- Asturias, Jose, Sewon Hur, Timothy J Kehoe, and Kim J Ruhl.** 2023. “Firm entry and exit and aggregate growth.” American Economic Journal: Macroeconomics, 15(1): 48–105.
- Axsäter, Sven, and Jan Olhager.** 1985. “The impact of capacity investments on work-in-process and inventories.” Engineering Costs and Production Economics, 9(1-3): 119–124.
- Bachmann, Rüdiger, Kai Carstensen, Stefan Lautenbacher, and Martin Schneider.** 2021. “Uncertainty and Change: Survey Evidence of Firms’ Subjective Beliefs.” CESifo Working Paper No. 9394.
- Bachmann, Rüdiger, Steffen Elstner, and Atanas Hristov.** 2017. “Surprise, surprise – Measuring firm-level investment innovations.” Journal of Economic Dynamics and Control, 83: 107–148.
- Baker, Scott R, Nicholas Bloom, and Steven J Davis.** 2016. “Measuring economic policy uncertainty.” The quarterly journal of economics, 131(4): 1593–1636.
- Balduzzi, Pierluigi, Emanuele Brancati, Marco Brianti, and Fabio Schiantarelli.** 2020. “Populism, political risk and the economy: Lessons from Italy.”
- Baley, Isaac, and Andrés Blanco.** 2021. “Aggregate dynamics in lumpy economies.” Econometrica, 89(3): 1235–1264.
- Bank of Italy.** 2024. “Survey on Industrial and Service Firms, 2002-2023.” Data retrieved from Bank of Italy, Source: Bank of Italy.
- Bansal, Ravi, Mariano Max Croce, Wenxi Liao, and Samuel Rosen.** 2019. “Uncertainty-induced reallocations and growth.” National Bureau of Economic Research.
- Baum, Christopher F, Mustafa Caglayan, and Oleksandr Talavera.** 2008. “Uncertainty determinants of firm investment.” Economics Letters, 98(3): 282–287.
- Belo, Frederico, Xiaoji Lin, and Santiago Bzdresch.** 2014. “Labor Hiring, Investment, and Stock Return Predictability in the Cross Section.” Journal of Political Economy, 122(1): 129–177.
- Berndt, Ernst R, and Melvyn A Fuss.** 1982. “Productivity measurement using capital asset valuation to adjust for variations in utilization.” National Bureau of Economic Research.
- Bloise, Francesco, Irene Brunetti, and Valeria Cirillo.** 2021. “Firm strategies and distributional dynamics: labour share in Italian medium-large firms.” Economia Politica, 1–33.
- Bloom, Nicholas.** 2009. “The Impact of Uncertainty Shocks.” Econometrica, 77(3): 623–685.
- Bloom, Nicholas, Steven J Davis, Lucia S Foster, Scott W Ohlmacher, and Itay Saporta-Eksten.** 2022. “Investment and Subjective Uncertainty.” National Bureau of Economic Research.
- Bloom, Nick, Stephen Bond, and John Van Reenen.** 2007. “Uncertainty and Investment Dynamics.” The Review of Economic Studies, 74(2): 391–415.
- Blundell, Richard, Stephen Bond, Michael Devereux, and Fabio Schiantarelli.** 1992. “Investment and Tobin’s Q: Evidence from company panel data.” Journal of econometrics, 51(1-2): 233–257.

- Bo, Hong, and Robert Lensin.** 2005. “Is the investment–uncertainty relationship nonlinear? An empirical analysis for the Netherlands.” Economica, 72(286): 307–331.
- Bokpin, Godfred A, and Joseph M Onumah.** 2009. “An empirical analysis of the determinants of corporate investment decisions: Evidence from emerging market firms.” International Research Journal of Finance and Economics, 33(33): 134–141.
- Bolton, Patrick, Neng Wang, and Jinqiang Yang.** 2019. “Investment under uncertainty with financial constraints.” Journal of Economic Theory, 184: 104912.
- Bond, Stephen R, and Domenico Lombardi.** 2006. “To buy or not to buy? Uncertainty, irreversibility, and heterogeneous investment dynamics in Italian company data.” IMF Staff Papers, 53: 375–400.
- Bond, Stephen R, Giacomo Rodano, and Nicolas Andre Benigno Serrano-Velarde.** 2015. “Investment dynamics in Italy: Financing constraints, demand and uncertainty.” Bank of Italy Occasional Paper, , (283).
- Bontempi, Maria Elena, Roberto Golinelli, and Giuseppe Parigi.** 2010. “Why demand uncertainty curbs investment: Evidence from a panel of Italian manufacturing firms.” Journal of Macroeconomics, 32(1): 218–238.
- Boot, Arnoud, and Vladimir Vladimirov.** 2019. “(Non-) precautionary cash hoarding and the evolution of growth firms.” Management Science, 65(11): 5290–5307.
- Bottazzi, Giulio, Angelo Secchi, and Federico Tamagni.** 2008. “Productivity, profitability and financial performance.” Industrial and Corporate Change, 17(4): 711–751.
- Bottazzi, Giulio, Angelo Secchi, and Federico Tamagni.** 2014. “Financial constraints and firm dynamics.” Small Business Economics, 42: 99–116.
- Brown, Ross, and Suzanne Mawson.** 2013. “Trigger points and high-growth firms: A conceptualisation and review of public policy implications.” Journal of Small Business and Enterprise Development, 20(2): 279–295.
- Bruno, Giuseppe, Leandro D’Aurizio, and Raffaele Tartaglia-Polcini.** 2014. “Remote processing of business microdata at the Bank of Italy.” Statistical Methods and Applications from a Historical Perspective: Selected Issues, 239–249.
- Bugamelli, Matteo, Francesca Lotti, Monica Amici, Emanuela Ciapanna, Fabrizio Colonna, Francesco D’Amuri, Silvia Giacomelli, Andrea Linarello, Francesco Manaresi, Giuliana Palumbo, et al.** 2018. “Productivity growth in Italy: a tale of a slow-motion change.” Bank of Italy Occasional Paper, , (422).
- Bull, Martin J.** 2020. “In the eye of the storm: the Italian economy and the eurozone crisis.” In Italy Transformed. 13–28. Routledge.
- Busetti, Fabio, Claire Giordano, and Giordano Zevi.** 2015. “Main Drivers of the Recent Decline in Italy’s Non-Construction Investment.” Bank of Italy Occasional Paper, , (276).

- Butters, R. Andrew.** 2020. “Demand volatility, adjustment costs, and productivity: An examination of capacity utilization in hotels and airlines.” American Economic Journal: Microeconomics, 12(4): 1–44.
- Calligaris, Sara, Massimo Del Gatto, Fadi Hassan, Gianmarco IP Ottaviano, and Fabiano Schivardi.** 2018. “The productivity puzzle and misallocation: an Italian perspective.” Economic Policy, 33(96): 635–684.
- Campello, Murillo, John R. Graham, and Campbell R. Harvey.** 2010. “The real effects of financial constraints: Evidence from a financial crisis.” Journal of Financial Economics, 97(3): 470–487. The 2007-8 financial crisis: Lessons from corporate finance.
- Carpenter, Robert E, and Laura Rondi.** 2000. “Italian corporate governance, investment, and finance.” Empirica, 27(4): 365–388.
- Ciapanna, Emanuela, Sauro Mocetti, Alessandro Notarpietro, et al.** 2020. The effects of structural reforms: Evidence from Italy. Banca d’Italia Eurosistema.
- Coad, Alex, and Max Planck.** 2012. “Firms as bundles of discrete resources—towards an explanation of the exponential distribution of firm growth rates.” Eastern Economic Journal, 38: 189–209.
- Coad, Alex, and Stjepan Srhoj.** 2020. “Catching Gazelles with a Lasso: Big data techniques for the prediction of high-growth firms.” Small Business Economics, 55(3): 541–565.
- Coad, Alex, Clemens Domnick, Florian Flachenecker, Peter Harasztosi, Mario Lorenzo Janiri, Rozalia Pal, and Mercedes Teruel.** 2021. “Capacity constraints as a trigger for high growth.” Small Business Economics, 1–31.
- Corrado, Carol, and Joe Matthey.** 1997. “Capacity Utilization.” Journal of Economic Perspectives, 11(1): 151–167.
- Cucculelli, Marco, and Valentina Peruzzi.** 2020. “Post-crisis firm survival, business model changes, and learning: evidence from the Italian manufacturing industry.” Small Business Economics, 54: 459–474.
- Dagdeviren, Hulya.** 2016. “Structural constraints and excess capacity: an international comparison of manufacturing firms.” Development Policy Review, 34(5): 623–641.
- Dangl, Thomas.** 1999. “Investment and capacity choice under uncertain demand.” European Journal of Operational Research, 117(3): 415–428.
- Daruich, Diego, Sabrina Di Addario, and Raffaele Saggio.** 2023. “The effects of partial employment protection reforms: Evidence from Italy.” Review of Economic Studies, 90(6): 2880–2942.
- deB. Harris, Frederick H.** 1988. “Capital Intensity and the Firm’s Cost of Capital.” The Review of Economics and Statistics, 70(4): 587–594.
- Dellas, Harris, and Ana Fernandes.** 2006. “The Allocation of Resources under Uncertainty.” Available at SSRN 952688.

- Di Domenico, Lorenzo.** 2023. “Multiplicity and not necessarily heterogeneity: implications for the long-run degree of capacity utilization.” Journal of Economic Interaction and Coordination, 18(4): 835–877.
- Dixit, Avinash K, and Robert S Pindyck.** 1994. Investment under uncertainty. Princeton university press.
- Doms, Mark, and Timothy Dunne.** 1998. “Capital adjustment patterns in manufacturing plants.” Review of economic dynamics, 1(2): 409–429.
- Dosi, Giovanni, Marco Grazzi, Chiara Tomasi, and Alessandro Zeli.** 2012. “Turbulence underneath the big calm? The micro-evidence behind Italian productivity dynamics.” Small Business Economics, 39: 1043–1067.
- Dotsey, Michael, and Thomas Stark.** 2005. “The relationship between capacity utilization and inflation.” Federal Reserve Bank of Philadelphia Business Review, 2: 8–17.
- Fagnart, Jean-François, Omar Licandro, and Franck Portier.** 1999. “Firm heterogeneity, capacity utilization, and the business cycle.” Review of Economic Dynamics, 2(2): 433–455.
- Fazzari, Steven M, and Bruce C Petersen.** 1988. “Financing Constraints and Corporate Investment.” Brookings Papers on Economic Activity, 1988(1): 141–195.
- Fernald, John, and Robert Inklaar.** 2020. “Does disappointing european productivity growth reflect a slowing trend? weighing the evidence and assessing the future.” International Productivity Monitor, 2020(38): 104–135.
- Ferreira, Miguel H, Timo Haber, and Christian Rorig.** 2023. “Financial Constraints and Firm Size: Micro-Evidence and Aggregate Implications.”
- Fuss, Catherine, and Philip Vermeulen.** 2008. “Firms’ investment decisions in response to demand and price uncertainty.” Applied Economics, 40(18): 2337–2351.
- Giachetti, Claudio, and Gianluca Marchi.** 2017. “Successive changes in leadership in the worldwide mobile phone industry: The role of windows of opportunity and firms’ competitive action.” Research Policy, 46(2): 352–364.
- Gourio, Francois, and Anil K Kashyap.** 2007. “Investment spikes: New facts and a general equilibrium exploration.” Journal of Monetary Economics, 54: 1–22.
- Grazzi, Marco, Nadia Jacoby, and Tania Treibich.** 2016. “Dynamics of investment and firm performance: Comparative evidence from manufacturing industries.” Empirical Economics, 51: 125–179.
- Greene, W.H.** 2018. Econometric Analysis. Econometric Analysis, Pearson.
- Greenwood, Jeremy, Zvi Hercowitz, and Gregory W. Huffman.** 1988. “Investment, Capacity Utilization, and the Real Business Cycle.” The American Economic Review, 78(3): 402–417.
- Grullon, Gustavo, and David L Ikenberry.** 2021. “Excess Capacity, Marginal Q and Corporate Investment.” Marginal q and Corporate Investment (July 1, 2021).

- Gufler, Ivan, Michael Donadelli, Marta Castellini, et al.** 2020. “On the Impact of COVID-19-Related Uncertainty.” INNOVATION IN BUSINESS, ECONOMICS & FINANCE, 219–232.
- Guiso, Luigi, and Giuseppe Parigi.** 1999. “Investment and Demand Uncertainty*.” The Quarterly Journal of Economics, 114(1): 185–227.
- Hosmer Jr, David W, Stanley Lemeshow, and Rodney X Sturdivant.** 2013. Applied logistic regression. John Wiley & Sons.
- Howes, Cooper.** 2023. “Financial constraints, sectoral heterogeneity, and the cyclicalities of investment.” Review of Economics and Statistics, 1–45.
- Hulten, Charles R.** 1986. “Productivity change, capacity utilization, and the sources of efficiency growth.” Journal of Econometrics, 33(1): 31–50.
- Huynh, Japan.** 2024. “Banking uncertainty and corporate financial constraints.” International Journal of Finance & Economics.
- Im, Hyun Joong, Colin Mayer, and Oren Sussman.** 2020. “Heterogeneity in investment spike financing.”
- ISTAT.** 2018. “Report on competitiveness.” Italian National Institute of Statistics (ISTAT).
- Kamber, Günes, Christie Smith, and Christoph Thoenissen.** 2015. “Financial frictions and the role of investment-specific technology shocks in the business cycle.” Economic Modelling, 51: 571–582.
- Koenig, Evan F.** 1994. “Capacity utilization and the evolution of manufacturing output: a closer look at the \"bounce-back effect.\"”
- Krugman, Paul.** 1994. “Defining and measuring productivity.” The Age of diminishing Expectations.
- Lamorgese, Andrea, Andrea Linarello, Megha Patnaik, and Fabiano Schivardi.** 2021. “Management practices and resilience to shocks: Evidence from COVID-19.”
- Leal Filho, Walter, Maria Alzira Pimenta Dinis, Arminda do Paço, Vera Herédia-Colaço, Lucas Veiga Avila, Karen Dennis, David Tyler, Samara Neiva, and Olena Liakh.** 2023. “COVID-19 and sustainability in textile, apparel and fashion use: an assessment of trends.” Textile research journal, 93(3-4): 674–690.
- Lee, Keun, and Franco Malerba.** 2017. “Catch-up cycles and changes in industrial leadership: Windows of opportunity and responses of firms and countries in the evolution of sectoral systems.” Research Policy, 46(2): 338–351.
- Locatelli, Andrea, Libero Monteforte, and Giordano Zevi.** 2016. “Heterogeneous fall in productive capacity in Italian Industry during the 2008-13 double-dip recession.” Bank of Italy occasional paper, , (303).
- Loderer, Claudio, René Stulz, and Urs Waelchli.** 2017. “Firm rigidities and the decline in growth opportunities.” Management Science, 63(9): 3000–3020.

- Mauer, David C, and Steven H Ott.** 1995. “Investment under uncertainty: The case of replacement investment decisions.” Journal of financial and quantitative analysis, 30(4): 581–605.
- Michelacci, Claudio, and Fabiano Schivardi.** 2013. “Does idiosyncratic business risk matter for growth?” Journal of the European Economic Association, 11(2): 343–368.
- Minetti, Raoul, and Susan Chun Zhu.** 2011. “Credit constraints and firm export: Microeconomic evidence from Italy.” Journal of International Economics, 83(2): 109–125.
- Mobley, R Keith.** 2011. Maintenance fundamentals. Elsevier.
- Mohades, Siavash, and Maria Savona.** Forthcoming. “Investments in Data and R&D Activities.” Working Paper.
- Mohades, Siavash, Giulia Piccillo, and Tania Treibich.** 2024. “Unpacking Economic Uncertainty—Measuring the Firm, Sector and Aggregate Components.” CESifo Working Paper.
- Morrison, Catherine J.** 1985. “On the economic interpretation and measurement of optimal capacity utilization with anticipatory expectations.” The Review of Economic Studies, 52(2): 295–309.
- Nelson, Randy A.** 1989. “On the measurement of capacity utilization.” The Journal of Industrial Economics, 273–286.
- Nguyen, Phan Dinh.** 2013. “Determinants of corporate investment decisions: The case of Vietnam.”
- Nikiforos, Michalis.** 2011. “On the desired rate of capacity utilization.” The New School for Social Research, Working Paper, , (16).
- Nikiforos, Michalis.** 2013. “The (normal) rate of capacity utilization at the firm level.” Metroeconomica, 64(3): 513–538.
- OECD.** 2021. “OECD Economic Surveys: Italy.” [Accessed 30-11-2023].
- OECD.** 2024. “Production: Industry: Total Industry Excluding Construction for Italy [ITAPROINDQISMEI].” Accessed on April 13, 2024.
- Penrose, E., and C. Pitelis.** 2009. The Theory of the Growth of the Firm. OUP Oxford.
- Pozzi, Andrea, and Fabiano Schivardi.** 2012. “What determines firm growth? The role of demand and TFP shocks.” RAND Journal of Economics.
- Pozzi, Andrea, and Fabiano Schivardi.** 2016. “Demand or productivity: what determines firm growth?” The RAND Journal of Economics, 47(3): 608–630.
- Ray, Subhash C, John Walden, and Lei Chen.** 2021. “Economic measures of capacity utilization: A nonparametric short-run cost function analysis.” European Journal of Operational Research, 293(1): 375–387.
- Saltari, Enrico, and Giuseppe Travaglini.** 2001. “Financial constraints and investment decisions.” Scottish Journal of Political Economy, 48(3): 330–344.
- Sarkar, Sudipto.** 2009. “A real-option rationale for investing in excess capacity.” Managerial and Decision Economics, 30(2): 119–133.

- Schiantarelli, Fabio.** 1996. “Financial Constraints and Investment: Methodological Issues and International Evidence.” Oxford Review of Economic Policy, 12(2): 70–89.
- Schivardi, Fabiano, Enrico Sette, and Guido Tabellini.** 2022. “Credit misallocation during the European financial crisis.” The Economic Journal, 132(641): 391–423.
- Secchi, Angelo, Federico Tamagni, and Chiara Tomasi.** 2016. “Financial constraints and firm exports: accounting for heterogeneity, self-selection, and endogeneity.” Industrial and Corporate Change, 25(5): 813–827.
- Shapiro, Matthew D, Robert J Gordon, and Lawrence H Summers.** 1989. “Assessing the Federal Reserve’s measures of capacity and utilization.” Brookings Papers on Economic Activity, 1989(1): 181–241.
- Ughetto, Elisa.** 2008. “Does internal finance matter for R&D? New evidence from a panel of Italian firms.” Cambridge Journal of Economics, 32(6): 907–925.
- Zamagni, V.** 2018. The Italian Economy. World economies, Agenda Publishing Limited.
- Zeli, Alessandro, Matilde Bini, and Leopoldo Nascia.** 2022. “A longitudinal analysis of Italian manufacturing companies’ labor productivity in the period 2004–2013.” Industrial and Corporate Change, 31(4): 1004–1030.

A Additional Graphs and Tables

A.1 Capacity Utilisation Heterogeneity at the Sector Level

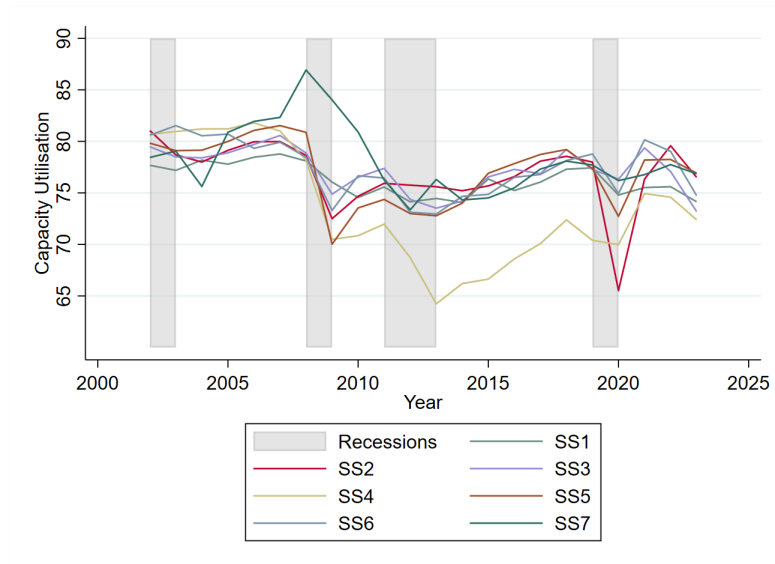


Figure 9: Annual Average of Italian Manufacturing Sectors' Capacity Utilisation from 2002-2023. Data source: INVIND, [Bank of Italy \(2024\)](#)

A.2 Market Dynamics in Italy

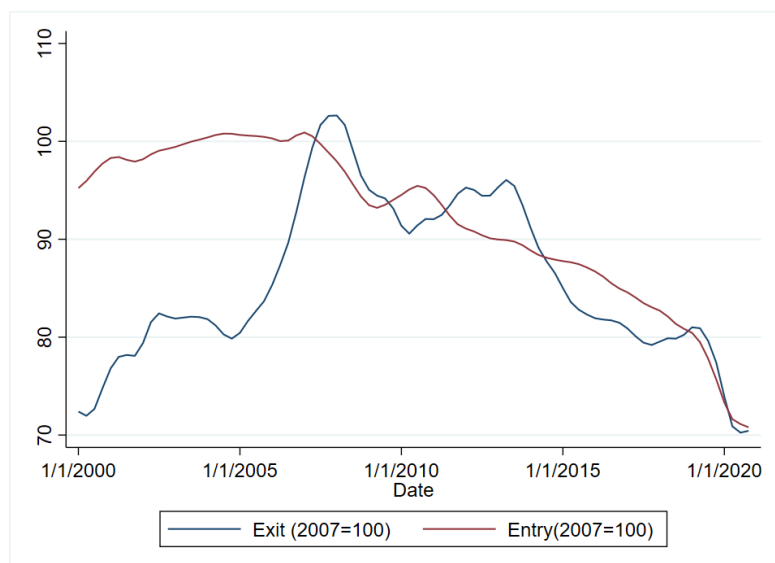


Figure 10: Quarterly Exit and Entry (2007=100) of Italian Firms. Data source: OECD

A.3 Sectoral Division of INVIND Data

Variable	Values	Description	ATECO 2002	ATECO 2007
settor11	SS1	Food industries, beverages, and tobacco products	DA	10, 11, 12
	SS2	Textiles, clothing, and hide, leather, and footwear products	DB, DC	13, 14, 15
	SS3	Coke manufacturing, chemical industry, rubber, and plastics	DF, DG, DH	19, 20, 21, 22
	SS4	Processing of non-metallic minerals	DI	23
	SS5	Metal engineering industry	DJ, DK, DL, DM	24, 25, 26, 27, 28, 29, 30, 33
	SS6	Other manufacturing industries	DD, DE, DN	16, 17, 18, 31, 32
	SS7	Other industries excluding construction	CA, CB, CE	05, 06, 07, 08, 09, 35, 36, 37, 38, 39
	SS8	Wholesale and retail commerce	G	45, 46, 47
	SS9	Hotels and restaurants	H	55, 56
	SS10	Transport and communications	I	49, 50, 51, 52, 53, 58, 59, 60, 61, 62, 63
	SS11	Real estate activities, IT, etc.	K	68, 69, 70, 71, 72, 73, 74, 75, 77, 78, 79, 80, 81, 82

Table A.1: Sector categorisation for the *settor11* variable in the INVIND dataset and its comparability to ATECO by ISTAT.

A.4 Autocorrelation of Uncertainty Measures

	Overall Uncertainty ($U_{i,t}$)	Sectoral Volatility ($U_{s,t}$)	$SU_{i,t}$	$FCE_{i,t}$
L1	0.1736*	0.1600*	0.2252*	0.0068
L2	0.0732*	0.0870*	0.0929*	0.0081
L3	0.0550*	0.2228*	0.0647*	0.0003
L4	0.0625*	0.0882*	0.0679*	0.0022

Table A.2: Autocorrelation Table of Uncertainty Measures

A.5 Summary Statistics for Growth Window and Non-Growth Window

Variable	Growth Window Firms			Non-Growth Window Firms		
	Mean	SD	N	Mean	SD	N
Raw Variables						
$Y_{i,t}^S$	297,904.70	3,822,909.00	28,850	310,292.90	3,815,190.00	23,424
$I_{i,t}$	8,184.02	74,619.18	23,063	8,413.82	87,432.89	15,862
$L_{i,t}$	327.34	874.17	28,850	302.90	1,577.68	23,424
$CU_{i,t}$	86.13	7.31	28,850	64.97	13.80	23,424
Constructed Variables						
$I_{i,t}/\bar{Y}_{i,t-1}^S$	3.57	20.97	23,063	3.75	33.53	15,859
$\mathcal{S}_{i,t}$	0.10	0.30	23,063	0.10	0.30	15,859
$CU_{i,t} - CU_{i,t}^\psi$	9.44	7.36	28,850	-11.63	13.71	23,424
$FC_{i,t}$	0.05	0.21	28,850	0.09	0.28	23,424
$\pi_{i,t}$	0.76	0.43	26,607	0.62	0.49	21,277
$U_{s,t}$	4.45	5.55	27,518	5.07	7.51	22,898
$L_{i,t-1}^G$	0.01	0.11	18,955	-0.01	0.11	15,932

Table A.3: Summary Statistics for Growth Window Firms and Non-Growth Window Firms

A.6 First-stage Regression

Dependent Variable: $FC_{i,t}$	
$FC_{i,t-1}$	0.334***
	0.051
Estimation Method	Clogit
Time FE	✓
Firm FE	✓
Observations	12,144

Table A.4: Regression Results For Equation 2.3

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

B Further Analysis - Regression Tables

Dependent Variable				
	Inv Rate		Spike	
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.002		0.009	
	(0.004)		(0.005)	
$\mathcal{O}_{i,t}$		0.072		0.241*
		(0.109)		(0.114)
$\pi_{i,t-1}$	0.272*	0.271*	0.175	0.172
	(0.122)	(0.122)	(0.130)	(0.130)
$L_{i,t-1}^G$	0.780	0.785	0.498	0.492
	(0.487)	(0.486)	(0.459)	(0.459)
Constant	2.910***	2.870***		
	(0.099)	(0.115)		
Estimation Method	OLS	OLS	Clogit	Clogit
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Firms	2,920	2,920	-	-
Observations	11,546	11,546	3,308	3,308

Table B.1: Regression Results For Equation 2.1

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

	Dependent Variable	
	Inv Rate	Spike
$(CU_{i,t} - CU_{s,\tau}^\psi)$	-0.007 (0.005)	0.002 (0.004)
$\mathcal{O}_{i,t}$	0.267* (0.125)	0.022 (0.094)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times \mathcal{O}_{i,t}$	0.012 (0.010)	0.007 (0.007)
$\pi_{i,t-1}$	0.299** (0.103)	0.268*** (0.079)
$L_{i,t-1}^G$	0.618 (0.383)	0.501 (0.266)
Constant	2.753*** (0.111)	
Estimation Method	OLS	Clogit
Time FE	✓	✓
Firm FE	✓	✓
Firms	3,575	
Observations	23,954	8,815

Table B.2: Regression Results For Equation 2.1 with Intensity

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Investment Rate				
	(1)	(2)	(3)	(4)
$\mathcal{O}_{i,t}$	0.282*	0.281*	0.417**	0.407**
	(0.124)	(0.124)	(0.141)	(0.140)
$(CU_{i,t} - CU_{s,\tau}^\psi)$	-0.005	-0.004	-0.004	-0.004
	(0.006)	(0.006)	(0.005)	(0.005)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times \mathcal{O}_{i,t}$	0.010	0.009	0.009	0.009
	(0.010)	(0.010)	(0.010)	(0.010)
$FC_{i,t}$	-0.570**	-0.510*	-0.618*	-0.495
	(0.211)	(0.206)	(0.262)	(0.257)
$U_{s,t}$	0.018*	0.018*	0.033***	0.030**
	(0.007)	(0.007)	(0.010)	(0.009)
$\mathcal{O}_{i,t} \times FC_{i,t}$			0.093	-0.085
			(0.327)	(0.320)
$FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi)$	0.013	0.005		
	(0.010)	(0.010)		
$\mathcal{O}_{i,t} \times U_{s,t}$			-0.031*	-0.026*
			(0.013)	(0.013)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times U_{s,t}$	0.001	0.001		
	(0.001)	(0.001)		
$U_{s,t} \times FC_{i,t}$	-0.022	-0.028	-0.029	-0.024
	(0.024)	(0.023)	(0.024)	(0.023)
$\pi_{i,t-1}$	0.292**	0.300**	0.291**	0.299**
	(0.103)	(0.103)	(0.103)	(0.103)
$L_{i,t-1}^G$	0.628	0.650	0.625	0.644
	(0.380)	(0.379)	(0.380)	(0.379)
$U_{s,t-1}$	0.034***	0.034***	0.034***	0.033***
	(0.007)	(0.007)	(0.007)	(0.007)
Constant	2.551***	2.548***	2.485***	2.486***
	(0.124)	(0.123)	(0.129)	(0.128)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Estimation Method	OLS	OLS	OLS	OLS
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Firms	3,575	3,575	3,575	3,575
Observations	23,951	23,951	23,951	23,951

Table B.3: Regression Results For Equation 2.2 considering Intensity Effect

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

	Dependent Variable: Probability of Spike			
	(1)	(2)	(3)	(4)
$\mathcal{O}_{i,t}$	0.033 (0.095)	0.033 (0.095)	0.116 (0.106)	0.124 (0.104)
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.004 (0.004)	0.004 (0.004)	0.002 (0.004)	0.003 (0.004)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times \mathcal{O}_{i,t}$	0.006 (0.007)	0.006 (0.007)	0.007 (0.007)	0.006 (0.007)
$FC_{i,t}$	-0.481** (0.165)	-0.355* (0.165)	-0.409 (0.219)	-0.186 (0.201)
$U_{s,t}$	0.012* (0.005)	0.012** (0.005)	0.022** (0.008)	0.019** (0.007)
$\mathcal{O}_{i,t} \times FC_{i,t}$			-0.062 (0.276)	-0.388 (0.263)
$FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi)$	-0.003 (0.009)	-0.004 (0.008)		
$\mathcal{O}_{i,t} \times U_{s,t}$			-0.017 (0.010)	-0.015 (0.009)
$(CU_{s,\tau}^\psi) \times U_{s,t}$	0.001 (0.001)	0.001 (0.001)		
$U_{s,t} \times FC_{i,t}$	0.004 (0.014)	0.001 (0.016)	-0.005 (0.015)	0.006 (0.016)
$\pi_{i,t-1}$	0.258** (0.079)	0.266*** (0.079)	0.258** (0.079)	0.265** * (0.079)
$L_{i,t-1}^G$	0.505 (0.266)	0.519 (0.266)	0.492 (0.267)	0.516 (0.266)
$U_{s,t-1}$	0.011** (0.004)	0.011* (0.004)	0.011** (0.004)	0.011* (0.004)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Estimation Method	OLS	OLS	OLS	OLS
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Observations	8,813	8,813	8,813	8,813

Table B.4: Regression Results For Equation 2.2 considering Intensity Effect

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

C Robustness Checks - Regression Tables

Dependent Variable: Realised Investment Rate				
	(1)	(2)	(3)	(4)
$\mathcal{O}_{i,t}$			0.252*	0.216
			(0.126)	(0.125)
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.001	-0.002		
	(0.004)	(0.004)		
$FC_{i,t}$	-0.409	-0.672**	-0.112	-0.588*
	(0.231)	(0.223)	(0.286)	(0.278)
$U_{s,t}$	-0.005	-0.005	0.004	0.002
	(0.008)	(0.008)	(0.011)	(0.010)
$\mathcal{O}_{i,t} \times FC_{i,t}$			-0.500	-0.206
			(0.357)	(0.346)
$FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi)$	-0.010	0.011		
	(0.010)	(0.011)		
$\mathcal{O}_{i,t} \times U_{s,t}$			-0.017	-0.014
			(0.015)	(0.015)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times U_{s,t}$	-0.001	-0.001		
	(0.001)	(0.001)		
$U_{s,t} \times FC_{i,t}$	-0.012	-0.022	-0.016	-0.020
	(0.026)	(0.025)	(0.026)	(0.025)
$\pi_{i,t-1}$	-0.003	-0.015	-0.022	-0.033
	(0.112)	(0.111)	(0.111)	(0.111)
$L_{i,t-1}^G$	2.960***	2.940***	2.915***	2.909** *
	(0.412)	(0.412)	(0.412)	(0.412)
$U_{s,t-1}$	0.020**	0.020**	0.021**	0.020**
	(0.007)	(0.007)	(0.007)	(0.007)
Constant	3.254***	3.290***	3.120***	3.174***
	(0.105)	(0.105)	(0.126)	(0.124)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Estimation Method	OLS	OLS	OLS	OLS
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Firms	3,575	3,575	3,575	3,575
Observations	24,166	24,166	24,166	24,166

Table C.1: Regression Results For Equation 2.2 with Realised investment rates

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Realised Probability of Spike				
	(1)	(2)	(3)	(4)
$\mathcal{O}_{i,t}$			0.206*	0.191*
			(0.082)	(0.082)
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.008**	0.007*		
	(0.003)	(0.003)		
$FC_{i,t}$	0.012	-0.556***	0.061	-0.560**
	(0.212)	(0.163)	(0.250)	(0.206)
$U_{s,t}$	0.010*	0.008	0.011	0.009
	(0.005)	(0.005)	(0.007)	(0.006)
$\mathcal{O}_{i,t} \times FC_{i,t}$			-0.087	0.008
			(0.254)	(0.265)
$FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi)$	-0.005	0.006		
	(0.008)	(0.009)		
$\mathcal{O}_{i,t} \times U_{s,t}$			-0.002	-0.001
			(0.008)	(0.008)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times U_{s,t}$	-0.001	-0.001		
	(0.001)	(0.001)		
$U_{s,t} \times FC_{i,t}$	-0.036	0.002	-0.036	0.002
	(0.037)	(0.015)	(0.036)	(0.015)
$\pi_{i,t-1}$	-0.083	-0.092	-0.078	-0.086
	(0.076)	(0.076)	(0.076)	(0.076)
$L_{i,t-1}^G$	1.505***	1.507***	1.503***	1.511** *
	(0.264)	(0.264)	(0.265)	(0.265)
$U_{s,t-1}$	0.011**	0.011**	0.012**	0.011**
	(0.004)	(0.004)	(0.004)	(0.004)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Estimation Method	CLogit	CLogit	CLogit	CLogit
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Observations	8,830	8,830	8,830	8,830

Table C.2: Regression Results For Equation 2.2 with Realised Spike

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

	Dependent Variable			
	Realised Inv Rate		Realised Spike	
$(CU_{i,t} - CU_{s,\tau}^\psi)$	-0.003 (0.004)		0.007** (0.002)	
$\mathcal{O}_{i,t}$		0.127 (0.102)		0.183** (0.068)
$\pi_{i,t-1}$	0.003 (0.114)	-0.023 (0.114)	-0.084 (0.076)	-0.079 (0.076)
$L_{i,t-1}^G$	2.987*** (0.421)	2.940*** (0.421)	1.489*** (0.264)	1.489*** (0.264)
Constant	3.308*** (0.090)	3.250*** (0.103)		
Estimation Method	OLS	OLS	Clogit	Clogit
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Firms	3,575	3,575		
Observations	24,170	24,170	8,833	8,833

Table C.3: Regression Results For Equation 2.1 using Realised investments

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Investment Rate ^M				
	(1)	(2)	(3)	(4)
$\mathcal{O}_{i,t}$			0.395*** (0.118)	0.403*** (0.118)
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.005 (0.004)	0.006 (0.004)		
$FC_{i,t}$	-0.656** (0.216)	-0.562** (0.210)	-0.657* (0.268)	-0.505 (0.263)
$U_{s,t}$	-0.008 (0.007)	-0.007 (0.007)	0.004 (0.010)	0.003 (0.010)
$\mathcal{O}_{i,t} \times FC_{i,t}$			0.029 (0.334)	-0.115 (0.327)
$FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi)$	0.009 (0.010)	0.002 (0.010)		
$\mathcal{O}_{i,t} \times U_{s,t}$			-0.023 (0.014)	-0.020 (0.014)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times U_{s,t}$	-0.001 (0.001)	-0.001 (0.001)		
$U_{s,t} \times FC_{i,t}$	-0.001 (0.024)	-0.015 (0.024)	-0.007 (0.024)	-0.013 (0.024)
$\pi_{i,t-1}$	0.325** (0.105)	0.333** (0.105)	0.309** (0.105)	0.318** (0.105)
$L_{i,t-1}^G$	1.103** (0.389)	1.124** (0.389)	1.077** (0.388)	1.097** (0.388)
$U_{s,t-1}$	-0.000 (0.007)	-0.001 (0.007)	0.000 (0.007)	-0.000 (0.007)
Constant	3.013*** (0.099)	3.000*** (0.099)	2.794*** (0.118)	2.779*** (0.117)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Estimation Method	OLS	OLS	OLS	OLS
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Firms	3,575	3,575	3,575	3,575
Observations	23,951	23,951	23,951	23,951

Table C.4: Regression Results For Equation 2.2

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Probability of Spike ^M				
	(1)	(2)	(3)	(4)
$\mathcal{O}_{i,t}$			0.253** (0.084)	0.251** (0.083)
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.009** (0.003)	0.009** (0.003)		
$FC_{i,t}$	-0.337 (0.178)	-0.425** (0.164)	-0.282 (0.227)	-0.327 (0.204)
$U_{s,t}$	0.003 (0.005)	0.001 (0.005)	0.009 (0.008)	0.005 (0.007)
$\mathcal{O}_{i,t} \times FC_{i,t}$			-0.075 (0.268)	-0.200 (0.258)
$FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi)$	-0.002 (0.009)	-0.002 (0.009)		
$\mathcal{O}_{i,t} \times U_{s,t}$			-0.010 (0.010)	-0.008 (0.009)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times U_{s,t}$	-0.001 (0.001)	-0.001 (0.001)		
$U_{s,t} \times FC_{i,t}$	-0.012 (0.023)	0.012 (0.016)	-0.016 (0.022)	0.015 (0.017)
$\pi_{i,t-1}$	0.310*** (0.078)	0.313*** (0.078)	0.316*** (0.078)	0.319** * (0.078)
$L_{i,t-1}^G$	0.889*** (0.259)	0.900*** (0.258)	0.880*** (0.259)	0.898** * (0.258)
$U_{s,t-1}$	0.001 (0.004) (0.004)	0.001 (0.004) (0.006)	0.001 (0.004) (0.005)	0.001 (0.004) (0.007)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Estimation Method	CLogit	CLogit	CLogit	CLogit
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Observations	9,053	9,053	9,053	9,053

Table C.5: Regression Results For Equation 2.2

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval,

* Significant with 90% Confidence Interval

	Dependent Variable			
	Inv Rate ^M		Spike ^M	
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.004 (0.003)		0.008** (0.002)	
$\mathcal{O}_{i,t}$		0.303** (0.097)		0.204** (0.068)
$\pi_{i,t-1}$	0.355*** (0.107)	0.336** (0.107)	0.322*** (0.078)	0.327** * (0.077)
$L_{i,t}^G$	1.204** (0.398)	1.171** (0.398)	0.902*** (0.258)	0.901** * (0.258)
Constant	2.915*** (0.085)	2.755*** (0.098)		
Estimation Method	OLS	OLS	CLogit	CLogit
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Firms	3,575	3,575		
Observations	23,956	23,956	9,057	9,057

Table C.6: Regression Results For Equation 2.1 using more lags for investment rates

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Investment Rate ^L				
	(1)	(2)	(3)	(4)
$\mathcal{O}_{i,t}$			0.170*** (0.020)	0.178*** (0.020)
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.007*** (0.001)	0.007*** (0.001)		
$FC_{i,t}$	-0.228*** (0.038)	-0.249*** (0.037)	-0.275*** (0.048)	-0.222*** (0.047)
$U_{s,t}$	-0.005*** (0.001)	-0.005*** (0.001)	-0.003 (0.002)	-0.004* (0.002)
$\mathcal{O}_{i,t} \times FC_{i,t}$			0.083 (0.058)	-0.047 (0.057)
$FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi)$	0.004* (0.002)	-0.001 (0.002)		
$\mathcal{O}_{i,t} \times U_{s,t}$			-0.003 (0.002)	-0.002 (0.002)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times U_{s,t}$	-0.001 (0.001)	-0.001 (0.001)		
$U_{s,t} \times FC_{i,t}$	-0.002 (0.004)	0.004	-0.003 (0.004)	0.004
$\pi_{i,t-1}$	0.231*** (0.018)	0.233*** (0.018)	0.235*** (0.018)	0.238** * (0.018)
$L_{i,t-1}^G$	0.257*** (0.067)	0.260*** (0.067)	0.266*** (0.067)	0.271** * (0.067)
$U_{s,t-1}$	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)
Constant	6.156*** (0.017)	6.155*** (0.017)	6.059*** (0.020)	6.054*** (0.020)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Estimation Method	OLS	OLS	OLS	OLS
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Firms	3,461	3,461	3,461	3,461
Observations	22,697	22,697	22,697	22,697

Table C.7: Regression Results For Equation 2.2 with normalisation by labour

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval,
* Significant with 90% Confidence Interval

Dependent Variable: Probability of Spike ^L				
	(1)	(2)	(3)	(4)
$\mathcal{O}_{i,t}$			0.267** (0.095)	0.293** (0.094)
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.007* (0.003)	0.008* (0.003)		
$FC_{i,t}$	-0.059 (0.303)	-0.640** (0.212)	-0.009 (0.361)	-0.164 (0.259)
$U_{s,t}$	-0.006 (0.005)	-0.007 (0.005)	0.002 (0.008)	-0.001 (0.007)
$\mathcal{O}_{i,t} \times FC_{i,t}$			-0.094 (0.333)	-0.995** (0.346)
$FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi)$	-0.002 (0.011)	-0.023* (0.011)		
$\mathcal{O}_{i,t} \times U_{s,t}$			-0.015 (0.010)	-0.014 (0.010)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times U_{s,t}$	-0.001 (0.001)	-0.001 (0.001)		
$U_{s,t} \times FC_{i,t}$	-0.072 (0.063)	0.009 (0.024)	-0.072 (0.063)	0.017 (0.026)
$\pi_{i,t-1}$	0.680*** (0.097)	0.681*** (0.097)	0.679*** (0.097)	0.680** * (0.097)
$L_{i,t-1}^G$	0.495 (0.329)	0.525 (0.331)	0.477 (0.330)	0.512 (0.331)
$U_{s,t-1}$	0.001 (0.005)	0.000 (0.005)	0.001 (0.005)	-0.000 (0.005)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Estimation Method	CLogit	CLogit	CLogit	CLogit
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Observations	6,799	6,799	6,799	6,799

Table C.8: Regression Results For Equation 2.2 with normalisation by labour

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

	Dependent Variable			
	Inv Rate ^L		Spike ^L	
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.007*** (0.001)		0.006* (0.003)	
$\mathcal{O}_{i,t}$		0.172*** (0.016)		0.205* (0.080)
$\pi_{i,t-1}$	0.243*** (0.018)	0.247*** (0.018)	0.696*** (0.097)	0.694** * (0.096)
$L_{i,t-1}^G$	0.273*** (0.067)	0.284*** (0.067)	0.522 (0.328)	0.512 (0.328)
Constant	6.107*** (0.014)	6.009*** (0.016)		
Estimation Method	OLS	OLS	CLogit	CLogit
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Firms	3,461	3,461		
Observations	22,701	22,701	6,803	6,803

Table C.9: Regression Results For Equation 2.1 with normalisation by labour

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Investment Rate				
	(1)	(2)	(3)	(4)
$\mathcal{O}_{i,t}^{\text{alt}}$			0.174 (0.101)	0.169 (0.101)
$(CU_{i,t} - CU_{i,\tau}^{\psi})$	0.010* (0.005)	0.010* (0.005)		
$FC_{i,t}$	-0.542* (0.212)	-0.540** (0.206)	-0.774** (0.260)	-0.637* (0.257)
$U_{s,t}$	0.017* (0.007)	0.016* (0.007)	0.028** (0.010)	0.024** (0.009)
$\mathcal{O}_{i,t} \times FC_{i,t}$			0.363 (0.315)	0.167 (0.312)
$FC_{i,t} \times (CU_{i,t} - CU_{i,\tau}^{\psi})$	0.015 (0.012)	0.004 (0.013)		
$\mathcal{O}_{i,t} \times U_{s,t}$			-0.023 (0.013)	-0.017 (0.013)
$(CU_{i,t} - CU_{i,\tau}^{\psi}) \times U_{s,t}$	-0.001 (0.001)	-0.001 (0.001)		
$U_{s,t} \times FC_{i,t}$	-0.029 (0.024)	-0.023 (0.023)	-0.025 (0.024)	-0.022 (0.024)
$\pi_{i,t-1}$	0.302** (0.103)	0.310** (0.103)	0.305** (0.103)	0.313** (0.103)
$\mathbb{L}_{i,t-1}^G$	0.630 (0.380)	0.658 (0.379)	0.662 (0.379)	0.680 (0.379)
$U_{s,t-1}$	0.034*** (0.007)	0.033*** (0.007)	0.034*** (0.007)	0.033*** (0.007)
Constant	2.765*** (0.097)	2.762*** (0.096)	2.676*** (0.109)	2.672*** (0.108)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Estimation Method	OLS	OLS	OLS	OLS
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Firms	3,575	3,575	3,575	3,575
Observations	23,951	23,951	23,951	23,951

Table C.10: Regression Results For Equation 2.2

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

	Dependent Variable: Probability of Spike			
	(1)	(2)	(3)	(4)
$\mathcal{O}_{i,t}^{\text{alt}}$			0.141 (0.072)	0.139 (0.071)
$(CU_{i,t} - CU_{i,\tau}^{\psi})$	0.006 (0.003)	0.007* (0.003)		
$FC_{i,t}$	-0.462** (0.165)	-0.354* (0.164)	-0.486* (0.223)	-0.359 (0.210)
$U_{s,t}$	0.012* (0.005)	0.012** (0.004)	0.016* (0.007)	0.015* (0.006)
$\mathcal{O}_{i,t} \times FC_{i,t}$			0.032 (0.268)	0.002 (0.259)
$FC_{i,t} \times (CU_{i,t} - CU_{i,\tau}^{\psi})$	0.002 (0.011)	-0.004 (0.011)		
$\mathcal{O}_{i,t} \times U_{s,t}$			-0.008 (0.009)	-0.007 (0.008)
$(CU_{i,t} - CU_{i,\tau}^{\psi}) \times U_{s,t}$	-0.001 (0.001)	-0.001 (0.001)		
$U_{s,t} \times FC_{i,t}$	0.001 (0.014)	0.001 (0.016)	0.001 (0.014)	0.002 (0.016)
$\pi_{i,t-1}$	0.262*** (0.079)	0.271*** (0.079)	0.263*** (0.079)	0.271** * (0.079)
$L_{i,t-1}^G$	0.513 (0.266)	0.527* (0.266)	0.508 (0.267)	0.521 (0.266)
$U_{s,t-1}$	0.011** (0.004)	0.011** (0.004)	0.011** (0.004)	0.011** (0.004)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Estimation Method	CLogit	CLogit	CLogit	CLogit
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Observations	8,813	8,813	8,813	8,813

Table C.11: Regression Results For Equation 2.2

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable				
	Inv Rate		Spike	
$(CU_{i,t} - CU_{i,\tau}^\psi)$	0.003 (0.003)		0.004 (0.003)	
$\mathcal{O}_{i,t}^{\text{alt}}$		0.074 (0.078)		0.098 (0.057)
$\pi_{i,t-1}$	0.312** (0.103)	0.312** (0.103)	0.274*** (0.079)	0.275** * (0.079)
$L_{i,t-1}^G$	0.636 (0.383)	0.641 (0.383)	0.515 (0.266)	0.509 (0.266)
Constant	2.962*** (0.082)	2.920*** (0.089)		
Estimation Method	OLS	OLS	CLogit	CLogit
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Firms	3,575	3,575	3,575	3,575
Observations	23,954	23,954	23,954	23,954

Table C.12: Regression Results For Equation 2.1

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Investment Rate				
	(1)	(2)	(3)	(4)
$\mathcal{O}_{i,t}$			0.451*** (0.103)	0.482*** (0.103)
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.008* (0.004)	0.009** (0.004)		
$FC_{i,t}$	-0.612** (0.208)	-0.526** (0.201)	-0.671* (0.264)	-0.362 (0.264)
$FCE_{i,t}$	19.424*** (2.337)	25.683*** (4.300)	54.175*** (2.324)	66.861*** (4.356)
$\mathcal{O}_{i,t} \times FC_{i,t}$			0.237 (0.344)	-0.022 (0.334)
$FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi)$	0.015 (0.010)	0.002 (0.010)		
$\mathcal{O}_{i,t} \times FCE_{i,t}$			-35.429*** (3.554)	-36.731*** (3.567)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times FCE_{i,t}$	-0.637*** (0.047)	-0.632*** (0.047)		
$FCE_{i,t} \times FC_{i,t}$	1.178 (14.083)	-21.164 (12.295)	-9.573 (14.082)	-42.525*** (12.353)
$\pi_{i,t-1}$	0.276* (0.108)	0.276* (0.108)	0.279** (0.108)	0.277* (0.108)
$L_{i,t-1}^G$	0.378 (0.395)	0.393 (0.394)	0.401 (0.395)	0.403 (0.395)
$FCE_{i,t-1}$	11.492*** (2.176)	11.558*** (2.175)	11.142*** (2.180)	11.276*** (2.179)
Constant	2.797*** (0.090)	2.793*** (0.090)	2.449*** (0.105)	2.431*** (0.104)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Estimation Method	OLS	OLS	OLS	OLS
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Firms	3,093	3,093	3,093	3,093
Observations	20,008	20,008	20,008	20,008

Table C.13: Regression Results For Equation 2.2 using FCE instead of U as uncertainty proxy

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval,
* Significant with 90% Confidence Interval

Dependent Variable: Probability of Spike				
	(1)	(2)	(3)	(4)
$\mathcal{O}_{i,t}$			0.191*	0.214*
			(0.095)	(0.095)
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.010**	0.010**		
	(0.003)	(0.003)		
$FC_{i,t}$	-0.365	-0.299	-0.441	-0.127
	(0.195)	(0.196)	(0.257)	(0.242)
$FCE_{i,t}$	15.028***	15.619***	19.243***	19.208***
	(4.392)	(4.465)	(5.041)	(5.018)
$\mathcal{O}_{i,t} \times FC_{i,t}$			0.143	-0.327
			(0.301)	(0.288)
$FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi)$	0.002	-0.002		
	(0.010)	(0.010)		
$\mathcal{O}_{i,t} \times FCE_{i,t}$			-6.362	-5.081
			(7.628)	(7.737)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times FCE_{i,t}$	-0.234	-0.218		
	(0.184)	(0.186)		
$FCE_{i,t} \times FC_{i,t}$	-8.002	-14.029	-8.215	-14.710
	(14.961)	(15.978)	(14.737)	(16.200)
$\pi_{i,t-1}$	0.219*	0.224*	0.232**	0.236**
	(0.088)	(0.088)	(0.087)	(0.087)
$L_{i,t-1}^G$	0.367	0.374	0.379	0.390
	(0.290)	(0.290)	(0.289)	(0.289)
$FCE_{i,t-1}$	1.058	1.083	0.891	0.910
	(1.238)	(1.245)	(1.260)	(1.266)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Estimation Method	CLogit	CLogit	CLogit	CLogit
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Observations	7,150	7,150	7,150	7,150

Table C.14: Regression Results For Equation 2.2 using FCE instead of OU as uncertainty proxy

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Investment Rate				
	(1)	(2)	(3)	(4)
$\mathcal{O}_{i,t}$			0.364 (0.334)	0.389 (0.335)
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.001 (0.009)	0.002 (0.009)		
$FC_{i,t}$	0.111 (0.711)	-0.493 (0.537)	-0.034 (0.728)	-0.391 (0.576)
$SU_{i,t}$	0.000 (0.064)	0.009 (0.093)	0.075 (0.096)	0.066 (0.113)
$\mathcal{O}_{i,t} \times FC_{i,t}$			0.414 (0.369)	-0.297 (0.374)
$FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi)$	0.013 (0.011)	-0.015 (0.011)		
$\mathcal{O}_{i,t} \times SU_{i,t}$			-0.130 (0.121)	-0.121 (0.121)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times SU_{i,t}$	0.001 (0.003)	0.002 (0.003)		
$SU_{i,t} \times FC_{i,t}$	-0.233 (0.256)	-0.028 (0.183)	-0.269 (0.256)	-0.003 (0.184)
$\pi_{i,t-1}$	0.419*** (0.118)	0.423*** (0.118)	0.431*** (0.118)	0.434** * (0.118)
$\mathbb{L}_{i,t-1}^G$	1.238** (0.472)	1.271** (0.471)	1.258** (0.471)	1.301** (0.471)
$SU_{i,t-1}$	0.000 (0.064)	0.001 (0.066)	0.075 (0.096)	0.065 (0.095)
Constant	2.942*** (0.226)	2.928*** (0.229)	2.727*** (0.303)	2.714*** (0.301)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Estimation Method	OLS	OLS	OLS	OLS
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Firms	2,220	2,220	2,220	2,220
Observations	8,888	8,888	8,888	8,888

Table C.15: Regression Results For Equation 2.2 using subjective uncertainty

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Probability of Spike				
	(1)	(2)	(3)	(4)
$\mathcal{O}_{i,t}$			0.635 (0.408)	0.680 (0.411)
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.003 (0.013)	0.004 (0.013)		
$FC_{i,t}$	0.718 (1.100)	-0.477 (0.954)	0.861 (1.122)	0.330 (1.018)
$SU_{i,t}$	-0.007 (0.075)	-0.017 (0.075)	0.181 (0.123)	0.175 (0.126)
$\mathcal{O}_{i,t} \times FC_{i,t}$			0.043 (0.444)	-0.616 (0.454)
$FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi)$	0.001 (0.015)	-0.017 (0.014)		
$\mathcal{O}_{i,t} \times SU_{i,t}$			-0.281 (0.149)	-0.283 (0.151)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times SU_{i,t}$	-0.003 (0.005)	-0.003 (0.005)		
$SU_{i,t} \times FC_{i,t}$	-0.397 (0.419)	0.064 (0.340)	-0.452 (0.425)	-0.121 (0.350)
$\pi_{i,t-1}$	0.497*** (0.140)	0.494*** (0.141)	0.496*** (0.140)	0.492** * (0.140)
$L_{i,t-1}^G$	0.612 (0.548)	0.627 (0.550)	0.631 (0.547)	0.642 (0.550)
$SU_{i,t-1}$	0.001 (0.073) (0.004)	-0.002 (0.073) (0.006)	-0.002 (0.074) (0.005)	-0.007 (0.074) (0.007)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Estimation Method	CLogit	CLogit	CLogit	CLogit
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Observations	2,497	2,497	2,49781	2,497

Table C.16: Regression Results For Equation 2.2

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

	Dependent Variable			
	Inv Rate		Spike	
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.011 (0.007)		0.008 (0.005)	
$\mathcal{O}_{i,t}$		0.322 (0.205)		0.123 (0.150)
Large	0.358 (0.200)	0.420 (0.233)	0.301* (0.134)	0.295 (0.166)
Large $\times (CU_{i,t} - CU_{s,\tau}^\psi)$	-0.009 (0.008)		-0.003 (0.006)	
Large $\times \mathcal{O}_{i,t}$		-0.093 (0.227)		0.016 (0.166)
$\pi_{i,t-1}$	0.312** (0.104)	0.297** (0.104)	0.255** (0.079)	0.259** (0.079)
$L_{i,t}^G$	0.621 (0.386)	0.589 (0.386)	0.487 (0.269)	0.478 (0.269)
Constant	2.678*** (0.177)	2.496*** (0.206)		
Estimation Method	OLS	OLS	Clogit	Clogit
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Firms	3,568	3,568	3,568	3,568
Observations	23,753	23,753	23,753	23,753

Table C.17: Regression Results For Equation 2.1

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

	Dependent Variable: Investment Rate			
	(1)	(2)	(3)	(4)
$\mathcal{O}_{i,t}$			0.485*	0.476*
			(0.215)	(0.214)
$(CU_{i,t} - CU_{s,\tau}^{\psi})$	0.012	0.013		
	(0.007)	(0.007)		
$FC_{i,t}$	-0.253	-0.853*	-0.251	-0.793*
	(0.365)	(0.360)	(0.400)	(0.395)
$U_{s,t}$	0.013	0.013	0.033	0.029
	(0.017)	(0.017)	(0.019)	(0.019)
Large	0.367	0.290	0.441	0.357
	(0.217)	(0.217)	(0.251)	(0.251)
Large $\times (CU_{i,t} - CU_{s,\tau}^{\psi})$	-0.008	-0.008		
	(0.008)	(0.008)		
Large $\times \mathcal{O}_{i,t}$			-0.071	-0.065
			(0.225)	(0.225)
Large $\times FC_{i,t}$	-0.458	0.454	-0.443	0.443
	(0.397)	(0.392)	(0.397)	(0.392)
Large $\times U_{s,t}$	0.005	0.005	0.000	0.001
	(0.018)	(0.018)	(0.018)	(0.018)
$\mathcal{O}_{i,t} \times FC_{i,t}$			0.026	-0.109
			(0.329)	(0.322)
$FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^{\psi})$	0.009	0.004		
	(0.010)	(0.010)		
$\mathcal{O}_{i,t} \times U_{s,t}$			-0.032*	-0.026
			(0.014)	(0.014)
$(CU_{i,t} - CU_{s,\tau}^{\psi}) \times U_{s,t}$	0.001	0.001		
	(0.001)	(0.001)		
$U_{s,t} \times FC_{i,t}$	-0.023	-0.028	-0.031	-0.025
	(0.024)	(0.023)	(0.024)	(0.023)
$\pi_{i,t-1}$	0.304**	0.315**	0.291**	0.303**
	(0.103)	(0.103)	(0.103)	(0.103)
$L_{i,t-1}^G$	0.632	0.652	0.603	0.624
	(0.382)	(0.382)	(0.382)	(0.382)
$U_{s,t-1}$	0.034***	0.033***	0.034***	0.034***
	(0.007)	(0.007)	(0.007)	(0.007)
Constant	2.473***	2.522***	2.176***	2.234***
	(0.199)	(0.199)	(0.232)	(0.230)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Estimation Method	OLS	OLS	OLS	OLS
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Firms	3,568	3,568	3,568	3,568
Observations	23,750	23,750	23,750	23,750

Table C.18: Regression Results For Equation 2.2 controlling for Large Firms

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

	Dependent Variable: Probability of Spike			
	(1)	(2)	(3)	(4)
$\mathcal{O}_{i,t}$			0.230 (0.159)	0.101 (0.136)
$(CU_{i,t} - CU_{s,\tau}^\psi)$	0.010 (0.005)	0.010 (0.005)		
$FC_{i,t}$	-0.289 (0.270)	-0.847** (0.314)	-0.185 (0.313)	-0.711* (0.313)
$U_{s,t}$	0.017 (0.011)	0.016 (0.011)	0.029* (0.013)	0.021 (0.012)
Large	0.338* (0.145)	0.275 (0.146)	0.350* (0.176)	
Large $\times (CU_{i,t} - CU_{s,\tau}^\psi)$	-0.004 (0.006)	-0.004 (0.006)		
Large $\times \mathcal{O}_{i,t}$			0.004 (0.167)	0.165 (0.136)
Large $\times FC_{i,t}$	-0.276 (0.314)	0.699* (0.334)	-0.268 (0.313)	0.745* (0.329)
Large $\times U_{s,t}$	-0.005 (0.012)	-0.003 (0.012)	-0.007 (0.012)	0.001 (0.011)
$\mathcal{O}_{i,t} \times FC_{i,t}$			-0.120 (0.281)	-0.361 (0.264)
$FC_{i,t} \times (CU_{i,t} - CU_{s,\tau}^\psi)$	-0.004 (0.009)	-0.002 (0.009)		
$\mathcal{O}_{i,t} \times U_{s,t}$			-0.017 (0.010)	-0.013 (0.009)
$(CU_{i,t} - CU_{s,\tau}^\psi) \times U_{s,t}$	-0.001 (0.001)	-0.001 (0.001)		
$U_{s,t} \times FC_{i,t}$	0.001 (0.014)	-0.002 (0.016)	-0.009 (0.015)	0.002 (0.016)
$\pi_{i,t-1}$	0.247** (0.080)	0.258** (0.080)	0.251** (0.080)	0.262** * (0.080)
$L_{i,t-1}^G$	0.489 (0.269)	0.501 (0.269)	0.463 (0.269)	0.507 (0.268)
$U_{s,t-1}$	0.011** (0.004)	0.011** (0.004)	0.011** (0.004)	0.011** (0.004)
Instrument for $FC_{i,t}$	-	$FC_{i,t-1}$	-	$FC_{i,t-1}$
Estimation Method	CLogit	CLogit	CLogit	CLogit
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Observations	8,876	8,876	8,876	8,876

Table C.19: Regression Results For Equation 2.2

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

D Questions Used from the INVIND Survey

- Investments in plant and machinery: “What is the extent of your investment in plant, machinery and equipment in the current year and next year”?
- Capacity Utilisation: Capacity Utilisation
- Sales: Turnover from year’s sales of goods
- Profits: Please describe the firm’s operating result (1 = large profit; 2 = small profit; 3 = broad balance; 4 = small loss; 5 = large loss.)
- Employees: Average Workforce

Construction of the Dummy FC

A major disagreement in the economic literature concerns the identification of firms that are financially constrained.¹⁸ For instance, the work of [Amiti and Weinstein \(2018\)](#) uses a comprehensive, matched, lender-borrower data set covering all loans received from all sources by every listed Japanese firm over the period 1990–2010.¹⁹ By surveying a group of CFOs in various geographical locations and using the matching procedure²⁰, [Campello, Graham and Harvey \(2010\)](#) find that constrained firms planned deeper cuts in tech spending, employment, and capital spending during the Great Recession of 2008. By measuring financial constraints using official credit ratings among a sample of Italian firms, [Bottazzi, Secchi and Tamagni \(2014\)](#) find that constrained firms have a lower chance of growth as well as more chance of experiencing higher volatilities. While their analysis demonstrates a substantial relationship between firms size and the chance of being constrained, [Ferreira, Haber and Rorig \(2023\)](#) find the presence of constrained firms²¹ across the entire size distribution among Portuguese firms. However, they show that the presence of constraints among larger firms amplifies the impact of a financial shock.

In the INVIND questionnaire, firms are asked multiple questions that can be interpreted as being financially constrained [or not].

These questions include:

1. Please indicate whether during 2021, at the interest rate and collateral terms applied to your firm, you wanted to increase your debt with banks or other financial intermediaries (Yes/No)

¹⁸[Schiantarelli \(1996\)](#) reviews the empirical approaches in measuring and identifying financial constraints using economic theory and/or firm-level data up to 1996. While the expectation of the author was that there would exist a common agreement in empirically proxying for financial constraints over time, there are still various ways to measure and identify the constraints that firms face while financing their expansion projects.

¹⁹Their paper provides evidence that supply-side financial shocks - regardless of their measurement - greatly impact firms’ investment.

²⁰Instead of comparing the average difference in policy outcomes across all of the constrained and all of the unconstrained firms, they compare the differences in average outcomes of firms that are quite similar (i.e., matched) except for the “marginal” dimension of CFO-reported financial constraints. This yields an estimate of the differential effect of financial constraints on corporate policies across “treated” firms and their “counterfactuals”

²¹Similar to [Bottazzi, Secchi and Tamagni \(2014\)](#), they also use credit information data to distinguish between constrained and unconstrained firms

2. you were willing to accept more stringent loan terms (e.g. higher interest rate or more collateral) in order to increase the amount of borrowing (Yes/No)

3. in 2021, did you actually apply for new loans from banks or other financial intermediaries (Yes/No)?

If 3 is yes:

(a) you received the amount requested (Yes/No)

(b) you were granted only part of the amount requested (Yes/No)

(c) you were given no loan because the financial intermediaries contacted were not willing to grant the loan (Yes/No)

(d) no loan was obtained for other reasons (e.g. cost or collateral considered to be excessive) (Yes/No)

If 3 is no:

(a) we didn't contact banks or other intermediaries because we were convinced they would reject the application (Yes/No)

Chapter 3

Data, Uncertainty and R&D Investments¹

Abstract

We investigate whether investing in data affects investments in research and development (R&D) and whether these two investments are complementary for productivity. We formulate and test the hypothesis that investments in data decrease firms' uncertainty due to the opportunity to gather strategic information on markets. Lowering market uncertainty might lead firms to better cope with the intrinsic uncertainty of R&D and, therefore, positively affect R&D investments. Based on Italian firm-level data covering 2002-2023 and using GDPR as an instrument for data in an IV setting, we provide empirical evidence on the causal relationship between investments in data and R&D. We find that higher investments in data and analytics increase firms' R&D activity. We also find that investments in R&D and data are complementary for productivity, meaning that investing in data and R&D and combining these assets increases firms' short- and long-term productivity.

Keywords: Data, R&D, Digitalisation, Innovation, Productivity, Uncertainty

¹This chapter is co-authored with Maria Savona

3.1 Introduction

"Some economists believe that the Greeks' work ethic and thrift can pull them through. But the classical virtues can do nothing to offset the dearth of innovation that plagues the economy."
— Edmund Phelps

Digital transformation and higher data-intensive processes underpin the 4th Industrial Revolution² and have brought substantial benefits for firms, such as production efficiency (Björkdahl, 2020). The economic literature on innovation and firm dynamics has primarily focused on robotics and digital automation (Cirillo et al., 2023). "Digitalisation", according to Aronica, Piacentino and Picone (2022), is the transformation of information into digital components that are easier to access, transfer, analyse and benefit from. "Datafication" refers to the collection, databasing, stock, and software analysis of information. An example is self-driving car producers, who collect, store and analyse millions of data points online from free (users' cars) and paid (traffic) sources. Therefore, digital automation has increased exponentially the amount and uses of data as a production factor to increase firm strategic knowledge, service optimisation, and economic value-generation (Mayer-Schönberger and Cukier, 2013; Klein et al., 2022). However, using this digital automation in the data-driven business environment might impact firms' willingness to engage in Research and Development (R&D). Furthermore, combining higher data intensity and more R&D can impact firms' performance through higher efficiency.

In this paper, we tackle these two issues. First, we study whether higher investments in data lead to more R&D investments. The increased investments in data and software or datafication result in increased information, which feeds firms' Machine Learning (ML) usage (Savona, 2019). ML tools are the primary instruments for efficiently converting raw data into information. The opportunity to gain information at a higher scale and faster pace than it occurred for previous ICT waves allows for lower uncertainties in inventory management, marketing and using inputs in production (Frankel and Kamenica, 2019). Less uncertainty translates into lower marginal costs, increased firm size, higher market concentration and lower market uncertainty. While it is well known that larger firms invest more in R&D, it is relatively less explored whether this evidence is related to the lowering of the uncertainty that investments in data allow, with notable exceptions (Eeckhout and Veldkamp, 2022; Brynjolfsson, Jin and Wang, 2023). We argue that the augmentation in R&D investments is facilitated by improved risk spread and enhanced uncertainty management, further fuelled by the advantages of size and scalability. It is plausible to argue that digital transformation and datafication provide substantially higher opportunities to reduce or better cope with uncertainty, with better and faster access to information stimulating R&D activities via lowering the intrinsic uncertainty of investments in innovation. We conjecture that each firm faces a portfolio of different varieties of uncertainty. While investing in data does not directly reduce R&D uncertainty - which stems from uncertain outcomes of innovation - it can reduce a firm's overall uncertainty via lower market uncertainties (Groh and Pfäuti, 2023) and hence, leave more room for R&D uncertainty.

However, the causal link between investments in data and R&D is not obvious. R&D generates data, which can then feed itself. Furthermore, some R&D activities are data-intensive and would directly need data. Therefore, studying data and R&D may be subject to endogeneity issues. As a solution, we use the General Data Protection Regulation (GDPR) and its implementation in Italy³ in

²Or interchangeably Industrie 4.0 due to Kagermann, Wahlster and Helbig (2013); Björkdahl (2020)

³Decreto Legislativo 10 Agosto 2018, n. 101

2018 as an exogenous instrument for data. GDPR limits firms in how they can collect and store data to protect individual's privacy, and therefore, can influence business models by potentially increasing the number of value chain functions, impacting customer service, and requiring changes in data handling practices, which may increase costs and affect the value formula (Lindgren, 2016). On the other hand, such regulations can improve incentives to buy and store data since they reduce the risk of high future judiciary costs for firms, which could occur under less transparent regimes (EU Commission, 2020).

Second, we study whether investing in data and R&D - as two different classes of intangible assets⁴ are complementary for labour productivity, meaning that their combination improves labour productivity, following Corrado et al. (2021). Previous efforts primarily point out such complementarity of intangible assets at the country (Nonnis, Bounfour and Kim, 2023) and firm and industry level (Ennen and Richter, 2010; Delbecque, Bounfour and Barreneche, 2015; Añón Higón, Gómez and Vargas, 2017), mainly across managerial and organisational capital and innovation.⁵ It has been found that the adoption and use of Artificial Intelligence (AI) - among other intangible assets - yield strong complementarity and learning spillover effects between AI investments and existing expertise in data management, networking, and communication (Igna and Venturini, 2023).

The role of intangible assets in affecting productivity (Goodridge and Haskel, 2023) and growth has been the object of a growing literature over the last decades, see Arrow (1972); Machlup (1962); Romer (1990); Hall (2011). The literature has previously shown that, among intangible investments, R&D activities impact firms the most (Griliches, 1984; Hall, Mairesse and Mohnen, 2010; Shefer and Frenkel, 2005; Siliverstovs et al., 2016; Mohnen, 2019). For R&D activities to bear fruit, firms need to invest in complementary elements (Teece, 1986), including tangible capital such as new machinery in the lab and intangible capital such as organisational capital, human capital or Information and Communication Technologies (Calvino et al., 2022). Akerman, Gaarder and Mogstad (2015), Truant, Broccardo and Dana (2021) and Cirillo et al. (2023) found a positive impact of digital technology adoption on labour productivity due to cost-reduction mechanisms. Nucci, Puccioni and Ricchi (2022)'s findings confirm the productivity gains of digital transformation but stress the importance of firm-level heterogeneity.

By putting these two conjectures of the study, this paper poses the questions: "Does investing in data affect R&D investments?" secondly, "Is there a cumulative effect on the productivity of sole investments in data, sole investments in R&D and the joint investments in data and R&D?"

In this paper, we contribute to the literature on firm dynamics by studying the relationship between investments in data and analytics, R&D and productivity. We build on Hall (2011) and explicitly consider the high heterogeneity in intangible assets and the conjecture related to trade-offs in types of firm risks. Hence, we test the presence of a causal relationship between investments in data and R&D activities. We do so by applying an instrumental variable (IV) approach to firms present in the INVIND survey of the Bank of Italy. In addition, we test the potential complementarity of investments in data and R&D to labour productivity. Specifically, we examine whether sole or multiplied data and R&D investments impact firm labour productivity performance. We find that not only do firms

⁴R&D and data and analytics are usually classified as "intangible" assets. This class of assets comprises a collection of intangible resources crucial for producing and marketing new or improved products and processes. It encompasses both internally generated resources like innovative knowledge, designs, blueprints, brand equity, in-house software, and construction projects, as well as externally acquired resources like technology licenses, patents, copyrights, and economic competencies gained through the procurement of management and consulting services (Arrighetti, Landini and Lasagni, 2014).

⁵In the latter, the complementarity depends on the asset type. For instance, a firm might be unwilling to spend resources on extensive marketing campaigns and instead use data analytics to provide tailored products to its clients (Grabowska and Saniuk, 2021; Piccarozzi, Aquilani and Gatti, 2018). In this case, acquiring one intangible asset (data and software) can substitute the other (marketing).

experience within-year gains in their productivity by investing in both data and R&D, but they may continue to gain in the longer term. In this section, we argued that the impact of datafication on firms' performance or value is still fairly under-researched. We fill this gap by exploring whether and how investments in database and software complement existing investments in R&D. We test this complementarity conjecture empirically by analysing the multiplier effect and how combining these investments can improve firms' labour productivity.

Our results can be of value to both academic and policy debates, including on data regulation. For instance, our first-stage estimations showed that implementing GDPR might have increased investments in data and analytics. Hence, regulations commonly considered as constraining firm investments and performance might instead be leverage for firms to comply in different - not necessarily detrimental - ways. Furthermore, we showed that more investments in data encourage investments in R&D, and both of these investments and their combination improve firms' performance. In this line, we provide empirical evidence for the benefits of B2B data sharing for more R&D and higher labour productivity, following (Goos and Savona, 2024; Savona, 2024).

The remainder of the paper is structured as follows. First, we introduce a stylised theoretical framework in Section 3.2. In Section 3.3, we motivate the choice of Italy as a study context and describe the dataset of interest, INVIND and the constructed variables we use in our study. Section 3.4 details our empirical strategy and Section 3.5 presents the results. Section 3.6 provides the robustness checks. Finally, Section 3.7 concludes by summarising our findings and discussing their implications.

3.2 Theoretical Framework

In this section, we present the theoretical framework to analyse the causality between data and R&D and the complementarity of these assets for productivity.

3.2.1 From Data to R&D

We propose a mechanism through which investing in data and software can lead to higher research and development activities. The previous literature has shown that data can improve R&D outcomes. Furthermore, a firm holding big data can provide novel decision-making possibilities, which support firms' innovation process (Niebel, Rasel and Viete, 2019). We argue that by increasing firms' information set through deeper analytics and more data, firms face less market uncertainty about the future, can maintain growth through lower costs, gain market power, and invest more in R&D (Del Monte and Papagni, 2003; Shefer and Frenkel, 2005).

Throughout our study, we treat data and analytics as information. Doing so opens gateways to see reducing uncertainty as a primary gain from these investments (CFOTech, 2023; Wondrium, 2016). From a firm's perspective, uncertainty operates as a hidden cost and translates into delays in investment and hiring (Cuervo-Cazurra and Annique Un, 2010; Bloom et al., 2022), impeding firms' expansion. While business-cycle-related developments affect firms' uncertainty, data and analytics can partially help reduce these uncertainties (Simon, 2013). Reducing uncertainty via having more data affects the firm size and the composition of the goods firms produce, all of which affect markups (Eeckhout and Veldkamp, 2022) and firm's value (Salvi et al., 2021). Firms that fail to undertake correct practices in analysing outside and aggregate data are more likely to underperform through

higher uncertainty about any phenomenon outside of their control (Wu et al., 2017; Tysiac, 2015).⁶ Therefore, having more data translates into a more extensive information set, which could reduce overall uncertainty across a portfolio of activities and motivate the firm to engage in other high-uncertainty activities such as R&D. Previously, the literature found that increased data and software capabilities significantly impact R&D and its management. Big data has the potential to inform, enable, and transform R&D strategies, processes, and technologies across various industries (Blackburn et al., 2017). Digitalisation, including virtual experimentation, collaboration, and big data, is expected to enhance consumer intimacy, accelerate time-to-market, and promote open innovation in R&D (Farrington and Alizadeh, 2017).

We are aware of multiple ways that the direct relationship between data and R&D could be exposed to severe endogeneity issues. First, R&D generates data and in-house software, which might be a potential reverse causality. Furthermore, firms' R&D investment plans are affected by uncertainty (Bloom, 2007) and increase uncertainty (Fung, 2006). For our framework to identify the causal direction and to overcome the issues of simultaneity and the possibility of reverse causality between data investments and R&D,⁷ we use the General Data and Privacy Regulation (GDPR) as an exogenous instrument for investments in data. GDPR may be expected to limit investments in data and analytics by increasing the costs of storing and using data through more strict regulations on privacy. The literature has previously studied multiple ways in which GDPR can affect firms' venture investments (Jia, Jin and Wagman, 2020), revenues (Goldberg, Johnson and Shriver, 2024; Koski and Valmari, 2020), data sharing (Gal and Aviv, 2020), and the cost of acquiring data, especially for smaller firms (Demirer et al., 2024). Therefore, it can be considered an instrument for data as it directly affects firms' willingness to acquire data while it remains uncorrelated to R&D investments.

3.2.2 Complementarity of Data and R&D

A firm rich in data and analytics and performing innovative activities like R&D might combine these intangible assets and gain higher productivity. To assess these potential benefits, we study the potential complementarity of these investments for labour productivity by estimating a model that measures the contribution of investments and firm characteristics to productivity.

So far, we have shown that investing in data can potentially lead to more R&D, and through more data-driven decision-making and innovative production processes, the firm may become increasingly efficient, and labour productivity grows. Previously, Müller, Fay and Vom Brocke (2018) documented that big data and analytics assets are associated with 3-7% improvement in firm productivity, but the effect varies by industry. Bakhshi, Bravo-Biosca and Mateos-Garcia (2014) found that more extensive data use improves firms' productivity, and its combination with other organisational attributes such as organisational structure and innovation contributes positively to firm performance. Longitudinal studies have demonstrated positive returns on R&D investment, with company-financed research yielding higher returns and basic research providing a productivity premium (Lichtenberg and Siegel, 1991). These findings underscore the critical role of data and software in enhancing R&D capabilities and driving innovation and productivity. By combining these insights with those provided previously,

⁶Mathematical models with uncertainty are usually solved using deep-learning via big-data instead of prior unrealistic assumptions on beliefs and information sets, including the work of Ning and You (2019) who introduced a new setting to solve optimisation problems using machine learning. Differently, Bayesian methods incorporate updating beliefs as the available dataset to the agent grows, as demonstrated in Van Oijen (2017).

⁷In some sectors, data and R&D are used slightly differently. For instance, imagine how data and R&D develop differently in industries like pharmaceutical or medical research, where data is both an input and outcome of R&D.

we show that there might be a loop in which firms become data-intensive, get further involved in innovative activities, and become more productive. In the next section, we introduce the dataset and variables corresponding to the elements of the theoretical framework we have developed.

3.3 The Scope of Our Study: Choice of Italy, Data & Variables

This section presents the rationale for selecting Italy as the focus of our study, outlines the dataset used, and details the key variables constructed for our analysis

3.3.1 Benchmarking Italy: Intangible Investments and Productivity

In contrast to evidence on firm-level gains from digitalisation and innovation-based activities, the diffusion of digitalisation and high-tech activities has been simultaneous with the slowdown of aggregate productivity growth, bringing up the so-called productivity puzzle of the 21st century (Acemoglu et al., 2014; Coyle, 2019). Besides lower aggregate investments, the Italian economy has suffered significantly from this slowdown - especially during the Great Recession and European debt crisis (Busetti, Giordano and Zevi, 2015). In addition, its economy has featured one of the lowest proportions of intangible investments out of all investments among advanced economies (Crass and Peters, 2014), potentially driven by low investments in big data and e-commerce and not management technologies or cloud computing (Calvino et al., 2022), despite the recent upward trend in investments in data and software, as shown in Figure 3.1.

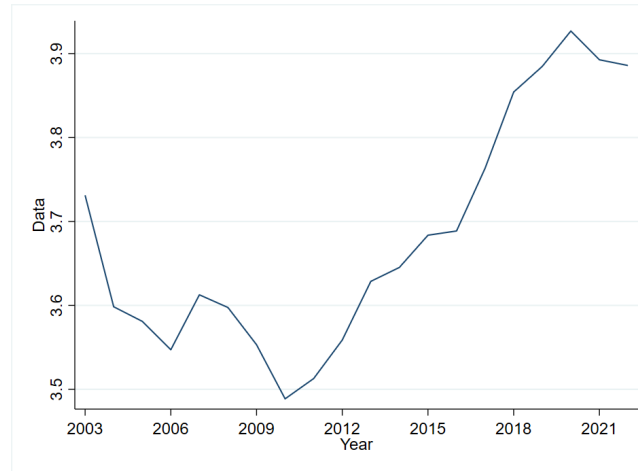


Figure 3.1: Annual average of log of investments in data and software among Italian firms. Data source: INVIND, Bank of Italy

As a result of low intangible asset accumulation, Italy has received particular recommendations from the EU Commission to incentivise investments in innovation and other intangibles. In this direction, tax incentives motivated by these recommendations have contributed to an increase in investments in innovation and research (Briguglio et al., 2019). Interestingly, there is sizeable heterogeneity across the type of assets invested and the firms involved - for instance, larger firms have become more innovative following Impresa 4.0 starting in 2014 (ISTAT, 2018).

Previously, Bontempi and Mairesse (2008) found that intangible assets were at least as productive as tangible assets for Italian manufacturing firms. Also, adopting high-end digital technologies has been found to increase the chances of internationalising Italian manufacturing firms (Ernesto, Linda and Marco, 2016). Therefore, these assets can be an ideal source of growth and productivity gains for

Italian firms as they require lower costs and less external financing (Marrocu, Paci and Pontis, 2012). Therefore, in an economy that has faced decades of low productivity growth and insufficient intangible accumulation, ensuring the presence of a positive role of datafication, R&D, and their combination on firm performance can provide two primary benefits. First, it motivates policy design and academic studies to stress the importance of such activities more than before. Secondly, it leaves enough room for scholars to dig deeper into other potential drivers of the productivity growth decline, for instance, population ageing or other demographics (Maestas, Mullen and Powell, 2023; Kotschy and Bloom, 2023).

3.3.2 Data and Variables

In this work, we use a multi-year version of the firm-level survey dataset “INVIND”⁸ conducted and collected annually by Bank of Italy since early 1980’s. This survey is administered to the manufacturing and services firms with at least 20 employees. Since the 1990s, the INVIND surveys have included questions on firms’ actual and expected values on multiple variables, including turnover, different classes of investments, and capacity utilisation. After the Great Recession of 2007-08, the survey was expanded by including questions on firms’ financing conditions and constraints. Furthermore, the survey comprises year-by-year questions on interesting economic research topics, such as the Russo-Ukrainian war or the energy crises.⁹

We limit our sample to start from 2002 as the number of participating firms surges from 1,500 to around 4,000 this year by including firms in the services sector and expanding to more industries within manufacturing. We observe 11,960 unique firms in our sample between 2002 and 2023, with 90,347 corresponding observations or, on average, 7.55 annual survey responses per firm. We use the data on firms’ characteristics, such as location, industry, number of employees, sales, capacity utilisation, profitability, year of establishment, and other classes of investments. Firms answer directly the amount they spend on data and software acquisition as well as their R&D expenditures - independently from the other¹⁰. However, measuring data investments in other samples can prove challenging for multiple reasons.

First, data is typically an intermediate and experience good, meaning its value is not observable before consumption (information asymmetry in the seller’s favour). Furthermore, the intermediate hand in the market (collectors that sell to final good producers) holds substantial market power (Ichihashi, 2021). Second, data is non-rival (Determann, 2018; Jones and Tonetti, 2020). As a result, there may be social gains to data being used broadly across firms, even with privacy considerations. Fearing creative destruction, firms may choose to accumulate their data, leading to the inefficient use of non-rival data. In addition, data depreciation is not trivial and differs from the normal wear and tear of other assets (OECD, 2020). Instead, its declining value is usually due to data becoming irrelevant or obsolete, even though the old data can produce new values through data fusion (Li, Nirei and Yamana, 2019), bringing a further layer of complexity to the depreciation data. Lastly, a large

⁸“Indagine sugli investimenti delle imprese manifatturiere” (Inquiry into investments of manufacturing firms) (Bank of Italy, 2024)

⁹The dataset is available through an online platform that supports code execution in R and STATA. The output is restricted to text (.txt format), preventing us from analysing variables’ distributional properties, such as histograms or Kernel densities. Additionally, direct access to the data and individual observations is limited due to confidentiality restrictions, preventing us from viewing the variables’ percentiles, minimum and maximum.

¹⁰The explicit terms used for these variables are: “Total expenditure on software & databases and mineral explorations” and “Expenditure on R&D; design and test products”

proportion of data is collected via free services (e.g., Google Ads, Facebook, and search engines). Therefore, recording and monetising free transactions of data has made its measurement as an input for digital firms even more complicated. In our dataset, firms report the values for investments in data directly. Therefore, we expect these values only to include the direct expenditure of firms on data and, hence, represent a conservative (lower bound) of firms' acquisitions of data and software.

We construct firm performance (labour productivity) and data intensity measures using the variables provided in INVIND. We exploit the available insights into firms' sales and capacity utilisation in our dataset, which allows us to calculate a firm's production capacity per labour. In this way, we do not rely on unavailable production data as supply-side labour productivity, and our analyses would be kept safe from purely demand-driven business cycles if we were to use sales per labour.

We can compute the production capacity (Y^{PC}) of the firms using the information on total sales (Y^S) within each year divided by the capacity utilisation ($0 < CU < 1$) reported by the firms:

$$CU_{i,t} \equiv \frac{Y_{i,t}^S}{Y_{i,t}^{PC}} \quad \therefore \quad Y_{i,t}^{PC} = \frac{Y_{i,t}^S}{CU_{i,t}} \quad (3.1)$$

We then calculate labour productivity as production capacity per labour as follows:

$$\text{Lab Prod}_{i,t} = \frac{\log(Y_{i,t}^{PC})}{\log(\text{EMP}_{i,t})} \quad (3.2)$$

To accurately identify the impact of data investments on R&D investments through using data regulation as an instrument, we need to construct a treatment and a control group by their data intensity, e.g., to what extent they would be sensitive to regulation on data privacy as their share of investments in data would determine such sensitivity. Ideally, one would calculate data intensity as the share of investments in data out of all intangible investments. However, our dataset does not offer insights into all intangible investments, such as marketing, goodwill, or copyrights. Therefore, we introduce our identification strategy based on sectoral "data intensity" calculated as:

$$\text{Data Intensity}_{i,t} = \frac{\text{Investment in Data}_{i,t}}{\text{Total Tangible Investments}_{i,t}}$$

This measure normalises the values using total tangible investments instead of total or intangible investments, e.g., measuring the intensity in proportional terms.

In our analyses, we control for various firm characteristics as follows. Initially, we use the log of age, which is calculated as the difference between the observation year and the establishment year. Firms located in central and north of Italy take the dummy North = 1. We also calculate employee growth (Emp Growth) and sales growth (Sales Growth) and use lagged values as a control in our analysis, in line with the literature. The profitability dummy controls for the firms' recent performance and potential internal funding of investments - constructed based on a categorical variable regarding the size of profits and losses. We categorise firms that reported large and moderate profits as profitable, those without profit or losses, and significant losses as non-profitable. We also construct a dummy for observations with higher than 25% of their investments in high tech and assign them as High Tech firms. Lastly, we follow [Mohades et al. \(Forthcoming\)](#) and construct a financial constraints dummy based on the questions included after 2009. Firm-year observations that entirely, partially or by self-selection¹¹ did not receive their required funds, are categorised as financially constrained.

¹¹These firms did not apply for funds as they were sure their request would be rejected.

The descriptive statistics of the variables taken directly from the survey and those that we constructed are presented in Table 3.3.1:

Variable	Mean	SD	N
R&D Investments (in log)	5.15	2.23	19,908
Data and Software Investments (in log)	3.76	2.09	54,926
Sales Growth	-0.01	0.26	70,570
Employees Growth	0.01	0.12	70,583
Age (in log)	3.32	0.74	89,754
Data Intensity	0.75	49.31	81,460
Financial Constraints	0.06	0.24	37,350
Profit	0.68	0.47	83,608
North	0.66	0.48	90,347
High Tech	0.32	0.47	23,270
Lab Prod	6.26	1.55	52,072
Observations per Sector			
Manufacturing			64,934
Services			25,413
All Sample			90,347

Table 3.3.1: Descriptive Statistics

Overall, more than 65% of our firm-year observations are from firms located in the centre and north of Italy. More than 68% have recorded profits, while only 32% have had significant high-tech investments. Furthermore, our firms are, on average, more than 27 years old ($\exp(3.32) = 27.66$). We can also see that while our sample displays an interestingly low standard deviation in R&D and data investments, the data intensity measure features a high variation across our sample.

3.4 Empirical Strategy

We develop our empirical analyses to study the conceptual framework explained in detail in Section 3.2. We discussed our hypothesis that investing in data, software, and analytics allows firms to manage and afford less uncertainty, providing the leeway to redirect resources toward high-risk ventures such as R&D. In this section, we abstract from including all the mediating steps in our analysis and aim to estimate a direct causal relationship from investing in data to investing in R&D. After that, we study the potential complementarity of investing in these intangible assets for labour productivity.

In what follows, we first present our identification strategy and motivation for using GDPR as an instrument for investments in data in Section 3.4.1. Second and in Section 3.4.2, we develop an IV analysis to study the relationship between investments in data and analytics and R&D activities. Lastly and in Section 3.4.3, we explore the potential complementarity between investments in data and R&D by studying whether there are gains to productivity due to combining these two investments.

3.4.1 GDPR as an Instrument for Data

In 2016, the EU introduced the General Data and Privacy Regulation (GDPR), which aims to protect the personal data of all EU residents regardless of the processing location by anonymising the users and hiding any Personally Identifiable Information (PII). PII refers to any information that can identify an individual alone or combined with other sources. The application of GDPR requires member states to adopt the legislation and adapt it to their needs, geographical institutions, and business environment. In Italy, this was done via the data protection process and governance decree.¹² In this framework, personal data is information that directly or indirectly can identify an individual and includes explicitly online identifiers such as IP addresses, cookies, digital fingerprinting and location data (Goddard, 2017).

Given the broad scope and significant potential impacts of GDPR on multiple aspects of business and life, numerous attempts have been made to measure and identify the impact of GDPR on different outcomes in various fields (Johnson, 2023). Firms might want to collect data on user preferences, searches, and even relocations through cookies when data is collected online. With GDPR's emphasis on data minimisation, limited storage and other protective principles, many firms may abstain from purchasing more data as the potential fines and costs set by local member states might exceed the benefit of collecting data, with penalties that can go up to up to 4% of the annual worldwide turnover of the preceding financial year.¹³ In short, companies are required to ask for users' consent before collecting and sharing personal data, and by law, users now have the right to gain access to the personal information collected about them, which can translate into higher costs in collecting and analysing data from the business point of view (Urban et al., 2018). According to Layton and Elaluf-Calderwood (2019), firms face many challenges in complying with costly and complex rules, broad definitions of PII, and increased risk of fees and lawsuits for violations, vulnerabilities, and compliance. On the other hand, GDPR has the potential to provide a transparent and legal framework for firms to deepen their investments in data and analytics due to the global standards being set by the EU (EU Commission, 2020; Voss and Houser, 2019).¹⁴

In this context, GDPR might influence businesses' willingness to gather data, thus indirectly affecting their uncertainty and willingness to invest in high-risk activities like R&D and, consequently, can serve as an instrument for data acquisition. First, if data privacy legislation is driven by factors unrelated to firm-level decisions about R&D investment, such as political or societal concerns, GDPR can be considered an exogenous instrument. Second, if the legislation unintentionally affects R&D due to issues like consumer protection or national security, its exogeneity regarding R&D is plausible. Finally, highlighting the primary policy goals of data privacy laws, such as protecting consumer rights or enhancing cybersecurity, underscores that these laws are not explicitly intended to impact R&D activities directly, thus reinforcing their exogeneity.

Our identification strategy is based on firms' data intensity across industries to construct control and treatment groups based on how GDPR has impacted each sector's data intensity. Previously, the literature has found substantial heterogeneity among sectors in their digitalisation and use of big data (Calvino and Criscuolo, 2019). Sectors with the highest rates of innovation have shown to be more data-intensive (Cong et al., 2022). Therefore, these sectors with high data intensity may be more

¹²Decreto Legislativo 10 Agosto 2018, n. 101

¹³See more details on this penalty [here](#)

¹⁴Scholars in other fields have also studied the effect of GDPR on web privacy (Degeling et al., 2018) and content sharing (Lefrere et al., 2020), anonymisation in legal systems (Bolognini and Bistolfi, 2017) and future technology narrative (Li, Yu and He, 2019), among others.

sensitive to data regulation. On the other hand, sectors that do not use data and analytics will be less affected by any regulation regarding data and privacy. In our identification strategy, sectors relatively unaffected by GDPR are classified as the control group. Suppose the data intensity within a sector remains low before and after a data-related regulation. In that case, it is potentially less specialised in activities that require data or has not adapted data-intensive production methods. To identify the sectors most affected by GDPR, we use the data intensity measure previously introduced in Section 3.3. We average data intensity at the sector level pre and post-regulation years (2016 for European and 2018 for Italian regulation announcements). This process consists of first calculating the average data intensity in each sector in each year as:

$$\text{Data Intensity}_{s,t} = \frac{1}{N_s} \sum_{i=1}^{N_s} \text{Data Intensity}_{i,t} \quad (3.3)$$

Second, we construct these averages for years pre and post-GDPR (2016) and its implementation in Italy (2018). In this case, our sample is splitted in two parts for each year-criterion:

$$\text{Data Intensity}_{s,t \geq \tau} = \frac{1}{T} \sum_{i=1}^T \text{Data Intensity}_{s,t \geq \tau} \quad \text{for } \tau = 2016, 2018 \quad (3.4)$$

These averaged data intensity measures capture the average data intensity in each sector in the years before and after GDPR (2016) and the Italian adaptation of this regulation (2018). We then use the following rule of thumb to identify sectors as data-intensive, namely if their data intensity doubled pre- and post-2016 (EU-level GDPR legislation) and 2018 (Italian decree on data):

$$\begin{aligned} &\text{if } \text{Data Intensity}_{s,t > 2016} \geq 2 \times \text{Data Intensity}_{s,t < 2016} \\ &\& \text{Data Intensity}_{s,t > 2018} \geq 2 \times \text{Data Intensity}_{s,t < 2018} \\ &\quad \text{Sector falls in the treatment group} \end{aligned}$$

Table 3.4.1 below presents each sector's sector-level average of data intensity pre and post-2016 and 2018. The last column indicates whether the sector falls into a treatment (T) or control (C) by using our rule of thumb above. Furthermore, it displays the values for manufacturing and services in each subsample and across the whole sample. Importantly, we can see that the two column-pairs exhibit similar patterns in the data intensity of sectors.

Sectors	Pre-2018	Post-2018	Pre-2016	Post-2016	All Sample	Result
Manufacturing	0.21	1.45	0.21	1.13	0.53	-
1 - Food industries, beverages and tobacco products	0.12	0.07	0.12	0.07	0.10	C
2 - Textiles, clothing, and hide, leather and footwear products	0.22	0.50	0.21	0.42	0.28	T
3 - Coke manufacturing, chemical industry, rubber and plastics	0.11	0.19	0.11	0.16	0.13	C
4 - Processing of non-metallic minerals	0.11	0.08	0.12	0.08	0.10	C
5 - Metal engineering industry	0.22	0.31	0.22	0.27	0.24	C
6 - Other manufacturing industries	0.26	0.23	0.26	0.22	0.23	C
7 - Other industries excluding construction (energy and extraction)	0.71	18.41	0.60	14.21	6.63	T
Services	0.58	3.24	0.57	2.63	1.33	-
8 - Wholesale and retail commerce	0.31	3.24	0.30	2.61	1.09	T
9 - Hotels and restaurants	0.12	0.12	0.13	0.11	0.12	C
10 - Transport and communications	0.66	4.46	0.65	3.57	1.88	T
11 - Real estate activities, etc.	1.20	2.40	1.18	2.12	1.49	C
All Industries	0.31	1.97	0.30	1.56	0.75	-

Table 3.4.1: Sector-Level Averages of Data Intensity, Taken From INVIND Database in Annual Frequency for 2002-2022 - Divided by Pre and Post GDPR and Decreto Legislativo 10 Agosto 2018, n. 101 .

Table 3.4.1 above presents each sector's average data intensity pre and post-2016 and 2018. The last column indicates whether the sector falls into a treatment (T) or control (C) by using our rule of thumb above. Furthermore, it displays the values for manufacturing and services in each subsample and across the whole sample.

We note that sectors 2,7,8 and 10 were impacted according to our rule of thumb and are grouped as treatment (T). Hence, we take other sectors as our control since they should not be sensitive to a policy that concerns data acquisitions. Our finding is of particular interest. In sectors 8 and 10 of Table 3.4.1, we see that wholesale and retail commerce and transport and communications sectors have had a much larger data intensity, especially after the implementation of GDPR. Our finding is similar to [Haddara, Salazar and Langseth \(2023\)](#), who find substantial impacts of GDPR on the retail sector - mainly e-commerce - in the Netherlands.

We use the implementation of GDPR in Italy in 2018 as our baseline specification and define the GDPR dummy applied to the highly data-intensive sectors post (and including) 2018. This choice does not take the anticipation effect of EU-level legislation away, but it focuses on the application of the directive in Italy. Our identification strategy leaves us with a treatment group of 7,871 observations and 82,476 in the control group. Lastly, we see that not only was the services sector more likely to be affected by the policy, but it also represented higher data intensity across all subsamples compared to the manufacturing sector. After all, the difference in the values in the post and pre-GDPR years can result from low investments in data at the beginning of our sample. Table A.2 in Appendix A shows that even removing the first ten years of our sample does not dramatically change our results from the table above, and we can take the same sectors as control and treatment even in a sample of 2012-2023. Furthermore, Table A.1 in Appendix A shows that our treatment sectors featured seemingly higher raw values for investments in data across the whole sample. Therefore, rather than the textiles and clothing sector (group 2), higher data intensity in the treatment sectors cannot be driven by meagre investments in tangible assets; instead, it can be driven by higher investments in data.

We also calculate the R&D intensity as the proportion out of total tangible investments - similar to our measure of data intensity and find that the pre-2016 era hints at higher R&D intensity in the manufacturing sector than in services, see Table A.3 in Appendix A. However, since 2016, a sharp positive trend in the R&D intensity of firms in the services sector can be observed, primarily driven by firms in transport, communications, and real estate activities. Therefore, the higher intangible asset intensity of the Italian economy is primarily driven by the firms in the services sector that also act more in line with the pillars of digitalisation and datafication.

3.4.2 More data, more R&D?

In this section, we assess a potential causal relationship between investments in data and R&D as also described in Sections 3.1 & 3.2.1. To address potential endogeneity issues discussed in Section 3.1 and establish a robust causal relationship, we employ an instrumental variable (IV) setting using GDPR as an instrument for data acquisition. This IV approach helps to mitigate biases in the relationships between variables and provides more credible estimates.

In our case, the instrument (Z) is that the firm is in a sector highly exposed to GDPR, as defined in the previous section. The treatment (T) is the firms' investments in data and analytics, and the outcome (Y) is the R&D investment. Hence, this analysis tests the potential causality of investing more in data, resulting in higher R&D.

We follow intuitions from Angrist, Imbens and Rubin (1996); Angrist (2022) and construct a two-step control function model similar to a 2SLS. The first step is to predict the values of investments in data using a GDPR dummy, fixed effects ($\mu_i + \lambda_t$) and a small set of control variables ($X_{i,t}$). These controls are the sales growth and employee growth (Grazzi, Jacoby and Treibich, 2016; Farboodi et al., 2019), lagged one year and firm's age (Fortune and Shelton, 2014). Formally:

$$\text{Data}_{i,t} = \gamma \text{GDPR}_{s,2018} + \mu_i + \lambda_t + X_{i,t}\Lambda + u_{i,t} \quad (3.5)$$

Where μ_i and λ_t are firm and time-fixed effects, respectively. This estimation yields the predicted values for data as

$$\widehat{\text{Data}}_{i,t} = \gamma \text{GDPR}_{s,2018} + X_{i,t}\Lambda + \mu_i + \lambda_t \quad (3.6)$$

We now insert the predicted values in our main regression equation featuring the sector-fixed effects (ψ_s) and run separate regressions for $\tau = 0, 1, 2, 3, 4, 5$ to assess the impact of investments in data on R&D over time as follows:

$$\text{R\&D}_{i,t} = \alpha \widehat{\text{Data}_{i,t-\tau}} + X_{i,t}\beta + \psi_s + \varepsilon_{i,t} \quad (3.7)$$

In the second stage regression, we extend our analysis and control for firms' lagged high-tech investments as they provide leeways to more investments in other intangibles (Schivardi and Schmitz, 2019). Furthermore, we control for financial constraints since they can curb firms' investments in R&D. We also control whether large firms or those in the North drive the majority of R&D investments. The regression results are provided in Section 3.5.2.

3.4.3 Estimating Complementarity of Data and R&D Investments for Productivity

So far, we discussed how to establish whether firms with high data investments invest more in R&D. However, to understand whether combining these two investments contributes positively to the firm's productivity, we need to introduce a different setting in line with the framework described in Section 3.2.2. Our dependent variable of interest is the production capacity per labour as a proxy for productivity. However, we are interested in both the within-period (year) effect and the lagged effect since the use and combination of these investments may take more than one year to affect firms' performance (Capasso, Treibich and Verspagen, 2015).

Therefore, our estimation equation of interest is as follows:

$$\text{Lab Prod}_{i,t} \equiv \frac{\log(Y_{i,t}^{PC})}{\log(\text{EMP}_{i,t})} = \alpha_1 \text{Data}_{i,t-\tau} + \alpha_2 \text{R\&D}_{i,t-\tau} + \alpha_3 \text{Data}_{i,t-\tau} \times \text{R\&D}_{i,t-\tau} + X_{i,t}\beta + \psi_s + \varepsilon_{i,t} \quad (3.8)$$

The functional form above includes the interaction between R&D and data investments and runs separate regressions for $\tau=0,1,2,3,4,5$. The positive coefficient of the interaction term (α_3) will be a sign of potential complementarity and the presence of complementarity effect. $X_{i,t}$ includes controls such as all other investments, age, North and large dummies to control for the frontier firms' high productivity and lagged high tech investments.

3.5 Results

This section presents the results from the empirical settings developed in Section 3.4. First and in Section 3.5.1, we display the results of our first-stage regression. Next and in Section 3.5.2, we present the results on the relationship between investing in data and R&D expenditures. We lastly move to Section 3.5.3 and emphasise the findings on the potential complementarity between R&D and data investments to labour productivity. Lastly, we show that these gains vary over short and long-term horizons.

3.5.1 First-stage Regression: GDPR and Investments in Data

Table 3.5.1 provides results on our first-stage regression (Equation 3.5) introduced in Section 3.4.2. We find that controlling for firm and time-fixed effects, GDPR is positively correlated with investments in data:

Dependent Variable: Log of Data Investments	
$GDPR_{2018} = 1$	0.205*** (0.032)
Sales Growth $_{t-1}$	0.136*** (0.029)
Age	0.004*** (0.001)
Emp Growth $_{t-1}$	0.332*** (0.063)
Constant	3.598*** (0.034)
Firms	6,498
Observations	34,941
Time FE	✓
Firm FE	✓
Treated Sectors: Energy and extraction - Wholesale and retail commerce - Textiles, clothing, and hide, leather and footwear products - Transport and communications	

Table 3.5.1: Regression Results for the first-stage regression in Equation 3.5

*** Significant at 99% CI, ** at 95% CI, * at 90% CI

This result can be due to more transparency for firms in acquiring data and the legal and judiciary frameworks they would operate within, as emphasised by [EU Commission \(2020\)](#). This - somewhat counterintuitive - result can be of interest in studying firms' heterogeneity in perceiving privacy regulation as suggested by [Taufick \(2021\)](#). Previously, it has been shown that smaller firms may be more affected negatively by regulations that affect the business environment ([Schiffer and Weder, 2001](#)). In our analysis, where we study the impact of GDPR among larger firms represented in INVIND, more regulation might reduce the risk of future issues with the judiciary system, media, and large data in-

intermediary firms. Furthermore, these large firms have likely internalised these risks prior to regulation and have already put better data governance practices in effect and, hence, after regulation, gain more market power (Sabatino and Sapi, 2022). In short, a transparent legal system may benefit larger firms as they can fit all their processes within the framework with lower costs and be less uncertain about the data they collect, buy or generate.

However, our analysis shows that the large manufacturing Italian firms - who constitute our sample - acquired more data after implementing GDPR. This first stage regression shows that GDPR might have offered lower judiciary costs and higher transparency, reducing the risk of future privacy-related concerns when holding clients' data, in line with EU Commission (2020).

However, our study abstains from studying the direct causality of GDPR on data investments. This means that our capture is unable to capture the causal impact of this regulation on firms' acquisitions of data and software. Consequently, caution is needed regarding the empirical strategy used to identify the critical mechanisms between GDPR and investments in data. For instance, a difference-in-difference (DiD) strategy can more rigorously identify GDPR's impact on firms' data investments, as our result may be affected by a quadratic trend in data investments. Furthermore, these approaches can better identify the impact of data and privacy regulations on data investments rather than studying the relationship as we did in this section.

3.5.2 Second stage regression - Data and R&D

Table 3.5.2 presents the results from estimating the second-stage regression in Equation 3.6. We find support to our hypotheses introduced in Sections 3.2.1 & 3.4.2: more investments in data and analytics lead to higher expenditures in R&D. It is also clear that a weak negative relationship between employees' growth and R&D investments exists. However, profits in the previous year seem to contribute to ongoing R&D investments positively. Larger firms and those operating in the north of Italy are more active in R&D investments, and those with larger proportions of high-tech-related investments also were detected to be more actively engaging in research and development. Our results also confirm the finding of IMF (2015); Altomonte et al. (2022) that financial constraints and poor credit access have remained problematic barriers for private investments in Italy, especially among intangible assets. However, across a small sample including additional controls - especially the High Tech dummy with few observations, the positive effect of data disappears; see regressions (5) and (6) below.

Dependent Variable: Log of R&D Investments						
Instrument for Data Investments: GDPR Implementation 2018						
	(1)	(2)	(3)	(4)	(5)	(6)
Log of Data Investments	0.922*** (0.242)	1.543*** (0.247)	1.699*** (0.246)	1.686*** (0.246)	-0.160 (0.305)	-0.159 (0.305)
Sales Growth _{t-1}	-0.216*** (0.058)	-0.297*** (0.059)	-0.310*** (0.059)	-0.308*** (0.059)	-0.045 (0.080)	-0.046 (0.080)
Age	0.006*** (0.002)	0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	0.003 (0.002)	0.003 (0.002)
Emp Growth _{t-1}	0.091 (0.137)	-0.121 (0.138)	-0.311* (0.139)	-0.310* (0.139)	-0.029 (0.181)	-0.030 (0.181)
Profit _{t-1}	0.116*** (0.029)	0.122*** (0.029)	0.110*** (0.029)	0.106*** (0.029)	0.069 (0.042)	0.069 (0.042)
Large			1.065*** (0.049)	1.065*** (0.049)	1.096*** (0.068)	1.096*** (0.068)
North			0.377*** (0.069)	0.373*** (0.069)	0.460*** (0.093)	0.460*** (0.093)
Fin Cons				-0.090* (0.042)		-0.018 (0.070)
High Tech _{t-1}					0.093** (0.033)	0.093** (0.033)
Constant	0.855 (0.874)	-2.100* (0.894)	-3.170*** (0.890)	-3.106*** (0.891)	3.928*** (1.107)	3.925*** (1.107)
Sector FE		✓	✓	✓	✓	✓
Firms	3,554	3,554	3,552	3,552	1,902	1,902
Observations	13,276	13,276	13,233	13,233	5,173	5,173

Treated Sectors: Energy and extraction - Wholesale and retail commerce - Textiles, clothing, and hide, leather and footwear products - Transport and communications

Table 3.5.2: Regression Results For the second-stage regression in Equation 3.7

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

By combining our results from this section and Section 3.5.1, we can establish the empirical results of the relationship between GDPR, investing in data and R&D. So far, we have shown that GDPR years come with firms' larger investments in data and analytics, and more data and analytics have increased firms "within-year" R&D investments. Lastly, we present our results for $\tau > 0$ in Equation 3.7 in Figure 3.2:

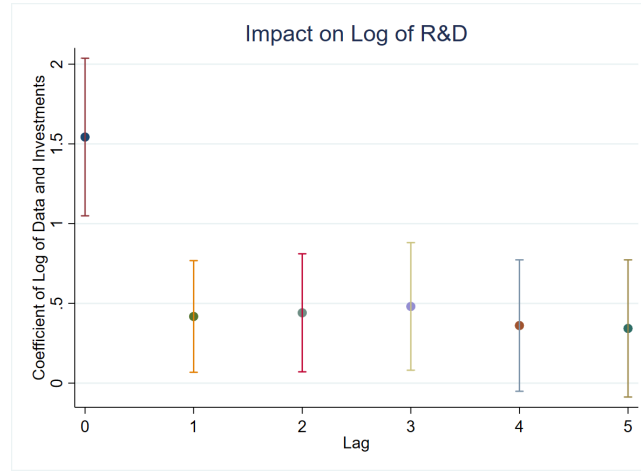


Figure 3.2: The effect of data on R&D over time for: $\tau = 0, 1, 2, 3, 4, 5$ in Equation 3.7

We document that data's simultaneous effect on R&D is significantly larger than the lagged effects up to 5 years. Furthermore, it is only significantly positive for up to three years and then becomes insignificant. This result should motivate policymakers to encourage more investments in data and provide easier data sharing via B2B data transfer, as it can increase firms' engagement in innovative activities in the short term. The short-lasting effect of these investments can also hint at a depreciation or obsolescence feature of data as an asset. Investments in data need to be persistent enough to affect firms' innovation activities continuously. The regression results for $\tau > 0$ can be found in Table A.5.

3.5.3 Complementarity Effects

The estimation of Equation 3.8 provides key insights into the potential complementarity of data and R&D for labour productivity. Our findings in Table 3.5.3 show that both R&D and data investments positively contribute to productivity. In addition, their interaction displays a positive contribution to productivity in all specifications of (1) to (6), confirming our hypothesis on complementarity in Sections 3.2.2 & 3.4.3. It is important to note that these effects reported in Table 3.5.3 are within the year when $\tau = 0$: investing in data and R&D in any given year leads to low but significantly positive labour productivity gains.

Furthermore, older, large firms and those based in the north are more productive, and investments in high-tech-related resources do not display simultaneous productivity gains when we account for investments in R&D and data. Our results align with Kumbhakar et al. (2012): firms that invest more in high tech gain the most by combining data and R&D - even if their productivity might be lower on average due to early-stage market presence. Crass and Peters (2014) found that intangible assets are usually complementary, but they obtained mixed results when studying the combinations of R&D and other intangible assets. Our results identify an important complementary intangible asset to R&D. Data acts as information by reducing firms' uncertainty about markets. Firms can afford to bear more innovation uncertainty and engage more in R&D by freeing up space in a portfolio of various uncertainties. Combining these two assets may increase efficiency and thus improve labour productivity.

Dependent Variable: Labour Productivity						
	(1)	(2)	(3)	(4)	(5)	(6)
Log of Data Investments	0.045*** (0.007)	0.046*** (0.007)	0.042*** (0.008)	0.040*** (0.008)	0.019* (0.008)	-0.018 (0.015)
Log of R&D Investments	0.033*** (0.006)	0.038*** (0.006)	0.024*** (0.006)	0.027*** (0.006)	0.013* (0.006)	0.008 (0.011)
Log of Data Investments × Log of R&D Investments	0.003** (0.001)	0.003** (0.001)	0.003* (0.001)	0.003* (0.001)	0.005*** (0.001)	0.013*** (0.002)
Log of Other Investments			0.104*** (0.004)	0.101*** (0.004)	0.093*** (0.004)	0.100*** (0.006)
Age					-0.008 (0.013)	0.100*** (0.024)
North					0.534*** (0.041)	0.447*** (0.056)
Large					0.554*** (0.021)	0.647*** (0.034)
High Tech _{t-1}						-0.029* (0.013)
Constant	5.954*** (0.037)	6.236*** (0.061)	5.513*** (0.046)	5.774*** (0.064)	5.295*** (0.078)	4.802*** (0.129)
Sector FE		✓		✓	✓	✓
Firms	3,127	3,127	2,607	2,607	2,601	1,499
Observations	12,333	12,333	10,046	10,046	9,998	3,964

Table 3.5.3: Regression Results of Equation 3.8 with Random Effects and $\tau = 0$

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

The regression results for $\tau > 0$ are provided in Appendix B. Below and in Figure 3.3, we provide a graphical representation of the coefficients of lagged investments. These graphs provide insight into the effect of investments in each of the data and R&D and their complementarity effect over time:

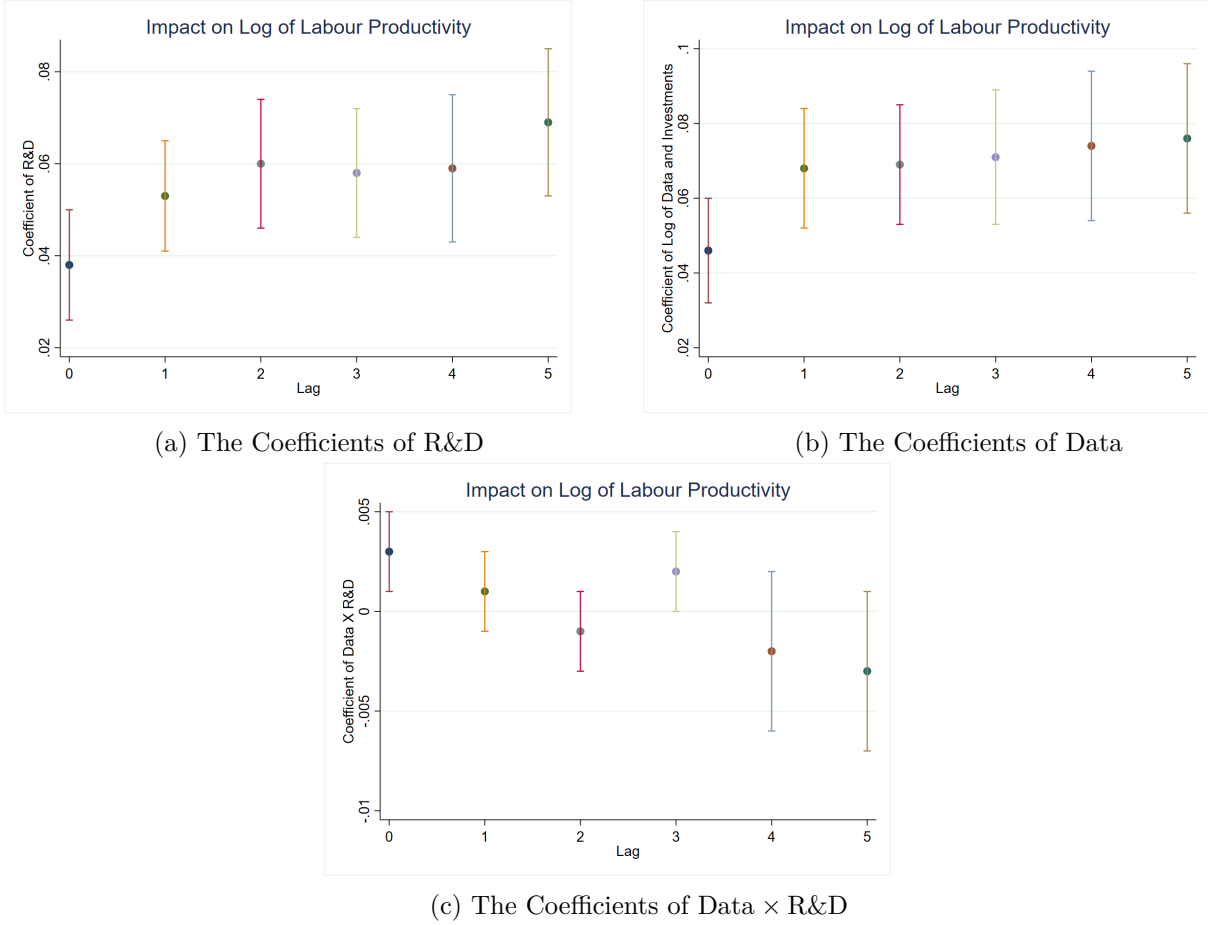


Figure 3.3: Coefficients for: $\tau = 0, 1, 2, 3, 4, 5$ in Equation 3.8

We can observe that the returns of R&D are larger after one period, similar to the findings of [Aydin, Alrajhi and Jouini \(2018\)](#); [Guellec and Van Pottelsberghe de la Potterie \(2004\)](#). Similarly, the productivity gains of investing in data slightly increase over time. However, combining these two investments yields a decreasing impact in the first two years ($\tau = 0, 1, 2$), then increases for one year and returns to smaller values afterwards. Therefore, the multiplier effect is present and statistically significant, yet it is reasonably small compared to the sole impact and does not increase over time. However, this analysis shows that analysing solely within-year effects might be incomplete as the firms' labour productivity continues to be positively affected over time as they invest in data and R&D and combine them.

3.6 Robustness Checks

As robustness checks, we analyse whether the inclusion or exclusion of fixed effects can affect our analysis. We then study whether our rule of thumb regarding dividing sectors into control and treatment drives our results by dividing our sample into control and treatment by services and manufacturing, respectively. We also show that our results will be unaffected if we use the European timeline of GDPR (in 2016 instead of 2018) as an instrument. The next robustness check includes repeating our analysis by using data and R&D per labour and studying whether normalising the investments per labour affects our main takeaways. We also estimate our model for the multiplier effect by using our identification strategy to predict the data investment values by GDPR and run the first-stage regression before Equation 3.8 to address any potential endogeneity among our independent

variables. We also extend our analysis and study whether GDPR affected the existing firms before its introduction by estimating our model pre and post-GDPR.

3.6.1 Inclusion of Fixed Effects

Our baseline specification in studying the impact of data investments on R&D in Table 3.5.2 does not include a time or firm-fixed effect. One might speculate that these results may be driven by firms' time-invariant characteristics or omitted variable bias. For instance, a few large firms in northern Italy might drive the results. Furthermore, our result can be an outcome of a positive trend in both data investments and R&D. In Table C.1, we present the results with time and firm-fixed effects. We can observe that even though the magnitude of the impact of data investments on R&D is reduced, the sign and significance of this coefficient remain robust. More data means more R&D, even after controlling for multiple firm characteristics and multi-dimensional fixed effects.

Table C.5 attempts to scrutinise our results of the complementarity effect analysis presented in 3.4.3 by adding time and fixed effects. In summary, firms that decide to invest in data may have unobservable characteristics that affect productivity, such as long-term good quality management or unaccounted human capital. If these characteristics are not controlled for, it can lead to selection bias. However, we find that our take on the complementarity between these investments is robust to adding multi-dimensional fixed effects.

3.6.2 Manufacturing vs Services

Our analysis also accounts for the bias caused by using the rule of thumb in Table 3.4.1 and how it may be only driven by a few sectors. Therefore, we modify our analysis by using the manufacturing sector as the control group and services as treatment, as the services have been mainly impacted and were more data intensive compared to manufacturing firms across the whole sample. The results of regressions are provided in Table C.2 for $\tau = 0$. We find that changing treatment and control groups to broader sectors does not affect our main takeaways from the analyses.

3.6.3 GDPR 2016 instead of 2018

We also suspect whether the legal adaptation of GDPR was expected by firms in 2018, as a European framework was introduced two years before, in 2016. Therefore, firms might adjust their behaviour pre-shock. We extend our analysis in Table C.3 by using the European announcement of GDPR in 2016 as an instrument - instead of the application in Italy in 2018. We find that our instrumental variable analysis remains robust to this specification as well, and more investments in data have had a simultaneous positive effect on R&D investments.

3.6.4 Data and R&D Per Labour

We are also aware that the intensity degree of investing in data and software and innovative activities such as R&D can be a potentially more rigorous measure than the raw values in the log scale. To understand whether our results might be impacted by our choice of working with raw values rather than intensity per labour, we repeat our analyses but use the values of investments in data and R&D per unit of labour each year. The results of these analyses are provided in Tables C.4 and C.6. We show that the causality direction from data to R&D holds firmly, but our results regarding the multiplier effect display mixed intuitions. It also indicates that the impact of R&D per labour

diminishes - at least in $\tau = 0$. However, we understand that future research can address these issues by potentially studying the long-term effect or using different model specifications to uncover such mixed outcomes.

3.6.5 Estimating Complementarity via Instrumental Variable

While both investments in data and R&D appear as independent variables in analysis in Section 3.4.3, one may raise the issue that not addressing the endogeneity of data to R&D can overestimate their role in labour productivity. Therefore, we use the same approach in Section 3.4.2 in implementing GDPR as an instrument for investments in data and then use the predicted values of the first stage in Equation 3.5 into Equation 3.8 for $\tau = 0$. In this way, we check whether our baseline specification overestimates the impact of independent variables on labour productivity. The results of this application are provided in Table C.7. We show that investments in data display positive contributions to labour productivity, while the evidence on R&D's contribution and the multiplier effect is weak and has a vague direction.

3.6.6 Pre-Post GDPR Subsample Study

In this section, we extend our analysis to understand whether the effect of investing in data on R&D and their multiplier effect differ pre and post-GDPR. The results of these analyses are provided in Tables C.8 & C.9. We can see that the impact of data on R&D and productivity might have been lowered post GDPR. However, in the future, with more waves of the survey available, one can measure these impacts more comparably, as our post-GDPR sample is smaller than the pre-GDPR sample.

3.7 Conclusion

In this paper, we shed light on the complex relationship between data and R&D investments and their joint impact on labour productivity using the firm-level survey data of the Bank of Italy. Our findings show the significance of data investments in nurturing innovation and enhancing firm performance through R&D activities. The empirical evidence presented in this paper demonstrates that investments in data and analytics lead to increased R&D activity within firms, thereby contributing to their productivity. We also found that introducing the GDPR may have positively affected firms' data investments. This result suggests that regulation may have increased transparency of the legal framework, encouraging larger firms to invest more in data. Our research also suggests a complementarity effect of combining data and R&D investments for labour productivity. This indicates that firms benefit from investing in data in the short and long term, emphasising the importance of a holistic approach to investment strategies in the era of data-driven decision-making.

These findings carry significant implications for academic research and policy-making, highlighting the need for a comprehensive and nuanced approach to policy design, considering the complex interplay between data, R&D, productivity, and sector-specific dynamics. Firstly, our research highlights the potentially positive impact of the GDPR on firms' data investments. This suggests that a clear and transparent legal framework can encourage firms, especially larger ones, to invest more in data and analytics. Policymakers can leverage this insight to design regulations that encourage data investments, thereby driving innovation and productivity growth. Secondly, the significant positive correlation between data and R&D investments and labour productivity underscores the importance

of a comprehensive investment strategy. Policymakers should, therefore, promote policies that encourage firms to invest in both data and R&D, as these investments appear to have a multiplier effect on productivity. Third, a potentially important issue is the creation of incentives for B2B data sharing, which would allow the distribution of the benefits of data investments and avoid detrimental market concentration in data-intensive markets and barriers to entry for small players.

Future research could delve deeper into the mechanisms through which data regulation impacts firms' investment decisions. Management quality, human capital, IT infrastructure and other firm characteristics can indicate significant heterogeneity in the perception of privacy regulation and optimising data use in encouraging R&D and improving productivity. Additionally, the role of different factors influencing the relationship between data, R&D investments, and labour productivity could be explored. This would further enrich our understanding of the dynamics at play in this complex and rapidly evolving field.

References

- Acemoglu, Daron, David Autor, David Dorn, Gordon H. Hanson, and Brendan Price.** 2014. “Return of the Solow Paradox? IT, Productivity, and Employment in US Manufacturing.” American Economic Review, 104(5): 394–99.
- Akerman, Anders, Ingvil Gaarder, and Magne Mogstad.** 2015. “The skill complementarity of broadband internet.” The Quarterly Journal of Economics, 130(4): 1781–1824.
- Altomonte, Carlo, Peter Bauer, Alberto Maria Gilardi, and Chiara Soriolo.** 2022. “Intangible assets, industry performance and finance during crises.” Baffi Carefin Centre Research Paper, , (2022-173).
- Angrist, Joshua D.** 2022. “Empirical strategies in economics: Illuminating the path from cause to effect.” Econometrica, 90(6): 2509–2539.
- Angrist, Joshua D, Guido W Imbens, and Donald B Rubin.** 1996. “Identification of causal effects using instrumental variables.” Journal of the American statistical Association, 91(434): 444–455.
- Añón Higón, Dolores, Jaime Gómez, and Pilar Vargas.** 2017. “Complementarities in innovation strategy: do intangibles play a role in enhancing firm performance?” Industrial and Corporate Change, 26(5): 865–886.
- Aronica, Martina, Davide Piacentino, and Pasquale Massimo Picone.** 2022. “Big data analytics and internationalisation in Italian firms.” Developments in Entrepreneurial Finance and Technology, , ed. David Audretsch, Maksim Belitski, Nada Rejeb and Rosa Caiazza, Chapter 4, 67–81. Edward Elgar Publishing.
- Arrighetti, Alessandro, Fabio Landini, and Andrea Lasagni.** 2014. “Intangible assets and firm heterogeneity: Evidence from Italy.” Research Policy, 43(1): 202–213.
- Arrow, Kenneth Joseph.** 1972. Economic welfare and the allocation of resources for invention. Springer.
- Aydin, Necati, Ahmed Naser Alrajhi, and Jamel Habib Jouini.** 2018. “Estimating The Impact Of R&D Spending On Total Factor Roductivity For OECD Countries: Pooled Mean Group Approach.” The Journal of Developing Areas, 52(2): 159–168.
- Bakhshi, Hasan, Albert Bravo-Biosca, and Juan Mateos-Garcia.** 2014. “The analytical firm: Estimating the effect of data and online analytics on firm performance.” Nesta Working Paper.
- Bank of Italy.** 2024. “Survey on Industrial and Service Firms, 2002-2023.” Data retrieved from Bank of Italy, Source: Bank of Italy.
- Björkdahl, Joakim.** 2020. “Strategies for digitalization in manufacturing firms.” California management review, 62(4): 17–36.
- Björkdahl, Joakim.** 2020. “Strategies for Digitalization in Manufacturing Firms.” California Management Review, 62(4): 17–36.

- Blackburn, Michael, Jeffrey Alexander, J David Legan, and Diego Klabjan.** 2017. “Big Data and the Future of R&D Management: The rise of big data and big data analytics will have significant implications for R&D and innovation management in the next decade.” Research-Technology Management, 60(5): 43–51.
- Bloom, Nicholas, Steven J Davis, Lucia S Foster, Scott W Ohlmacher, and Itay Saporta-Eksten.** 2022. “Investment and Subjective Uncertainty.” National Bureau of Economic Research Working Paper 30654.
- Bloom, Nick.** 2007. “Uncertainty and the Dynamics of R&D.” American Economic Review, 97(2): 250–255.
- Bolognini, Luca, and Camilla Bistolfi.** 2017. “Pseudonymization and impacts of Big (personal/anonymous) Data processing in the transition from the Directive 95/46/EC to the new EU General Data Protection Regulation.” Computer law & security review, 33(2): 171–181.
- Bontempi, Maria Elena, and Jacques Mairesse.** 2008. “Intangible capital and productivity: an exploration on a panel of Italian manufacturing firms.” National Bureau of Economic Research.
- Briguglio, Damiano, Lazaros Dimitriadis, Virginia Maestri, and Gianluca Papa.** 2019. “Private Investment in Italy.” European Economy Discussion Papers.
- Brynjolfsson, Erik, Wang Jin, and Xiupeng Wang.** 2023. “Information technology, firm size, and industrial concentration.” National Bureau of Economic Research.
- Busetti, Fabio, Claire Giordano, and Giordano Zevi.** 2015. “Main Drivers of the Recent Decline in Italy’s Non-Construction Investment.” Bank of Italy Occasional Paper, , (276).
- Calvino, Flavio, and Chiara Criscuolo.** 2019. “Business dynamics and digitalisation.”
- Calvino, Flavio, Stefano DeSantis, Isabelle Desnoyers-James, Sara Formai, Ilaria Goretti, Silvia Lombardi, Francesco Manaresi, and Giulio Perani.** 2022. “Closing the Italian digital gap: The role of skills, intangibles and policies.”
- Capasso, Marco, Tania Treibich, and Bart Verspagen.** 2015. “The medium-term effect of R&D on firm growth.” Small Business Economics, 45: 39–62.
- CFOTech.** 2023. “Data reduces uncertainty and drives better decisions .” Shannon Williams. Available at: <https://cfotech.com.au/story/data-reduces-uncertainty-and-drives-better-decisions>, Accessed: 31 July 2023.
- Cirillo, Valeria, Lucrezia Fanti, Andrea Mina, and Andrea Ricci.** 2023. “New digital technologies and firm performance in the Italian economy.” Industry and Innovation, 30(1): 159–188.
- Cong, Lin William, Wenshi Wei, Danxia Xie, and Longtian Zhang.** 2022. “Endogenous growth under multiple uses of data.” Journal of Economic Dynamics and Control, 141: 104395.
- Corrado, Carol, Chiara Criscuolo, Jonathan Haskel, Alexander Himbert, and Cecilia Jona-Lasinio.** 2021. “New evidence on intangibles, diffusion and productivity.”
- Coyle, Diane.** 2019. “Do-it-yourself Digital: the Production Boundary, the Productivity Puzzle and Economic Welfare.” Economica, 86: 750–774.

- Crass, Dirk, and Bettina Peters.** 2014. “Intangible assets and firm-level productivity.” ZEW-Centre for European Economic Research Discussion Paper, , (14-120).
- Cuervo-Cazurra, Alvaro, and C Annique Un.** 2010. “Why some firms never invest in formal R&D.” Strategic Management Journal, 31(7): 759–779.
- Degeling, Martin, Christine Utz, Christopher Lentzsch, Henry Hosseini, Florian Schaub, and Thorsten Holz.** 2018. “We value your privacy... now take some cookies: Measuring the GDPR’s impact on web privacy.” arXiv preprint arXiv:1808.05096.
- Delbecque, Vincent, Ahmed Bounfour, and Andrés Barreneche.** 2015. “Intangibles and Value Creation at the Industrial Level: Delineating Their Complementarities.” Intangibles, Market Failure and Innovation Performance, , ed. Ahmed Bounfour and Tsutomu Miyagawa, 27–56. Cham:Springer International Publishing.
- Del Monte, Alfredo, and Erasmo Papagni.** 2003. “R&D and the growth of firms: empirical analysis of a panel of Italian firms.” Research policy, 32(6): 1003–1014.
- Demirer, Mert, Diego J Jiménez Hernández, Dean Li, and Sida Peng.** 2024. “Data, privacy laws and firm production: Evidence from the GDPR.” National Bureau of Economic Research.
- Determann, Lothar.** 2018. “No one owns data.” Hastings LJ, 70: 1.
- Eeckhout, Jan, and Laura Veldkamp.** 2022. “Data and market power.” National Bureau of Economic Research.
- Ennen, Edgar, and Ansgar Richter.** 2010. “The whole is more than the sum of its parts—or is it? A review of the empirical literature on complementarities in organizations.” Journal of Management, 36(1): 207–233.
- Ernesto, Cassetta, Meleo Linda, and Pini Marco.** 2016. “The role of digitalization in the internationalization of Italian manufacturing firms.” L’industria, , (2): 305–328.
- EU Commission.** 2020. “Two years of the GDPR: Questions and answers.” Accessed: 21 December 2023.
- Farboodi, Maryam, Roxana Mihet, Thomas Philippon, and Laura Veldkamp.** 2019. “Big data and firm dynamics.” Vol. 109, 38–42, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.
- Farrington, Ted, and Azar Alizadeh.** 2017. “On the Impact of Digitalization on R&D: R&D practitioners reflect on the range and type of digitalization’s likely effects on R&D management.” Research-Technology Management, 60(5): 24–30.
- Fortune, Annetta, and Lois M Shelton.** 2014. “Age Matters: Disentangling the Effect of R&D Investment in the Global Chemical Products Industry.” Business Management Dynamics, 3(11).
- Frankel, Alexander, and Emir Kamenica.** 2019. “Quantifying Information and Uncertainty.” American Economic Review, 109(10): 3650–80.
- Fung, Michael K.** 2006. “R&D, knowledge spillovers and stock volatility.” Accounting and Finance, 46(1): 107 – 124. Cited by: 15.

- Gal, Michal S, and Oshrit Aviv.** 2020. "The competitive effects of the GDPR." Journal of Competition Law & Economics, 16(3): 349–391.
- Goddard, Michelle.** 2017. "The EU General Data Protection Regulation (GDPR): European regulation that has a global impact." International Journal of Market Research, 59(6): 703–705.
- Goldberg, Samuel G, Garrett A Johnson, and Scott K Shriver.** 2024. "Regulating privacy online: An economic evaluation of the GDPR." American Economic Journal: Economic Policy, 16(1): 325–358.
- Goodridge, Peter, and Jonathan Haskel.** 2023. "Accounting for the slowdown in UK innovation and productivity." Economica, 90(359): 780–812.
- Goos, Maarten, and Maria Savona.** 2024. "The governance of artificial intelligence: Harnessing opportunities and mitigating challenges." Research Policy, 53(3): 104928.
- Grabowska, Sandra, and Sebastian Saniuk.** 2021. Modern marketing for customized products under conditions of fourth industrial revolution. Routledge: New York, NY, USA.
- Grazzi, Marco, Nadia Jacoby, and Tania Treibich.** 2016. "Dynamics of investment and firm performance: Comparative evidence from manufacturing industries." Empirical Economics, 51: 125–179.
- Griliches, Zvi.** 1984. "Introduction to" R & D, Patents, and Productivity"." In R&D, Patents, and Productivity. 1–20. University of Chicago Press.
- Groh, Carl-Christian, and Oliver Pfäuti.** 2023. "Monetary Policy and Uncertainty in the Data Economy." In Monetary Policy and Uncertainty in the Data Economy: Groh, Carl-Christian|uPfäuti, Oliver. [SI]: SSRN.
- Guellec, Dominique, and Bruno Van Pottelsberghe de la Potterie.** 2004. "From R&D to productivity growth: Do the institutional settings and the source of funds of R&D matter?" Oxford bulletin of economics and statistics, 66(3): 353–378.
- Haddara, Moutaz, Arra Salazar, and Marius Langseth.** 2023. "Exploring the Impact of GDPR on Big Data Analytics Operations in the E-Commerce Industry." Procedia Computer Science, 219: 767–777.
- Hall, Bronwyn H.** 2011. "Innovation and productivity." National bureau of economic research.
- Hall, Bronwyn H, Jacques Mairesse, and Pierre Mohnen.** 2010. "Measuring the Returns to R&D." In Handbook of the Economics of Innovation. Vol. 2, 1033–1082. Elsevier.
- Ichihashi, Shota.** 2021. "Competing data intermediaries." The RAND Journal of Economics, 52(3): 515–537.
- Ignà, Ioana, and Francesco Venturini.** 2023. "The determinants of AI innovation across European firms." Research Policy, 52(2): 104661.
- IMF.** 2015. Private Investment: What's the Holdup? Washington, D.C.:International Monetary Fund. Chapter 4 in World Economic Outlook: Uneven Growth—Short- and Long-Term Factors, April 2015.

- ISTAT.** 2018. “Report on competitiveness.” Italian National Institute of Statistics (ISTAT).
- Jia, Jian, Ginger Zhe Jin, and Liad Wagman.** 2020. “GDPR and the Localness of Venture Investment.” Available at SSRN 3436535.
- Johnson, Garrett A.** 2023. “Economic Research on Privacy Regulation: Lessons from the GDPR and Beyond.” The Economics of Privacy. University of Chicago Press.
- Jones, Charles I, and Christopher Tonetti.** 2020. “Nonrivalry and the Economics of Data.” American Economic Review, 110(9): 2819–2858.
- Kagermann, Henning, Wolfgang Wahlster, and Johannes Helbig.** 2013. “Securing the future of German manufacturing industry: Recommendations for implementing the strategic initiative INDUSTRIE 4.0.” Industry-Science Research Alliance, Germany. Final report of the Industrie 4.0 Working Group.
- Klein, Tobias J, Madina Kurmangaliyeva, Jens Prufer, and Patricia Prufer.** 2022. “How important are user-generated data for search result quality? Experimental evidence.” In . SSRN.
- Koski, Heli, and Nelli Valmari.** 2020. “Short-term Impacts of the GDPR on Firm Performance.” ETLA Working Papers.
- Kotschy, Rainer, and David E Bloom.** 2023. “Population aging and economic growth: From demographic dividend to demographic drag?”
- Kumbhakar, Subal C, Raquel Ortega-Argilés, Lesley Potters, Marco Vivarelli, and Peter Voigt.** 2012. “Corporate R&D and firm efficiency: evidence from Europe’s top R&D investors.” Journal of Productivity Analysis, 37: 125–140.
- Layton, Roslyn, and Silvia Elaluf-Calderwood.** 2019. “A Social Economic Analysis of the Impact of GDPR on Security and Privacy Practices.” 1–6.
- Lefrere, Vincent, Logan Warberg, Cristobal Cheyre, Veronica Marotta, Alessandro Acquisti, et al.** 2020. “The impact of the GDPR on content providers.”
- Lichtenberg, Frank R, and Donald Siegel.** 1991. “The impact of R&D investment on productivity—New evidence using linked R&D–LRD data.” Economic inquiry, 29(2): 203–229.
- Li, He, Lu Yu, and Wu He.** 2019. “The impact of GDPR on global technology development.”
- Lindgren, Peter.** 2016. “GDPR regulation impact on different business models and businesses.” Journal of Multi Business Model Innovation and Technology, 4(3): 241–254.
- Li, Wendy CY, Makoto Nirei, and Kazufumi Yamana.** 2019. Value of data: there’s no such thing as a free lunch in the digital economy. RIETI.
- Machlup, Fritz.** 1962. The production and distribution of knowledge in the United States. Vol. 278, Princeton university press.
- Maestas, Nicole, Kathleen J. Mullen, and David Powell.** 2023. “The Effect of Population Aging on Economic Growth, the Labor Force, and Productivity.” American Economic Journal: Macroeconomics, 15(2): 306–332.

- Marrocu, Emanuela, Raffaele Paci, and Marco Pontis.** 2012. “Intangible capital and firms’ productivity.” Industrial and corporate change, 21(2): 377–402.
- Mayer-Schönberger, Viktor, and Kenneth Cukier.** 2013. Big data: A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt.
- Mohades, Siavash, Giulia Piccillo, Maria Savona, and Tania Treibich.** Forthcoming. “Firms’ Investments & Capacity Utilisation: The Role of Financial Constraints and Uncertainty.”
- Mohnen, Pierre.** 2019. R&D, innovation and productivity. Springer.
- Müller, Oliver, Maria Fay, and Jan Vom Brocke.** 2018. “The effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics.” Journal of management information systems, 35(2): 488–509.
- Niebel, Thomas, Fabienne Rasel, and Steffen Viete.** 2019. “BIG data–BIG gains? Understanding the link between big data analytics and innovation.” Economics of Innovation and New Technology, 28(3): 296–316.
- Ning, Chao, and Fengqi You.** 2019. “Optimization under uncertainty in the era of big data and deep learning: When machine learning meets mathematical programming.” Computers & Chemical Engineering, 125: 434–448.
- Nonnis, Alberto, Ahmed Bounfour, and Keungoui Kim.** 2023. “Knowledge spillovers and intangible complementarities: Empirical case of European countries.” Research Policy, 52(1): 104611.
- Nucci, Francesco, Chiara Puccioni, and Ottavio Ricchi.** 2022. “Digital Technologies and Productivity: A Firm-Level Investigation for Italy.” Government of the Italian Republic (Italy), Ministry of Economy and Finance, Department of the Treasury Working Paper, , (3).
- OECD.** 2020. “A roadmap toward a common framework for measuring the digital economy.” Report for the G20 Digital Economy Task Force, available at.
- Piccarozzi, Michela, Barbara Aquilani, and Corrado Gatti.** 2018. “Industry 4.0 in management studies: A systematic literature review.” Sustainability, 10(10): 3821.
- Romer, Paul M.** 1990. “Endogenous technological change.” Journal of political Economy, 98(5, Part 2): S71–S102.
- Sabatino, Lorien, and Geza Sapi.** 2022. “Online privacy and market structure: Theory and evidence.” Information Economics and Policy, 60: 100985.
- Salvi, Antonio, Filippo Vitolla, Michele Rubino, Anastasia Giakoumelou, and Nicola Raimo.** 2021. “Online information on digitalisation processes and its impact on firm value.” Journal of Business Research, 124: 437–444.
- Savona, Maria.** 2019. “The Value of Data: Towards a Framework to Redistribute It.” Science Policy Research Unit, University of Sussex SWPS 2019-21. Research Note.
- Savona, Maria.** 2024. “Data Governance: Main Challenges.” Vol. 25, 28–31, CESifo.

- Schiffer, Mirjam, and Beatrice Weder.** 2001. Firm size and the business environment: Worldwide survey results. Vol. 43, World Bank Publications.
- Schivardi, Fabiano, and Tom Schmitz.** 2019. "The IT Revolution and Southern Europe's Two Lost Decades." Journal of the European Economic Association, 18(5): 2441–2486.
- Shefer, Daniel, and Amnon Frenkel.** 2005. "R&D, firm size and innovation: an empirical analysis." Technovation, 25(1): 25–32.
- Siliverstovs, Boriss, et al.** 2016. "R&D and non-linear productivity growth." Research policy, 45(3): 634–646.
- Simon, Phil.** 2013. Too big to ignore: the business case for big data. Vol. 72, John Wiley & Sons.
- Taufick, Roberto D.** 2021. "The underdeterrence, underperformance response to privacy, data protection laws." Technology in Society, 67: 101752.
- Teece, David J.** 1986. "Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy." Research Policy, 15(6): 285–305.
- Truant, Elisa, Laura Broccardo, and Léo-Paul Dana.** 2021. "Digitalisation boosts company performance: an overview of Italian listed companies." Technological Forecasting & Social Change, 173: 121173.
- Tysiac, Ken.** 2015. "Data analytics helps auditors gain deep insight." Journal of Accountancy, 219(4): 52.
- Urban, Tobias, Dennis Tatang, Martin Degeling, Thorsten Holz, and Norbert Pohlmann.** 2018. "The unwanted sharing economy: An analysis of cookie syncing and user transparency under GDPR." arXiv preprint arXiv:1811.08660.
- Van Oijen, Marcel.** 2017. "Bayesian methods for quantifying and reducing uncertainty and error in forest models." Current Forestry Reports, 3: 269–280.
- Voss, W Gregory, and Kimberly A Houser.** 2019. "Personal data and the GDPR: providing a competitive advantage for US companies." American Business Law Journal, 56(2): 287–344.
- Wondrium.** 2016. "How Gathering Data can Reduce Uncertainty ." Connel Fullenkamp. Available at: <https://www.wondriumdaily.com/reducing-uncertainty-creating-information>, Accessed: 31 July 2023.
- Wu, Kuo-Jui, Ching-Jong Liao, Ming-Lang Tseng, Ming K Lim, Jiayao Hu, and Kimhua Tan.** 2017. "Toward sustainability: using big data to explore the decisive attributes of supply chain risks and uncertainties." Journal of Cleaner Production, 142: 663–676.

A Additional Tables

Sectors	Inv in Data	Inv in Tangible Assets	# Obs
Manufacturing	454.93	6,507.04	64,934
1 - Food industries, beverages and tobacco products	201.15	3,632.03	9,565
2 - Textiles, clothing, and hide, leather and footwear products	132.34	1,295.42	7,290
3 - Coke manufacturing, chemical industry, rubber and plastics	429.42	8,698.17	7,410
4 - Processing of non-metallic minerals	88.82	3,510.76	4,130
5 - Metal engineering industry	348.23	5,302.40	25,956
6 - Other manufacturing industries	120.84	2,288.67	7,326
7 - Other industries excluding construction (energy and extraction)	4,046.33	44,518.13	3,257
Services	2,471.35	12,341.30	25,413
8 - Wholesale and retail commerce	285.14	4,540.99	10,816
9 - Hotels and restaurants	122.51	2,007.69	1,940
10 - Transport and communications	7,642.24	33,794.41	7,391
11 - Real estate activities, etc.	569.47	2,059.40	5,266
All Industries	1,022.11	8,148.12	90,347

Table A.1: Average investments in data and tangible assets, in constant thousand-euros, taken from INVIND database in annual frequency for 2002-2023

Sectors	2012-2017	2018-2023
Manufacturing	0.30	1.45
1 - Food industries, beverages and tobacco products	0.11	0.07
2 - Textiles, clothing, and hide, leather and footwear products	0.25	0.50
3 - Coke manufacturing, chemical industry, rubber and plastics	0.14	0.19
4 - Processing of non-metallic minerals	0.12	0.08
5 - Metal engineering industry	0.31	0.31
6 - Other manufacturing industries	0.24	0.23
7 - Other industries excluding construction (energy and extraction)	1.34	18.41
Services	0.59	3.25
8 - Wholesale and retail commerce	0.32	3.24
9 - Hotels and restaurants	0.09	0.12
10 - Transport and communications	0.90	4.46
11 - Real estate activities, etc.	0.96	2.40
All Industries	0.38	1.97

Table A.2: Sector-Level Averages of Data Intensity, Taken From INVIND Database in Annual Frequency for 2012-2023 - Divided by Pre and Post GDPR and Decreto Legislativo 10 agosto 2018, n. 101 .

Sectors	Pre-2018	Post-2018	Pre-2016	Post-2016	All Sample
Manufacturing	1.03	0.95	0.93	1.01	1.02
1 - Food industries, beverages and tobacco products	0.13	0.19	0.12	0.19	0.16
2 - Textiles, clothing, and hide, leather and footwear products	1.80	1.55	2.04	1.45	1.67
3 - Coke manufacturing, chemical industry, rubber and plastics	0.43	0.66	0.44	0.62	0.55
4 - Processing of non-metallic minerals	0.33	0.18	0.33	0.22	0.28
5 - Metal engineering industry	1.70	1.38	1.42	1.41	1.53
6 - Other manufacturing industries	0.65	1.04	0.64	1.44	1.08
7 - Other industries excluding construction (energy and extraction)	0.17	0.39	0.04	0.48	0.30
Services	0.73	7.11	0.81	5.44	3.46
8 - Wholesale and retail commerce	0.18	0.45	0.19	0.36	0.29
9 - Hotels and restaurants	0.01	0.04	0.01	0.03	0.02
10 - Transport and communications	1.72	3.10	2.02	2.60	2.30
11 - Real estate activities, etc.	0.72	30.05	0.67	22.55	14.17
All Industries	0.95	2.74	0.90	2.29	1.71

Table A.3: Sector-Level Averages of R&D Intensity, Taken From INVIND Database in Annual Frequency for 2002-2023 - Divided by Pre and Post GDPR and Decreto Legislativo 10 Agosto 2018, n. 101 .

Condition	Count	Unique Firms
Data-intensive pre-GDPR & R&D-intensive pre-GDPR	4,684	1,802
Not data-intensive pre-GDPR & R&D-intensive pre-GDPR	7,084	2,649
Data-intensive pre-GDPR & not R&D-intensive pre-GDPR	8,485	3,542
Not data-intensive pre-GDPR & not R&D-intensive pre-GDPR	30,455	6,682
Data-intensive pre-GDPR & not data-intensive Post GDPR	959	828
Not data intensive pre-GDPR and data-intensive post GDPR	0	0
R&D-intensive pre-GDPR & not R&D-intensive Post GDPR	0	0
Not data intensive pre-GDPR and R&D-intensive post GDPR	2,186	1,393
Data-intensive pre-GDPR & data-intensive post-GDPR	19,639	6,171
R&D-intensive pre-GDPR & R&D-intensive post-GDPR	11,768	3,375
Data-intensive & R&D-intensive pre- & post-GDPR	4,493	1,753
Not data-intensive & not R&D-intensive pre- & post-GDPR	28,797	6,476
Data-intensive pre- & post-GDPR & not R&D-intensive pre- & post-GDPR	7,572	3,263
Not data-intensive pre- & post-GDPR & R&D-intensive pre- & post-GDPR	7,084	2,649

Table A.4: Firms' state Pre- and Post-GDPR

Dependent Variable: Log of R&D Investments						
Instrument for Data Investments: GDPR Announcement 2018						
Lag (τ)	0	1	2	3	4	5
Log of Data Investments	1.543*** (0.247)	0.418* (0.175)	0.441* (0.185)	0.481* (0.200)	0.361 (0.206)	0.343 (0.215)
No Lags						
Sales Growth $_{t-1}$	-0.297*** (0.059)	-0.090 (0.051)	-0.097 (0.054)	-0.144* (0.057)	-0.140* (0.059)	-0.151* (0.060)
Age	0.002 (0.002)	0.008*** (0.001)	0.009*** (0.002)	0.009*** (0.002)	0.008*** (0.002)	0.007*** (0.002)
Emp Growth $_{t-1}$	-0.121 (0.138)	0.383** (0.123)	0.339* (0.132)	0.345* (0.146)	0.338* (0.156)	0.360* (0.162)
Profit $_{t-1}$	0.122*** (0.029)	0.137*** (0.031)	0.120*** (0.033)	0.123 *** (0.035)	0.113** (0.036)	0.106** (0.037)
Constant	-2.100* (0.894)	1.902** (0.640)	1.789** (0.680)	1.643* (0.735)	2.182** (0.760)	2.288** (0.794)
Sector FE	✓	✓	✓	✓	✓	✓
Firms	3,554	3,085	2,699	2,384	2,347	2,181
Observations	13,276	11,552	10,144	8,993	8,442	7,913

Table A.5: Regression Results: Causal analysis in Equation 3.7 for $\tau > 0$ with random effects

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Log of R&D Investments						
Instrument for Data Investments: GDPR Announcement 2018						
Lag (τ)	0	1	2	3	4	5
Log of Data Investments	0.769** (0.270)	0.006 (0.186)	0.143 (0.194)	0.266 (0.209)	0.228 (0.216)	0.308 (0.224)
No Lags						
Sales Growth $_{t-1}$	-0.210*** (0.061)	-0.120* (0.053)	-0.114* (0.055)	-0.155* * (0.059)	-0.148* (0.060)	-0.167** (0.062)
Age	0.009*** (0.003)	0.015*** (0.003)	0.017*** (0.003)	0.021*** (0.003)	0.013*** (0.003)	0.010*** (0.003)
Emp Growth $_{t-1}$	0.051 (0.145)	0.330** (0.128)	0.243 (0.137)	0.240 (0.153)	0.295 (0.163)	0.304 (0.169)
Profit $_{t-1}$	0.068* (0.031)	0.082* (0.033)	0.063 (0.035)	0.063 (0.037)	0.063 (0.038)	0.055 (0.040)
Constant	3.409** (1.124)	5.111*** (0.963)	4.452*** (0.989)	3.810*** (1.036)	4.270*** (1.112)	3.749** (1.250)
Sector FE	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Firms	3,554	3,085	2,699	2,384	2,347	2,181
Observations	13,276	11,552	10,144	8,993	8,442	7,913

Table A.6: Regression Results: Causal analysis in Equation 3.7 for $\tau > 0$ with fixed effects

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

B Lagged Productivity Gains

Lag (τ)	Dependent Variable: Labour Productivity					
	0	1	2	3	4	5
Log of Data Investments	0.046*** (0.007)	0.068*** (0.008)	0.069*** (0.008)	0.071*** (0.009)	0.074*** (0.010)	0.076*** (0.010)
Log of R&D Investments	0.038*** (0.006)	0.053*** (0.006)	0.060*** (0.007)	0.058*** (0.007)	0.059*** (0.008)	0.069*** (0.008)
Log of Data Investments $\times$ Log of R&D Investments	0.003** (0.001)	0.001 (0.001)	-0.001 (0.001)	0.002 (0.001)	-0.002 (0.002)	-0.003* (0.002)
Constant	6.236*** (0.061)	6.051*** (0.067)	6.087*** (0.069)	6.136*** (0.071)	6.083*** (0.074)	6.096*** (0.078)
Sector FE	✓	✓	✓	✓	✓	✓
Firms	3,127	2,659	2,532	2,370	2,200	2,000
Observations	12,333	10,032	9,029	8,088	7,124	6,231

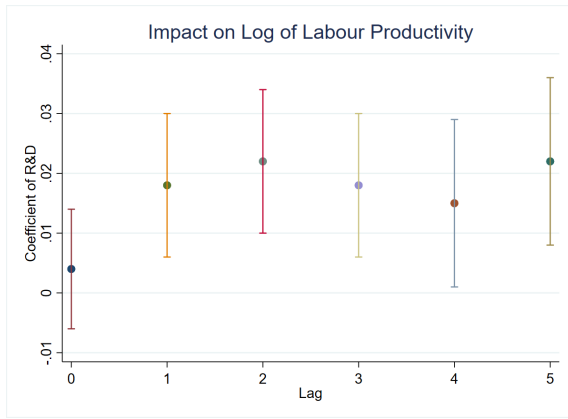
Table B.1: Regression Results: Complementarity analysis in Equation 3.8 for $\tau > 0$ with random effects

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

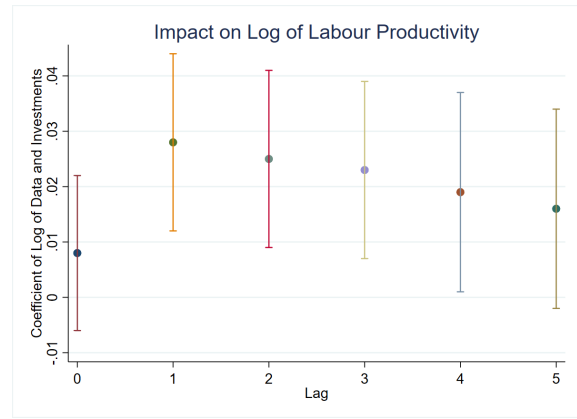
Lag (τ)	Dependent Variable: Labour Productivity					
	0	1	2	3	4	5
Log of Data Investments	0.008 (0.007)	0.028*** (0.008)	0.025** (0.008)	0.023** (0.008)	0.019* (0.009)	0.016 (0.009)
Log of R&D Investments	0.004 (0.005)	0.018** (0.006)	0.022*** (0.006)	0.018** (0.006)	0.015* (0.007)	0.022** (0.007)
Log of Data Investments $\times$ Log of R&D Investments	0.003** (0.001)	0.000 (0.001)	-0.002 (0.001)	0.002 (0.001)	-0.002 (0.001)	-0.003* (0.001)
Constant	6.933*** (0.246)	6.918*** (0.370)	7.259*** (0.160)	6.719*** (0.202)	6.751*** (0.209)	6.846*** (0.348)
Sector FE	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Firms	3,127	2,659	2,532	2,370	2,200	2,000
Observations	12,333	10,032	9,029	8,088	7,124	6,231

Table B.2: Regression Results: Complementarity analysis in Equation 3.8 for $\tau > 0$ with fixed effects

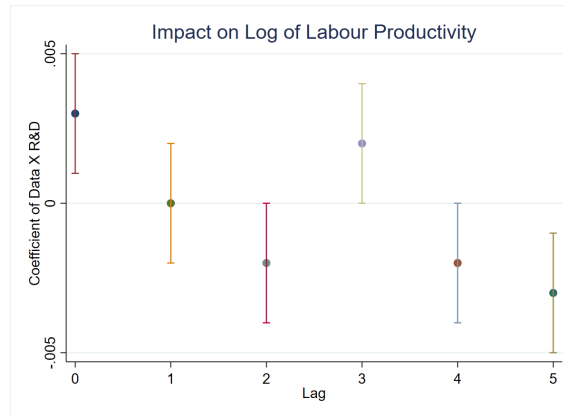
*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval



(a) The Coefficients of R&D



(b) The Coefficients of Data



(c) The Coefficients of **Data** $\times$ R&D

Figure 4: The Lagged Coefficients - $\tau = 1, 2, 3, 4, 5$ in Equation 3.8

C Robustness Checks - Regression Tables

Dependent Variable: Log of R&D Investments						
Instrument for Data Investments: GDPR Implementation 2018						
Log of Data Investments	0.734** (0.268)	0.769** (0.270)	0.737** (0.270)	0.733** (0.270)	-0.234 (0.321)	-0.239 (0.321)
Sales Growth _{t-1}	-0.207*** (0.061)	-0.210*** (0.061)	-0.205*** (0.061)	-0.205*** (0.061)	-0.033 (0.084)	-0.032 (0.084)
Age	0.009*** (0.003)	0.009*** (0.003)	0.009*** (0.003)	0.009*** (0.003)	-0.005 (0.004)	-0.005 (0.004)
Emp Growth _{t-1}	0.071 (0.144)	0.051 (0.145)	0.023 (0.146)	0.023 (0.146)	0.006 (0.189)	0.009 (0.190)
Large			0.243** (0.076)	0.243** (0.076)	0.266* (0.110)	0.264* (0.110)
North			-0.765 (0.466)	-0.763 (0.466)	-3.944*** (0.681)	-3.944*** (0.681)
Profit _{t-1}	0.066* (0.031)	0.068* (0.031)	0.065* (0.031)	0.062* (0.031)	0.033 (0.044)	0.034 (0.044)
Fin Cons				-0.064 (0.044)		0.037 (0.074)
High Tech _{t-1}					0.066 (0.034)	0.066 (0.034)
Constant	1.983* (0.959)	3.409** (1.124)	3.882** (1.183)	3.903*** (1.183)	9.902*** (1.351)	9.920*** (1.352)
Sector FE		✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Firms	3,554	3,554	3,552	3,552	1,902	1,902
Observations	13,276	13,276	13,233	13,233	5,173	5,173
Treated Sectors: Energy and extraction - Wholesale and retail commerce - Textiles, clothing, and hide, leather and footwear products - Transport and communications						

Table C.1: Regression Results for the second stage regression in Equation 3.7 with firm and time FE

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Log of R&D Investments						
Instrument for Data Investments: GDPR Implementation 2018						
Log of Data Investments	0.987** (0.365)	2.506*** (0.375)	2.751*** (0.372)	2.730*** (0.372)	-0.965* (0.447)	-0.966* (0.447)
Sales Growth _{t-1}	-0.224** (0.069)	-0.427*** (0.070)	-0.452*** (0.070)	-0.450*** (0.070)	0.066 (0.092)	0.065 (0.092)
Age	0.005** (0.002)	-0.002 (0.002)	-0.007*** (0.002)	-0.007*** (0.002)	0.006* (0.003)	0.006* (0.003)
Emp Growth _{t-1}	0.073 (0.163)	-0.430** (0.166)	-0.648*** (0.166)	-0.644*** (0.166)	0.229 (0.208)	0.228 (0.208)
Profit _{t-1}	0.115*** (0.029)	0.119*** (0.029)	0.107*** (0.029)	0.102*** (0.029)	0.070 (0.042)	0.070 (0.042)
Large			1.065*** (0.049)	1.065*** (0.049)	1.096*** (0.068)	1.096*** (0.068)
North			0.378*** (0.069)	0.375*** (0.069)	0.457*** (0.093)	0.456*** (0.093)
Fin Cons				-0.089* (0.042)		-0.020 (0.070)
High Tech _{t-1}					0.093** (0.032)	0.093** (0.032)
Constant	0.626 (1.316)	-5.561*** (1.350)	-6.950*** (1.340)	-6.858*** (1.340)	6.821*** (1.614)	6.826*** (1.615)
Sector FE		✓	✓	✓	✓	✓
Firms	3,554	3,554	3,552	3,552	1,902	1,902
Observations	13,276	13,276	13,233	13,233	5,173	5,173
Treated Sectors: All in Services						

Table C.2: Regression Results for Equation 3.7 for Manufacturing and Services sectors as control and treatments, respectively

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Log of R&D Investments						
Instrument for Data Investments: GDPR Announcement 2016						
Log of Data Investments	1.380*** (0.246)	2.118*** (0.254)	2.291*** (0.252)	2.277* ** (0.252)	0.483 (0.389)	0.487 (0.389)
Sales Growth _{t-1}	-0.278*** (0.059)	-0.375*** (0.059)	-0.390*** (0.059)	-0.389* ** (0.059)	-0.135 (0.087)	-0.136 (0.087)
Age	0.004* (0.002)	-0.000 (0.002)	-0.004** (0.002)	-0.004** (0.002)	-0.000 (0.002)	-0.000 (0.002)
Emp Growth _{t-1}	-0.062 (0.138)	-0.316* (0.139)	-0.512*** (0.140)	-0.510** * (0.140)	-0.246 (0.201)	-0.249 (0.201)
Large			1.067*** (0.049)	1.067*** (0.049)	1.095*** (0.068)	1.095*** (0.068)
North			0.381*** (0.069)	0.378*** (0.069)	0.462*** (0.093)	0.462*** (0.093)
Profit _{t-1}	0.113*** (0.029)	0.118*** (0.029)	0.105*** (0.029)	0.101 *** (0.029)	0.070 (0.042)	0.069 (0.042)
High Tech _{t-1}					0.091** (0.033)	0.090** (0.033)
Fin Cons				-0.087* (0.042)		-0.021 (0.070)
Constant	-0.796 (0.889)	-4.158*** (0.916)	-5.292*** (0.911)	-5.224*** (0.912)	1.614 (1.406)	1.606 (1.407)
Sector FE		✓	✓	✓	✓	✓
Firms	3,554	3,554	3,552	3,552	1,902	1,902
Observations	13,276	13,276	13,233	13,233	5,173	5,173

Table C.3: Regression Results for Equation 3.7 for GDPR 2016 as Instrument

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Log of R&D Investments per Labour Instrument for Data Investments per Labour: GDPR 2018						
Log of Data Investments per Labour	1.037*** (0.288)	2.071*** (0.295)	2.075*** (0.296)	2.069*** (0.296)	-0.283 (0.385)	-0.288 (0.385)
Sales Growth _{t-1}	-0.047*** (0.013)	-0.078*** (0.014)	-0.078*** (0.014)	-0.078* ** (0.014)	-0.007 (0.019)	-0.006 (0.019)
Age	-0.000 (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.000 (0.000)	-0.000 (0.000)
Emp Growth _{t-1}	-0.024 (0.023)	-0.034 (0.023)	-0.026 (0.024)	-0.026 (0.024)	-0.060 (0.033)	-0.060 (0.033)
Large			-0.043*** (0.010)	-0.043*** (0.010)	-0.071*** (0.014)	-0.071*** (0.014)
North			0.007 (0.014)	0.007 (0.014)	0.025 (0.018)	0.025 (0.018)
Profit _{t-1}	0.024*** (0.006)	0.026*** (0.006)	0.027*** (0.006)	0.026 *** (0.006)	0.016 (0.009)	0.016 (0.009)
High Tech _{t-1}					0.013 (0.007)	0.013 (0.007)
Fin Cons				-0.008 (0.009)		0.011 (0.015)
Constant	0.229 (0.198)	-0.541** (0.203)	-0.532** (0.203)	-0.526** (0.203)	1.202*** (0.265)	1.204*** (0.265)
Sector FE		✓	✓	✓	✓	✓
Firms	3,554	3,554	3,552	3,552	1,902	1,902
Observations	13,276	13,276	13,233	13,233	5,173	5,173

Table C.4: Regression Results for the second stage regression in Equation 3.7 with R&D and data both normalised for labour

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Labour Productivity						
Log of Data Investments	0.009 (0.007)	0.008 (0.007)	0.008 (0.007)	0.007 (0.007)	0.002 (0.007)	0.006 (0.014)
Log of R&D Investments	0.005 (0.005)	0.004 (0.005)	0.001 (0.006)	-0.000 (0.006)	-0.003 (0.006)	-0.002 (0.011)
Log of Data Investments × Log of R&D Investments	0.003** (0.001)	0.003** (0.001)	0.002* (0.001)	0.003* (0.001)	0.003** (0.001)	0.001 (0.002)
Log of Other Investments			0.041*** (0.004)	0.041*** (0.004)	0.039*** (0.004)	0.027*** (0.006)
Age					-0.024 (0.013)	-0.018 (0.030)
North					-0.066 (0.121)	-0.978*** (0.294)
Large					0.273*** (0.022)	0.221*** (0.037)
High Tech _{t-1}						0.001 (0.012)
Constant	6.456*** (0.031)	6.933*** (0.246)	6.406*** (0.042)	6.888*** (0.311)	6.836*** (0.327)	7.903*** (0.425)
Sector FE		✓		✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓
Firms	3,127	3,127	2,607	2,607	2,601	1,499
Observations	12,333	12,333	10,046	10,046	9,998	3,964

Table C.5: Regression Results: Complementarity analysis in Equation 3.8 with time and firm fixed effects

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Labour Productivity						
Log of Data Investments per Labour	0.159*** (0.042)	0.154*** (0.041)	0.223*** (0.052)	0.214*** (0.052)	0.185*** (0.051)	0.208* (0.088)
Log of R&D Investments per Labour	0.026 (0.030)	0.031 (0.030)	0.023 (0.038)	0.032 (0.038)	0.040 (0.037)	0.055 (0.064)
Log of Data Investments per Labour × Log of R&D Investments per Labour	-0.067 (0.035)	-0.063 (0.035)	-0.137** (0.045)	-0.130** (0.044)	-0.116** (0.043)	-0.124 (0.073)
Log of Other Investments			0.114*** (0.004)	0.110*** (0.004)	0.102*** (0.004)	0.108*** (0.006)
Age					0.000 (0.013)	0.102*** (0.025)
North					0.575*** (0.044)	0.505*** (0.060)
Large					0.553*** (0.021)	0.650*** (0.034)
High Tech _{t-1}						-0.017 (0.013)
Constant	6.254*** (0.042)	6.511*** (0.074)	5.712*** (0.053)	5.926*** (0.072)	5.293*** (0.084)	4.770*** (0.140)
Sector FE		✓		✓	✓	✓
Firms	3,127	3,127	2,607	2,607	2,601	1,499
Observations	12,333	12,333	10,046	10,046	9,998	3,964

Table C.6: Regression Results: Complementarity analysis in Equation 3.8 with both R&D and data normalised for labour

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Labour Productivity Data Investments instrumented by GDPR 2018						
Log of Data Investments	0.283*	0.350**	0.204	0.275*	0.218	0.186
	(0.111)	(0.112)	(0.132)	(0.132)	(0.132)	(0.229)
Log of R&D Investments	-0.117	-0.056	-0.110	-0.043	-0.062	-0.093
	(0.065)	(0.066)	(0.075)	(0.075)	(0.074)	(0.133)
Log of Data Investments ×Log of R&D Investments	0.042*	0.027	0.038	0.022	0.026	0.042
	(0.017)	(0.017)	(0.020)	(0.020)	(0.019)	(0.035)
Log of Other Investments			0.098***	0.095* **	0.090***	0.085***
			(0.004)	(0.004)	(0.004)	(0.006)
Age					-0.038*	0.025
					(0.019)	(0.034)
North					0.574***	0.479***
					(0.044)	(0.060)
Large					0.504***	0.619***
					(0.022)	(0.035)
High Tech _{t-1}						-0.029*
						(0.013)
Constant	5.006***	4.955***	4.919***	4.851***	4.636***	4.375***
	(0.421)	(0.431)	(0.499)	(0.502)	(0.495)	(0.863)
Sector FE		✓		✓	✓	✓
Firms	2,478	2,478	2,255	2,255	2,251	1,499
Observations	12,333	12,333	10,046	10,046	9,998	1,356

Table C.7: Regression Results: Complementarity analysis in Equation 3.8 with data predicted using the first-stage regression in Equation 3.5

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Log of R&D Investments				
Instrument for Data Investments: GDPR Announcement 2018				
Period	Pre GDPR		Post GDPR	
Log of Data Investments	1.146*** (0.020)	0.927*** (0.035)	-0.292 (0.465)	1.095*** (0.046)
Sales Growth _{t-1}	-0.242*** (0.064)	-0.200** (0.063)	-0.011 (0.090)	-0.207* ** (0.061)
Age	0.004** (0.001)	0.004** (0.001)	0.008*** (0.002)	0.004* (0.002)
Emp Growth _{t-1}	-0.178 (0.148)	-0.065 (0.147)	0.377 (0.210)	-0.103 (0.147)
Profit _{t-1}	0.179*** (0.039)	0.194*** (0.039)	0.021 (0.041)	0.012 (0.041)
Sector FE		✓		✓
Firms	2,898	2,898	1,828	1,828
Observations	8,334	8,334	4,942	4,942

Table C.8: Regression Results of Equation 3.7 for pre and post GDPR Analysis

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Dependent Variable: Labour Productivity				
Period	Pre GDPR		Post GDPR	
Log of Data Investments	0.045*** (0.009)	0.047*** (0.009)	0.031* (0.013)	0.032* (0.013)
Log of R&D Investments	0.036*** (0.007)	0.044*** (0.007)	0.038*** (0.010)	0.044*** (0.010)
Log of Data Investments × Log of R&D Investments	0.004** (0.001)	0.004** (0.001)	0.007*** (0.002)	0.007*** (0.002)
Constant	6.035*** (0.047)	6.275*** (0.072)	5.813*** (0.063)	6.152*** (0.090)
Sector FE		✓		✓
Firms	2,490	2,490	1,823	1,823
Observations	7,290	7,290	5,043	5,043

Table C.9: Regression Results of Equation 3.8 for pre and post GDPR Analysis

*** Significant with 99% Confidence Interval, ** Significant with 95% Confidence Interval, * Significant with 90% Confidence Interval

Summary

English Summary: This thesis delves into the multifaceted nature of economic uncertainty and its prominence for both macroeconomic theory and firm-level decisions. The first chapter in this thesis develops a novel method based on firm-level sales volatility decomposition to identify uncertainty components at the firm, the sector and the aggregate levels. This chapter also introduces the concept of “uncertainty heterogeneity”, which refers to how uncertainty varies across and within firms. The results indicate that macroeconomic conditions are the primary drivers of uncertainty; nonetheless, variation exists across firms and sectors. Unlike other aggregate and firm uncertainty measures in the literature, this approach can generate granular uncertainty measures for most economies without much data hunger.

In the second chapter, we study how firms’ investments are affected by demand uncertainty, capacity utilisation, i.e. the ratio of actual to potential output, and financial constraints. Italian firms are of interest given their lower capacity utilisation and investments compared to European counterparts. Our results show that even though firms operating at high capacity are not sensitive to uncertainty, the investments of those operating at low capacity are positively affected by uncertainty. Therefore, these results underscore the heterogeneity of the impact of uncertainty on firms’ investments, namely between firms at and outside of the “growth window”, i.e. firms operating at a higher capacity than desired. Accordingly, these results demonstrate the need for targeted policies focusing on firms in this critical growth window to stimulate investment and foster economic growth, especially during economic downturns. Furthermore, we find that firms’ investments are bound by financial constraints as most financing occurs via external funds, stressing the role of policy in easing and allocating access to funds for investments across high-growth potential but constrained firms.

The third chapter examines the effects of firms’ investments in data on their Research and Development (R&D) investments. Additionally, we investigate how R&D investments and “datafication”, i.e. the process of collecting, quantifying and analysing data, affect firms’ labour productivity. Following our earlier discourse, we expect datafication to be associated with reducing uncertainty in areas such as inventory management and marketing and, therefore, increasing R&D investments. Investments in data reduce marginal costs and hence increase the firm size and market concentration, which together encourage more R&D investments. To disentangle the endogeneity between datafication and R&D investments, we utilise the introduction of GDPR as an instrumental variable and establish a positive causal link between data investments and R&D. Furthermore, data can be used to enhance firms’ productivity. As a result, we argue that a virtuous cycle exists between datafication, R&D investments, and firm performance, which must be considered, primarily when regulating data is the core of policy.

Riassunto in Italiano: Questa tesi approfondisce la natura multidimensionale dell'incertezza economica e la sua importanza sia per la teoria macroeconomica sia per le decisioni a livello aziendale. Il primo capitolo propone un nuovo metodo per identificare le componenti dell'incertezza economica comuni a livello aziendale, settoriale e aggregato. Questo nuovo metodo si basa sulla decomposizione della volatilità delle vendite a livello aziendale. Inoltre in questo capitolo si introduce anche il concetto di “eterogeneità dell'incertezza”, riferito alla variabilità dell'incertezza da impresa a impresa e all'interno della stessa. I risultati indicano che le condizioni macroeconomiche sono alla base del principale componente di incertezza; tuttavia, esistono variazioni tra imprese e settori. A differenza di altre misure di incertezza (sia a livello aggregato sia a livello aziendale) presenti in letteratura, questo approccio può generare misure di incertezza granulari per la maggior parte dei Paesi senza bisogno di molti dati.

Nel secondo capitolo, studiamo come gli investimenti delle imprese siano influenzati dall'incertezza della domanda, dall'utilizzo della capacità produttiva, ovvero dal rapporto tra produzione effettiva e potenziale, e dai vincoli finanziari. Le imprese italiane sono interessanti in quanto il loro utilizzo della capacità produttiva e gli investimenti sono più bassi rispetto alle controparti europee. I risultati mostrano che anche se le imprese che operano ad alta capacità non reagiscono significativamente all'incertezza, gli investimenti di quelle che operano a bassa capacità ne sono influenzati positivamente. Questi risultati sottolineano l'eterogeneità dell'impatto dell'incertezza sugli investimenti delle imprese, in particolare tra le imprese che si trovano nella “finestra di crescita” e quelle che ne sono al di fuori, ossia le imprese che operano a una capacità superiore a quella desiderata. Di conseguenza, questi risultati dimostrano la necessità di politiche mirate che si concentrino sulle imprese in questa finestra di crescita critica per stimolare gli investimenti e favorire la crescita economica, soprattutto durante le fasi di recessione economica. Inoltre, si scopre che gli investimenti delle imprese sono soggetti a vincoli finanziari, in quanto la maggior parte dei finanziamenti avviene attraverso fondi esterni, sottolineando il ruolo della politica nell'agevolare e permettere l'accesso ai fondi per gli investimenti nelle imprese ad alto potenziale di crescita ma soggette a vincoli.

Il terzo articolo esamina gli effetti degli investimenti delle imprese nei dati sui loro investimenti in Ricerca e Sviluppo (R&S). Inoltre, studiamo come gli investimenti in R&S e la “datificazione”, ovvero il processo di raccolta, quantificazione e analisi dei dati, influenzino la produttività del lavoro delle imprese. Seguendo il discorso precedente, ci aspettiamo che la datificazione sia associata alla riduzione dell'incertezza in aree come la gestione delle scorte e marketing, portando quindi all'aumento degli investimenti in R&S. Gli investimenti in dati riducono i costi marginali accrescendo le dimensioni dell'azienda e la concentrazione del mercato, che insieme incoraggiano maggiori investimenti in R&S. Per risolvere l'endogeneità tra la datificazione e gli investimenti in R&S, utilizziamo l'introduzione del GDPR come variabile strumentale e stabiliamo un nesso causale positivo tra gli investimenti in dati e la R&S. Inoltre, i dati possono essere utilizzati per migliorare l'efficienza produttiva delle imprese. Di conseguenza sosteniamo l'esistenza di un circolo virtuoso tra datificazione, investimenti in R&S e performance delle imprese, che deve essere preso in considerazione soprattutto quando la regolamentazione dei dati è al centro delle politiche.

Nederlandse Samenvatting: Dit proefschrift onderzoekt de veelzijdige aard van economische onzekerheid en de rol die deze onzekerheid speelt in macro-economische theorie en bedrijfsbeslissingen. Het eerste hoofdstuk ontwikkelt een nieuwe methode, gebaseerd op de decompositie van de verkoopvolatiliteit, die de onzekerheid op bedrijfsniveau ontleedt in componenten gerelateerd aan bedrijfs-, sector- en geaggregeerde niveaus. Daarnaast introduceert dit hoofdstuk ook het concept van “onzekerheidsheterogeniteit” dat verwijst naar de variatie in onzekerheid tussen en binnen bedrijven. De resultaten wijzen macro-economische omstandigheden aan als de voornaamste bron van onzekerheid; hoewel er aanzienlijke variatie is in de effecten tussen bedrijven en sectoren. In tegenstelling tot andere maatstaven voor (bedrijfs)onzekerheid, is deze aanpak in staat om verschillende onzekerheidsuitkomsten te genereren in de meeste gevallen zonder de noodzaak van uitgebreide data.

In het tweede hoofdstuk onderzoeken we hoe onzekerheid over de bezettingsgraad, i.e. de verhouding tussen feitelijke en potentiële productie, en financiële beperkingen de investeringen van bedrijven beïnvloeden. Italiaanse bedrijven zijn interessant gezien hun lagere bezettingsgraad en investeringen in vergelijking met andere Europese bedrijven. Onze resultaten tonen aan dat bedrijven met een hoge capaciteit niet gevoelig zijn voor onzekerheid, terwijl bedrijven met een lage capaciteit juist positief worden beïnvloed door onzekerheid. Daarnaast ondersteunen onze resultaten de heterogeniteit in de impact van onzekerheid op de bedrijfsinvesteringen. Deze bevindingen ondersteunen de noodzaak van gericht beleid op bedrijven met groeipotentie om investeringen te stimuleren en economische groei te bevorderen, vooral tijdens economische recessies. Verder benadrukken wij het belang van beleidsmaatregelen die fondsen vrijmaken en toewijzen aan bedrijven met een hoog groeipotentieel, aangezien deze vaak te maken hebben met financiële beperkingen die hun groei beperken.

Het derde hoofdstuk onderzoekt de effecten van bedrijfsinvesteringen in data op hun Research and Development (R&D) investeringen. Daarnaast onderzoeken we hoe deze R&D investeringen en “dataficatie”, de arbeidsproductiviteit van bedrijven beïnvloeden. In navolging van onze eerdere argumenten, stellen we de hypothese dat dataficatie leidt tot een verhoging in R&D investeringen door een vermindering in onzekerheid. Investeringen in data verlagen de marginale kosten, vergroten de bedrijfsgrootte en marktconcentratie, en stimuleren daarmee verdere investeringen in R&D. Om de endogeniteit tussen dataficatie en R&D investeringen aan te pakken, gebruiken we de invoering van De Algemene Verordening Gegevensbescherming (AVG) als instrument voor datainvesteringen en stellen we een positief causaal verband vast tussen datainvesteringen en R&D. Bovendien kan data worden gebruikt om de arbeidsproductiviteit van bedrijven te verhogen. Kortom, we concluderen dat er een opwaartse spiraal bestaat tussen dataficatie, R&D investeringen en bedrijfsprestaties, wat cruciaal is om in overweging te nemen bij het opstellen van beleid rond dataregulering.