**Europa-Universität Flensburg**

International Institute for Management and Economic Education

# Essays in Industrial Organization: Firm Boundaries and Market Power

**Author:** Ali Taleb

Student ID: 556723

Submitted in fulfillment of the requirements for the degree of

**Doctor of Economics (Dr. rer. pol.)**

**Reviewers:**

Prof. Dr. Marcel Smolka

Prof. Dr. Nils Karl Sørensen

**Date of Defense:** 3 July 2025

**Year of Publication:** 2025

**Ali ibn Abi Talib**

*"There is no wealth like intellect, no poverty like lack of knowledge, no inheritance like good manners, and no support like consultation."*

# Acknowledgements

I would like to express my heartfelt thanks to everyone who supported me throughout my academic journey. First and foremost, I am very grateful to Prof. Marcel Smolka, my supervisor. It has been an incredible experience, and I have learned so much along the way. I also want to thank Prof. Nils Karl, my second supervisor. We first met when I started my PhD, and I enjoyed the many interesting discussions we had over lunch. A special thanks to Prof. Bart Los for inviting me to the Department of Economics at the University of Groningen. The research visit there was a great opportunity, and I truly appreciated the experience. I would like to thank Prof. Ulrich Glassmann and his partner for being like family to me, offering support and friendship, particularly during difficult times.

I am grateful to my friend Robin Konietzny, whom I met during my research visit. Our shared love for cooking, delicious meals, and engaging conversations made our time together truly memorable. Yolla Ahmad, my best friend, has been with me since I started my academic journey in Europe. She has always been there for me, supporting me and making me believe in myself. I truly appreciate her pride in me. I would also like to thank my friend Max Ahlemann, whom I first met in Glasgow. Over the years, he has been a great source of support, and we've had a lot of fun traveling together across Europe and within Germany during my PhD. Thanks to Mohamad Ebadi, who has always supported me in my studies and been a great friend.

Finally, a big thank you also goes to my siblings and the rest of my family, who have always believed in me.

# Dedication

*To Professor Toufic Salloum, whose constant support has been the foundation of my academic journey and personal growth. His guidance has shaped not only my academic thinking but also the way I approach life.*

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Industrial organization examines various aspects of market structure and firm behavior, with firm boundaries and market power being two key concepts. Firm boundaries define the scope of activities conducted internally (vertical integration) versus those outsourced to external suppliers. Studying the determinants of firm boundaries traces back to Coase (1937), which established the economic foundation for analyzing the factors influencing firm organization and structure. Since then, economists have sought to explain how firms strategically structure their operations to achieve efficiency and competitive advantage. Theories of firm boundaries and ownership structures have evolved, particularly in response to antitrust policies, to clarify the determinants of these decisions. Among them, the Property Rights Theory (PRT), developed by Grossman and Hart (1986) and Hart and Moore (1990), provides a formal framework to analyze the trade-offs involved in vertical integration and contractual arrangements.

The PRT approach focuses on incomplete contracts and ex-post bargaining, predicting their impact on ex-ante investment in noncontractible assets. It links the allocation of property rights to changes in investment. However, vertical integration isn't always a solution to asset specificity and may reduce investment below market levels. The PRT suggests that vertical integration boosts investment for only one party, enhancing its ex-post bargaining power and surplus. Ultimately, the benefits of integration depend on the importance of each party's ex-ante investments for the final output.

A key question in this context is: What determines a firm's decision to internalize certain transactions while outsourcing others? This decision is influenced by a range of factors, including transaction costs (as emphasized by Transaction Cost Economics (TCE)), asset specificity, market power, risk and uncertainty, and the regulatory and institutional environment. However, in this thesis, I specifically focus on examining how technological advancements, contractual frictions, and financial constraints influence firms' decisions regarding organizational structure.

As technology evolves, firms may seek to internalize production processes in order to capture the potential efficiency gains from innovation, reduce costs, and sustain competitive advantages. Recognizing that innovation and technological change can significantly influence firms' decisions on vertical integration versus outsourcing, it is crucial to understand how technological advancements drive these organizational choices. Additionally, incomplete contracts or enforcement issues can prompt firms to integrate vertically to mitigate inefficiencies or potential disputes. However, to address hold-up problems, firms often need to make transfers as a cost of integrating with other parties. In the absence of sufficient financial resources, these transfers can become prohibitively expensive, potentially deterring firms from pursuing vertical integration.

Another important aspect of industrial organization that I explore in this thesis is market power and firm markups. Market power refers to a firm's ability to influence prices and output levels within a market. As such, dominant firms with substantial market power can set prices above competitive levels, either by raising markups or benefiting from declining marginal costs. Data from the five largest European countries over the period from 1998 to 2019 highlight two main trends. First, there has been an increase in overall industry concentration, indicating that fewer firms now hold larger market shares across different sectors. Second, a larger share of industries is becoming highly concentrated, with a small number of firms dominating the market within those industries. Specifically, the share of industries where the largest four firms control at least 50% of the market has doubled from 16% to 37% (Koltay et al., 2023).[1]

Moreover, technological advancements have intensified this trend by benefiting early adopters and larger firms, enabling them to scale more efficiently, boost productivity, and strengthen their pricing power. Firms that fail to adapt to technological change risk losing market share and profitability. In this context, automation is increasingly transforming industrial sectors by enhancing operational efficiency and reducing costs. Through the adoption of robotics, artificial intelligence (AI), and machine learning, firms can streamline production, improve product quality, and strengthen their competitive positioning. This raises a crucial question: What factors drive markups, and how do automation shape pricing and cost structures?

This thesis addresses the aforementioned questions across three chapters, each examining

[1]The five largest European countries are: Germany, Spain, France, Italy, and the United Kingdom.

a key aspect of firm behavior and strategic decision-making. The first two chapters focus on identifying the critical factors that determine a firm's decision regarding the trade-off between vertical integration and outsourcing. The third chapter analyzes the second question and explores the impact of automation on pricing, marginal costs, and markups.

In Chapter I: ***"Vertical Integration, Technology, and Domestic Versus Foreign Supply"***, I study how technology intensity and the importance of both domestic and foreign inputs in production affect vertical integration decisions. In the model of Acemoglu et al. (2010), input cost shares, which reflect input importance, are calculated based solely on domestic inputs, without considering the role of foreign suppliers. However, firms source inputs both domestically and internationally. Ignoring foreign suppliers can obscure the impact of input cost shares on investment decisions and vertical integration. To address this, I adopt the vertical integration model from Acemoglu et al. (2010) and distinguish between domestic and foreign inputs, focusing on the decisions of domestic manufacturing firms. Thus, I examine how the domestic trade-off between vertical integration and outsourcing is influenced by the presence of foreign suppliers.

The contribution of this paper is twofold. First, I provide a more nuanced interpretation of the domestic trade-off between vertical integration and outsourcing by considering the cost shares of both domestic and foreign inputs. While Acemoglu et al. (2010) focus on domestic sourcing and Berlingieri et al. (2021) on international sourcing using trade data, my analysis offers a more comprehensive understanding of how the importance of domestic versus foreign inputs, alongside technology intensity, influences organizational decisions.

Second, I contribute to the literature by analyzing German manufacturing firms from 2009 to 2018. While previous studies have examined countries such as the UK, Canada, and France or taken a broader cross-country perspective, this paper focuses on Germany, one of Europe's leading economies. Understanding whether the patterns of vertical integration observed elsewhere hold for German firms is essential, given the country's strong industrial base.

To empirically test the effects of technology intensity and input cost share on vertical integration, I use three distinct data sources: plant-level data from Germany's Official Firm Data (AFiD), the OECD's harmonized input-output table for Germany, and the OECD's Analytical Business Enterprise Research and Development (ANBERD) database for R&D expenditures, which serves as a proxy for technology intensity. My empirical analysis shows that input cost shares significantly interact with R&D intensity in shaping vertical integration decisions.

Specifically, a higher domestic input cost share increases the likelihood of vertical integration when the manufacturer's R&D intensity is greater than that of the supplier. Conversely, an increase in the foreign input cost share reduces the likelihood of vertical integration when the manufacturer's R&D intensity surpasses that of the supplier.

Moreover, I performed several robustness checks, including a panel analysis (2009–2018) to capture time trends, as well as an alternative measure of technology intensity using physical investment intensity instead of R&D. These checks provide strong support for the main findings, as the key relationships remain consistent across various specifications. Although some effects lose statistical significance, the overall directionality of the results remains in line with theoretical expectations, reinforcing the robustness of the analysis. These results align with prior research on vertical integration, particularly regarding input importance and investment intensity (Acemoglu et al., 2010; Berlingieri et al., 2021; Alfaro et al., 2024; Egger et al., 2023), and offer important insights into how domestic and foreign input sources, alongside technological factors, influence organizational decisions in manufacturing firms. While Chapter I's theoretical model assumes perfect financial markets, Chapter II relaxes this assumption by exploring financial imperfections, including credit constraints, that influence the organizational decisions of contracting parties.

Contractual and financial factors play a crucial role in shaping firm boundaries, yet the impact of these factors on firm structure remains inconclusive. Chapter II: ***"Vertical Integration under Contractual and Financial Frictions"*** examines how contractual and financial market frictions affect vertical integration decisions by the contracting parties. I build on a model of vertical integration under contractual and financial frictions proposed by Acemoglu et al. (2009) and apply it empirically to analyze German data. The model investigates the supplier's decision to integrate forward with the manufacturer, where integration incurs fixed transaction costs that depend on the quality of contractual institutions and the level of financial development. These costs play a significant role in shaping whether firms choose integration or outsourcing.

I contribute to the literature by analyzing both firm- and plant-level data. Since some firms in the data operate multiple plants with distinct characteristics, using plant-level data allows for a more granular analysis of integration choices. Moreover, while financially developed economies generally benefit from easier access to capital, they are still vulnerable to financial market imperfections. The 2008 financial crisis demonstrated how even advanced financial systems can face disruptions, leading firms to encounter credit constraints, higher interest rates, and stricter lending conditions. To move beyond broad cross-country comparisons of financial development, I focus on specific financial frictions within the German manufacturing sector, using the financial crisis as an exogenous shock to credit supply. Additionally, previous studies have relied on the external financial dependence (EFD) measure introduced by Rajan and Zingales (1998); in contrast, I introduce the financial dependence ratio derived from the input-output table, offering an alternative industry-level measure. Finally, by employing a Triple-difference approach, I provide causal evidence on how financial constraints, particularly in industries reliant on external finance, shape firm boundaries.

The central question addressed in this chapter is: What is the combined effect of contractual and financial frictions on vertical integration decisions among firms in Germany? To answer this question, I empirically apply the model from Acemoglu et al. (2009), providing new evidence on the role of these frictions in shaping vertical integration decisions in a developed economy. When contracting costs increase, the potential for opportunistic behavior rises, making vertical integration a tool to mitigate the holdup problem. However, financial frictions can discourage vertical integration by limiting firms' ability to invest or effectively manage the transaction costs associated with it.

For the empirical analysis, I use micro-level data from Germany's Federal Statistical Office (Amtliche Firmendaten in Deutschland AFiD) for the years 2003-2018 and combine it with the OECD input-output table to create a firm-level measure of vertical integration. This is done by matching firms with their plants and interacting input cost shares with an indicator for firms operating plants in both supplying and manufacturing industries. To measure contractual frictions, I use the Product Complexity Index (PCI), which gauges industry-level complexity and technological advancement, obtained from the Growth Lab at Harvard University. To assess an industry's external financial dependence, I calculate the financial dependence ratio using input-output tables. Specifically, I divide the input from the financial services sector by the total cost share of the industry, obtaining the share of financial services in production costs. I primarily use data from before 2008 to avoid the effects of financial shocks and endogeneity. Finally, to obtain a time-invariant measure, I compute the median financial dependence ratio for each industry at the two-digit level classification across all available years.

I investigate the impact of financial frictions by using the financial crisis as an exogenous treatment variable. The treated group consists of firms within industries heavily reliant on external finance, while the control group comprises firms in industries less dependent on external finance. Another exogenous source of variation is the product complexity measure, which is based on diverse technological advancements and productive know-how across various countries, reflecting an external framework beyond the direct control of individual firms. Using a Triple-difference framework, I interact the financial crisis dummy with the financial dependence ratio and PCI to provide empirical evidence on the combined effect of contractual and financial frictions on vertical integration decisions in Germany.

The theoretical predictions of the model show an ambiguous relationship between product complexity and vertical integration. Empirically, the results confirm that the effect of contractual frictions on integration varies depending on model specifications, suggesting that this relationship is not straightforward. However, the model predicts a clear negative relationship between financial frictions and vertical integration, with empirical evidence supporting this prediction. Higher financial constraints discourage firms from integrating, as they lack the capital to cover transaction costs associated with integration decisions.

Furthermore, when both contractual and financial frictions are present, firms may find that the combined transaction costs of integration outweigh the potential efficiency gains. In these cases, firms prefer outsourcing. Using the Triple-difference approach, the empirical results show that a one standard deviation increase in product complexity, combined with a financial crisis, leads to a stronger negative effect on vertical integration, particularly for firms in industries highly dependent on external finance.

I conduct the analysis first at the firm level and then at the plant level, finding similar results at both levels. For robustness checks, I use the EFD measure instead of the financial dependence ratio, incorporating the industry-level EFD measure from Eppinger and Neugebauer (2022), which provides a similar conclusion. Additionally, I extend the crisis period to 2008-2010 and find consistent results.

Chapter III, ***"Competition, Markups, and Automation"*** (with Marcel Smolka) examines firm pricing strategies, shifting the focus from firm boundaries to market power. In the first part of the paper, we analyze the evolution of market shares and market concentration from 1990 to 2018 using firm-level survey data along with official industry data. The firms in our survey data set (*Encuesta Sobre Estrategias Empresariales, (ESEE)*) are free to define markets on their own terms, allowing for a more flexible and comprehensive view of market dynamics. We supplement this with industry-level data from the Spanish National Statistics Office (INE), which includes information on the number of firms, firm size distribution, and revenues. Our approach differs from other types of analyses by considering how firms define their own market boundaries, whether by product type, customer groups, or other characteristics. By combining these different data sources, we gain a macro view of the evolution of market shares and market concentration over the past three decades.

Our findings show that markets targeted by Spanish firms have become significantly more competitive over the period from 1990 to 2018. Furthermore, the top firms in the manufacturing sector have experienced stagnation in their market shares over time, which stands in contrast to the trend observed in other advanced economies. In particular, unlike in the United States (US), where market concentration has increased due to globalization and technological advancements reinforcing the power of top firms, Spain's top firms have not been able to capture larger market shares.

In the second part of the paper, we examine the evolution of firm-level markups, focusing on how output prices and marginal costs have changed from 1991 to 2018. We decompose changes in markups into changes in output prices and marginal costs using data from the *ESEE* survey, which allows us to track output price changes directly. This is possible due to the availability of the first-order condition of a variable input in the firm's cost minimization problem. By disentangling price and cost changes, we aim to better understand the underlying factors driving markup changes over time. We also explore how markups have evolved not just on average, but also at the top of the markup distribution, to investigate potential "superstar effects" in the Spanish manufacturing sector.

The results indicate that output prices and marginal costs for firms have evolved in tandem since 1991, such that the average markup (price over marginal cost) in the manufacturing industry has remained relatively stable. Importantly, we do not observe rising markups among top firms in Spain over the period from 1991 to 2018, which again contrasts with trends observed in other countries, particularly the U.S., where rising markups among the largest firms have been well-documented.

In the final part of the paper, we focus on the relationship between automation and markups, analyzing firm-level data on the use of robots in the production process over nearly three decades. The *ESEE* survey data allows us to track automation adoption and its impact on firm-level prices, costs, and markups. We investigate the impact of automation on a firm's marginal cost and output price, emphasizing how automation influences the firm's markup. Specifically, we examine whether automation leads to reductions in marginal costs and whether these reductions lead to higher or lower markups.

Our findings show that automation drives productivity gains, which in turn lead to higher markups at the firm level. Specifically, we find that automation reduces a firm's marginal cost, which translates into a reduction in output prices. However, the price reduction is less than proportional, meaning that the firm's markup rises. This highlights that automation is the driving force behind the increase in markups, with productivity improvements being the channel through which automation affects the firm's pricing strategy.

## Bibliography

**Acemoglu, Daron, Rachel Griffith, Philippe Aghion, and Fabrizio Zilibotti**, "Vertical Integration and Technology: Theory and Evidence," *Journal of the European Economic Association*, 2010, *8* (5), 989–1033. Publisher: Oxford University Press.

**_ , Simon Johnson, and Todd Mitton**, "Determinants of Vertical Integration: Financial Development and Contracting Costs," *The Journal of Finance*, 2009, *64* (3), 1251–1290. Publisher: [American Finance Association, Wiley].

**Alfaro, Laura, Nick Bloom, Paola Conconi, Harald Fadinger, Patrick Legros, Andrew F Newman, Raffaella Sadun, and John Van Reenen**, "Come Together: Firm Boundaries and Delegation," *Journal of the European Economic Association*, February 2024, *22* (1), 34–72.

**Berlingieri, Giuseppe, Frank Pisch, and Claudia Steinwender**, "Organizing Global Supply Chains: Input-Output Linkages and Vertical Integration," *Journal of the European Economic Association*, June 2021, *19* (3), 1816–1852.

**Coase, R. H.**, "The Nature of the Firm," *Economica*, 1937, *4* (16), 386–405. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0335.1937.tb00002.x.

**Egger, Peter H., Katharina Erhardt, and Gerard Masllorens**, "Backward versus forward integration of firms in global value chains," *European Economic Review*, 2023, *153*, 104401. Publisher: Elsevier.

**Eppinger, Peter S. and Katja Neugebauer**, "External financial dependence and firms' crisis performance across Europe," *Empirical Economics*, February 2022, *62* (2), 887–904.

**Grossman, Sanford J. and Oliver D. Hart**, "The costs and benefits of ownership: A theory of vertical and lateral integration," *Journal of political economy*, 1986, *94* (4), 691–719. Publisher: The University of Chicago Press.

**Hart, Oliver and John Moore**, "Property Rights and the Nature of the Firm," *Journal of political economy*, 1990, *98* (6), 1119–1158. Publisher: The University of Chicago Press.

**Koltay, Gábor, Szabolcs Lorincz, and Tommaso Valletti**, "Concentration and Competition: Evidence From Europe and Implications For Policy," *Journal of Competition Law & Economics*, October 2023, *19* (3), 466–501.

**Rajan, Raghuram G. and Luigi Zingales**, “Financial Dependence and Growth,” *The American Economic Review*, 1998, *88* (3), 559–586. Publisher: American Economic Association.

# Chapter 2

# Vertical Integration, Technology, and Domestic Versus Foreign Supply

### Abstract

In this paper, I explore the domestic trade-off between outsourcing and vertical integration. I adapt a standard model of vertical integration and differentiate between domestic and foreign inputs. Using Property Rights Theory (PRT), the model predicts that vertical integration depends on the technological intensity of contracting parties and the relative importance of domestic versus foreign inputs. Empirically, I test these predictions using plant-level data from Germany between 2009 and 2018. The findings show that a higher domestic input cost share and greater R&D intensity of the manufacturer increase the likelihood of vertical integration, while a higher foreign input cost share and greater R&D intensity of the manufacturer increase the likelihood of outsourcing. This result is reversed for the supplying industry. Overall, the results in this paper emphasize the significance of the PRT framework in understanding the vertical integration decisions of the contracting parties.

**Keywords:** Investment intensity, R&D intensity, input cost share, vertical integration, outsourcing, foreign supply.

# 1 Introduction

The procurement of inputs is crucial for a firm's performance and plays a pivotal role in shaping its boundaries. When determining where to obtain these inputs, firms face a two-dimensional decision problem: where to source the inputs and how to structure ownership. On the one hand, firms may opt to acquire inputs from external sources through market transactions, which can be done locally (domestic outsourcing) or internationally (foreign outsourcing). On the other hand, they might choose to bring input production in-house, either through foreign direct investment (FDI) for overseas production or through vertical integration for domestic production.

In this paper, I argue that the domestic trade-off between vertical integration and outsourcing is influenced by the presence of foreign suppliers. In their model, Acemoglu et al. (2010) take into account that inputs flow domestically when calculating input cost shares, and do not focus on the role of the foreign suppliers. In fact, firms access both domestic and foreign suppliers.

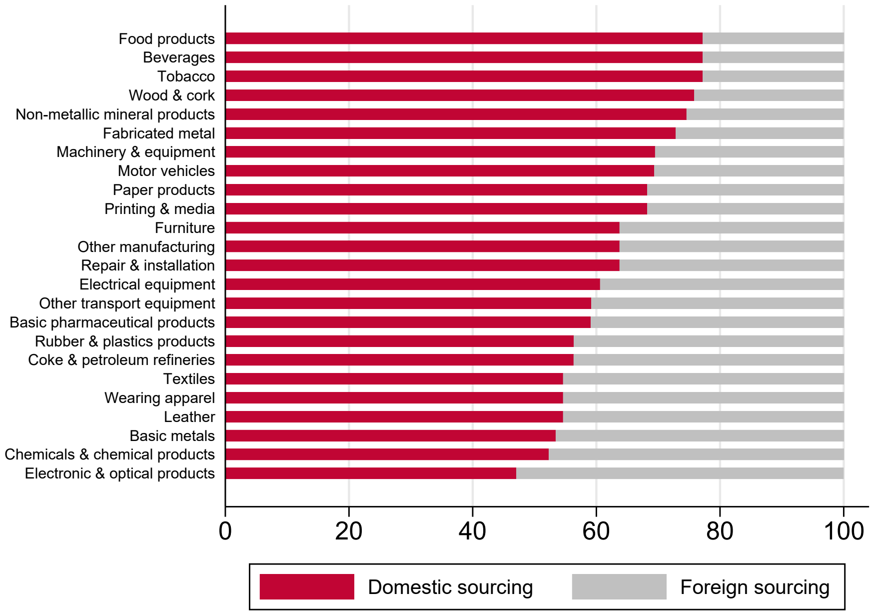
Figure 1 presents the average share of domestic and foreign sourcing across manufacturing industries taken from the German input-output tables between 2009 and 2018. The figure shows that all manufacturing industries in Germany procure intermediate inputs from both domestic and international suppliers, albeit with varying shares. Interestingly, these proportions reflect heterogeneity across industries, as some industries are more globalized than others. Among these industries, those with the highest reliance on foreign inputs include *Electronic & optical products*, with a remarkable 53% of inputs sourced from abroad, closely followed by *Chemicals & chemical products* at 48%. These sectors demonstrate a significant dependency on foreign suppliers for their production processes. Conversely, industries with the lowest reliance on foreign inputs include *Wood & Cork* and *Tobacco*, with averages of 24% and 23%, respectively, sourced from foreign suppliers. [1]

Therefore, the empirical data presented in Figure 1 reveals that producers obtain inputs from suppliers that may be located either within the country or abroad. Considering domestic inputs while neglecting the availability of accessing foreign suppliers would obscure the impact of input cost shares on the optimal level of investment and thus on vertical integration decisions. Since the manufacturer sources its inputs domestically and abroad, the importance of the input supplied must be put into the right perspective. Therefore, what should matter is the importance of domestic inputs compared to foreign inputs in relation to technology intensity and vertical integration.

The Property Rights Theory (PRT) focuses on how ownership of assets and control rights influence ex-ante investment incentives, highlighting both the benefits and costs of vertical integration. The theo-

[1] Note that some industries in the input-output tables, such as *Tobacco*, *Food products*, and *Beverages*, are grouped together. The data available for these industries is not at the sector level but at the group level. Hence, the values of these industries are similar and represent the average value of sourcing across the three industries within the same group.

**Figure 1:** Share of inputs sourced domestically and abroad across industries in Germany (%)




*Notes*: Figure 1 shows the percentage of inputs sourced domestically and abroad across industries in Germany, averaged over the years 2009–2018.
*Source*: Based on the author's calculations using the German input-output tables.

retical framework examines the impact of the investment intensity of the contracting parties and the importance of their products for production on vertical integration. While the model assumes relationship-specific investments, it emphasizes the role of technology intensity—particularly R&D. This is based on the presumption that firms engaging in technology investments, particularly in R&D, face similar challenges of potential holdup and opportunistic behavior—problems highlighted by both the PRT and Transaction Cost Economics (TCE) approaches.

Acemoglu et al. (2010)'s model leads to key predictions underlying the property-rights framework. First, technology intensities of the manufacturer and the supplier have contrasting effects on the probability of vertical integration. Second, the more important the supplier's input in the relationship, reflected in the proportion of the manufacturer's costs that are attributable to the supplier's inputs, the more likely vertical integration is to occur. Third, there is an important interaction effect between input cost share and technology intensity. In particular, higher technology intensity of the manufacturer, lower technology intensity of the supplier and higher input cost share increase the likelihood of vertical integration.[2]

In this paper, I adapt Acemoglu et al. (2010)'s model by incorporating foreign suppliers. Specif-

[2]In addition, the model predicts that the degree of competition in the industry has a negative impact on vertical integration—a dimension that I do not explore in this paper.

ically, I distinguish between inputs sourced from domestic suppliers and those imported from foreign suppliers. This approach allows for a more nuanced understanding of how domestic manufacturers make decisions about their organizational structure decisions. Manufacturers may choose to source inputs from both local and international markets depending on factors such as the importance of the input for production of the final good. The importance of inputs not only affects the integration decision but also influences investment incentives under different organizational structures. If the domestic input is less crucial to the production of the final output compared to the foreign input, the manufacturer will invest more under outsourcing. Conversely, the importance of domestic inputs tends to encourage vertical integration, whereas the significance of foreign inputs favors outsourcing. Most importantly, there are notable interaction effects between input cost share and the investment intensity of the contracting parties.

This study aligns with a body of literature examining the relationship between the investment intensity of contracting parties and firm boundaries through the property rights approach. It also connects closely with research on the role of input cost shares in firms' decisions to vertically integrate or outsource. Acemoglu et al. (2010) provide empirical support for the theoretical predictions discussed earlier by analyzing cross-sectional data on all UK manufacturing plants and integrating it with the UK input-output table. Their findings indicate that the R&D intensity of producing industries is positively correlated with the likelihood of vertical integration, while the R&D intensity of supplier industries is negatively correlated. These correlations are more pronounced when the supplying industry constitutes a significant portion of the manufacturer's costs.

Lileeva and Van Biesebroeck (2013) build on the framework established by Acemoglu et al. (2010), focusing on the relative investment intensity—differences in investment levels between suppliers and manufacturers. They also introduce a measure of technological interrelatedness, which underscores the strong link between the investments of the two parties involved. The arguement is that firms with similar production technologies are more likely to experience technological spillovers. Analyzing data from a sample of large Canadian manufacturing firms, the authors find a positive relationship between the manufacturer's investment intensity and vertical integration, particularly when transactions involve unrelated technologies.

More recently, Liu (2021) and Egger et al. (2023) use cross-country firm-to-firm linkage data to differentiate between forward and backward integration and outsourcing. Building on Grossman and Hart (1986) and the frameworks of Antràs (2003) and Antràs and Helpman (2004), Liu (2021) introduce a bi-directional integration model that examines how both buyer and seller firms influence ownership structures. They differentiate between homogeneous and heterogeneous firms and explore three key ownership structures: forward integration, backward integration, and outsourcing. Empirical analysis of

209,062 buyer-seller relationships across 154 countries supports the models and explains the coexistence of backward and forward integration. The results indicate that an increase in the buyer's relative R&D intensity is associated with a 6.8% higher likelihood of choosing buyer integration and a 2.4% lower likelihood of choosing seller integration. Additionally, high-productivity firms are more likely to choose buyer integration when the buyer's relative R&D intensity is high and seller integration when the R&D intensity is low, compared to low-productivity firms.

Egger et al. (2023) present a modified ,model based on Acemoglu et al. (2010)'s framework to explore the determinants of asset ownership. They explicitly account for the fixed costs associated with vertical integration decisions. Consistent with Acemoglu et al. (2010), Egger et al. (2023) find that vertical integration (both backward and forward) is influenced by the relative investment intensities of the supplier and the manufacturer. They also show that higher input dependence leads to increased vertical integration. Furthermore, Egger et al. (2023) demonstrate that bilateral investment treaties, which serve as a proxy for investment costs, result in more integration in both forward and backward directions.

Furthermore, Berlingieri et al. (2021) and Alfaro et al. (2024) focus on the technological importance of inputs, as reflected by the input cost share. Berlingieri et al. (2021) examine the influence of the technological importance of inputs on a multinational firm's choice to source them from either an international affiliate or an independent supplier. Utilizing French import data, they reveal that inputs crucial for a multinational's output are significantly more likely to be internally sourced. Additionally, they demonstrate that inputs with greater technological importance are more frequently acquired from affiliated entities. They also explore potential factors influencing the outsourcing decision, finding that the relationship between input importance and vertical integration probability is amplified by the contracting environment as well as headquarters intensity.

In a recent study, Alfaro et al. (2024) investigate the decision of firms to vertically integrate suppliers, outsource, or delegate. Their model generates two key predictions: first, as an input becomes more valuable to a firm, headquarters are more likely to delegate responsibilities to integrated suppliers providing that input; second, suppliers of inputs with higher value are more prone to integration. Using firm-level data from 20 countries, the authors find that manufacturers of final goods typically integrate suppliers of more valuable inputs. Additionally, within integrated suppliers, those providing more valuable inputs were given more autonomy.

My contribution to this literature is as follows. First, I provide additional insights into the trade-off between domestic vertical integration and outsourcing by interpreting the domestic input cost share in relation to the foreign input cost share. I address this issue by differentiating between domestic and foreign inputs. While Acemoglu et al. (2010) focus on domestic sourcing, Berlingieri et al. (2021) focus solely on international sourcing using international trade data. However, their findings remain robust regardless of whether a firm predominantly sources its inputs from within the EU (domestic sourcing) or from international suppliers. My analysis aims to provide a comprehensive understanding of the impact of both domestic and foreign inputs on vertical integration decisions. By incorporating both aspects into my analysis, I can assess the relative importance of each source of input and how they influence optimal investment levels and decisions regarding vertical integration. Additionally, I focus on how the interaction between domestic and foreign input cost shares and the technology intensity of both the manufacturer and its supplier affects vertical integration and outsourcing decisions.

Second, previous studies on the relationship between contracting parties' investment and vertical integration have focused on countries such as the UK (Acemoglu et al., 2010), Canada (Lileeva and Van Biesebroeck, 2013), and France (Berlingieri et al., 2021), as well as on global cross-country data (Liu, 2021; Egger et al., 2023; Alfaro et al., 2024). This paper, however, examines the case of manufacturing firms in Germany between 2009 and 2018. Given Germany's status as one of the strongest economies in Europe, it is crucial to investigate whether the patterns of vertical integration observed in other countries also apply to German firms. Understanding this sheds light on the unique features of Germany's industrial structure and its implications for ownership structures.

To examine the effects of technology intensity and input cost share on vertical integration, I employ three distinct data sources. First, I utilize comprehensive data on plant operations across various manufacturing sectors sourced from the Official Firm Data in Germany (Amtliche Firmendaten in Deutschland (AFiD)). Second, I incorporate the OECD's harmonized national input-output tables for Germany, which describes the relationships among industries regarding input and output flows. It presents inputs from domestic and international suppliers for each manufacturing sector. The domestic matrix shows domestically sourced inputs from domestic suppliers, while the import matrix displays imported inputs from foreign suppliers used by domestic manufacturers. This provides insights into the percentages of domestically sourced versus imported inputs, aiding in distinguishing between domestic and foreign input cost shares. By combining plant-level data with the input-output tables, I construct a measure for vertical integration. Specifically, the index takes the value of one only if the firm has plants in both the manufacturing industry and the supplying industry, and zero otherwise. Lastly, I use the Analytical Business Enterprise R&D (ANBERD) database, which offers data on R&D expenditures at the industry level. R&D intensity, which proxies for technology intensity, is measured as the ratio of R&D expenditures to total sales using data from the pre-sample period.

In general, I find that there are important interaction effects between input cost shares and R&D intensity. I also document that input-output linkages play an important role in understanding organizational decisions across firms. These results are robust to a variety of different specifications. Precisely, the empirical analysis reveals the following:

1. A higher domestic input cost share increases the likelihood of vertical integration, especially for manufacturers with greater R&D intensity. Meanwhile, a higher supplier R&D intensity is associated with a reduced likelihood of vertical integration.

2. I find a strikingly similar pattern and an opposite effect when using the foreign input cost share. Specifically, an increase in the foreign input cost share is associated with a reduced likelihood of vertical integration for manufacturers with high R&D intensity, while a higher foreign input cost share increases the likelihood of vertical integration for manufacturers with low R&D intensity.

For example, consider the *Motor vehicle* industry, which is highly R&D intensive, sourcing inputs from the *Basic metals* industry, a sector with relatively low R&D intensity (see Figure 3). The motor vehicle manufacturer sources inputs both domestically and from foreign suppliers, including basic metal suppliers outside Germany. According to the findings, as the domestic input cost share increases, the motor vehicle manufacturer is more likely to opt for vertical backward integration with the domestic supplier. However, when the foreign input cost share surpasses the domestic input cost share, the manufacturer is more likely to outsource. This pattern demonstrates how input cost shares, in combination with R&D intensity, significantly influence the decision to integrate or outsource.

The rest of the paper is organized as follows. In Section 3, I describe the modified theoretical model and the main testable predictions. In Section 3, I discuss the data sources, and describe the main variables. I describe the empirical framework in Section 4. In Section 5, I present the main results along with additional analyses and robustness checks. I conclude in Section 6.

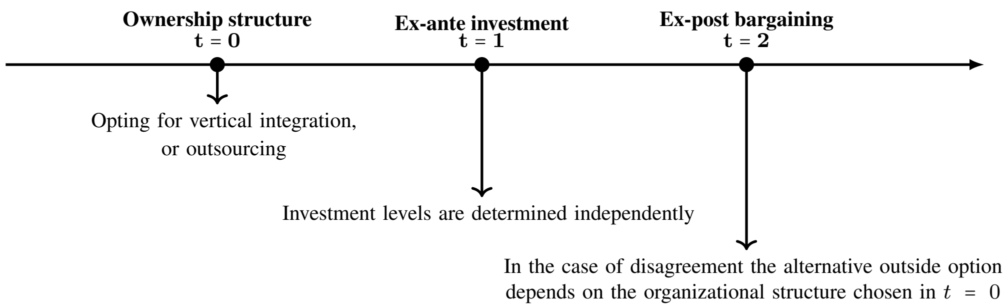
## 2 Theoretical model

According to the Property Rights Theory (PRT) approach, in situations where contracts are incomplete, parties engage in negotiations to determine the distribution of returns from investments. The significance of ownership lies in its role of enhancing a party's disagreement payoff and thus influencing its bargaining position. In this section, I describe a vertical integration model developed by Acemoglu et al. (2010) in line with the property rights approach proposed by Hart and Moore (1990). Within this framework, I examine the relationship between the investments undertaken by two risk neutral contracting parties, an upstream supplier and a downstream manufacturer, and their decision to vertically integrate (backward) or remain independent.

Due to data limitations, this study can only empirically examine vertical backward integration, as compared to outsourcing.[3] Therefore, despite Acemoglu et al. (2010) addressing both types, I focus solely on vertical backward integration, where the manufacturer integrates with the supplier.

Consider a domestic market where there is a supplier, $s$, who provides an intermediate input, $L_s$, that is used by a manufacturer, $j$, to produce a final product. The model assumes a sequence of events across three periods, which is depicted in Figure 2.

**Figure 2:** Time assumption




Source: Author's illustration.

**Ownership structure ($t = 0$).** It is assumed that the contracts are incomplete. In this sense, it is not possible for the two parties to conclude contracts specifying the ex-post division of the revenue or to provide detailed information about the relationship-specific investments. However, it is possible to create the right incentives for non-contractible investments by specifying the allocation of property rights, which ensures both parties are motivated to maximize and appropriately divide the surplus of the relationship. As such, the input supplier (seller) and the manufacturer (buyer) can choose to remain independent, or to integrate backward (where the manufacturer integrates the supplier).

The timing assumption requires that the two contracting parties decide on the organizational form of their relationship prior to investment and production. In this period, the supplier receives an offer from the manufacturer proposing a particular organizational structure $n$, where $n \in \{$Vertical Backward Integration (I), Outsourcing (O)$\}$. The offer also specifies the amount of ex-ante transfer received by the manufacturer $(T_j(n))$ and the supplier $(T_s(n))$, such that $T_j(n) + T_s(n) = 0$. One can interpret this transfer a participation cost. In the initial stage ($t = 0$), the producer proposes the ownership structure and production terms, which the supplier can accept or reject. For instance, if $n =$ I, the manufacturer gains possession of the input and does not need the supplier's approval to use it. However, if the supplier rejects the offer, the game ends with both parties receiving their respective outside options.

[3]This contrasts with recent studies, such as Liu (2021) and Egger et al. (2023), which differentiate between forward and backward integration.

**Ex-ante investment ($t = 1$).** After determining the ownership structure in stage $t = 0$, the manufacturer and the supplier decide on their respective investment levels, $i_j$ and $i_s$, simultaneously, where $i_j, i_s \geq 0$. Each party has its own resources (e.g. physical and human capital) and makes its own investment. It is not possible for the manufacturer to take over the supplier's investment and vice versa. These investments are specific to their respective relationship, which creates a risk of a holdup problem.

The relationship-specific production technology is as follows:

$$Q(L_s, i_j, i_s) = \phi L_s(\lambda_j i_j + \lambda_s i_s + 1) + (1 - \phi)(\lambda_j i_j + 1). \tag{2.1}$$

This production function depends on whether the supplier provides its input to the manufacturer, in this case $L_s = 1$, and $L_s = 0$ otherwise. Additionally, $\phi \in [0, 1]$ is the share of inputs provided by the supplier in the manufacturer's total costs (total input cost share). Moreover, the value added of the relationship depends on the level of investment of the manufacturer $i_j$ and the supplier $i_s$. Finally, the parameters $\lambda_j$ and $\lambda_s$ indicate the marginal product of the two parties' investments.[4]

I interpret the parameter $\phi$ the share of the manufacturer's total input costs attributable to inputs provided by the domestic supplier. The remaining fraction, $1-\phi$, represents inputs sourced from foreign suppliers in international markets. Thus, the parameter $\phi$ reflects the degree of reliance on domestic suppliers for the necessary inputs in the production process, while $1 - \phi$ accounts for inputs sourced from abroad. In this context, $\phi$ serves as a measure of the input cost share, highlighting the relative importance of domestically supplied inputs. A higher value of $\phi$ suggests that domestic inputs are relatively more important in the production process compared to imported inputs. Conversely, a lower value of $\phi$ indicates that the production process relies more heavily on imported inputs than on domestic ones. I assume there is a continuum of foreign suppliers providing inputs used by domestic manufacturers, with the foreign supplier's investment considered exogenous. Additionally, the manufacturer imports standardized inputs from abroad while relying on domestic suppliers for more relationship-specific inputs.

Equation 2.2 represents a simple quadratic form for the investment costs ($\omega$) incurred by the two parties and are given by:

$$\omega_j = \frac{i_j^2}{2} \quad \text{and} \quad \omega_s = \phi\frac{i_s^2}{2}. \tag{2.2}$$

Here, the investment cost function of the domestic supplier depends on $\phi$. This is to avoid implicit economies of scale in the supplier's cost function.

[4]For simplicity, the model assumes that there is no complementarity between $\lambda_j$ and $\lambda_s$. Thus, increasing investment levels by one party does not necessarily increase the marginal value by increasing the level of investment of the other party.

**Ex-post bargaining ($t$ = 2).** At this stage of the game, investments are sunk. In the case of a conflict arising between the two parties, they engage in bargaining to resolve it. If there is a disagreement, the manufacturer and the supplier possess an outside option denoted by $Z_j^n$ and $Z_s^n$, respectively. These options depend on the organizational structure chosen in $t = 0$. Consequently, each party has its unique outside option as follows:

**1. Backward vertical integration (I):** Under this particular organizational structure, the manufacturer possesses residual control rights over the supplier's input. The supplier exits the relationship without any post-breakup entitlements. Thus, its outside option is zero. Although the manufacturer keeps all the assets and the relationship-specific input for itself, it will not keep the entire effective value of the investment made by the supplier. This can be attributed, for instance, to a coordination issue between the two parties, where the supplier is redundant in their cooperation with the manufacturer. Under vertical integration, the outside options are

$$Z_j^I(i_j, i_s) = Q(L_s = 1, i_j,\ (1-\sigma)i_s)$$

$$Z_s^I(i_j, i_s) = 0,$$

where $\sigma \in [\ 0, 1)$ is the proportion of the supplier's investment that cannot be retained by the manufacturer.

**2. Outsourcing (O):** Under outsourcing, the manufacturer cannot access the relationship-specific investment in case of disagreement. Consequently, the disagreement leads to a reduction in output for products that rely on those particular inputs. For the supplier, selling the input outside the relationship is costly, as it loses part of the revenue, $\theta$, due to the specificity of its input. In this model, $\theta$ is treated as an exogenous factor, which is empirically represented by the ratio of manufacturers to suppliers in the market; the larger the number of manufacturers (potential buyers) in the market, the smaller the loss ($\theta$).[5] The outside options under outsourcing for both the manufacturer and the supplier are:

$$Z_j^O(i_j, i_s) = Q(L_s = 0, i_j, i_s) = (1-\phi)(\lambda_j i_j + 1)$$

$$Z_s^O(i_j, i_s) = \theta\phi(\lambda_s i_s + 1). \tag{2.3}$$

In the context of these negotiations, the two parties then bargain over sharing the revenue based

[5] In the empirical analysis, I abstain from directly measuring or examining $\theta$. Acemoglu et al. (2010) investigates $\theta$ by obtaining the relative number of manufacturers to suppliers. Due to data limitations, I lack information on the specific number of firms within the manufacturing and supplying industries.

on the organizational form $n$ according to the Nash bargaining solution. Trading in this relationship provides benefits for both parties, leading to an agreement being struck. Ultimately, after reaching an agreement, production takes place, and the manufacturer and the supplier share the specified output. The ex-post total output is determined by the outside options of the parties involved and is based on the symmetric Nash bargaining solution:

$$y_c^n(i_j, i_s) = Z_c^n(i_j, i_s) + \frac{1}{2}[\,Q(L_s = 1, i_j, i_s) - Z_j^n(i_j, i_s) - Z_s^n(i_j, i_s)], \tag{2.4}$$

where $y_c^n(i_j, i_s)$ denotes the payoff that remains with party $c \in \{j, s\}$ under organizational form $n$ after symmetric Nash bargaining. Bargaining occurs over the relationship-specific surplus: $[\,Q(L_s = 1, i_j, i_s) - Z_j^n(i_j, i_s) - Z_s^n(i_j, i_s)]$. Each party's share of the revenue increases with its own outside option and decreases with the other party's outside option. This is a direct consequence of the Nash bargaining solution, which allocates the surplus in proportion to the bargaining power of the parties, determined by their outside options. Finally, the surplus that each party generates depends on its output, investment cost, and transfers under each organizational structure:

$$\pi_c^n(y_c(i_j, i_s), i_c) = y_c^n(i_j, i_s) - \omega_c + T_c(n) \tag{2.5}$$

## 2.1 Equilibrium

I begin by defining the total surplus $\Pi$ under each ownership structure in equilibrium as

$$\begin{aligned}\Pi^n &= \pi_j^n + \pi_s^n \\ &= \pi_j^n(y_j^n(i_j^*(n), i_s^*(n)), i_j^*(n)) + \pi_s^n(y_s^n(i_j^*(n), i_s^*(n)), i_s^*(n)),\end{aligned}$$

where terms with an asterisk $(*)$ represent equilibrium. Under equilibrium, $i_j^*(n)$ and $i_s^*(n)$ represent the optimal investment of the manufacturer and the supplier, respectively, given the organizational structure. Additionally, since $T_j(n) + T_s(n) = 0$, the total surplus is:

$$\Pi^n = Q(L_s = 1, i_j^*(n), i_s^*(n)) - \omega_j(n) - \omega_s(n).^{6} \tag{2.6}$$

Eventually, the subgame perfect equilibrium will choose the ownership structure that maximizes the total surplus, where $\Pi^{n^*} \geq \Pi^n$ for all $n \in$ {I, O}, and $n^*$ is the equilibrium organizational structure where

$$n^* = \underset{n \in \{I, O\}}{\operatorname{argmax}} \Pi^n$$

[6] These costs are associated with the equilibrium investment of both parties: $\omega_j(i_j^*(n))$ and $\omega_s(i_s^*(n))$.

The equilibrium investment level depends on the organizational structure, $i_c^*(n)$, and is determined by:

$$i_j^*(n) = \max_{i_j} \{y_j^n(i_j, i_s^*(n)) - \omega_j\}$$

$$i_s^*(n) = \max_{i_s} \{y_s^n(i_j^*(n), i_s) - \omega_s\}$$

Profit maximization is determined by the Nash equilibrium investment levels. For each party, the level of investment increases as its ownership of all assets increases. These investment levels form the basis for explaining each party's decision regarding the optimal ownership structure.

Table 1 displays the investment levels in equilibrium for each party under each organizational structure. Clearly, the ownership structure has different effects on investment incentives. In equilibrium, the manufacturer invests the most under vertical backward integration ($i_j^*(I) > i_j^*(O)$). While vertical backward integration promotes the manufacturer's investment by increasing its share of the surplus, it reduces the supplier's incentive to invest ($i_s^*(O) > i_s^*(I)$) because its outside option is reduced.

**Table 1:** Optimal investment under different organizational structures

| **Organizational structure** | **Investment levels in equilibrium** | |
|---|---|---|
| | $i_j^*$ | $i_s^*$ |
| Backward vertical integration (I) | $\lambda_j,$ | $\frac{\sigma}{2}\lambda_s$ |
| Outsourcing (O) | $\left(\frac{2-\phi}{2}\right)\lambda_j,$ | $\frac{1+\theta}{2}\lambda_s$ |

*Notes*: This table shows the different investment levels in equilibrium for each of the manufacturer ($i_j^*$) and supplier ($i_s^*$) based on the organizational structure: vertical backward integration, and outsourcing.
*Source*: Author's derivations based on the theoretical model.

Under outsourcing, the manufacturer's optimal investment level is influenced by the input cost share, $\phi$. The incentive to invest decreases as $\phi$ increases, reflecting the growing importance of domestic, relationship-specific inputs for production. That is, as domestic inputs grow in importance relative to foreign-imported inputs, the investment level under domestic outsourcing declines. By contrast, a higher foreign input cost share, $1 - \phi$, leads to increased investment by the manufacturer. As for the supplier, the optimal investment level depends on the marketability of the customized input, $\theta$. When remaining independent, the supplier's investment increases in $\theta$: the higher the marketability of the input, the better the outside option. This is because higher $\theta$ means that the supplier can retain a large fraction of the returns on its investment.

It should be noted that the relationship between the input cost share and firm boundaries is in line with the traditions of transaction cost economics (TCE) in which the manufacturer wants to avoid ex-post inefficiencies by backward integrating with the upstream supplier. This contradicts the PRT prediction that high input cost share favors outsourcing production to independent suppliers. According to PRT, outsourcing can mitigate potential underinvestment by upstream suppliers. In these models, the focus is on examining the upfront investment made by a specific upstream supplier in relation to the investment made by the downstream firm in producing the input provided by the supplier. When the upstream investment offers a higher marginal contribution compared to the downstream investment, it indicates a likelihood of increased outsourcing. Moreover, when crucial investments are involved, the problem of supplier underinvestment becomes more pronounced, leading to a higher likelihood of outsourcing. However, the suppliers's investment cost in Equation 2.2 is convex, i.e., increasing marginal cost of investment. By introducing the term $\phi$ in the supplier's cost function the cost structure is adjusted to prevent implicit economies of scale. This adjustment ensures that as the supplier's investment increases ($i_s$), the cost does not decrease too rapidly, which would typically be due to economies of scale. This discourages the supplier from increasing its investment under any organizational structure. Thus, integration becomes more likely and less costly: the problem of underinvestment by the supplier no longer arises as long as the ownership structure does not matter for the investment.

## 2.2 Comparative statics

The total surplus $\Pi^n$ for all $n \in \{\mathrm{I, O}\}$ is obtained by substituting $i_j^*(n)$ and $i_s^*(n)$ in Equation 2.6. By subtracting the total surplus generated by vertical integration from the total surplus generated if the two parties remained separated, I obtain the additional surplus generated by backward vertical integration relative to outsourcing as follows:

$$\Delta^I = \frac{\lambda_j^2(\phi)^2}{8} - \frac{(1+\theta-\sigma)(3-\theta-\sigma)\phi\lambda_s^2}{8}$$

The model operates by analyzing the optimal investments made by the two contracting parties, where the equilibrium organizational structure depends on the relative returns to investment, defined as $\beta = \frac{\lambda_j}{\lambda_s}$, with $\lambda_j$ and $\lambda_s$ denoting the productivity of the manufacturer and supplier, respectively.

A key threshold value, denoted by $\beta^I$, characterizes the point at which the total surplus generated under outsourcing equals that under vertical integration. That is, when $\beta = \beta^I$, the manufacturer is indifferent between the two ownership structures. This threshold is given by:

$$\beta^I = \sqrt{\frac{(1+\theta-\sigma)(3-\theta-\sigma)}{\phi}}.$$

- **Prediction 1:** $\frac{\partial \Delta^I}{\partial \lambda_j} > 0, \frac{\partial \Delta^I}{\partial \lambda_s} < 0$

The first result shows that the relative returns to investment of the manufacturer and the supplier determines the probability of integration.[7] Differentiating the additional surplus achieved by backward integration, $\Delta^I$, shows that it increases with the relative marginal return to investment. This implies that the manufacturer's investment adds more value to the production of the final product. In addition, the importance of the input to the manufacturer has greater implications on the ownership structure. Given the other parameters, the contracting parties choose $n$ based on $\lambda_j$ and $\lambda_s$: vertical backward integration becomes the organizational structure when $\lambda_j$ is high (or $\lambda_s$ is low). The reason is that the manufacturer's investment has a higher marginal value in this case. Backward integration increases the outside option for the manufacturer, while that of the supplier decreases.

- **Prediction 2:** $\frac{\partial \beta^{BI}}{\partial \phi} < 0$

The relative importance of domestic inputs compared to foreign inputs affects the optimal investment levels in the absence of vertical integration. This highlights the need to understand the role of the foreign supplier in the manufacturer's decision to integrate with domestic suppliers, which essentially depends on the relative importance of the domestic input. The interpretation of $\phi$ (and subsequently $1 - \phi$) in this paper considers that the inputs required to produce the final product can be provided by both domestic and foreign suppliers. Accordingly, a fraction of the total input cost share is from domestic sources, while the other part is from foreign sources. Generally, the lower the importance of the input to the manufacturer, reflected in the input cost share, the higher the investment under outsourcing. Furthermore, a higher input cost share means that the manufacturer is highly dependent on the supplier's inputs, which increases the likelihood that the manufacturer will be held up by the supplier. Backward integration becomes a tool to avoid the hold-up problem. If the importance of the domestic input rises, the likelihood of opting for vertical integration increases, while an increase in the importance of the foreign input diminishes this probability. Thus, the availability of a more important input obtained from a foreign supplier not only increases the level of investment under outsourcing, but also makes vertical backward integration less likely.

- **Prediction 3:** $\frac{\partial^2 \Delta^I}{\partial \lambda_j \partial \phi} > 0, \frac{\partial^2 \Delta^I}{\partial \lambda_s \partial \phi} < 0.$

The model predicts that the magnitude of these effects is amplified by the interaction between the input cost shares and the return on investment of the contracting parties. This prediction suggests that when the relationship between the manufacturer and the supplier is more crucial, the investment intensities of

[7] Finding the derivative of the additional surplus with respect to the relative returns to investment could be also expressed as $\frac{\partial \Delta^I}{\partial \beta^I} > 0$.

both parties are likely to play a more substantial role in integration decisions. Furthermore, the higher the fraction of inputs coming from domestic (foreign) suppliers, and the higher the relative investment intensity of the manufacturer, the higher (lower) the probability of backward vertical integration.

# 3 Data and descriptive analyses

To evaluate the predictions of the stylized model, I use three datasets for the empirical analysis. First, I rely on the official firm-level data in Germany Amtliche Firmendaten in Deutschland (AFiD), provided by the German Federal Statistical Office and the Offices of the Laender (2021). This dataset, which has not been used in previous research on firm boundaries, provides comprehensive plant-level information. Second, I incorporate the OECD's harmonized national input-output tables for Germany. This dataset helps in constructing measures of vertical integration and provides details on input cost shares. Finally, I utilize the Analytical Business Enterprise Research and Development (ANBERD) database, which offers information on R&D expenditures. This data allows me to calculate R&D intensity, serving as a proxy for technology intensity.

## 3.1 Main data (AFiD-panels)

I use plant-level data (Industrial plants or *"Industriebetriebe"*) that contain information on a wide range of plant characteristics. The AFiD panels combine various individual sources to provide information on plant economic activities.[8] These surveys provide information on total production, sales, employment, wages, and business investment in tangible and intangible assets, among other variables. Participation in the survey is mandatory and covers all plants of firms with more than 20 employees. The data spans the period 2009-2018.[9]

The dataset primarily consists of plant-level observations, each associated with its respective firm. Plants are classified based on their major four-digit industrial activity according to the International Standard Industrial Classification of All Economic Activities Rev. 4 (ISIC).[10] Although the dataset includes mining and quarrying plants, my analysis focuses specifically on manufacturing plants, to ensure meaningful comparisons among manufacturing establishments. Additionally, I generate standard control variables commonly utilized in the literature, namely, firm age and size. Firm age denotes the number of years since the firm's first plant was recorded in the data. Firm size is derived from the firm's number

[8]The AFiD plant-level panel surveys are: (i) the Monthly Report for Establishments (MBB), (ii) the Annual Report for Plants (JBB), (iii) the Investment Survey (IEB), and (iv) the Production Survey (PE).

[9]For the analysis period, I use data from 2009 to 2018. However, for the construction of R&D intensity, I use pre-sample data from 2003 to 2008 as I will explain below.

[10]In the appendix, Table A2 shows the industries at the two-digit level, as classified according to ISIC Rev. 4.

of employees, calculated as the sum of all active employees across all firm plants as of September of the respective year.

The dataset used contains a total of 9,170,112 observations over the panel period. For the main empirical framework, I focus on a cross-sectional analysis comprising 1,355,448 observations. There are 61516 plants owned by 52636 firms. Furthermore, the dataset shows that single-plant firms predominated between 2009 and 2018, accounting for 74% of all firms. Typically, smaller firms outnumber larger ones, with the latter often being multi-plant firms, commonly engaged in vertical integration. This observation is not unique to the German manufacturing sector. Previous research has indicated that across various countries, firms with multiple plants constitute only a small fraction of the total number of firms within the sector. For instance, Bloom et al. (2012) analyze plant-level data from the United Kingdom, revealing that 84% of firms operate with a single plant. Similarly, Alfaro et al. (2016) utilize the WorldBase database, finding that 96% of firms are single-plant entities.

## 3.2 Input-output tables

The input-output tables I use span the period 2009-2018 and contain information on 24 manufacturing industries and 576 pairs of manufacturing (downstream) and supplying (upstream) industries for each year. These industries are reported according to the two-digit ISIC Rev. 4 classification, consistent with industry classification in the plant-level data. The tables provide data regarding the total monetary value of output produced by suppliers used in the final product's production by the manufacturers. Additionally, they obtain information on inputs from both domestic and international suppliers for each manufacturing sector. Specifically, the domestic matrix outlines domestic intermediate input flows—the value of domestically sourced inputs utilized in domestic production. Similarly, the import matrix portrays the value of imported inputs from foreign suppliers utilized by domestic manufacturers. Thus, I know the percentages of inputs purchased domestically and those imported from abroad. This information enables the distinction between domestic and foreign input cost shares, $\text{Domestic}_{\text{jst}}$ and $\text{Foreign}_{\text{jst}}$, respectively. These variables represent the proportion of input costs between the manufacturer $j$ and supplier $s$ relative to the total production cost (Total $\text{cost}_{jt}$) incurred by the manufacturer in producing its final product:

$$\text{Domestic}_{\text{jst}} = \frac{\text{Domestic input}_{\text{jst}}}{\text{Total cost}_{\text{jt}}}, \text{ and } \text{Foreign}_{\text{jst}} = \frac{\text{Foreign input}_{\text{jst}}}{\text{Total cost}_{\text{jt}}} \tag{2.7}$$

where Total $\text{cost}_j$ encompasses both domestic and foreign inputs. It should be noted that values in the input-output tables are expressed in dollars, whereas the AFiD data are provided in euros. However, by dividing the costs of both domestic and foreign inputs by the total production cost, I'm essentially normalizing the data to solely focus on the domestic and foreign input cost shares, regardless of the currency used for data collection. Finally, I construct an input-output tables for Germany by computing the averages for the years 2009-2018 across industries so that the input cost shares are given by $\text{Domestic}_{\text{js}}$ and $\text{Foreign}_{\text{js}}$.

## 3.3 R&D intensity

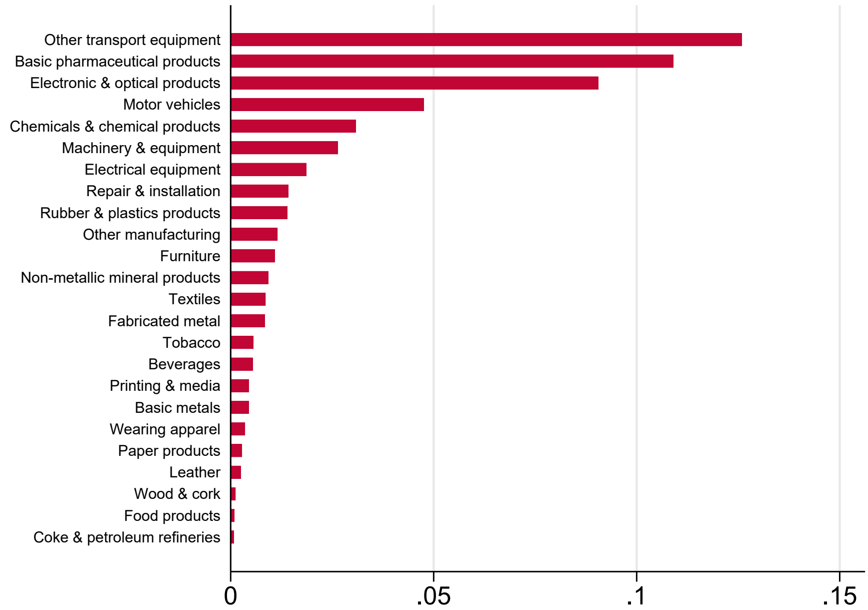
The theoretical model above explains the impact of the investment productivities of the manufacturing and supplying industries, $\lambda_j$ and $\lambda_s$, respectively, on the probability of vertical integration. Measuring these variable is difficult because these productivity are not observable in the data. Following the literature, I focus on industry-level technology intensity proxied by R&D intensity of the manufacturing and supplying indutries. I use the ANBERD database which contains annual industry-level R&D expenditure data at the two/three-digit ISIC Rev 4 industry classification. I calculate the average R&D intensity across the years 2003 to 2008. This measure is defined as the ratio of R&D investment to total sales, where sales are taken from the plant-level in the AFiD data and aggregated to the industry level. Additionally, to minimize the impact of outliers, I disregard the lowest 1% of firms with very low R&D intensity and the highest 1% with exceptionally high R&D intensity.

Figure 3 depicts the average R&D intensity of industries in Germany, revealing significant variation across different sectors. The *Other Transport Equipment* and *Basic Pharmaceutical Products* industries exhibit the highest R&D intensity, exceeding 10%. This underscores the critical importance of innovation and technological advancement in these sectors. Conversely, the *Food Products* and *Coke & Petroleum Refineries* industries show the lowest R&D intensity, with less than 1% of sales revenue devoted to R&D.

## 3.4 Vertical integration measure

To study the trade-off between vertical integration and outsourcing, I construct a measure of vertical integration utilizing the AFiD plant-level data. My focus lies on examining the situation from the manufacturer's standpoint, assessing the likelihood of the manufacturer vertically integrating with its supplier. I match the plant industry codes to the two-digit industry codes in the input-output tables and aggregate observations when more than one plant within the same firm belongs to the same industry in the table. This process ensures a unique plant per firm per year in the dataset. I set $V_{ijst} = 1$ if $j = s$, and the firm operates more than one plant in the four-digit industries within the two-digit industry in the input-output tables each year. Since the main analysis is cross-sectional, I aggregate the panel data by computing averages for the years 2009-2018. Each observation in the final dataset corresponds to a specific firm-

**Figure 3:** Average R&D intensity (2003-2008)




*Notes*: Figure 3 shows the average R&D intensity (as a decimal ratio) across industries in Germany between 2003 and 2008.
*Source*: Based on the author's calculations using the ANBERD and AFiD databases.

industry pair, where the vertical integration measure ($V_{ijs}$) is a dummy for whether the firm is vertically integrated in an industry pair at any time over the period 2009-2018.

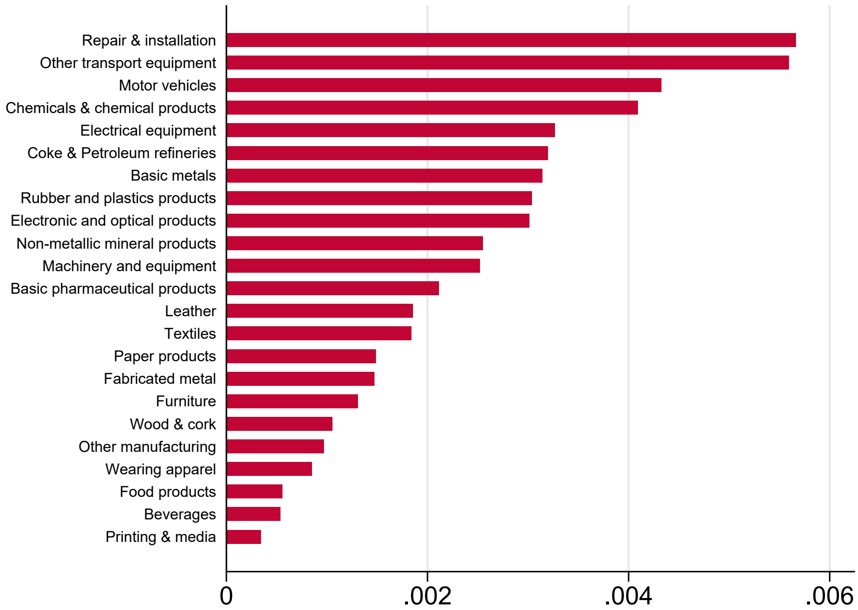
The average vertical integration is 0.0022. The reason for the small mean is that the majority of the observations consist of single-plant firms and there exist 24 supplying industries, resulting in a vertical integration measure of zero for the majority of cases.[11] Figure 4 illustrates that the degree of vertical integration varies significantly across industries. The top three most vertically integrated industries are *Repair and installation of machinery and equipment*, *Other transport equipment*, and *Motor vehicles*. Conversely, the least vertically integrated industries include *Food products*, *Beverages*, and *Printing and media*. These variations show that industry-specific characteristics, such as the importance of inputs and technology intensity, shape the extent of vertical integration.

## 4 Empirical framework

The theoretical predictions outlined in this paper offer insights into the relationship between the technology intensities of the manufacturer of a final product and its input supplier. It explores how these

[11]Table A1 in the Appendix presents descriptive statistics for the variables used in the analysis.

**Figure 4:** Average vertical integration




*Notes*: Figure 4 shows the average vertical integration across industries in Germany.
*Source*: Based on the author's calculations using the AFiD database.

intensities and the input cost shares influence decisions regarding ownership structure. In this section, I assess whether the theoretical predictions are supported by the data.

In this analysis, I adopt a cross-sectional approach and aggregate the data by taking the average over the years 2009-2018 to examine the relationship described by Equation 2.8. This decision is motivated by the observation that R&D intensity exhibit minimal variation within industries over the specified time period. R&D intensity represents factors that are relatively stable across years and are not subject to significant temporal fluctuations within my study context. However, as part of robustness checks, I also conduct a regression analysis using the full panel dataset, exploiting the time variation over the period from 2009 to 2018. I rely on the Linear Probability Model (LPM) to estimate the effect of input cost shares (domestic and foreign) and R&D intensity on the probability of vertical integration:

$$\begin{aligned} Pr(V_{ijs}=1) = \beta_0 + \beta_1\, Domestic_{js} + \beta_2\, Foreign_{js} + R\&D_j \left(\beta_3 + \beta_5\, Domestic_{js} + \beta_7\, Foreign_{js}\right) \\ + R\&D_s \left(\beta_4 + \beta_6\, Domestic_{js} + \beta_8\, Foreign_{js}\right) + \beta_9 X_{ij} + \beta_{10} X_j + \beta_{11} X_s + \gamma_i + \epsilon_{ijs} \end{aligned} \tag{2.8}$$

The coefficients $\beta_1$ and $\beta_2$ measure the impact of the domestic and foreign input cost shares, respectively. I expect a positive value of $\beta_1$ indicating an increase in vertical integration probability with a higher domestic input cost share, and a negative $\beta_2$ indicating a negative correlation between vertical integration probability and foreign input cost share. The coefficients $\beta_3$ and $\beta_4$ in my regression model capture the impact of R&D intensity within the manufacturing industry ($R\&D_j$) and the supplying industry ($R\&D_s$), respectively, on the likelihood of vertical integration. I expect a positive effect for the manufacturing industry and a negative effect for the supplying industry.

However, the key aspect is centered on the coefficients of the interaction terms as depicted in Equation 2.8. These coefficients tell how the combined effect of input cost share and industry-specific R&D intensity contribute to vertical integration. Specifically, $\beta_5$ and $\beta_7$ represent the effects of domestic and foreign input cost shares in producing industries with high R&D intensity. It is expected that their effects are opposite: while a high R&D-intensive manufacturer is expected to vertically integrate with a domestic supplier when the domestic input cost share is relatively high (yielding a positive $\beta_5$), it is expected to seek input outsourcing from domestic suppliers when the foreign input cost share is relatively high (resulting in a negative $\beta_7$). Regarding the interactions with the supplying industry's R&D intensity and input cost shares, the expected signs and effects are opposite: I expect $\beta_6$ to be negative, while $\beta_8$ is expected to be positive. Consequently, when the R&D intensity of the manufacturing industry surpasses that of the supplying industry, *and* when domestic input cost shares outweigh foreign input cost shares, there is a higher likelihood of vertical integration.

As control variables, I include the natural logarithm of firm age and size at both the firm-industry level and the industry level. Specifically, $X_{ij}$ comprises firm size and age at the firm-(manufacturing)-industry level, as well as their averages at the manufacturing and supplying industry levels, $X_j$ and $X_s$. Finally, $\epsilon_{ijs}$ represents the error term.

Firm characteristics may influence decisions regarding vertical integration. To examine the relationship between R&D intensity and vertical integration within specific firms, I include firm fixed effects ($\gamma_i$) later in the analysis, where I examine the within-firm variation. Doing so implies that I only keep multi-plant firms in the sample, which could result in a selection bias issue. Following Acemoglu et al. (2010), I employ a standard Heckman selection model to assess the probability of a firm being a multi-plant entity. By first estimating a selection equation to understand the factors influencing the firm's decision to operate multiple plants, and then correcting for bias in the outcome equation, the Heckman model ensures accurate estimates of the factors driving this firm behavior. Finally, I adjust standard errors for clustering at the industry-pair (manufacturer-supplier) level. This adjustment is made to address potential similarities or correlations among firms within the same industry-pair

The demand for intermediate inputs encompasses both inputs procured domestically and from overseas. The key identification assumption is that the input-output relationship specified in the input-output tables is exogenous to the organizational form chosen by firms. I assume that the variation in these shares is entirely determined by the technological importance of an input. This means that German firms have no influence on the industry-level production technologies. Thus, the decision to vertically integrate with a supplier is driven by the input-output linkages, which are the source of heterogeneity across inputs. However, there are considerations of transfer pricing within firms that may also affect the input cost shares in the input-output tables (Alfaro et al., 2016). The necessary granular data needed to adequately tackle this issue is not available for this study.

# 5 Results

In this section, I outline the findings from the baseline analysis along with the robustness checks.

## 5.1 Domestic supply and R&D intensity

I start by replicating the main analysis of Acemoglu et al. (2010) using data on German manufacturing plants. Here I do not distinguish between domestic and foreign input cost shares. This aims to establish a comparative baseline and aligns my work with existing literature. In Equation 2.9 below, the input cost share, $Cost\, share_{js}$, is defined similarly to Acemoglu et al. (2010)'s approach, where only domestic suppliers are considered. More precisely, it is measured as the proportion of input transactions between the manufacturer and the domestic supplier, relative to the total cost incurred from obtaining these inputs (domestically) .

$$\begin{aligned} Pr(V_{ijs} = 1) = \beta_0 + \beta_1\, Cost\, share_{js} + R\&D_j(\beta_2 + \beta_4\, Cost\, share_{js}) + R\&D_s(\beta_3 + \beta_5\, Cost\, share_{js}) \\ + \beta_6 X_{ij} + \beta_7 X_j + \beta_8 X_s + \epsilon_{ijs} \end{aligned} \tag{2.9}$$

where,

$$\text{Cost share}_{\text{js}} = \frac{\text{Domestic input}_{\text{js}}}{\text{Total domestic cost}_{\text{j}}}$$

Table 2 presents results as specified in Equation 2.9. Column 1 shows the main effect of the input cost share, the manufacturer's R&D intensity, and the supplier's R&D intensity. In Column 2, interaction terms are added, and Column 3 incorporates control variables into the regression. Column 1 demonstrates that a higher input cost share significantly increases the likelihood of vertical integration, with a statistically significant coefficient of 0.035. This effect remains consistent across the other specifications in Columns 2 and 3. The main effects of the manufacturer's R&D intensity and the supplier's R&D intensity do not exhibit the opposite effects predicted by the theoretical model (except in Column 3 where the sign turns negative). But they are statistically insignificant when the interaction terms and

**Table 2:** Domestic supply and R&D intensity

| Dep. Var: $V_{ijs}$ | (1) | (2) | (3) |
|---|---|---|---|
| Cost share$_{js}$ | 0.035*** | 0.026*** | 0.026*** |
| | (0.006) | (0.007) | (0.007) |
| R&D$_{j}$ | 0.038*** | 0.014 | -0.011 |
| | (0.011) | (0.011) | (0.013) |
| R&D$_{s}$ | 0.011** | 0.008 | 0.003 |
| | (0.005) | (0.005) | (0.008) |
| Cost share$_{js}$ × R&D$_{j}$ | | 0.557* | 0.582** |
| | | (0.298) | (0.285) |
| Cost share$_{js}$ × R&D$_{s}$ | | 0.029 | -0.011 |
| | | (0.344) | (0.329) |
| Observations | 1,355,448 | 1,355,448 | 1,350,648 |
| R-squared | 0.004 | 0.004 | 0.007 |
| Control variables | No | No | Yes |

*Notes:* The dependent variable, $V_{ijs}$, is a binary variable that equals 1 if a firm operates plants in both manufacturing and supplying industries. Cost share$_{js}$ is the input share of cost and is calculated as the ratio of input transactions from the domestic supplier to the manufacturer relative to the total cost of domestic inputs, obtained from the input-output tables for Germany. R&D$_{j}$ and R&D$_{s}$ denote the R&D intensities of the manufacturing and supplying industries, respectively, calculated from the AFiD plant-level data and ANBERD database as the R&D expenditure to sales. Standard errors are clustered at the industry-pair level and reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
*Source:* Author's calculations based on AFiD, input-output tables, and ANBERD datasets.

control variables are included in Columns 2 and 3. However, the primary focus of the analysis is on the combined effect. The interaction between the manufacturer's R&D intensity and the input cost share (Cost share$_{js}$ × R&D$_{j}$) leads to a higher probability of vertical integration, and this effect remains stable after adding control variables in Column 3. However, no statistically significant effect is found for the interaction between the input cost share and the supplier's R&D intensity (Cost share$_{js}$ × R&D$_{j}$).

It is important to note that the results in Table 2 do not fully align with those of Acemoglu et al. (2010). While I find similar results regarding the input cost share and its interaction with the manufacturer's R&D intensity, the interaction effect for the supplier, though displaying the expected opposite sign, is not statistically significant. However, this regression excludes the foreign input cost share, focusing solely on the relationship between vertical integration, domestic input cost share, and R&D intensity. If firms also rely on foreign suppliers, but this factor is not accounted for, the domestic input cost share variable may capture some of this effect, potentially overstating its positive impact on vertical integration. Consequently, the estimated coefficient on domestic input cost share is likely biased upward.

Introducing foreign input cost share helps mitigate this bias by incorporating an additional channel that influences vertical integration decisions, namely, globalization. As I will show in the next subsection, including the foreign input cost share significantly alters the regression results.

## 5.2 Domestic vs. foreign supply and R&D intensity

How does considering foreign supply change the picture of the vertical integration decision? I answer this question by examining Equation 2.8. The results are reported in Table 3. In column 1, I exclude the foreign input cost share and control variables. The coefficient of the domestic input cost share is positive and statistically significant at the 1% level, with a value of 0.013. This suggests that an increase in the share of domestic inputs in the production process is associated with a higher probability of vertical integration. This result aligns with the hypothesis that firms facing higher domestic input costs may find it more beneficial to internalize their supply chain through vertical integration. While the coefficients on both the manufacturer's R&D investment ($R\&D_j$) and the supplier's R&D investment ($R\&D_s$) are positive, neither of these variables is statistically significant. This suggests that for very low domestic input cost shares, R&D investments by manufacturers or suppliers do not have a significant impact on the likelihood of vertical integration. The coefficients remain statistically insignificant in all specifications.

The main focus of this analysis is on the interplay between the importance of input costs and the R&D intensity of each party involved. Specifically, the interaction effect between the domestic input cost share and the manufacturer's R&D investment ($Domestic_{js} \times R\&D_j$) is positive and statistically significant at the 5% level. This interaction indicates that the probability of vertical integration increases when both the domestic input cost share is high and the manufacturer is heavily investing in R&D. Although the sign of the interaction between the domestic input cost share and the supplier's R&D investment ($Domestic_{js} \times R\&D_s$) is negative, as predicted by theory, it is statistically insignificant.

In column 2, I introduce the foreign input cost share into the regression. The coefficient for the domestic input cost share remains positive and statistically significant, while the coefficient for the foreign input cost share is negative but lacks statistical significance. The interactions of domestic input cost share with the manufacturer's R&D intensity remain unchanged in terms of signs and significance. In this specification, the negative interaction between the supplier's R&D intensity and the domestic input cost share becomes statistically significant. Interestingly, as predicted by the theoretical model, the interaction between the foreign input cost share and the manufacturer's R&D intensity shows the opposite sign compared to the domestic input cost share estimate, and it is statistically significant. I also find an opposite sign for the interaction between the supplier's R&D intensity and the foreign input cost share, which shows a positive sign and is statistically significant. These results remain stable even after controlling for firm age and size in column 3.

**Table 3:** Domestic vs. foreign supply and R&D intensity

| Dep. Var: $V_{ijs}$ | (1) | (2) | (3) |
|---|---|---|---|
| $\text{Domestic}_{js}$ | 0.013*** | 0.017*** | 0.017*** |
| | (0.003) | (0.004) | (0.004) |
| $\text{Foreign}_{js}$ | | -0.009 | -0.009 |
| | | (0.008) | (0.007) |
| $\text{R\&D}_{j}$ | 0.003 | 0.001 | -0.002 |
| | (0.004) | (0.003) | (0.004) |
| $\text{R\&D}_{s}$ | 0.002 | 0.001 | -0.001 |
| | (0.002) | (0.001) | (0.002) |
| $\text{Domestic}_{js} \times \text{R\&D}_{j}$ | 0.397** | 0.811*** | 0.748*** |
| | (0.196) | (0.260) | (0.238) |
| $\text{Domestic}_{js} \times \text{R\&D}_{s}$ | -0.232 | -0.706** | -0.680** |
| | (0.193) | (0.304) | (0.291) |
| $\text{Foreign}_{js} \times \text{R\&D}_{j}$ | | -0.535*** | -0.496** |
| | | (0.206) | (0.197) |
| $\text{Foreign}_{js} \times \text{R\&D}_{s}$ | | 0.631*** | 0.601*** |
| | | (0.237) | (0.230) |
| Observations | 1,355,448 | 1,355,448 | 1,350,648 |
| R-squared | 0.004 | 0.004 | 0.007 |
| Control variables | No | No | Yes |

Notes: The dependent variable, $V_{ijs}$, is a binary variable that equals 1 if a firm operates plants in both producing and supplying industries. $\text{Domestic}_{js}$ and $\text{Foreign}_{js}$ are the domestic and foreign input cost shares, respectively, obtained from the input-output tables for Germany. $\text{R\&D}_{j}$ and $\text{R\&D}_{s}$ denote the R&D intensities of the manufacturing and supplying industries, respectively, calculated from the AFiD plant-level data and ANBERD database as the R&D expenditure to sales. Standard errors are clustered at the industry-pair level and reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's calculations based on AFiD, input-output tables, and ANBERD datasets.

**Within-firm variation.** Firm fixed effects play a crucial role in controlling for unobserved heterogeneity across firms, allowing for more accurate and reliable estimates by isolating the effects of the variables of interest from firm-specific factors. This justifies the focus on within-firm variation. However, by the definition used in this paper, single-plant firms are not involved in vertical integration, making them irrelevant for the within-firm analysis. This selective focus results in a significant drop in the number of observations, with approximately 85% of the observations lost (in Acemoglu et al. (2010) there is a drop of around 70% of the observations when adding firm fixed effects).

**Table 4:** Firm fixed effects

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Dep. Var: $V_{ijs}$ | | | | = 1 if multi-plant | |
| $\text{Domestic}_{js}$ | 0.292*** | 0.287*** | 0.302*** | -0.0189 | 0.3117*** |
| | (0.085) | (0.082) | (0.088) | (0.379) | (0.089) |
| $\text{Foreign}_{js}$ | -0.010 | 0.007 | -0.016 | 0.0297 | -0.0356 |
| | (0.139) | (0.139) | (0.146) | (0.455) | (0.145) |
| $\text{R\&D}_j$ | 0.038 | 0.040 | -0.090 | -2.2753*** | -0.1505 |
| | (0.060) | (0.074) | (0.238) | (0.401) | (0.198) |
| $\text{R\&D}_s$ | 0.049* | -0.001 | -0.045 | 0.7252* | 0.0555* |
| | (0.028) | (0.045) | (0.061) | (0.371) | (0.030) |
| $\text{Domestic}_{js} \times \text{R\&D}_j$ | 10.167** | 9.777** | 9.130** | 3.6783 | 9.7433** |
| | (4.118) | (3.918) | (3.918) | (13.346) | (4.037) |
| $\text{Domestic}_{js} \times \text{R\&D}_s$ | -7.801 | -7.727 | -7.559 | -4.9903 | -8.0488 |
| | (5.729) | (5.651) | (5.532) | (15.440) | (5.563) |
| $\text{Foreign}_{js} \times \text{R\&D}_j$ | -6.969** | -6.778** | -5.602* | -14.8796 | -5.9392* |
| | (3.321) | (3.325) | (3.297) | (13.599) | (3.234) |
| $\text{Foreign}_{js} \times \text{R\&D}_s$ | 7.616* | 7.466* | 6.731* | 16.1023 | 7.1965* |
| | (4.044) | (4.081) | (4.017) | (15.632) | (3.934) |
| Observations | 200,424 | 198,816 | 198,816 | 1,350,648 | 198,816 |
| R-squared | 0.024 | 0.027 | 0.069 | | 0.067 |
| Control variables | No | Yes | Yes | Yes | No |
| Firm FE | No | No | Yes | No | Yes |

Notes: The dependent variable, $V_{ijs}$, is a binary variable that equals 1 if a firm operates plants in both producing and supplying industries. $\text{Domestic}_{js}$ and $\text{Foreign}_{js}$ are the domestic and foreign input cost shares, respectively, obtained from the input-output tables for Germany. $\text{R\&D}_j$ and $\text{R\&D}_s$ denote the R&D intensities of the manufacturing and supplying industries, respectively, calculated from the AFiD plant-level data and ANBERD database as the R&D expenditure to sales. Standard errors are clustered at the industry-pair level and reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's calculations based on AFiD, input-output tables, and ANBERD datasets.

In columns 1 through 3 of Table 4, the analysis focuses exclusively on multi-plant firms, replicating the analysis from Column 2 of Table 3, but without firm fixed effects. In column 2, I add control variables, and in column 3, I include firm fixed effects. Interestingly, in all columns, the interaction terms display the signs predicted by the theoretical model and are statistically significant (except for the interaction between the domestic input cost share and the R&D intensity of the supplier). Specifically, manufacturers with a high domestic input cost share are more likely to integrate with their suppliers, particularly when they operate in high-R&D intensity industries. Conversely, the likelihood of integration decreases when the foreign input cost share is higher than the domestic cost share, especially in R&D-intensive industries. This pattern holds true for suppliers as well. The opposite signs of the domestic and foreign input cost shares further validate the theoretical predictions.

Nevertheless, solely analyzing multi-plant firms in the assessment of within-firm variation could introduce a selection bias problem. To examine whether these estimates are affected by selection bias, I employ the Heckman two-stage procedure. In Column 4, I conduct a regression where the dependent variable is a dummy indicating whether the plant is part of a multi-plant firm or a single-plant firm (coded as 0) using a probit model, and all covariates are included. Of the key variables of interest, only the R&D intensity of the manufacturer and the supplier are statistically significant, suggesting potential selection bias in these variable. The other variables do not exhibit selection bias issues, especially the interaction terms. It's worth noting that Acemoglu et al. (2010) find that only the manufacturing industry is prone to bias. In Column 5, I introduce firm fixed effects and present the second stage of the Heckman model, focusing on whether a firm is vertically integrated, conditioned on being a multi-plant firm. Following Acemoglu et al. (2010), the second stage does not include control variables. Interestingly, the interaction terms maintain the same sign and statistical significance as those in column 3.

Overall, the findings in this paper align with existing literature on how vertical integration is influenced by the importance of inputs. This is consistent with studies by Acemoglu et al. (2010), Berlingieri et al. (2021), and Alfaro et al. (2024). The results also support previous research on the effect of contracting parties' investment intensity on vertical integration, as discussed in works by Acemoglu et al. (2010) and Egger et al. (2023). Furthermore, the inclusion of foreign suppliers introduces a new factor that could impact a firm's decision to integrate vertically—a dimension not fully explored in prior studies.

## 5.3 Additional analyses

**Excluding top and bottom quartiles.** Large or small firms of the size distribution might have disproportionate impacts on the results due to their unique characteristics. In Table 5 I excluded the top quartile and the bottom quartile of firm size. Interestingly, the results remain consistent with the previous findings from the main analysis. The interaction terms exhibit the expected signs and are statistically significant, except for the interaction between R&D intensity of the supplying industry and domestic and foreign cost shares. When excluding the bottom quartiles, I find similar fingerings.

**Total input cost share.** The main results above suggest that $\text{Domestic}_{js}$ and $\text{Foreign}_{js}$ have opposite effects on vertical integration, especially when these variables interact with R&D intensity. However,

**Table 5:** Excluding top and bottom quartiles

| | Excluding top quartiles | | | Excluding bottom quartiles | | |
|---|---|---|---|---|---|---|
| Dep. Var: $V_{ijs}$ | (1) | (2) | (3) | (4) | (5) | (6) |
| $\text{Domestic}_{js}$ | 0.033*** | 0.026 | 0.029* | 0.045*** | 0.040* | 0.043** |
| | (0.011) | (0.017) | (0.016) | (0.014) | (0.022) | (0.022) |
| $\text{Foreign}_{js}$ | | 0.019 | 0.015 | | 0.015 | 0.012 |
| | | (0.030) | (0.030) | | (0.037) | (0.037) |
| $\text{R\&D}_j$ | 0.015 | 0.011 | -0.011 | 0.016 | 0.010 | -0.010 |
| | (0.011) | (0.010) | (0.012) | (0.014) | (0.013) | (0.016) |
| $\text{R\&D}_s$ | 0.008 | 0.007 | 0.001 | 0.011* | 0.009 | 0.000 |
| | (0.005) | (0.005) | (0.007) | (0.007) | (0.006) | (0.010) |
| $\text{Domestic}_{js} \times \text{R\&D}_j$ | 1.197** | 2.256*** | 2.094*** | 1.360** | 2.743*** | 2.539*** |
| | (0.525) | (0.737) | (0.678) | (0.653) | (0.924) | (0.855) |
| $\text{Domestic}_{js} \times \text{R\&D}_s$ | -0.031 | -1.088 | -1.102 | -0.103 | -1.520 | -1.510 |
| | (0.563) | (0.963) | (0.937) | (0.713) | (1.221) | (1.197) |
| $\text{Foreign}_{js} \times \text{R\&D}_j$ | | -1.415** | -1.287** | | -1.887** | -1.730** |
| | | (0.650) | (0.626) | | (0.848) | (0.822) |
| $\text{Foreign}_{js} \times \text{R\&D}_s$ | | 1.128 | 1.129 | | 1.638 | 1.617 |
| | | (0.779) | (0.768) | | (1.007) | (1.000) |
| Observations | 1,018,368 | 1,018,368 | 1,014,072 | 1,017,984 | 1,017,984 | 1,014,624 |
| R-squared | 0.001 | 0.001 | 0.001 | 0.005 | 0.005 | 0.007 |
| Control variables | No | No | Yes | No | No | Yes |

Notes: The dependent variable, $V_{ijs}$, is a binary variable that equals 1 if a firm operates plants in both producing and supplying industries. $\text{Domestic}_{js}$ and $\text{Foreign}_{js}$ are the domestic and foreign input cost shares, respectively, obtained from the input-output tables for Germany. $\text{R\&D}_j$ and $\text{R\&D}_s$ denote the R&D intensities of the manufacturing and supplying industries, respectively, calculated from the AFiD plant-level data and ANBERD database as the R&D expenditure to sales. Standard errors are clustered at the industry-pair level and reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
*Source:* Author's calculations based on AFiD, input-output tables, and ANBERD datasets.

it is worth noting that there is a strong positive correlation between the two variables, with a correlation coefficient of around 0.7339. This means that industries that source a large share of their inputs domestically also tend to rely significantly on foreign inputs. This suggests that firms do not exclusively rely on either domestic or foreign suppliers but instead balance their sourcing strategies between both. Therefore, in this part of the analysis, I focus exclusively on the total domestic input cost share. According to the input-output tables I use, on average 66% of the inputs are source domestically, while the rest are obtained from abroad. The total domestic and foreign input cost shares are complementary, as their sum always equals 100%: this means that an increase in one corresponds to a decrease in the other. Therefore, I can exclude the total foreign input cost share variable from the regression and its impact can be interpreted based on its inverse relationship with domestic input costs.

To do this, I calculate $\text{Domestic}_\text{j}$ as the total of the input cost shares from all individual suppliers, divided by the manufacturer's total input expenditure. This variable represents the proportion of the manufacturer's input costs that come from domestic sources across all suppliers. An increase in $\text{Domestic}_\text{j}$ is expected to result in higher levels of vertical integration. Therefore, a positive coefficient for $\text{Domestic}_\text{j}$ indicates that greater domestic input costs are linked to more vertical integration. This implies that $\text{Foreign}_\text{j}$ has a negative relationship with vertical integration, as an increase in the domestic share typically corresponds to a decrease in reliance on foreign inputs. I run the following regression and present the results in Table 6.

$$\begin{aligned} Pr(V_{ijs}=1) = \beta_0 + \beta_1\, Domestic_j + \beta_2\, R\&D_s + \beta_3\, R\&D_j \times Domestic_j + \beta_4\, R\&D_s \times Domestic_j \\ + \beta_5\, X_{ij} + \beta_6\, X_j + \beta_7\, X_s + \gamma_i + \epsilon_{ijs}. \end{aligned} \tag{2.10}$$

The coefficient for the total domestic input cost share is negative across all specifications but becomes statistically insignificant after adding control variables in column 3. This finding does not align with the predictions of the model. It suggests that higher domestic input usage may lead to lower vertical integration. However, when I interact the domestic input cost share with R&D intensity, I obtain the expected signs predicted by the model. Nonetheless, none of the variables, particularly in the preferred specification in column 3, are statistically significant.

Interestingly, I find more robust results, consistent with earlier findings regarding the impact of domestic versus foreign input cost shares and R&D intensity on vertical integration, when I use the full panel time duration rather than focusing on cross-sectional data, as shown in Table 7. Here, too, the total domestic input cost share is negatively related to vertical integration, but its interactions with R&D intensity exhibit the expected signs. Specifically, in highly R&D-intensive manufacturing sectors, a higher domestic input cost share leads to greater integration. This result is statistically significant in column 2. In column 3, while the coefficient becomes insignificant after controlling for firm size and age, it retains the positive sign. This suggests that while an increase in the importance of domestic inputs tends to promote vertical integration, an increase in the reliance on foreign inputs is associated with less integration. Additionally, both specifications in columns 2 and 3 show statistically significant and negative coefficients for the interaction between total domestic input cost share and the supplier's R&D intensity.

**Table 6:** Total input cost share

| Dep. Var: $V_{\mathrm{ijs}}$ | (1) | (2) | (3) |
|---|---|---|---|
| $\mathrm{Domestic}_{\mathrm{j}}$ | -0.008** | -0.012*** | -0.005 |
| | (0.004) | (0.004) | (0.004) |
| $\mathrm{R\&D}_{\mathrm{j}}$ | 0.020 | -0.182** | -0.061 |
| | (0.015) | (0.072) | (0.070) |
| $\mathrm{R\&D}_{\mathrm{s}}$ | 0.004 | 0.067 | 0.055 |
| | (0.006) | (0.045) | (0.047) |
| $\mathrm{Domestic}_{\mathrm{j}} \times \mathrm{R\&D}_{\mathrm{j}}$ | | 0.363*** | 0.125 |
| | | (0.128) | (0.129) |
| $\mathrm{Domestic}_{\mathrm{j}} \times \mathrm{R\&D}_{\mathrm{s}}$ | | -0.093 | -0.086 |
| | | (0.064) | (0.064) |
| Observations | 1,356,024 | 1,356,024 | 1,351,224 |
| R-squared | 0.004 | 0.004 | 0.007 |
| Control variables | No | No | Yes |

Notes: The dependent variable, $V_{\mathrm{ijs}}$, is a binary variable that equals 1 if a firm operates plants in both producing and supplying industries. $\mathrm{Domestic}_{\mathrm{j}}$ represents the total input cost shares and obtained from the input-output tables for Germany. $\mathrm{R\&D}_{\mathrm{j}}$ and $\mathrm{R\&D}_{\mathrm{s}}$ denote the R&D intensities of the manufacturing and supplying industries, respectively, calculated from the AFiD plant-level data and ANBERD database as the R&D expenditure to sales. Standard errors are clustered at the industry-pair level and reported in parentheses. $^{*}$, $^{**}$, and $^{***}$ indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's calculations based on AFiD, input-output tables, and ANBERD datasets.

## 5.4 Robustness checks

**Panel analysis.** I now use panel data spanning from 2009 to 2018 to capture potential time trends and account for any temporal variation that may influence the organizational structure. In this part of the analysis, R&D is computed using data from the full time period rather than relying on pre-sample data. The results presented in Table 8, derived from the panel data analysis, are consistent with those obtained from a cross-sectional approach, suggesting that the inclusion of time-series variation does not significantly alter the key relationships observed in the data. It should be noted, however, that while the interaction between the supplier's R&D intensity and the domestic foreign input cost share does not exhibit statistical significance, the signs of the estimated coefficients are consistent with the theoretical predictions. This suggests that, although the relationship may not be statistically robust, the directionality of the effects remains in line with theory.

**Table 7:** Total input cost share - panel analysis

| Dep. Var: $V_{\text{ijst}}$ | (1) | (2) | (3) |
|---|---|---|---|
| $\text{Domestic}_{\text{jt}}$ | -0.007** | -0.008** | -0.005 |
| | (0.003) | (0.003) | (0.003) |
| $\text{R\&D}_{\text{jt}}$ | 0.011 | -0.109** | -0.063 |
| | (0.010) | (0.045) | (0.043) |
| $\text{R\&D}_{\text{st}}$ | 0.009 | 0.082** | 0.067* |
| | (0.006) | (0.038) | (0.037) |
| $\text{Domestic}_{\text{jt}} \times \text{R\&D}_{\text{jt}}$ | | 0.221*** | 0.097 |
| | | (0.082) | (0.076) |
| $\text{Domestic}_{\text{jt}} \times \text{R\&D}_{\text{st}}$ | | -0.109** | -0.097* |
| | | (0.052) | (0.051) |
| Observations | 9,170,112 | 9,170,112 | 9,122,808 |
| R-squared | 0.002 | 0.003 | 0.005 |
| Control variables | No | No | Yes |
| Time dummies | Yes | Yes | Yes |

Notes: The dependent variable, $V_{\text{ijst}}$, is a binary variable that equals 1 if a firm operates plants in both producing and supplying industries at time t. $\text{Domestic}_{\text{jt}}$ represents the total input cost shares at time t and obtained from the input-output tables for Germany. $\text{R\&D}_{\text{jt}}$ and $\text{R\&D}_{\text{st}}$ denote the R&D intensities of the manufacturing and supplying industries at time t, respectively, calculated from the AFiD plant-level data and ANBERD database as the R&D expenditure to sales. Standard errors are clustered at the industry-pair level and reported in parentheses. $^{*}$, $^{**}$, and $^{***}$ indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's calculations based on AFiD, input-output tables, and ANBERD datasets.

**Physical investment intensity.** The measurement of R&D intensity is closely linked to investment in new technologies within manufacturing and supplying industries. However, in the context of R&D spending, it is typical to find that a few industries allocate a disproportionately large share of funds to R&D, while many others receive relatively smaller allocations (a right-skewed distribution). Furthermore, there might exist a bidirectional relationship between R&D expenditures and vertical integration, as they are frequently observed in industries dominated by large firms, which are more inclined to invest heavily and tend to be vertically integrated (Acemoglu et al., 2010). Therefore, I consider physical investment intensity as an alternative measure of industry-level technology intensity. From the investment survey (IEB) in the AFiD data, I use plant-level information on tangible capital expenditure encompassing property, plant, and equipment from the pre-sample period, 2003-2008. By dividing investment by total sales, I derive investment intensity, which is aggregated from the plant level to the industry level,

**Table 8:** Pooled panel analysis

| Dep. Var: $V_{\text{ijst}}$ | (1) | (2) | (3) |
|---|---|---|---|
| $\text{Domestic}_{\text{jst}}$ | 0.023** | 0.018 | 0.020 |
| | (0.009) | (0.013) | (0.013) |
| $\text{Foreign}_{\text{jst}}$ | | 0.012 | 0.011 |
| | | (0.022) | (0.021) |
| $\text{R\&D}_{\text{jt}}$ | 0.010 | 0.006 | -0.019** |
| | (0.008) | (0.007) | (0.008) |
| $\text{R\&D}_{\text{st}}$ | 0.011** | 0.008* | 0.001 |
| | (0.005) | (0.004) | (0.005) |
| $\text{Domestic}_{\text{jst}} \times \text{R\&D}_{\text{jt}}$ | 0.822** | 1.665*** | 1.533*** |
| | (0.406) | (0.527) | (0.473) |
| $\text{Domestic}_{\text{jst}} \times \text{R\&D}_{\text{st}}$ | -0.060 | -0.837 | -0.850 |
| | (0.464) | (0.662) | (0.621) |
| $\text{Foreign}_{\text{jst}} \times \text{R\&D}_{\text{jt}}$ | | -1.168** | -1.041** |
| | | (0.461) | (0.418) |
| $\text{Foreign}_{\text{jst}} \times \text{R\&D}_{\text{st}}$ | | 0.956* | 0.905* |
| | | (0.537) | (0.499) |
| Observations | 9,165,288 | 9,165,288 | 9,117,984 |
| R-squared | 0.002 | 0.003 | 0.005 |
| Control variables | No | No | Yes |
| Time dummies | Yes | Yes | Yes |

Notes: The panel data spans the period 2009-2018. The dependent variable, $\text{V}_{\text{ijst}}$, is a binary variable that equals 1 if a firm operates plants in both producing and supplying industries at time $t$. $\text{Domestic}_{\text{jst}}$ and $\text{Foreign}_{\text{jst}}$ are the domestic and foreign input cost shares at time $t$, respectively, obtained from the input-output tables for Germany. $\text{R\&D}_{\text{jt}}$ and $\text{R\&D}_{\text{st}}$ denote the R&D intensities of the manufacturing and supplying industries at time $t$, respectively, calculated from the AFiD plant-level data and ANBERD database as the R&D expenditure to sales, using data from 2009 to 2018. Standard errors are clustered at the industry-pair level and reported in parentheses. $^*$, $^{**}$, and $^{***}$ indicate statistical significance at the 10%, 5%, and 1% levels, respectively. *Source:* Author's calculations based on AFiD, input-output tables, and ANBERD datasets.

averaged over years.

In Table 9, the coefficients for both domestic and foreign input cost shares are positive, but statistically insignificant, with the exception of the domestic input cost share in the first column. The investment intensity of the manufacturer remains negative and statistically insignificant across all specifications. In contrast, the investment intensity of the supplier aligns with the theoretical expectation, being negative and statistically significant in all models. The interaction between the manufacturer's investment inten-

**Table 9:** Physical investment intensity

| Dep. Var: $V_{ijs}$ | (1) | (2) | (3) |
|---|---|---|---|
| $\text{Domestic}_{js}$ | 0.068** | 0.016 | 0.021 |
| | (0.030) | (0.052) | (0.050) |
| $\text{Foreign}_{js}$ | | 0.143 | 0.132 |
| | | (0.094) | (0.093) |
| $\text{InvInt}_{j}$ | -0.020 | -0.010 | -0.021 |
| | (0.031) | (0.029) | (0.029) |
| $\text{InvInt}_{s}$ | -0.005*** | -0.005*** | -0.006*** |
| | (0.001) | (0.001) | (0.002) |
| $\text{Domestic}_{js} \times \text{InvInt}_{j}$ | -0.263 | 0.930 | 0.738 |
| | (0.992) | (1.546) | (1.495) |
| $\text{Domestic}_{js} \times \text{InvInt}_{s}$ | -0.228** | -0.259* | -0.206 |
| | (0.113) | (0.145) | (0.135) |
| $\text{Foreign}_{js} \times \text{InvInt}_{j}$ | | -3.488 | -3.211 |
| | | (2.689) | (2.639) |
| $\text{Foreign}_{js} \times \text{InvInt}_{s}$ | | 0.177 | 0.156 |
| | | (0.157) | (0.154) |
| Observations | 1,345,656 | 1,345,656 | 1,340,880 |
| R-squared | 0.003 | 0.004 | 0.007 |
| Control variables | No | No | Yes |

Notes: The dependent variable, $V_{ijs}$, is a binary variable that equals 1 if a firm operates plants in both producing and supplying industries. $\text{Domestic}_{js}$ and $\text{Foreign}_{js}$ are the domestic and foreign input cost shares, respectively, obtained from the input-output tables for Germany. $\text{InvInt}_{j}$ and $\text{InvInt}_{s}$ denote the physical investment intensities of the manufacturing and supplying industries, respectively, calculated from the AFiD plant-level data as the ratio of tangible capital expenditure to sales. Standard errors are clustered at the industry-pair level and reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's calculations based on AFiD, input-output tables, and ANBERD datasets.

sity and the domestic input cost share is positive, especially after adding control variables, though it remains statistically insignificant. For the supplier, the interaction term is negative and statistically significant, except when control variables are included, where the effect becomes insignificant. Although neither of the interaction terms between foreign input cost share and investment intensity for the manufacturer and the supplier is statistically significant, the signs of both coefficients are in the direction predicted by theory.

In Table 10, I run the regressions with firm fixed effects as in Table 4, focusing on multi-plant firms. In columns 1 through 3, only the domestic input cost share, the supplier's investment intensity, and the interactions with the domestic and foreign input cost shares (except in column 3 when fixed effects are added) are statistically significant and exhibit the expected signs. In column 4, both the manufacturer's and the supplier's investment intensity suffer from selection bias, but the interaction terms do not. In the second stage of Heckman selection procedure, only the domestic input cost share, the supplier's investment intensity, and their interaction with the domestic input cost share remain statistically significant.

**Table 10:** Investment intensity with firm fixed effects

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Dep. Var: $V_{ijs}$ | | | = 1 if multi-plant | | |
| $Domestic_{js}$ | 0.853*** | 0.842*** | 0.831*** | -1.6313 | 0.8370*** |
| | (0.277) | (0.275) | (0.275) | (1.720) | (0.280) |
| $Foreign_{js}$ | -0.269 | -0.268 | -0.262 | 3.6460 | -0.2590 |
| | (0.440) | (0.446) | (0.429) | (2.235) | (0.446) |
| $InvInt_j$ | -0.138 | -0.167 | 0.113 | 5.7340*** | 0.2723 |
| | (0.195) | (0.193) | (0.734) | (2.002) | (0.762) |
| $InvInt_s$ | -0.012*** | -0.014*** | -0.014*** | -0.1304*** | -0.0155*** |
| | (0.003) | (0.003) | (0.004) | (0.036) | (0.004) |
| $Domestic_{js} \times InvInt_j$ | -11.765 | -11.905* | -11.781 | 46.2848 | -11.6249 |
| | (7.253) | (7.146) | (7.266) | (47.911) | (7.397) |
| $Domestic_{js} \times InvInt_s$ | -3.391*** | -3.072*** | -2.680*** | -1.3720 | -2.8705*** |
| | (0.946) | (0.932) | (0.958) | (5.545) | (0.972) |
| $Foreign_{js} \times InvInt_j$ | 3.354 | 3.856 | 3.727 | -88.3103 | 3.5510 |
| | (11.576) | (11.641) | (11.417) | (61.859) | (11.837) |
| $Foreign_{js} \times InvInt_s$ | 6.111* | 5.763* | 5.215 | -17.2716 | 5.0832 |
| | (3.283) | (3.096) | (3.208) | (16.843) | (3.340) |
| Observations | 200,424 | 198,816 | 198,816 | 1,350,648 | 198,816 |
| R-squared | 0.025 | 0.029 | 0.070 | | 0.069 |
| Control variables | No | Yes | Yes | Yes | No |
| Firm FE | No | No | Yes | No | Yes |

Notes: The dependent variable, $V_{ijs}$, is a binary variable that equals 1 if a firm operates plants in both producing and supplying industries. $Domestic_{js}$ and $Foreign_{js}$ are the domestic and foreign input cost shares, respectively, obtained from the input-output tables for Germany. $InvInt_j$ and $InvInt_s$ denote the physical investment intensities of the manufacturing and supplying industries, respectively, calculated from the AFiD plant-level data as the ratio of tangible capital expenditure to sales. Standard errors are clustered at the industry-pair level and reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's calculations based on AFiD, input-output tables, and ANBERD datasets.

The difference in the results between using R&D intensity and investment intensity reflects the distinct roles these variables play in the empirical model. Moreover, physical investment intensity is not a perfect indicator of the overall technology intensity of a sector. Technology intensity is also influenced by intangible assets such as intellectual property and patents, which may not involve significant capital expenditure but can greatly enhance a sector's technological capability. Moreover, substantial investment in areas like infrastructure or marketing does not necessarily correlate with increased technology intensity. Although R&D intensity and investment intensity are positively correlated, the correlation is relatively weak—0.250. This suggests that each measure captures different aspects or dimensions of technology intensity.[12]

# 6 Conclusion

The relationships between producing and supplying industries have become increasingly sophisticated. This complexity arises because production processes often involve multiple stages and require relationship-specific inputs. In an open economy, these inputs are supplied by both domestic and foreign sources. Firms then decide whether to outsource these inputs or produce them in-house, and these decisions ultimately shape their organizational structure.

In this study, I build on the standard vertical integration framework developed by Acemoglu et al. (2010) to investigate the factors influencing vertical integration among manufacturing firms in Germany, with a specific emphasis on the role of foreign suppliers. The model introduced by Acemoglu et al. (2010) operates more closely within the property-rights framework and suggests that a firm's decision to integrate vertically depends on the investment intensities of both manufacturers and suppliers, as well as the share of total costs accounted for by supplier inputs. Additionally, their framework predicts that the effects of input cost shares on vertical integration are shaped by the interaction with technology intensity. However, their analysis primarily considers domestic input flows when calculating input cost shares and does not examine the impact of foreign suppliers. In this paper, I argue that the domestic trade-off between vertical integration and outsourcing is significantly influenced by the availability of foreign suppliers.

To empirically assess these dynamics, I utilize plant-level data from Germany's official firm data (AFiD), alongside the OECD's input-output tables for Germany and the Analytical Business Enterprise R&D (ANBERD) database. By merging these datasets, I construct a measure of vertical integration that captures whether a firm owns production facilities in both the manufacturing and supplying industries.

The results indicate that a higher share of domestic input costs increases the likelihood of vertical

[12]Note that Acemoglu et al. (2010) report a correlation of 0.251 between R&D intensity and investment intensity.

integration when the manufacturer has a higher R&D intensity than its supplier. Conversely, when the manufacturer's R&D intensity is lower than that of the supplier, a greater domestic input cost share reduces the probability of vertical integration.

A similar but opposite pattern emerges when considering foreign input costs. Specifically, an increase in the foreign input cost share decreases the likelihood of vertical integration when the manufacturer's R&D intensity exceeds that of the supplier, whereas a greater reliance on foreign inputs raises the probability of integration when the manufacturer's R&D intensity is lower. These findings remain consistent across multiple model specifications, underscoring the pivotal role of input cost structures and technology intensity in shaping firms' organizational choices.

This paper provides additional evidence on the issue of incomplete contracts and the relationship between contracting parties in the market. Incomplete contracts can lead to the hold-up problem, which may cause conflicts and production disruptions. Furthermore, it highlights the continued importance of understanding inefficiencies from underinvestment, a key issue emphasized by the property rights theory.

## Bibliography

**Acemoglu, Daron, Rachel Griffith, Philippe Aghion, and Fabrizio Zilibotti**, "Vertical Integration and Technology: Theory and Evidence," *Journal of the European Economic Association*, 2010, *8* (5), 989–1033. Publisher: Oxford University Press.

**Alfaro, Laura, Nick Bloom, Paola Conconi, Harald Fadinger, Patrick Legros, Andrew F Newman, Raffaella Sadun, and John Van Reenen**, "Come Together: Firm Boundaries and Delegation," *Journal of the European Economic Association*, February 2024, *22* (1), 34–72.

**__, Paola Conconi, Harald Fadinger, and Andrew F. Newman**, "Do Prices Determine Vertical Integration?," *The Review of Economic Studies*, July 2016, *83* (3), 855–888. Publisher: Oxford Academic.

**Antràs, Pol**, "Firms, Contracts, and Trade Structure*," *The Quarterly Journal of Economics*, November 2003, *118* (4), 1375–1418.

**__ and Elhanan Helpman**, "Global Sourcing," *Journal of Political Economy*, 2004, *112* (3), 552–580. Publisher: The University of Chicago Press.

**Berlingieri, Giuseppe, Frank Pisch, and Claudia Steinwender**, "Organizing Global Supply Chains: Input-Output Linkages and Vertical Integration," *Journal of the European Economic Association*, June 2021, *19* (3), 1816–1852.

**Bloom, Nicholas, Raffaella Sadun, and John Van Reenen**, "Americans Do IT Better: US Multinationals and the Productivity Miracle," *American Economic Review*, February 2012, *102* (1), 167–201.

**Egger, Peter H., Katharina Erhardt, and Gerard Masllorens**, "Backward versus forward integration of firms in global value chains," *European Economic Review*, 2023, *153*, 104401. Publisher: Elsevier.

**Grossman, Sanford J. and Oliver D. Hart**, "The costs and benefits of ownership: A theory of vertical and lateral integration," *Journal of political economy*, 1986, *94* (4), 691–719. Publisher: The University of Chicago Press.

**Hart, Oliver and John Moore**, "Property Rights and the Nature of the Firm," *Journal of political economy*, 1990, *98* (6), 1119–1158. Publisher: The University of Chicago Press.

**Lileeva, Alla and Johannes Van Biesebroeck**, "Outsourcing when investments are specific and interrelated," *Journal of the European Economic Association*, 2013, *11* (4), 871–896. Publisher: Oxford University Press.

**Liu, Mengxiao**, “The missing option in firm boundary decisions,” *European Economic Review*, February 2021, *132*, 103602.

**RDC of the Federal Statistical Office and Statistical Offices of the Federal States of Germany**, `https://doi.org/10.21242/42111.2021.00.01.1.1.0` 2021. Own calculations.

# Appendix

## Classification of industries and descriptive statistics

Table A2 displays industries categorized based on the two-digit ISIC Rev. 4 classification. Meanwhile, Table A1 provides the mean, standard deviation, and the number of observations for the main variables used in the analysis.

**Table A1:** Classification and description of industries

| Division | Description |
|---|---|
| 10 | Food products |
| 11 | Beverages |
| 12 | Tobacco |
| 13 | Textiles |
| 14 | Wearing apparel |
| 15 | Leather |
| 16 | Wood & cork |
| 17 | Paper products |
| 18 | Printing & media |
| 19 | Coke & Petroleum refineries |
| 20 | Chemicals & chemical products |
| 21 | Basic pharmaceutical products |
| 22 | Rubber and plastics products |
| 23 | Non-metallic mineral products |
| 24 | Basic metals |
| 25 | Fabricated metal |
| 26 | Electronic and optical products |
| 27 | Electrical equipment |
| 28 | Machinery and equipment |
| 29 | Motor vehicles |
| 30 | Other transport equipment |
| 31 | Furniture |
| 32 | Other manufacturing |
| 33 | Repair & installation |

*Notes*: The table lists industries based on the two-digit ISIC Rev. 4 classification.
*Source*: Author's presentation based on OECD input-output tables.

**Table A2:** Descriptive Statistics

| Variable | Mean | Std. Dev. | Obs |
|---|---|---|---|
| $V_{ijs}$ | .0021667 | .0464975 | 9,170,112 |
| $IO_{js}$ | .0416667 | .0910874 | 9,170,112 |
| $Domestic_{js}$ | .0278763 | .0616858 | 9,170,112 |
| $Foreign_{js}$ | .0137904 | .0364535 | 9,170,112 |
| R&D | .0175231 | .0213387 | 9,170,112 |
| Investment | .0337141 | .0087504 | 9,101,784 |
| $Age_i$ | 7.763 | 2.290 | 9,170,112 |
| $Employment_i$ | 210.567 | 2,084.348 | 9,111,288 |

**Notes:** The dependent variable, $V_{ijs}$, is a binary variable that equals 1 if a firm operates plants in both producing and supplying industries. $Domestic_{js}$ and $Foreign_{js}$ represent the domestic and foreign input cost shares, respectively, obtained from Germany's input-output tables. R&D denotes industry-level R&D intensity, calculated from AFiD plant-level data and the ANBERD database as the ratio of R&D expenditure to sales. Investment represents industry-level physical investment intensity, derived from AFiD plant-level data as the ratio of tangible capital expenditure to sales. Age refers to firm age, measured as the number of years since the firm's first recorded plant. Employment is the firm's total number of employees, calculated as the sum of all active employees across all firm plants as of September in the respective year.
**Source:** Author's calculations based on AFiD, input-output tables, and ANBERD datasets.

# Chapter 3

# Vertical Integration Under Contractual and Financial Frictions


## Abstract

This paper examines how contractual and financial frictions shape firm boundaries, using data from the German manufacturing sector. I build on a model of vertical integration that incorporates both contractual and financial constraints. In my empirical analysis, I measure contractual frictions using product complexity and measure financial frictions through the impact of the 2008 financial crisis, especially on industries reliant on external finance. The firm-level findings reveal that product complexity alone does not significantly influence vertical integration. However, financial frictions reduce vertical integration. Furthermore, using a triple-difference framework, I explore the combined effects of these frictions by exploiting the exogenous shock to credit supply caused by the financial crisis. The results indicate that the interaction between product complexity and financial constraints further discourages vertical integration. These findings are robust across multiple sensitivity checks and also hold in plant-level analysis.



**Keywords:** Contractual frictions, product complexity, financial frictions, financial crisis, vertical integration

# 1 Introduction

Contractual and financial factors are key institutional determinants shaping firm boundaries. Prominent theories of the firm emphasize the substantial impact of contractual frictions on firm integration (Gibbons, 2005). In addition, developments in international trade and production fragmentation have triggered theoretical and empirical investigations into the implications of financial constraints on the organization of global supply chains. Moreover, when production processes are dispersed across multiple sectors and countries, disruptions in one part of the supply chain, such as those caused by a shock to credit supply, can trigger ripple effects that disrupt production in other sectors and countries, too. This interconnectedness amplifies the negative impacts of financial frictions, such as those arising from credit supply shocks, as problems in one country and/or sector can spread to others through the global production network (Deardorff, 2001).

In this paper, I build on the model of vertical integration under contractual frictions and financial market imperfections proposed by Acemoglu et al. (2009). The model explores the supplier's decision to integrate forward with the manufacturer, where the integration incurs fixed costs. These costs are shaped by the quality of contractual institutions and the level of financial development within the economy. Consequently, these transaction costs significantly influence the decision to integrate versus outsourcing.

While financially developed countries typically have improved access to capital and lower transaction costs, they are not immune to credit market imperfections, much like less-developed economies. Even well-established financial systems can experience disruptions, as evidenced by the 2008 financial crisis. Furthermore, credit market frictions, such as higher interest rates, more stringent lending criteria, or limited access to financing, can still pose significant challenges for firms. My focus in this paper is on financial frictions, specifically taking a microeconomic perspective rather than concentrating on the broader macroeconomic development of financial systems in countries. The primary question I address empirically is: What is the combined effect of contractual and financial frictions on vertical integration among firms in Germany? In other words, given the presence of contractual incompleteness, how do the financial constraints faced by firms affect their organizational structure, particularly in terms of vertical integration? To answer this question, I empirically apply the model from Acemoglu et al. (2009), providing additional evidence on how these frictions impact vertical integration in a developed economy.

Despite their fundamental significance and the extensive literature on the topic, theories concerning firm boundaries lack definitive conclusions regarding the impact of changes in these factors on the organizational structure of the firm. On the one hand, even with the rigorous and formal Property Rights Theory (PRT) framework, clarity on the effect of contractual frictions on integration remains elusive. On the other hand, there remains no definitive consensus on the direct impact of financial development on vertical integration. Various studies have yielded inconclusive or contradictory findings regarding this relationship.[1]

Why do contradictory findings persist, and what factors contribute to the lack of consensus in theoretical and empirical studies? This question serves not as an endpoint but rather motivates further exploration into the effects of these factors on vertical integration. Table 1 provides an overview of the most related previous literature investigating this relationship, revealing contrasting theoretical predictions as well as empirical results. While a review of related literature will be presented in Section 2, my primary focus here is to shed light on the role of contextual variables (or moderating factors) through which contractual and financial institutions affect vertical integration.

Column 1 in Table 1 lists the studies, and column 2 distinguishes between institutional factors: contractual, financial, and the two factors combined in the first, second, and third panels, respectively. Column 3 presents the results of the integration, indicating the likelihood or extent of integration, and Column 4 presents the contextual factors.

The contextual factors examined in each study appear to play a significant role in explaining the variations in the impact of these frictions on integration decisions. In the first panel of Table 1, I list studies that exclusively focus on the main effects of contractual institutions. The inconclusiveness in findings is evident within and across studies, which can be attributed to various contextual variables. For example, while one study may find that contractual frictions lead to more integration, it could also be the case that the same or another study may find the opposite. This relationship varies depending on factors such as the maturity level of the product, contractual frictions within the headquarter, buyer, and seller (both domestic and foreign), productivity levels, degree of upstreamness, input substitutability, and relationship specificity.

In the second panel, I list studies focusing on the impact of financial development on vertical integration. Here, too, differences in the extent of integration arise due to contextual factors such as firm size distribution, external financial dependence, productivity, headquarters intensity, and sequential production (complements versus substitutes). Another body of literature, as shown in the last panel, examines the interaction between contractual and financial institutions in relation to vertical integration and also finds contradicting effects. This is a relatively small body of literature that focuses on the combined effects of these factors.[2] According to the table, the main difference between the last two studies is their

[1] The literature, including Acemoglu et al. (2009), Eppinger and Kukharskyy (2021), and Macchiavello (2012), has highlighted the challenge of precisely defining the impact of contractual and/or financial institutions on vertical integration.

[2] Alquist et al. (2019) examine mergers and acquisitions (M&A) in the context of foreign direct investment (FDI) in emerging markets, with a focus on financial development and institutions. The authors use corruption indices as a key measure of institutional weakness. While M&A transactions do not specifically focus on vertical integration, they could indeed involve vertical integration if the acquired firms operate in different stages of the supply chain.

**Table 1:** The effect of contractual and financial institutions on vertical integration

| Study | Institutional factors | Vertical integration | Contextual factors |
|---|---|---|---|
| Antràs (2005) | Good contractual institutions (lower frictions, better contractibility) | (+) | Standardized products |
| | | (-) | Less-standardized products |
| Antràs and Helpman (2008) | | (-) | Domestic sourcing (headquarter's frictions) |
| | | (+) | Domestic sourcing (supplier's frictions) |
| | | (+) | Foreign sourcing (south's frictions) |
| Macchiavello (2012) | | (+) | Input markets |
| Corcos et al. (2013) | | (+) | Most productive firms |
| | | (-) | Least productive firms |
| Antràs and Chor (2013) | | (+) | Supplier's industry contractibility |
| | | (-) | Buyer's industry contractibility |
| Alfaro et al. (2019) | | (+) | Upstream input's friction and demand elasticity |
| Eppinger and Kukharskyy (2021) | | (+) | In highly relationship-specific industries |
| Macchiavello (2012) | High financial development | (-) | Industries dominated by small firms |
| | | (+) | Industries dominated by large firms |
| Shen (2017) | | (-) | Financially dependent industries |
| | | (+) | Total factor productivity and headquarter intensity |
| Choi (2020) | | (-) | Sequential complements in production |
| | | (+) | Sequential substitutes in production |
| Acemoglu et al. (2007) | Weak contractual institutions & Low financial development | (+) | |
| Carluccio and Fally (2012) | | (-) | Open economy |
| Acemoglu et al. (2009) | Weak contractual institutions & High financial development | (+) | Closed economy |

*Notes*: These results stem from theoretical predictions validated by empirical analysis, with the exception Acemoglu et al. (2007) and Antràs and Helpman (2008), who exclusively offer theoretical findings. The symbols (+) and (-) indicate a higher and lower vertical integration, respectively.
*Source*: Author's investigation of literature.

application of the model to open versus closed economies.

Moreover, further examination reveals that, first, except for Acemoglu et al. (2009) and Macchiavello (2012), the studies in Table 1 depart from the Grossman and Hart (1986); Hart and Moore (1990)'s PRT approach. The inclusion of fixed costs in the theoretical framework of Acemoglu et al. (2009) and Macchiavello (2012), especially in their study of transaction costs of vertical integration, suggests a closer alignment with Transaction Cost Economics (TCE) rather than the PRT. Second, Eppinger and Kukharskyy (2021) and Carluccio and Fally (2012) employ IV strategy and Difference-in-Differences approach, respectively, allowing for a causal interpretation of their results, unlike the rest of the studies which provide mere correlations.[3]

This paper is closely aligned with the work of Carluccio and Fally (2012). However, the choice of using Acemoglu et al. (2009)'s model stems from its particular suitability for analyzing data from the German manufacturing sector, where firms face the decision of whether to outsource or integrate production domestically. Since Acemoglu et al. (2009)'s model is designed to describe firm boundaries in a closed economy, it is well-suited for my empirical analysis. In contrast, Carluccio and Fally (2012) examines a scenario involving a manufacturer and an overseas supplier, focusing on the financial development of the foreign supplier's country.

In addition, my analysis specifically emphasizes the supplier's costs and financial responsibilities during the integration process. Transfers between the supplier and the manufacturer occur only when they choose to integrate, with the supplier making payments to the manufacturer. These payments, along with the costs associated with ensuring contract compliance due to contractual incompleteness, are considered fixed costs of integration. In the outsourcing scenario, contractual frictions do not affect the supplier's profits, as there are no costs associated with monitoring or ensuring compliance. The manufacturer of the final product operates independently, and the supplier does not share in the revenues from the final product's sales. However, in the case of vertical integration, these costs directly impact the supplier's profits, as the final goods manufacturer becomes part of the integrated firm. Thus, Acemoglu et al. (2009)'s model emphasizes the financial burden that these fixed integration costs place on the supplier, a feature that is not present in Carluccio and Fally (2012)'s theoretical framework.

If incomplete contracts and financial frictions are significant factors influencing firm boundaries, what exactly is the mechanism underlying their interaction? One mechanism to understand the interaction effects between these two factors is that, to alleviate contractual frictions, parties might opt for vertical integration. This involves one firm taking control of another at a different production stage, thereby reducing the impact of costly and complex contracts. However, vertical integration demands significant financial resources to cover transaction costs. Thus, under weak contractual institutions, the

[3]Carluccio and Fally (2012) explore the interaction between contractual costs and financial development, whereas Eppinger and Kukharskyy (2021) focus solely on contractual institutions.

role of credit constraints and financial system becomes more important in determining vertical integration.

Another mechanism involves financially imperfect markets, where the supplier of an input may face financial constraints. Initiating investments demands liquidity beyond their means, prompting them to request funding from the manufacturer. However, if the products are complex, meaning that they are difficult to enforce, the manufacturer may not provide the liquidity requested by the supplier. This reluctance stems from the fact that the supplier retains a larger share of the ex-post surplus and may behave opportunistically by holding up the manufacturer. In cases where the supplier of a complex input is financially constrained, integration helps alleviate their financial constraints. In the context of the model employed in this paper, the former mechanism is more prevalent.

This paper is closely related to the literature that investigates the effects of contractual and financial institutions on the organizational structure of firms. Besides Acemoglu et al. (2009), Macchiavello (2012), and Carluccio and Fally (2012) present theoretical and empirical insights on how the quality of financial and contractual institutions impacts vertical integration. Acemoglu et al. (2009) use cross-country data on firms and show an apparent lack of a systematic relationship between financial development and vertical integration. As for contractual frictions, they find an ambiguous effect on vertical integration. They demonstrate that countries with poorer contractual institutions exhibit higher degrees of vertical integration but find no evidence supporting this result within industries. Additionally, cross-country differences in financial development and contracting institutions are correlated with more vertical integration. In line with their theoretical predictions, they find a positive and statistically significant interaction effect between contracting frictions and greater financial development on vertical integration.

In the model presented by Macchiavello (2012), firms (entrepreneurs) can either vertically integrate or outsource the production of final goods. Macchiavello (2012) emphasizes industry-specific characteristics, such as firm size distribution, and the importance of credit market imperfections, whereas Acemoglu et al. (2009) concentrate on country-level factors and the effects of financial development in environments with weak contract enforcement. Macchiavello (2012) considers firm entry into the market as a key mechanism that influences decisions regarding vertical integration and its associated effects. Using cross-country data, the author explains that financial development, by increasing access to credit and fostering the entry of firms, enhances competition. This competition reduces vertical integration among large firms, while prompting smaller firms to exit the market. Financial development has a heterogeneous impact on vertical integration, reducing it in industries where small firms generate a higher share of revenue, and increasing it in industries where large firms dominate.

Carluccio and Fally (2012)'s model focuses on backward vertical integration, where the manufacturer makes the integration offer to the supplier. In this scenario, financial constraints are significant, particularly for the supplier, who may lack the capital capacity to cover the cost of initial investments in complex inputs. Carluccio and Fally (2012) emphasize the manufacturer's role in mitigating these constraints for the supplier. Thus, ex-ante transfers occur under both sourcing modes (outsourcing and integration) and can be either negative (the manufacturer paying the supplier) for basic inputs or positive (the supplier paying the manufacturer) for complex ones.

Specifically, Carluccio and Fally (2012) use Antràs and Helpman (2008)'s framework and incorporate financial frictions to investigate the interaction effects of contractual incompleteness and the suppliers' financial constraints on vertical integration. Using data from the import transactions of French multinational firms, their findings indicate that multinational firms integrate their suppliers of complex goods in countries with lower financial development to mitigate the hold-up problem and alleviate their suppliers' financial constraints. Thus, in the case of complex inputs, outsourcing is more likely only when inputs are imported from suppliers located in a financially developed foreign country. Moreover, vertical integration is more likely when the supplier is located in a financially less developed foreign country. Complex inputs amplify this effect.

For the empirical analysis, I use micro-level data from the Federal Statistical Office of Germany (Amtliche Firmendaten für Deutschland, (AFiD)), matched with the input-output table for Germany obtained from the OECD database. By combining these datasets, I create a sample of firms matched with their respective plants, enabling the construction of a measure of vertical integration at the firm-level. Specifically, I interact the input cost shares of each industry pair with an indicator variable that equals one if the firm owns plants in both the supplying and producing industries.

The empirical literature often measures contractual and financial factors in various ways. In particular, to measure contractual frictions, previous studies have employed indicators such as the quality of the judicial system, product contractibility (Corcos et al., 2013), the 'rule of law' index (Nunn, 2007; Nunn and Trefler, 2014; Eppinger and Kukharskyy, 2021), procedural complexity index, contract enforcement procedures, legal formalism (Acemoglu et al., 2009), and contractual needs (Levchenko, 2007; Macchiavello, 2012).[4]

However, another body of literature uses product complexity as an indicator of contractual frictions. Complex products, featuring complex tasks and components, lead to incomplete contracts due to the challenges in fully describing them within contractual terms. Furthermore, while complexity measures are often based on qualitative data, R&D intensity has been employed as a quantitative measure for complexity (Carluccio and Fally, 2012). To measure the contractual frictions at the industry level, I

[4]The authors utilize a measure of contractual needs within industries, represented by the inverse Herfindahl index of input shares. This index, derived from input-output tables, measures the concentration of intermediate input usage.

use the Product Complexity Index (PCI) obtained from the Growth Lab at Harvard University. The PCI provides a comprehensive measure of how complex and advanced the production of a product is, reflecting the cumulative expertise and technological development across different industries.

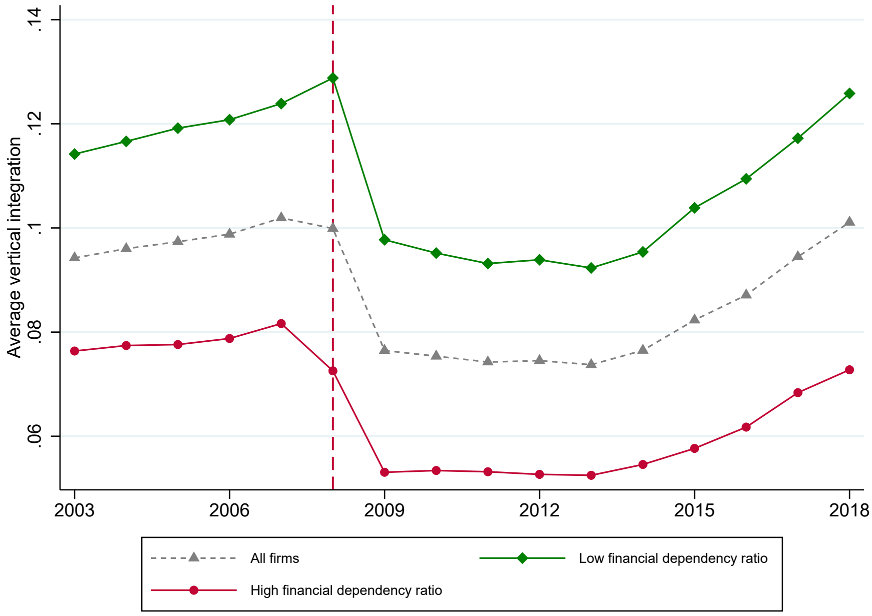
Additionally, financial (under)development is often measured in empirical research by the ratio of credit from private-sector banks and financial intermediaries to GDP (Acemoglu et al., 2009; Carluccio and Fally, 2012; Macchiavello, 2012; Shen, 2017; Alquist et al., 2019; Choi, 2020). Alternatively, in some instances, the net interest margin provides a measure of banking sector efficiency, hence assessing financial market imperfections (Carluccio and Fally, 2012). To assess an industry's external financial dependence, I calculate the financial dependence ratio using input-output tables. Specifically, I divide the input from the financial services sector by the total cost share of the industry, obtaining the share of financial services in production costs. I primarily use data from before 2008 to avoid the effects of financial shocks and endogeneity. Finally, to obtain a time-invariant measure, I compute the median financial dependence ratio for each industry at the two-digit level classification across all available years.

Figure 1 illustrates the evolution of vertical integration over time, showing both the overall trend for all firms and distinguishing between those with high and low external financial dependence. Prior to the 2008 financial crisis, vertical integration generally followed an upward trajectory for firms in both categories. However, in 2008, firms with high financial dependence experienced a decline in vertical integration, while firms with low financial dependence continued their upward trend. After the crisis, both groups saw a reduction in vertical integration, though the gap between them remained significant. From 2009 onward, firms with high financial dependence fell below the average vertical integration level of 0.06 and remained stable until 2014, when they began to show an increasing trend again. In contrast, firms with high financial dependence consistently had a higher level of vertical integration across all years, although they too saw a decline following the crisis and also witnessed an increase in the vertical integration average starting in 2014. This motivates the empirical analysis in this paper, as it highlights the role of financial frictions in shaping organizational structures.

Moreover, I present Figure 2 as a preliminary insight into the general pattern of the relationship between product complexity, financial dependence, and vertical integration. Figure 2a shows a positive relationship between vertical integration and PCI, suggesting that when contractual frictions are high, firms are more likely to internalize transactions to mitigate the risks of incomplete contracts and safeguard against opportunistic behavior. In contrast, Figure 2b reveals a negative relationship between vertical integration and financial dependence, indicating that greater financial dependence discourages firms from pursuing integration, likely due to the costs and risks associated with reliance on external capital.

The theoretical framework in this paper examines the effects of contractual and financial frictions, as

**Figure 1:** Evolution of vertical integration over time



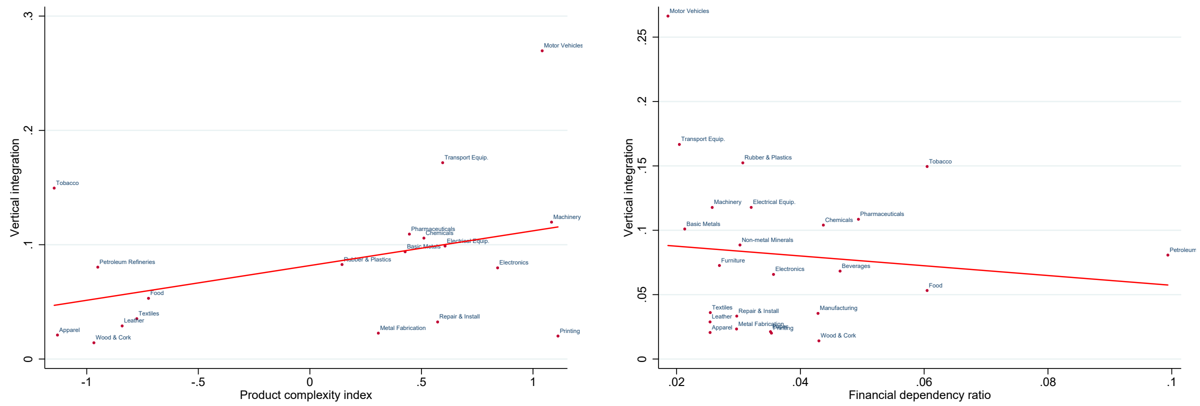

*Notes*:Figure 1 illustrates the evolution of vertical integration from 2003 to 2018, highlighting the period before and after the financial crisis, which is marked by a vertical line in 2008.
*Source*: Based on the author's calculations using the AFiD data and OECD input-output tables.

well as their interaction. The model predicts that there is an ambiguous relationship between contractual frictions (represented by product complexity) and vertical integration. Consistent with this theoretical prediction, the empirical results show variations in both the sign and statistical significance across different model specifications, indicating that the effect is not robust to model choices. In contrast, the theory predicts a negative relationship between vertical integration and financial frictions, implying that greater financial constraints discourage firms from pursuing integration. This is further supported by the empirical results, which demonstrate a negative association between financial frictions and vertical integration.

Additionally, when both financial and contractual frictions are simultaneously present, firms may find that the transaction costs associated with both types of frictions outweigh the efficiency gains from integration, leading them to prefer outsourcing. Empirically, the Triple-difference framework shows that, in the presence of incomplete contracts, financial frictions reduce the extent of vertical integration. Firms facing higher financial frictions may lack the capital necessary to cover the transaction costs of integration decisions.

My research contributes to the empirical literature investigating the effects of contractual and financial institutions on firm boundaries in several ways. One key contribution of this paper is its examination

**Figure 2:** Relationship between product complexity, financial dependence, and vertical integration




**(a)** Product complexity and vertical integration

**(b)** Financial dependence ratio and vertical integration

*Notes*: Figure 2a shows the relationship between product complexity and vertical integration across different industries. Figure 2b shows the relationship between the financial dependence ratio and vertical integration across different industries.
*Source*: Based on the author's calculations using the AFiD data, OECD input-output tables, and the Growth Lab at Harvard University database.

of both firm- and plant-level data, providing a detailed view of variation across individual plants within a firm. Although firm-level analysis offers a broad view of decision-making and integration, it may overlook plant-level variations. I highlight within-firm heterogeneity, as different plants within the same firm can vary in resources, product complexity, size, capital intensity, or the market concentration they face.

While financial development typically focuses on the overall growth and sophistication of a financial system, financial frictions highlight the specific obstacles or inefficiencies that hinder its smooth functioning. A more developed financial system suggests high levels of financial intermediation and broad access to services, but the presence of financial frictions signals underlying challenges that can impede efficient resource allocation and investment. Previous research has considered a country's financial system development as a factor determining financial constraints. Rather than focusing on cross-country comparisons of financial development, my study shifts the lens to the specific financial frictions across industries within the German manufacturing sector. By incorporating the 2008 financial crisis—an angle that has not been explored before—I contribute to the body of literature on the effect of financial frictions on vertical integration. This approach highlights the efficiencies within Germany's financial landscape, providing insights that differ from the broader discussions on a country's financial development.

Additionally, another contribution of this study is the use of the financial dependence ratio to assess the financial dependence of industries, relying solely on the input-output table. Previous studies have often used the external financial dependence (EFD) measure proposed by Rajan and Zingales (1998), which I employ as a robustness check in this study. Moreover, I present empirical evidence of a relationship that is more closely interpreted as a causal relationship rather than a mere correlation. The main identification strategy in this paper employs a Triple-difference approach, exploiting the exogeneity of contractual and financial frictions. In a previous study, Carluccio and Fally (2012) employed a Difference-in-Differences framework, where the differential effects arise between complex and basic goods. In contrast, in my paper, the differential effects primarily stem from exposure to the financial crisis. To examine the effects of financial frictions on vertical integration, I use the 2008 financial crisis and interact it with the industry's dependence on external finance to signify financial constraints. The model presented in this paper emphasizes the significance of financial frictions in shaping firm organizational structure. The impact of these frictions might be more pronounced in sectors heavily reliant on external financing, as Rajan and Zingales (1998) argue. The intuition is that firms within high external financial dependence industries are more susceptible to the credit shock. Incorporating the financial crisis as an exogenous shock to credit supply supports the interpretation of causal relationships in the empirical analysis, as it functions as a quasi-natural experiment.

Specifically, I investigate the impact of the financial crisis, which represents the exogenous treatment variable. The treated group comprises firms within industries heavily reliant on external finance, while the control group consists of firms in industries less dependent on external finance. Another exogenous source of variation is the product complexity measure, which is based on diverse technological advancements and productive know-how across various countries, reflecting an external framework beyond the direct control of individual firms. Using a Triple-difference framework, I interact the financial crisis dummy with the financial dependence ratio and PCI to provide empirical evidence of the effect of contractual and financial frictions on vertical integration decisions in Germany.

However, there are concerns regarding reverse causality. Erel et al. (2015) find that when a large firm purchases a smaller counterpart, it typically enhances the financial condition of the latter, especially if the smaller firm was facing financial constraints prior to the acquisition. Moreover, Görg and Kersting (2017) find that integrated suppliers linked with multinational corporations primarily finance their operations using internal resources, reducing reliance on external funds. This makes financing their operations less responsive to changes in their home country's financial frictions. Nonetheless, the quasi-experimental setup applied in this paper helps address these concerns.

Finally, while Acemoglu et al. (2009) have focused on cross-country data, my contribution lies in testing the applicability of the model in a different setting. Studying Germany is particularly interesting due to its status as the leading economy in Europe.

The paper is structured as follows. In Section 2, I provide a review of the related literature. In Section 3, I present the theoretical model and the mechanisms through which contractual and financial factors influence vertical integration. In Section 4, I present my data sources and describe the main variables. In Section 5, I present my empirical framework and findings, and conduct robustness checks to assess the sensitivity of my results to various model specifications. Finally, I conclude in Section 6.

## 2 Related literature

This paper is related to the literature studying the impact of contractual and financial institutions on vertical integration. The models developed to understand the relationship between contractual and financial frictions are fundamentally based on the property rights framework. These models emphasize the limitations in drafting complete contracts due to unforeseeable contingencies, highlighting issues related to non-contractibility.

A well-known departure from the property rights model of firm boundaries is the model developed by Antràs (2003), which assumes that all inputs are non-contractible. Antràs (2003) emphasizes the relative importance of total inputs provided by headquarters compared to those from suppliers, focusing primarily on the impact of factor intensity and resource endowment on intrafirm imports. This model has been widely used and extended by subsequent researchers.

Expanding on this framework, Antràs and Helpman (2004) expands on how variations in productivity within sectors influence firms' organizational choices. They highlight that only the most productive firms opt for integration. In contrast, Antràs and Helpman (2008) adopts the notion that contractibility varies across inputs and countries, as proposed by Acemoglu et al. (2007), shifting the focus to incentivizing non-contractible investments. Consequently, the relative importance of non-contractible investments made by headquarters versus those made by suppliers becomes crucial. Specifically, differences in contracting costs influence whether firms integrate or outsource. Higher contractual frictions in headquarter services tend to lead to integration, while outsourcing is more likely when dealing with suppliers of components. Moreover, weaker institutions in the Global South disproportionately affect the contractibility of inputs over headquarter services, influencing firms' decisions regarding foreign sourcing. In sectors where weak institutions in the Global South more strongly affect the contractibility of components than headquarter services, firms may find outsourcing more advantageous when offshoring production. Relocating production to foreign locations, particularly in regions with weaker institutions, may exacerbate contractual issues related to intermediate inputs.

Corcos et al. (2013) conduct empirical analysis motivated by the theoretical predictions of Antràs

(2003) and Antràs and Helpman (2004, 2008).[5] Utilizing French customs data, the authors examine the trade-off that firms encounter between intra-firm and arm's-length trade. They find a positive correlation between well-functioning judicial institutions in foreign countries and insourcing, especially for the most productive firms. Contract enforcement causes, nonetheless, the least productive ones to outsource. In line with Carluccio and Fally (2012), they find empirical support for a positive relationship between complexity of inputs and final goods and vertical integration.

In their study on North-South integration, Shen (2017) builds on the work of Antràs and Helpman (2004) by examining the relationship between sectoral productivity and headquarter intensity. The author extends the model to include country-level financial development and industry-level financial dependence. Shen (2017) notes that increased financial development tends to shift the balance towards outsourcing compared to integration. Using a sample of US intra-firm imports from 156 exporting countries, the empirical findings suggest that improvements in financial development are associated with a notable decrease in the median share of US intra-firm imports. This effect is more pronounced in sectors heavily reliant on external finance, though mitigated by higher productivity and headquarter intensity. While increases in Total Factor Productivity (TFP) lead to a moderate rise in the share of intra-firm imports, the most significant impact is observed with higher headquarter intensity. Firms with a stronger concentration of headquarter activities exhibit a substantial increase in intra-firm imports.

The primary emphasis of Eppinger and Kukharskyy (2021) involves investigating the impact of contracting costs and institutional quality on firm boundaries. Analyzing ownership share data from over 200,000 firm pairs, their results reveal a positive relationship between ownership shares and the high quality of contracting institutions. Furthermore, they indicate that better contracting institutions promote greater integration, particularly in industries characterized by relationship-specific investments. This implies that as contracting frictions decrease, the probability of integration increases.

Antràs (2005) proposes a model in which the early stages of a product's life cycle, involving extensive testing and marketing, require close oversight, leading to initial production in the same country where development occurs. In the early stages, contractual frictions prevent the relocation of production to foreign countries with lower wages. As the product matures and production becomes standardized, manufacturing shifts to low-wage countries. The decision to either integrate production within the firm or outsource it depends on the product's maturity level at the time of shifting. Outsourcing is more feasible when a product is mature and the threshold maturity level is high and a more standardized product, while integration through wholly owned foreign affiliates is more likely during the earlier stages of a

[5] Nunn and Trefler (2013) analyze intra-firm and arm's-length US imports data, covering 5705 products imported from 220 countries. In line with the theoretical predictions of Antràs (2003) and Antràs and Helpman (2004, 2008), their results suggest that intra-firm trade is largest where headquarter inputs are important and productivity is high.

product's life cycle, when the threshold maturity level is low. In general, when shifting production to the South, lower maturity levels lead to integrated production abroad within the firm's own organizational boundaries, with outsourcing occurring later in the product's life cycle.

Antràs (2005) argues that while contracts for revenues are legally binding in the North, they lack enforceability in the South. Building upon this framework, Basco (2013) expands the analysis by assuming that contracts are enforceable in both contexts, albeit with varying financial systems: differences in financial development play a crucial role in determining firm boundaries. Basco (2013) finds that financial development in the South (a financially less developed region relative to the North) plays a crucial role in facilitating offshoring, particularly for less standardized goods. This shift is driven by the trade-off between lower wages and contractual distortions. Improving Southern financial institutions boosts labor demand and reduces the Northern supplier's comparative advantage. This allows for cheaper offshoring, leading to a slight rise in Southern wages but more production shifting South due to better contracts. Nevertheless, the model does not address vertical integration decisions. Instead, the author focuses on the effect of financial development on offshoring. Using data on 145 trading partners of the US, the empirical findings corroborate the model, indicating that industries with higher R&D intensity (indicating a less standardized good) are more influenced by financial development in their offshoring decisions.

Acemoglu et al. (2007) constructs a theoretical framework, where tasks are partially contractible, to investigate the connection among incomplete contracts, technological complementarities, and the adoption of technology. This framework is then employed to explore additional implications of incomplete contracts, particularly in the context of vertical integration and outsourcing.[6] The model suggests that elevated contracting costs, stemming from weak institutions and credit market imperfections, result in a greater tendency towards vertical integration. In addition, greater complementarity between inputs increases vertical integration.

Antràs and Chor (2013) introduces the concept of sequentiality in production stages to Acemoglu et al. (2007)'s model, where high demand elasticity and strong input complementarity favor outsourcing upstream stages and vertically integrating downstream ones. In particular, firms integrate upstream stages when demand for the final product is relatively inelastic and production inputs are sequential complements. In such cases, integrating these upstream stages incentivizes investment by suppliers, which benefits downstream production. Moreover, in their empirical analysis, they control for a number of industry characteristics, among which are contractual frictions between the buyer and the supplier. Using data on US related party trade shares, and consistent with Antràs and Helpman (2008), they find that inputs that are contractible tend to be transacted more within firm boundaries, whereas a higher

[6] The authors demonstrate through their model framework that the interaction between contractual frictions and the choice of technology can significantly influence cross-country income disparities and international trade patterns.

degree of contractibility within buyer industries is linked with a decreased likelihood of integration. The authors, however, do not explicitly state whether the decision to integrate or outsource, influenced by contractual frictions or financial development, could yield different outcomes between the upstream and downstream stages of production. This becomes clearer in Alfaro et al. (2019) and Choi (2020).

Alfaro et al. (2019) extends Antràs and Chor (2013)'s framework and introduces contractual asymmetries across different input production stages. Essentially, the degree to which contracts can be enforced affects whether firms decide to outsource or integrate, considering factors like demand elasticity. They argue that the degree of contractibility of inputs significantly affects firms' ownership decisions, particularly in relation to upstream inputs. Using cross-country firm-level data, they find that when upstream inputs are more contractible, firms are more likely to integrate these inputs rather than outsourcing them, and the tendency to integrate upstream inputs is particularly pronounced when the firm faces high demand elasticity.

On the other hand, drawing from Antràs and Chor (2013)'s analysis, Choi (2020) integrates financial development into the analysis and expands upon their research by introducing the potential interaction effects between financial constraints, downstreamness, and sequential complements. Through an examination of US intra-firm import shares between 2000 and 2010, the author observes a correlation between credit market imperfections and the propensity for vertical integration. Specifically, when input suppliers reside in financially underdeveloped countries, firms show a greater inclination towards vertically integrating these suppliers when their inputs are sequential complements (this result is similar to Acemoglu et al. (2007)'s predictions for input complementarity). However, they are more inclined to opt for outsourcing when the inputs are sequential substitutes.

In a related context, the empirical literature also examines these frictions. In the context of the airline industry, known for its incomplete contracts, complex transactions, and frequent ex-post adaptations, Januszewski Forbes and Lederman (2009) explore the inclination towards vertical integration in situations where adaptation decisions are more common and expensive. They combine data from the US airline sector with precipitation and weather data. Their analysis reveals a robust correlation between adaptation decisions and vertical integration. They propose that firms are motivated to vertically integrate to reduce transaction costs associated with incomplete contracts and to negotiate post-execution adaptation decisions when needed.

In addition, Minetti et al. (2019) explores the relationship between firms' access to bank credit and their involvement in supply chains. Utilizing data from a 2010 survey of over 7,000 Italian firms, the study uncovers that companies experiencing bank credit rationing and having less robust relationships with banks are more likely to participate in supply chains. The study finds that firms with limited access to bank credit show weak evidence of increased participation in domestic supply chains but strong evidence of increased participation in international supply chains. Focusing on determining the optimal foreign ownership structure in international mergers and acquisitions (M&A), Alquist et al. (2019) present a model for cross-border acquisitions, where the foreign acquirer's choice of ownership is influenced by a trade-off between alleviating the target's financial constraints and managing the challenges and costs of operating in a low institutional quality environment. The authors find that complete foreign acquisitions are more prevalent in sectors heavily dependent on external funding, countries with high financial market imperfections, and those with better institutional quality. In addition, the impact of country-level financial development and institutional quality is more pronounced in sectors with a greater dependence on external finance.

# 3 Theoretical model

In this section, I introduce a theoretical model of vertical integration that considers contractual frictions and credit market imperfections, as proposed by Acemoglu et al. (2009). Essentially, incomplete contracts arise because it is often not possible to perfectly specify quality and payments, aligning with the TCE approach (Williamson, 1975, 1985) as well as with the PRT approach (Grossman and Hart, 1986; Hart and Moore, 1990). The focus in Acemoglu et al. (2009) is not on contractual incompleteness due to technological reasons, but rather on incompleteness caused by contract enforcement problems and institutional factors. In this paper, I explore contractual frictions that arise due to technological factors, such as product complexity, rather than institutional factors (e.g., the quality of courts). These factors make the manufacturer-supplier relationship challenging, particularly in terms of crafting and defining contracts.

I start with a scenario involving two contracting parties: a supplier $s$ and a manufacturer $m$.[7] The supplier provides an input $r$, which the manufacturer then uses to produce and sell output valued at $r$. Both are risk-neutral and aim to maximize their expected profits. The skills required for producing the input are supplier-specific, while producing the final product requires manufacturer-specific skills. Because of this division of labor, the final product cannot be produced without the involvement of both either the supplier or the manufacturer, as each provides specialized skills at different stages of the production process. Moreover, since only the supplier makes non-contractible investments, the manufacturer may behave opportunistically. Vertical integration addresses this by transferring control and property rights to the supplier, reducing the risk of holdup. In addition, the outside options of both parties are normalized to zero.

In the subsequent model descriptions, the supplier proposes offers under both outsourcing and verti-

[7] In this paper, I will refer to the supplier as 'he' and the manufacturer as 'she' for clarity.

cal integration. Under outsourcing, the supplier offers to supply an input to the manufacturer. However, should the supplier decide to integrate the manufacturer, he proposes to acquire the manufacturer by transferring a certain payment to her. Under vertical integration, where the supplier owns the manufacturer, he offers to pay her a wage for her services within the integrated relationship. These services refer to the various tasks or expertise that the manufacturer completes as part of her role within the integrated firm, which could encompass anything from coordination and management tasks to providing specialized knowledge or skills. Therefore, I consider below two possible organizational structures: outsourcing and vertical integration.

**Under outsourcing.** In the context of outsourcing, where the two parties are independent, the game is outlined as follows:

1. At time $t = 0$:
   - The supplier and the manufacturer are two independent firms and operate separately.
2. At time $t = 1$:
   - The game begins with the supplier proposing a contract to the manufacturer, offering the delivery of an input of specified quality $r_c$ at a price $p_c$ to be paid by the manufacturer.
   - The decision to accept the offer rests with the manufacturer.
3. At time $t = 2$:
   - Upon acceptance of the contract, the supplier chooses the actual input quality $r$ to be produced, but remains uncertain about whether the contract will be honored or not.
   - Following the contract acceptance:
     - There's a probability $\sigma$ that it will be upheld.
       * If the input matches $r_c$, the manufacturer receives the input, and the supplier is paid the agreed-upon price.
       * If the agreed-upon quality is not met, the supplier receives no payment, though the manufacturer still receives the input.
     - If the contract is not upheld after the offer has been accepted (probability $1 - \sigma$), the supplier and manufacturer engage in bargaining over the price that the manufacturer has to pay for the input of quality $r$ produced by the supplier.
       * Both parties have zero (ex-post) outside options and engage in asymmetric Nash/Rubinstein bargaining, with the supplier's bargaining power denoted by $\rho$, and that of the manufacturer by $(1 - \rho)$, where $\rho \in [0, 1]$.

  * When outside options are normalized to zero, the influence of alternative opportunities that each party might have outside the negotiation is eliminated. This simplification allows the focus to be solely on the dynamics of the bargaining process itself. Additionally, normalizing outside options to zero makes the asymmetric Nash/Rubinstein bargaining solution—considering a specific sequence of moves and discount factors—similar to the solution obtained by the asymmetric Nash bargaining model, which I use in the subsequent model description.

4. At time $t = 3$:

   - The supplier determines the quality $r$ once this uncertainty is unveiled.
   - Ultimately, transactions occur, after which the manufacturer proceeds with producing and selling the final product.

Acemoglu et al. (2009) interpret $\sigma$ as an indicator of the quality of legal institutions and the effectiveness of contract enforcement. In this paper, $\sigma$ measures the degree of product complexity inherent in the relationship between the supplier and the manufacturer, potentially amplifying contractual frictions. Specifically, when $\sigma = 1$, the probability that contracts are upheld is high, and the product is a basic good. Conversely, when $\sigma = 0$, the product is a complex good, and contracts are incomplete.

Product complexity refers to the degree of difficulty in designing, manufacturing, and using a product. It encompasses various factors such as the number of components, the complexity of the design, the level of customization, and the sophistication of the technology involved. Higher product complexity often requires more advanced skills, specialized knowledge, and detailed coordination among different processes and teams. Furthermore, complex products typically involve more uncertainties (such as production timelines, quality control, resource allocation, and cost management, among others) and potential for errors, making their production and management more challenging.

**Under (forward) vertical integration.** The supplier offers the manufacturer a wage ($w$ for her services and determines his own investment in quality $r$. The wage to the manufacturer is paid once revenues from production are realized. In the event that the services are not rendered, the manufacturer will not be entitled to receive the wage. However, to verify that the manufacturer has indeed completed the work, the supplier faces additional costs ($\tau(\sigma)$) associated with overseeing the manufacturer's services.[8] This pertains to the expenses the supplier must bear upon integration to oversee the manufacturer and ensure adherence to quality protocols, internal standards, and contractual agreements.

[8] The fixed cost of vertical integration ($\tau(\sigma)$) is considered as given, aligning this approach with the TCE rather than the PRT framework.

This cost depends on the product complexity: with a complex input, there may be a greater need for stringent quality control measures throughout the production process. This could include more frequent inspections, testing, and monitoring to ensure that the complex input is being used correctly and that the final product meets the required quality standards. The simpler the supplier's product, the less it costs to ensure compliance under vertical integration. Therefore, the cost (represented by $\tau$) may decrease as product complexity decreases. The intuition behind this is that enforcing contracts is easier when products are basic (less complex).

## 3.1 Equilibrium

The extent of social gains derived from vertical integration hinges on the difference between the surplus achieved through integration versus that of outsourcing. To illustrate this, I will begin by outlining the equilibrium reached through vertical integration first, followed by the equilibrium under outsourcing, then finally determining the total social surplus.

## Equilibrium under integration

The supplier's objective function is given as the revenue $r$ net of the cost of operation $c(r)$, the wage $w$ proposed by the supplier to the manufacturer, and $\tau(\sigma)$, which represents the fixed cost associated with the effort of the supplier to ensure that the manufacturer adheres to the contract terms in the case of vertical integration:

$$\pi_s(r, w, a^m) = (r - c(r) - w - \tau(\sigma))a^m.$$

Additionally, it is assumed that the supplier's cost function, denoted by $c(\cdot)$, is strictly increasing, convex, and differentiable with $c(0) = 0$. Furthermore, the cost function satisfies two Inada conditions: $c'(0) = 0$ and $\lim_{r\to\infty} c'(r) = \infty$. Moreover, $a^m$ (which can be either 0 or 1) indicates whether the manufacturer agrees to the offer. Given that her outside option is normalized to zero, she will agree to the offer ($a^m = 1$) if $w \geq 0$. Under vertical integration ($V$), the supplier's optimal contract includes

$$w^V = 0 \quad \text{and} \quad r^V = r^*.$$

Hence, vertical integration attains the best-quality level ($r^*$) chosen by the supplier. Note that due to the convexity of the cost function, $r^*$ is unique and determined by

$$c'(r^*) = 1, \tag{3.1}$$

where $c'$ is the marginal cost of the supplier.

Under vertical integration the supplier's and manufacturer's profits are:

$$\pi_s^V = r^* - c(r^*) - \tau(\sigma)$$

$$\pi_m^V = 0.$$

The strict convexity of the cost function, along with Equation 3.1, indicate that $r^* - c(r^*) > 0$. This implies that the supplier's profit depends on the magnitude of the transaction cost associated with vertical integration, $\tau(\sigma)$. In equilibrium, if $r^* - c(r^*) < \tau(\sigma)$, vertical integration does not occur since $\pi_s^V < 0$.

## Equilibrium under outsourcing

At time $t = 1$ in the described sequence of events, the manufacturer decides whether to accept or refuse the offer. Nonetheless, even if the manufacturer accepts the contract, there is still uncertainty for the supplier about whether the contract will be upheld. To analyze this situation, I start with backward induction, first considering the subgame inwhich the contracts are upheld.

The contract should be structured to incentivize the manufacturer to accept the offer. On the one hand, if the price ($p_c$) paid by the manufacturer is higher than the reward or benefit it gets from $r_c$, the manufacturer will not have an incentive to accept the offer (i.e., when $p_c > r_c$). On the other hand, if $p_c < r_c$, the supplier could potentially increase profits by raising the price for a given $r_c$. Therefore, to balance these incentives, the contract should specify that $p_c = r_c$. Thus, when contracts are upheld, the supplier chooses input quality $r = r_c$ and the manufacturer pays $p_c$. Contract enforcement leads to having $p_c = r_c$: the supplier makes a profit of $r_c - c(r_c)$ and the manufacturer makes zero profits.

Second, I now consider the scenario where contracts are not upheld and the quality of the input equals $r$. In this situation, the supplier engages in negotiations with the manufacturer to exchange the input. However, the input price is now weighted by the bargaining power of the supplier. This is due to asymmetric Nash bargaining at this stage, with zero outside options. The two parties agree to exchange $r$ at a price given by

$$p = \rho r.$$

To maximize profits, the supplier needs to find the optimal quality $r$ that maximizes $\rho r - c(r)$. I denote the optimal quality as $\hat{r}_\rho$. Thus:

$$c'(\hat{r}_\rho) = \rho. \tag{3.2}$$

Since $c(.)$ is convex, $\hat{r}_\rho$ is unique and is increasing in $\rho$. In case the contract is not upheld, the supplier will underinvest under outsourcing and choose a lower input quality. This happens because when $\rho < 1$, the adjusted quality level $\hat{r}_\rho$ falls below the optimal level $r^*$ achieved under vertical integration.

The expected profits of the supplier and manufacturer depends on the probability of whether contracts are upheld or not. Specifically, until the uncertainty surrounding this probability is resolved, the supplier's and manufacturer's expected profits under outsourcing ($N$) can be expressed as follows:

$$\pi_s^N = \sigma(r_c - c(r_c)) + (1-\sigma)(\rho\hat{r}_\rho - c(\hat{r}_\rho))$$

$$\pi_m^N = (1-\sigma)(1-\rho)\hat{r}_\rho.$$

The supplier is thus constrained to optimize the contractually specified quality level $r_c$.[9] Maximizing $\pi_s^N$ with respect to $r_c$ implies that

$$p_c = r_c = r^*.$$

This means that the quality specified in the contract must be equal to the efficient quality as in 3.1. When contracts are upheld, the profit received by the supplier is $r_c - c(r_c)$ with probability $\sigma$. When contracts are not upheld, the supplier makes $\rho\hat{r}_\rho - c(\hat{r}_\rho)$ with probability $(1-\sigma)$. Additionally, due the convexity of the cost function and since $\hat{r}_\rho > 0$, as a result, the ex-ante expected profits of the two firms are

$$\begin{aligned} \pi_s^N &= \sigma(r^* - c(r^*)) + (1-\sigma)(\rho\hat{r}_\rho - c(\hat{r}_\rho)) > 0 \\ \pi_m^N &= (1-\sigma)(1-\rho)\hat{r}_\rho > 0 \end{aligned} \tag{3.3}$$

## Social surplus

The total surplus from vertical integration is given by

$$\Delta\pi^V \equiv (\pi_s^V + \pi_m^V) - (\pi_s^N + \pi_m^N)$$

$$\Delta\pi^V = (1-\sigma)[(r^* - \hat{r}_\rho) - (c(r^*) - c(\hat{r}_\rho)] - \tau(\sigma).$$

Thus, substantial transaction costs under vertical integration reduce the profits. Opting for vertical integration will generate more profits compared to outsourcing only if $\tau(\sigma)$ is very low. Therefore, the costs of ensuring compliance by the contracting party are one of the key components of transaction costs. They can influence the choice of governance structures, such as vertical integration, in economic transactions. Ultimately, different governance structures are chosen based on the trade-off between minimizing these transaction costs and achieving economic efficiency.

[9] This implies that the supplier's decision-making is limited to choosing the revenue $r_c$ based on the quality level specified in the contract. This is his only choice variable.

## 3.2 Forward vertical integration

The decision to pursue vertical integration results in financial transactions occurring between the involved contracting parties. At time $t = 0$ in the outlined sequence under outsourcing, the supplier and the manufacturer operate independently. At this point, the supplier offers to integrate the manufacturer by paying a transfer $T$. If the manufacturer declines the offer, she does not receive the payment. Then both parties engage in the outsourcing game as described above starting at time $t = 1$. If the manufacturer agrees to the supplier's offer, two outcomes follow: she receives a payment of $T$ and vertical integration occurs.

However, there is a possibility that the supplier might have limitations on credit. To execute the transaction, the supplier faces expenses associated with securing financial support. Additionally, credit market imperfections can hinder the efficient allocation of capital and the smooth operation of borrowing and lending activities. Consequently, every euro allocated at the game's onset incurs a cost of $(1 + \theta)$ euros for the supplier, where $\theta \geq 0$ represents the degree of financial market frictions encountered by the supplier when vertical integration decision is taken. Thus, a higher $\theta$ indicates greater credit market frictions, making financing more expensive. The expected payoffs of the supplier and manufacturer before the organization structure is determined can be expressed as follows:

$$\Pi_s(A_m, T) = (1 - A_m)\pi_s^N + A_m(\pi_s^V - T(1 + \theta))$$

$$\Pi_m(A_m, T) = (1 - A_m)\pi_m^N + A_m(\pi_m^V + T).$$

This surplus depends on the manufacturer's acceptance of the offer, i.e. $A_m = 1$, and the specific transfer amount $T$. If the manufacturer decides to turn down the offer, i.e. $A_m = 0$, profits return to the levels indicated in Equation 3.3. The manufacturer's decision to accept the offer depends on whether the profits achievable under vertical integration plus the transfer amount from the supplier are at least equivalent to her profits under outsourcing:

$$\pi_m^V + T \geq \pi_m^N \equiv \hat{T} \geq (1 - \sigma)(1 - \rho)\hat{r}_\rho,$$

where $\hat{T}$ represents the specific payment that the supplier must make to the manufacturer for vertical integration to occur. For the supplier, its profitability is directly affected by the financial frictions linked to this payment:

$$\pi_s^V - \hat{T}(1 + \theta) \geq \pi_s^N.$$

In equilibrium, vertical integration will occur if:

$$(1-\sigma)[(r^* - c(r^*)) - (\hat{r}_\rho - c(\hat{r}_\rho)] - \theta(1-\sigma)(1-\rho)\hat{r}_\rho - \tau(\sigma) \geq 0. \tag{3.4}$$

This statement shows that the optimal organizational form will be vertical integration if the benefits of integration, which include efficiency gains ($(1-\sigma)[(r^* - c(r^*)) - (\hat{r}_\rho - c(\hat{r}_\rho)]$), outweigh the associated costs ($\theta(1-\sigma)(1-\rho)\hat{r}_\rho - \tau(\sigma)$). These organizational costs involve credit market imperfections ($\theta$) and transactional costs associated with vertical integration decision, $\tau(\sigma)$. When there are imperfections in credit markets or costs associated with integration, outsourcing can occur in equilibrium. Nevertheless, the absence of these costs makes vertical integration the logical option as it attains the optimal input quality level.

In light of this analysis, it is possible to derive two theoretical predictions.

Prediction 1 (Main effect): *Vertical integration is less likely (more likely) in situations where credit market imperfections (indicated by $\theta$) are high (low). The effect of contractual frictions ($\sigma$) is ambiguous.*

Inspecting the left-hand-side of inequality 3.4 (henceforth, L), one can see that an increase in $\theta$ reduces the value of the left-hand side of the inequality (but has no impact on the right-hand side). As a result, it becomes less likely that the inequality will hold, making the specified condition less likely to be satisfied. Moreover, the ambiguity of the effect of $\sigma$ arises because its effect on $L$ depends on $\theta$ and the difference between $r^*$ and $\hat{r}_\rho$.

Prediction 2 (Interaction effect): *When contractual frictions are severe, higher financial frictions make vertical integration less likely. Thus, vertical integration is more likely when both $\sigma$ and $\theta$ are lower, and less likely when they are higher.*

The second prediction of the model pertains to the interaction effects between the contractual and financial frictions. Vertical integration is most likely when both contractual frictions and financial frictions are lower, enabling firms to manage relationships efficiently and finance the necessary investments for integration. Using expression 3.4, I have:

$$\frac{\partial^2 L}{\partial\theta\partial\sigma} = (1-\rho)\hat{r}_\rho > 0.$$

This means that when $\theta$ and $\sigma$ are lower, the efficiency gains will be greater than the costs associated with vertical integration (in the left-hand side of expression 3.4 ).

# 4 Data and main variables

To confront the theoretical predictions with empirical evidence, I primarily rely on mirco-level data provided by the German Federal Statistical Office and the Offices of the Laender (2021a, (2021b). Specifically, I merge plant-level data (*AFiD-Panel Industriebetriebe*) with firm-level data (*AFiD-Panel Industrieunternehmen*). The dataset encompasses all manufacturing firms in Germany with a workforce of more than 20 employees. In addition, the data include information on firm activities, such as employment, capital expenditure, and sales.

One key benefit of this dataset is its comprehensive information about firms, their affiliated plants, and most importantly, the specific economic activities they engage in. Thus, I am able to establish a connection between each plant and its corresponding firm. Firms and plants are categorized at the four-digit level based on the International Standard Industrial Classification of All Economic Activities (ISIC). However, it is worth noting that the industry classification system in Germany underwent changes after 2008, leading to the reclassification of certain industries. To maintain a consistent industrial classification across different time periods, I use the conversion tables provided by Dierks et al. (2020). As a result, the final dataset comprises firms and plants classified according to ISIC Rev.4.[10] I focus on firms operating in manufacturing sectors.[11]

The dataset comprises detailed information on 65,304 manufacturing firms, encompassing a total of 73,574 plants, over the period 2003–2018. Within this dataset, approximately 80% are single-plant operations.[12] Some firms operate multiple plants within the same four-digit industry code. For this analysis, I consolidate all such plants into a single unit when they share the same industry classification. Finally, Table A1 in the Appendix presents descriptive statistics of the main variables used in this analysis.

## 4.1 Vertical integration measure

Assessing vertical integration necessitates data on whether a plant obtains inputs from another plant within the same firm. Typically, researchers do not observe this specific information. Input-output tables have been extensively employed in the literature to overcome this issue. These tables outline how the output of one industry becomes the input for another. I use input-output tables from the Organisation for Economic Co-operation and Development (OECD) database. These comprehensive tables provide

[10]The industry classification for the years 2003-2008 in the AFiD data is based on ISIC Rev. 3. In contrast, for the years 2009-2018, the industry classification follows ISIC Rev. 4.

[11]Table A2 in the Appendix presents a complete list of the ISIC Rev. 4 two-digit industry divisions and their corresponding descriptions.

[12]The observation that the majority of firms operate as single-plant firms aligns with earlier research findings. For example, this is similar to the findings of Bloom et al. (2012) and Alfaro et al. (2016), which also indicate a predominance of single-plant firms.

necessary information about the interrelationships between different sectors of the German economy and span the period from 2003 to 2018. Additionally, industries in the dataset are classified at the two-digit level according to ISIC Rev. 4.

To construct a measure of vertical integration at the firm-level, I begin by matching the AFiD dataset to the input-output table using the two-digit industry classification. Then, I create a binary variable indicating whether a firm possesses plants in both upstream and downstream industries. This variable, denoted as $D_{fkjt}$, equals one if firm $f$ owns plants in both sector $k$ (which supplies essential inputs) and sector $j$ (where the final product is produced) at time $t$. Otherwise, it is set to zero. Next, the input-output table indicates the euro of output from sector $k$ required to produce one euro of output in sector $j$ at time $t$. I use the input cost share $IO_{kjt}$ for the sector pair $kj$. The underlying assumption is that firms owning plants in both sectors can internally supply all the necessary inputs $k$ needed to produce $j$. From the perspective of the supplying firm, the higher the flow of inputs $k$ to output $j$ within a firm-owned $k$-producing plant, the more vertically integrated the firm is considered to be in the production of $j$.

The vertical integration index is calculated by summing the product of $D_{fkjt}$ and $IO_{kjt}$. I compute vertical integration at the firm-plant-time level. Equation 3.5 represents this measure, which reflects the degree of integration for plant $k$ owned by firm $f$ at time $t$:

$$V_{fkt} = \sum_{j} D_{fkjt} \times IO_{kjt} \tag{3.5}$$

This approach to construct the vertical integration index was first introduced in Fan and Lang (2000) using US input-output tables and later applied by Acemoglu et al. (2009), Macchiavello (2012), Alfaro and Chen (2012), and Alfaro et al. (2016).

In Table 2, I present an example to illustrate how I construct the vertical integration measure. Consider a firm $f$ operating in three industries: *Motor vehicles*, *Electrical equipment*, and *Fabricated metal*. According to the German input-output table, each of these industries supplies an input $k$ for the production of product $j$. Additionally, column 3 displays the input-output coefficients between these industries, showing the value of the input flow between industries for the year 2018. Since this firm does not own a plant in *Other transport equipment*, the dummy variable $D_{fkjt}$ is equal to zero for this industry. For each plant $k$, I calculate the measure $V_{fkt}$ as the sum of the product of $D_{fkjt}$ and $IO_{kjt}$ over all product sectors $j$.

In the firm-level analysis, I use the total vertical integration index $V_{ft}$ as in the last column of Table 2. This variable is computed by summing over all plants within the firm, and dividing by the total

**Table 2:** Construction of vertical integration measure

| Firm $f$ | | | | | |
|---|---|---|---|---|---|
| Input $k$ | Product $j$ | $IO_{kj}$ | $D_{fkj}$ | $V_{fkt}$ | $V_{ft}$ |
| Electrical equipment | Electrical equipment | 0.3251 | 1 | 0.3665 | 0.6135 |
| Electrical equipment | Fabricated metal | 0.0343 | 1 | 0.3665 | 0.6135 |
| Electrical equipment | Motor vehicles | 0.0070 | 1 | 0.3665 | 0.6135 |
| Electrical equipment | Other transport equipment | 0.0076 | 0 | 0.3665 | 0.6135 |
| Fabricated metal | Electrical equipment | 0.0228 | 1 | 0.2958 | 0.6135 |
| Fabricated metal | Fabricated metal | 0.2669 | 1 | 0.2958 | 0.6135 |
| Fabricated metal | Motor vehicles | 0.0061 | 1 | 0.2958 | 0.6135 |
| Fabricated metal | Other transport equipment | 0.0058 | 0 | 0.2958 | 0.6135 |
| Motor vehicles | Electrical equipment | 0.1317 | 1 | 1.1783 | 0.6135 |
| Motor vehicles | Fabricated metal | 0.1822 | 1 | 1.1783 | 0.6135 |
| Motor vehicles | Motor vehicles | 0.8644 | 1 | 1.1783 | 0.6135 |
| Motor vehicles | Other transport equipment | 0.0149 | 0 | 1.1783 | 0.6135 |

*Notes*: The table shows the construction of the vertical integration measure at the firm-level.
*Source*: Author's calculations using the OECD input-output tables for Germany for the year 2018.

number of sectors $S_{ft}$ in which firm $f$ operates at time $t$ as shown in Equation (3.6):

$$V_{ft} = \frac{\sum_k V_{fkt}}{S_{ft}} \tag{3.6}$$

## 4.2 Contractual frictions

Product complexity is a key determinant of incomplete contracts because investments in these products are too complex to fully specify in contracts (Hart and Moore, 1990). To measure contractual frictions, I use the Product Complexity Index (PCI) from the Growth Lab at Harvard University.[13] The Product Complexity Index (PCI) measures a product's sophistication based on two factors: (i) the number of countries that can produce it, and (ii) the economic complexity of those countries.

Complex products, like specialized machinery and electronics, are typically produced by a few advanced economies with high technological capabilities. In contrast, simpler products, such as raw materials, can be made by a broader range of countries, including less complex ones. The PCI ranks products by the diversity and sophistication of the productive capabilities required to produce them, reflecting

[13]The Product Complexity Index (PCI) is made available by the Growth Lab at Harvard University via `https://atlas.cid.harvard.edu/rankings/product` (accessed on November 15, 2024).

both technological advancement and the global distribution of production capacity. In this analysis, the exogeneity of product complexity comes from how the PCI is constructed. Product complexity is not directly controlled by individual firms but is instead influenced by a broader, external framework of global capabilities and technological development.

The PCI data is classified at the industry level using four-digit Harmonized System (HS) Codes. I convert these codes to ISIC Rev 3 using the Product Concordance tables from WITS.[14] Then, using conversion tables from Dierks et al. (2020), I further convert from ISIC Rev 3 to ISIC Rev 4. Finally, I match the PCI data with the main dataset using the industry classification of the plant ($k$) at the four-digit level. As a result of this conversion, the analysis includes fewer four-digit industries than in the raw PCI data set. At the two-digit level, this corresponds to 19 industries as depicted in Figure 3.

On average, across the years, the PCI exhibits substantial variation across industries. The lowest average PCI is observed in the *Manufacture of wearing apparel* (ISIC 1012), with a value of -1.4585, indicating relatively simpler products. In contrast, the highest average PCI is found in the *Manufacture of chemicals and chemical products* (ISIC 2822), at 1.3668, reflecting a higher level of product complexity. Figure 3 illustrates the average product complexity across manufacturing industries from 2003 to 2018 at the two-digit level. There is noticeable heterogeneity among industries in terms of product complexity. Industries such as *Tobacco*, *Wearing apparel*, and *Wood & cork* exhibit the least complexity, while *Machinery and equipment*, *Motor vehicles*, and *Electronic and optical products* feature the most complex products.

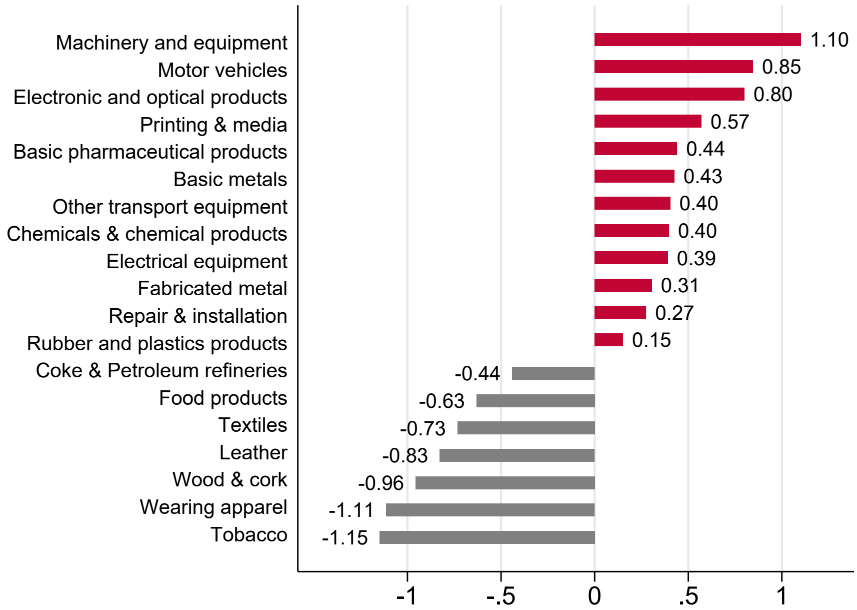
The complexity of products in these industries can be attributed to several factors. *Machinery and equipment* manufacturing, for instance, often involves complex designs, cutting-edge technologies, and specialized components, contributing to high complexity. Conversely, industries with lower complexity index typically produce goods with simpler designs and fewer technological requirements. These products often rely on traditional manufacturing techniques and may have straightforward compositions and functionalities, resulting in lower complexity levels compared to their counterparts in more technologically intensive sectors. This variation across industries helps in understanding the impact of product complexity on integration decisions.

## 4.3 Financial frictions

To measure financial constraints, I exploit the event of the 2008-2009 financial crisis as an exogenous shock to the credit market supply. While the financial crisis itself may not act as a direct indicator of a nation's financial development, it functions as a useful proxy for assessing the financial frictions and

[14]The World Integrated Trade Solution (WITS) offers product concordance tables and information on various product classifications. These tables are accessible via: `https://wits.worldbank.org/product_concordance.html`.

**Figure 3:** Average product complexity by industry (2003-2018)




*Notes*: Figure 3 illustrates the varying levels of product complexity across different industries.
*Source*: Based on the author's calculations using the database of the Growth Lab at Harvard University database.

constraints faced by firms to obtain credit. Firms in certain industries may heavily rely on external finance, while those in other industries may have a lower dependence on such funding sources. Some manufacturing industries that require significant investment in infrastructure, equipment, and technology often need more external funding. Firms in these industries are more exposed to the burden of financial crises compared to those in industries where firms rely more on internal funding. This diversity in financial dependency levels can significantly influence not only their investment decisions but also their input sourcing strategies.

During financial crises, industries heavily reliant on external finance often face severe vulnerability. Firms within these industries encounter challenges such as restricted access to credit, increased default risk, and disruptions in capital markets. Tightening credit conditions and supply chain disruptions further compound their difficulties. Unlike firms with greater internal financial resources, those heavily dependent on external financing may lack the resilience needed to withstand sudden economic shocks. Consequently, they are perceived to be more severely affected by financial crises, as their reliance on external funding magnifies their exposure to economic instability and market imperfections.

Typically, firm financial dependence is calculated based on several financial indicators, including the firm's debt ratio, the volatility of its stock returns, and its ability to generate internal cash flows.

Nevertheless, the AFiD data does not include this information that provides insights into how much a firm relies on external sources of funding relative to internal funds. Thus, I employ two measures to evaluate the external financial needs of the industry. First, I use the financial dependence ratio derived from the input-output table, as described below. Additionally, I incorporate the industry-level external financial dependence (EFD) measure from Eppinger and Neugebauer (2022) as a robustness check, providing an alternative perspective on the extent to which firms within an industry rely on external financing.

To analyze the extent of financial service usage across manufacturing industries, I use the OECD input-output table, specifically focusing on how much manufacturing relies on financial services. *Financial service activities, except insurance and pension funding* industry encompasses monetary intermediation, such as banking, and other financial services like investment management and securities trading. Given that the OECD input-output table aggregates financial services into a single category (combining sectors 64 to 66), I take the average across these categories.[15] I then assume that this measure mainly captures financial service activities, excluding insurance and pension funding.

To calculate the financial dependence ratio, I first compute the ratio of financial services inputs to the total output of the manufacturing sector, which includes financial services inputs. Next, I restrict the analysis to observations from years prior to 2008. Using pre-crisis data helps exclude immediate financial shocks and avoid endogeneity issues, providing a clearer and more stable measure of financial dependence. Finally, I compute the median value of the financial services input for each industry. This ratio is then a time-invariant, industry-specific measure, and is expressed as:

$$\text{Fin} = \frac{\text{Financial Services Input}}{\text{Total Output}}.$$

In the robustness checks section, I use the external financial dependence measure. The EFD measure is an index used to measure a firm's reliance on external financing sources, such as bank loans and bonds, as opposed to internal funds (profits and retained earnings) for investments and operational needs. I use the industry ranking of EFD provided by Eppinger and Neugebauer (2022). This index is particularly significant for my study because it offers an EFD measure specific to Germany. The authors construct this index utilizing information gathered from the "European Firms in a Global Economy" project in 2010.[16] This specific time frame chosen is crucial for the EFD measure as it eliminates the immediate disruptions from the 2008-2009 financial crisis. By 2010, industries had stabilized and adjusted to the new economic conditions. Trhe EFD index is based on firms' responses to the survey question

[15] Sectors 64 to 66 include: 64 - Financial service activities, except insurance and pension funding; 65 - Insurance, reinsurance, and pension funding; 66 - Activities auxiliary to financial services and insurance.

[16] See Eppinger and Neugebauer (2022) for comprehensive details on the data survey and the construction of the EFD index.

within each German industry. The survey asks: "*In the industry your firm operates in, how reliant are companies on external financing?*" with response options ranging from 1 (not dependent at all) to 5 (extremely dependent). The EFD measure is then calculated as the arithmetic mean of the firms' responses in each industry.

Industries in the provided EFD data are classified using the NACE Rev 1.1 system at the two- or three-digit level. Converting industries from NACE 1.1 to NACE 2 at the two-digit level is difficult with EFD data because there is not a direct one-to-one correspondence. As a result, I converted them manually. However, this process led to the exclusion of the industry *Repair and installation of machinery and equipment* from the analysis due to the lack of a clear mapping.[17]

Initially, I create a dummy variable that equals 1 for the years 2008-2009 and 0 for all other years. Then, I interact this financial crisis dummy with the financial decadency ratio. Industries with a high share of financial services are typically those that require significant financing to support their production processes. Similarly, I interact the crisis dummy with the EFD measure inthe robustness checks part. The interaction term reveals the financial frictions experienced by the industry. In times of financial stress—the 2008 financial crisis in his case—firms in these industries are more vulnerable to financial frictions. As access to external finance tightens, manufacturing firms with higher financial service needs face increased challenges in securing the funding required for their operations. In this analysis, firms within high financial dependency industries are designated as the treatment group, while firms in low financial dependency industries act as the control group.

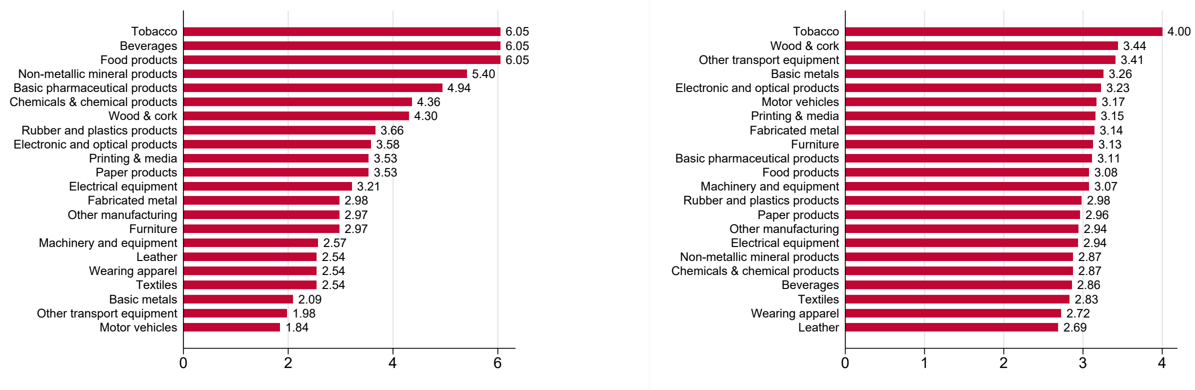
Figure 4 illustrates the varying levels of external financial dependence across different manufacturing industries in Germany, focusing on both the financial dependence ratio and the EFD. The financial dependence ratio shows considerable variation across sectors. The *Tobacco* industry stands out with the highest financial dependence ratio of 6%, indicating a significant reliance on external financing.[18] In contrast, the *Motor Vehicles* industry exhibits the lowest financial dependence ratio at 1.84%, signifying relatively low external financing reliance. Regarding the EFD measure, the *Tobacco* industry has the highest EFD, too, while the *Leather* industry shows the lowest EFD. Comparing the two measures, the *Tobacco* industry ranks among the most financially dependent in both the financial dependence ratio and the EFD measure, while the *Leather* industry exhibits low dependence in both metrics.

However, notable differences emerge in the overall financial dependency ranking across industries, highlighting variation between the two measures. These differences should not be problematic, as each

[17]Table A3 in the Appendix shows the manually mapped two-digit industry classification between the provided NACE Rev 1.1 and the corresponding NACE Rev 2, along with a ranking of the industries based on their external financial dependence. I further validated this matching by comparing it with the final industry mapping from the AFiD data, as outlined in Dierks et al. (2020). I found that all industries are correctly matched except for Repair and installation of machinery and equipment, which I have excluded from the analysis.

[18]The data for the *Tobacco*, *Beverages*, and *Food Products* industries are combined, which is why they share the same financial dependence ratio.

**Figure 4:** Financial dependence by industry in Germany




**(a)** Financial dependence ratio (%)    **(b)** External financial dependence

*Notes*: Figure 4a shows the financial dependence ratio in percentage for each manufacturing industry in Germany calculated using data between 2003 and 2008. Figure 4b shows the arithmetic mean of the ranking of industries according to their external financial dependence in 2010.
*Source*: Based on the author's calculations using the database of the Growth Lab at Harvard University and the external financial dependence measure (EFD) from Eppinger and Neugebauer (2022).

measure captures different dimensions of financial dependency, offering a broader view of an industry's reliance on external finance. First, the ratio captures financial dependency based on actual usage of financial services (e.g., banking, credit) by the industry, while the EFD is calculated from survey data in which firms rate their industry's dependence on external financing. These differences imply that each measure provides unique insights into financial dependence. Second, the correlation between the financial dependence ratio and the EFD is weak, with a non-statistically significant correlation coefficient of -0.1 (a p-value of 0.5833). This indicates that the two measures capture distinct aspects of financial dependence. Therefore, I use both measures in separate regressions, as they offer complementary insights. Since the two measures rank industries differently, comparing results across both can serve as a robustness check. If the effects of integration are consistent across both regressions, it enhances the credibility of the findings.

It should be noted that I have dropped the *Manufacture of coke and refined petroleum products* both in the figure above and in the empirical analysis because of extreme values. In these values, the variable Fin ranks this industry as the highest in terms of external financial dependence, while the EFD ranks it as the lowest.[19]

[19] However, I have also run the regressions including *Manufacture of coke and refined petroleum products* and obtained similar results.

# 5 Empirical framework and results

According to the model, the main effects of contractual frictions are ambiguous. However, a positive (negative) relationship is expected between lower (higher) financial imperfections and vertical integration. Furthermore, the combined impact of contractual and financial frictions is expected to reduce integration. Specifically, given the complexity of products (indicative of higher contractual frictions), the financial crisis is expected to lead to less integration.

In this section, I first examine the main effects of contractual and financial frictions, followed by their interaction effects. In the first part of my empirical analysis, I focus on the plant-level variation in vertical integration, providing a more granular examination. In the subsequent section, I transition to an aggregate analysis at the firm-level.

## 5.1 Firm-level analysis

I start with firm-level analysis and match the PCI data to the main AFiD dataset based on the four-digit industry classification of the firm's industry, where $f \in k$. I then match the $Fin$ variable to the two-digit industry classification of the firm. Additionally, the control variables are computed at the four-digit industry level. The measure of vertical integration is now used at the firm-level in the regression below, as computed in Equation 3.6.

**Main effects.** I examine the main effect of product complexity on vertical integration by conducting a simple regression analysis with fixed effects for firm and year, as depicted in Equation 3.7:

$$V_{ft} = \beta_0 + \beta_1 \mathbf{PCI}_{kt} + \beta_2 Emp_{ft} + \beta_3 CapInt_{kt} + \beta_4 Conc_{kt} + \delta_f + \delta_t + \epsilon_{ft}, \;\; f \in k, \tag{3.7}$$

where $V_{ft}$ represents vertical integration of firm $f$ at time $t$. The parameter $\beta_1$ captures the effects of contractual frictions. PCI$_{kt}$ measures the complexity of product $k$ at the four-digit industry level at time $t$. The supplying firm faces higher contractual frictions when the input is more complex. According to the Transaction Cost Economics (TCE) framework, these frictions incentivize suppliers to integrate with manufacturers in order to avoid hold-up problems, suggesting a positive relationship between product complexity and vertical integration. While the theoretical predictions leave the sign of $\beta_1$ ambiguous, this regression aims to provide empirical evidence and further insights on this matter.

Additionally, I include control variables at both the firm-level and the four-digit industry level to account for other factors that may influence vertical integration. Specifically, $Emp_{ft}$, account for firm size and computed using the natural logarithm of the number of employees. $CapInt_{kt}$ is industry-level capital intensity calculated as the natural logarithm of total capital expenditure relative to sales. I also account for the degree of competition by adding a variable that captures market concentration ($Conc_{kt}$) within a specific sector over time, utilizing sales data from all firms operating in that sector. I measure this variable using Herfindahl-Hirschman Index (HHI), which is calculated by summing the squares of the market shares of all firms in the sector for each time period, as follows:

$$\text{Conc}_{kt} = \sum_{f=1}^{N} \left( \frac{\text{Sales}_{fkt}}{\text{Sales}_{kt}} \right)^2$$

where $N$ is the number of firms in sector $k$.

Furthermore, $\delta_f$ is firm fixed effects which control for characteristics unique to each firm. However, I observe that firms in the dataset do not experience changes in their core industry over time. Therefore, adding industry (of the firm) fixed effects becomes redundant, as this characteristic is time-invariant and will be accounted for through firm fixed effects.[20] $\delta_t$ controls for any common factors or shocks affecting all firms within a particular year. These fixed effects help mitigate potential biases arising from omitted variables, time-invariant factors, and time-variant factors that affect all firms in the same way across different time periods. Since the data is panel data, observations from the same firm over time are not independent. This correlation within a firm can bias standard error estimates if not addressed. To fix this, I adjust for correlations in the error terms within each firm by clustering standard errors at the firm-level. Finally, $\epsilon_{ft}$ represents the error term.

Table 3 presents the main effects of contractual frictions derived from the outcomes of Equation 3.7. In column 1, where the model is estimated without control variables or firm fixed effects, the coefficient on PCI is positive and statistically significant. This implies that a one standard deviation increase in PCI is associated with a 0.734 unit change in the dependent variable. But this result is not stable across other specifications. In columns 2 and 3, where I add firm size, industry capital intensity, and concentration, and then firm fixed effects, respectively, both coefficients remain positive but are statistically insignificant.

Now I estimate the effect of financial constraints on vertical integration by running following regression:

$$V_{ft} = \alpha_0 + \alpha_1 \, \mathbf{FC}_t \times \mathbf{Fin}_k + \alpha_2 \, Emp_{ft} + \alpha_3 \, CapInt_{kt} + \alpha_4 \, Conc_{kt} + \delta_f + \delta_t + \epsilon_{ft}. \qquad (3.8)$$

The key distinction between this regression equation and Equation 3.7 lies in the interaction term $FC_t \times Fin_k$, which represents the financial market frictions and captures the effect of the financial

[20] Note that in the empirical analysis, I have 18 industries at the two-digit level, which are covered by the variables PCI and FIN. In Section 6 of the Appendix, I list all the industries included in the analysis and explain the reasons for excluding the other industries.

**Table 3:** Main effects of contractual frictions: firm-level analysis

| Dep. Var: $V_{ft}$ | (1) | (2) | (3) |
|---|---|---|---|
| PCI | 0.734*** | 0.056 | 0.014 |
| | (0.070) | (0.068) | (0.085) |
| Emp | | 4.241*** | 1.241*** |
| | | (0.372) | (0.147) |
| CapInt | | -0.084 | -0.004 |
| | | (0.137) | (0.080) |
| Conc | | -3.549*** | -0.801 |
| | | (0.838) | (0.608) |
| Observations | 200,145 | 198,741 | 195,462 |
| R-squared | 0.004 | 0.136 | 0.896 |
| Control variables | No | Yes | Yes |
| Year FE | Yes | Yes | Yes |
| Firm FE | No | No | Yes |

*Notes*: This table displays the means and standard deviations of the firm- and industry-specific variables for the period 2003-2018. Standard errors, in parentheses, are clustered at the firm-level. The variable $V_{ft}$ captures the integration level of firm $f$ at time $t$, derived from AFiD and input-output tables. It is derived by combining input cost shares with a binary variable indicating if $f$ owns plants in both upstream and downstream sectors. $PCI$ is the industry-level product complexity index obtained from the Growth Lab at Harvard University. From the AFiD data: *Emp* indicates firm size and is computed as the natural logarithm of the number of employees, *CapInt* is the natural logarithm of capital expenditure to total sales at the industry level, and *Conc* measures industry competition and is computed using firm sales according to the Herfindahl index approach. Industries are classified at the four-digit level according to ISIC Rev. 4. $^{*}$, $^{**}$, and $^{***}$ indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
*Source*: Author's calculations based on the AFiD data, OECD input-output tables, and the Growth Lab at Harvard University datasets.

crisis on vertical integration. In this regression, the industry in $Fin_k$ is defined at the two-digit level. The focus of this analysis is on understanding how financial market conditions, particularly the financial crisis of 2008-2009, influenced vertical integration decisions. To do this, I treat the financial crisis as a "treatment" factor, assigning firms in industries with high financial dependence (i.e., those with a high financial dependence ratio) to the "treated" group.

The interaction term $FC_t \times Fin_k$ is central to the analysis. Here, $FC_t$ is a dummy variable indicating the presence of the financial crisis (2008-2009), while $Fin_k$ measures the extent to which an industry is reliant on external finance. This interaction term allows me to assess how the effect of the financial crisis varies across industries with different levels of financial vulnerability. Essentially, this model employs a Differences-in-Differences (DiD) method, which compares the pre- and post-crisis changes in vertical integration. I expect the coefficient $\alpha_1$ to be negative, which suggests that the financial crisis led to a reduction in vertical integration, particularly for firms in industries that are highly dependent on external finance.

Table 4 presents the results following the same specifications as those in Table 3. In column 1, the coefficient for financial dependence is negative and statistically significant at the 1% level, suggesting that a one standard deviation increase in financial dependence is associated with a reduction of 0.717 units in vertical integration. The interaction term between FC and Fin is $-0.129$, indicating that during a financial crisis, being financially dependent further reduces vertical integration by an additional 0.129 units. In other words, firms that are financially dependent experience a greater reduction in vertical integration when faced with a financial crisis, compared to firms that are less financially dependent. This highlights that financial market frictions, such as a crisis, exacerbate the negative relationship between financial dependence and vertical integration. These results show the average effect of financial frictions on vertical integration, while in Table 5, this effect varies depending on the level of product complexity.

**Table 4:** Main effects of financial frictions: firm-level analysis

| Dep. Var: $V_{ft}$ | (1) | (2) | (3) |
|---|---|---|---|
| Fin | -0.717*** | -0.314*** | -4.345*** |
| | (0.100) | (0.072) | (0.968) |
| FC × Fin | -0.129*** | -0.129*** | -0.089*** |
| | (0.048) | (0.048) | (0.032) |
| Emp | | 4.222*** | 1.249*** |
| | | (0.365) | (0.147) |
| CapInt | | -0.187 | -0.044 |
| | | (0.132) | (0.082) |
| Conc | | -3.245*** | -1.037* |
| | | (0.792) | (0.622) |
| Observations | 200,145 | 198,741 | 195,462 |
| R-squared | 0.004 | 0.137 | 0.896 |
| Control variables | No | Yes | Yes |
| Year FE | Yes | Yes | Yes |
| Firm FE | No | No | Yes |

*Notes*: This table displays the means and standard deviations of the firm- and industry-specific variables for the period 2003-2018. Standard errors, in parentheses, are clustered at the firm-level. The variable $V_{ft}$ captures the integration level of firm $f$ at time $t$, derived from AFiD and input-output tables. It is derived by combining input cost shares with a binary variable indicating if $f$ owns plants in both upstream and downstream sectors. $Fin$ is the financial dependence ratio and is calculated as the median of the ratio of financial services inputs to total output for the manufacturing sector, using pre-2008 data. FC is a binary indicator, taking the value of 1 to denote the occurrence of a financial crisis during the years 2008-2009, and 0 otherwise. From the AFiD data: *Emp* indicates firm size and is computed as the natural logarithm of the number of employees, *CapInt* is the natural logarithm of capital expenditure to total sales at the industry level, and *Conc* measures industry competition and is computed using firm sales according to the Herfindahl index approach. Industries are classified at the four-digit level according to ISIC Rev. 4, except for the $Fin$ variable, which is defined at the two-digit industry level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
*Source*: Author's calculations based on the AFiD data and OECD input-output tables.

**Interaction effects.** The second prediction posits important interaction effects between contractual and financial frictions. The identification strategy in this section hinges on the exogeneity of the PCI, the measures of reliance on external finance, and the financial crisis event. This exogeneity is crucial for interpreting results causally rather than as mere correlations. To evaluate this prediction, I employ a Triple-difference method (Difference-in-Differences-in-Differences) as depicted in Equation 3.9:

$$\begin{aligned} V_{ft} = \mu_0 + \mu_1\, PCI_{kt} + \mu_2\, FC_t + \mu_3\, Fin_k + \mu_4\, (FC_t \times Fin_k) + \mu_5\, (FC_t \times PCI_{kt}) \\ + \mu_6\, (PCI_{kt} \times Fin_k) + \mu_7\, (\mathbf{PCI_{kt}} \times \mathbf{FC_t} \times \mathbf{Fin_k}) + \mu_8\, Emp_{ft} + \mu_9\, CapInt_{kt} \\ + \mu_{10}\, Conc_{kt} + \delta_f + \delta_t + \epsilon_{ft}, \;\; f \in k. \end{aligned} \tag{3.9}$$

The main effects hold even after incorporating the triple interaction in Table 5. Particularly, the main effect of PCI continues to be spurious. In contrast, the effect of financial frictions is persistent, continuing to be negative and statistically significant. More importantly, the triple interaction demonstrates a negative relationship with vertical integration and it is statistically significant coefficient, particularly in columns 2 and 3, where controls and fixed effects are included. In the preferred specification in column 3, the results suggest that a one standard deviation increase in product complexity, combined with a financial crisis, leads to a stronger negative effect on vertical integration, particularly for firms in industries highly dependent on external finance. This indicates that firms with high product complexity and high financial dependency experience a reduction in vertical integration, particularly in times of financial crisis.

## Firm-level robustness checks

The literature typically uses the external financial dependence (EFD) measure to assess the level of financial dependence across industries. In Table 6, I use the EFD variable instead of the financial dependence ratio in the regression from Equation 3.9. Overall, I observe similar patterns in the main effects. The triple interaction aligns with the results obtained using the financial dependence ratio. Across all specifications, the coefficients are negative and statistically significant. Therefore, even with an alternative measure of financial dependence, I demonstrate that the results remain robust, showing a negative effect of both contractual and financial frictions on vertical integration.[21]

Additionally, the timing of the financial crisis varies across the literature. While it commenced in the US in 2007, it spread to other countries, persisting for several years. According to the World

[21] Table 6 contains fewer observations due to the exclusion of the industry *Repair and installation of machinery and equipment*. To ensure consistency in the industry rankings between the variables Fin ad EFD, I conducted the analysis excluding this industry when using the Fin variable as well. The results remained consistent across both specifications.

**Table 5:** Interaction effects of contractual and financial frictions: firm-level analysis

| Dep. Var: $V_{ft}$ | (1) | (2) | (3) |
|---|---|---|---|
| PCI | 0.865*** | -0.348** | -0.194 |
| | (0.147) | (0.148) | (0.164) |
| Fin | -1.487*** | -0.870*** | -3.070** |
| | (0.268) | (0.266) | (1.291) |
| FC × Fin | -0.319*** | -0.260** | -0.180*** |
| | (0.121) | (0.126) | (0.065) |
| FC × PCI | -0.072 | 0.042 | 0.086** |
| | (0.067) | (0.076) | (0.037) |
| PCI × Fin | -1.193*** | -0.322 | -0.271* |
| | (0.179) | (0.215) | (0.164) |
| PCI × FC × Fin | -0.115 | -0.175** | -0.173*** |
| | (0.084) | (0.083) | (0.048) |
| Emp | | 6.534*** | 1.611*** |
| | | (1.091) | (0.179) |
| CapInt | | -0.187 | 0.152 |
| | | (0.437) | (0.107) |
| Conc | | -6.658*** | -1.713* |
| | | -2.524 | (0.976) |
| Observations | 200,145 | 198,741 | 195,462 |
| R-squared | 0.003 | 0.057 | 0.933 |
| Control variables | No | Yes | Yes |
| Year FE | Yes | Yes | Yes |
| Firm FE | No | No | Yes |

*Notes*: This table displays the means and standard deviations of the firm- and industry-specific variables for the period 2003-2018. Standard errors, in parentheses, are clustered at the firm-level. The variable $V_{ft}$ captures the integration level of firm $f$ at time $t$, derived from AFiD and input-output tables. It is derived by combining input cost shares with a binary variable indicating if $f$ owns plants in both upstream and downstream sectors. $PCI$ is the industry-level product complexity index obtained from the Growth Lab at Harvard University. $Fin$ is the financial dependence ratio and is calculated as the median of the ratio of financial services inputs to total output for the manufacturing sector, using pre-2008 data. FC is a binary indicator, taking the value of 1 to denote the occurrence of a financial crisis during the years 2008-2009, and 0 otherwise. From the AFiD data: *Emp* indicates firm size and is computed as the natural logarithm of the number of employees, *CapInt* is the natural logarithm of capital expenditure to total sales at the industry level, and *Conc* measures industry competition and is computed using firm sales according to the Herfindahl index approach. Industries are classified at the four-digit level according to ISIC Rev. 4, except for the $Fin$ variable, which is defined at the two-digit industry level. $^{*}$, $^{**}$, and $^{***}$ indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
*Source*: Author's calculations based on the AFiD data, OECD input-output tables, and the Growth Lab at Harvard University datasets.

Bank's Global Financial Development Database (GFDD, Cihak et al., 2012), the onset of the banking crisis began in Europe by 2008, persisting until 2010. Although the primary crisis period is typically recognized as 2008-2009, I analyze the effects of these frictions by considering the crisis period to 2008-2010 and present the results in Table 7. The estimated impact of PCI remains non-robust and

**Table 6:** Using the EFD measure: firm-level analysis

| Dep. Var: $V_{ft}$ | (1) | (2) | (3) |
|---|---|---|---|
| PCI | 0.743*** | 0.091 | -0.044 |
| | (0.071) | (0.072) | (0.089) |
| EFD | -0.140 | 0.143 | 0.288 |
| | (0.111) | (0.109) | (0.217) |
| FC × EFD | -0.111 | -0.203*** | -0.183*** |
| | (0.069) | (0.067) | (0.049) |
| FC × PCI | 0.068 | 0.104** | 0.102*** |
| | (0.046) | (0.045) | (0.031) |
| PCI × EFD | -0.044 | -0.185*** | 0.015 |
| | (0.064) | (0.058) | (0.083) |
| PCI × FC × EFD | -0.088** | -0.099*** | -0.097*** |
| | (0.037) | (0.038) | (0.026) |
| Emp | | 4.350*** | 1.229*** |
| | | (0.389) | (0.150) |
| CapInt | | -0.340* | -0.163* |
| | | (0.186) | (0.088) |
| Conc | | -3.462*** | -1.407** |
| | | (0.877) | (0.711) |
| Observations | 190,868 | 189,539 | 186,393 |
| R-squared | 0.004 | 0.138 | 0.899 |
| Control variables | No | Yes | Yes |
| Year FE | Yes | Yes | Yes |
| Firm FE | No | No | Yes |

*Notes*: This table displays the means and standard deviations of the firm- and industry-specific variables for the period 2003-2018. Standard errors, in parentheses, are clustered at the firm-level. The variable $V_{ft}$ captures the integration level of firm $f$ at time $t$, derived from AFiD and input-output tables. It is derived by combining input cost shares with a binary variable indicating if $f$ owns plants in both upstream and downstream sectors. $PCI$ is the industry-level product complexity index obtained from the Growth Lab at Harvard University. *EFD* measures external financial dependence and is obtained from Eppinger and Neugebauer (2022), and calculated as the arithmetic mean of the firms' responses in each industry. FC is a binary indicator, taking the value of 1 to denote the occurrence of a financial crisis during the years 2008-2009, and 0 otherwise. From the AFiD data: *Emp* indicates firm size and is computed as the natural logarithm of the number of employees, *CapInt* is the natural logarithm of capital expenditure to total sales at the industry level, and *Conc* measures industry competition and is computed using firm sales according to the Herfindahl index approach. Industries are classified at the four-digit level according to ISIC Rev. 4, except for the *EFD* variable, which is defined at the two-digit industry level. $^{*}$, $^{**}$, and $^{***}$ indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
*Source*: Author's calculations based on AFiD data, OECD input-output tables, Eppinger and Neugebauer (2022), and the Growth Lab at Harvard University database.

continues to show inconsistent effects across specifications. The variable Fin itself remains negative and statistically significant; however, the crisis dummy does not appear to have an additional significant effect. Nevertheless, the triple interaction term continues to exhibit a negative and statistically significant effect in Column (3), even after the inclusion of both fixed effects and control variables.

## 5.2 Plant-level analysis

Although firm-level analysis provides a broad perspective on decision-making and firm-wide factors such as integration, it may overlook important variations that occur at the plant-level. In this section, I shift to a plant-level analysis, which allows me to capture more granular effects that may not be evident

**Table 7:** Financial crisis years 2008-2010: firm-level analysis

| Dep. Var: $V_{ft}$ | (1) | (2) | (3) |
|---|---|---|---|
| PCI | 0.537*** | -0.235** | -0.133 |
| | (0.081) | (0.118) | (0.083) |
| Fin | -1.296*** | -0.912*** | -4.149*** |
| | (0.202) | (0.165) | (0.951) |
| FC × Fin | -0.111 | -0.096 | -0.076 |
| | (0.094) | (0.088) | (0.056) |
| FC × PCI | -0.008 | 0.013 | 0.050 |
| | (0.045) | (0.046) | (0.031) |
| PCI × Fin | -1.033*** | -0.494*** | -0.367** |
| | (0.115) | (0.087) | (0.149) |
| PCI × FC × Fin | -0.038 | -0.039 | -0.067* |
| | (0.062) | (0.058) | (0.038) |
| Emp | | 4.194*** | 1.250*** |
| | | (0.370) | (0.147) |
| CapInt | | -0.035 | -0.044 |
| | | (0.140) | (0.082) |
| Conc | | -2.990*** | -1.005 |
| | | (0.822) | (0.619) |
| Observations | 200,145 | 198,741 | 195,462 |
| R-squared | 0.011 | 0.139 | 0.896 |
| Control variables | No | Yes | Yes |
| Year FE | Yes | Yes | Yes |
| Firm FE | No | No | Yes |

*Notes*: This table displays the means and standard deviations of the firm- and industry-specific variables for the period 2003-2018. Standard errors, in parentheses, are clustered at the firm-level. The variable $V_{ft}$ captures the integration level of firm $f$ at time $t$, derived from AFiD and input-output tables. It is derived by combining input cost shares with a binary variable indicating if $f$ owns plants in both upstream and downstream sectors. $PCI$ is the industry-level product complexity index obtained from the Growth Lab at Harvard University. $Fin$ is the financial dependence ratio and is calculated as the median of the ratio of financial services inputs to total output for the manufacturing sector, using pre-2008 data. FC is a binary indicator, taking the value of 1 to denote the occurrence of a financial crisis during the years 2008-2010, and 0 otherwise. From the AFiD data: *Emp* indicates firm size and is computed as the natural logarithm of the number of employees, *CapInt* is the natural logarithm of capital expenditure to total sales at the industry level, and *Conc* measures industry competition and is computed using firm sales according to the Herfindahl index approach. Industries are classified at the four-digit level according to ISIC Rev. 4, except for the $Fin$ variable, which is defined at the two-digit industry level. $^*$, $^{**}$, and $^{***}$ indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
*Source*: Author's calculations based on AFiD, and OECD input-output table datasets.

at the firm-level. This disaggregation ensures that the observed relationships hold across different levels of analysis.

I merge the variables of interest for this analysis based on the plant's industry classification, using the four-digit level (except for the Fin variable, which is at the two-digit level), as outlined in Section 4. For brevity, I do not present the two separate equations where I estimate the main effects of product complexity and financial frictions individually. In Equation 3.10, the measure of vertical integration is now used at the firm-level in the regression below, as computed in Equation 3.5.

$$\begin{aligned} V_{fkt} = \gamma_0 + \gamma_1\, PCI_{kt} + \gamma_2\, FC_t + \gamma_3\, Fin_k + \gamma_4\, (FC_t \times Fin_k) + \gamma_5\, (FC_t \times PCI_{kt}) \\ + \gamma_6\, (PCI_{kt} \times Fin_k) + \gamma_7\, (\mathbf{PCI}_{kt} \times \mathbf{FC}_t \times \mathbf{Fin}_k) + \gamma_8\, Emp_{ft} + \gamma_9\, CapInt_{kt} \\ + \gamma_{10}\, Conc_{kt} + \delta_f + \delta_k + \delta_t + \epsilon_{fkt}, \end{aligned} \tag{3.10}$$

where $V_{fkt}$ represents the vertical integration of the plant in industry $k$, owned by firm $f$, at time $t$. Moreover, the equation includes plant industry fixed effects ($\delta_k$) at the four-digit industry classification level to control for industry-specific factors at the plant-level. Tables 8, 9, and 10 present the results showing the main effect of contractual frictions, the main effect of financial frictions, and the combined effect of these frictions, respectively.

Table 8 reports the effects of contractual frictions. In Column 1, where no control variables or firm fixed effects are included, the PCI coefficient is positive and statistically significant. This means that a one standard deviation increase in PCI corresponds to a 1.6 unit change in the dependent variable. In Column 2, after adding control variables, and in Column 3, with firm fixed effects, the PCI coefficient remains positive and statistically significant. This confirms that the estimated effect holds even when controlling for firm heterogeneity. Finally, when plant industry fixed effects are introduced in Column 4, the PCI coefficient remains statistically significant but flips to negative sign.

This suggests that as product contracting costs increase, firms tend to engage in more vertical integration, which aligns with the theory that firms vertically integrate to mitigate the hold-up problem by reducing reliance on external suppliers and gaining greater control over the production process. However, when plant industry fixed effects are introduced, the relationship becomes less stable, as indicated by a change in the sign of the coefficient. This suggests that the initial relationship between contracting costs and vertical integration may have been influenced by unobserved industry-specific factors at the plant-level, which are captured by the fixed effects.

Table 9 presents the results for estimating the main effect of financial frictions. In column 1, the coefficient for financial dependence ($Fin$) is $-0.016$, which is statistically significant at the 1% level. This suggests a negative relationship between financial dependence and $V_{fkt}$. However, the coefficient for the interaction term between financial crisis ($FC$) and the financial dependence ratio is not statistically

**Table 8:** Main effects of contractual frictions: plant-level analysis

| Dep. Var: $V_{fkt}$ | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| PCI | 0.016*** | 0.004*** | 0.014** | -0.004* |
| | (0.002) | (0.001) | (0.007) | (0.002) |
| Emp | | 0.068*** | 0.016*** | 0.016*** |
| | | (0.006) | (0.002) | (0.002) |
| CapInt | | -0.001 | 0.002 | -0.005* |
| | | (0.002) | (0.006) | (0.003) |
| Conc | | 0.010 | 0.048* | 0.004 |
| | | (0.010) | (0.026) | (0.004) |
| Observations | 206,000 | 204,540 | 201,200 | 201,200 |
| R-squared | 0.007 | 0.156 | 0.628 | 0.651 |
| Control variables | No | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| Firm FE | No | No | Yes | Yes |
| Plant industry FE | No | No | No | Yes |

*Notes*: This table displays the means and standard deviations of the firm- and industry-specific variables for the period 2003-2018. Standard errors, in parentheses, are clustered at the firm-level. The variable $V_{fkt}$ captures the integration level of firm $f$ in plant $k$ at time $t$, derived from AFiD and input-output tables. It is derived by combining input cost shares with a binary variable indicating if $f$ owns plants in both upstream and downstream sectors. $PCI$ is the industry-level product complexity index obtained from the Growth Lab at Harvard University. From the AFiD data: *Emp* indicates firm size and is computed as the natural logarithm of the number of employees, *CapInt* is the natural logarithm of capital expenditure to total sales at the industry level, and *Conc* measures industry competition and is computed using firm sales according to the Herfindahl index approach. Industries are classified at the four-digit level according to ISIC Rev. 4. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
*Source*: Author's calculations based on the AFiD, OECD input-output tables, and Growth Lab at Harvard University datasets.

significant in this specification as well as in column 2. In column 3, when both control variables and firm fixed effects are included, the interaction term becomes statistically significant. This significance persists in column 4, where plant industry fixed effects are further incorporated. The results suggest that a one-standard-deviation increase in financial dependence is associated with a slight decrease in vertical integration during the crisis (by 0.124, compared to a decrease of 0.123 in non-crisis periods). The results for the main effects of both contractual and financial frictions do not hold across all specifications, which is also observed in Acemoglu et al. (2009) (especially for the contracting costs). However, a clear effect is anticipated in the interaction between these two frictions, as I will demonstrate next.

In Table 10, I report the combined effect of contractual and financial frictions. Both the main effects and the triple interaction effect of these frictions are consistent with the firm-level analysis. More importantly, the three-way interaction term is negative and statistically significant in all specifications. This

**Table 9:** Main effects of financial frictions: plant-level analysis

| Dep. Var: $V_{fkt}$ | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Fin | -0.016*** | -0.008*** | -0.123*** | |
| | (0.002) | (0.001) | (0.038) | |
| FC × Fin | -0.001 | -0.002 | -0.001** | -0.001* |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Emp | | 0.068*** | 0.017*** | 0.016*** |
| | | (0.006) | (0.002) | (0.002) |
| CapInt | | -0.002 | 0.002 | -0.005* |
| | | (0.002) | (0.006) | (0.003) |
| Conc | | 0.009 | 0.052* | 0.004 |
| | | (0.010) | (0.027) | (0.005) |
| Observations | 206,000 | 204,540 | 201,200 | 201,200 |
| R-squared | 0.007 | 0.157 | 0.629 | 0.651 |
| Control variables | No | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| Firm FE | No | No | Yes | Yes |
| Plant industry FE | No | No | No | Yes |

*Notes*: This table displays the means and standard deviations of the firm- and industry-specific variables for the period 2003-2018. Standard errors, in parentheses, are clustered at the firm-level. The variable $V_{fkt}$ captures the integration level of firm $f$ in plant $k$ at time $t$, derived from AFiD and input-output tables. It is derived by combining input cost shares with a binary variable indicating if $f$ owns plants in both upstream and downstream sectors. $Fin$ is the financial dependency ratio and is calculated as the median of the ratio of financial services inputs to total output for the manufacturing sector, using pre-2008 data. FC is a binary indicator, taking the value of 1 to denote the occurrence of a financial crisis during the years 2008-2009, and 0 otherwise. From the AFiD data: *Emp* indicates firm size and is computed as the natural logarithm of the number of employees, *CapInt* is the natural logarithm of capital expenditure to total sales at the industry level, and *Conc* measures industry competition and is computed using firm sales according to the Herfindahl index approach. Industries are classified at the four-digit level according to ISIC Rev. 4, except for the financial dependency ratio variable, which is defined at the two-digit industry level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
*Source*: Author's calculations based on AFiD, and OECD input-output tables datasets.

result suggests that a one standard deviation increase in PCI leads to a slightly more negative effect on vertical integration during periods of financial crisis, particularly when the financial dependence ratio is higher.

## Plant-level robustness checks

Similar to the firm-level analysis, I perform two robustness checks using the EFD measure and defining the financial crisis as occurring between 2008 and 2010. In Table 11, I substitute the financial dependence ratio with the EFD variable. In general, the main effects exhibit similar trends, and the triple

**Table 10:** Interaction effects of contractual and financial frictions: plant-level analysis

| Dep. Var: $V_{fkt}$ | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| PCI | 0.012*** | -0.002 | 0.008* | -0.004* |
| | (0.002) | (0.001) | (0.005) | (0.002) |
| Fin | -0.027*** | -0.020*** | -0.115*** | |
| | (0.003) | (0.002) | (0.035) | |
| FC × Fin | -0.004* | -0.004** | -0.004** | -0.002* |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| FC × PCI | -0.001 | -0.001 | -0.001 | -0.000 |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| PCI × Fin | -0.021*** | -0.013*** | -0.007 | 0.002 |
| | (0.002) | (0.001) | (0.008) | (0.002) |
| PCI × FC × Fin | -0.002* | -0.002** | -0.002** | -0.001** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Emp | | 0.067*** | 0.017*** | 0.016*** |
| | | (0.006) | (0.002) | (0.002) |
| CapInt | | 0.002 | -0.000 | -0.005* |
| | | (0.002) | (0.005) | (0.003) |
| Conc | | 0.015 | 0.052* | 0.005 |
| | | (0.010) | (0.027) | (0.005) |
| Observations | 206,000 | 204,540 | 201,200 | 201,200 |
| R-squared | 0.018 | 0.161 | 0.629 | 0.651 |
| Control variables | No | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| Firm FE | No | No | Yes | Yes |
| Plant industry FE | No | No | No | Yes |

*Notes*: This table displays the means and standard deviations of the firm- and industry-specific variables for the period 2003-2018. Standard errors, in parentheses, are clustered at the firm-level. The variable $V_{fkt}$ captures the integration level of firm $f$ in plant $k$ at time $t$, derived from AFiD and input-output tables. It is derived by combining input cost shares with a binary variable indicating if $f$ owns plants in both upstream and downstream sectors. $PCI$ is the industry-level product complexity index obtained from the Growth Lab at Harvard University. $Fin$ is the financial dependency ratio and is calculated as the median of the ratio of financial services inputs to total output for the manufacturing sector, using pre-2008 data. FC is a binary indicator, taking the value of 1 to denote the occurrence of a financial crisis during the years 2008-2009, and 0 otherwise. From the AFiD data: *Emp* indicates firm size and is computed as the natural logarithm of the number of employees, *CapInt* is the natural logarithm of capital expenditure to total sales at the industry level, and *Conc* measures industry competition and is computed using firm sales according to the Herfindahl index approach. Industries are classified at the four-digit level according to ISIC Rev. 4, except for the financial dependency ratio variable, which is defined at the two-digit industry level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
*Source*: Author's calculations based on the AFiD data, OECD input-output tables, and the Growth Lab at Harvard University datasets.

interaction effect is consistent with the results from the financial dependence ratio. Across all model specifications, the coefficients remain negative and statistically significant. In Table 12, the three-way interaction term is negative and statistically significant, especially after including control variables in column 2 and firm fixed effects in column 3.

**Table 11:** Using the EFD measure: plant-level analysis

| Dep. Var: $V_{fkt}$ | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| PCI | 0.017*** | 0.005*** | 0.012* | -0.004** |
| | (0.002) | (0.001) | (0.007) | (0.002) |
| EFD | -0.002 | 0.004*** | 0.027 | |
| | (0.002) | (0.002) | (0.020) | |
| FC × EFD | -0.002 | -0.003** | -0.003*** | -0.002*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| FC × PCI | -0.000 | 0.001 | 0.001* | 0.001 |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| PCI × EFD | -0.000 | -0.002** | -0.004 | -0.003 |
| | (0.001) | (0.001) | (0.007) | (0.002) |
| PCI × FC × EFD | -0.001** | -0.001** | -0.002*** | -0.001** |
| | (0.001) | (0.001) | (0.001) | (0.000) |
| Emp | | 0.071*** | 0.017*** | 0.017*** |
| | | (0.007) | (0.002) | (0.002) |
| CapInt | | -0.008*** | -0.006 | -0.006** |
| | | (0.002) | (0.005) | (0.003) |
| Conc | | 0.011 | 0.045 | 0.003 |
| | | (0.011) | (0.027) | (0.006) |
| Observations | 196,203 | 194,830 | 191,628 | 191,628 |
| R-squared | 0.007 | 0.160 | 0.638 | 0.661 |
| Control variables | No | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| Firm FE | No | No | Yes | Yes |
| Plant industry FE | No | No | No | Yes |

*Notes*: This table displays the means and standard deviations of the firm- and industry-specific variables for the period 2003-2018. Standard errors, in parentheses, are clustered at the firm-level. The variable $V_{fkt}$ captures the integration level of firm $f$ in plant $k$ at time $t$, derived from AFiD and input-output tables. It is derived by combining input cost shares with a binary variable indicating if $f$ owns plants in both upstream and downstream sectors. $PCI$ is the industry-level product complexity index obtained from the Growth Lab at Harvard University. *EFD* measures external financial dependence and is sourced from Eppinger and Neugebauer (2022), and calculated as the arithmetic mean of the firms' responses in each industry. FC is a binary indicator, taking the value of 1 to denote the occurrence of a financial crisis during the years 2008-2009, and 0 otherwise. From the AFiD data: *Emp* indicates firm size and is computed as the natural logarithm of the number of employees, *CapInt* is the natural logarithm of capital expenditure to total sales at the industry level, and *Conc* measures industry competition and is computed using firm sales according to the Herfindahl index approach. Industries are classified at the four-digit level according to ISIC Rev. 4, except for the *EFD* variable, which is defined at the two-digit industry level. $^{*}$, $^{**}$, and $^{***}$ indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
*Source*: Author's calculations based on AFiD data, OECD input-output tables, Eppinger and Neugebauer (2022), and the Growth Lab at Harvard University database.

**Table 12:** Financial crisis years 2008-2010: plant-level analysis

| Dep. Var: $V_{fkt}$ | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| PCI | 0.013*** | -0.002 | 0.008 | -0.004* |
| | (0.002) | (0.001) | (0.005) | (0.002) |
| Fin | -0.027*** | -0.021*** | -0.115*** | |
| | (0.003) | (0.002) | (0.035) | |
| FC × Fin | -0.005** | -0.004** | -0.004*** | -0.002* |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| FC × PCI | -0.002* | -0.000 | -0.000 | 0.000 |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| PCI × Fin | -0.022*** | -0.013*** | -0.007 | 0.002 |
| | (0.002) | (0.001) | (0.008) | (0.002) |
| PCI × FC × Fin | -0.002 | -0.003** | -0.002** | -0.001 |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Emp | | 0.067*** | 0.017*** | 0.016*** |
| | | (0.006) | (0.002) | (0.002) |
| CapInt | | 0.002 | -0.000 | -0.005* |
| | | (0.002) | (0.005) | (0.003) |
| Conc | | 0.014 | 0.052* | 0.004 |
| | | (0.010) | (0.027) | (0.005) |
| Observations | 206,000 | 204,540 | 201,200 | 201,2 |
| R-squared | 0.018 | 0.161 | 0.629 | 0.651 |
| Control variables | No | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| Firm FE | No | No | Yes | Yes |
| Plant industry FE | No | No | No | Yes |

*Notes*: This table displays the means and standard deviations of the firm- and industry-specific variables for the period 2003-2018. Standard errors, in parentheses, are clustered at the firm-level. The variable $V_{fkt}$ captures the integration level of firm $f$ in plant $k$ at time $t$, derived from AFiD and input-output tables. It is derived by combining input cost shares with a binary variable indicating if $f$ owns plants in both upstream and downstream sectors. $PCI$ is the industry-level product complexity index obtained from the Growth Lab at Harvard University. $Fin$ is the financial dependency ratio and is calculated as the median of the ratio of financial services inputs to total output for the manufacturing sector, using pre-2008 data. FC is a binary indicator, taking the value of 1 to denote the occurrence of a financial crisis during the years 2008-2010, and 0 otherwise. From the AFiD data: *Emp* indicates firm size and is computed as the natural logarithm of the number of employees, *CapInt* is the natural logarithm of capital expenditure to total sales at the industry level, and *Conc* measures industry competition and is computed using firm sales according to the Herfindahl index approach. Industries are classified at the four-digit level according to ISIC Rev. 4, except for the financial dependency ratio variable, which is defined at the two-digit industry level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
*Source*: Author's calculations based on the AFiD data, OECD input-output tables, and the Growth Lab at Harvard University datasets.

## 6 Conclusion

This paper empirically examines the impact of contractual and financial frictions on firm boundaries following the theoretical framework developed by Acemoglu et al. (2009). Using data on German manufacturing firms from 2003 to 2018, it analyzes how financial constraints—particularly during the 2008 financial crisis—affect firms in industries highly dependent on external finance, as well as how product complexity shapes integration decisions. Additionally, it explores the interaction between these frictions in shaping firm organizational structure.

In the empirical analysis, I conduct the study at both firm- and plant-levels to ensure robustness across different levels of aggregation. The theoretical framework reveal an ambiguous effect of contractual frictions on vertical integration. This is supported empirically as product complexity does not exhibit a stable or persistent impact on vertical integration. In contrast, financial frictions consistently demonstrate a negative and stable effect across different specifications, in line with the theoretical prediction on the main effect of financial constraints on integration.

Moreover, the interaction effect suggests that the combined impact of product complexity and the presence of a financial crisis results in lower integration, especially for firms in industries highly dependent on external finance. In the plant-level analysis, I observe similar patterns for both the main effects and the interaction effects of these frictions.

These findings highlight that financial frictions reduce vertical integration by limiting firms' ability to absorb the fixed costs associated with integration, with this effect being more pronounced for complex products. This underscores the compounded challenges firms face when managing both contractual and financial constraints, emphasizing the critical role of financial resources in facilitating integration decisions.

## Bibliography

**Acemoglu, Daron, Pol Antràs, and Elhanan Helpman**, "Contracts and technology adoption," *American Economic Review*, 2007, *97* (3), 916–943. Publisher: American Economic Association.

__ **, Simon Johnson, and Todd Mitton**, "Determinants of Vertical Integration: Financial Development and Contracting Costs," *The Journal of Finance*, 2009, *64* (3), 1251–1290. Publisher: [American Finance Association, Wiley].

**Alfaro, Laura and Maggie Xiaoyang Chen**, "Surviving the Global Financial Crisis: Foreign Ownership and Establishment Performance," *American Economic Journal: Economic Policy*, 2012, *4* (3), 30–55. Publisher: American Economic Association.

__ **, Davin Chor, Pol Antràs, and Paola Conconi**, "Internalizing Global Value Chains: A Firm-Level Analysis," *Journal of Political Economy*, April 2019. Publisher: The University of Chicago Press-Chicago, IL.

__ **, Paola Conconi, Harald Fadinger, and Andrew F. Newman**, "Do Prices Determine Vertical Integration?," *The Review of Economic Studies*, July 2016, *83* (3), 855–888. Publisher: Oxford Academic.

**Alquist, Ron, Nicolas Berman, Rahul Mukherjee, and Linda L. Tesar**, "Financial constraints, institutions, and foreign ownership," *Journal of International Economics*, May 2019, *118*, 63–83.

**Antràs, Pol**, "Firms, contracts, and trade structure," *The Quarterly Journal of Economics*, 2003, *118* (4), 1375–1418. Publisher: MIT Press.

__ , "Incomplete Contracts and the Product Cycle," *The American Economic Review*, 2005, *95* (4), 1054–1073. Publisher: American Economic Association.

__ **and Davin Chor**, "Organizing the global value chain," *Econometrica*, 2013, *81* (6), 2127–2204. Publisher: Wiley Online Library.

__ **and Elhanan Helpman**, "Global Sourcing," *Journal of Political Economy*, June 2004, *112* (3), 552–580. Publisher: The University of Chicago Press.

__ **and** __ , *The Organization of Firms in a Global Economy*, Cambridge, MA: Harvard University Press, 2008. Pages: 9-54.

**Basco, Sergi**, "Financial development and the product cycle," *Journal of Economic Behavior & Organization*, October 2013, *94*, 295–313.

**Bloom, Nicholas, Raffaella Sadun, and John Van Reenen**, “Americans Do IT Better: US Multinationals and the Productivity Miracle,” *American Economic Review*, February 2012, *102* (1), 167–201.

**Carluccio, Juan and Thibault Fally**, “Global Sourcing Under Imperfect Capital Markets,” *The Review of Economics and Statistics*, 2012, *94* (3), 740–763. Publisher: The MIT Press.

**Choi, Jaerim**, “The global value chain under imperfect capital markets,” *The World Economy*, 2020, *43* (2), 484–505. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/twec.12897.

**Corcos, Gregory, Delphine M. Irac, Giordano Mion, and Thierry Verdier**, “The Determinants of Intrafirm Trade: Evidence from French Firms,” *The Review of Economics and Statistics*, 2013, *95* (3), 825–838. Publisher: The MIT Press.

**Deardorff, Alan V.**, “Financial Crisis, Trade, and Fragmentation,” in Leonard K. Cheng and Henryk Kierzkowski, eds., *Global Production and Trade in East Asia*, Boston, MA: Springer US, 2001, pp. 9–31.

**Dierks, Steffi, Alexander Schiersch, and Jan Stede**, “Industry Conversion Tables for German Firm-Level Data,” *Jahrbücher für Nationalökonomie und Statistik*, October 2020, *240* (5), 677–690. Publisher: De Gruyter Oldenbourg.

**Eppinger, Peter and Bohdan Kukharskyy**, “Contracting institutions and firm integration around the world,” *European Economic Review*, August 2021, *137*, 103815.

**Eppinger, Peter S. and Katja Neugebauer**, “External financial dependence and firms’ crisis performance across Europe,” *Empirical Economics*, February 2022, *62* (2), 887–904.

**Erel, Isil, Yeejin Jang, and Michael S. Weisbach**, “Do Acquisitions Relieve Target Firms’ Financial Constraints?,” *The Journal of Finance*, 2015, *70* (1), 289–328. Publisher: [American Finance Association, Wiley].

**Fan, Joseph PH and Larry HP Lang**, “The measurement of relatedness: An application to corporate diversification,” *The Journal of Business*, 2000, *73* (4), 629–660. Publisher: JSTOR.

**Forbes, Silke Januszewski and Mara Lederman**, “Adaptation and Vertical Integration in the Airline Industry,” *American Economic Review*, December 2009, *99* (5), 1831–1849.

**Gibbons, Robert**, “Four formal(izable) theories of the firm?,” *Journal of Economic Behavior & Organization*, October 2005, *58* (2), 200–245.

**Grossman, Sanford J. and Oliver D. Hart**, "The costs and benefits of ownership: A theory of vertical and lateral integration," *Journal of political economy*, 1986, *94* (4), 691–719. Publisher: The University of Chicago Press.

**Görg, Holger and Erasmus Kersting**, "Vertical integration and supplier finance," *Canadian Journal of Economics/Revue canadienne d'économique*, 2017, *50* (1), 273–305. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/caje.12257.

**Hart, Oliver and John Moore**, "Property Rights and the Nature of the Firm," *Journal of political economy*, 1990, *98* (6), 1119–1158. Publisher: The University of Chicago Press.

**Levchenko, Andrei A.**, "Institutional Quality and International Trade," *The Review of Economic Studies*, 2007, *74* (3), 791–819. Publisher: [Oxford University Press, Review of Economic Studies, Ltd.].

**Macchiavello, Rocco**, "Financial Development and Vertical Integration: Theory and Evidence," *Journal of the European Economic Association*, 2012, *10* (2), 255–289. Publisher: Oxford University Press.

**Minetti, Raoul, Pierluigi Murro, Zeno Rotondi, and Susan Chun Zhu**, "Financial Constraints, Firms' Supply Chains, and Internationalization," *Journal of the European Economic Association*, April 2019, *17* (2), 327–375.

**Nunn, N.**, "Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade," *The Quarterly Journal of Economics*, May 2007, *122* (2), 569–600.

**Nunn, Nathan and Daniel Trefler**, "Incomplete contracts and the boundaries of the multinational firm," *Journal of Economic Behavior & Organization*, October 2013, *94*, 330–344.

__ **and** __ , "Chapter 5 - Domestic Institutions as a Source of Comparative Advantage," in Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, eds., *Handbook of International Economics*, Vol. 4 of *Handbook of International Economics*, Elsevier, January 2014, pp. 263–315.

**Rajan, Raghuram G. and Luigi Zingales**, "Financial Dependence and Growth," *The American Economic Review*, 1998, *88* (3), 559–586. Publisher: American Economic Association.

**RDC of the Federal Statistical Office and Statistical Offices of the Federal States of Germany**, `https://doi.org/10.21242/42111.2021.00.01.1.1.0` 2021. Own calculations.

__ , `https://doi.org/10.21242/42221.2021.00.01.1.1.0` 2021. Own calculations.

**Shen, Leilei**, “Global sourcing and credit constraints,” *Canadian Journal of Economics/Revue canadienne d’économique*, 2017, *50* (3), 778–803. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/caje.12278.

**Williamson, Oliver E.**, “Markets and hierarchies: analysis and antitrust implications: a study in the economics of internal organization,” *University of Illinois at Urbana-Champaign’s Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*, 1975.

__ , *The economic institutions of capitalism : firms, markets, relational contracting*, paperback ed. ed., London : Collier Macmillan, 1985.

# Appendix

## Descriptive statistics

**Table A1:** Descriptive statistics

| Variable | Mean | Std. Dev. | Obs |
|---|---|---|---|
| $V_{ft}$ | 0.09 | 0.250 | 5655096 |
| PCI | 0.495 | 0.739 | 5655096 |
| Fin | 0.034 | 0.012 | 5655096 |
| EFD | 3.102 | 0.130 | 5353608 |
| Employment | 1322.872 | 10184.140 | 5613264 |
| CapInt | -3.649 | 0.426 | 5655096 |
| Conc | 0.002 | 0.002 | 5655096 |

*Notes*: This table presents the mean, standard deviation, and the number of observations for the variables used in the analysis. The variable $V_{ft}$ captures the integration level of firm $f$ at time $t$, derived from AFiD and input-output tables. It is derived by combining input cost shares with a binary variable indicating if $f$ owns plants in both upstream and downstream sectors. $PCI$ is the industry-level product complexity index obtained from the Growth Lab at Harvard University. $Fin$ is the financial dependency ratio and is calculated as the median of the ratio of financial services inputs to total output for the manufacturing sector, using pre-2008 data. *EFD* measures external financial dependence and is obtained from Eppinger and Neugebauer (2022), and calculated as the arithmetic mean of the firms' responses in each industry. From the AFiD data: *Emp* indicates firm size and is computed as the natural logarithm of the number of employees, *CapInt* is the natural logarithm of capital expenditure to total sales at the industry level, and *Conc* measures industry competition and is computed using firm sales according to the Herfindahl index approach.
*Source*: Author's calculations based on AFiD data, OECD input-output tables, Eppinger and Neugebauer (2022), and the Growth Lab at Harvard University database.

## Included industries

Table 2 presents industries classified according to the two-digit ISIC Rev. 4. Certain industries are omitted due to the absence of a direct correspondence in the conversion tables. Additionally, Coke & Petroleum refineries was identified as an outlier in the external financial dependence variable and excluded from the analysis. The industries considered in the analysis include Tobacco, Wearing apparel, Wood & cork, Leather, Textiles, Food products, Rubber and plastics products, Repair & installation of machinery and equipment, Fabricated metal, Electrical equipment, Chemicals & chemical products, Other transport equipment, Basic metals, Basic pharmaceutical products, Printing & media, Electronic and optical products, Motor vehicles, and Machinery and equipment—resulting in a total of 18 industries when applying the Triple-difference framework. However, in the robustness check using the external financial dependence (EFD) measure, Repair & installation of machinery and equipment lacks a direct match and is therefore excluded.

**Table A2:** Classification and description of industries

| Division | Description |
|---|---|
| 10 | Food products |
| 11 | Beverages |
| 12 | Tobacco |
| 13 | Textiles |
| 14 | Wearing apparel |
| 15 | Leather |
| 16 | Wood & cork |
| 17 | Paper products |
| 18 | Printing & media |
| 19 | Coke & Petroleum refineries |
| 20 | Chemicals & chemical products |
| 21 | Basic pharmaceutical products |
| 22 | Rubber and plastics products |
| 23 | Non-metallic mineral products |
| 24 | Basic metals |
| 25 | Fabricated metal |
| 26 | Electronic and optical products |
| 27 | Electrical equipment |
| 28 | Machinery and equipment |
| 29 | Motor vehicles |
| 30 | Other transport equipment |
| 31 | Furniture |
| 32 | Other manufacturing |
| 33 | Repair & installation |

*Notes*: The table lists industries based on the two-digit ISIC Rev. 4 classification.
*Source*: Author's presentation based on OECD input-output tables.

## External financial dependence (EFD) classification and mapping

In Table A3, I present the industry-level external financial dependence measure (EFD) sorted in descending order. Industries in the original data are classified according to the Statistical Classification of Economic Activities in the European Community, NACE 1.1. I manually match them to their corresponding industries under NACE 2 (the equivalent of the International Standard Industrial Classification, ISIC Rev. 4) classification. The manufacturing industry Repair and installation of machinery and equipment is excluded from the analysis because there is no one-to-one matching. Additionally, to ensure accuracy, I cross-validated this mapping with the final industry classification from the AFiD data, as described in Dierks et al. (2020). The comparison confirmed that all industries were correctly matched.

**Table A3:** Mapping of industries for external financial dependence

| Industry (NACE 1.1) | Industry (NACE 2) | EFD |
|---|---|---|
| Tobacco | Tobacco | 4.0000 |
| Wood products, except furniture | Wood & cork | 3.4386 |
| Other transport equipment | Other transport equipment | 3.4073 |
| Office machinery and computers | Electronic and optical products | 3.3431 |
| Basic metals | Basic metals | 3.2563 |
| Motor vehicles | Motor vehicles | 3.1681 |
| Publishing and printing | Printing & media | 3.1547 |
| Fabricated metal products | Fabricated metal | 3.1429 |
| Furniture | Furniture | 3.1252 |
| Pharmaceuticals | Basic pharmaceutical products | 3.1121 |
| Radio, TV and communication equipment | Electronic and optical products | 3.1101 |
| Food | Food products | 3.0753 |
| Machinery and equipment | Machinery and equipment | 3.0743 |
| Rubber and plastic products | Rubber and plastics products | 2.9824 |
| Pulp, paper and paper products | Paper products | 2.9631 |
| Other manufacturing | Other manufacturing | 2.9420 |
| Electrical machinery and apparatus | Electrical equipment | 2.9376 |
| Non-metalic mineral products | Non-metallic mineral products | 2.8727 |
| Other chemicals and chemical products | Chemicals & chemical products | 2.8719 |
| Beverages | Beverages | 2.8582 |
| Textiles | Textiles | 2.8296 |
| Wearing apparel and fur | Wearing apparel | 2.7240 |
| Leather and footwear | Leather | 2.6854 |
| Refined petroleum products | Coke & Petroleum refineries | 2.5695 |

*Notes*: This table shows the industry-level external financial dependence measure (EFD) sorted in descending order. It displays the industry classification based on the original data (NACE 1.1) along with the manually matched industries under NACE 2.
*Source*: Provided by Eppinger and Neugebauer (2022).

# Chapter 4

# Competition, Markups, and Automation

**Marcel Smolka and Ali Taleb**[1]

University of Flensburg

## Abstract

In this paper, we provide new long-run evidence on market concentration, markups, and automation by focusing on the case of Spain. We present three main insights derived from firm-level data. First, markets targeted by Spanish firms have become significantly more competitive over the period 1990–2018, and the top firms in the manufacturing sector have seen their market shares stagnate over time. Second, output prices and marginal costs of firms have evolved in tandem since 1991, such that the average markup in the manufacturing industry has been rather flat. In contrast to the evidence found for other countries, especially the U.S., markups among the top firms have not been rising over the period 1991–2018. Finally, automation comes with higher average markups, an effect that, as we show, is fully explained by the productivity gains associated with automation. Specifically, we find that automation reduces the firm's marginal cost, which translates into a lower output price, but the price reduction is less than proportional, so that the firm's markup rises.

**Keywords:** Markups, market power, competition, automation, technology, productivity

[1]We gratefully acknowledge funding from the European Union's H2020 research and innovation programme GI-NI under the grant agreement number 101004494. The contents of this publication are the sole responsibility of the GI-NI project consortium and do not necessarily reflect the opinion of the European Union.

# 1 Introduction

Several influential papers have uncovered declining labor shares and rising markups in the US and other advanced economies over the last few decades (De Loecker et al., 2020; De Loecker and Eeckhout, 2021). These observations have been linked to the phenomenon of "superstar firms" triggered by globalization and technology (Autor et al., 2020). As shown by Acemoglu and Restrepo (2018, 2022) and Acemoglu et al. (2020), automation technologies have the potential to widen inequality between both workers and firms. Recently, new evidence on the link between automation technologies and rising concentration and markups has started to emerge (Hubmer and Restrepo, 2024; Firooz et al., 2023).

In this paper we provide new long-run evidence on market concentration, markups, and automation by focusing on the case of Spain. The central insights from our paper are as follows. First, markets targeted by Spanish firms have become significantly more competitive over the period 1990-2018, and the top firms in the manufacturing sector have not been able to capture larger market shares over time. Second, output prices and marginal costs of firms have evolved in tandem since 1991, such that the average markup (price over marginal cost) in the manufacturing industry has been rather flat. Importantly, we do not find evidence for rising markups among the top firms over the period 1991-2018. Finally, automation comes with higher markups at the firm level, an effect explained fully by the productivity gains associated with automation. Specifically, we find empirically that automation reduces the firm's marginal cost, which translates into a lower output price, but the price reduction is less than proportional, so that the firm's markup rises.

To arrive at these results, we conduct a variety of analyses bringing together data from different sources. In particular, our paper is structured in three parts. In the first part, we consult firm-level survey data along with industry-level administrative data in order to provide a broad macro view on the evolution of market shares and market concentration since the 1990s. The picture we provide is different from other types of analyses as the firms in our survey data set (ESEE) are free to define markets on their own terms (by products, customer groups, or other characteristics firms find relevant). To complete the picture we obtain based on the ESEE survey data, we draw in official industry-level data on the evolution of the number of firms, the firm size distribution, and revenues from the Spanish National Statistics Office (INE).

In the second part of the paper, we compute, and explore, changes in firm-level markups over the period 1991-2018. A novelty of our analysis relative to the existing literature is that the ESEE data set allows us to decompose changes in markups into changes in output prices and changes in marginal costs. This is made possible by the fact that output price changes are observed directly in the data, and by exploiting the first-order condition of a variable input in the firm's cost minimization problem

(Hall, 1988). Disentangling price and cost changes seems important in order to understand and pin down the sources of markup changes over time. A focus of this part of the analysis is also on how markups have evolved, not just on average, but also at the top of the markup distribution, to investigate potential "superstar effects".

In the final part of the paper, we zoom in explicitly on the micro-level of our data to investigate the relationship between markups and automation over time. Apart from being able to compute firm-level changes in markups, prices, and marginal costs based on the ESEE survey data, we also observe a direct measure of automation over nearly three decades (viz. whether the firm uses robots in its production process). To the best of our knowledge, our paper is the first to analyse empirically the relationship between automation on the one hand and markups, prices, and marginal costs on the other hand over a long time period capturing the adoption and diffusion of automation technology in the manufacturing industry.

The rest of the paper is organized as follows. In Section 2 we introduce the data that we use and provide descriptive analyses. In Section 3 we explain how we compute markups and we document some patterns in the data. In Section 4 we present our micro-level analysis on automation and markups, prices, and marginal costs. Section 5 concludes.

## 2 Data and descriptive analyses

Our data derive from the Encuesta Sobre Estrategias Empresariales (ESEE). This is an annual survey of roughly 1,900 firms active in the manufacturing sector in Spain.[2] We analyze the data for the period 1990-2018 and are thus able to paint a detailed picture of the evolution of the manufacturing sector and firm activities over almost three decades. The sampling of the data in 1990 had a two-tier structure taking into account the size of firms measured by the number of employees.[3] Specifically, survey questionnaires were sent out to all large firms that had more than 200 employees, and to a sample of small firms that had between 10 and 200 employees. Small firms were chosen based on stratified, proportional and systematic sampling with a random seed, where a firm's industry affiliation and size class were used as stratification variables. The survey distinguishes between 20 different industries, where each industry is given by a set of 3-digit NACE Rev. 2 products. The sampling of the data is done in such a way that the ESEE data are representative of the Spanish manufacturing sector at large when the sampling properties are taken into account. To do so, we use sampling weights that reflect the inverse sampling probability of a firm relative to the population of firms by industry-size stratum in 2010, based on data from the

[2]The ESEE is conducted by the SEPI foundation (Sociedad Estatal de Participaciones Industriales). See `https://www.fundacionsepi.es/investigacion/esee/en/spresentacion.asp` for details on the survey and how to access the data.

[3]Over time, refreshment samples are incorporated so that new firms are added to the survey as other firms exit.

Spanish Instituto Nacional de Estadística (INE).

**Is there evidence for superstar firms in the ESEE survey data?**—In the first part of the analysis, we portray the evolution of market shares and market concentration between 1990 and 2018. To do so, we exploit survey questions on *(i)* firms' market shares in different markets; and *(ii)* the number of competitors and their market shares in these markets. The purpose of this part of the analysis is to see whether there has been a tendency in the data for rising market shares at the top of the sales distribution and an overall decline in the degree of competition with fewer companies and stronger market positions of a few dominant firms ("superstar firms").

Specifically, firms in the survey report their own market shares as well as the market shares of their competitors in their most important markets (up to five markets that together make up at least 50% of the firm's total sales). As an important difference to the use of administrative data, in the ESEE data firms are free to define a market on their own terms, that is, they can define a market along the lines of products, types of customers, or other characteristics they deem relevant.

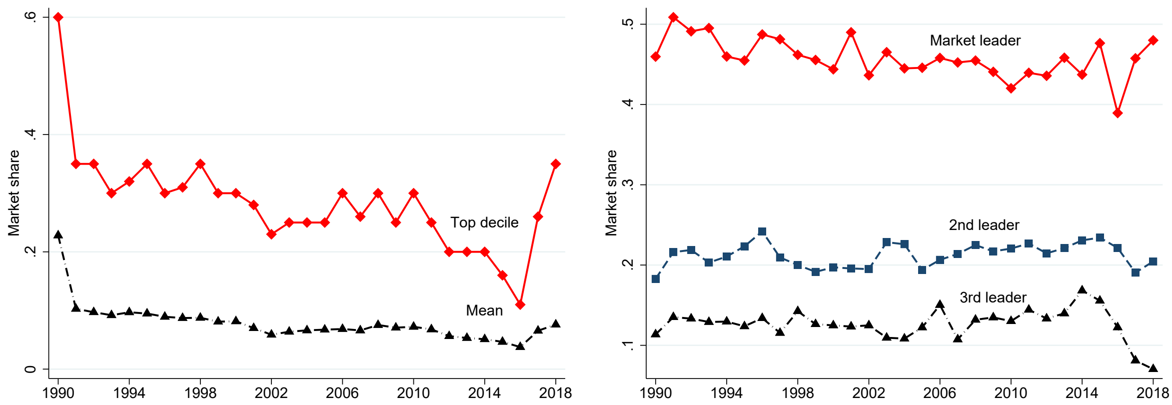
Figure 1a shows that, overall, the average market share of firms in the single most important market they serve has been on a declining trend over the period 1990-2016. Only in the last two years of our sample period, between 2016 and 2018, do we see an increase in the average market share reported by firms. In the beginning of the 1990s the average market share stood at around 10%, and came down to 3.8% in 2016. We observe a similar trend for the top decile of market shares, with a declining trend between 1990 and 2016, and a rebound between 2016 and 2018. To dig deeper into the possibility that the most dominant firms have been able to crowd out their competitors, Figure 1b shows the evolution of market shares depending on the reporting firm's own position in the market (1st, 2nd, or 3rd market leader). That is, the figure focuses exclusively on the top-performing firms in the single most important market reported by the firm. We observe rather stable market shares for these firms over the entire period of analysis, and no indication of rising market shares at the top of the sales distribution. We get a similar impression when looking at the data for the other markets reported by the firms in our sample, that is, the 2nd, 3rd, 4th, and 5th most important markets. Figure A1 in the Appendix shows average market shares and market shares for the top decile of the distribution; Figure A2 shows average market shares for just the market leaders in their respective markets. If anything, the data show that market shares at the top have been declining, rather than rising, over time.

Figure 2 exploits data on the number of competitors in the single most important market reported by firms. Firms report whether there are 10 or less competitors, 11-25 competitors, more than 25 competitors, or whether the market is "atomized" (i.e. competitive).[4] The figure shows that the share of firms reporting a competitive market has been rising from less than 20% in 1991 to more than 50% in 2017

[4]This last option was added to the survey in 1991 and was not available in 1990, which is why the figure omits the relevant data points in 1990.

(and even 80% in 2018). The rise has been particularly steep since 2010. All other categories, including the one with 10 ore less competitors, have been declining over time. We see strikingly similar pictures when looking at the firms' 2nd, 3rd, 4th, and 5th most important markets; see Figure A3 in the Appendix. Overall, we thus observe a clear shift towards more competitive markets in Spanish manufacturing since 1990, and no evidence for superstar firms becoming more dominant players in their markets over time.

**Figure 1:** Evolution of market shares (1990-2018).†




**(a)** Average and top decile

**(b)** Top three market leaders

†*Note:* This figure shows the evolution of market shares as reported by firms in the ESEE data. Panel (a) shows average market shares as well as market shares for the top decile of firms. Panel (b) shows market shares of the top three firms in their respective markets. Sampling weights apply. *Source:* Authors' illustration using ESEE data.

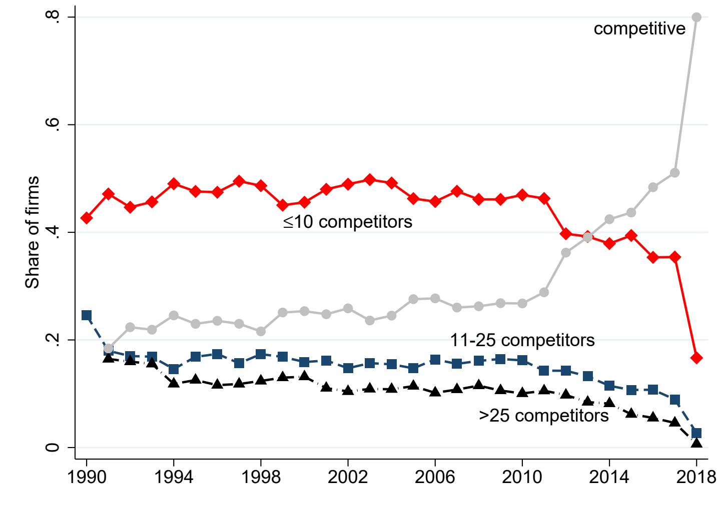
**Is there evidence for superstar firms in administrative data?**—The data we use above are survey data from a sample of manufacturing firms. To see whether the absence of evidence for superstar effects is confirmed in administrative data, we resort to industry-level information available from the Spanish Instituto Nacional de Estadística (INE).

We start by inspecting data on the number of manufacturing firms by industry. These data from the Structural Business Statistics are available for the period 1993-2022.[5] Due to a change in industry classification from NACE rev. 1 to NACE rev. 2, we show the data separately for two periods, viz. 1993-2007 and 2008-2022.[6] In Figure 3a we compute the inverse number of firms by industry and year, and then take the (revenue-weighted) average across industries for each year. The data show a decline in this measure over time. If markets were equally apportioned across firms, this would imply a decline

[5]They can be accessed at `https://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica_C&cid=1254736143952&menu=ultiDatos&idp=1254735576715`.

[6]For the first (second) period, we retain 94 (119) industries for which complete data are available across all years.

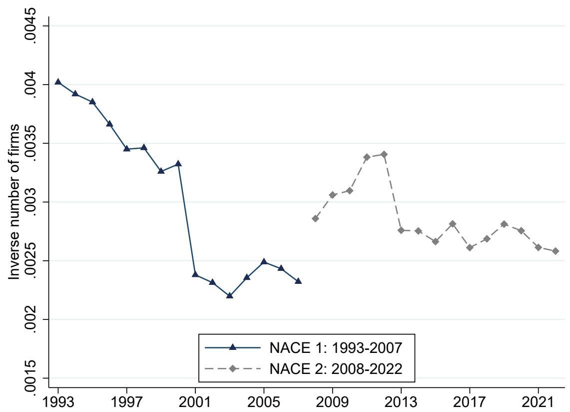
**Figure 2:** Evolution of market concentration (1990-2018).†



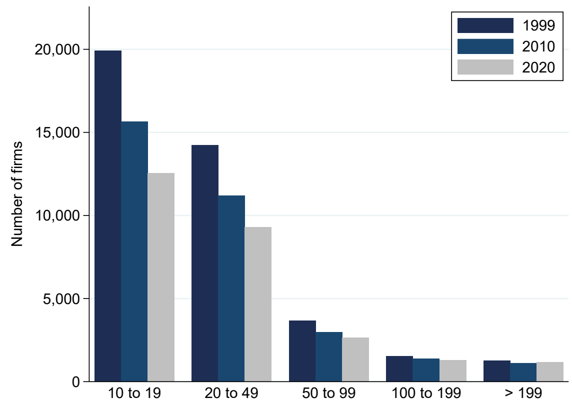

†*Note:* This figure shows the evolution of market concentration for the most important market as reported by firms in the ESEE data. Sampling weights apply. *Source:* Authors' illustration using ESEE data.

in the average market share over time.

**Figure 3:** Average inverse number of firms and firm size distribution.†




**(a)** Average inverse number of firms (1993-2022)




**(b)** Firm size distribution

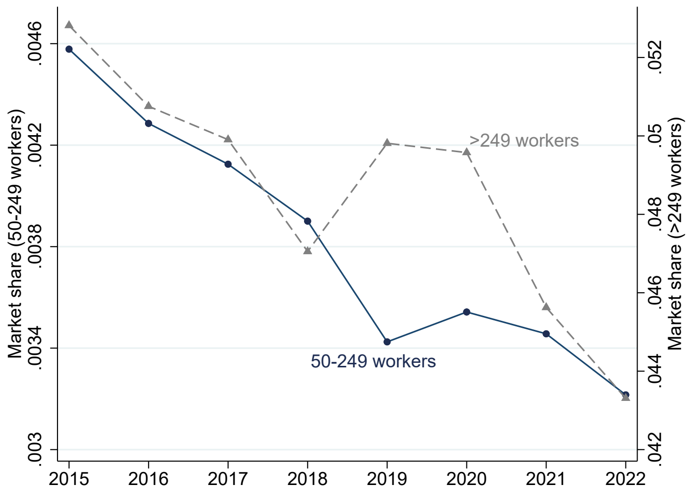
†*Note:* Panel (a) shows the inverse number of firms averaged across 94 NACE rev. 1 industries (1993-2007) and across 119 NACE rev. 2 industries (2008-2022), respectively. Panel (b) shows the number of firms by size group (measured by the number of employees) across all manufacturing industries in 1999, 2010, and 2020, respectively. *Source:* Authors' illustration using INE data.

We next look into the evolution of the firm size distribution in Spanish manufacturing over time.

For this purpose, we use data available since 1999 and compare the firm size distribution in the initial year with the ones in 2010 and 2020, respectively.[7] Firm size groups are defined by the number of employees.[8] Figure 3b shows that in each year the firm size distribution is strongly right-skewed with many small firms and few large firms. Over time, the number of firms has been declining in all size groups except for the biggest firms (those with more than 199 employees) whose number increased between 2010 and 2020 (but was still smaller in 2020 than in 1999).

Unfortunately, we do not have revenue data available by size group for the entire period. This makes it impossible to paint a comprehensive picture of how market shares at the top of the firm size distribution have evolved. However, we have the necessary revenue data available for the period 2015-2022. Since firm-level market shares cannot be computed exactly, we proceed as follows: We compute the revenue share of each size group by industry, and then apportion this share equally across firms falling into the respective size group. We finally compute a (revenue-weighted) average across industries for each size group. Figure 4 shows the evolution of the thus defined market shares for the big firms (50-249 employees) and the biggest firms (more than 249 employees). We find a clear negative trend in market shares for these firms, and no evidence for superstar effects.

**Figure 4:** Average market share by size group (2015-2022).[†]




[†]*Note:* This figure shows the evolution of average market shares by firm size group for the period 2015-2022. *Source:* Authors' illustration using INE data.

[7]These data from the Central Business Register (CBR) can be accessed at `https://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica_C&cid=1254736160707&menu=ultiDatos&idp=1254735576550`.

[8]We discard firms with less than 10 employees and only keep industries that we can consistently map across the three years.

# 3 Computing markups

## 3.1 Framework

Our central object of interest is the firm-level markup (as a measure of market power) defined as $\mu_{it} \equiv P_{it}/\lambda_{it}$, where $P_{it}$ is the output price of firm $i$ at time $t$ and $\lambda_{it}$ is the firm's marginal cost. A markup greater than one means that the firm charges a price that exceeds its marginal cost and the firm has some market power. Specifically, we are interested in the *change* in the markup over time:

$$\hat{\mu}_{it} = \hat{P}_{it} - \hat{\lambda}_{it}, \tag{4.1}$$

where a 'hat' indicates percentage changes. Our data includes explicit information on annual output price changes at the firm level (i.e., $\hat{P}_{it}$), while changes in marginal cost ($\hat{\lambda}_{it}$) are unobserved. To get an expression for $\hat{\lambda}_{it}$ that can be obtained from our data, we derive the *level* of the firm's marginal cost (i.e., $\lambda_{it}$) as the shadow value in the firm's cost minimization problem (Hall, 1988). Specifically, we assume that production of firm $i$ at time $t$ takes place according to the following production function:

$$Q_{it} = Q(L_{it}, M_{it}, K_{it}; \Omega_{it}), \tag{4.2}$$

where $Q_{it}$ is units of output, $L_{it}$, $M_{it}$, and $K_{it}$ are units of labor, intermediates, and capital, respectively, and $\Omega_{it}$ is a vector describing the firm's underlying production technology. This technology vector includes both Hicks-neutral and factor-biased productivity terms, as well as other parameters linked to the use of automation technology, in particular the range of automatable tasks and the complexity of the production process (Acemoglu and Restrepo, 2018). Importantly, *all* technology parameters in $\Omega_{it}$ are allowed to be $it$-specific so that each firm is allowed to have its own technology trajectory in all relevant dimensions.

We denote exogenous input prices corresponding to labor, intermediates, and capital by $p^L$, $p^M$, and $p^K$, respectively. We assume that intermediates $M_{it}$ are *flexible*, that is, they can be adjusted instantaneously at zero cost.[9] In contrast, the labor input $L_{it}$ and the capital input $K_{it}$ are both treated as *fixed* inputs. Hence, the optimal input demand choice of the firm at time $t$ that we derive below is going to be *conditional* on both the labor input and the available capital stock at time $t$, and it will be sufficient to study the firm's cost minimization problem in terms of a static optimization problem. Specifically, to

[9] Alternatively, we can treat both labor and capital as fixed inputs in the production process.

produce a certain level of output $\overline{Q}_{it}$, the firm minimizes its production cost $C_{it}$ by solving

$$\min_{M_{it}} \left\{ C_{it} = p_{it}^{L} L_{it} + p_{it}^{M} M_{it} + p_{it}^{K} K_{it} \right\} \tag{4.3}$$

subject to $Q(L_{it}, M_{it}, K_{it}; \Omega_{it}) \geq \overline{Q}_{it}$.[10] This cost-minimization problem can be re-written as:

$$\min_{M_{it}, \lambda_{it}} \left\{ \mathcal{L}_{it} = p_{it}^{L} L_{it} + p_{it}^{M} M_{it} + p_{it}^{K} K_{it} + \lambda_{it} \left[ \overline{Q}_{it} - Q(\cdot) \right] \right\} \tag{4.4}$$

where $\mathcal{L}_{it}$ is the Lagrangian function and $\lambda_{it}$ is the Lagrange multiplier, respectively. We can then write down the first-order condition for intermediates as follows:

$$\frac{\partial \mathcal{L}_{it}}{\partial M_{it}} = p_{it}^{M} - \lambda_{it} \frac{\partial Q(\cdot)}{\partial M_{it}} \overset{!}{=} 0 \quad \Rightarrow \quad \frac{\partial Q(\cdot)}{\partial M_{it}} \frac{M_{it}}{Q_{it}} = \frac{p_{it}^{M} M_{it}}{\lambda_{it} Q_{it}}, \tag{4.5}$$

where $\lambda_{it}$ is the shadow price in the cost minimization problem, that is, the marginal cost of production $\frac{\partial C_{it}}{\partial Q(\cdot)}$. Note that the left-hand side of the last equation is the output elasticity with respect to intermediates. We denote this variable by $\theta_{it}^{M}$. Optimal firm behavior thus dictates that $\theta_{it}^{M}$ be equal to the respective input share when evaluating output at marginal cost. Let $\phi_{it}^{M}$ denote the total expenses on intermediates. By re-arranging terms, we can then re-write (4.5) as:

$$\hat{\lambda}_{it} = \hat{\phi}_{it}^{M} - \hat{Q}_{it} - \hat{\theta}_{it}^{M}, \tag{4.6}$$

where a 'hat' indicates percentage changes. Equation (4.6) relates changes in the firm's marginal cost to changes in *(i)* the firm's factor bill $\left(\hat{\phi}_{it}^{M}\right)$, *(ii)* the firm's output $\left(\hat{Q}_{it}\right)$, and *(iii)* the firm's output elasticity $\left(\hat{\theta}_{it}^{M}\right)$. From our data, we can compute *(i)* and *(ii)* directly, while *(iii)* is often estimated using a production function approach. For a Cobb-Douglas production function, as is often assumed in the literature (Peters, 2020; Crouzet and Eberly, 2021; Meier and Reinelt, 2024), the output elasticity is constant so that *(iii)* is zero and (4.6) simplifies to $\hat{\lambda}_{it} = \hat{\phi}_{it}^{\ell} - \hat{Q}_{it}$. Note that even under the Cobb-Douglas assumption the literature can usually not identify $\hat{\lambda}_{it}$, as firm-level changes in *real* output (i.e., $\hat{Q}_{it}$) are rarely observed in the data.[11] In contrast, since we observe $\hat{P}_{it}$ along with the value of output $V_{it} \equiv P_{it} Q_{it}$, we can compute $\hat{Q}_{it} = \hat{V}_{it} - \hat{P}_{it}$ for use in (4.6).[12]

In the following, we exploit the rich information in our data to compute the output elasticity directly without having to estimate production functions. Specifically, we compute $\theta_{it}^{M}$ (and thus $\hat{\theta}_{it}^{M}$) directly

[10] The virtue of this approach is that it relies only on cost-minimizing behavior of firms, but does not impose any assumptions on the type of competition the firm faces.

[11] The literature typically uses an industry-level price index to deflate output. This negates firm-level heterogeneity in the evolution of output prices.

[12] Note that $V_{it}$ refers to the value of *production*, as observed in the data, not to the *revenue* of the firm. Note also that, as usual, neither the output price level $P_{it}$ nor the physical units of output $Q_{it}$ are observed in the data.

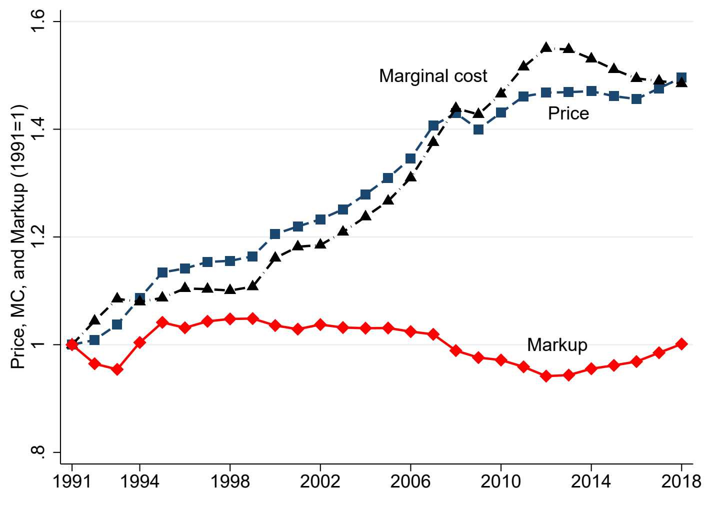
from the data as the firm's material cost share. This requires assuming constant returns to scale, but prevents us from having to estimate production functions, which would come with additional assumptions. Importantly, we can depart from the assumption, customarily made in the literature, that firms in the same industry share the same underlying technology vector (leaving aside between-firm differences in Hicks-neutral productivity). Instead we allow firms to have heterogeneous technology paths in all dimensions, by tracking their changes in cost shares (and thus output elasticities) over time. Moreover, we can track separately the evolution of prices, marginal costs, and markups, which implies that we can attribute changes in the markup to changes in the firm's output price on the one hand, and changes in its marginal cost on the other hand. We can also compute the level of the markup, but not the level of the price and the marginal cost, respectively.

## 3.2 Evolution of prices, marginal costs, and markups

We now turn to a brief exploration of the thus obtained time series on prices, marginal costs, and markups. In Figure 5 we normalize all series to equal one in 1991, and then link the average annual percentage changes through to 2018. We see a pretty strong comovement of rising prices and marginal costs over time, implying a rather flat markup evolution. Price and cost increases were steep in the years leading up to the financial crisis, and more moderate since then. Prices and marginal costs were both about 50% higher (in nominal terms) in 2018 than in 1991, implying that the markup was of the same magnitude at the end of our sample period as it was in the beginning. Overall, there is no evidence for a rise in average markups over the three decades that we cover with our sample.

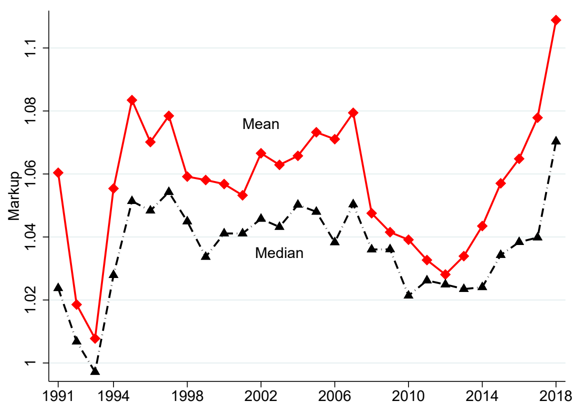
Our data also allow us to compute markup *levels*, even though a firm's price level and its level of marginal cost is unobserved. In Figure 6 we show the evolution of markup levels over the period 1991-2018. To focus first on the average and the typical markup in the manufacturing sector in Spain, panel (a) displays the (sales-weighted) average and median markup levels. We see that both series fluctuate somewhat between 1.0 and 1.1. Overall, it is difficult to discern a clear trend in the data for either average or median markups. Interestingly, the average markup has been increasing since 2012, reaching an all-time high close to 1.12 in the last year of the sample period. However, it had previously been going down for several consecutive years in the wake of the 2008 financial crisis. In panel (b), we depict the evolution of markups at the top of the markup distribution, i.e., the markup levels for the 80th, 90th, and 95th percentiles. As before, there is no clear trend visible in the data. This demonstrates that, over the long period we consider here, markups have not been subject to a rising trend at the very top of the distribution.

**Figure 5:** Evolution of prices, marginal costs, and markups (1991-2018)

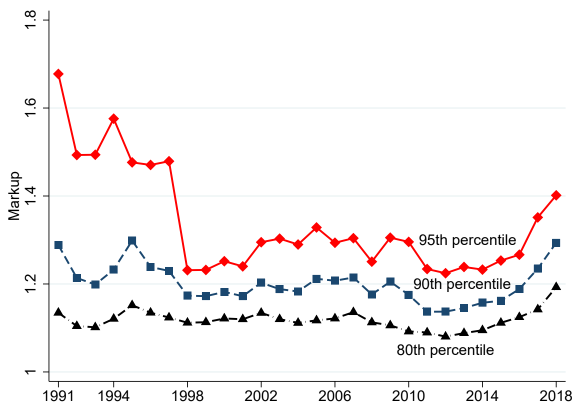



*Note:* The figure shows the evolution of nominal prices, nominal marginal costs, and markups from 1991 to 2018 computed from the ESEE data as described in the text. The price and marginal cost series are computed by linking the average annual percentage changes through time. The price and marginal cost data are winsorized at the top and bottom 1 percent of observations in each year before computing the markup. All three series are sales-weighted and normalized to one in 1991. Sampling weights apply. *Source:* Authors' illustration using ESEE data.

**Figure 6:** Evolution of markup levels (1991-2018).†




**(a)** Average and median markup levels




**(b)** Top markup levels

†*Note:* This figure shows the evolution of markup levels from 1991 to 2018 computed from the ESEE data as described in the text. The data are trimmed at the top and bottom 1 percent of observations in each year and sales-weighted. Sampling weights apply. *Source:* Authors' illustration using ESEE data.

# 4 Markups and automation

We now turn to our micro-level empirical application, whether automated firms have higher markups and whether markups change when firms start automating. This application is made possible by the fact that the ESEE data set includes a direct question on whether firms use robots in their production process or not. For further details on the survey question and some descriptive explorations of the variable, see Koch et al. (2021). In the following, we first discuss cross-sectional results, and then turn to the time series dimension of our data where we explore whether markups change when firms start using robots in their production process.

**Do automated firms have different markups?**—Since we have firm-specific markups, we can use a simple regression framework to investigate the relationship between a firm's markup and its automation status, that is, whether or not it is using robots in its production process. Our focus is the markup *difference* (in percent) between automated and non-automated firms. We run the following regression to explore this issue:

$$\ln \mu_{it} = \alpha + \beta \text{Automated}_{it} + \boldsymbol{X}_{it}^{\top}\gamma + \varepsilon_{it}, \tag{4.7}$$

where $\text{Automated}_{it}$ is a 0/1 indicator for the use of robots[13], $\alpha$ is a constant, $\boldsymbol{X}_{it}^{\top}$ is a row vector of control variables, $\gamma$ is the corresponding column vector of parameters to be estimated, and $\varepsilon_{it}$ is the error term. The central parameter of interest is $\beta$, which captures the relationship between automation and markups and, more precisely, gives the percentage markup premium for automated firms. As controls, we include the firm's input variables in logs (labor, capital, and materials) to account for size differences between firms as well as differences in factor intensities. In addition, we include different sets of fixed effects to take out variation across industries, regions, and years, as well as combinations thereof. This means that we control for markup shocks even if they are specific to industries and regions in Spain. We emphasize that we do not lend a causal interpretation to $\beta$, but rather test whether automated firms have, on average, different markups than non-automated firms.

Table 1 reports the results based on varying sets of fixed effects. We find positive and highly significant estimates of $\beta$ demonstrating significant markup differences between automated and non-automated firms. The estimates imply a markup premium for automated firms in the vicinity of one percent. The differences across the various fixed effect specifications are rather small and indicate that the markup premium cannot be explained by industry- or region-specific trends. We should like to stress that differences in firm size and factor intensities cannot explain the markup premium either, as they are controlled for throughout. The markup premium of one *percent* translates almost one by one into an *absolute*

[13] The raw data reports the variable every four years since 1990. To keep all firm-year observations in the analysis, we interpolate the variable by carrying forward the last value of the variable for three years.

markup difference of 0.01, as we find an average markup around $\mu = 1$ for non-automated firms (captured by the constant $\alpha$ included in (4.7)).

**Table 1:** Pooled sample (OLS)

| | Dependent variable: Markup (in logs) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Automated | 0.0120*** | 0.0119*** | 0.0116*** | 0.00975*** |
| | (0.00296) | (0.00290) | (0.00292) | (0.00307) |
| Industry-year FE | No | Yes | Yes | Nested |
| Region-year FE | No | No | Yes | Nested |
| Industry-region-year FE | No | No | No | Yes |
| Observations | 41460 | 41460 | 41458 | 39859 |
| R-squared | 0.029 | 0.098 | 0.112 | 0.221 |

Note: The dependent variable in all regressions is the firm-level markup (in logs). All regressions include the firm's inputs in logs (labor, capital, and materials) as control variables. Robust standard errors are clustered at the level of the individual firm and are given in parentheses. *,**,*** denote significance at the 10%, 5%, 1% levels, respectively.

A candidate explanation for the level difference in markups found above are productivity differences between firms. Automated firms are, on average, more productive than non-automated firms (Koch et al., 2021), and more productive firms exhibit lower marginal costs and can charge higher markups. To explore this possibility, we augment the model in (4.7) with a firm-level measure of productivity. Specifically, we use the estimation routine developed by Ackerberg et al. (2015) (henceforth ACF) to estimate a flexible output-based translog production function with three inputs (labor, capital, and materials) and robots as an additional input variable. We then run the same set of regressions as before with the firm-level measure of total factor productivity included as an additional control variable. Table 2 reports the results. We find much smaller point estimates of $\beta$ than before, such that the markup premium of automated firms is no longer significantly different from zero. We conclude from this exercise that the markup premium of automated over non-automated firms found in Table 1 indeed reflects productivity differences between firms.

**Markup heterogeneity within automated firms.**—Above we document average markup differences between automated and non-automated firms. However, our data also allow us to discriminate within the group of automated firms, viz. among those that start automating at some point during our sample period, those that stop automating, and those that automate throughout (i.e., those using robots in each and every year they appear in our sample). This opens up another perspective on the relationship between markups and automation that exploits the time dimension in our data. It also allows us to

**Table 2:** Pooled sample controlling for productivity (OLS)

| | Dependent variable: Markup (in logs) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Automated | 0.00368 | 0.00270 | 0.00225 | 0.00122 |
| | (0.00450) | (0.00439) | (0.00432) | (0.00462) |
| Industry-year FE | No | Yes | Yes | Nested |
| Region-year FE | No | No | Yes | Nested |
| Industry-region-year FE | No | No | No | Yes |
| Observations | 16271 | 16271 | 16265 | 15360 |
| R-squared | 0.049 | 0.112 | 0.144 | 0.253 |

Note: The dependent variable in all regressions is the firm-level markup (in logs). All regressions include the firm's inputs in logs (labor, capital, and materials) and total factor productivity in logs as control variables. Robust standard errors are clustered at the level of the individual firm and are given in parentheses. *,**,*** denote significance at the 10%, 5%, 1% levels, respectively.

form expectations about the future evolution of markups within firms, as more and more firms opt to use robots in their production processes over time. We run the following regression, which is similar to the model in (4.7), but sorts firms into different (mutually exclusive) automation categories:

$$\ln \mu_{it} = \alpha + \beta_1 \text{Start}_{it} + \beta_2 \text{Stop}_{it} + \beta_3 \text{Always}_i + \boldsymbol{X}_{it}^{\mathsf{T}} \gamma + \varepsilon_{it}, \tag{4.8}$$

where $\text{Start}_{it}$ is a 0/1 indicator for a firm that uses robots *for the first time* in year $t$, $\text{Stop}_{it}$ is a 0/1 indicator for an automated firm that stops using robots in year $t$, and $\text{Always}_i$ is a time-constant 0/1 indicator for firms that use robots throughout the sample period. Importantly, the variable $\text{Start}_{it}$ is equal to one not just in the first year the firm automates, but also in all subsequent years, and accordingly for $\text{Stop}_{it}$. For convenience, we drop all firms that first switch into and then out of automation (or the other way around).

Table 3 reports the estimation results for the key parameters $\beta_1$, $\beta_2$, and $\beta_3$. As expected, firms automating throughout have a higher average markup than non-automated firms. The specification with the most demanding set of fixed effects in column (4) indicates a markup premium of 1.2 percent. More importantly, firms that start automating also show a markup increase relative to non-automated firms. Their markup premium stands at around 1.5 percent (column (4)). Firms that stop automating exhibit a negative markup premium, but the estimates are not different from zero (in a statistical sense) when looking at column (4) with the full set of fixed effects. These are novel results that speak to the markup dynamics when firms switch from a non-automated production process to an automated one. Importantly, when running the same set of regressions with firm-level productivity (based on ACF) included as a control variable, we do not find significant markup differences for any of the different automation categories; see Table 4. This highlights, again, that productivity is the main driving force behind the markup differences that we find.

**Table 3:** Start automating — Pooled sample (OLS)

| | Dependent variable: Markup (in logs) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Start | 0.0183*** | 0.0164*** | 0.0161*** | 0.0147*** |
| | (0.00503) | (0.00490) | (0.00492) | (0.00550) |
| Stop | -0.0153* | -0.0167** | -0.0177** | -0.0136 |
| | (0.00790) | (0.00762) | (0.00753) | (0.00877) |
| Always | 0.0142*** | 0.0166*** | 0.0161*** | 0.0121** |
| | (0.00519) | (0.00528) | (0.00529) | (0.00562) |
| Industry-year FE | No | Yes | Yes | Nested |
| Region-year FE | No | No | Yes | Nested |
| Industry-region-year FE | No | No | No | Yes |
| Observations | 32741 | 32741 | 32739 | 31012 |
| R-squared | 0.026 | 0.093 | 0.109 | 0.235 |

Note: The dependent variable in all regressions is the firm-level markup (in logs). All regressions include the firm's inputs in logs (labor, capital, and materials) as control variables. Robust standard errors are clustered at the level of the individual firm and are given in parentheses. *,**,*** denote significance at the 10%, 5%, 1% levels, respectively.

**Differential effects of automation on prices and marginal costs.**—We have so far analyzed the relationship between automation and markup *levels*. We have documented novel facts about markup differences between automated and non-automated firms. However, a key virtue of our framework is that we can also disentangle the effects of automation on output prices and marginal costs. We can do this by focusing on *changes* in markups, prices, and marginal costs over time, as detailed in Section 3.2. This is important, as it helps us in understanding, and pinning down, the sources of the markup differences documented above. To do so, we adapt our model in (4.7) as follows:

$$\hat{\Omega}_{it} = \alpha + \beta \text{Automated}_{it} + \boldsymbol{X}_{it}^{\top}\gamma + \varepsilon_{it}, \tag{4.9}$$

where $\hat{\Omega}_{it}$ is the percentage change in the markup $\mu_{it}$, the output price $P_{it}$, or the marginal costs $\lambda_{it}$. We stress again that neither the price level nor the level of the marginal costs is observed, but that we can compute the percentage changes from our data (as explained in Section 3.2). Important is also the different interpretation of the estimated coefficients relative to the previous regressions. Since we regress

**Table 4:** Start automating — Pooled sample controlling for productivity (OLS)

| | Dependent variable: Markup (in logs) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Start | 0.00698 | 0.00487 | 0.00511 | 0.00555 |
| | (0.00726) | (0.00734) | (0.00745) | (0.00815) |
| Stop | -0.00189 | -0.00277 | -0.00462 | -0.000905 |
| | (0.0104) | (0.0108) | (0.0104) | (0.0113) |
| Always | 0.00858 | 0.00860 | 0.00811 | 0.00421 |
| | (0.00805) | (0.00809) | (0.00812) | (0.00888) |
| Industry-year FE | No | Yes | Yes | Nested |
| Region-year FE | No | No | Yes | Nested |
| Industry-region-year FE | No | No | No | Yes |
| Observations | 12851 | 12851 | 12837 | 11896 |
| R-squared | 0.047 | 0.109 | 0.140 | 0.258 |

Note: The dependent variable in all regressions is the firm-level markup (in logs). All regressions include the firm's inputs in logs (labor, capital, and materials) and total factor productivity in logs as control variables. Robust standard errors are clustered at the level of the individual firm and are given in parentheses. *,**,*** denote significance at the 10%, 5%, 1% levels, respectively.

a growth rate on our automation indicator variable, the coefficient $\beta$ now traces out differences between automated and non-automated firms in how markups, prices, and marginal costs evolve over time rather than differences in levels.

Table 5 reports the results. In all regressions we include the full set of industry-region-year fixed effects as well as the firm's input factor variables in logs (in first differences, to relate changes in outcomes to changes in input use). We find a significant premium in the markup growth rate for automated firms equal to 0.45 percentage points; see column (1). More importantly, we find negative effects of automation on output prices in column (2) and on marginal costs in column (3), implying that output prices and marginal costs both decrease for automated firms relative to non-automated firms. The relative decrease is stronger for the marginal costs, explaining the markup premium. In columns (4) to (6) we also include the ACF-based measure of firm-level total factor productivity in logs (in first differences). As expected from our previous results, the estimated coefficients become smaller and turn insignificant for the markup and the marginal costs.

Finally, we again exploit the time dimension of our data to investigate whether automated firms that start or stop using robots in the production process show similar patterns when it comes to the evolution of markups, prices, and marginal costs. Table 6 runs the same set of regressions as before, but now we include $\text{Start}_{it}$, $\text{Stop}_{it}$, and $\text{Always}_i$ as right-hand side variables. The results can be summarized

**Table 5:** Changes in markups, prices, and marginal costs

| | Dependent variable: Percentage change in the variable given below | | | | | |
|---|---|---|---|---|---|---|
| | Markup (1) | Price (2) | MC (3) | Markup (4) | Price (5) | MC (6) |
| Automated | 0.0045*** | -0.0026*** | -0.0072*** | 0.0003 | -0.0032*** | -0.0035 |
| | (0.0014) | (0.0007) | (0.0015) | (0.0023) | (0.0011) | (0.0025) |
| Ind.-region-year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| ln TFP (first diff.) | No | No | No | Yes | Yes | Yes |
| Observations | 32999 | 32999 | 32999 | 13023 | 13023 | 13023 |
| R-squared | 0.198 | 0.254 | 0.201 | 0.208 | 0.246 | 0.211 |

Note: The dependent variable is the percentage change in the firm-level markup (columns (1) and (4)), output price (columns (2) and (5)), and marginal cost (columns (3) and (6)), respectively. All regressions include the firm's inputs in logs (labor, capital, and materials) and first differences. Robust standard errors are clustered at the level of the individual firm and are given in parentheses. *,**,*** denote significance at the 10%, 5%, 1% levels, respectively.

**Table 6:** Start automating — Changes in markups, prices, and marginal costs

| | Dependent variable: Percentage change in the variable given below | | | | | |
|---|---|---|---|---|---|---|
| | Markup (1) | Price (2) | MC (3) | Markup (4) | Price (5) | MC (6) |
| Start | 0.0036** | -0.0007 | -0.0043** | 0.0040 | 0.0006 | -0.0034 |
| | (0.0017) | (0.0009) | (0.0019) | (0.0027) | (0.0015) | (0.0030) |
| Stop | -0.0000 | 0.0012 | 0.0012 | 0.0060** | 0.0028 | -0.0031 |
| | (0.0017) | (0.0011) | (0.0021) | (0.0028) | (0.0019) | (0.0034) |
| Always | 0.0034* | -0.0029** | -0.0063*** | 0.0003 | -0.0032* | -0.0035 |
| | (0.0020) | (0.0012) | (0.0024) | (0.0029) | (0.0018) | (0.0035) |
| Ind.-region-year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| ln TFP (first diff.) | No | No | No | Yes | Yes | Yes |
| Observations | 32999 | 32999 | 32999 | 13023 | 13023 | 13023 |
| R-squared | 0.198 | 0.254 | 0.201 | 0.208 | 0.246 | 0.211 |

Note: The dependent variable is the percentage change in the firm-level markup (columns (1) and (4)), output price (columns (2) and (5)), and marginal cost (columns (3) and (6)), respectively. All regressions include the firm's inputs in logs (labor, capital, and materials) and first differences. Robust standard errors are clustered at the level of the individual firm and are given in parentheses. *,**,*** denote significance at the 10%, 5%, 1% levels, respectively.

as follows. For firms that use robots throughout, we find a pattern of effects that is consistent with the findings in Table 5: prices decrease, but marginal costs decrease even further, so that markups rise (relative to non-automated firms). For firms that start using robots, the effects are similar, although the price decrease is not significantly different from zero. It thus appears that price reductions materialize with some delay following robot adoption. For firms that stop using robots, we find no effects relative to non-automated firms. Finally, and similar to our previous results, once we control for firm-level productivity, most effects turn insignificant.

## 5 Conclusion

In this paper we have provided new long-run evidence on market concentration, markups, and automation for the interesting case of Spain. The main insights derived from our analysis are as follows. First, firms in Spain find themselves caught up in a tougher and more competitive environment (measured by the number of competing firms). Since this perspective derives from self-reported survey data where firms are free to define the boundaries of the markets they serve, this observation can be explained by a stronger integration of markets across countries (with an increasing total number of competitors in spite of fewer domestic competitors). Importantly, the top firms in the manufacturing sector have seen their market shares stagnate over time. This is consistent with our second finding on rather stable markups for the typical firm in the sample as well as for firms located at the top of the markup distribution. The evidence that we find is, hence, going against some of the evidence found for other countries, especially the U.S., where markups among the top firms have been rising (with significant public and academic concern). Finally, automation comes with higher average markups, an effect that, as we show, is fully explained by the productivity gains associated with automation. Specifically, we find that automation reduces the firm's marginal cost, which translates into a lower output price, but the price reduction is less than proportional, so that the firm's markup rises. This last finding is consistent with less elastic demand and lower passthroughs for more productive firms (Baqaee and Farhi, 2020), an assumption that plays a central role in recent literature.

# Bibliography

**Acemoglu, Daron and Pascual Restrepo**, "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment," *American Economic Review*, June 2018, *108* (6), 1488–1542.

— **and** — , "Tasks, Automation, and the Rise in U.S. Wage Inequality," *Econometrica*, 2022, *90* (5), 1973–2016.

— **, Claire Lelarge, and Pascual Restrepo**, "Competing with robots: Firm-level evidence from france," in "AEA Papers and Proceedings," Vol. 110 2020, pp. 383–88.

**Ackerberg, Daniel A, Kevin Caves, and Garth Frazer**, "Identification properties of recent production function estimators," *Econometrica*, 2015, *83* (6), 2411–2451.

**Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen**, "The fall of the labor share and the rise of superstar firms," *The Quarterly Journal of Economics*, 2020, *135* (2), 645–709.

**Baqaee, David Rezza and Emmanuel Farhi**, "Productivity and Misallocation in General Equilibrium*," *The Quarterly Journal of Economics*, 09 2020, *135* (1), 105–163.

**Crouzet, Nicolas and Janice Eberly**, "Intangibles, markups, and the measurement of productivity growth," *Journal of Monetary Economics*, 2021, *124*, S92–S109.

**De Loecker, Jan and Jan Eeckhout**, "Global market power," Technical Report, KU Leuven 2021.

— **,** — **, and Gabriel Unger**, "The rise of market power and the macroeconomic implications," *The Quarterly Journal of Economics*, 2020, *135* (2), 561–644.

**Firooz, Hamid, Zheng Liu, and Yajie Wang**, "Automation and the Rise of Superstar Firms," Technical Report 2023.

**Hall, Robert E**, "The relation between price and marginal cost in US industry," *Journal of Political Economy*, 1988, *96* (5), 921–947.

**Hubmer, Joachim and Pascal Restrepo**, "Not a Typical Firm: Capital-Labor Substitution and Firms' Labor Shares," Technical Report 2024.

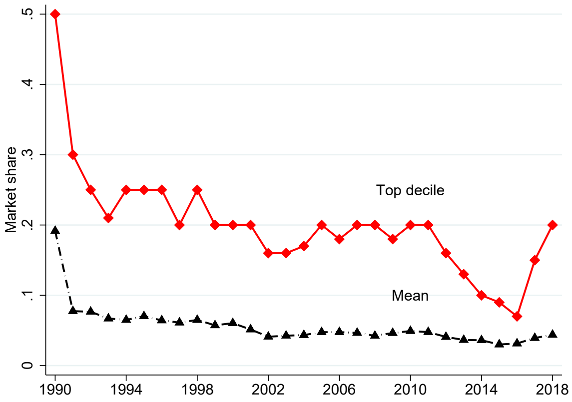
**Koch, Michael, Ilya Manuylov, and Marcel Smolka**, "Robots and Firms," *The Economic Journal*, 2021, *131* (638), 2553–2584.

**Meier, Matthias and Timo Reinelt**, "Monetary Policy, Markup Dispersion, and Aggregate TFP," *The Review of Economics and Statistics*, 07 2024, *106* (4), 1012–1027.

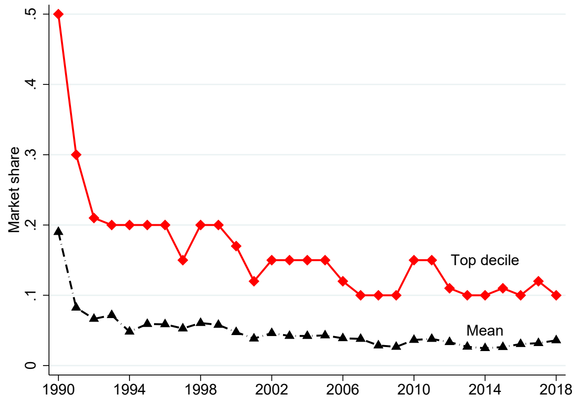
**Peters, Michael**, "Heterogeneous Markups, Growth, and Endogenous Misallocation," *Econometrica*, 2020, *88* (5), 2037–2073.

# Appendix

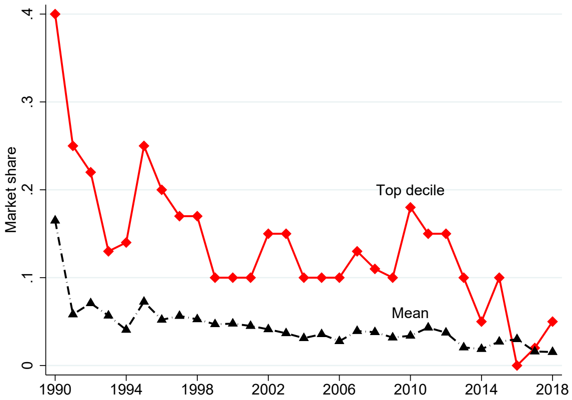
**Figure A1:** Evolution of market shares (1990-2018).†



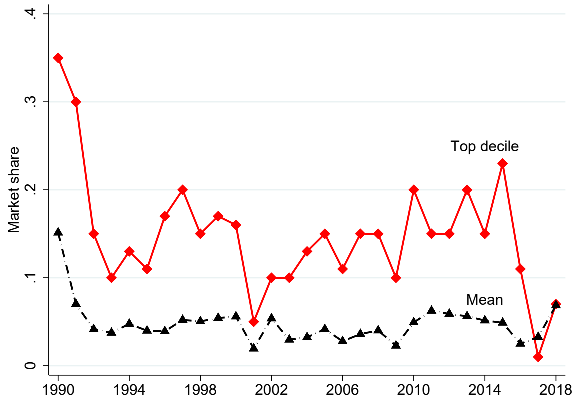

**(a)** 2nd market




**(b)** 3rd market



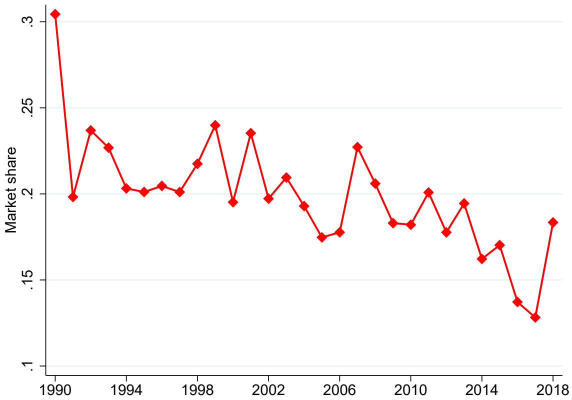

**(c)** 4th market



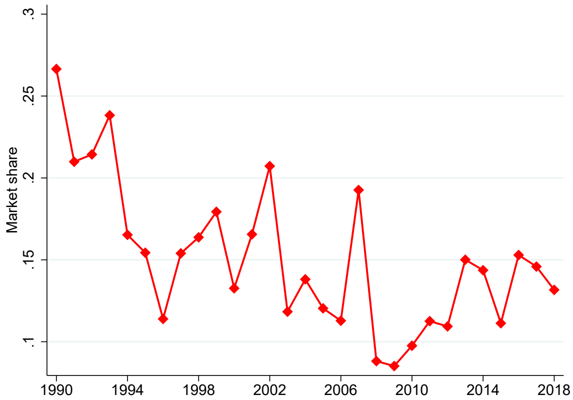

**(d)** 5th market

†*Note:* This figure shows the evolution of market shares as reported by firms in the ESEE data. Panels (a) to (d) refer to the firms' 2nd, 3rd, 4th, and 5th most important markets, respectively. Sampling weights apply. *Source:* Authors' illustration using ESEE data.

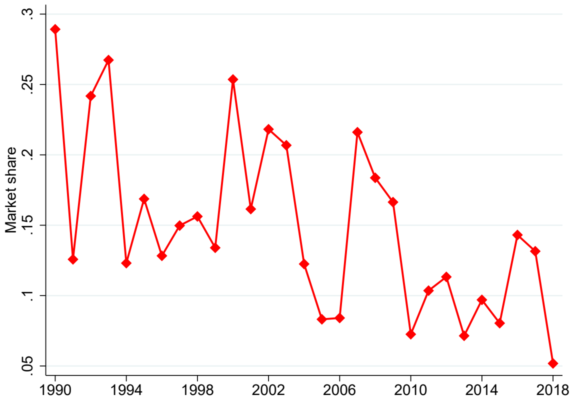
**Figure A2:** Evolution of market shares of market leaders (1990-2018).†



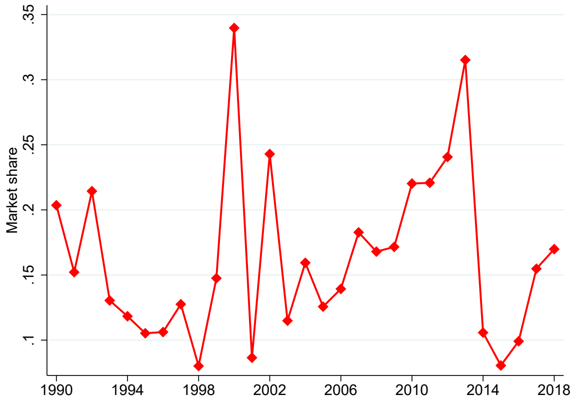

**(a)** 2nd market




**(b)** 3rd market



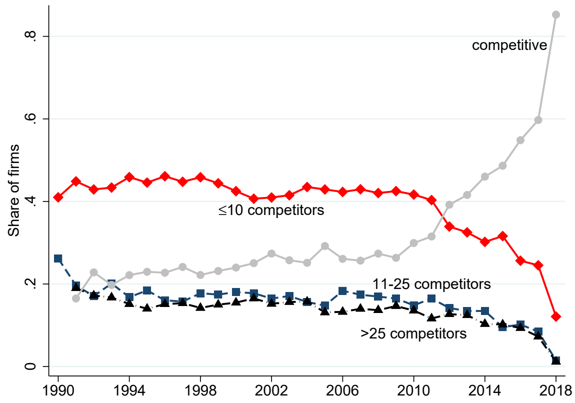

**(c)** 4th market



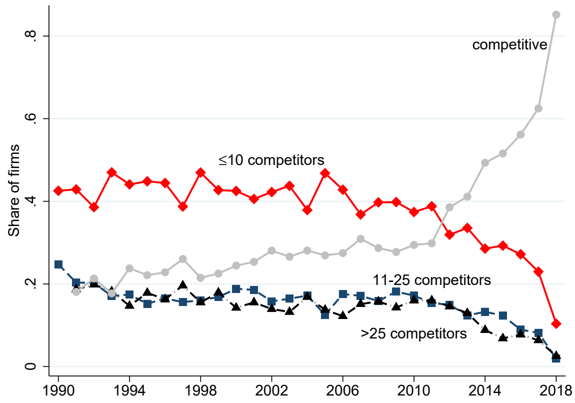

**(d)** 5th market

†*Note:* This figure shows the evolution of market shares as reported by firms in the ESEE data. All firms report to be the market leaders in the respective markets. Panels (a) to (d) refer to the firms' 2nd, 3rd, 4th, and 5th most important markets, respectively. Sampling weights apply. *Source:* Authors' illustration using ESEE data.

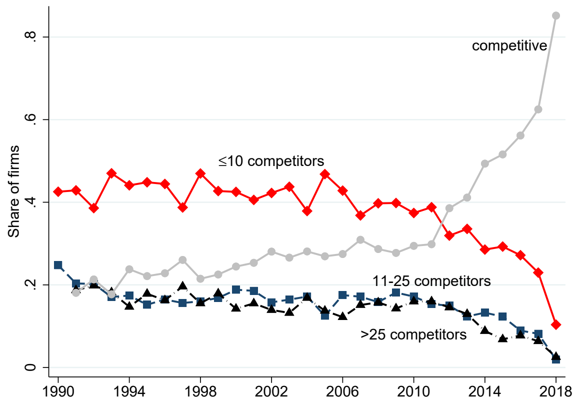
**Figure A3:** Evolution of market concentration (1990-2018).[†]



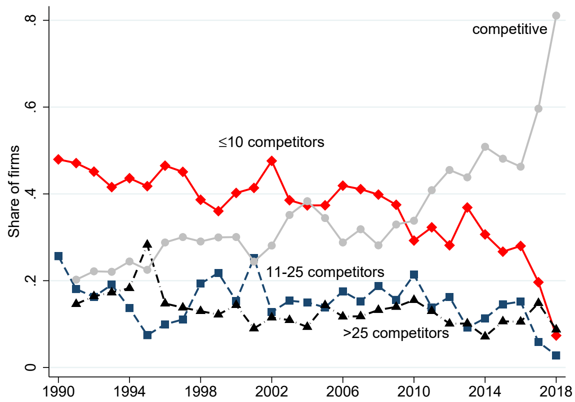

**(a)** 2nd market




**(b)** 3rd market




**(c)** 4th market




**(d)** 5th market

[†] *Note:* This figure shows the evolution of market concentration as reported by firms in the ESEE data. Panels (a) to (d) refer to the firms' 2nd, 3rd, 4th, and 5th most important markets, respectively. Sampling weights apply. *Source:* Authors' illustration using ESEE data.

# Chapter 5

# Conclusion

This thesis examines the determinants of firm boundaries and market power. It focuses on the role of technology, contractual frictions, and financial constraints in shaping firms' organizational structures, and on how automation affects market power. By analyzing vertical integration decisions and the implications of automation on firm markups, the research contributes to the literature of industrial organization.

The first chapter explores how domestic and foreign input cost shares, along with technological intensity, influence firm vertical integration decisions. The empirical findings suggest that a higher domestic input cost share increases the likelihood of vertical integration when the manufacturer's R&D intensity surpasses that of the supplier. Conversely, the presence of foreign suppliers alters the traditional trade-off between vertical integration and outsourcing, underscoring the importance of considering both domestic and international sourcing strategies.

The second chapter examines the impact of contractual and financial frictions on vertical integration. It highlights the crucial role of financial constraints in limiting firms' ability to integrate, particularly during periods of financial distress. The analysis provides causal evidence using a Triple-difference approach, showing that industries heavily reliant on external finance are more affected by financial crises, reducing their propensity to integrate vertically.

The third chapter investigates the relationship between automation, pricing strategies, and market power. We find that, first, between 1990 and 2018, markets served by Spanish firms became more competitive, with leading firms in the manufacturing sector seeing no growth in their market shares. Second, output prices and marginal costs have evolved similarly since 1991, leading to stable average markups in the industry. This contrasts with rising markups seen among top firms in countries like the U.S., where globalization and technological progress have strengthened market power. Finally, automation is linked to higher markups, driven entirely by productivity gains, as it reduces marginal costs but leads to a less than proportional drop in output prices.

Finally, it is important to acknowledge several data limitations across the three papers in this thesis. In the first chapter, a key limitation is the inability to precisely identify whether plants operate in upstream or downstream activities. A more detailed classification would improve the investigation of vertical integration by allowing for a clearer distinction between manufacturers' and suppliers' R&D intensity. Additionally, the study is constrained by data limitations that prevent examining both forward and backward integration simultaneously.

In the second chapter, a more disaggregated industry-level input-output table would enhance the measurement of both vertical integration and financial dependence. Currently, financial dependence is constructed at the two-digit industry level, which may obscure within-industry heterogeneity. More detailed industry classifications could provide a finer understanding of the financial constraints faced by firms. In the third chapter, a limitation arises from the lack of data on advanced technologies such as artificial intelligence and digitalization over the entire study period. Furthermore, while the dataset captures whether a firm uses robots, it does not provide information on the intensity of robot usage, limiting the ability to assess the extent of automation.

Future research could focus on how emerging technologies, particularly artificial intelligence (AI), automation, and digitalization, influence vertical integration decisions within firms. As these technologies continue to evolve, they offer firms new ways to streamline operations, reduce costs, and improve efficiency. This raises important questions about how advancements in automation and AI might encourage firms to reorganize their supply chains, either by integrating more vertically or by relying on external suppliers. Specifically, automation could reduce the need for manual labor in certain processes, prompting firms to either expand control over those processes (vertical integration) or outsource them more effectively using advanced technology. Additionally, AI's role in optimizing decision-making could lead firms to reevaluate their internal capabilities versus the benefits of external partnerships. This research could offer valuable insights into how new technologies not only change how firms operate internally but also reshape the strategic decisions regarding the scope and boundaries of firms in an increasingly automated and AI-driven world.