Universitat Pompeu Fabra
Departament d'Economia i Empresa

DOCTORAT EN ECONOMIA, FINANCES I EMPRESA

Essays in Firm Dynamics and Misallocation

Tesi Doctoral de:
Sampreet Singh Goraya

Directors de la tesi:
Julian di Giovanni
Manuel García Santana

TESI DOCTORAL UPF / 2021

ACKNOWLEDGEMENTS

I am deeply indebted to my advisors Julian di Giovanni, Manuel García Santana and co-advisor Alessandro Tarozzi for their guidance, encouragement, and patience throughout this research project. I would also like to thank Andrea Caggese and Edouard Schaal for all their help, especially throughout the job market.

I was fortunate to learn from Davide Debortoli, David Nagy, Giacomo Ponzetto, Jan Eeckhout, Jaume Ventura and all the participants of the CREI Macroeconomics Lunch, and the CREI International Lunch for their helpful comments and suggestions. Marta Araque and Laura Agustí provided amazing support throughout the PhD, especially throughout the job market. I thank the Spanish Ministry of Economy for the generous financial support throughout the studies.

I am indebted to my mother Krishna and father Sardool, my brother Gurpreet, my sister Monika, my sisters and brother in law Nancy, Raman, and Santokh. A special thanks to my cousin Lali who stood by me throughout all these years.

I would like to thank my friends in Amritsar who has been a source of inspiration and support, Abhishek, Rohit and Tarun. The PhD would have been a lonely experience without companions to embark on this long journey. Juan and Philipp, you were the best office mates. A special thanks to Akhil, Laurenz, Alaitz, Chiavari, Fabiani, Willy, Jura, Sofia, Elisa and Sonja for all the amazing memories.

SAMPREET SINGH GORAYA
Barcelona
June 2021

Abstract

This dissertation consists of three essays that investigate the drivers of misallocation of resources and its macroeconomic implications. The first chapter highlights the importance of the caste system in India in explaining capital misallocation. Using a quantitative model, it finds that the difference in capital allocation across castes is explained by differences in access to credit and that such asymmetries reduce aggregate TFP by 6% to 10%. The second chapter studies the properties of intangible capital, which includes investments in research and development, and intellectual property product and how they influence the empirical measures of misallocation. It shows that intangible capital entails higher investment adjustment costs than traditional capital. Using a model of firm dynamics, this chapter links the increasing input share of intangible capital to the rising average firm size, dispersion in total factor productivity revenue, and increasing industry concentration. The final chapter provides evidence of the market power of firms in public procurement. Firms that access procurement contracts increase their markups and profits relative to others. Limited competition in public procurement is a potential driver behind these findings and this generates markup dispersion across firms.

RESUM

Aquesta dissertació consta de tres assajos que investiguen els factors que impulsen la mala assignació de recursos i les seves implicacions macroeconòmiques. El primer capítol destaca la importància de el sistema de castes a l'Índia per explicar la mala assignació de capital. Utilitzant un model quantitatiu, es troba que la diferència en l'assignació de capital entre les castes s'explica per les diferències en l'accés a l'crèdit i que aquestes asimetries redueixen la TFP agregada entre un 6 % i un 10 %. El segon capítol estudia les propietats de l'capital intangible, que inclou inversions en recerca i desenvolupament, i productes de propietat intel·lectual i com influeix en les mesures empíriques de mala assignació. Mostra que el capital intangible implica majors costos d'ajust de la inversió que el capital tradicional. Utilitzant un model de dinàmica d'empreses, aquest capítol vincula la creixent participació de les entrades de capital intangible amb l'augment de la grandària mitjana de les empreses, la dispersió dels ingressos totals per productivitat dels factors i la creixent concentració de la indústria. El capítol final proporciona evidència de el poder de mercat de les empreses en la contractació pública. Les empreses que accedeixen als contractes d'adquisició augmenten els seus marges i guanys en relació amb altres. La competència limitada en la contractació pública és un motor potencial darrere d'aquestes troballes i genera una dispersió de marges entre les empreses.

RESUMEN

Esta disertación consta de tres ensayos que investigan los factores que impulsan la mala asignación de recursos y sus implicaciones macroeconómicas. El primer capítulo destaca la importancia del sistema de castas en la India para explicar la mala asignación de capital. Utilizando un modelo cuantitativo, se encuentra que la diferencia en la asignación de capital entre las castas se explica por las diferencias en el acceso al crédito y que tales asimetrías reducen la TFP agregada entre un 6 % y un 10 %. El segundo capítulo estudia las propiedades del capital intangible, que incluye inversiones en investigación y desarrollo, y productos de propiedad intelectual y cómo influye en las medidas empíricas de mala asignación. Muestra que el capital intangible implica mayores costos de ajuste de la inversión que el capital tradicional. Utilizando un modelo de dinámica de empresas, este capítulo vincula la creciente participación de los insumos del capital intangible con el aumento del tamaño medio de las empresas, la dispersión de los ingresos totales por productividad de los factores y la creciente concentración de la industria. El capítulo final proporciona evidencia del poder de mercado de las empresas en la contratación pública. Las empresas que acceden a los contratos de adquisición aumentan sus márgenes y ganancias en relación con otras. La competencia limitada en la contratación pública es un motor potencial detrás de estos hallazgos y genera una dispersión de márgenes entre las empresas.

Preface

This dissertation consists of three essays that investigate the drivers of misallocation of resources and its macroeconomic implications. I employ micro-level datasets in conjunction with cutting-edge econometric techniques to document stylized facts and examine underlying mechanisms using quantitative models of producer heterogeneity and firm dynamics.

In the first chapter, I investigate the relative importance of the caste system in explaining resource misallocation in India and quantify its impact on aggregate productivity. I document three main stylized facts. First, firms of historically disadvantaged castes have a higher average revenue product of capital, $arpk$, relative to firms owned by high castes, whereas no significant differences in the average revenue product of labor, $arpl$, exist. Second, across-caste dispersion in $arpk$ is primarily driven by small and young firms. Third, the majority of this dispersion is concentrated in financially underdeveloped regions in India. In a quantitative model of entrepreneurship, I find that the majority of across-caste dispersion in arpk is explained by differences in access to credit and that such asymmetries reduce aggregate total factor productivity by 6% to 10%.

In the second chapter, co-authored with Andrea Chiavari, we study the rise of intangible capital. In the last few decades, intangible investment has increased dramatically, and by 2015 it accounted for more than 30% of aggregate investment. However, we know still little about its importance in the production process and its associated properties. We estimate the firm-level production function finding that intangible capital is an important factor for production, whose share increased from 0.03 to 0.12, at the expense of the labor. We label this phenomenon Intangible Capital Biased Technological Change (IBTC). Further, we provide novel empirical evidence showing that the investment process of intangible capital is associated with higher sunk costs, meaning that it entails higher investment adjustment costs relative to tangible capital. Finally, using a model of firms and investment dynamics, we show that IBTC can jointly explain most of the trends witnessed in the US economy since the 1980s. Specifically, it quantitatively explains the rise in the average firm size and concentration, the changes in aggregate factor shares, the increase in the profit rate, the decline in the tangible capital investment rate, and the decrease in allocative efficiency.

In the third chapter, co-authored with Jurica Zrnc, we provide evidence of market power in public procurement. Firms with ex-ante similar markups witness a 10% increase after winning a public procurement contract relative to firms that sell only to the private sector. Parallel with this increase in markups, profits increase as well, while labor share declines. Limited competition in public procurement is a potential driver behind these findings, given that 35% of contracts are awarded in a single bid procurement procedure. We show that single bid procurement consistently incurs higher prices than expected by the procuring body relative to a multi-bid setting, even for standardized products. Similarly, single bid procedures are also associated with an increase in firm-level markups.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1

HOW DOES CASTE AFFECT ENTREPRENEURSHIP? BIRTH VERSUS WORTH

1.1 INTRODUCTION

A large body of literature has argued that the misallocation of resources explains a substantial fraction of cross-country differences in aggregate productivity (see, e.g., Banerjee and Duflo 2005, Restuccia and Rogerson 2008, Guner et al. 2008, and Hsieh and Klenow 2009a).[1] A number of market-oriented distortions, such as financial frictions, labor market regulation and size-dependent policies, among others, have been proposed as being responsible for resource misallocation. However, we lack systematic evidence about the quantitative importance of informal institutions, which profoundly shape individuals' economic outcomes in developing countries, in generating aggregate misallocation.

This paper quantifies the effects of one such institution – the caste

[1] See Hall and Jones (1999) and Caselli and Feyrer (2007) for detailed analyses on cross-country productivity differences.

system in India – on aggregate productivity. In particular, I explore the hypothesis that "birth and not worth" – that is, the caste instead of productivity of individuals – determines the way in which resources are allocated in the economy. Historically, the caste system sorted people into different occupations and restrained any mobility, suppressing the entrepreneurial prowess of a vast section of society. While mobility restrictions for dominant castes have weakened over time, the caste system remains a salient feature of India.[2]

I use firm-level data to provide novel empirical evidence that is consistent with the presence of high levels of *caste-driven* resource misallocation. First, I show that the allocation of capital across entrepreneurs is influenced by their caste. In particular, low-caste (LC) and middle-caste (MC) entrepreneurs have a higher average revenue product of capital, $arpk$, relative to high-caste (HC) entrepreneurs, whereas no such dispersion is visible in the average revenue product of labor, $arpl$. Furthermore, the majority of the cross-caste dispersion in $arpk$ is driven by small and young firms and is concentrated in financially underdeveloped regions in India.

Motivated by these facts, I develop a quantitative model of entrepreneurship to evaluate the relative importance of productivity, technology, and access to credit in explaining the cross-caste dispersion in $arpk$. Through the lens of the model, the majority of differences in $arpk$ are explained by stringent borrowing constraints for non-HC individuals, whereas asymmetries in technology and productivity play a minor role. I find that raising the borrowing capacity of non-HC firms to the level of HC firms would increase aggregate total factor productivity (TFP) by 6%, whereas eliminating such asymmetries in technology and productivity increases it by an additional 4%. More-

[2]See Munshi (2016). Traditionally, entrepreneurship and financial intermediation belonged to one group called "Vaishyas"; however, these occupations have spilled over to other high castes such as "Brahmins" and "Kshatriyas"; see Damodaran (2008). The high castes represent 35% of the total population.

over, the model allows me to decompose TFP gains along two margins. First, among active entrepreneurs, a reduction in the misallocation of capital due to differential access to credit across firms of different castes is responsible for 75% of the TFP gains. Second, a reduction in the misallocation of talent in the economy, where productive but poor non-HC entrepreneurs can enter while unproductive but wealthy HC entrepreneurs exit, explains the rest.

I test the model predictions by exploiting the heterogeneity in financial development across various states in India. This approach helps me to evaluate how limited access to credit affects the performance of firms owned by non-HC individuals and its overall welfare implications for the non-HC population. My model explains most of the variation in the cross-caste dispersion in $arpk$ across states. I find that, consistent with the model predictions, moving from the least to the most financially developed state in India, $arpk$ differences between LC and HC entrepreneurs decline from $+40\%$ to essentially zero, whereas non-HC households' consumption and asset-holdings increase substantially and converge toward that of HC households.

My empirical analysis exploits data from the Micro, Small and Medium Enterprises (MSME) census of 2006. This dataset provides exhaustive balance sheet information, along with the caste of the enterprise owner and employees, a feature missing in other commonly used firm-level datasets in India.[3] This approach allows me to dissect the data along the caste-dimension and compute various measures of firm performance. Using this dataset, I establish three main stylized facts.

First, within a sector, LC and MC entrepreneurs have 30% and 13% higher $arpk$ relative to HC entrepreneurs, respectively. Moreover, such differences are 6 percentage points higher in rural areas relative to that of urban areas. This evidence is consistent with the fact that, to this day, the caste system is strictly enforced in rural areas, where the majority

[3]The most commonly used datasets are the Annual Survey of Industries and Prowess. They do not include the caste of the enterprise owner.

of the Indian population resides.

Second, I find that most of the cross-caste $dispersion$ in $arpk$ is driven by small and young firms. In particular, moving from the smallest to the largest firm in the economy, $arpk$ for LC firms declines from being 30% higher than that of HC firms to essentially nil. A similar convergence is also documented over firm age; however, in this case, substantial $arpk$ differences remain even for older firms in the sample.

Finally, cross-caste $arpk$ differences negatively correlate with regional financial development. In particular, I construct a credit-to-output ratio for each state and use it as a measure of financial development. The observed differences in $arpk$ across castes fall as the credit-output ratio increases. I observe that LC firms have an $arpk$ that is double the value of HC firms in the states with the lowest financial development such as Bihar, Jharkhand, and Uttar Pradesh, whereas no such differences are observed in states with well-functioning financial markets, such as Maharashtra. More interestingly, $arpk$ in absolute terms declines for non-HC entrepreneurs but remains relatively flat for HC entrepreneurs.

A high $arpk$ for non-HC entrepreneurs is compatible with other facts in the data: these entrepreneurs are relatively poorer (i.e., their financial wealth is limited), more likely to operate in labor-intensive sectors and use conventional means of production, and less likely to be linked with institutional (bank) financing.

To rationalize these facts, I build a quantitative model of entrepreneurship in which agents from different castes can choose to become either entrepreneurs or workers in the context of *caste-dependent technology, productivity, and access to credit*. Moreover, the model allows for intertemporal savings to capture the self-financing channel. The model serves two main purposes. First, it helps me to disentangle the impact of fundamentals such as technology and productivity from that of potential cross-caste heterogeneity in financial frictions on the $arpk$ dispersion. Second, it helps me to evaluate the welfare implications of such asymmetries at both the extensive and intensive margins.

The quantitative predictions crucially depend on the identification of four sets of parameters, namely, the technology that determines the scale of operation, the dispersion and persistence of the productivity distribution, and the degree of financial frictions. I exploit data from multiple sources to precisely estimate all of these parameters. In particular, first, the moments of the income distribution for each caste are used to pin down span-of-control parameters, which determine the scale of operation of firms. Second, dispersion of the productivity distribution in the model is pinned down by employment distribution for each caste. Third, the degree of financial frictions is estimated by matching credit-to-output ratios. Fourth, the persistence of productivity over time, a crucial parameter that controls the efficacy of the self-financing channel (see Midrigan and Xu 2014), is calibrated with the autocorrelation of output. It is important to note that all of these parameters are jointly estimated in the stochastic steady state.[4]

The model estimates substantial differences in access to credit and dispersion in the ability distribution, whereas estimated values for span-of-control parameters and persistence in ability are quite similar across castes. In particular, the model identifies a smaller span-of-control parameter, a less persistent and less dispersed productivity dispersion, and stricter borrowing constraints for non-HC firms relative to those of HC firms. Furthermore, the model can explain around 80% of the value computed in the data for the cross-caste dispersion in $arpk$. In particular, the $arpk$ of LC and MC entrepreneurs is 21% and 13% higher than that of HC entrepreneurs, respectively, and these differences are primarily, around 70%-80%, driven by limited access to credit for non-HC entrepreneurs. The model identifies financial frictions to be 34%

[4]For example, less stringent financial frictions reduce the dispersion in the employment distribution, whereas, the thicker tail of the ability distribution makes it more dispersed. Therefore, a combination of parameters is identified together in the stochastic steady state.

and 23% more stringent for LC and MC entrepreneurs relative to those of HC entrepreneurs, respectively. Such constraints not only lower the borrowing capacity of incumbents but also hinder the entry of non-HC entrepreneurs. As a result, the model estimates a firm ownership rate of 16% and 42% for LC and MC individuals, respectively, and these values are quite close to the ones found in the data (17% for LC and 46% for MC).

The model is able to capture the heterogeneity in the life-cycle dynamics of firms that are owned by different castes. In particular, non-HC entrepreneurs not only enter with a smaller firm size, but also grow slower over time relative to HC entrepreneurs. These differences arise from a combination of two forces: limited borrowing capacity and the small scale of production technology. Therefore, similar to what I find in the data, the size of firms diverges across castes as firms become older. Further, in line with the stylized facts, the model captures a declining trend in the cross-caste dispersion in $arpk$ over firm age and firm size; however, even among the oldest firms, substantial differences remain, owing to the fact that older non-HC firms – that are most likely to be unconstrained – have a high $arpk$ because they use small-scale production technology.

In the data, I document that more financially developed states also distribute credit more efficiently across castes (i.e., cross-caste differences in the credit-to-output ratio decline). In the same spirit, I solve the model with different levels of financial development to replicate different states in India. The model predicts a steeply declining $arpk$ for LC entrepreneurs over states' credit-to-output ratios. Enhanced borrowing capacity causes LC firms to become more capital intensive as the shadow cost of capital declines. Meanwhile, the efficient allocation of capital increases states' output per capita, which in turn increases household consumption and spurs savings. Such improvement in financial markets primarily benefits marginalized individuals (i.e., non-HC agents in the model).

In what follows, I use the model to conduct various counterfactual exercises. First, I allow non-HC entrepreneurs to have a borrowing capacity that is similar to their HC counterparts. The model identifies gains of 6% in aggregate TFP and 8% in income per capita. Second, an additional 4% of TFP gains are realized once I allow non-HC entrepreneurs to have a technology and ability distribution that is similar to that of HC entrepreneurs. Further, I use the model to decompose TFP gains at the extensive and intensive margins. First, the reallocation of capital from unproductive HC entrepreneurs to more productive non-HC entrepreneurs increases the allocative efficiency of the economy; therefore, as a result, the dispersion in $arpk$ declines by 13%. These changes at the intensive margin explain 75% of the TFP gains. Second, the reduction in borrowing constraints induces the entry of more non-HC entrepreneurs. The share of LC enterprises increases from 16% in the benchmark economy to 29%, whereas the share of MC entrepreneurs decreases from 42% to 35% – that is exactly proportional to their respective population weights. Moreover, because of the excess entry of entrepreneurs, demand for capital and labor increases. This implies a 47% higher interest rate, which further led to the exit of unproductive HC firms and explains the rest of the TFP gains. In the end, I conclude that these TFP gains may represent 15% of the overall gains mentioned in Hsieh and Klenow (2009a).[5]

Finally, I provide evidence to rule out alternative explanations that are most likely to predict $arpk$ dispersion across castes but are not directly linked to financial frictions, such as imperfect competition and the heterogeneous output elasticity of capital. I show that markup heterogeneity across castes is unable to account for the $arpk$ dispersion. Further, I compute a quantity-based measure of the average product of

[5]Hsieh and Klenow (2009a) argue that if capital and labor were efficiently allocated in India, then TFP would be around 40%-60% higher in the manufacturing sector. However, it is important to note that the model used in this paper is different from the one used in their paper.

capital, apk, and document that it is even higher for non-HC firms than $arpk$. In the end, I also take into account the variation in the output elasticity of capital and show that it does not explain the majority of the cross-caste $arpk$ dispersion.

Literature review: This paper contributes to the literature on the misallocation of resources. Banerjee and Duflo (2005), De Mel et al. (2008), and Hsieh and Klenow (2009a) document large dispersions in the marginal product of capital across establishments in developing countries.[6] More specifically, a number of papers relate ethnic heterogeneity and misallocation. Hsieh et al. (2019) argue that race-based and gender-based distortions affect the allocation of talent in the US. Erosa et al. (2017) argue that misallocation of talent across occupations has significant aggregate effects on productivity. Hjort (2014) explores the role of ethnic heterogeneity in distorting the allocation of resources within an establishment. Banerjee and Munshi (2004) document inefficiencies in the allocation of capital across communities in the knitted garment industry in Tirupur (India), and Villanger (2015) evaluates the role of the caste system on entrepreneurship in rural Nepal. I contribute to this literature by quantifying the *aggregate* effects of caste-specific misallocation of capital and talent.[7]

This paper also builds on the work of Thorat and Sadana (2009), Iyer et al. (2013) and Deshpande et al. (2013), who document substantial caste differences in entrepreneurship rates, employment, and growth rates in India. I take their analysis one step further and document caste disparities in average products in the MSME sector. Jodhka (2010) reports borrowing constraints as a major obstacle for the low-caste entrepreneurs (self reported by the respondents). Fisman et al. (2017) *provide evidence on the importance of caste match between lender and borrower*

[6]Papers on misallocation in India include Hsieh and Klenow (2014), Garcia-Santana and Pijoan-Mas (2014), and Asturias et al. (2019).

[7]This paper also relates to the long-standing literature that explores the role of ethnic heterogeneity and economic prosperity; (e.g., Easterly and Levine 1997, Alesina and Ferrara 2005, and Montalvo and Reynal-Querol 2005).

for the access to credit. They find that a lender of a certain caste increases credit access and reduces collateral requirements for a borrower of the same caste. In general, it is more likely that an owner of a bank, a bank manager or a loan officer is an HC individual.[8] This implies that LC and MC individuals are more likely to face unfavorable loan conditions. This paper formalizes the idea of caste specific borrowing limits or financial constraints in a parsimonious way.

This paper also contribute to the literature that quantifies the impact of financial frictions on aggregate TFP (see, e.g., Banerjee and Moll 2010, Buera et al. 2011, Buera and Shin 2013, Midrigan and Xu 2014, Hopenhayn 2014, Moll 2014 and Buera et al. 2015). In particular, this paper tries to quantify the role of heterogeneity in the degree of financial frictions faced by different types of entrepreneurs.

The remainder of the paper is organized as follows. Section 1.2 describes the institutional setup, and Section 3.2 describes the data. The stylized facts are documented in Section 1.4. Section 1.5 presents the theoretical framework. Section 1.6 describes the quantification exercise and discusses the main results and Section 1.7 provides an analysis of firm-specific factors, other than financial frictions, that may cause $arpk$ dispersion across castes. I summarize the findings in Section 1.8.

1.2 INSTITUTIONAL SETUP: THE CASTE SYSTEM

The caste system is a form of social stratification that divides people into rigid hierarchical groups based on their occupation. For centuries, caste dictated customary social interaction, exclusion and endogamy.[9] In order of hierarchy, these are the Brahmins (priests and teachers), Kshatriyas (rulers and soldiers), Vaishyas (merchants and traders) and

[8] In their sample, 74% of the lenders belong to high caste.

[9] Bidner and Eswaran (2015) have describe the caste system as a 3,500 year old system within the context of the four principal castes also known as varnas (Deshpande 2010). Figure 1.B.1 in Appendix 1.A provides the caste structure in detail.

the Sudras (laborers and artisans); see figure 1.B.1. Further, there are two additional groups that fall outside the caste system. The first one embodies the group of people traditionally known as Dalits.[10] The second group of people is known as Scheduled Tribes. They have been subject to various forms of discrimination including *barriers to access capital and firm creation*.

For the remainder of the paper, I ignore the micro structure of the caste system and primarily focus on a very broad definition; that is, low-caste individuals are denoted by "LC," which includes the Schedules Castes and Scheduled Tribes; middle-caste individuals are denoted by "MC," which includes the Sudras (also known as Other Backward Castes, OBC, which fall between the traditional upper castes and the lowest), and the high caste is denoted by "HC," which includes the rest.[11] The castes differ in many dimensions; however, I focus on one particular margin: access to credit markets. This paper, using a general equilibrium setting, argues that low-caste individuals face stricter borrowing limits because of imperfect access to credit markets. Tighter borrowing limits could be a result of statistical or taste-based discrimination.[12]

[10]In the Indian constitution, Dalits have fallen under the category of Scheduled Castes since 1947. Scheduled Castes is an officially designated group of historically disadvantaged people.

[11]The rest includes Brahmins, Kshatriyas, and Vaishyas (as well as several religions). Traditionally, the caste system has been part of Hinduism, but in modern India we also find its presence in other religions. Neuman (1981) describes the caste and social stratification among Muslims in India. Jodhka (2004) and Puri (2003) study the caste system in Sikhism. Recently, the Catholic Church also acknowledged the presence of caste based discrimination in their report: Policy of Dalit Empowerment in the Catholic Church in India: An Ethical Imperative to Build Inclusive Communities.

[12]This includes a lack of entrepreneurial networks that could provide trade credit or lack of own-caste non-institutional financial intermediaries.

1.3 DESCRIPTION OF THE DATA

In this paper I use the Economic Census of India (EC) 2005, Micro, Small and Medium Enterprises (MSME) census of 2006-2007, Indian Human Development Survey (IHDS) 2005, Annual Survey of Industries (ASI) 2006 and National Sample Survey (NSS) 2006. Most of the new empirical facts are drawn from the MSME 2006-2007, therefore, I provide its details below, and the details of other datasets are in Appendix 1.A. The main advantage of using the MSME and EC is that they provide the caste of the enterprise owner.[13]

MSME Census: The MSME dataset consists of two parts: a census of registered MSMEs and a survey of unregistered MSMEs.[14] In particular, the dataset provides the geographical information, industry classification, balance sheet variables, and the caste of the owner. There are two measures of capital stock in the data: the original value of investment in plant and machinery, and the market values of fixed assets. The total wage bill includes salaries and wages, allowances, bonuses, and so on. The measure of output is gross value added. The amount of loan outstanding captures all the loans from formal and informal sources, where informal sources include local moneylenders, friends and relatives. There are 1.4 million observations left after the cleaning

[13]MSME and EC are not commonly used as the ASI or the CMIE Prowess. Researchers do not use the Economic Census more frequently primarily because it does not provide balance sheet information of the enterprise and also lacks a panel dimension. The MSME census does provide balance sheet information; however, it omits large firms and does not have a panel dimension either.

[14]Registration under Factories Act 1948-"Registration of manufacturing units is mandatory under Sectors 2m (i) and 2m (ii) of the Factories Act. Section 2m (i) refers to units engaging 10 or more workers and using power whereas 2m (ii) refers to units engaging 20 or more workers and not using power. Besides, some of the State Governments notify certain industrial activities for mandatory registration, although they do not conform to the criteria laid down under Sectors 2m (i) and 2m (ii). Such registrations are done under Section 85 (i) or Section 85 (ii) by the concerned State Governments. Section 85 (i) refers to units engaging less than 10 workers and using power and Section 85 (ii) refers to units engaging less than 20 workers and not using power."

process, which is described in detail in Appendix 1.A. The descriptive statistics are provided in table 1.1.

MSME Synthetic Panel 2005-2007: The MSME census also provides retrospective information on output for the firms that survive upto 2007. This allows me to construct a balanced panel of MSMEs for the three-year period, 2005-2007. This synthetic panel allows me to compute statistics such as the auto correlation of firms' output, which is crucial to pinning down the persistence in the ability distribution; (see Appendix 1.A for more details).

Table 1.1: Summary Statistics- MSME 2006

	HC		MC		LC		Overall
	All	%	All	%	All	%	
Observations (000s)	742	49%	550	40%	145	11%	1437
Employees	2.9	-	2.3	-	1.9	-	2.5
Output (000s)	219	-	104	-	81	-	146
Wage-bill (000s)	79	-	55	-	35	-	61
Capital (000s)	516	-	172	-	115	-	298
Credit (000s)	110	-	42	-	16	-	65
Age	6.7	-	6.1	-	6.7	-	6.5

Notes: Summary statistics for MSME census 2006. Employees is mean employment, capital is mean value of fixed-assets, output is mean value-added and credit is mean amount of outstanding loans. Age is mean years since initial year of production. Percentages indicate percentage of enterprises in a group with respect to all enterprises. Sampling multipliers are applied to compute averages.

1.4 STYLIZED FACTS

This section illustrates the observed differences in the average product of capital, $arpk$, and average product of labor, $arpl$, across castes. In particular, it shows that $arpk$ is substantially different across castes, whereas such differences in $arpl$ are essentially nil. Further, I document that the majority of such differences are found among small and young firms. Finally, this section showcases a remarkable convergence in cross-caste $arpk$ over regional financial development.

Fact 1: arpk is high for LC and MC firms.

In this section, I divide firms by the caste of their owner. The average product of capital and labor for firm i in sector s with owner of caste c is described as

$$arpk_{isc} := \; ln(ARPK_{isc}) = ln(Y_{isc}) - ln(K_{isc}),$$

$$arpl_{isc} := \; ln(ARPL_{isc}) = ln(Y_{isc}) - ln(L_{isc}).$$

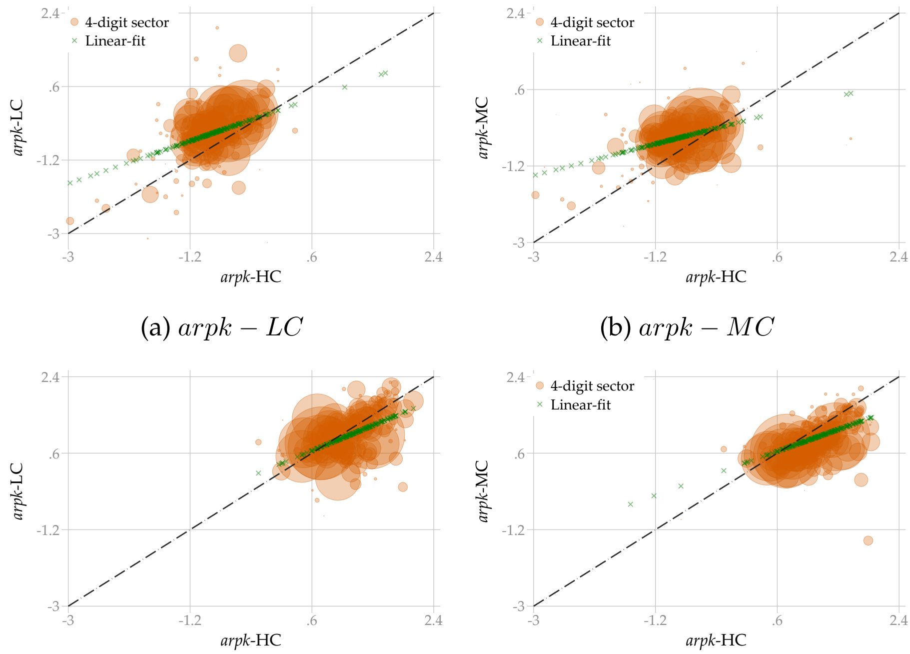
The variable Y_{isc} is gross value added, K_{isc} is capital, and L_{isc} is labor input, measured as wage bill. Employment is an imperfect of labor input because it fails to capture actual hours worked and quality; therefore, similar to Hsieh and Klenow (2009a), I use the wage bill. I compute the sectoral averages of $arpk$ and $arpl$ for each caste. The sectors are defined according to the National Industry Classification 2004 (NIC 2004), and there are 211 sectors at the 4-digit level. In Figure 1.1a, I plot the average $arpk$ of LC firms against that of HC firms. It is evident that the differences in $arpk$ exist in most of the sectors, while LC firm have a higher $arpk$ relative to HC firms. Similar results are documented for MC firms (see Figure 1.1b). Moreover, differences persist in sectors such as food products and beverages; tanning and dressing of leather; and manufacturing of luggage, handbags, saddlery, harness and footwear, and apparels, where the enterprise ownership of low castes is quite substantial (see Figure 1.1a). However, when I plot the $arpl$ of non-HC firms against that of HC firms, no such differences are documented (see Figure 1.1c for a comparison of LC and HC firms and Figure 1.1d for MC and HC firms).[15]

To evaluate within-sector $arpk$ differences and to control for regional heterogeneity, I run the following regression;

$$ln\, Y_i = \beta_0 + \beta_1 1_{L-CASTE} + \beta_2 1_{M-CASTE} + \Gamma + \varepsilon_i. \qquad (1.1)$$

[15]For robustness, see Figure 1.B.2 in Appendix 1.A, which provides 5-digit sector classification with 633 sectors.

FIGURE 1.1: $arpk$ and $arpl$: MSME 2006-2007



(a) $arpk - LC$ (b) $arpk - MC$

(c) $arpl - LC$ (d) $arpl - MC$

Notes: The orange circle represents a 4-digit sector (211 in total). The green cross(x) represents the linear fit between the variable on the x-axis (treated as independent var.) and the variable on the y-axis (treated as dependent var.). Black dashed represent in the 45 degree line. Figure a compares the weighted average $arpk$, where weights capital shares, for LC (low-caste) and HC (high-caste) firms. Figure b shows weighted average $arpk$, where weights capital shares, for MC (medium caste) and HC firms. Figure c compares the weighted average $arpl$, where weights labor input shares, for LC and HC firms. Figure d compares the weighted average $arpl$, where weights labor input shares, for MC and HC firms. The Sampling weights are applied.

The dependent variables are $\{arpk, arpl\}$. The main explanatory variables are the dummies for the low-caste firms, $1_{L-CASTE}$, and the middle-caste firms, $1_{M-CASTE}$, whose corresponding coefficients are β_1 and β_2. The estimators $\widehat{\beta_1}$ and $\widehat{\beta_2}$ are interpreted as the log points difference in the dependent variable between the low- and high-caste firms and the middle- and high-caste firms, respectively. The regressions include sector and state fixed effects. Additionally, there is a vector of controls, Γ, which includes gender and religion fixed effects.

The estimates suggest that MC and LC firms have 14% and 30% higher $arpk$, whereas $arpl$ is 5% lower for MC firms and no such differences for LC firms are observed relative to HC firms, respectively (see Table 1.2 specifications 1 and 3). In what follows, I explore how cross-caste dispersion in $arpk$ differs across rural and urban areas. To do so, I interact the caste dummies with the urban dummy, which allows me to disentangle the average effect of being in urban areas on non-HC firms' $arpk$ relative to that of HC firms. In particular, I run the following regression:

$$\begin{aligned} ln\, Y_i = &\gamma_0 + \gamma_1 1_{L-CASTE} + \gamma_2 1_{M-CASTE} + \gamma_3 1_{L-CASTE} \times 1_{Urban} \\ &+ \gamma_4 1_{M-CASTE} \times 1_{Urban} + \gamma_5 1_{Urban} + \Gamma + \varepsilon_i, \end{aligned} \tag{1.2}$$

where 1_{Urban} is the dummy variable for urban areas, and γ_3 and γ_4 represent the additional effect on the dependent variable of being in urban areas relative to rural areas for LC and MC firms, respectively. The estimates of γ_3 and γ_4 are negative, suggesting lower $arpk$ for non-HC firms (see Table 1.2).

The evidence that non-HC firms have a high $arpk$ is consistent with many models in which small firms with high returns are constrained from expanding. In these models, marginal products are proportional to average products under the assumption of a Cobb-Douglas production function; therefore, a high $arpk$ implies a high shadow cost of capital. It also suggests potential technological differences, in that HC firms are using modern capital-intensive techniques of production. Moreover, as expected, such constraints are more likely to bind in rural areas, where the caste system is more salient, relative to urban areas; therefore, high $arpk$ for non-HC firms in rural areas is consistent with this view. The next two facts will provide more corroborative evidence and strengthen

Table 1.2: ARPK and ARPL across castes

	(1)	(2)	(3)	(4)
Dep. Var.	*arpk*	*arpk*	*arpl*	*arpl*
MC	0.138	0.160	-0.0537	0.0044
	(0.019)	(0.022)	(0.015)	(0.022)
LC	0.299	0.311	-0.0063	0.0402
	(0.035)	(0.041)	(0.030)	(0.039)
URBAN		-0.0541		0.105
		(0.022)		(0.022)
MC $\times$ URBAN		-0.0591		-0.0994
		(0.027)		(0.029)
LC $\times$ URBAN		-0.0573		-0.0674
		(0.038)		(0.031)
Obs. (millions)	1.4	1.4	1.4	1.4
R-squared	0.45	0.45	0.45	0.45
State & NIC4 FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓

Notes: Results from the enterprise level regression using equations 1.1 and 1.2. Dependent variables are in logs and shown in column headings. The variables $arpk$ and $arpl$ are the average products of capital and labor, respectively. M-caste is the dummy variable for the middle-caste enterprises. L-caste is the dummy variable for the low-caste enterprises. The vector of controls, Γ, includes region, gender and religion fixed effects. Robust standard errors are in parentheses, clustered at caste, region and sector levels.

this view.

Finally, given the fact that $arpl$ is not different across castes, from here onward, I will primarily focus on the observed dispersion in $arpk$ and explore its various facets.

Fact 2: $arpk$ differences across castes decline with size and age.

Here I document the evolution of $arpk$ differences over firm size and firm age. I divide enterprises into five different size bins, defined by employment and compute β_1 for each size bin using the regression model described in equation 1.1. As shown in Figure 1.2a, differences persist among smaller enterprises, but they are essentially nil for large firms. In fact, if one looks at enterprises with more than 100 employees, the LC entrepreneurs have a lower $arpk$ relative to HC entrepreneurs. This evidence suggests that the technology differences across castes or the presence of capital adjustment costs cannot explain all the dispersion in $arpk$.[16]

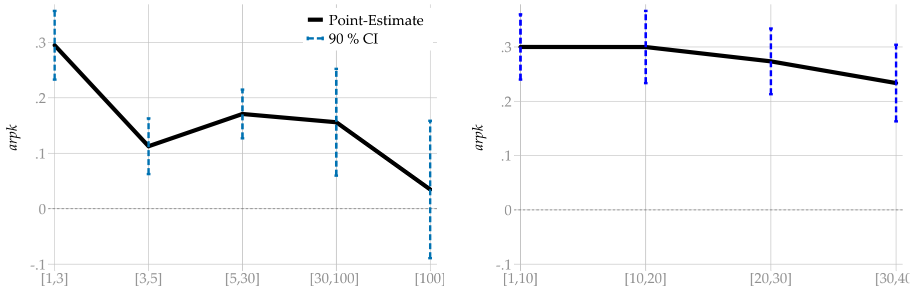
Furthermore, $arpk$ differences are heterogeneous across firms of different age groups. First, I find that the mean age is similar for LC and HC firms, whereas MC firms are half a year younger than other firms (the median age is 5 years for all castes; see Figure 1.B.3 in Appendix 1.A for the age distribution). Further, I pair enterprises into five different age groups and compute β_1 using the regression model described in equation 1.1 for each group. Figure 1.2b shows that $arpk$ is highest for young LC entrepreneurs, +35% relative to that of HC firms, and it declines over age. This result is consistent with the existing evidence on financial frictions and the firm life-cycle (e.g., see Hadlock and Pierce 2010) that suggests that size and age are good predictors of financial constraints such that young and small entrepreneurs are most constrained, and large and old firms are least likely to be constrained.[17] However, even among the oldest firms in the data, the $arpk$ for LC

[16]The presence of non-convexity in the production technology in capital usage could also lead to differences in $arpk$ across large and small firms. The presence of non-convex capital adjustment costs, including time-to-build, create dispersion in $arpk$ among small and large firms. However, such forces can not explain differences within a size group.

[17]I use a cross section to provide evidence on $arpk$ over firm age. It is plausible that the LC enterprises that are born in different years are inherently different from each other. The data do not allow me to rule out such a hypothesis.

firms is substantially higher than that of HC firms (+25%), pointing toward potential technological differences.

FIGURE 1.2: LCs' ARPK over Size and Age



(a) Employment (b) AGE

Notes: The blue line represents value of the L-caste dummy in regression specification 1.1. White bands represent 95% confidence intervals.

Fact 3: $arpk$ differences across castes decline over regional financial development.

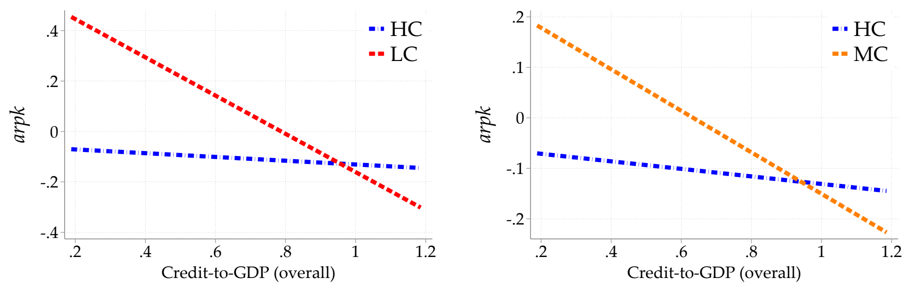
In this section, I further explore the effect of financial development on the cross-caste dispersion in $arpk$. Ayyagari et al. (2014) document large and persistent differences in financial development across states in India. Meanwhile, Munshi (2016) documents low mobility in India. Therefore, concerns regarding endogenous spatial sorting of entrepreneurs are quite low.

In what follows I construct a credit-to-output ratio for each state and use it as a measure of financial development.[18] The observed differences in $arpk$ across castes fall as the credit-to-output ratio increases (see Figure 1.3). In particular, LC firms have an $arpk$ that is 40% higher than that of HC firms in states with the lowest financial development

[18]State-wise indicators on GDP and domestic credit are taken from RBI's Handbook of Statistics on Indian states.

such as Bihar, Jharkhand, and Uttar Pradesh, but no such differences are observed in states such as Maharashtra (see Figure 1.3a). More interestingly, $arpk$ in absolute terms declines for LC entrepreneurs but is relatively stable for HC entrepreneurs. Similar trends are documented for MC firms as well (see Figure 1.3b). This evidence is consistent with the models of financial frictions in which better functioning financial markets improve credit allocation and selection among entrepreneurs in the economy.[19]

FIGURE 1.3: ARPK and External Financing



(a) Low-Caste Firms

(b) Middle-Caste Firms

Note. All lines represent linear regressions including sector fixed effects and control variable Γ. LC = low caste, MC = middle castes, and HC = high Castes. Sampling weights are applied. Regression details are provided in Table 1.B.1 in Appendix 1.A.

In this section, I have documented three stylized facts that suggest misallocation of capital across castes. Historically, LC and MC individuals had limited access to entrepreneurship and capital markets. Even today, they are relatively poorer (i.e., their financial wealth is limited), more likely to operate in labor-intensive sectors and use conventional means of production, and less likely to be linked with lending organizations (see Table 1.1). Furthermore, such constraints are most likely to affect small and young non-HC firms and Fact 2 is in line with this

[19]In these models, credit is allocated based on financial wealth or cash flows and not on productivity. Therefore, wealthy but unproductive individuals create firms and are also leveraged.

notion. I further provided suggestive evidence on the role played by underdeveloped financial markets in explaining such dispersion by exploiting the heterogeneity in regional financial development. However, the welfare implications of this dispersion are not straight forward and could be a result of differences in fundamentals such as the productivity process and technology, as well as heterogeneity in access to credit. To shed light on different sources of the $arpk$ dispersion, in the next section, I build a general equilibrium model in which agents of different castes make occupation choices in the context of caste-specific borrowing limits, technology and productivity. Through the lens of the model, I link the dispersion in $arpk$ partly to the misallocation of capital and quantify the role played by fundamentals and access to credit in explaining such dispersion. Finally, I evaluate the welfare implications of cross-caste dispersion in $arpk$ in the model.

1.5 THEORETICAL FRAMEWORK

The model is an extension of the framework used in Buera et al. (2011) and Buera and Shin (2013). Time is discrete and there is a measure M of infinitely lived agents that are heterogeneous across productivity z, assets a, and caste c. Every period, agents choose to become either workers or entrepreneurs based on their wealth a and entrepreneurial productivity z, and this occupational choice is represented by o_t. Financial wealth is determined endogenously by the consumption-saving problem described below, whereas productivity follows a stochastic process such that agents retain their last-period productivity z_t with probability ψ_c, and with probability $1-\psi_c$, they draw their new productivity from a Pareto distribution with scale parameter η_c. The parameter ψ_c represents the persistence, whereas η_c captures the dispersion in the productivity process. If $\psi_c = 1$, then there is no uncertainty, and hence productivity is the sole determinant of the agent's saving behavior and occupational choice. On the contrary, when $\psi_c = 0$, the productivity process is a random walk.

Preferences: Agents' utility functions are strictly increasing, concave and satisfy standard Inada conditions. Agents discount their future utility at a discount rate ρ and at any point in time t, their preferences are represented by the following function:

$$\mathbb{E}\sum_{t=o}^{\infty}\rho^{t}\frac{\zeta_t^{1-\gamma}-1}{1-\gamma}.$$

The entrepreneurs have access to a decreasing returns to scale production function $f(z,k,l) = z\,(k^{\alpha}l^{\beta})^{1-\nu_c}$ where $\alpha+\beta=1$ and $1-\nu$ is the span-of-control parameter that varies between 0 and 1. The output price is normalized to one. An entrepreneur rents capital k in the financial market (more discussion follows below) and hires labor l to produce y units of a single good. Also, entrepreneurs need to pay a per-period fixed cost of operation κ.

Financial Markets: There is a perfectly competitive intermediary that receives deposits from savers and lends these funds to entrepreneurs. There is no intermediation cost; that is, the deposit rate is equal to the borrowing cost. The rental rate of capital is $r_t+\delta$ in period t, where δ is a time invariant depreciation cost and r_t is the deposit rate. The financial markets are incomplete in a way that entrepreneurs' ability to borrow capital is proportional to their asset base.[20] Specifically, the capital constraints take the following forms:

$$k_t \leqslant \lambda_c\, a_t; \quad a_t \geqslant 0,$$

[20]Recently, financial frictions based on cash flows rather than collateral are being used in the literature, such as Buera et al. (2011). However, I argue that collateral-based financial constraints are more common in India as the majority of loans are based on collateral and not on cash flows. For example, more than 84% of the loans required collateral in India in 2014 according to the World Enterprise Survey 2014, World Bank.The micro-data for the WES are available: `http://microdata.worldbank.org/index.php/catalog/2225/get_microdata`.

where λ_c measures the degree of credit constraints and varies from 1 to ∞.[21] Individuals of certain caste c with $\lambda_c = 1$ will operate in a zero-credit environment (financial autarky), whereas $\lambda_c = \infty$ will allow individuals to borrow according to their productivity and not based on their financial wealth.[22]

Recursive Formulation of Individuals' Problem: Agents maximize their expected utility for a given set of factor prices $\{w, r\}$, their asset base a, productivity z, and a vector of probabilities corresponding to future productivity z' given by $d\Upsilon(z'|c)$, such that the resource constraint always binds. The value function that agents maximize is

$$V(a, z, c) = \max\{V^w(a, z, c), V^e(a, z, c)\}. \tag{1.3}$$

The workers' value function is given by

$$\begin{aligned} V^w(a, z, c) = \max_{\zeta, a' \geqslant 0} u(\zeta) + \rho\{\psi_c V(a', z, c) + (1 - \psi_c) \int_{z'} V(a', z', c) d\Upsilon(z'|c)\} \\ \text{s.t.} \quad \zeta + a' \leqslant w + (1 + r)a. \end{aligned} \tag{1.4}$$

The entrepreneurs' value function is given by

$$\begin{aligned} V^e(a, z, c) = \max_{\zeta, a' \geqslant 0} u(\zeta) + \rho\{\psi_c V^e(a', z, c) + (1 - \psi_c) \int_{z'} V(a', z', c) d\Upsilon(z'|c)\} \\ \text{s.t.} \quad \zeta + a' \leqslant z \left(k^\alpha l^\beta\right)^{1-\nu_c} - wl - (r + \delta)k - \kappa + (1 + r)a \\ k \leqslant \lambda_c a. \end{aligned} \tag{1.5}$$

[21] This type of financial frictions can be micro-founded with the following limited enforcement problem. Entrepreneurs deposit their financial wealth a and can borrow capital k from financial intermediaries at the beginning of the period, whereas, financial institutions can recover upto $\frac{1}{\lambda_c}$ times the rented capital in case of default. The entrepreneur will lose financial wealth a but is included in the future economic activity. In this scenario, financial intermediaries will lend upto $\lambda_c a$, which makes default incentive incompatible.

[22] Similar to Buera and Shin (2013), I rule out any borrowing for intertemporal consumption smoothing by assuming $a_t \geqslant 0$. This constraint is binding for workers, whereas it does not matter for entrepreneurs as they need to have a sufficiently large asset base to fund their capital requirements.

1.5.1 *Recursive Competitive Equilibrium*

Equilibrium: At time 0, given the distribution $\Lambda_0(a,z,c)$, the equilibrium of the economy is characterized by a sequence of allocations $\{o_t, \zeta_t, a_{t+1}, k_t, l_t\}_{t=0}^{\infty}$, factor prices $\{w_t, r_t\}_{t=0}^{\infty}$, and $\Lambda_t(a,z,c)_{t=1}^{\infty}$ such that

1. $\{o_t, \zeta_t, a_{t+1}, k_t, l_t\}_{t=0}^{\infty}$ solves the individuals' policy functions for given factor prices $\{w_t, r_t\}_{t=0}^{\infty}$;

2. The capital, labor and goods markets clear in each period:

$$\int_{o_t(a,z,c)=e} k_t \, d\Lambda_t(a,z,c) - \int a_t \, d\Lambda_t(a,z,c) = 0,$$

$$\int_{o_t(a,z,c)=e} l_t \, d\Lambda_t(a,z,c) - \int_{o_t(a,z,c)=w} d\Lambda_t(a,z,c) = 0,$$

$$\int_{o_t(a,z,c)=e} [z_t(k_t^{\alpha} l_t^{\beta})^{1-\nu_c} - \kappa] d\Lambda_t(a,z) = \int \zeta_t d\Lambda_t(a,z,c) + \delta K;$$

3. The joint distribution of productivity, assets for each caste $\Lambda_t(a,z,c)_{t=1}^{\infty}$ evolve according to the equilibrium mapping:

$$\begin{aligned}\Lambda_{t+1}(a,z,c) = \psi_c \int_{\{z, a_{t+1}(a,z,c)<a\}} & \Lambda_t(da,dz,c) + \\ (1-\psi_c) \int_{\{z' \leqslant z, a'(a,z,c) \leqslant a\}} & \Lambda_t(da,dz,c) d\Upsilon_t(z'|c).\end{aligned}$$

1.6 QUANTITATIVE ANALYSIS

In this section, I evaluate the role of technology, productivity, and access to credit in explaining the cross-caste dispersion in the $arpk$. Further, I quantify the TFP losses due to resource misallocation generated by such asymmetries at the extensive and intensive margins. I begin by calibrating the model to the manufacturing sector of India using data from multiple sources.[23] I then conduct various experiments to validate the model. In particular, I vary the parameters that change the credit-to-output ratio across castes to mimic various regions in India and to draw verifiable predictions from the model. Finally, I evaluate the losses from financial and fundamental asymmetries (includes technology and productivity asymmetries) in aggregate TFP.

1.6.1 *Calibration*

The calibration strategy is based on Buera and Shin (2013) and Midrigan and Xu (2014). Overall, I need to specify values for 17 parameters: span-of-control of production technologies, dispersion and persistence in ability distributions, degree of financial frictions, fixed cost of operation, discount factor, coefficient of risk aversion, capital depreciation rate, and physical capital share. The parameterization proceeds in two steps. First, I fix a set of parameters that are fixed outside the model (e.g., the depreciation rate). The fixed parameters are difficult to identify with the available data, so I use the values that are commonly used in the literature. Second, given the values of these fixed parameters I choose the remaining parameters to match the salient features of the economy (e.g., the distribution of employment and business income, among others).

1.6.1.1 Data for Empirical Moments

As discussed in Section 3.2, the MSME dataset, which contains variables such as capital stock, the wage bill, output, and credit, only represents firms below a certain threshold level of capital, while the Economic

[23] I choose the manufacturing sector for two reasons. First, because of restrictions on the data side, I can evaluate the proportion of output that is linked to MSMEs in the manufacturing sector but not in the service sector. This is important as I match this moment in the model to compute comparable statistics. Second, Hsieh and Klenow (2009a) also evaluate the role of misallocation in the manufacturing sector in India, thus, their analysis helps me to gauge my results with respect to their findings.

census 2005 contains a universe of firms but only provides information on the caste of the owner and the number of employees. Therefore, I use the Economic Census 2005 to compute the employment distribution for the overall economy and for each caste.[24] Meanwhile, the serial correlation of output and the credit-to-output ratio for each caste are evaluated using the MSME dataset. For this reason, I need to define MSME firms in the model, such that I can compute the model counterparts of the empirical moments.

To define MSME firms in the model, I need to compute the MSME capital threshold stock $\bar{K}$ in the model. I evaluate $\bar{K}$ such that the total output produced by firms with capital stock below $\bar{K}$ is matched to its data counterpart. The share of output produced by MSMEs in the data in 2006 is computed using the ASI-NSS 2006 dataset. According to the reports generated by the Ministry of Micro, Small and Medium Enterprises in India, the threshold to be defined as MSME is in "cumulative investment in plant and machinery (original cost)" (see Garcia-Santana and Pijoan-Mas 2014 for more details).[25] This variable is available in the ASI-NSS dataset as *value of plant and machinery owned by the firm*. I use the MSME threshold to compute the share of output produced by firms that are below this threshold, and it stands at 41%.[26]

The income distributions and population shares for each caste are computed using IHDS 2005 data.[27]

[24] I use EC 2005 because it is not available for 2006. I compute the employment distribution for ASI-NSS 2006 and Economic census 2013-2014, and it is quite stable over time (see Appendix 1.A).

[25] The limit is 100 million in Indian rupees.

[26] Because of data restrictions, I assumed that the share of MSMEs output as the fraction of total output in 2007 is the same as in 2006.

[27] The business income distribution is computed for the year 2005 in the IHDS because of the absence of data for the year 2007. However, data are available for the year 2012 and the income distribution is very stable over time (see Table 1.B.6 in Appendix 1.A).

1.6.1.2 Identification

Fixed Parameters: A model period is one year. The structural parameters α, γ, and δ are the same across castes. The annual depreciation rate for capital is $\delta = 0.06$, the capital income share is $\alpha = 0.33$, and the coefficient of risk aversion is $\gamma = 1.5$, following Hsieh and Klenow (2009a) and Cagetti and De Nardi (2006).[28]

Fitted Parameters: Given the parameters α, γ, and δ, the model is solved to match certain moments in the data. The discount factor is set at $\rho = 0.844$, the same for every caste, to match the annual interest rate of 5.682% in 2007.[29] The fixed cost of operation is set at $\kappa = 0.10$ to match the relative size of entrants with respect to the incumbents.[30]

The span-of-control parameters $\{1-\nu_{lc}, 1-\nu_{mc}, 1-\nu_{hc}\}$ are set such that the business income share of the bottom 95% of entrepreneurs is the same in the data and the model for each caste. A lower $1-\nu_c$ implies a larger scale of operation and higher profits for entrepreneurs, which further makes the income distribution of entrepreneurs more dispersed; therefore, it is informative regarding the production technology used by each caste. Meanwhile, the scale parameters of the ability distribution $\{\eta_l, \eta_m, \eta_h\}$ are set such that the employment share of the bottom 95% of enterprises is matched in the data and the model for each caste. A higher value of a η_c means a thicker tail of the productivity distribution, which further implies a greater employment generation by large firms. This helps the model to distinguish η_c for each caste.

[28] The capital income share α is assumed to be the same across castes because of non-availability of data that could be used to measure it credibly (see Section 1.7 for more detail).

[29] The annual real interest rate in India varies from 2% to 8.34% between 1999 and 2010.

[30] I measure the relative entrant size in the ASI-NSS 2006 dataset that encompasses the manufacturing sector (see Appendix 1.A for more details). I do not allow for heterogeneity in κ, which matters for the misallocation at the extensive margin, across castes in order to focus on financial friction as the sole driver of misallocation. Moreover, as shown in Table 1.5, the model closely matches the percentage of firms owned by each caste without requiring any heterogeneity in κ.

Table 1.3: Parameters Value

Parameter	Value	Description
Fixed:		
δ	0.060	Annual depreciation rate physical capital
α	0.330	Physical capital share
γ	1.500	Coefficient of risk aversion
Fitted:		
ρ	0.844	Discount factor
$1-\nu_h$	0.761	Span of control for HC
$1-\nu_m$	0.745	Span of control for MC
$1-\nu_l$	0.745	Span of control for LC
ψ_h	0.927	Persistence in productivity for HC
ψ_m	0.922	Persistence in productivity for MC
ψ_l	0.918	Persistence in productivity for LC
λ_h	1.760	Degree of financial frictions for HC
λ_m	1.370	Degree of financial frictions for MC
λ_l	1.160	Degree of financial frictions for LC
η_h	4.520	Scale parameter of ability distribution for HC
η_m	4.700	Scale parameter of ability distribution for MC
η_l	4.890	Scale parameter of ability distribution for LC
κ	0.100	Fixed cost of operation

To discipline the persistence of the productivity process $\{\psi_{lc}, \psi_{mc}, \psi_{hc}\}$, I match the one-year autocorrelation of output in the data with its counterpart in the model. A persistent productivity process increases the serial correlation of output and reduces the impact of financial frictions on capital misallocation in the stochastic steady state (see Midrigan and Xu 2014 and Moll 2014). Therefore, heterogeneity in ψ_c allows me to disentangle the effects of borrowing constraints from that of the productivity process.

The parameters related to financial frictions $\{\lambda_{lc}, \lambda_{mc}, \lambda_{hc}\}$ are fixed such that the overall credit-to-output ratio in the economy is the same in the model and the data. Furthermore, the credit-to-output ratios for LC and MC relative to HC are also matched.[31] A higher λ_c implies a

[31] In the MSME dataset, the credit-output ratios for the high, middle and low castes are 0.30, 0.17, and 0.09, respectively.

Table 1.4: Model Moments

Targeted Moments	Model	Data
HC:		
One-year autocorrelation of output	0.94	0.94
Employment share of bottom 95%	0.62	0.56
Income share of bottom 95%	0.64	0.67
Population share	0.35	0.35
MC:		
One-year autocorrelation of output	0.96	0.96
Employment share of bottom 95%	0.66	0.72
Income share of bottom 95%	0.67	0.71
Population share	0.36	0.36
Credit/output rel. HC	0.56	0.56
LC:		
One-year autocorrelation of output	0.96	0.94
Employment share of bottom 95%	0.72	0.78
Income share of bottom 95%	0.72	0.72
Population share	0.29	0.29
Credit/output rel. HC	0.27	0.27
Overall Economy:		
Annual interest rate	5.7%	5.7%
Entrants' relative size	0.32	0.23
Credit/output	0.45	0.45
Share of MSME sector	0.41	0.41
Additional moments:		
Overall employment share of bottom 95%	0.64	0.64
Overall employment share of bottom 90%	0.52	0.52
Overall income share of bottom 95%	0.67	0.68
Overall income share of bottom 90%	0.54	0.55

Notes: Employment distribution in the data is evaluated with Economic Census 2005 and income distribution is evaluated with IHDS 2005. The overall credit-to-output ratio is taken from statistics published by the Reserve Bank of India. The credit-to-output ratio for LC and MC relative to HC is computed using the MSME data. One-year autocorrelation of output is computed using the synthetic panel in the MSME data while controlling for sectoral and regional heterogeneity.

larger supply of credit in the economy and hence higher leverage in the economy.

The parameters are estimated using the following routine. For arbitrary values of the vector of parameters, $\Xi = (1-\nu_{lc}, 1-\nu_{mc}, 1-\nu_{hc}, \eta_{lc}, \eta_{mc}, \eta_{hc}, \psi_{lc}, \psi_{mc}, \psi_{hc}, \lambda_{lc}, \lambda_{mc}, \lambda_{hc}, \kappa, \rho)$, I solve the recursive competitive equilibrium and evaluate the stationary distribution in $\{a, z, c\}$. Using this distribution, I compute the business income distributions, employment distributions, and the $\bar{K}$, the capital threshold for MSMEs, such that the share of output produced by them is matched to its data counterpart. I evaluate respective credit-to-output ratios. Furthermore, I draw from the stationary distribution to simulate the economy for three periods, construct a balanced panel of MSME firms, and compute the serial correlation of output in the same spirit of the empirical counterpart. I denote all these simulated moments by $\Omega(\Xi)$ and estimate the fitted parameters $\widehat{\Xi}$ using a minimum distance criterion given by

$$\mathcal{L}(\Xi) = \min_{\Xi}(\widehat{\Omega} - \Omega(\Xi))'\mathbf{W}(\widehat{\Omega} - \Omega(\Xi)). \tag{1.6}$$

I set the weighting matrix $\mathbf{W} = \mathbf{I}$ and use grid search to find the minimum.

1.6.2 *Results*

The fitted parameters from the simulated method of moments are listed in Table 1.3, and the implied moments of the model, in comparison to their data counterparts, are presented in Table 1.4. The model closely matches the set of targeted moments as well as a set of additional moments that captures the overall income and size distribution of firms. The model identifies different parameters of technology, productivity, and financial constraints for each caste, and its implications are discussed below.

The scale parameter of the ability distribution is lower for HC individuals compared to that of MC and LC individuals. This stems from the fact that the employment distribution of HC is skewed toward the right, which is achieved in the model by having a thick right tail of the productivity distribution. This further implies that the mean ability is higher for HC individuals relative to that of MC and LC individuals. These differences could be interpreted as differences in human capital (proxied by years of schooling) or cognitive abilities, which stems from centuries of discrimination against low castes. Further examination of these differences is beyond the scope of this paper and is left for future research.[32] The persistence of the productivity process is also identified to be lower for LC and MC individuals relative to that of HC individuals. This is the result of their higher serial correlation of output in the data (see Table 1.4). A lower persistence would also imply higher income uncertainty as there is more occupational switching among non-HC individuals.[33]

The model identifies that MC and LC entrepreneurs operate their firms with a production technology with smaller span-of-control parameter relative to HC entrepreneurs. This is because the business income distribution is relatively more skewed towards right in the case of HC compared to other castes. This implies a small size for LC and MC firms, but with high profitability, a prediction that is verified in the data as well (see Appendix 1.A). Finally, the model implies limited access to credit for LC and MC relative to HC individuals. In particular, the model identifies λ_{lc} to be 34% smaller and λ_{mc} to be 23% smaller than λ_{hc}. This is driven by lower credit-to-output ratio of LC and MC firms relative to that of HC firms.

[32] Fehr and Hoff (2011) and Hoff and Pandey (2006) argue that caste affects cognitive task performance and responses to economic opportunities by young boys in villages.

[33] Higher uncertainty about future productivity could stem from various sources – for example, an absence of entrepreneurial networks that could help sustain bad shocks, institutional discrimination, among others. Asker et al. (2014) discuss the impact of higher volatility in productivity on $arpk$ and its potential causes.

1.6.2.1 Misallocation across Castes

The main objective of this paper is to quantify the misallocation of resources across castes and its impact on aggregate TFP. The literature has stressed the role of financial frictions on two different margins of misallocation: the extensive margin and the intensive margin.

The intensive margin refers to the overall dispersion in $arpk$; however, this paper is primarily concerned with the $arpk$ dispersion across castes.[34] The model predicts that LC and MC agents have 13% and 21% higher $arpk$ than that of HC agents. This captures around 70%-80% of the values observed in the data (see Table 1.5). The dispersion in $arpk$ is driven by two factors in my model. First, difference in access to credit make LC and MC relatively more constrained in the stochastic steady state, thereby increasing the shadow cost of capital.[35] Further, I evaluate the role of asymmetric access to credit in explaining the cross-caste dispersion in $arpk$. To do so, I equalize the parameters governing the degree of financial frictions across castes, keeping all else constant. I find that this leads to a reduction of around 70%-80% of the differences in $arpk$ across HC and non-HC firms.

Second, differences in fundamentals such as technology and the productivity process, exacerbate the dispersion in $arpk$ and explain the remaining 20%-30% of differences in $arpk$ across castes. Multiple forces are at work here; (i) heterogeneity in the persistence of the productivity

[34] The literature refers to the dispersion in the marginal revenue product of capital (MRPK) as a misallocation of capital (Hsieh and Klenow 2009a). However, my model implies that the MRPK is directly proportional to the ARPK, $MRPK = \alpha(1-\nu)ARPK$.

[35] The consumption Euler equation for constrained entrepreneurs contains Θ_{t+1}, the shadow cost of capital:

$$\zeta_t^{-\gamma} = \rho \int_{z'} \left\{ \zeta_{t+1}^{-\gamma}(1 + r_{t+1} + \lambda_c \Theta_{t+1}) \right\} d\Pi(z'|z),$$

$$\Theta_{t+1} = \max[f_k(\lambda_c a_{t+1}, z_{t+1}) - (r_{t+1} + \delta), 0].$$

Table 1.5: Results

	Model	Data
Intensive-margin		
$arpk - MC$	$+13\%$	$+22\%$
$arpk - LC$	$+21\%$	$+34\%$
$k/l - MC$	-11%	-31%
$k/l - LC$	-28%	-58%
Extensive-margin		
$\%\ of\ firms$-LC	16%	17%
$\%\ of\ firms$-MC	42%	46%

Notes: The measures of $arpk$ and capital-intensity(k/l) are computed for MSMEs in both the data and the model and represent their respective values with respect to HC firms in the manufacturing sector. The percentage of firms owned by each caste in the data is computed using the Economic Census 2005. Sampling weights are applied.

process exacerbates $arpk$ differences across castes in conjunction with financial constraints, as lower persistence dampens the channel of self-financing. For non-HC entrepreneurs, past productivity is not a good predictor of future productivity; therefore, it enhances the volatility of business income and induces more occupation switching. (ii) A lower span-of-control implies a higher $arpk$ for LC and MC firms independent of the degree of financial constraints; however, such dispersion in $arpk$ does not directly imply misallocation.[36]

The extensive margin refers to the distorted occupation choice in the context of the model. In particular, in this economy, the productivity threshold of entry $\overline{z}(a, \lambda_c)$ is decreasing in a and increasing in λ_c.[37]

[36]However, in models of technology adoption under financial constraints, this could be a potential source of misallocation (see Section 1.6.4 for detailed discussion).

[37]The max operator in equation 1.3 pins down the occupation choice $o(a, z, c)$ for each agent such that, whenever the value of being an entrepreneur is greater (lower) than that of being a worker, agents decide to be an entrepreneur (worker).

Under such circumstances, non-HC individuals' productivity threshold $\overline{z}(a, \lambda_c)$ is higher than that of HC individuals, which implies higher labor-force participation and a lower entrepreneurial rate for the former group. The model does a good job in matching the number of firms owned by each group (see Table 1.5). For more detailed results on enterprise ownership across castes in the Economic Census 2005, see Table 1.B.2 in Appendix 1.A. Another important implication is that the average productivity of the low-caste entrepreneurs should be relatively higher than that of HC entrepreneurs because only the very productive non-HC agents can operate profitably. A model-consistent measure of firm productivity is given by

$$tfpr := ln(TFPR_{ic}) = ln(Y_{ic}) - \alpha(1-\nu_c)ln(K_{ic}) - \beta(1-\nu_c)ln(L_{ic}),$$

where i represents the firm, c stands for caste and $tfpr$ is revenue productivity.[38] Capital is demoted by K_{ic} and L_{ic} is the wage bill, which is consistent with Section 2.3.1. In the data, I find HC entrepreneurs to be less productive than LC entrepreneurs (see Table 1.B.3 in Appendix 1.A), and this effect is captured well by the model, where LC firms are 5% more productive than HC firms. However, the model does not imply any significant difference in the productivity between HC and MC entrepreneurs.

Furthermore, such a reduction in entrepreneurship means lower demand for factors of production, which in turn implies lower factor prices and higher profits for incumbent entrepreneurs. This allows the entry of more HC firms that are marginally unproductive. As a

For a certain asset base a, there exists a productivity threshold $\overline{z}(a, \lambda_c)$ such that $\pi(a, \overline{z}(a, \lambda_c), \lambda_c; r, w) = w$. A wealthy but unproductive agent is more likely to enter into entrepreneurship than a poor but productive one. This creates misallocation of talent or misallocation at the extensive margin. An agent with zero wealth can never be an entrepreneur, $\overline{z}(0) \to \infty$, whereas an agent with infinite wealth can be an entrepreneur only if he or she is productive enough, $\overline{z}(\infty) = \overline{z}$.

[38] This measure includes firm-level prices and therefore encapsulates the dispersion in markups as well. However, I do provide more evidence on quantity-based measures and markups (see Section 1.7).

result, the overall TFP of the economy goes down. I will postpone the discussion on welfare implications until Section 1.6.4.

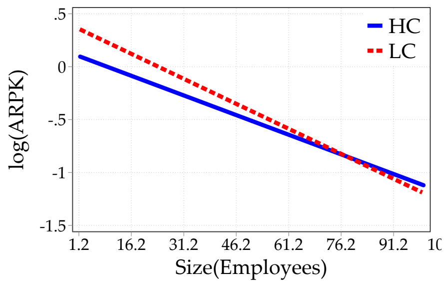
1.6.2.2 Firm Entry and Life Cycle Dynamics

In this section, I discuss the life cycle of firms and shed light on its differences across castes. Similar to Fact 3 documented above, I found that the model predicts declining differences in $arpk$ over the firm life cycle. In particular, Figure 1.4a shows that the $arpk$ of HC and LC firms declines over employment; however, the decline is faster for LC firms such that no difference in $arpk$ emerges for larger firms in the size distribution. Meanwhile, Figure 1.4b depicts a similar decline; however, differences do not completely disappear even for older firms in the sample. This model prediction is consistent with what I documented above as Fact 3 in Section 1.4. These results are driven by the technology differences across castes. LC firms use small scale technology, which implies a higher $arpk$ than HC firms, irrespective of financial frictions.

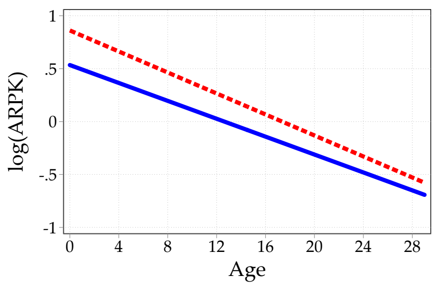
Finally, Figure 1.4c shows how firm output grows over its age. Firms enter small because of their limited borrowing capacity; however, with time they accumulate financial wealth and grow in size. It is clear that LC firms enter small and then grow slower relative to HC firms. This is driven by a combination of two forces: limited borrowing capacity and the use of small-scale technology by LC firms. As a result, firm size diverge over age. A similar pattern of life-cycle dynamics emerges in the data as well (see Figure 1.B.4 in Appendix 1.A).

Until now, I have shown that the cross-caste dispersion in $arpk$ is driven by differences in fundamentals and access to credit. In particular, the model estimates stringent borrowing limits for non-HC entrepreneurs and large span-of-control parameter for HC entrepreneurs. The untargeted moments such as $arpk$ and firm ownership for non-HC castes are very well matched to their data counterparts. In what follows, I use the model to evaluate how the dispersion in $arpk$ evolves over

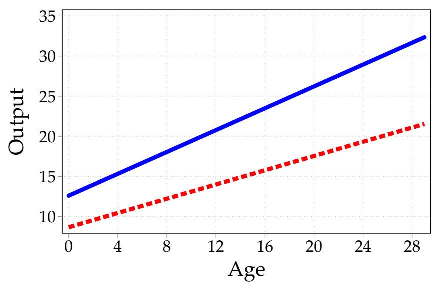
FIGURE 1.4: ARPK and Firm Life Cycle: Model



(a) ARPK and Size



(b) ARPK and Age



(c) Output and Age

Note. All lines represent a linear-fit of the data produced with a benchmark calibration in the model.

regional financial development.

1.6.3 *Regional Financial Development and the arpk Dispersion*

In this section, I will revisit Fact 2 in Section 1.4 in the spirit of the model. First, I document that the states that are more financially developed (i.e., those with a high credit-to-output ratio) distribute credit more efficiently across castes as well. In particular, Figure 1.5a depicts a positive correlation between the regional credit-to-output ratio and the credit-to-output ratio of LC firms relative to HC firms.[39] Second, I exploit this heterogeneity to further validate the findings of the model by solving the model for different levels of financial development and predicting the dispersion in $arpk$ across castes. In particular, I move the parameter that governs the degree of financial frictions in the model λ_{lc} such that the credit-to-output ratio for LC entrepreneurs relative to that of HC entrepreneurs increases and, thus, the overall overall credit-to-output ratio increases as well (see Figure 1.6a).[40]

The model predicts a steeply declining $arpk$ for LC entrepreneurs, whereas a mild increase is captured for HC entrepreneurs, owing to the enhanced borrowing capacity of and a high interest rate in response to a high demand for capital. As a result, $arpk$ differences across castes decline from over $+30\%$ in the least financially developed regions to essentially nil in the most advanced ones (see Figure 1.6b). As a result, increasing borrowing capacity leads LC firms to become more capital intensive (see Figure 1.6c). I find support for these predictions in the data (see Figure 1.5b for $arpk$ and Figure 1.5c for capital intensity).

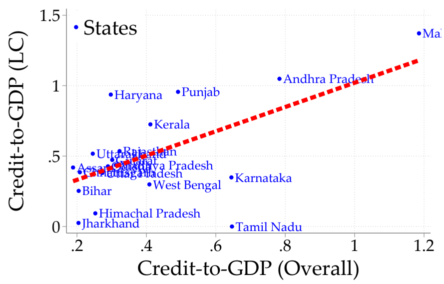
Furthermore, a high borrowing capacity allows the entry of more LC firms relative to the benchmark economy, which further implies a reduction in firm level $tfpr$ of the marginal entrant of LC individuals and a steep convergence toward the average level of $tfpr$ of HC individuals (see Figure 1.B.6 in Appendix 1.A). This prediction is consistent with the evidence that I find in the data; that is, productivity is 40% higher than that of HC firms in less financially developed states to below zero in states with well-functioning financial markets (see Figure 1.B.7 in Appendix 1.A).

I further use a regression model to pin down the elasticity of $\{arpk, k/l, tfpr\}$ to financial development for LC entrepreneurs. In what follows, I in-

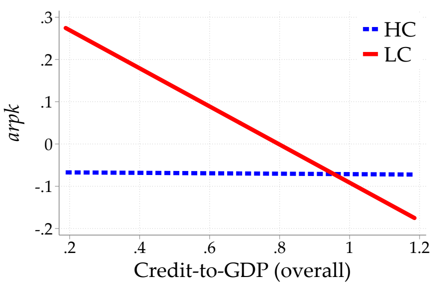
[39] See Figure 1.B.5 in Appendix 1.A for MC firms.

[40] The model is solved for several values of λ_{lc} between 1 and 2. The regional analysis captures the spirit of Buera et al. (2011), where the improved efficiency of financial sector increases overall welfare. Output per capita, capital intentsity and TFP increase over the regional credit-to-output ratio (see Figure 1.B.8 in Appendix 1.A).

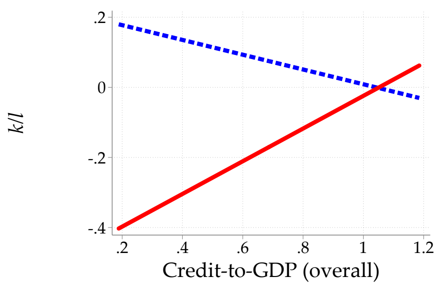
FIGURE 1.5: Financial Development and LC Firms-Data



(a) Access-to-Credit



(b) ARPK



(c) Capital-Intensity

Notes: The overall credit-to-GDP ratio is computed with statistics published by the Reserve Bank of India. The credit-to-GDP ratio for LC and MC firms is computed in MSME data. Graphs (b) and (c) present a linear fit of residualized $arpk$ and k/l. Sampling weights are applied

teract the caste dummies with the financial development of states Fd_s. The regression specification is

$$ln\,Y_i = \widehat{\gamma}_0 + \widehat{\gamma}_1 1_{L-CASTE} + \widehat{\gamma}_2 1_{L-CASTE} \times Fd_s + \widehat{\gamma}_3 Fd_s + \Gamma + \varepsilon_i, \tag{1.7}$$

where $\widehat{\gamma}_2$ represents the elasticity of the dependent variable to Fd_S with respect to HC entrepreneurs, respectively. The value of $\widehat{\gamma}_2$ is significantly negative for $arpk$ and $tfpr$ and positive for k/l, suggesting an improved allocation of credit across castes (see Table 1.6).

Table 1.6: Financial Development and LC Firms: Data

	(1)	(2)	(3)
Dep. Var.	$arpk$	k/l	$tfpr$
LC	0.538	-0.875	0.416
	(0.101)	(0.109)	(0.175)
Fd	-0.0254	-0.196	-0.254
	(0.085)	(0.104)	(0.073)
LC × Fd	-0.533	0.765	-0.454
	(0.175)	(0.187)	(0.228)
LC-population	1.072	-1.763	0.382
	(0.228)	(0.246)	(0.187)
Observations	624,987	624,987	605,255
R-squared	0.139	0.250	0.075
NIC4 FE	✓	✓	✓
Controls	✓	✓	✓

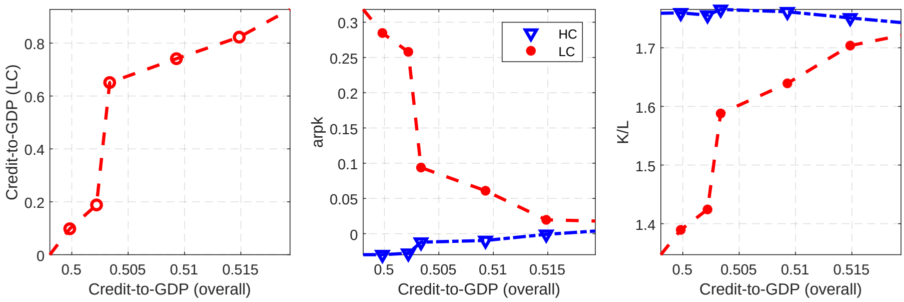
Notes: Results from the enterprise level regression using equation 1.7. Dependent variables are in logs and shown in column headings. Fd is an index of financial development across states. The vector of controls, Γ, includes region, gender and religion fixed effects. Robust standard errors are in parentheses, clustered at the caste, region and sector level. Sampling weights are applied.

The excess entry of LC entrepreneurs increases demand for capital and labor, which further implies high factor prices such as the interest rate and wages. Moreover, a high interest rate spurs saving and credit growth simultaneously, which increases the asset holdings of households (see Figure 1.B.9b in Appendix 1.A).[41] Meanwhile, high wages, in combination with more output because of the efficient allocation of capital, increase household consumption (Figure 1.B.9a in Appendix 1.A), and such an improvement in financial markets primarily benefits

[41]In this model, households consist of a single agent.

marginalized individuals, that is , non-HC agents in the model. As a result, one can see the convergence across castes in these economic variables. Similarly in the data, I find that the efficient allocation of credit in some states such as Maharashtra implies high household welfare (i.e., high consumption and assets for LC households; see Figure 1.B.10 in Appendix 1.A).[42]

FIGURE 1.6: Finance Development and LC Firms: Model



(a) Acces to Credit-LC (b) ARPK (c) Capital Intensity

1.6.4 *Counterfactual Analysis*

In the last section, I evaluated how firm performance increases for non-HC individuals over regional financial development. Moreover, I documented that overall economic performance, in terms of output and consumption per capita, improves as states' financial markets perform better. Therefore, it would be interesting to evaluate by how much economic welfare will improve if every state in India efficiently allocates credit across castes. I answer this question in two different counterfactual calibrations.

I discussed in Section 1.6.2 that non-HC individuals differ in terms of fundamentals such as technology and productivity, as well as their respective borrowing capacities. Therefore, in counterfactual analysis

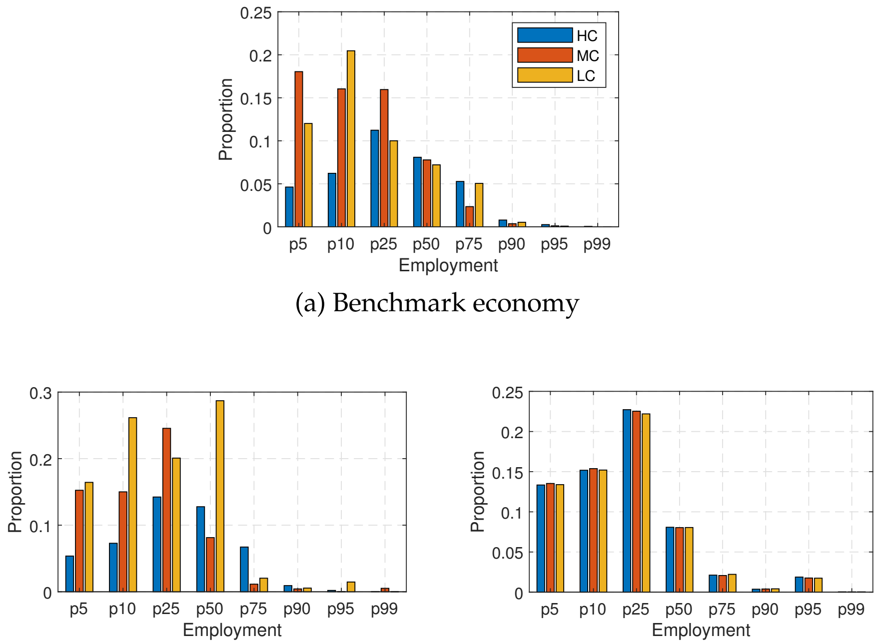
[42]In the data, the household assets represent household possessions and housing.

$CF1$, I allow the degree of financial frictions for non-HC individuals to be similar to that of HC individuals (i.e., $\lambda_{lc} = \lambda_{mc} = \lambda_{hc}$) and in the second counterfactual exercise $CF2$, along with symmetric access to credit, I enforce the same fundamentals for all castes. The model predicts TFP gains of 6% in $CF1$ and an additional gain of 4% in $CF2$ (see Table 1.7).

These gains come from two main sources in $CF1$: first, the reallocation of capital from unproductive HC entrepreneurs to more productive non-HC entrepreneurs increases the allocative efficiency of the economy; therefore, as a result, dispersion in $arpk$ declines by 13% and output per capita increases by 8%. Moreover, the economy becomes more capital intensive, with gains of $+7\%$. These gains mostly come from LC and MC entrepreneurs, which increase their capital-labor ratio by 15% and 13%, respectively. Second, the reduction in borrowing constraints induces the entry of more non-HC entrepreneurs. The share of LC enterprises increases from 16% in the benchmark economy to 29%, whereas the share of MC entrepreneurs decreases from 42% to 35% - that is, exactly proportional to their respective population weights. Moreover, because of the excess entry of entrepreneurs, demand for capital and labor increases. This implies a 47% higher interest rate in $CF1$, which further led to the exit of unproductive HC firms. The labor productivity gains for non-HC firms, as mentioned in Table 1.7, are driven by increments in wages.

In the $CF2$ economy I allow the span-of-control parameter and productivity distribution to be same for all castes (see Table 1.B.7 in Appendix 1.A for a full characterization of the parametric values). This economy experiences gains of 15% in output per capita and 13% in capital intensity, where most of these gains are driven by the improved performance of non-HC entrepreneurs (see column three of Table 1.7). In this economy, the gains come from three sources: (i) improved allocation of capital at the intensive margin; (ii) improved selection of entrepreneurs at the extensive margin; (iii) a large span-of-control and improved productivity distribution, which allows non-HC firms to operate on a larger scale and earn higher profits relative to the benchmark economy. The evolution of the employment distribution of three castes over three different economies is shown in Figure 1.7. It is evident that the employment distribution of non-HC firms is skewed toward the left in the benchmark economy, whereas it shifts toward right in the $CF1$ because of improved credit allocation, while no discernible differences remain in the $CF2$ economy.

FIGURE 1.7: Employment Distribution across Castes: Model



(a) Benchmark economy

(b) CF1

(c) CF2

Moreover, removing barriers to external financing for non-HC entrepreneurs has distributional consequences. In $CF2$, within caste wealth inequality will increase as non-HC entrepreneurs accumulate

Table 1.7: Gains in the Counterfactual Economy

	CF1 (%)	CF2 (%)
Overall Economy		
TFP	$+6$	$+10$
$\sigma(arpk)$	-13	-13
$Capital\text{-}Intensity$	$+7$	$+13$
$Output\text{-}per\text{-}worker$	$+8$	$+15$
LC		
$Capital\text{-}Intensity$	$+15$	$+22$
$Output\text{-}per\text{-}worker$	$+2$	$+10$
$\%\text{-}of\text{-}firms$	$+87$	$+81$
MC		
$Capital\text{-}Intensity$	$+13$	$+20$
$Output\text{-}per\text{-}worker$	$+2$	$+10$
$\%\ of\ firms$	-17	-17

more assets; however, cross-caste inequality will decrease. The evidence for this is already provided in Section 1.6.3, where states with well-functioning financial markets exhibit small differences in their asset holdings.

Finally, I perform two more counterfactual exercises to highlight the importance of misallocation at the extensive and intensive margins and disentangle the gains from these two sources. I start with the stationary distribution $\Lambda(z, a, c)$ in the benchmark economy. I then redistribute capital across entrepreneurs such that $arpk$ equalizes across castes, conditional on their financial wealth and productivity, while the distribution of firms, total capital, and labor supply are kept constant. The reallocation of capital from the unproductive HC entrepreneurs toward non-HC entrepreneurs account for 75% of the total TFP gains.

Next, I allow productive non-HC entrepreneurs to enter the market, along with an efficient allocation of capital across castes. These entrepreneurs could not produce profitably before because of stringent financial frictions. In the new steady state, firms as a proportion of population increase by 14%, while labor supply decreases. This, in conjunction with the enhanced borrowing capacity of non-HC firms, creates more demand for capital and labor, thereby increasing factor prices and further improving the selection of entrepreneurs. The TFP gains from removing talent misallocation at the extensive margin represents 25% of the total gains.

Hsieh and Klenow (2009a) argue that if capital and labor were efficiently allocated in India, then TFP would be around 40%-60% higher in the manufacturing sector. Therefore, I conclude that caste-specific distortions in India are important and could account for 15% of the overall gains mentioned in their paper. These results suggest that special attention is needed from policymakers to unleash the entrepreneurial prowess of non-HC individuals, which not only is important from a social justice point of view but also is an economically efficient thing to do. However, caste-specific distortions are not the whole story as far as misallocation in India is concerned. Potentially, other firm-level distortions are present in the Indian economy that drag productivity growth.

1.7 DISCUSSION AND ROBUSTNESS

In the last section, I have discussed potential TFP losses due to $arpk$ dispersion across castes in the context of the model. Some aspects, however, such as imperfect competition and the heterogeneous output elasticity of capital, are not covered in the model. Therefore, in this section, I discuss the potential impact of these forces on the validity of the results presented in Section 1.6 and provide further robustness checks.

1.7.1 *Markup Dispersion*

The model assumes perfect competition in the goods market, which does not allow me to consider product market distortion, which is potentially correlated with $arpk$ dispersion across castes. In principle, markup dispersion could be driven by financial frictions. For example, Chevalier and Scharfstein (1996) and Gilchrist et al. (2017) document a positive correlation between financial constraints and firm markups in times of low demand. In such a situation, the TFP losses mentioned in Table 1.7 are well identified. Furthermore, recent literature on markups, such as De Loecker et al. (2016), documents increasing markups over firm size. Given the fact that LC firms are small in size, the markups should downwardly bias my estimates of $arpk$ differences across castes. However, forces such as selection could drive up the markup for non-HC firms.[43] In what follows, I provide two different pieces of evidence to support my assumption in the baseline model. First, I compute the markups for each caste and include them as controls in the regression model presented in equation 1.1. Second, I compute a quantity based measure of average products that does not include the selling price of the goods produced.

1.7.1.1 Markup Estimation

The markup estimation requires a generalized production function, where the firm produces quantity $Q = AK^{\theta_k}L^{\theta_l}M^{\theta_v}$. The variable A is productivity, K is capital, L is labor, and M is intermediate input. The output elasticities of capital, labor, and material input are denoted by θ_k, θ_l, θ_m, respectively. This production function is different from the one assumed in Section 1.5 in one key aspects: output is in the quantity of the product produced rather than in value-added. Following the seminal paper on markup estimation by De Loecker and Warzynski

[43] Other frictions such as demand segmentation in the presence of imperfect competition generate markup dispersion (Goraya and Ilango 2020).

(2012), I compute firm markups as

$$Markup = \theta_v \frac{sales}{Variable\,cost}, \tag{1.8}$$

where θ_v is the sales elasticity of variable input and $variable\,cost$ is the cost of materials. A full characterization of the markup estimation is provided in Appendix 1.A.6. The output elasticity of materials is computed for each caste at the 4-digit sector classification (see Figure 1.B.11).[44] I use the regression specification in equation 1.1, with firm-level markup μ as an additional control variable. The markup μ is positively correlated with $arpk$;[45] however, it has a negligible effect on caste dummies, particularly in the case of the manufacturing sector (see specifications 3 and 4 in Table 1.8). Further, the $arpk$ differences presented here are remarkably close to the ones identified in the model (see Section 1.6), which reinforces the validity of the results presented above.

1.7.1.2 Measuring Average Product of Capital

In this paper I rely on a revenue based measure of average products, which includes firm-level prices. Here, I exploit one more dimension of the data to compute a quantity-based measure of average products for enterprises that produce only one product.[46] The average product of capital, apk, is defined as follows:

$$apk_{isc} := \; ln(APK_{isc}) = ln(Q_{isc}) - ln(K_{isc}).$$

[44]The output elasticity of material input is evaluated using cost-share methodology. This approach requires a constant returns to scale assumption for the production technology.

[45]The $arpk$ measure is based on gross-output as well to be consistent with the model of markup estimation.

[46]I use single product firms because the data are not detailed enough to compute apk for multi-product firms. In particular, the measure of capital and labor is at the firm level and not at the product level.

Table 1.8: ARPK and Markups

	All sectors		Manufacturing	
	(1)	(2)	(3)	(4)
Dep. Var.	$arpk$	$arpk$	$arpk$	$arpk$
MC	0.0728	0.0657	0.117	0.114
	(0.036)	(0.036)	(0.047)	(0.046)
LC	0.151	0.132	0.201	0.190
	(0.074)	(0.071)	(0.058)	(0.057)
μ		0.0759		0.0448
		(0.026)		(0.034)
Obs (millions)	1.4	1.4	1.0	1.0
R-squared	0.154	0.158	0.170	0.172
State & NIC4 FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓

Notes: Results from firm-level regression, presented in equation 1.1. Sector and state fixed effects are included. Sampling weights are applied. Standard errors are in parentheses.

I exploit the regression model in equation 1.1 with product-level fixed effects. The apk for LC and MC firms is 40% and 24% higher than that of HC firms (see Table 1.9). These estimates are larger than those observed while using $arpk$.

1.7.1.3 Output Elasticity of Capital

Using the framework from the previous section, I am able to compute the output elasticity of capital θ^k for each caste within a sector. In Figure 1.B.12 in Appendix 1.A, I compared the values of θ^k for LC and MC firms against those of HC firms and found no systematic bias.

Further, $arpk$ in this framework can be decomposed as follows;

$$arpk_{isc} = \mu_{isc} - \widehat{\theta}^k_{sc} + \tau_{isc} + constant,$$

Table 1.9: Average Product of Capital

Dep. Var.	All-Sectors apk	Manufacturing apk
MC	0.280	0.236
	(0.094)	(0.119)
LC	0.407	0.400
	(0.119)	(0.165)
Obs.	97,913	87,184
R-squared	0.553	0.383
State & Product FE	✓	✓
Controls	✓	✓

Notes: Results from firm-level regression, presented in equation 1.1. Dependent variable is average product of capital. The product and state fixed-effects are included. The sampling weights are applied and standard error are in parentheses.

where μ_{isc} is firm-level markup as defined above, $\widehat{\theta}^k_{sc}$ is log of output elasticity of capital, and τ_{isc} is firm-level distortions. For example, in the presence of only financial frictions, it represents the Lagrange multiplier, which increases as the firm becomes more constrained. The decomposition pins down a negative relation between $arpk$ and the output elasticity of capital; that is, those who use capital-intensive technologies should have a lower $arpk$, all else equal. Under the assumption that $\widehat{\theta}^k_{sc}$ is not correlated with τ_{isc}, I control for the output elasticity of capital using the regression model mentioned in equation 1.1.[47] I find no effect on the estimates of LC dummy but document negative effects on estimates of the MC dummy (see Table 1.B.8 in Appendix 1.A).

[47]In models of technology adoption in the presence of financial frictions, this assumption breaks down, and therefore the estimates of caste-dummies in Table 1.B.8 in Appendix 1.A will be biased downward.

1.8 Conclusion

It is well established that misallocation of resources can explain a large chunk of productivity differences across countries. However, its sources are still disputable and several firm-level distortions and frictions have been proposed. This paper suggests that the caste system in India is one example of such distortions and quantifies its importance in explaining aggregate TFP losses, as mentioned in Hsieh and Klenow (2009a).

This paper takes a different perspective in dissecting firm-level data in India. Instead of using firm performance measures, I use the caste of the firm owner as a defining feature. I document a large dispersion in $arpk$ across firms of different castes, whereas no such dispersion is visible for $arpl$. Further, contrary to the previous literature that documents a high $arpk$ for large firms in India, I find that firms owned by historically disadvantaged castes while small in size, exhibit a high $arpk$.

I then use a quantitative model of entrepreneurship, based on Buera and Shin (2013), to decompose the effects of fundamentals such as productivity and technology, as well as the availability of external financing, on the cross-caste $arpk$ dispersion. The model identifies a very high degree of financial constraints for non-HC entrepreneurs. Meanwhile, the productivity distribution is characterized by a lower dispersion, and the scale of production technology is smaller for LC and MC individuals relative to HC individuals.

In this paper, I exploit the heterogeneity in financial development across various states in India to identify the impact of limited access to credit on the performance of firms owned by non-HC agents and its overall welfare implications for the non-HC population. The $arpk$ difference vanishes over regional financial development. Meanwhile, the welfare of non-HC individuals increases substantially; in particular, household consumption and asset holdings converge toward the level of HC individuals.

I use the model to perform various counterfactual experiments. First, I homogenize the degree of financial frictions across castes, which delivers TFP gains of 6%. Second, an additional 4% of TFP gains are realized when I impose the productivity process and technology of HC individuals on that of non-HC individuals. In the counterfactual economy, gains come from three sources. The first source is the improved allocation of capital across castes at the intensive margin. The second is the improved selection of entrepreneurs at the extensive margin, particularly the entry of productive but poor non-HC entrepreneurs and the exit of unproductive but wealthy HC entrepreneurs. The third source is the use of large scale production technology and an improved productivity distribution, which allows non-HC firms to operate at a larger scale and earn higher profits relative to the benchmark economy.

Given the findings of this paper, a natural next step would be to understand the implication of the caste system on long-run growth. Furthermore, understanding the causes of productivity and technological differences across castes is important for establishing well-guided policies and therefore is a promising avenue for future research.

1

APPENDIX

1.A DATA AND MEASUREMENT

1.A.1 *MSME Dataset*

The MSME census is based on MSME sector which is defined by the Micro, Small and Medium Enterprise Developmemt (MSMED) act of 2006, spans the non-agricultural enterprises of the economy that are below a certain threshold of size (size in terms of original value of investment in plant of machinery). The investment limit for enterprises engaged in the manufacturing or production of goods is Indian rupees (INR) 100 million whereas for those providing or rendering in services is INR 50 million. According to the 4th MSME census of India 2006, the MSME sector accounts for 41% of the manufacturing output and 40% of the total exports of the country.[48] The sector is estimated to employ about 59 million individuals in over 26.1 million units throughout the country. Further, 1.5 million (5.94%) are registered MSMEs and 24.5 mil-

[48] These statistics are mentioned in MSME Annual report 2010-11 (Page 211), `https://msme.gov.in/relatedlinks/annual-report-ministry-micro-small-and-medium-enterprises`

lion (94.06 %) are unregistered MSMEs that employ 16.62 % and 83.38 % of the workforce respectively. Overall, 29 % of them are manufacturing and 71 % are service enterprises and provide employment to 51% and 49 % of the total labor force (in the MSME sector) respectively. The Scheduled Castes and Scheduled Tribes (LC), OBC's (MC) and Others (HC) own and operate 2.9 (11 %), 10.4 (40 %) and 11.4 (44 %) million MSMEs.

Unlike ASI and Prowess datasets, the economic census and the MSME datasets are able to capture small enterprises that are more likely to face financially constraints. Such effects may go unnoticed in datasets with predominantly large enterprises. Meanwhile, in the absence of large enterprises, this dataset may also upward bias the effect of caste differences. It could be that, in the overall economy, the share of such constrained enterprises is minuscule and hence caste specific frictions do not matter. I take into account such concerns while discussing the empirical results and calibration strategy and try to minimize such biases.

The measure of *profitability*, which is defined as the ratio of profits to value-added.The profitability is defined as $\pi_i = \frac{Y_i - RK_i - wL_i}{Y_i}$, where R is the cost of capital interest rate and assumed to be 5.682%. In my data, there are many observation with negative profitability. I use a IHS transformation of the profits as suggested in Bellemare et al. (2013). I find the low-caste entrepreneurs to be 9% higher profitability relative to the high-caste entrepreneurs. Such evidence suggests that very selected low-caste agents are entering the market.

1.A.2 *Economic Census 2005*

The 5th Economic census in 2005 covered agricultural (excluding crop-production and plantation) and non-agricultural activities within the geographical boundary of India. In total, there are 42 million enterprises employing 99 million individuals. The manufacturing and services sectors represent 84.7 % of all the enterprises that employ 88.5 % of the total labor force. As far as the caste-based firm ownership is concerned, the Scheduled Castes and Scheduled Tribes (LC) own and operate 5.67 million of the firms, the middle caste (MC) operates more than 18 million of them and, similarly, 18 million of the enterprises are owned by the high caste (HC).[49]

The enterprise ownership across castes is measured with the Economic census of 2005. The caste of the private enterprise is identified with the caste of its owner (public firms are dropped). I use the population census of 2001 and the National Sample Survey 66th Round 2009-10 to compute the low caste and the middle caste population shares respectively. The first two columns of Table 1.B.2 show that the low-caste individuals represent 24% of the total population, while they only own 13 % of all non-agricultural enterprises. Moreover, as shown in columns 3 and 4, low caste individuals own 14 % of the single employee enterprises, 1 percentage point higher than their overall ownership, and own 10% of the enterprises that hire labor outside of their family. In terms of employment, column 5, low castes employ around 11% of the total labor force.

The entrepreneurship intensity is measured by the ratio of share of enterprises of a certain caste group to its share in the population. In 2005, entrepreneurship intensity was 0.57, 1 and 1.3 for LC, MC and HC respectively. Given that, as argued in the literature (Deshpande et al. 2013), self-employment can be more of a survival activity rather than entrepreneurship, I also compute the entrepreneurship rates excluding single employee enterprises. Then, the entrepreneurship intensity is 0.46, 0.96 and 1.43 for LC, MC and HC respectively. While entrepreneurship intensity is significantly lower than one for low caste agents in

[49]Following Iyer et al. (2013), I keep 19 large states of India that constitute 95 % of all the enterprises and 96 % of the population.The states include:- Andhra Pradesh, Assam, Bihar, Chhattisgarh, Gujarat, Haryana, Himachal Pradesh, Jharkhand, Karnataka, Kerala, Maharashtra, Madhya Pradesh, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, Uttarakhand and West-Bengal.

all the states, there are some regional differences: Assam (1.06), West Bengal (0.79), Odisha (0.79), Himachal Pradesh (0.70) and Maharashtra (0.69) are the states with the highest entrepreneurial rate whereas Gujarat (.31), Jharkhand (0.34), Bihar (0.40), Rajasthan (0.45) and Madhya Pradesh (0.45) are the lowest.

1.A.3 *ASI-NSS 2006*

The firm-level dataset for the manufacturing sector India is provided by Annual Survey of Industries(ASI),which covers registered manufacturing. However, this dataset does not include small firms or unregistered firms. In particular, according to India's Factories Act of 1948 as explained in Section 3.2, establishment with more than 10 workers, in case they use electricity and 20 workers, in case they do not use electricity are required to registered. Hence, the ASI provides a truncated size distribution. I use National Sample Survey (NSS) 2006, which covers production units in the unorganized sector in India. The employment distribution is provided in table 1.B.5.

1.A.4 *Household Surveys 2005 & 2012*

IHDS surveys are household surveys that includes information on consumption, assets, wages, business-income, Desai et al. (2018). The Income distribution is provided in table 1.B.6.

1.A.5 *Winsorization*

The financial variable such as market value of fixed assets, gross value-added, total wage-bill, employment, amount of loan-outstanding, gross output, total cost of variable inputs and net-worth are winsorized at 1 and 99th percentile. Furthermore, the variables used in regressions $arpk$, $arpl$, $tfpr$, k/l are winsorized at 1 and 99th percentile.

1.A.6 *Markup Estimation*

Consider the following cost minimization problem;

$$\begin{aligned} \min_{K,LM} \quad & RK + wL + P^m M + \kappa \\ \text{s.t.} \quad & Q = AK^{\theta_k} L^{\theta_l} M^{\theta_v}, \ and \ \sum_j \theta_j = 1, \end{aligned} \tag{1.9}$$

where, Q is quantity produced, K is capital, L is labor, M is intermediate input. Further, total cost for firm is composed cost of capital RK, where R cost of capital and assumed to be $r + \delta$, i is the interest rate and δ is the depreciation rate; wL is wage-bill; $P^m M$ cost of materials with P_m being firm-specific purchasing price; and κ is fixed cost of operating. The markup estimation does not require cosntant returns o scale assumption, however, it is necessary to estimate output elasticities as discussed below.[50]

Further, it is assumed that capital K is chosen in the presence of frictions, including markups; and material and labor choices are undistorted except for the markup.[51] The markup can be computed, using cost of any input, as long as it is not fixed. Therefore, I use material input as labor and capital are quasi-fixed in the Indian context.

The first order conditions form cost minimization problem defined above gives me markup μ_{isc}, of firm i in sector s with owner of caste c that is equal to;

$$\mu_{isc} = \theta^v_{sc} \frac{P_{isc} Q_{isc}}{P^m_{isc} M_{isc}} \tag{1.10}$$

I use cost share technique to compute elasticities, see Foster et al. (2008). This technique requires all inputs to be free, however averaging out across the sample can get rid of this concern. Moreover, one need to specify cost of capital. I use $r = 0.0568$ and $\delta = 0.06$. Finally, it requires constant returns to scale technology. The cost share is;

[50] Other popular approches that are available to estimate output elasticities demand panel data.

[51] This method allows for any distortion in the input market.

$$\theta^v_{isc} = \frac{P^m_{isc} M_{isc}}{P_{isc} M_{isc} + wL_{isc} + RK_{isc}} \tag{1.11}$$

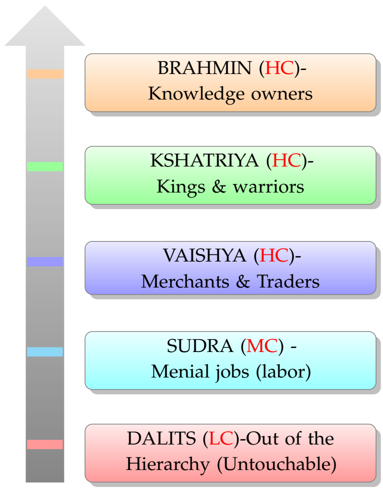
The sectoral output elasticities of material inputs are computed as $\theta^v_{sc} = median_{i \in s}\{\theta_{isc}\}$. The comparison of θ^v_{sc} across caste is presented in figure 1.B.11. No systematic bias is evident from the respective estimates. The $arpk$, in this setting, is defined as;

$$arpk_{isc} := \ ln(ARPK_{isc}) = ln(P_{isc}Q_{isc}) - ln(K_{isc}).$$

1.B FIGURES AND TABLES

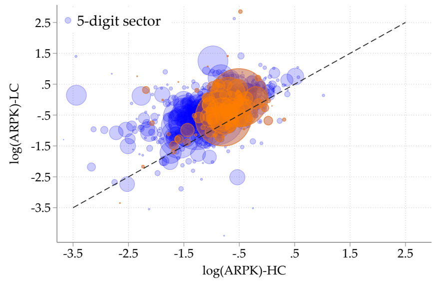
1.B.1 *Figures*

FIGURE 1.B.1: The Caste System

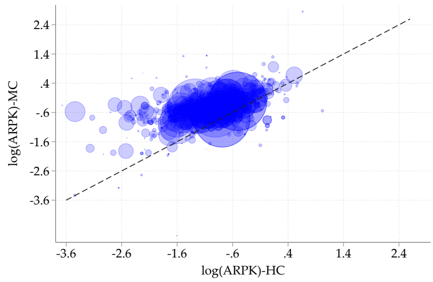


1.B.2 *Tables*

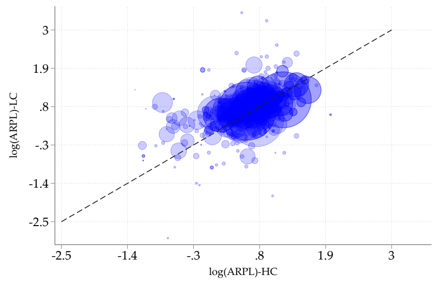
FIGURE 1.B.2: ARPK & ARPL: Data



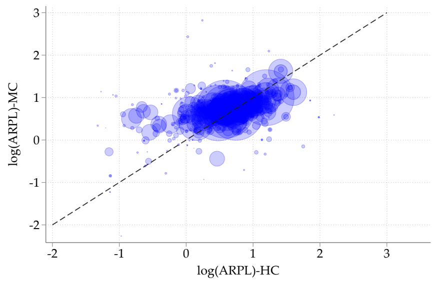
(a) ARPK-LC



(b) ARPK-MC



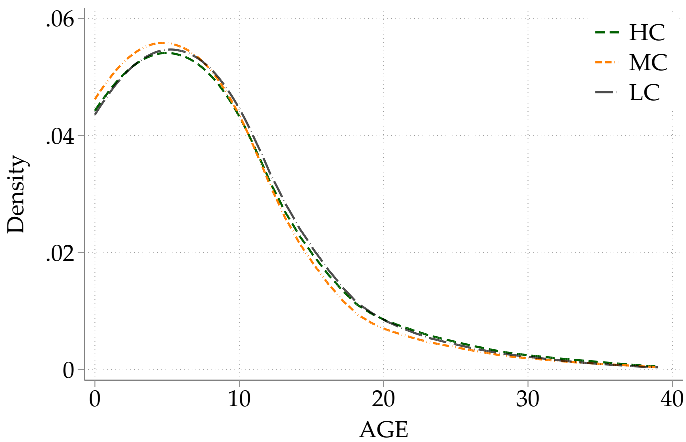
(c) ARPL-LC



(d) ARPL-MC

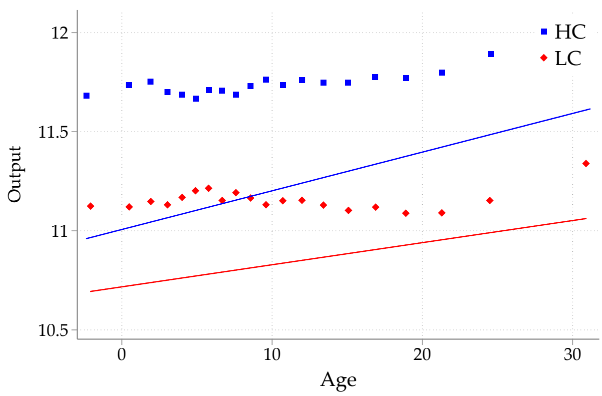
Note. Each blue circle represents a 5 digit sector (633 in total). The orange circles represent sectors such as food products and beverages (NIC-15), tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear, apparels or furniture (NIC-18,19). Sampling weights are applied.

Figure 1.B.3: Age Distribution



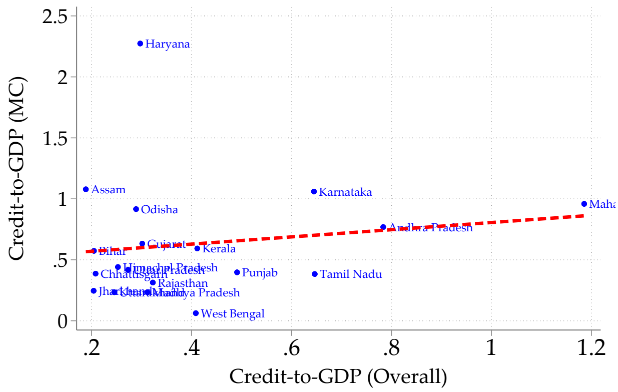
Note. sampling weights applied.

Figure 1.B.4: Output and Age: Data



Notes: Binscatter plot with age on x-axis and output on y-axis. Each square and diamond represent mean of the x-axis and y-axis variables within equally-sized bin of variable in x-axis for HC and LC firms, respectively. Sector and state FE included. sampling weights applied.

FIGURE 1.B.5: Regional Financial Development-Data



Notes: The overall credit-to-GDP ratio is computed with statistics published by the Reserve Bank of India. The credit-to-GDP ratio for MC firms is computed in MSME data. Sampling weights are applied

FIGURE 1.B.6: Finance and TFPR: Model

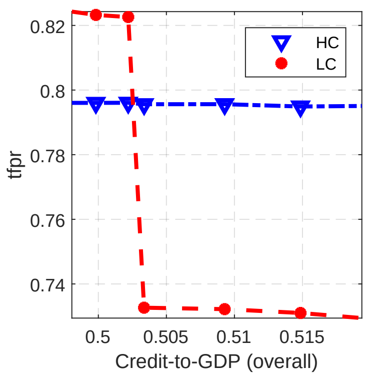
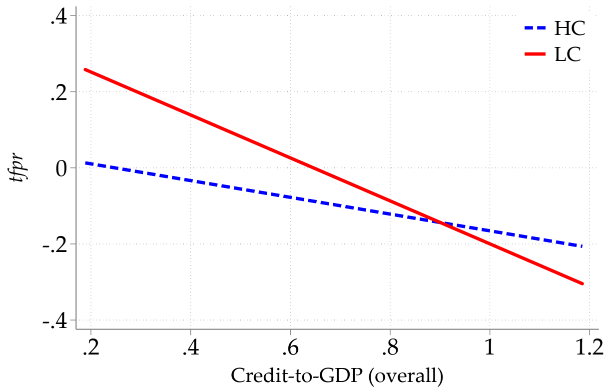
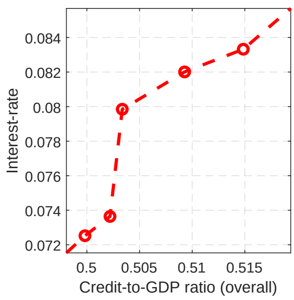


FIGURE 1.B.7: Finance and TFPR: Data

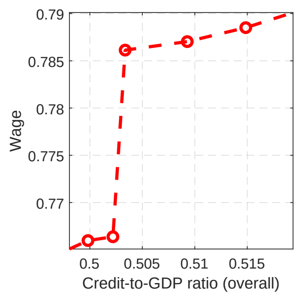


Notes: Linear-regression fit plot with credit-to-output ratio on x-axis and $tfpr$ on y-axis. Sector and state FE included. sampling weights applied.

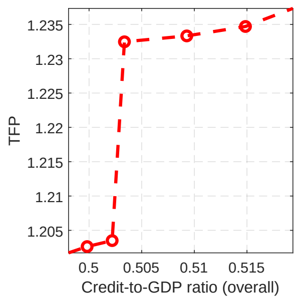
FIGURE 1.B.8: Finance and Regional Development-Model



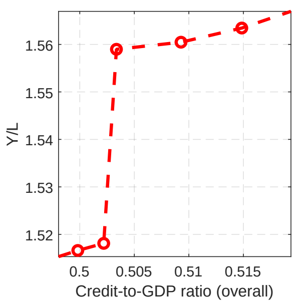
(a) Interest-Rate



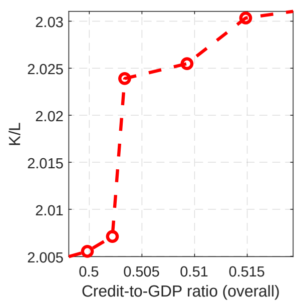
(b) Wages



(c) TFP

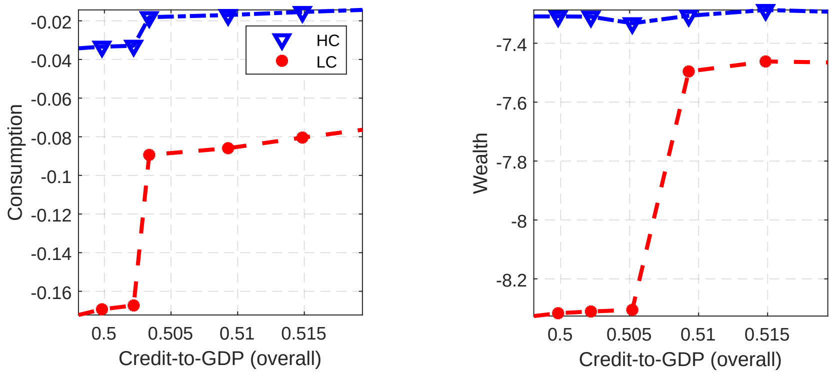


(d) GDP-per-Capita



(e) Capital-Intensity

FIGURE 1.B.9: Financial Development & LC Households-Model

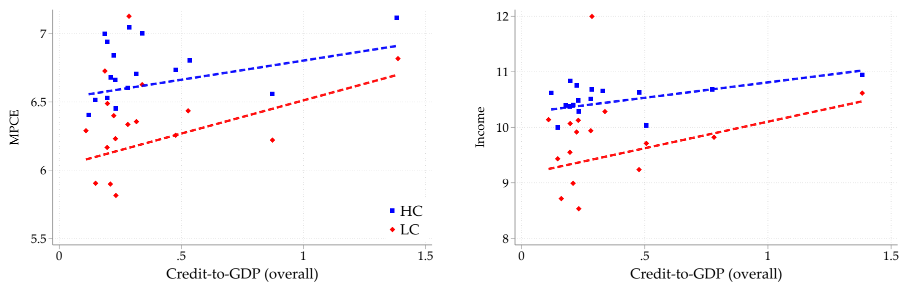


(a) Comsumption

(b) Financial Wealth

Note. Coefficients of the low caste dummy from regressions of log(k/l) (column 1) and log(MRPK) (column) using specification 2 and 5 in Table **??** for each age bin on the X-axis. Rows represent the employment bins.

FIGURE 1.B.10: Financial Development & LC Households-Data

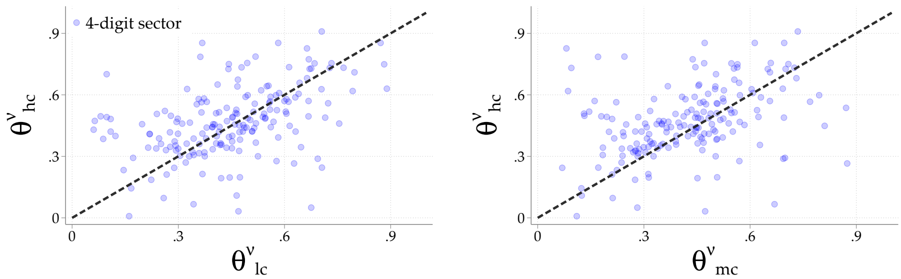


(a) Per-capita Comsumption

(b) Household Assets

Notes: Binscatter plot with credit-to-output ratio on x-axis and y-axis; (a) MPCE and (b) Household Assets . Each square and diamond represent mean of the x-axis and y-axis variables within equally-sized bin of variable in x-axis for HC and LC firms, respectively. sampling weights applied.

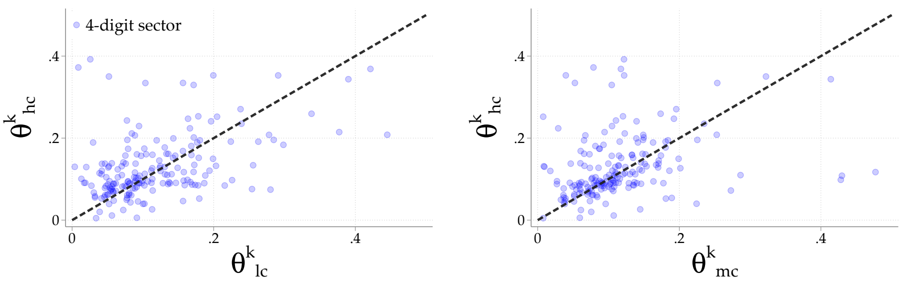
FIGURE 1.B.11: Output elasticity of Variable Input-Data



(a) LC (b) MC

Note. Each blue dot represent a 4-digit sector. θ^{ν} is the output elasticity of material input. The respective subscripts represent caste. Sampling weights applied.

FIGURE 1.B.12: Output elasticity of Capital-Data



(a) LC (b) MC

Note. Each blue dot represent a 4-digit sector. θ^{k} is the output elasticity of capital input. The respective subscripts represent caste. Sampling weights applied.

Table 1.B.1: $arpk$ and Financial Development

VARIABLES	(1) HC	(2) MC	(3) LC
Fd_s	-0.0746	-0.411	-0.759
	(0.153)	(0.099)	(0.152)
Constant	-0.0563	0.261	0.598
	(0.075)	(0.049)	(0.061)
Observations	719,313	547,316	134,561
R-squared	0.101	0.101	0.135
Control	✓	✓	✓
SIC3 FE	✓	✓	✓

Notes: Results from the enterprise level regression. Dependent variables are in logs and shown in column headings. Fd is index of financial development across states. The vector of controls, Γ, that includes region, gender and religion FE. Robust standard errors in parentheses. Clustered at sector level. Sampling weights are applied.

Table 1.B.2: Share of Population and Non-agricultural Enterprises across castes

	(1)	(2)	(3)	(4)	(5)
			Enterprises with		
Caste	Population	Enterprises	One employee	Outside labor	Employment
LC	29%	13%	14%	10%	11%
MC	35%	43%	44%	10%	39%
HC	36%	44%	42%	50%	50%

Notes: The enterprise ownership rates are computed with non-agricultural enterprises in the Economic census 2005. The population statistics for the low- and middle-caste are drawn from IHDS 2005 survey. Outside labor means labor outside the household.

Table 1.B.3: $arpk$ in Manufacturing Sector: Data

	(1)	(2)	(3)
Dep. Var.	$arpk$	k/l	$tfpr$
MC	0.223	-0.313	0.239
	(0.054)	(0.060)	(0.042)
LC	0.340	-0.582	0.245
	(0.073)	(0.081)	(0.069)
Constant	-0.173	10.54	2.926
	(0.030)	(0.037)	(0.025)
Observations	975,983	975,983	939,459
R-squared	0.176	0.263	0.115
State FE	✓	✓	✓
NIC4 FE	✓	✓	✓
Controls	✓	✓	✓

Notes: Results from the enterprise level regression using equation 1.1 and 1.2. Dependent variables are in logs and shown in column headings. $arpk$ and $arpl$ are average products of capital and labor, respectively. M-caste is the dummy variable for the middle-caste enterprises. L-caste is the dummy variable for the low-caste enterprises. The vector of controls, Γ, that includes region, gender and religion FE. Robust standard errors in parentheses. Clustered at Caste, region and sector level. Sampling weights are applied.

Table 1.B.4: Employment Distribution-Economic Census 2005

Firms	p50 (%)	p75 (%)	p90 (%)	p95 (%)	p99 (%)
All	0.36	0.36	0.53	0.64	0.74
HC	0.24	0.32	0.48	0.56	0.68
MC	0.21	0.45	0.63	0.72	0.81
LC	0.23	0.54	0.66	0.78	0.86

Table 1.B.5: Employment Distribution-ASI & NSS 2006

Firms	p50 (%)	p75 (%)	p90 (%)	p95 (%)	p99 (%)
All	0.42	0.42	0.60	0.64	0.76

Table 1.B.6: Income Distribution-Economic Census 2005

Firms	p50 (%)	p75 (%)	p90 (%)	p95 (%)
IHDS 2005				
All	0.13	0.33	0.55	0.68
HC	0.17	0.36	0.56	0.68
MC	0.14	0.36	0.60	0.71
LC	0.13	0.34	0.61	0.72
IHDS 2012				
All	0.13	0.31	0.50	0.63
HC	0.12	0.29	0.49	0.60
MC	0.14	0.33	0.55	0.65
LC	0.15	0.37	0.63	0.73

Table 1.B.7: Parameter Values

Parameter	BM-Value	CF1-Value	CF2 -Value	Description
Fixed:				
δ	0.060	0.060	0.060	Annual depreciation rate physical capital
α	0.330	0.330	0.330	Physical capital share
γ	1.500	1.500	1.500	Coefficient of risk aversion
ρ	0.844	0.844	0.844	Discount factor
Fitted:				
$1-\nu_h$	0.761	0.761	0.761	Span of control for HC
$1-\nu_m$	0.745	0.745	0.761	Span of control for MC
$1-\nu_l$	0.745	0.745	0.761	Span of control for LC
ψ_h	0.927	0.927	0.927	Persistence in productivity for HC
ψ_m	0.922	0.922	0.927	Persistence in productivity for MC
ψ_l	0.918	0.918	0.927	Persistence in productivity for LC
λ_h	1.760	1.760	1.760	Degree of financial frictions for HC
λ_m	1.370	1.760	1.760	Degree of financial frictions for MC
λ_l	1.160	1.760	1.760	Degree of financial frictions for LC
η_h	4.520	4.520	4.520	Scale parameter of ability distribution for HC
η_m	4.700	4.700	4.520	Scale parameter of ability distribution for MC
η_l	4.890	4.890	4.520	Scale parameter of ability distribution for LC
c_f	0.100	0.100	0.100	Fixed cost of Operation

Table 1.B.8: ARPK and Markups

	All sectors			Manufacturing		
Dep. Var.	$arpk$	$arpk$	$arpk$	$arpk$	$arpk$	$arpk$
MC	0.0728	0.0674	0.0349	0.117	0.115	0.0779
	(0.036)	(0.036)	(0.038)	(0.047)	(0.047)	(0.046)
LC	0.151	0.132	0.189	0.201	0.188	0.204
	(0.074)	(0.071)	(0.061)	(0.058)	(0.057)	(0.056)
mu		0.0769	0.0785		0.0463	0.0477
		(0.026)	(0.026)		(0.034)	(0.034)
θ^k_{sc}			-2.886			-3.615
			(0.912)			(0.978)
Constant	0.453	0.418	0.762	0.477	0.459	0.839
	(0.026)	(0.029)	(0.116)	(0.026)	(0.029)	(0.105)
Obs (millions)	1.4	1.4	1.4	1.0	1.0	1.0
R-squared	0.154	0.158	0.162	0.170	0.172	0.176
State & NIC4 FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓

Notes: Results from firm-level regression, presented in 1.1. Sector and state fixed-effects are included. Sampling weights applied. Standard error in parentheses.

2

THE RISE OF INTANGIBLE CAPITAL AND THE MACROECONOMIC IMPLICATIONS

Joint with Andrea Chiavari

2.1 INTRODUCTION

In the last decades, new technological improvements have reshaped the production process of US firms. Nowadays, investments in research and development, intellectual property products, and computerized information − commonly known as intangible capital − account for more than 30% of aggregate investment.This novel capital shows different characteristics compared to the tangible capital, such as equipment and structures.Specifically, it is usually immaterial, specific to the firm that uses it, and often internally produced rather than acquired.The rise of intangible capital, with its unique characteristics, has ramifications for competition, for antitrust policy, for allocative efficiency, and hence for economic well-being more broadly.

Despite the rising importance of intangible capital, we know little about its properties and the implied consequences of its rise.In this article, our goal is to document the properties of intangible capital and to shed light on the macroeconomic implications of its rise as an input in production for US firms.First, exploiting firm-level data, we document the changing nature of the production technology. We show that the input share of intangible capital has seen a sizeable increase since the 1980s, going from approximately 0.03 to 0.12. This rise has come at the expense of the labor share in production. We label this phenomenon *Intangible Capital Biased Technological Change* (IBTC).

Then, we provide novel empirical evidence on the behavior of investment in intangible capital.In the data, we see that the firm-level intangible capital investment process is lumpier compared to tangible capital investment, as it is characterized by long periods of inaction and by a high serial correlation.To rationalize these empirical findings, we use a general equilibrium model of firms and investment dynamics enlarged with intangible capital.We also allow for a flexible specification of adjustment costs associated with the investment process of both capitals.The model attributes higher adjustment costs − particularly fixed adjustment costs − to intangible capital investment relative to tangible capital investment.These findings confirm the view that intangible capital investment is inherently different from tangible capital investment.For example, the implementation of the Just In Time (JIT) production process by the US manufacturing sector required large investments beforehand, due to the presence of inherent indivisibilities in its implementation, and took long set up times, due to workers' retraining and the restructuring of the production procedures.[1]

[1] The JIT production process, pioneered by Japanese manufacturers, gained momentum among US firms in the 1980s. Nakamura, Sakakibara, and Schroeder (1998) wrote "Because so many different aspects of plant operation are involved, the transfer of JIT requires a substantial effort on the part of U.S. manufacturers" and Fullerton, McWatters, and Fawson (2003) wrote "Investment returns from JIT adoption are not immediately observable, due to the long-run nature of its implementation process."

Finally, we use the calibrated model to quantify the effects of the IBTC on changes that the US economy experienced in the average firm size, concentration, aggregate factor shares, tangible capital investment rate, profit rate, and allocative efficiency of the economy.Our model shows that the IBTC can go a long way in explaining the above transformations, highlighting the central role of intangible capital for recent macroeconomic events.Moreover, our findings suggest that a significant fraction of these transformations are the outcome of the efficient response of the economy to changes in firm-level production technology.

The elusive nature of intangible capital makes its measurement a difficult task. Corrado, Hulten, and Sichel (2009), Corrado and Hulten (2010), McGrattan and Prescott (2010b), McGrattan and Prescott (2014), and Koh, Santaeulàlia-Llopis, and Zheng (2020) have made large strides forward in documenting the aggregate value of intangible capital in the US economy.However, we are interested in the microeconomics properties of this capital, and consequently, we take a different route compared to the above papers: We leverage the Compustat dataset, which encompasses all US publicly traded firms, where we can observe the balance sheet and different expenditure invoices reported by these firms. However, the lack of US GAAP in fully accounting for intangible capital in the firms' balance sheet makes the firm-level measurement of intangible capital complicated, as reported by Lev and Gu (2016). In particular, our baseline measure of intangible capital, between 1980 and 2015, will be made of two components: (i) Internally generated intangible capital, through research and development expenditure, and (ii) identifiable intangible capital booked in the balance sheet. Finally, to validate our measure, we compare it with the one provided by Koh, Santaeulàlia-Llopis, and Zheng (2020) which yields similar trends to ours.

In this paper, using the above firm-level data, we make three main contributions. First, we estimate an augmented firm-level Cobb-Douglas production function with three inputs: Tangible capital, intangible capital, and labor. To do so we follow two approaches: (i) We follow the empirical industrial organization literature and adopt the control function approach developed by Olley and Pakes (1996), Levinsohn and Petrin (2003) and Ackerberg, Caves, and Frazer (2015) and (ii) we use the cost shares approach adopted by Foster, Haltiwanger, and Syverson (2008).All these methodologies find that intangible capital is an important factor in production and whose importance increased from 0.03 in 1980 to 0.12 in 2015. Moreover, most of this rise happened at the expense of the labor input in production. This finding is robust to different estimation techniques, production function specifications, and levels of disaggregation. We interpret this finding as a technological transformation that US firms are experiencing, where intangible capital is becoming a more prominent input in the production process at the cost of labor. We label this as *Intangible Capital Biased Technological Change* (IBTC).

Second, to study the properties of the investment process of intangible capital, we build a model of firms and investment dynamics in the spirit of Clementi and Palazzo (2016b).In the model, firms use a Cobb-Douglas production function using tangible capital, intangible capital, and labor.Moreover, the model features entry and exit of firms and a flexible structure of investment adjustment costs for both capitals. In particular, the adjustment costs associated with both capitals have two components:(i) A convex cost disciplining the intensive margin of investment and (ii) a fixed cost disciplining the extensive margin.Therefore, the model allows us to identify the technological differences in the investment process of both capitals.

The predictions of the model related to the investment process of both tangible and intangible capital depend on the precise identification of the two sets of parameters that discipline the convex and the fixed costs associated with the investment process.Following the seminal papers of Cooper and Haltiwanger (2006) and Asker, Collard-Wexler, and De Loecker (2014), we use inaction rates, defined as investment between $\pm 1\%$, to identify fixed costs of adjusting each type of capital.This moment is informative because higher fixed costs of adjustment increase the inaction rate, as firms prefer not to invest instead of paying these costs.Then, we use the autocorrelation of the investment rate process to discipline the convex costs associated with both capitals.High convex costs make the investment process serially correlated, forcing the firms to accumulate capital slowly.[2]

The calibrated model finds substantial differences in the investment process of these two types of capital, with intangible capital having higher adjustment costs compared to tangible capital $-$ particularly, fixed adjustment costs. This captures, to a certain extent, the notion of sunkness as described in Haskel and Westlake (2018):

> *If a business makes an intangible investment and later on decides it wants to back out, it's often hard to reverse the decision and try to get back the investment's cost by selling the created asset – and in general, it's harder than in the case of a tangible asset. [...] The [...] reason tangible investments are easier to sell is that they are less likely to be uniquely linked to the firm that owns them and its business.*

In the model, the high adjustment costs associated with intangible capital make this input slower to adjust relative to tangible capital when productivity shocks hit. To validate this prediction, we estimate the elasticity of the average revenue product of both capitals to productivity shocks.We find that this elasticity is higher for intangible capital, making the average revenue product of intangible capital, $ARPK_I$, more

[2] It is important to notice that these parameters are jointly calibrated. Moreover, the presence of counterbalancing forces since high fixed costs decrease the autocorrelation of the investment rates, whereas, high convex costs increase it is crucial for the correct identification of these two costs.

responsive relative to the average revenue product of tangible capital, $ARPK_T$.Moreover, we also document in the data that, consistently with the model predictions, the $ARPK_I$ is more dispersed relative to the $ARPK_T$ in most of the sectors.

As a final validation exercise, we exploit cross-sector variation in intangible intensity to provide reduced-form support for the model mechanism.In particular, we look at the intangible investment share, the tangible investment share, the labor share, the profit rate, industry concentration, and allocative efficiency, which are the objective of analysis when we study the consequences of IBTC.Overall, we find that the model captures the qualitative features of the data implying magnitudes of cross-sectoral correlations in line with the ones in the data.

In our third contribution, we use the calibrated model to quantify the effect of IBTC. Holding all other parameters fixed, we ask what are the effects of increasing the intangible capital share while decreasing the labor share, as in the data? We fix the firm-level returns to scale in production, thus isolating the role of shifting the input composition in production as a distinct factor relative to the returns to scale with which firms operate. While changes in the returns to scale in production are of interest on their own, as studied by Lashkari, Bauer, and Boussard (2018) and Chiavari (2021), much less is known about the effect of a change in the input shares in production, as the one documented in this project. This motivates our focus.

The IBTC can quantitatively explain most of the increase in the average firms' size and industry concentration, as measured by Herfindahl-Hirschman Index and by the employment share of firms with more than 250 employees (see Hopenhayn, Neira, and Singhania 2018), as observed in the data. This happens because, while intangible capital becomes more important in production, firms rely more on an input that entails higher adjustment costs, and as a result, the value of entry decreases, pushing up the threshold productivity of the marginal entrant. This implies that a relatively small but more productive mass of firms operate in the economy; thus, the average incumbent size increases. Furthermore, IBTC makes the growth of small firms costly, as they have to incur very high adjustment costs to build their stock of intangible capital, while it makes it easier for large firms to shrink, as the high depreciation rate of intangible capital favors its depletion. This mechanism, together with the above increase in selection, tilts a reallocation of sale shares towards the larger firms, leading to the rise in average firm size and industry concentration.

Further, through the lens of the model, we see that IBTC explains also most of the changes in the aggregate factor shares that have been emphasized in the literature. It explains the increase in the intangible capital share and the decline in both tangible capital and labor shares. This happens because micro-level technological change also affects the aggregate demand for each of the three inputs, favoring intangible capital in particular. Moreover, consistently with the findings in Koh, Santaeulàlia-Llopis, and Zheng (2020), we find that the labor share declines much less if intangible capital would be expensed instead of being capitalized.[3] Finally, as the selection process increases and only more productive firms operate in the market, we see an increase in the firm-level profit rate of a magnitude consistent with what emphasized in De Loecker, Eeckhout, and Unger (2020) and Barkai (2016).

Moreover, we find that IBTC can explain half of the decline in the tangible capital investment rate documented by Hall (2015), Gutiérrez and Philippon (2016), and Crouzet and Eberly (2019). This happens because firms with this new intangible-intensive technology tilt a greater share of their expenditure towards intangible capital. As a consequence investment in tangible capital declines in the new steady state. Hence, the model interprets half of the slack in the investment rate in the tan-

[3] Barro (2019) and Atkeson (2020) also show that part of the decline in the corporate non-financial labor share is due to the accounting procedures used by the Bureau of Economic Analysis of the US Department of Commerce.

gible capital as the by-product of a technological change that makes tangible capital a less relevant input in production.

Finally, the quantitative model shows that the IBTC can explain between 32% and 80% of the overall decline in allocative efficiency of the US economy as documented by Bils, Klenow, and Ruane (2020). This is driven by the fact that $TFPR$ in our framework is a weighted geometric mean of the average revenue product of inputs, where the weights are proportional to their output elasticities. Due to the presence of adjustment costs, dispersion in $TFPR$ is driven by dispersion in average products of both capitals. When the output elasticity of intangible capital increases, the dispersion in $ARPK_I$ becomes the primary driver of the dispersion in $TFPR$.[4] Therefore, dispersion in $TFPR$, which is our measure of allocative efficiency (where higher dispersion in $TFPR$ means lower allocative efficiency), increases. However, in our framework dispersion in $TFPR$ − as noticed already by Asker, Collard-Wexler, and De Loecker (2014) − cannot be interpreted as misallocation as done in Hsieh and Klenow (2009b) because the allocation still coincides with the planner ones.

Related Literature. This paper is related to the rising literature that measures intangible capital at the aggregate level as in Corrado, Hulten, and Sichel (2009), Corrado and Hulten (2010) McGrattan and Prescott (2010a), McGrattan and Prescott (2010b), McGrattan and Prescott (2014), and Koh, Santaeulàlia-Llopis, and Zheng (2020), and at firm-level as in Peters and Taylor (2017) and Ewens, Peters, and Wang (2019). Relative to them, we structurally estimate a Cobb-Douglas firm-level production function augmented with intangible capital. We document that intangible capital is an important input in production which is rising over time at the expense of labor.

[4] Adjustment costs associated with the investment process of input do not allow its marginal product to equalize across firms and hence generate dispersion in the average revenue product.

Furthermore, our paper is related to that extensive literature which examines lumpy investment dynamics pioneered by Abel and Eberly (1994), Abel and Eberly (1996), Doms and Dunne (1998) and Cooper and Haltiwanger (2006), highlighting the role of non-convex adjustment costs in the firm-level investment process. To the best of our knowledge, we are the first to highlight the presence of higher adjustment costs associated with the investment process of intangible capital relative to tangible capital.

Finally, Lashkari, Bauer, and Boussard (2018), Aghion, Bergeaud, Boppart, Klenow, and Li (2019), De Loecker and Mongey (2019), Hsieh and Rossi-Hansberg (2019), and Chiavari (2021) present different mechanisms, all associated with technological factors, behind some of the macroeconomic trends emphasized in this paper. De Ridder (2019) and Zhang (2019) both emphasized the role of intangible capital as a driving factor behind some recent trends. Relative to them we use firm-level data to inform our model about the production process and the properties associate with this new capital. We find that intangible capital is a dynamic input in production whose importance is rising and that its investment process is highly distorted by technological frictions like adjustment costs. Combining these novel insights with a quantitative model, to the best of our knowledge, we are the first to jointly explain the rise in average firm size and concentration, the changes in aggregate factor shares, the decline in the tangible investment rate, and the decline in allocative efficiency in the US economy between 1980 to 2015.

Outline. Section 2.2 briefly discusses data and shows the construction of our main variables. Section 2.3 documents the stylized facts. Section 2.4 presents our quantitative framework. Section 2.5 contains the calibration of the model and its external validation, whereas, section 2.7 presents the main results and discusses the implications of IBTC for the US economy. Section 2.8 concludes.

2.2 DATA AND MEASUREMENT

In this section, we present the main dataset used throughout the analysis. We explain (i) the construction of the variables, unrelated to intangible capital, used for the empirical analysis and (ii) the measurement of firm-level intangible capital, emphasizing the main challenges, its virtues, and its drawbacks.

2.2.1 *Main Measures*

The main data source is Compustat, a firm-level database with all the US publicly traded firms between 1980 to 2015. In this section, we discuss briefly the strengths and limitations of this dataset. We provide more details on the data cleaning process and the construction of the sample of analysis in Appendix 2.A.1.

The choice of the data is driven solely by its ability to cover the period of interest and the largest number of sectors. These characteristics make these data an excellent source of firm-level information to study technological changes in production undertaken by US firms.

Despite the fact that publicly traded firms are few relative to the total number of firms, as they tend to be the largest firms in the economy, they account for roughly 30% of US employment (see, Davis, Haltiwanger, Jarmin, Miranda, Foote, and Nagypal (2006)). The Compustat data contain information on firm-level financial statements including measures of sales, input expenditures, capital stock information, as well as a detailed industry activity classification.

However, despite its many virtues, these data present two main limitations: (i) The fact that it is impossible to distinguish quantity and prices, which makes the measurement of the production function elasticities significantly more challenging as extensively explained in the next section;[5] (ii) the possible selection issues arising from using

[5] This challenge is present in most of the production data.

only publicly traded firms. To address the first concern, we follow the methodologies explained in Appendix 2.A.2. Moreover, whenever possible we compare our results with additional data sources to isolate the potential bias of using only publicly traded firms.

We use as a measure of firm-level production firms' sales (`SALE`), as a measure of variable inputs used in production we use cost of goods sold (`COGS`), as a measure of firm-level employees we use (`EMP`), and as a measure of tangible capital we use gross capital (`PPEGT`). Summary statistics related to these variables are reported in Appendix 2.A.1.

2.2.2 *Intangible Capital Measurement*

The firm-level measurement of intangible capital is a challenging task as a substantial portion of it instead of being externally acquired is internally generated and US GAAP does not allow its capitalization in the balance sheet (see, Lev and Gu, 2016 and Ewens, Peters, and Wang, 2019). As a consequence, following the accounting standards in force, nearly all the internally generated intangible capital is recorded differently from tangible capital in the accounting books. In particular, all tangible investment is recorded in the balance sheet at its purchased price and then depreciated over its useful life; however, internally produced intangible investment, such as R&D, advertisement, or employee training, is fully expensed in the current period and hence it appears in the firms' income statement but not in the balance sheet. Only externally acquired intangible capital is directly booked in the balance sheet. For a more in-depth discussion about accounting standards and related challenges to firm-level intangible capital measurement, see Appendix 2.A.1.3.

In light of these considerations, our main measure in the paper is formed by two different components: (i) Internally generated intangible capital and (ii) externally acquired intangible capital. Internally generated intangible capital in our case is obtained through the capitalization of R&D expenditure (XRD). We do not include organizational capital in our benchmark measure as this is normally constructed through the capitalization of a sector-dependent share of Selling, General, and Administrative expenses (XSGA).[6] This item includes many expenditures that are not inherently related to intangible capital like CEO wages, rents for buildings, and capital adjustment costs, among others.[7] Capitalizing such a big expenditure item would heavily downward bias our estimates of the inaction rate as this expenditure item is never zero and even in periods of no investment in intangible capital, we would be capturing some unrelated overhead cost. Moreover, using organizational capital we would be capitalizing a part of incurred adjustment costs and hence we would artificially inflate our measure of intangible capital creating conceptual issues in the estimation of the production function. Finally, the imputation of a constant fraction across firms of SG&A as intangible investments would substantially increase the concerns related to potential firm-level measurement error.

Therefore, we use the perpetual inventory method on R&D expenditure to recover a firm-level measure of knowledge capital given by:

$$k_{R\&D,ft} = (1-\delta_s)k_{R\&D,ft-1} + \mathtt{XRD}_{ft}, \tag{2.1}$$

where XRD is gross investment in knowledge capital deflated by the IPP price deflator, the sector-level depreciation rate δ_s is taken from Ewens, Peters, and Wang (2019), and the initial stock is assumed to be zero.[8]

The second component of intangible capital is the externally acquired intangible capital, which is capitalized on the balance sheet at

[6]The organizational capital is used in Eisfeldt and Papanikolaou (2013), Peters and Taylor (2017), and Ewens, Peters, and Wang (2019).

[7]While working with Compustat data, it is often assumed that the capital adjustment costs are expensed in XSGA).

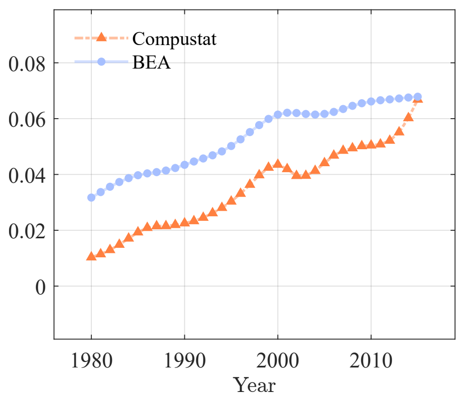
[8]For all our analysis, unless differently stated, we exclude all the observation in the first five years to avoid strong dependence of our results from our assumption on the initial condition for knowledge capital. Results are not sensitive to this. Moreover, results are similar if we use a different level of initial capital, for instance investment in the first period divided by its depreciation rate.

fair value under the variable INTAN in Compustat, according to the US GAAP under the guidelines provided in ASC 350. However, INTAN is *net* intangible capital and to get the gross measure, to be consistent with the measurement of both tangible capital and internally produce intangible capital, we use INTAN + AM, where AM is the amortization of balance sheet intangible capital. Finally, due to measurement issues explained extensively in Appendix 2.A.1.3, we drop goodwill form our measure of gross balance sheet intangible capital. Hence, our final measure of balance sheet intangible capital is:

$$k_{BS,ft} = \texttt{INTAN}_{ft} + \texttt{AM}_{ft} - \texttt{GDWL}_{ft}, \tag{2.2}$$

where all variables have been appropriately deflated with the IPP deflator.

FIGURE 2.1: Aggregate Intangible Investment Share: Compustat vs BEA



Note. The figure reports the evolution of the intangible investment share. Intangible investment share in Compustat (orange dashed line with triangles) is computed as the sum of total investment in intangible capital to the sum of total sales in a given year. Intangible investment share from the BEA corporate non-financial sector (light blue solid line with circles) is computed as the investment in intangible capital to GDP net of propriety income, taxes, and subsidies as computed by Koh, Santaeulàlia-Llopis, and Zheng (2020). The data are detrended with an Hpfilter with $\lambda = 6.25$.

Our final measure of firm-level intangible capital is given by the sum of internally produced and externally purchased intangible capital:

$$k_{I,ft} = k_{R\&D,ft} + k_{BS,ft}. \tag{2.3}$$

Figure 2.1 compares our total intangible capital investment share with the one reported by the BEA corporate non-financial sector and documented by Koh, Santaeulàlia-Llopis, and Zheng (2020).[9] We focus on the corporate non-financial sector as it is the most closely comparable to our Compustat dataset as we exclude financial firms as explained in the appendix. Overall, both data sources show a similar qualitative increase over the sample period. We see that the adjusted measure seems to better capture the level of the BEA. In Appendix 2.A.1.4 we show additional comparisons between our firm-level measure and aggregate measures from the national accounting measured at different levels of disaggregation. Concluding, we find that our firm-level measure performs reasonably well compared to national accounting data despite the data pitfalls and accounting limitations.

2.3 Empirical Analysis

This section presents the main empirical results of the paper. First, we show that the intangible capital share in production experienced a sizeable increase over the last decades and that this increase happened mostly at the expense of the labor input share. Second, we document that the investment rate distribution of intangible capital is qualitatively different from the investment rate distribution of tangible capital, suggesting a different underlying investment process.

[9] Intangible capital investment is the sum of internal investment in knowledge capital and investment in balance sheet capital. To calculate gross balance sheet capital investment, we assume a depreciation rate of 0.20, as mostly done in the literature, as there are no reliable estimates for this depreciation.

2.3.1 *Fact 1: Intangible Capital Share had a Fourfold Increase since 1980*

In this section, we investigate the importance of intangible capital as a new factor of production; to do so, we estimate a production function with three inputs: Tangible capital, intangible capital, and labor. Our estimates show that intangible capital is an important factor of production and that its importance had a fourfold increase since the 1980.

2.3.1.1 Production Function Estimation

In this section, we structurally estimate a production function with three inputs: Tangible capital, intangible capital, and labor. We estimate the log of a firm-level Cobb-Douglas production function given by:

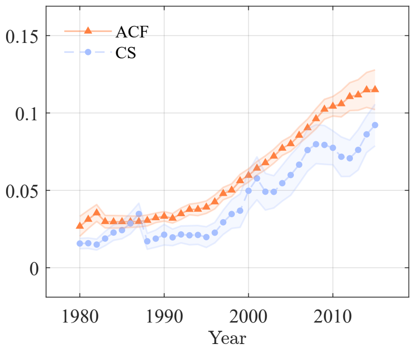
$$q_{ft} = \alpha k_{T,ft} + \nu k_{I,ft} + (1 - \alpha - \nu)\ell_{ft} + \omega_{ft} + \varepsilon_{ft}, \tag{2.4}$$

where q_{ft} is the log of output, $k_{T,ft}$ is the log of tangible capital, $k_{I,ft}$ is the log of intangible capital, ℓ_{ft} is the log of labor, ω_{ft} is the log of productivity and ε_{ft} is the error term.[10] Given that the objective of our analysis is to estimate variation in input shares over time we constraint the firm-level returns to scale to 1 and we assume that all the firms share a common technology; in the next section we show that these assumptions are inconsequential for our results. Estimating firm-level production functions is notoriously difficult as firm-level productivity ω_{ft} is un-observable to the econometrician but it is known to the firm at the moment of choosing its inputs. To address this endogeneity problem, we rely on two different estimation procedures proposed by the empirical industrial organization literature. In particular, we use the Cost Shares approach (CS) as in Foster, Haltiwanger, and Syverson (2008) and the Ackerberg-Caves-Frazer (ACF) approach

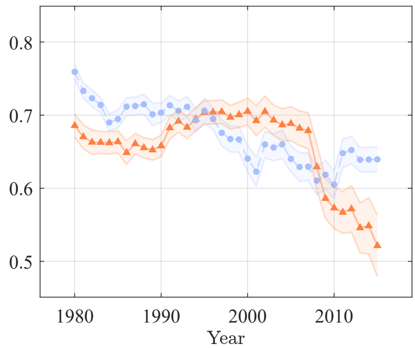
[10]Practically, as output we use firm's sales, as tangible capital we use gross property plant and equipment, as intangible capital we use the measure constructed in Section 2.2.2, and as labor we use the total firm-level number of employees.

Ackerberg, Caves, and Frazer (2015). We provide details regarding both methodologies in Appendix 2.A.2.

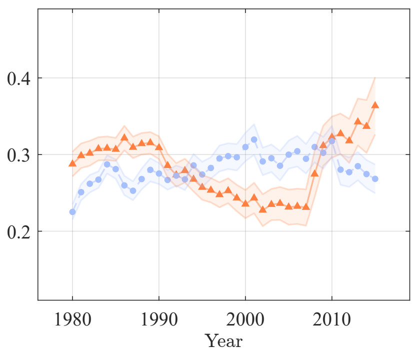
FIGURE 2.2: Trends in Input Shares



(a) Intangible Capital Share



(b) Labor Share



(c) Tangible Capital Input Share

Note. The figures present the output elasticities estimated with the Cost Shares (CS) approach (dashed light blue line with circles) and with the Ackerberg-Caves-Frazer (ACF) approach (solid orange line with triangles). The elasticities are estimating using a ten-year rolling windows over time. Bands around the point estimates report the 99% confidence interval.

To document the changes in the output elasticities of labor, intangible, and tangible capital in the production function, we estimate equation 2.4 with both methodologies between 1980 and 20015 using 10−years rolling windows. Figure 2.2 presents the results. Solid orange lines with triangles report the estimates from ACF with associated 99% confidence intervals, whereas, dashed light blue lines with circles report the estimates from CS with associated 99% confidence intervals. Regardless of the methodology chosen, all the action in the inputs shares comes from intangible and labor, as tangible capital does not show any obvious trend over the period.

In particular, intangible capital share with the CS approach goes from 0.016 in 1980 to 0.092 in 2015, whereas, with the ACF approach goes from 0.027 to 0.115. With the ACF approach – our preferred methodology – the intangible capital input share in 2015 is approximately four times as much as it was in 1980. Whereas, the CS approach estimates approximately a five times increase in the intangible capital input share over the same period. It is evident from these results that the Compustat firms, which represents a sizeable part of the US economy, have undergone a significant transformation in their production technology. We label this finding as *Intangible Capital Biased Technological Change* (IBTC).

Moreover, the labor share goes with CS approach from 0.759 to 0.639, whereas, with the ACF approach goes from 0.686 to 0.521. Therefore, we highlight that our estimates suggest a certain level of substitution between intangible capital and labor over time: While intangible capital share has increased the labor share has declined in the last decades. This is in line with the results from literature, for instance Elsby, Hobijn, and Şahin (2013), Karabarbounis and Neiman (2013), and Koh, Santaeulàlia-Llopis, and Zheng (2020), among others.

Given the results documented in this section, in the subsequent part of the paper, we interpret the rise in intangible capital as an exogenous technological change in the production technology biased towards intangible capital at the expense of the labor input.

2.3.1.2 Robustness

Here we document the extent of the robustness of our results relaxing most of the assumptions imposed on the benchmark specification. In particular, we look at the following specifications: (i) We re-estimate the production function in equation 2.4 leaving returns to scale unconstrained; (ii) we estimate equation 2.4 at two digit sector-level (NAICS 2), effectively allowing for sector-specific tenchnology; (iii) we estimate a Translong production function with constant returns to scale.

Figure 2.3 shows the results from the alternative specifications. Appendix 2.A.3 explains in detail the various specifications. Overall, the IBTC does not seem to be driven by the specific methodology applied and follows close patterns across the different specifications. The bottom line is that the findings from the benchmark specification are robust.

2.3.2 *Fact 2: Intangible Capital is More Lumpy than Tangible Capital*

Having in mind that intangible capital is an important factor in production and that its importance is growing over time we document in this section the salient differences in the investment behavior of firms between tangible capital and intangible capital. The investment rate of each capital is defined as:

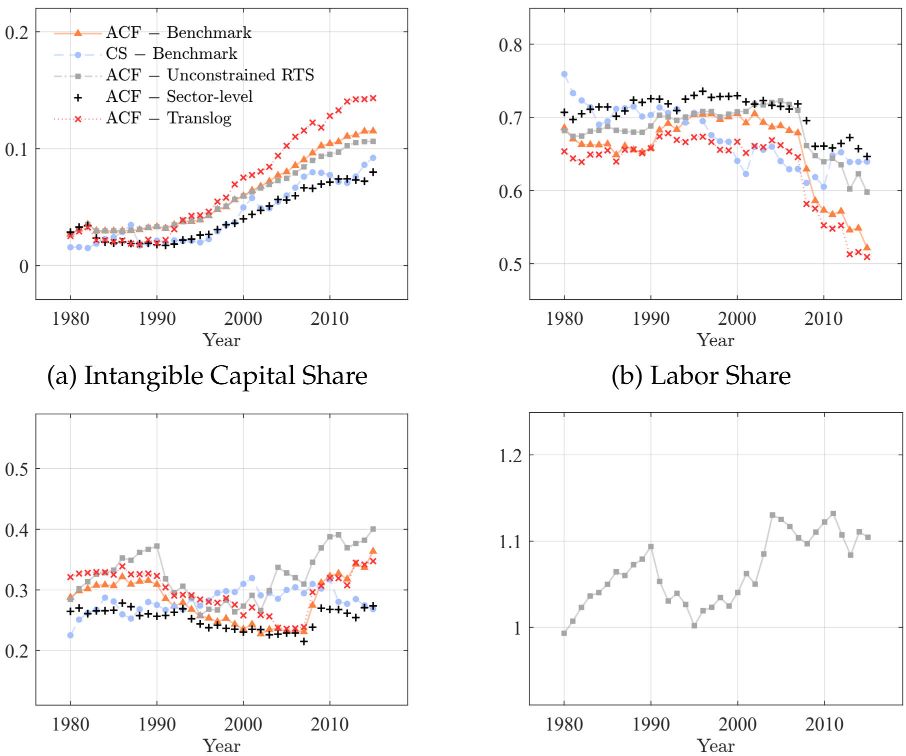
$$\frac{x_{j,ft}}{k_{j,ft-1}} \equiv \frac{k_{j,ft} - k_{j,ft-1}}{k_{j,ft-1}} + \delta_j, \quad j \in \{T, I\}, \tag{2.5}$$

where δ_j is the depreciation rate, $x_{j,ft}$ is investment, and $k_{j,ft}$ is capital.[11] Following Cooper and Haltiwanger (2006), and Clementi and Palazzo (2019), we construct a balanced panel of firms from 1980 to 1990 to study the properties of investment rates.[12] Following common practice, we also drop observations where total value of acquisitions

[11]The depreciation rate of tangible capital is 7%, whereas, the depreciation rate of intangible capital is firm dependent as explained in Section 2.2.2.

[12]This is doen to control for selection dynamics due to entry and exit in the data.

FIGURE 2.3: Trends in Input Shares – Robustness



(a) Intangible Capital Share

(b) Labor Share

(c) Tangible Capital Input Share

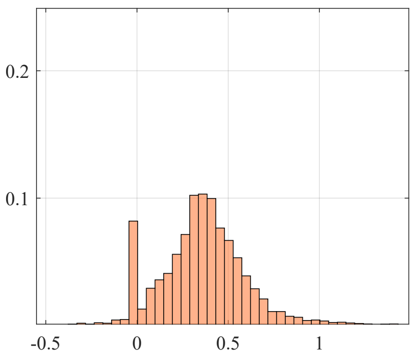
(d) Returns to Scale

Note. The figures present the output elasticities estimated with the Cost Shares (CS) approach (dashed light blue line with circles), with the Ackerberg-Caves-Frazer (ACF) approach (solid orange line with triangles), with the ACF approach and unconstrained returns to scale (dashed-dotted light gray line with squares), with the sector-level ACF approach (black plus), and with the Translog ACF approach (dotted red with crosses). The elasticities are estimating using a ten-year rolling windows over time.

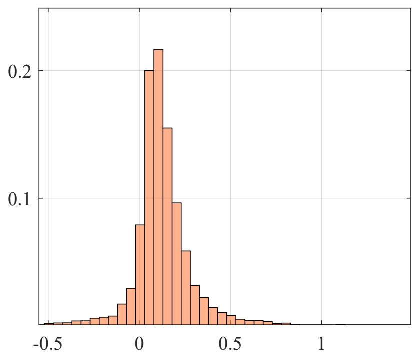
relative to total assets exceed 5%.[13] Finally, we drop those firms that have never invested in intangible capital, to prevent the overestimation of the inaction rate of intangible capital investment.

[13]This is done to avoid biases due to acquisitions, that is, given the accounting standards an acquisition would show up as a big investment for one firm but would not show up at all as a big disinvestment for the other. However, we notice that in our

FIGURE 2.4: Investment Rate Distributions



(a) Intangible Capital



(b) Tangible Capital

Note. The figures report the investment rate distributions of intangible and tangible capital for a balanced panel of firms between the years 1980 and 1990. Figure 2.4a shows the investment rate distribution for intangible capital. Figure 2.4b shows the investment rate distribution for tangible capital. The histograms are constructed dropping from the balanced panel all the firms that never invest in intangible capital and all the observations with investment rates above 2 or below -0.5. Results are robust to other winsorization schemes.

Figure 2.4a and Figure 2.4b plot the investment rate distribution for intangible and tangible capital, respectively. These distributions present two stark differences: First, the investment rate distribution for intangible capital presents a clear bi-modality, with a lot of mass at the mean and around zero. Meanwhile, the investment rate distribution for tangible capital is almost symmetric around the mean, and mimics closely the findings of Clementi and Palazzo (2019). Second, the investment rate distribution for intangible capital show a small amount of negative investments.[14]

We summarize the main moments of the investment rate distributions in Table 2.1. First, we notice that the average investment rate is much higher for intangible capital compared to tangible capital. This

balance panel this observations represent a small share of all entries.

[14]Notice that this is not by construction, that is, it is not entirely due to the capitalization of an expenditure voice like R&D, since our measure of intangible capital indeed contains balance sheet intangibles, which, given the depreciation, allows for negatives.

partly reflects a high depreciation rate for intangible capital that pushes the level of optimal investment above that of tangible capital. Second, as anticipated above, intangible capital has a much higher inaction rate, defined as the fraction of investment below 1% in absolute value; particularly, intangible capital inaction rate is 8% compared to 3% for tangible capital. This high inactivity in intangible capital suggests some underlying non-convexity in the investment process. Third, intangible capital seems to be more serially correlated over time. The autocorrelation in intangible investment is 0.31, much higher than the 0.11 exhibited by tangible capital. This suggests that, conditionally on investing in intangible capital the investment activity goes on for longer, hinting towards some slow adjustment process in the background.

In Appendix 2.A.4 we show that the investment rate distribution exhibits the same properties across sectors, suggesting that the results are not driven by sectoral heterogeneity. Moreover, we also shows that the investment rate distribution does not change over time, ruling out concerns related to potential time trends as underling factors of the documented bi-modality. Overall, we can say that the intangible capital investment process is robustly lumpy, that is, it entails long periods of inaction followed by booms of investment activity.

These findings are particularly informative on how should intangible capital be modeled as this capital appears to be neither a fixed cost nor a flexible input. Therefore, from here onward, we will think about intangible capital as a dynamic input in production which could in principle be subject to some adjustment frictions which will be quantified in the quantitative section of the paper.

2.4 THEORETICAL FRAMEWORK

In this section, to connect together all the three stylized facts from the previous sections, we introduce a quantitative general equilibrium model of investment dynamics with tangible and intangible capital, a

Table 2.1: Lumpiness

Investment rates	Intangible	Tangible
Average	0.35	0.13
Positive fraction, $i > 1$	0.89	0.87
Negative fraction, $i < -1$	0.03	0.10
Inaction rate	0.08	0.03
Spike rate, $\lvert i \rvert > 20$	0.75	0.25
Positive spikes, $i > 20$	0.73	0.22
Negative spikes, $i < -20$	0.02	0.03
Standard Deviation	0.30	0.22
Serial correlation, Corr(i_t, i_{t-1})	0.31	0.11

Note. This table shows the moments of the investment rate distribution of intangible and tangible capital. The statistics are computed for a balance panel of 5687 firm year observations between 1980 and 1990.

rich and flexible structure of investment adjustment costs, and endogenous firm entry and exit.

2.4.1 *Model Environment*

The model follows the spirit of Clementi and Palazzo (2016b). Time is discrete and indexed by $t = 1, 2, \ldots$. At time t a positive mass of price-taking firms produce an homogeneous good by means of the production function $y = e^z \left(k_T^\alpha k_I^\nu \ell^{(1-\alpha-\nu)}\right)^\omega$, with α, ω, ν in $(0, 1)$. Where k_T denotes tangible capital, k_I is intangible capital, ℓ is labor, and z is the idiosyncratic random productivity. Idiosyncratic productivity z is driven by the stochastic process:

$$z' = \rho_z z + \sigma_z \varepsilon',$$

where $\varepsilon \sim \mathcal{N}(0,1)$. The conditional distribution of z will be denoted by $\Gamma(z'|z)$.

Firms discount future profits by means of the time-invariant discount factor $\frac{1}{R}$, $R > 1$. Tangible capital depreciates at a rate $\delta_T \in (0,1)$, whereas, intangible capital depreciates at a rate $\delta_I \in (0,1)$. Adjusting tangible capital stock by x_T and intangible capital stock by x_I bears the cost:

$$\mathcal{C}(x_T, x_I; k_T, k_I) = \frac{\gamma_T}{2}\left(\frac{x_T}{k_T}\right)^2 k_T + \frac{\gamma_I}{2}\left(\frac{x_I}{k_I}\right)^2 k_I + \mathbf{1}\{x_T \neq 0\} f_T k_T + \mathbf{1}\{x_I \neq 0\} f_I k_I,$$

where $\gamma_T, \gamma_I, f_T, f_I \in \mathbf{R}^+$. We allow for two different kinds of adjustment costs: convex and fixed. We do not allow for irreversibilities in investment in the baseline version of the model. Generally, these non-convex costs of adjustment are intended to capture indivisibilities in capital, increasing returns in the installation of new capital, and increasing returns to retraining and restructuring of production activity. Moreover, this formulation of non-convex adjustment costs can be interpreted as a mild form of irreversibility, as disinvestment bears a cost in terms of output, which stems for the potential specific nature of capital. Specifically if capital is tailored to some particular needs of a firm it can in principle be difficult to resell it.[15] The convex costs capture overtime costs, inventory costs, and machine set-up costs. Furthermore, we assume that the capital adjustment costs are proportional to their respective capital stock: this is a common specification that is used to take care of the size effect. Finally, we assume that adjustment costs are paid in terms of final output.

We assume that the demand for firm's output and the supply of

[15] This idea that intangible capital is specific to the needs of the firm that uses it has been suggested by Haskel and Westlake (2018) and Edmond, Midrigan, and Xu (2018). However, we move here a step forward compared to them and we test this hypothesis concretely, specifying a flexible model to be tested in the data $-$ in principle our model could reject this estimating f_I to be close to zero.

both capitals are infinitely elastic and we normalize their prices to 1.[16] The supply of labor is given by $L(W) = W^{\psi}$, where $\psi > 0$ and $W \in \mathbf{R}^{+}$ is the real wage.[17]

Operating firms incur each period a fixed cost $c_f > 0$; this is usually interpreted as a per-period expense that firms must incur to operate, for instance to hire one unit of managerial activity. Firms that quit production cannot re-enter the market at a later stage and recoup the undepreciated part of their capital stocks, net of the adjustment cost.

Every period there is a constant exogenous mass $m > 0$ of prospective entrants, each of which receives an initial productivity s, with $s \sim \Lambda(s)$, a Pareto distribution with scale parameter η. Conditional on entry, the distribution of the idiosyncratic shock in the first period of operation is $\Gamma(z'|s)$, strictly increasing in s. Entrepreneurs that decide to enter must pay an entry cost $c_e \geq 0$.

Finally, in each period, the stationary distribution of operating firms over the three dimensions of heterogeneity is denoted by $\Omega(z, k_T, k_I; W)$. A comprehensive picture of timing in the model is presented in Figure 2.5.

2.4.2 *The Problem of the Incumbents*

Given idiosyncratic productivity z, tangible capital k_T and intangible capital k_I, the profits of an incumbent are given by:

$$\pi(z, k_T, k_I; W) = \max_{\ell} \; e^{z} \left(k_T^{\alpha} k_I^{\nu} \ell^{(1-\alpha-\nu)}\right)^{\omega} - W\ell. \tag{2.6}$$

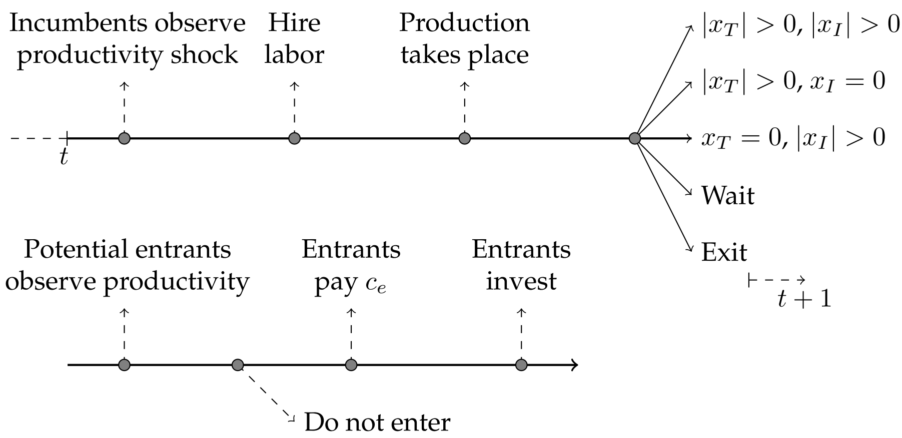
Upon exit, a firm obtains a value equal to the undepreciated portion of its tangible capital k_T and intangible capital k_I, net of the adjustment

[16] This is a standard assumption in the literature, see for example Khan and Thomas (2008) and Clementi and Palazzo (2016b).

[17] Effectively, we are assuming that the utility function of the representative household is given by:

$$u(C, L) = C - \frac{L^{1+1/\psi}}{1+1/\psi}.$$

FIGURE 2.5: Timing in the Model



cost it incurs to dismantle them:

$$\mathcal{V}_x(k_T, k_I) = (1-\delta_T)k_T + (1-\delta_I)k_I - \mathcal{C}(-(1-\delta_T)k_T, -(1-\delta_I)k_I; k_T, k_I).$$

Then, the start-of-period value of an incumbent firm is dictated by the function $\mathcal{V}(z, k_T, k_I; W)$ which solves the following functional equation:

$$\begin{aligned}\mathcal{V}(z, k_T, k_I; W) = \pi(z, k_T, k_I; W) \\ + \max\{&\mathcal{V}_x(k_T, k_I), \widetilde{\mathcal{V}}_1(z, k_T, k_I; W) - c_f, \\ &\widetilde{\mathcal{V}}_2(z, k_T, k_I; W) - c_f, \widetilde{\mathcal{V}}_3(z, k_T, k_I; W) - c_f, \\ &\widetilde{\mathcal{V}}_4(z, k_T, k_I; W) - c_f\};\end{aligned} \tag{2.7}$$

where the value of investing in both capital is given by:

$$\begin{aligned}\widetilde{\mathcal{V}}_1(z, k_T, k_I; W) = \max_{k'_T, k'_I} & -x_T - x_I - \mathcal{C}(x_T, x_I; k_T, k_I) + \\ & \frac{1}{R}\int \mathcal{V}(z', k'_T, k'_I; W)\Gamma(dz'|z), \\ \text{s.t.} \quad & k'_T = (1-\delta_T)k_T + x_T, \\ & k'_I = (1-\delta_I)k_I + x_I;\end{aligned} \tag{2.8}$$

the value of investing in only tangible capital is given by:

$$\begin{aligned}\widetilde{\mathcal{V}}_2(z, k_T, k_I; W) = \max_{k'_T} & -x_T - \mathcal{C}(x_T, 0; k_T, k_I) + \\ & \frac{1}{R}\int \mathcal{V}(z', k'_T, (1-\delta_I)k_I; W)\Gamma(dz'|z), \\ \text{s.t.} \quad & k'_p = (1-\delta_T)k_T + x_T;\end{aligned} \tag{2.9}$$

the value of investing in only intangible capital is instead given by:

$$\begin{aligned}\widetilde{\mathcal{V}}_3(z, k_T, k_I; W) = \max_{k'_I} & -x_I - \mathcal{C}(0, x_I; k_T, k_I) + \\ & \frac{1}{R}\int \mathcal{V}(z', (1-\delta_T)k_T, k'_I; W)\Gamma(dz'|z), \\ \text{s.t.} \quad & k'_i = (1-\delta_I)k_I + x_I;\end{aligned} \tag{2.10}$$

finally, the value of waiting is given by:

$$\widetilde{\mathcal{V}}_4(z, k_T, k_I; W) = \frac{1}{R}\int \mathcal{V}(z', (1-\delta_T)k_T, (1-\delta_I)k_I; W)\Gamma(dz'|z). \tag{2.11}$$

2.4.3 *The Problem of the Entrants*

The value of a potential entrant that draws an initial productivity s, where $s \sim \Lambda(s)$, is given by:

$$\mathcal{V}_e(s; W) = \max_{k'_T, k'_I} -k'_T - k'_I + \frac{1}{R}\int \mathcal{V}(z', k'_T, k'_I; W)\Gamma(dz'|s); \tag{2.12}$$

the potential entrant will invest and start operating if and only if $\mathcal{V}_e(s; W) \geq c_e$.

2.4.4 *Recursive Competitive Equilibrium*

The Recursive Competitive Equilibrium (RCE) consists of (i) value functions $\mathcal{V}(z, k_T, k_I; W)$, $\widetilde{\mathcal{V}}_1(z, k_T, k_I; W)$, $\widetilde{\mathcal{V}}_2(z, k_T, k_I; W)$, $\widetilde{\mathcal{V}}_3(z, k_T, k_I; W)$, $\widetilde{\mathcal{V}}_4(z, k_T, k_I; W)$ and $\mathcal{V}_e(s; W)$, (ii) policy functions $\ell(z, k_T, k_I; W)$, $x_T(z, k_T, k_I; W)$, $x_I(z, k_T, k_I; W)$, $k'_T(s; W)$, $k'_I(s; W)$, and (iii) an incumbents' measure $\Omega(z, k_T, k_I; W)$, and an entrants' measure $\mathcal{E}(z, k_T, k_I; W)$ such that:

1. $\mathcal{V}(z, k_T, k_I; W)$, $\widetilde{\mathcal{V}}_1(z, k_T, k_I; W)$, $\widetilde{\mathcal{V}}_2(z, k_T, k_I; W)$, $\widetilde{\mathcal{V}}_3(z, k_T, k_I; W)$, $\widetilde{\mathcal{V}}_4(z, k_T, k_I; W)$, $\ell(z, k_T, k_I; W)$, $x_T(z, k_T, k_I; W)$ and $x_I(z, k_T, k_I; W)$ solve (2.6), (3.2), (2.8), (2.9), (2.10), and (2.11);

2. $\mathcal{V}_e(s; W)$, $k'_T(s; W)$ and $k'_I(s; W)$ solve (2.12);

3. The labor market clears: $\int \ell(z, k_T, k_I; W) d\Omega(z, k_T, k_I; W) = L(W)$

4. For all Borel sets $\mathcal{Z} \times \mathcal{K}_T \times \mathcal{K}_I \subset \mathbf{R}^+ \times \mathbf{R}^+ \times \mathbf{R}^+$,

$$\mathcal{E}(\mathcal{Z} \times \mathcal{K}_T \times \mathcal{K}_I; W) = m \int_{\mathcal{Z}} \int_{\mathcal{B}_e(\mathcal{K}_T, \mathcal{K}_I; W)} \Lambda(ds)\Gamma(dz'|s),$$

   where $\mathcal{B}_e(\mathcal{K}_T, \mathcal{K}_I; W) = \big\{z \text{ s.t. } k'_T(s; W) \in \mathcal{K}_T, k'_I(s; W) \in \mathcal{K}_I \text{ and } \mathcal{V}_e(s; W) \geq c_e\big\}$;

5. For all Borel sets $\mathcal{Z} \times \mathcal{K}_T \times \mathcal{K}_I \subset \mathbf{R}^+ \times \mathbf{R}^+ \times \mathbf{R}^+$ and $\forall t \geq 0$,

$$\Omega(\mathcal{Z} \times \mathcal{K}_T \times \mathcal{K}_I; W) = \int_{\mathcal{Z}} \int_{\mathcal{B}(\mathcal{K}_T, \mathcal{K}_I; W)} \Omega(dz, dk_T, dk_I; W)\Gamma(dz'|z) + \mathcal{E}(\mathcal{Z} \times \mathcal{K}_T \times \mathcal{K}_I; W),$$

   where $\mathcal{B}(\mathcal{K}_T, \mathcal{K}_I; W) = \big\{(z, k_T, k_I) \text{ s.t. } \max\{\widetilde{\mathcal{V}}_1(z, k_T, k_I; W), \widetilde{\mathcal{V}}_2(z, k_T, k_I; W), \widetilde{\mathcal{V}}_3(z, k_T, k_I; W), \widetilde{\mathcal{V}}_4(z, k_T, k_I; W)\} - c_f \geq \mathcal{V}_x(k_T, k_I), (1-\delta_T)k_T + x_T(z, k_T, k_I; W) \in \mathcal{K}_T \text{ and } (1-\delta_I)k_I + x_I(z, k_T, k_I; W) \in \mathcal{K}_I\big\}$.

2.4.5 *Output Elasticities, Adjustment Costs, and Allocative Efficiency*

One of the main objects of interest for our analysis is going to be the evolution of allocative efficiency due to IBTC. To define a model consistent measure of allocative efficiency we leverage the work by Hsieh and Klenow (2009b) and define $TFPR$ in the model as:

$$\begin{aligned} TFPR_{ft} &= \frac{y_{ft}}{k_{T,ft}^{\alpha} k_{I,ft}^{\nu} \ell_{ft}^{(1-\alpha-\nu)}} \\ &\propto \left(ARPK_{T,ft}\right)^{\alpha} \left(ARPK_{I,ft}\right)^{\nu} \left(ARPL_{ft}\right)^{(1-\alpha-\nu)}, \end{aligned} \tag{2.13}$$

where $ARPK_{T,ft} = y_{ft}/k_{T,ft}$ is the average product of tangible capital, $ARPK_{I,ft} = y_{ft}/k_{I,ft}$ is the average product of intangible capital, and $ARPL_{ft} = y_{ft}/\ell_{ft}$ is the average product of labor.[18] Therefore, our measure of allocative efficiency in the economy is defined by:

$$\begin{aligned} Var(TFPR_{ft}) = \alpha^2 Var\left(ARPK_{T,ft}\right) + \nu^2 Var\left(ARPK_{I,ft}\right) \\ + 2\alpha\nu Cov(ARPK_{T,ft}, ARPK_{I,ft}) \end{aligned} \tag{2.14}$$

where $Var(\cdot)$ represents variance and $Cov(\cdot)$ is the covariance. This definition of allocative efficiency is the same as the one extensively used by the misallocation literature pioneered by Hsieh and Klenow (2009b). Notice that the allocative efficiency of this economy is independent of $ARPL$ as it is equalized across firms. Therefore, only the $ARPK_T$ and $ARPK_I$ are relevant to understand the evolution of allocative efficiency in our framework.

In the absence of adjustment costs to both capitals, their average product would equalize across firms, and hence allocative efficiency as measured by the dispersion in $TFPR$ would be zero – which is by definition the highest level of allocative efficiency achievable in the model. On the opposite, in the presence of adjustment

[18]Hopenhayn (2014) explains extensively how to define $TFPR$ in models of perfect competition.

costs to both capitals, their average product does not equalize anymore as the reallocation of both capitals is slowed down by the adjustment costs themself. Therefore, the effect of adjustment costs is to make $Var\left(ARPK_{T,ft}\right), Var\left(ARPK_{I,ft}\right) > 0$ and consequently $Var(TFPR_{ft}) > 0$.

Therefore, equation 2.14 clarifies the relation between IBTC and allocative efficiency in the model. An increase in intangible capital share, ν, relative to labor share, $1-\alpha-\nu$, increases the importance of $Var\left(ARPK_{I,ft}\right)$ and hence, all else equal, it increases overall dispersion in $TFPR$ and hence it lowers allocative efficiency in the model. This is just a by-product of the fact that IBTC lowers the reliance of firms on an undistorted input like labor while increases their reliance on a (potentially) highly distorted input like intangible capital.

2.5 QUANTITATIVE ANALYSIS

In this section, we use the structural framework presented in Section 2.4 to estimate the adjustment costs associated with tangible and intangible capital. Then, we validate our model with: (i) Non-targeted moments from the cross-sectional and age distributions; and (ii) on the empirical dispersion and responsiveness of the average revenue product of both capitals.

2.5.1 *Calibration*

The baseline calibration matches jointly the investment behavior of tangible and intangible capital at the micro-level and business dynamism in the overall US economy for the sample period 1980-1990. The parameterization proceeds in two steps. First, we fix a set of parameters that are estimated outside the model, for instance the parameters governing the production technology and the TFP process. Second, given the values of these fixed parameters, we choose the remaining parameters to match informative moments regarding firms' investment distribution and firms' life-cycle.

Fixed parameters. A model period is one year, so we set the interest rate $R = 1.05$. The annual depreciation rate for tangible capital is $\delta_T = 0.07$, which equals the value used to perform the empirical analysis above. We set the depreciation rate for intangible capital to be, $\delta_I = 0.29$, which is the average firm-level depreciation rate from our data. The production function parameters comes from the estimates reported in Section 2.3.1. The returns to scale ω is set 0.90 close the values used in the literature. Finally, the persistence of the idiosyncratic process is $\rho_z = 0.90$ and the standard deviation is $\sigma_z = 0.20$. This values are close to the empirical estimated reported in Foster, Haltiwanger, and Syverson (2008) and in Lee and Mukoyama (2015).

Fitted parameters. We choose the remaining parameters to match some moments from Table 2.1 and some moments on business dynamism from the BDS. Specifically, we use inaction rates, that is, investment rates that are within $\pm 1\%$, to discipline the parameters governing the fixed costs of investing in both tangible and intangible capital, f_T and f_I. This is particularly appealing since the model predicts that the fixed costs of adjusting directly influence the extensive margin of investment, that is, the amount of action and inaction in the investment of a given capital. We use the serial correlation of both investment rates to identify the convex costs of adjusting for both capitals, γ_T and γ_I. With high convex costs, firms adjust their capital stock more slowly over time, which in turn increases the autocorrelation of investment at the firm-level.[19] To identify the entry cost c_e, the operating cost c_f, and the parameter that governs the Pareto distribution of the productivity of potential entrants, η, we match respectively the entry rate, the average size of incumbents, and the average size of entrants. Finally, we set the

[19]David and Venkateswaran (2019) explains that using the autocorrelation in investment as an identifying moment makes the calibration robust to potential misspecification due to the absence of financial frictions.

measure of potential entrants to m to target an equilibrium wage of 1.

The parameters are estimated using the following routine. For arbitrary values of the vector of parameters, $\mathcal{P} = (\gamma_T, \gamma_I, f_T, f_I, c_e, c_f, \eta, m)$, the model is solved and the policy functions for investment in both capitals, for entry, and for exit are generated. Using these policy functions, the decision rules are simulated until the distribution of firms over $\{z, k_T, k_I\}$ is converged. We simulate the economy and construct a balanced panel of firms in the same spirit of the empirical analysis presented above. We compute the entry rate, the average size of entrants, and the average size of the incumbents from the stationary distribution. We compute the moments of the investment rates from the simulated balanced panel. We denote the vector of simulated moments as $\mathcal{M}(\mathcal{P})$. We estimate the fitted parameters $\widehat{\mathcal{P}}$ using a minimum distance criterion given by:

$$\mathcal{L}(\mathcal{P}) = \min_{\mathcal{P}} \left(\widehat{\mathcal{M}} - \mathcal{M}(\mathcal{P})\right)' \mathbf{W} \left(\widehat{\mathcal{M}} - \mathcal{M}(\mathcal{P})\right). \tag{2.15}$$

Following Asker, Collard-Wexler, and De Loecker (2014), we set the weighting matrix $\mathbf{W} = \mathbf{I}$ and use grid search to find the vector $\widehat{\mathcal{P}}$ that minimizes the objective function.

The fitted parameters from the grid search algorithm and the implied moments of the model are presented in Table 2.2. The model identifies different adjustment costs for tangible and intangible capital. Similar to Clementi and Palazzo (2019), our model imputes almost negligible fixed costs and low convex costs of investing into tangible capital. The reason being that the Compustat dataset contains disproportionately large firms.[20] Moreover, our model implies that intangible capital entails much higher adjustment costs relative to tangible capital. Therefore, the calibrated model shows that investment in intangible

[20]However, contrary to our results, Cooper and Haltiwanger (2006) use a more heterogeneous sample of plants from the confidential Census database and find larger adjustment costs for tangible capital. Another point of distinction is that our analysis is at firm-level. Therefore, we want to emphasize that our estimates can be interpreted as a lower bound to these costs for both capitals.

Table 2.2: Parameters and Moments

Fixed	Value	Description
R	1.05	Annual interest rate
δ_T	0.07	Annual depreciation rate tangible capital
δ_I	0.29	Annual depreciation rate intangible capital
α	0.28	Tangible capital share
ν	0.03	Intangible capital share
ω	0.90	Returns to scale
ρ_z	0.90	Autocorrelation idiosyncratic productivity
σ_z	0.20	Standard deviation idiosyncratic productivity

Fitted	Value	Description	Moments	Model	Data
γ_T	0.006	Convex adj. cost k_T	corr($x_{T,ft}, x_{T,ft-1}$)	0.13	0.12
γ_I	0.135	Convex adj. cost k_I	corr($x_{I,ft}, x_{I,ft-1}$)	0.30	0.31
f_T	2.5·e-3	Fixed adj. cost k_T	Inaction rate: x_T	0.03	0.03
f_I	0.021	Fixed adj. cost k_I	Inaction rate: x_I	0.08	0.08
c_e	3·e-4	Entry cost	Entry rate	0.13	0.13
c_f	2.540	Operating cost	Avg. firm size	23.52	20.49
η	2.025	Scale parameter	Avg. entrant size	4.80	6.07
m	6.2·e-3	Measure of potential entrants	Wage	1	–

capital is subject to higher technological frictions and hence is going to be more distorted relative to a frictionless benchmark.

Finally, to validate the plausibility of our parametrization we report that, in the model, the mean of the intangible capital investment rate and of the tangible capital investment rate are respectively 0.32 and 0.07, close to the empirical counterparts of 0.35 and 0.14. This s quite satisfactory as these are all untargeted moments. Moreover, we find that in the model the standard deviation of tangible capital to sales and intangible capital to sales are respectively 2.51 and 0.23, close to the empirical values of 2.47 and 0.36. This is a particularly relevant moment as partially identifies the persistency of the production process as explained by Clementi and Palazzo (2016a). We also notice that the model produces an employment share for firms with $250+$ employees of 0.49 compared to 0.51 the data.

2.5.2 *Validation*

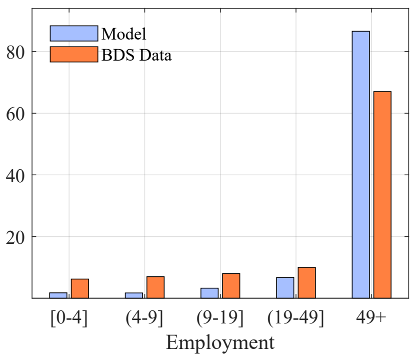
In this subsection we validate our model with: (i) Non-targeted moments from the cross-sectional and age distributions; (ii) on the empirical dispersion and responsiveness of the average revenue product of both capitals; and (iii) on the cross-sectoral implications. Additional validation exercises are shown in Appendix 2.B.1.

2.5.2.1 Model Cross-Section and Life-cycle

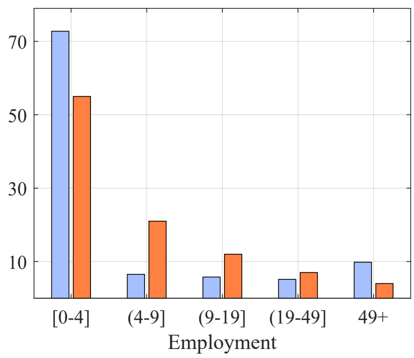
Here, we discuss the cross-sectional and life-cycle implications of the model. Figure 2.6 compares the distributions produced by the model with representative empirical distribution constructed using the BDS dataset. Similar to what is documented in the previous literature on firm dynamics, the model exhibits size and age distributions that are right-skewed.

Figure 2.6a shows that the model does a reasonably good job in matching the firm size distribution that is present in the data. This is not totally surprising as the average incumbent size and the average entrant size have been targeted in the calibration. Figure 2.6b shows that in the model the majority of firms are small, whereas, a large portion of employment is concentrated among the large firms, a feature well established in the data and visible in Figure 2.6a. Finally, the model predicts that around 70% of the firms are operating for more than 11 years and they account for around 80% of the employment share, which is slightly above to what we observe in the data (see Figure 2.6c for cohort-wise employment shares and Figure 2.6d for age distribution). Overall, the model does a satisfactory job in matching the empirical distributions of size and age despite the fact that most of these distributions were not a particular target in the calibration.

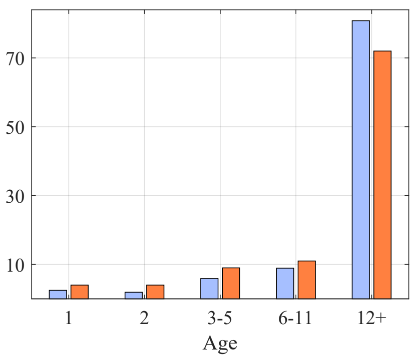
FIGURE 2.6: Size and Age Distribution



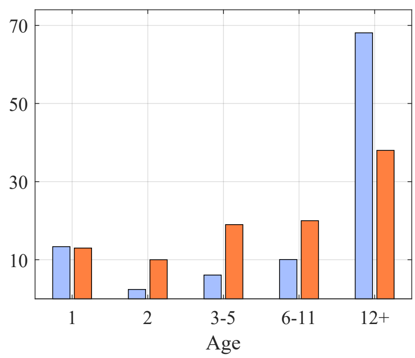
(a) Employment Distribution



(b) Firm Distributions



(c) Cohort Employment Share



(d) Age Distribution

Note. The figures show the size (employment) and age distribution of the firms, both in the model and in the data. Orange bars show the empirical distributions, light blue bars show the distributions from the model.The top-left figure shows the employment share across different employment categories. The top-right figure shows the share of firms across different employment categories. The bottom-left figure shows the employment share across different age bins. The bottom-right figure shows the share of firms across different age bins. Empirical distributions are from the BDS data.

2.5.2.2 Quasi-Fixed Inputs and Marginal Products

Here we discuss the consequences of adjustment costs when firms' are hit by productivity shocks; particularly we focus on two main things: (i) The dispersion in the average revenue product of both capitals and (ii) the responsiveness of the average revenue product of both capitals to productivity shocks. The fact that capital, in the presence of adjustment costs and time to build, is a quasi-fixed input leads to an environment where the average revenue product of each capital is not equalized across firms. This happens because, when a productivity shock hits, the firm cannot adjust the capital stock immediately to the desired frictionless level; therefore, the average revenue product of capital differs from the marginal cost, that is, the opportunity cost of holding the capital. Given that the calibration pointed out that intangible capital is more fixed compared to tangible capital as it subject to higher adjustment costs, the model predicts that the average revenue product of intangible capital is more dispersed as well. Moreover, this also implies that conditional a productivity shock intangible capital adjust less than tangible capital and, as a consequence, its average revenue product reacts by more.

To test the aforementioned predictions of the model in the data we need to compute the average product of both capitals. Under the assumption that all firms share the same Cobb-Douglas production technology within a sector, we can compute in the data the log of the average revenue product of both capitals, for firm f at time t as:

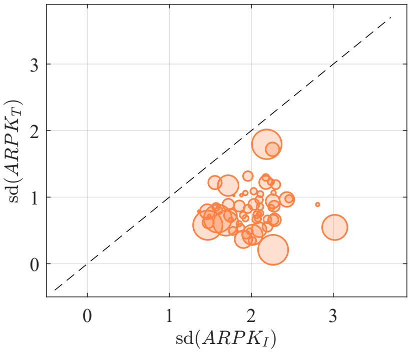
$$ARPK_{j,ft} = \log(y_{ft}) - \log(k_{j,ft}), \quad j \in \{T, I\}, \tag{2.16}$$

where y_{ft} is firm-level output and $k_{j,ft}$ is firm-level capital.

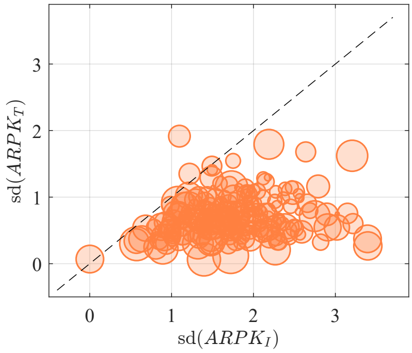
We compute the dispersion in the average v product of capital at the SIC2 and SIC3 level. Results are presented in Figure 2.7, where we scatter plot the sector-level standard deviation of $ARPK_I$ against the sector-level standard deviation of $ARPK_T$. It emerges from both figures that in the vast majority of sectors, both if we consider SIC2 or SIC3 level of disaggregation, the average revenue product of intangible capital is more dispersed than that of tangible capital, as predicted by the theory.

Furthermore, we test the second prediction of the model, namely the higher responsiveness to revenue productivity shocks of the average

FIGURE 2.7: Sector-Level Dispersion in $ARPK_I$ and $ARPK_T$



(a) SIC2−Level



(b) SIC3−Level

Note. The figures show the standard deviation of $ARPK_I$ on the $x-$axis and the standard deviation of $ARPK_T$ on the $y-$axis. Standard deviations are calculated within sectors and averaged across the years. Average products are constructed as described in the text. The black dashed line shows the 45 degree line. Figure 2.7a is constructed calculating standard deviations at SIC2−level, each circle represents a SIC2 sector, where the size of the circle is proportional to its size (sale-weighted) in Compustat. Figure 2.7b is constructed calculating standard deviations at SIC3−level, each circle represents a SIC3 sector, where the size of the circle is proportional to its size (sale-weighted) in Compustat.

revenue product of intangible capital than the average revenue product of tangible capital. To do so we perform the following regression:

$$ARPK_{j,ft} = \gamma_1 \varepsilon_{ft} + \gamma_2 k_{j,ft} + \gamma_1 TFPR_{ft-1} + \gamma_s + \gamma_t + \nu_{ft}, \quad j \in \{T, I\}, \tag{2.17}$$

where ε_{ft} is the innovation to log total factor productivity revenue.[21] The regression coefficient of interest is γ_1. In the absence of any adjustment cost, nor of time to build, the average revenue product of capitals should equalize across firms and be constant, hence the regression coefficient, γ_1, should be zero. On the contrary, the more distorted an

[21] To compute ε_{ft} we run the following regression:

$$TFPR_{ft} = \rho TFPR_{ft} + \gamma_s + \gamma_t + \nu_{ft}.$$

Then, the firm-level innovation to revenue productivity is calculated as $\varepsilon_{ft} = TFPR_{ft} - \widehat{\rho} \cdot TFPR_{ft}$.

input is the higher its average revenue product response to a revenue productivity shock and hence the higher the coefficient γ_1.

Table 2.3: Heterogeneous Response of Average Products to $TFPR$ shocks

	(1)	(2)	(3)	(4)
Dependent Variable	$ARPK_{T,ft}$	$ARPK_{I,ft}$	$ARPK_{T,ft}$	$ARPK_{I,ft}$
ε_{ft}	1.192***	1.592***	1.095***	1.239***
	(0.011)	(0.026)	(0.008)	(0.019))
$k_{T,ft}$			-0.111***	
			(0.001)	
$k_{I,ft}$				-0.399***
				(.001)
$TFPR_{ft-1}$			0.839***	0.940***
			(0.003)	(0.008)
Time dummies	✓	✓	✓	✓
Sector dummies	✓	✓	✓	✓
Observations	0.447	0.396	0.714	0.692
R-squared	89,967	89,967	89,967	89,967

Notes. We report the coefficients from the regressions of $ARPK_{T,ft}$ and $ARPK_{I,ft}$ on revenue productivity shock ε_{ft}. The controls include lagged revenue productivity, TFPR$_{ft-1}$, tangible capital, $k_{T,ft}$, and intangible capital, $k_{i,ft}$. Standard errors are in parentheses. *** p-value¡0.01, ** p-value¡0.05, * p-value¡0.1.

Results are presented in Table 2.3. As predicted by the theory, we find that the average product of both the tangible and the intangible capital reacts positively to revenue productivity shocks, as γ_1 is significantly greater than zero in all the specifications. Moreover, the average revenue product of intangible capital is more reactive to revenue productivity shocks relative to the average revenue product of tangible capital. This is in line with the prediction of the model that firms do not adjust their intangible capital as frequently as their tangible capital due to the presence of high adjustment costs.

2.6 INTANGIBLE BIASED TECHNOLOGICAL CHANGE AT WORK

In this section, we discuss the main mechanisms that drive our results. We describe the working of the model in detail and disentangle the partial and general equilibrium forces. Finally, we cross-validate the main mechanism in the data by exploiting cross-sector variation in the data.

2.6.1 *Main Mechanism*

Here, we analyze the underlying forces behind the main implications of IBTC. In the model, a rise in the output elasticity of intangible capital, at the cost of the labor elasticity, affects (i) the aggregate factor shares, (ii) the average firm size, profit rate, and concentration, and (iii) the allocative efficiency as measured by the dispersion in TFPR. This happens because, when a distorted input like intangible capital rises, it influences the demand of each input and many other equilibrium outcomes, such as equilibrium wages, firms' selection, firms' growth, and the allocation of capital across firms. Hence, the objective of this section is to uncover these forces and to link them with the IBTC.

The two fundamental forces that drive the aggregate changes are: (i) The change in the demand of the inputs due to the firm-level technological change and (ii) the endogenous change in the selection process of the firms due to the rise of a distorted input (intangible capital). First, IBTC makes production more intangible intensive at the expense of labor, this increases the demand for intangible capital while depresses the demand for labor. Therefore, this mechanically increases intangible investment share, while decreasing labor share.

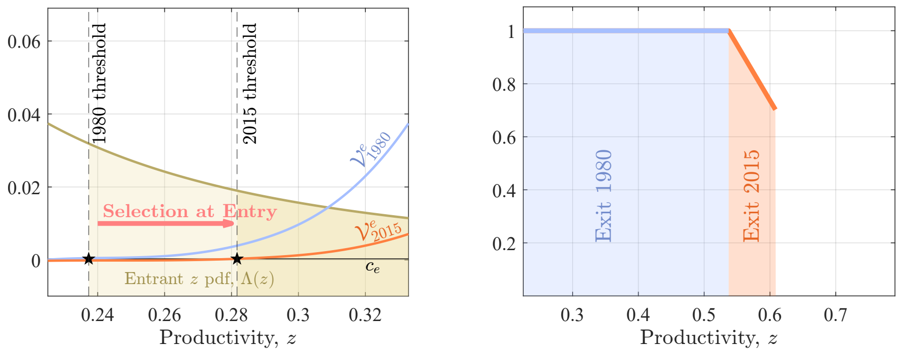
Second, this technological change commands the firm to invest more in a distorted input, subject to high adjustment costs, like intangible capital. Only sufficiently productive firms can do this, that is, firms that are productive enough to face a positive value of operating in the new intangible-intensive economy. Therefore, this affects selection both for entrant firms and for incumbent firms, as shown by Figure 2.8. Figure 2.8a on the left shows the entry decision for potential entrants 1980 (before IBTC) and in 2015 (after IBTC). Figure 2.8b on the right shows the exit probability of incumbent firms in both economies.

The IBTC lowers the value of entry as showed in Figure 2.8a ($\mathcal{V}^e_{1980} > \mathcal{V}^e_{2015}$). This triggers a rightward shift of the entry threshold, implying that by 2015 only more productive firms can enter. This happens because only more productive firms can pay the entry cost c_e. Moreover, as shown by Figure 2.8b on the right, incumbent firms are subject to a similar increase in selection. In the new economy, marginally more productive firms face a positive exit probability, as shown by the orange line. Overall, this means that IBTC increases both ex-ante and ex-post selection in the economy.

The rise in aggregate intangible investment share is lower than the counterfactual increase implied by the firm-level rise in its input share in a frictionless model (without adjustment costs), as can be seen quantitatively in Section 2.7.[22] This is because despite IBTC mechanically increases demand for intangible capital, it triggers a rise in the selection that favors more productive firms, dampening the aggregate rise. In the model, high productivity firms have a lower investment share as they expect to contract on average in the future due to the mean-reversion in the productivity process. Therefore, a redistribution towards high productivity firms translates into a redistribution towards low investment share firms. This composition effect dampens the rise of the aggregate intangible investment share. This same mechanism explains why the

[22] In a frictionless model changes in the firm-level input shares uniquely pin down the change in aggregate input shares.

FIGURE 2.8: IBTC and Firms' Selection



(a) Entry Decision (b) Exit Probability

Note. Figure 2.8a on the left shows graphically the entry problem of potential entrants both in the 1980 and 2015 calibrations. The 2015 calibration is shown in Section 2.7. The beige line in the background shows the productivity distribution of potential entrants, $\Lambda(z)$. The light blue curve and the orange curve show the value function of potential entrants for both calibrations, $\mathcal{V}^e_{1980}$ and $\mathcal{V}^e_{2015}$. The value of entry is lower in 2015 compared to 1980 because firms to grow in the intangible-intensive economy have to spend more resources on high adjustment costs. The black line shows the entry cost, c_e. The two black dashed vertical lines show the exit threshold in both 1980 and 2015, that is, the productivity level that satisfies $c_e = \mathcal{V}^e_t(z),\ t \in \{1980, 2015\}$. The shaded light beige area in the background shows the ex-post productivity distribution of entrants in 1980, whereas the shaded dark beige area in the background shows the ex-post productivity distribution of entrants in 2015.
Figure 2.8b on the right shows the exit probability of incumbent firms both for the 1980 and for the 2015 calibration. The light blue line shows the exit probability for incumbent firms in 1980. The orange line shows the exit probability in 2015. Firms with higher productivity in 2015 face a positive probability of exit because in the intangible-intensive economy is more difficult to operate as firms to respond to productivity shocks have to spend more on adjustment costs.

aggregate tangible investment share declines due to IBTC. In this case, because firm-level input share of tangible capital does not change over time, the composition effect drives this decline. Finally, the labor share declines only because of the change in the firm-level input share, as the composition effect has no impact on labor share that it is equalized across firms.

Moreover, IBTC raises the average firm size, profit rate, and concentration. This is because of two reasons: (i) It increases selection as explained above, and (ii) it favors the larger firms in the economy. IBTC makes the growth of small firms costly, as they have to incur very high adjustment costs to build their stock of intangible capital, while it makes it easier for large firms to shrink, as the high depreciation rate of intangible capital favors its depletion. This mechanism, together with the above increase in selection, triggers a reallocation of sale shares towards the larger firms, reinforcing the rise in average firm size, profit rate, and industry concentration.

Finally, the model predicts that the allocative efficiency in the model declines as intangible capital share increases. This is driven by the fact that $TFPR$ in our framework is a weighted geometric mean of the average revenue product of inputs, where the weights are proportional to their output elasticities. Due to the presence of adjustment costs, dispersion in $TFPR$ is driven by dispersion in average products of both capitals. This can be seen from equation 2.14. When the output elasticity of intangible capital increases, the dispersion in $ARPK_I$ becomes the primary driver of the dispersion in $TFPR$.[23] Therefore, dispersion in $TFPR$, which is our measure of allocative efficiency (where higher dispersion in $TFPR$ means lower allocative efficiency), increases.

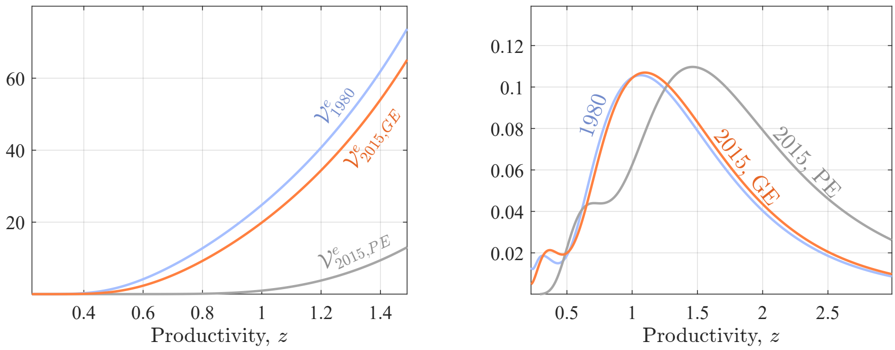
2.6.2 *General Equilibrium vs Partial Equilibrium*

Here, we highlight the consequences of IBTC on the economy and disentangle the partial and general equilibrium effects. To do so, we solve the model with the intangible-intensive production technology as estimated in 2015, while holding the wage constant, therefore we only capture the partial equilibrium effects of IBTC. As discussed above, IBTC lowers the value of entry as it makes it more difficult for firms to operate. As shown in Figure 2.9a, the partial equilibrium value of entry $\mathcal{V}^e_{2015,PE}$ is significantly lower than $\mathcal{V}^e_{1980}$, thus pushing up the productivity of the marginal entrant. A similar rise in selection is also

[23] Adjustment costs associated with the investment process of input do not allow its marginal product to equalize across firms and hence generate dispersion in the average revenue product.

observed for exiting firms. Due to this, the distribution of incumbent firms is shifted to the right, as shown in Figure 2.9b.

Figure 2.9: General vs Partial Equilibrium effects



(a) Entrant Value Function (b) Incumbents Distribution

Note. Figure 2.9a on the left shows the value of entry in 1980 and in 2015 both for the general equilibrium version of the model and for the partial equilibrium one. The light blue line shows the value of entry in 1980, the orange line shows the value of entry in 2015-GE, and the light grey line shows the value of entry in the 2015-PE. The value of entry declines between 1980 and 2015 because firms to grow in the intangible-intensive economy have to spend more resources on high adjustment costs. The value of entry decline more in PE relative to GE because in general equilibrium the wage decline and acts as a dampening force to the effect of IBTC.
Figure 2.9b on the right shows the endogenous distribution of firms in the economy in 1980 and in 2015 both for the general equilibrium version of the model and for the partial equilibrium one. The light blue line shows the distribution in 1980, the orange line shows the distribution in 2015-GE, and the light grey line shows the distribution in the 2015-PE. The distribution shifts to the right because of the increase in selection mentioned above. Again, the decline in wages dampens the PE effect resulting in a milder shift of the GE distribution towards the right.

However, once we allow the wages to adjust endogenously, the GE value of entry $\mathcal{V}^e_{2015,GE}$ increases and ends up being much higher relative to $\mathcal{V}^e_{2015,PE}$ level. This is due to a decline in the wages that arises due to an endogenous decline in the overall labor demand by firms. This is an artifact of the reduced firm entry, reduction in the output elasticity of labor at the firm-level, and the increase in the overall adjustment cost of investment faced by firms. This wage decline, which is a counter-balancing force to the selection effect due to IBTC, increases the value of entry and profit rates relative to the partial equilibrium level.

The model-implied decline in wage in conjunction with a decline in the labor supply between 1980 and 2015 is supported in the data. There is ample evidence of the stagnation of wages in the lower half of the distribution (median wage in 1982 prices increased from USD 330 to USD 345 from the 1980s). Furthermore, in the past three decades, labor force participation has declined from a peak of 67% to roughly 63%.

This exercise highlights the importance of general equilibrium effects in pinning down the overall macroeconomic implications of IBTC, without which one would have significantly overestimated the role of IBTC in explaining the recent trends that are the main focus of this paper.

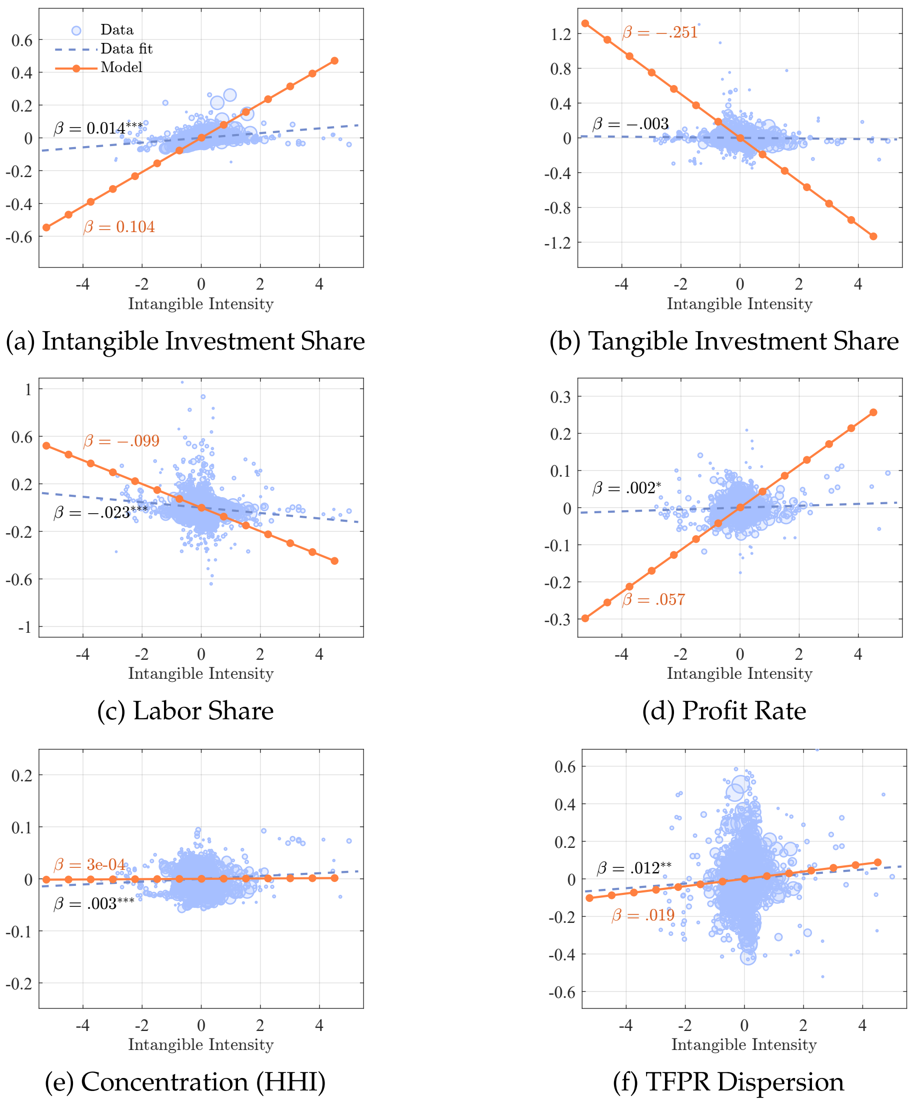
2.6.3 *Cross-Sectoral Validation*

This section contains a validation of the mechanism described in the previous sections. Here, we test the model predictions about how different intangible capital intensities in production, defined as the ratio of intangible capital share to labor share, shape sector-level factor shares, concentration, and allocative efficiency. However, as pure technological intangible capital intensity in production is difficult to measure at the sector level, we use a robust model prediction and we proxy it by the ratio of intangible capital to labor costs share.

Figure 2.10 shows the results. Dashed light blue lines show the data linear fit whereas orange lines with circles show the model predictions.[24] We focus on six main observables of interest: (i) Intangible investment share, (ii) tangible investment share, (iii) labor share, (iv) profit rate, (v) concentration, (vi) $TFPR$ dispersion.

[24]To obtain the model predictions, due to the high-level non-linearities and the different dispersion in intangible intensity between the model and data, we perturbed the model around the steady-state and infer the associated slope, then, we use the inferred slopes to extrapolate the overall tendency.

FIGURE 2.10: Sector-Level Correlations: Model vs Data



(a) Intangible Investment Share

(b) Tangible Investment Share

(c) Labor Share

(d) Profit Rate

(e) Concentration (HHI)

(f) TFPR Dispersion

Note. The figure shows the cross-sectoral correlations between intangible intensity, $k_I/w\ell$, and various measures of interest. Light blue bubbles show the sector-year observations net of sector and time fixed effects. Sectors are defined at SIC2-level. The light blue dashed line shows the empirical fit. The solid orange line with circles shows the model implied slope.

In the model, a rise in intangible capital intensity translates into a higher investment share of intangible capital relative to the other inputs whose shares instead decline. Therefore, the economy moves from labor which is a highly flexible input to intangible capital which is highly distorted due to the presence of technological frictions. This translates into a decrease in the allocative efficiency of the economy as measured by the dispersion in $TFPR$. Finally, as investing in intangible capital is a costly activity due to the associated high adjustment costs, selection increases that in turn increase both the market concentration, as measured by the HHI index, and the overall profit rate. Overall Figure 2.10 shows that all the qualitative predictions of the model are in line with the data from the cross-section of sectors.

2.7 INTANGIBLE CAPITAL BIASED TECHNOLOGICAL CHANGE AND THE MACROECONOMICS IMPLICATIONS

In this section, we study the quantitative implications of the IBTC as documented in Section 2.3.1. First, we document the firm-level and macroeconomic implications of the IBTC. Second, we document that our results are robust to alternative quantification exercises. Third, we discuss the relation between IBTC, market power, and the policy implications.

2.7.1 *Quantitative Implications*

Here, we study the quantitative implications of the IBTC documented in Section 2.3.1. In particular, we show the quantitative implications of a rise in intangible capital share in firm-level production from 0.03 to 0.12 and of an associated decline in labor share from 0.69 to 0.60 as estimated in the data for the period 1980−2015.[25]

Table 2.4: Quantitavie Implications of IBTC

	1980 S.S.	2015 S.S.	Change Model	Change Data
Firms Distribution				
Avg. firm size	23.523	26.121	+11%	+15%
Concentration	7.14e-04	9.88e-04	+38%	+33%
Employment share				
firms with 250+ employees	0.489	0.551	+6p.p.	+6p.p.
Aggregate Factor Shares				
Intangible				
investment share	0.014	0.055	+4p.p.	+4p.p.
Tangible				
investment share	0.078	0.070	−1p.p.	−2p.p.
Labor share	0.666	0.580	−9p.p.	−8p.p.
Labor share				
pre-revision	0.676	0.614	−6p.p.	−5p.p.
Profit rate (Compustat)	0.242	0.294	+5p.p.	+3p.p.
Profit rate (BEA)	0.242	0.294	+5p.p.	+5p.p.
Aggregate Investment Rate				
Tangible				
investment rate	0.052	0.041	−1p.p.	−2p.p.
Allocative Efficiency				
sd($TFPR$)	0.202	0.227	+12%	+38%
Adjusted sd($TFPR$)	0.202	0.227	+12%	+15%

Notes. All the variable are calculated coherently to their definitions as used in the data. The data sources are BDS, NIPA tables, and Compustat. To calculate the empirical moments from the 1980s we use the time window 1980-1990, whereas, for the empirical moments from 2015 we use simple the values in that year. The evolution of each trend is presented in Appendix 2.A.5.

Table 2.4 shows the results.[26] Looking at the firm-level moments, we

[25]We leave the tangible capital share unchanged as it does not show any particular trend over the period of interest.

[26]In Appendix 2.B.2 we document the evolution of the distribution of the firm-level

can see that the IBTC explains the majority of the observed rise in the average firm size and of the rise in concentration, both as measured by the HHI and as measured by the employment share of firms with $250+$ employees. These results are driven by the exogenous technological change in the production process and the endogenous rise of selection in the model as discussed in the previous section.

Then, we compare the quantitative implications of the IBTC with the changes in factor shares documented by Koh, Santaeulàlia-Llopis, and Zheng (2020). The model captures well the change in most of the factor shares in the non-financial corporate sector.[27] To study the implications of the rise of intangible capital on the decline in the labor share, we follow Koh, Santaeulàlia-Llopis, and Zheng (2020) and compute two different labor shares in the model given by:

$$LS \equiv \frac{WL}{Y} \quad \text{and} \quad LS\,\text{pre-revision} \equiv \frac{WL}{Y - X_I}, \tag{2.18}$$

where W is the wage, L is aggregate labor, Y is aggregate output net of adjustment costs and fixed costs, and X_I is the aggregate intangible investment. The labor share pre-revision is the counterfactual labor share that would emerge if intangible capital was not counted in the overall GDP calculation. Similar to the empirical evidence in Koh, Santaeulàlia-Llopis, and Zheng (2020), we find that the pre-revision labor share decline much less than the true labor share. This finding confirms the interpretation of the authors that rising intangible capital investment is a quantitative important factor in the decline of the labor share observed in the data. Moreover, the model can explain satisfactorily the rise in intangible capital investment share, the decline in the tangible investment share, and the rise in the profit rate. We find a lower increase in the profit rate in the data compared to De Loecker

intangible intensity and TFPR both in the model and in the data.

[27] We focus on the non-financial corporate sector as this is the best mapping with the Compustat data we used in the empirical analysis.

et al. (2020), this is because we also account for balance sheet intangible investment as documented in Appendix 2.A.5.

Finally, when looking at the overall allocative efficiency of the economy, we see that the IBTC can explain a substantial share of its downward trend − notice that an increase in the standard deviation of $TFPR$ translates into a decline in allocative efficiency. In the model, this is because when firms rely more on an input that is highly distorted as intangible capital relative to a flexible input like labor, then inputs become slower in reallocating towards more productive firms, and hence the overall allocation of resources worsen as measured by an increase in the standard deviation of $TFPR$. However, we emphasize that this cannot be considered as misallocation as the economy is fully efficient and the allocation of resources coincides with the one of the social planner. Concluding, the model does a good job in matching the quantitative decline in allocative efficiency, particularly in the case of the adjusted one.[28]

2.7.2 *Robustness Checks*

In this section, we perform two additional exercises to test the validity of our results from Table 2.4. In particular, we check (i) how the decline in the intangible capital investment price relative to the tangible capital investment price affect our baseline findings and (ii) how results change if we were to re-estimate the adjustment costs with moments investment rate distribution computed with the final part of the sample.[29]

We estimate the relative price of intangible capital as the ratio of the intangible capital deflator to the tangible capital deflator. We find that between 1980 and 2015 this relative price experienced a decline

[28] Adjusted allocative efficiency is measured as allocative efficiency net of a potential 60% measurement error as documented by Bils, Klenow, and Ruane (2020).

[29] All these additional robustness exercises are made on to the IBTC exercise. Meaning that we always re-estimate the model with IBTC together with one of the aforementioned changes.

of approximately 20%, suggesting that intangible capital is becoming cheaper relative to tangible capital. To introduce this decline into the model, we substitute into the value functions 2.8, 2.9, 2.10, 2.11, and 2.12 the relative price of intangible capital investment p such that the final intangible capital investment bill is px_I.[30]

Instead, to perform the second exercise we just re-estimate the capital adjustment costs to match moments from the investment rate distribution of both capitals between the periods 2000-2015. Table 2.A.2 shows the evolution of the investment rate distribution over time, whereas, Table 2.B.1 in Appendix 2.B.3 shows the new calibrated parameters and the associated targeted moments.

Table 2.5 shows the results. The first column shows the benchmark results as reported also in Table 2.4. The second and the third column show the results obtained re-estimating the adjustment costs and the results obtained by accommodating the decline in the relative price of intangible capital investment. The final column instead reports the values from the data. Both robustness exercises show the same qualitative patterns as in the benchmark case and, overall, results seem to be robust to these departures from the benchmark case. Moreover, we notice that even quantitatively results do not seem to deviate significantly from these alternative specifications. Effectively, what really matters for our results is that technology is just shifting towards an input whose sunk cost of adjusting it (or of using it) is relatively higher compared to the other inputs. Therefore, we conclude that our results are robust and that they just hinge on the main properties of estimated technology (both the production technology and the adjustment costs technology).

2.7.3 *IBTC, Market Power, and Policy Implications*

In our framework, as production technology becomes more intangible-intensive, firms invest more in an input that entails higher adjustment

[30]Therefore, in our quantitative experiment, we let the price p to go from 1 in 1980 to 0.8 in 2015.

Table 2.5: Quantitavie Implications of IBTC

	Change			
	Benchmark	Alternative Adj. Costs	Decline Rel. Price k_I	Data
Firms Distribution				
Avg. firm size	+11%	+13%	+5%	+15%
Concentration (HHI)	+38%	+36%	+26%	+33%
Employment share firms with 250+ employees	+6p.p.	+7%	+6p.p.	+6p.p.
Aggregate Factor Shares				
Intangible investment share	+4p.p.	+4p.p.	+4p.p.	+4p.p.
Tangible investment share	−1p.p.	−1p.p.	−0p.p.	−2p.p.
Labor share	−9p.p.	−9p.p.	−9p.p.	−8p.p.
Labor share pre-revision	−6p.p.	−6p.p.	−5p.p.	−5p.p.
Profit rate (Compustat)	+5p.p.	+5p.p.	+4p.p.	+3p.p.
Profit rate (BEA)	+5p.p.	+5p.p.	+4p.p.	+5p.p.
Aggregate Investment Rate				
Tangible investment rate	−1p.p.	−1p.p.	−1p.p.	−2p.p.
Allocative Efficiency				
sd($TFPR$)	+12%	+12%	+11%	+38%
Adjusted sd($TFPR$)	+12%	+12%	+11%	+15%

Notes. All the variable are calculated coherently to their definitions as used in the data. The data sources are BDS, NIPA tables, and Compustat. To calculate the empirical moments from the 1980s we use the time window 1980-1990, whereas, for the empirical moments from 2015 we use simple the values in that year.

costs. Although this technological change raises market concentration, firm size, and the aggregate profit rate, resources are still allocated efficiently across firms. The observed decline in allocative efficiency in the model is due to technological constraints, and therefore, the decentralized equilibrium allocation still coincides with the one provided by the social planner. Our paper suggests that a sizeable part of the macroeconomic changes that have been witnessed in the US economy are the by-product of an efficient technological change.

However, this conclusion does not exclude that other forces above and beyond the mechanism documented here are at play in the economy. For instance, consider a slightly different version of our baseline model. Instead of assuming that firms produce the same good, we could have allowed firms to produce differentiated goods aggregated a lá Kimball as in Edmond, Midrigan, and Xu (2018). In such a framework, markups would be positively correlated with firm size. Therefore, a technological change that favors larger firms would shift market shares toward high-markup firms and away from low-markup firms. The measured decline in allocative efficiency in the model would be magnified by the rise in the dispersion of markups on top of the one already generated by the IBTC. Moreover, even though in this alternative framework the decentralized allocation would not coincide with the one provided by the social planner, the implementation of the social planner allocation, through any potential optimal policy, would coincide with the allocation in our baseline framework. As a consequence, while extended frameworks could give rise to desirable policy interventions, our work suggests that at least a significant part of the macroeconomic trends that we observe in the US economy could be the by-product of efficient response to changes in the firm-level production technology.

2.8 CONCLUSION

In the last four decades, firm-level investment in intangible capital, such as research and development, intellectual property products, and computerized information, has dramatically increased in the US. However, still little is known about its intrinsic properties and its implications for the economy overall. In this paper, we made a step forward in the understanding of this new capital.

We estimate the firm-level production function finding that intangible capital is an important input in production and that its input share has gone from 0.03 in the 1980s to 0.12 in 2015. Moreover, we document that most of this rise has happened at the expense of the labor share in production. We interpret these findings as a paradigm shift in the production process of US firms, for instance see the importance that software and other intellectual property products have increasingly gained in the economy. We refer to transformation in the firm-level production process as Intangible Capital Biased Technological Change (IBTC).

Then, we document some novel properties of intangible capital, particularly, the fact that this new capital entails higher adjustment costs compared to tangible capital. This is consistent with the view that investments in intangible capital are plagued by inherent indivisibilities and are often sunk.

Finally, using a structural model of investment dynamics, we find that this technological change can jointly explain a sizeable fraction of the increase in average firm size, the increase in concentration, the change in the aggregate factor shares, the decline in the tangible capital investment rate, and the decline in allocative efficiency. Our findings bring intangible capital, its properties, and its trends at the center of the macroeconomic transformations that have been witnessed in the US economy. Therefore, we hope this paper will spur new exciting research on this topic.

2

APPENDIX

2.A EMPIRICAL APPENDIX

2.A.1 *Data*

2.A.1.1 Main Sample, Variables, and Summary Statistics

We use Compustat from 1980 to 2015. We linearly interpolate `SALE`, `COGS`, `XSGA`, `EMP`, `PPEGT`, `PPENT`, `XRD`, `INTAN`, `GDWL`, `AM`. We exclude utilities (SIC codes between $4900-4999$) because they are heavily regulated on prices and I also exclude financial firms (SIC codes between $6000-6999$) because their balance sheets are dramatically different from other firms.

For data quality, we interpret as mistakes if `SALE`, `PPEGT`, `PPENT`, `COGS`, `EMP`, or `XSGA` are zero, negative, or missing and we drop that observations, moreover, if `XSGA` is missing or negative we drop it as well. Finally, if `XRD`, `INTAN`, `AM`, or `GDWL` are negative or missing we treat them as zeros. To obtain a real measure of the main variables we deflate them with the GDP deflator, we deflate investment in tangible and intangible capital by the appropriate deflators.[31] The table below presents a few basic summary statistics for a few leading variables used in our analysis.

Table 2.A.1: Summary Statistics (1980−2015)

	Sales	Cost of Goods Sold	Employment	Tangible Capital Stock	Intangible Capital Stock
Mean	2,310,810	1,572,800	7,966	1,572,164	284,519
25$^{\text{th}}$ Percentile	27,495	14,880	131	8,004	2
Median	153,005	89,241	686	51,066	3,098
75$^{\text{th}}$ Percentile	809,728	510,199	3,625	349,551	34,060
No. Obs.	188,151	188,151	188,151	188,151	188,151

Note. Summary statistics of cleaned Compustat dataset between 1980 and 2015. All variables are in thousands US$. Sales and Costs of Goods Sold are deflated with the GDP deflator with base year 2012, whereas, both capital stocks are deflated using the appropriate investment deflator with base 2012.

2.A.1.2 User Cost of Tangible and Intangible Capital

One of the challenges of using the cost shares approach to estimate the firm-level production function is that it requires a measure of the user cost of capital. To this end, we define the user cost of capital as:

$$r_{j,t} = i_t - \mathbb{E}_t \pi_{t+1} + \delta_j, \quad j \in \{T, I\} \tag{2.19}$$

where i_t equals the nominal interest rate, $\mathbb{E}_t \pi_{t+1}$ is expected inflation at time t, and δ_j is the capital-specific depreciation rate. We take the annual Moody's Seasoned Aaa Corporate Bond Yield as an empirical proxy of the nominal interest rate, the annual growth rate of the Investment Nonresidencial Price Deflator to calculate expected inflation, the depreciation rate of tangible capital is calibrated to $\delta = 0.07$,

[31] Deflators are taken from the NIPA tables.

and the firm-level depreciation rate of intangible capital is computed as a weighted average of the depreciation rates used to construct the intangible capital stock.[32,33,34,35]

2.A.1.3 Intangible Capital Measurement and Accounting Standards

Measuring intangible capital is a difficult task as the accounting standards (US GAAP) are insufficient to satisfactorily book the intangible assets on the balance sheets. It is well established in the corporate finance literature that intangible assets are not fully captured on the firms' balance sheet due to the anachronism of the US GAAP. [36] In this section of the appendix we explains in detail which assumptions are

[32] Moody's Seasoned Aaa Corporate Bond Yield: https://fred.stlouisfed.org/series/AAA

[33] Investment Price Deflator: https://fred.stlouisfed.org/series/A008RD3Q086SBEA

[34] We estimate an AR(1) process on the annual growth rate of the Investment Nonresidential Price deflator and define the contemporaneous expected inflation as $\mathbb{E}_t \pi_{t+1} = \mu + \rho \pi_t$.

[35] The firm-level depreciation rate of intangible capital is computed as:

$$\delta_{I,ft} = \frac{k_{ft}^{R\&D}}{k_{ft}^{R\&D} + k_{ft}^{BS}} \delta_s^{R\&D} + \frac{k_{ft}^{R\&D}}{k_{ft}^{R\&D} + k_{ft}^{BS}} 0.20.$$

[36] Lev and Gu (2016) say:

> *Revolutionary changes, shifting economies and business enterprises from the industrial to the information age, started to profoundly affect the business models, operations, and values of companies in the 1980s, yet, amazingly, triggered no change in accounting. Entire industries, which are largely intangible (conceptual industries, as Alan Greenspan called them), including software, biotech, and internet services, came into being during the 1980s and 1990s. And for all other businesses, the major value drivers shifted from property, plant, machinery, and inventories, to patents, brands, information technology, and human resources. The latter set, all missing from companies? balance sheets because accountants treat intangible investments like regular expenses (wages, or interest), thereby distorts both the balance sheet and income statement. The constant rise in the importance of intangibles in companies? performance and value creation, yet suppressed by accounting and reporting practices, renders financial information increasingly irrelevant.*

needed to compute intangible capital at firm-level using balance sheet for stocks and income statements for flows.

To introduce our main measure, we have to clarify that intangible capital is intrinsically different from tangible capital as a significant part of it is internally generated by the firms. For nearly all internally generated intangible assets, such as knowledge and organizational capital, accounting standards differ significantly from tangible assets. All purchases of tangible assets are recorded on the balance sheet at their purchased price and depreciated over their useful life. On the contrary, internal intangible capital investments as firms' R&D expense, advertising, or training of employees are fully expensed in the period incurred.[37]

FIGURE 2.A.1: Advertising Expenses of Coca Cola

Selling, General and Administrative Expenses

The following table sets forth the significant components of selling, general and administrative expenses (in millions):

Year Ended December 31,		2016		2015		2014
Stock-based compensation expense	$	**258**	$	236	$	209
Advertising expenses		**4,004**		3,976		3,499
Selling and distribution expenses		**5,177**		6,025		6,412
Other operating expenses		**5,823**		6,190		7,098
Selling, general and administrative expenses	$	**15,262**	$	16,427	$	17,218

For instance, the Coca-Cola Company spends several billion dollars each year to maintain and promote its products, and brands, such as Coca-Cola and Dasani that are assets for the firm that are going to generate future benefits in the form of higher margins and increased sales volume. However, the Coca-Cola Company is not allowed to recognize these assets in its balance sheet. Figure 2.A.1 shows that Coca-Cola spent around four billion dollars in advertising in 2016. We also provide the example of Google Inc., that spent around sixteen

[37]However, there are some exceptions. For example, U.S. GAAP treats the development of computer software differently from other R&D costs. Following the ASC 985 (formerly FAS 2), once a software developer has reached technological feasibility, the developer must capitalize and amortize all development costs until the product becomes available for general release to consumers.

billion dollars in research and development and twelve billion dollars in sales and marketing (see Figure 2.A.2a and Figure 2.A.2b).

FIGURE 2.A.2: Intangible Investments by Google

Research and Development

The following table presents our R&D expenses (in millions):

	Year Ended December 31,		
	2015	2016	2017
Research and development expenses	$ 12,282	$ 13,948	$ 16,625
Research and development expenses as a percentage of revenues	16.4%	15.5%	15.0%

R&D expenses consist primarily of:

- Compensation expenses, including SBC, and facilities-related costs for employees responsible for R&D of our existing and new products and services; and
- Depreciation and equipment-related expenses.

(a) Research and Development Expenses

Sales and Marketing

The following table presents our sales and marketing expenses (in millions):

	Year Ended December 31,		
	2015	2016	2017
Sales and marketing expenses	$ 9,047	$ 10,485	$ 12,893
Sales and marketing expenses as a percentage of revenues	12.1%	11.6%	11.6%

Sales and marketing expenses consist primarily of:

- Advertising and promotional expenditures related to our products and services; and
- Compensation expenses, including SBC, and facilities-related costs for employees engaged in sales and marketing, sales support, and certain customer service functions.

(b) Marketing Expenses

Overall, what these figures prove is that there is a lot of intangible capital investment that is simply expensed by the firms, in accordance with the US GAAP, that does not show up as capital in the balance sheet. To overcome this limitation in the accounting standards we capitalize knowledge capital as explained in Section 2.2.2.

Externally acquired intangible capital can be capitalized on firms' balance sheet at the fair value according to the US GAAP under guidelines provided from ASC 350 (formerly FAS 142) and shows up in Compustat in the variable `INTAN`. According to Ewens, Peters, and Wang (2019), firms and accountants follow the guideline provided in

ASC 820 (formerly FAS 157) to mark externally acquired intangible capital in the balance sheet at the fair value at the time of the acquisition. Firms can choose among different methods to compute the fair value according to the US GAAP and firms' choice must be disclosed in the appraisal notes for intangibles in the buyer's financial statements. Firms have three option to appraise the value of intangible assets: (i) Estimating the replacement cost of the asset, (ii) comparing the asset to a similar asset whose price trades on the open market, or (iii) using the Discounted Cash Flow model, where earnings or cash flows are discounted by an appropriate discount rate. In particular, acquired intangible assets can be individually capitalized with the methodologies reported above if and only if they are identifiable, as documented in the ASC 805 notes. An intangible asset is identifiable if it meets (i) either the separability criterion, meaning it can be separated from the entity and sold or (ii) the contractual−legal criterion, meaning that the control of the future economic benefits arising from the intangible asset is warranted by contractual or legal rights. In other words, IIA prices reflect fair or public value, rather than value specific to the post−acquisition firm. Some examples of these identifiable intangible assets include brand names, customer lists, trademarks, internet domain names, royalty agreements, patented technologies, and trade secrets. Other intangibles with a non−zero value, such as corporate culture, advertising effectiveness, management quality, that fail to meet these criteria for identification are captured as goodwill in the buyer's balance sheet, `GDWL` in Compustat.

We give an example of Coca-Cola's externally purchased Intangibles in Figure 2.A.3. Coca-Cola says in their yearly report that:

> *We classify intangible assets into three categories: (1) Intangible assets with definite lives subject to amortization, (2) intangible assets with indefinite lives not subject to amortization and (3) goodwill.*

FIGURE 2.A.3: Coca-Cola's Externally Purchased Intangibles

Coca-Cola Co.

Balance sheet: goodwill and intangible assets

US$ in millions

	Dec 31, 2019	Dec 31, 2018	Dec 31, 2017	Dec 31, 2016	Dec 31, 2015
Trademarks	9,266	6,682	6,729	6,097	5,989
Bottlers' franchise rights	109	51	138	3,676	6,000
Goodwill	16,764	10,263	9,401	10,629	11,289
Other	110	106	106	128	164
Indefinite-lived intangible assets	**26,249**	**17,102**	**16,374**	**20,530**	**23,442**
Customer relationships	344	185	205	392	493
Bottlers' franchise rights	341	30	213	487	604
Trademarks	177	186	182	228	211
Other	55	88	94	179	97
Definite-lived intangible assets, gross carrying amount	**917**	**489**	**694**	**1,286**	**1,405**
Accumulated amortization	(400)	(321)	(432)	(688)	(715)
Definite-lived intangible assets, net	**517**	**168**	**262**	**598**	**690**
Intangible assets	**26,766**	**17,270**	**16,636**	**21,128**	**24,132**

Based on:10-K (filing date: 2020-02-24),10-K (filing date: 2019-02-21),10-K (filing date: 2018-02-23),10-K (filing date: 2017-02-24),10-K (filing date: 2016-02-25).

The goodwill and intangible assets with indefinite lives are subject to impairment test every period and their values are increased or decreased accordingly. As one can see, the balance sheet intangibles are the sum of heterogeneous assets, such as trademarks, franchise rights, customer relationships, among others.

Internally generated intangible capital: Potential issues. The fact that a sizeable fraction of intangible capital is internally produced and cannot be capitalized in firms' balance sheet potentially implies that there could be some concerns related to double-counting of some intangible assets. For example, when firm 1 produces its own intangible capital it will expense it in the income statement at the production cost x ; if this intangible capital gets then sold to firm 2 this will not show up in the income statement of firm 1 as a negative cost (or a negative investment), however, firm 2 will how this new intangible capital in its balance sheet at the fair value y, as it has been externally acquired. In this example, despite the fact that the overall amount of intangible capital has not changed in the economy $-$ as there has been just a transaction $-$ we would potentially observe an increase in the overall stock of intangible capital from x to $x + y$.

Despite the fact that this is in theory a concern, practically, we are confident that this is a rare situation and hence of little quantitative relevance. First, we know that often intangible capital is acquired through the acquisition of an entire firm.[38] Hence, as the target firm is acquired, it exits the sample and its intangible assets leave the sample as well − while now the acquiring firm will have an increase in its intangible capital in the balance sheet. Second, we also know that a lot of intangible capital is acquired as final goods from other firms (think for instance at software producers and advertisement/marketing companies) and in this case as well there is no double-counted as this is final production and not internal production for firms' own usage. Third, as showed in the previous section, internally produced intangible capital is a declining feature of our empirical measure, suggesting that this concern should be minor and declining over time. Therefore, we conclude that this issue is not quantitative appealing despite being of difficult solution.

Externally acquired intangible capital: Potential issues. Externally purchased intangible capital is almost often acquired through acquisitions of entire firms and this greatly influences the way it is capitalized in the firms' balance sheet. For example, imagine firm x buys firm y and it pays it p^y. At the moment of the acquisition, firm x has to place the acquired assets in its balance sheet. Normally, the procedure is the following: (i) Tangible assets are identified and capitalized at the fair value p^T, (ii) then identifiable intangible assets are capitalized at the fair value p^I, and (iii) the residual value is attributed to unidentifiable intangible assets (synergies, organizational culture, etc.) and is capitalized into goodwill. Therefore, in the data we have $\texttt{GDWL} = p^y - p^T - p^I$.

If a researcher thinks that firms acquire other firms to exercise future market power (and so firms are willing to pay high prices for them)

[38] Peters and Taylor (2017) and Ewens, Peters, and Wang (2019).

there can be the concern that these unidentifiable intangible assets are just the discounted expected sum of the value of future market power − and therefore balance sheet intangibles value goes up by more than its quantity. One way to address this concern is to use proper deflators, that is, to deflate intangible capital with the IPP deflator.[39] However, this takes care only of aggregate common trends and cannot account for the heterogeneity of firm-level input prices, and unfortunately more disaggregated investment deflators do not exist. We wish to emphasize that the inability to obtain firm-level investment deflators affects equally the measurement of tangible and intangible capital. Additionally, as a more appealing way to address these concerns, we remove goodwill from total balance sheet intangible capital as almost all the potential rise in prices related to unidentifiable assets is going to be capture exactly by a rise in goodwill. However, we want to acknowledge that Ewens, Peters, and Wang (2019) − using more detailed data than we have − have shown that at least 38% of firms' goodwill is indeed true intangible capital. Therefore, we see this solution as a necessary but imperfect solution.

Accounting standards for software: A special case. The accounting standards for expenditures in internal software development or in external purchases are different from that of other intangible assets. In particular, the FASB ASC subtopic 350−40 provides guidelines for the accounting of the costs for computer software developed or obtained for internal use and of the hosting arrangement obtained for internal use. The standards state that costs incurred during the development stage may be capitalized. Capitalization of the costs should cease in the post implementation stage. The FASB ASC subtopic 985−20 provides guidelines for the accounting of the costs incurred for software meant to be sold, leased or marketed. The standards state that costs incurred subsequent to the establishment of technological feasibility may be capitalized. Capitalization of the costs should cease when the software

[39]This is standard practice in empirical work based on firm-level data.

is available for the general release to customers.

FIGURE 2.A.4: Software Capitalization of Athena Health

6. CAPITALIZED SOFTWARE COSTS

Capitalized software consisted of the following:

	As of December 31, 2017	As of December 31, 2016
Capitalized internal-use software development costs	$ 113.9	$ 122.7
Acquired third-party software licenses for internal use	53.8	47.5
Total gross capitalized software for internal-use	167.7	170.2
Accumulated amortization	(74.8)	(82.9)
Capitalized internal-use software in process	46.8	38.5
Total capitalized software costs	$ 139.7	$ 125.8

Capitalized software amortization expense totaled $71.3 million, $73.5 million, and $53.4 million for the years ended December 31, 2017, 2016, and 2015, respectively. Future amortization expense for all capitalized software placed in service as of December 31, 2017 is estimated to be:

Years ending December 31,	Amount
2018	$ 50.5
2019	25.4
2020	10.7
2021	5.5
2022	0.8

To illustrate this, we provide an example of Athena Health Inc. software investments. The company has capitalized software development costs for USD 113.9 million in 2017 and report, external software acquisitions for USD 53 million.

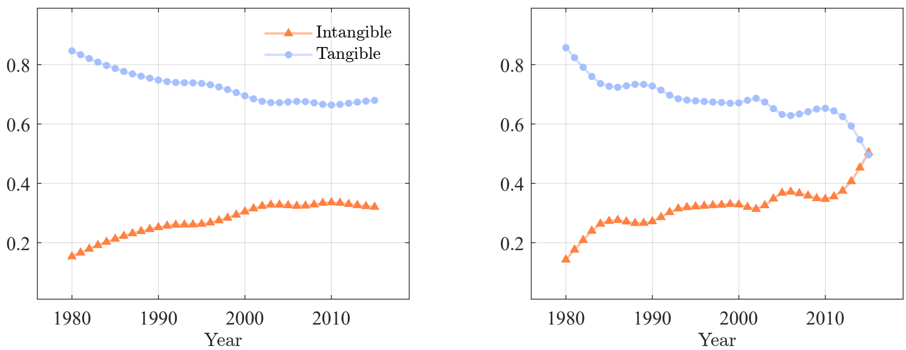
Software used in research and development are subject to the subtopic 730−10. In general, in case of software that are purchased from others and used for research and development activities and that have alternative futures/uses, should be capitalized and amortized as an intangible assets. However, the costs of software that are purchased from others for a particular research and development project and that have no alternative uses and therefore no separate economic values are considered research and development costs and have to be expensed at the time they are incurred.

In any case we would capture most of the intangible capital related to software in our measure throughout balance sheet intangible capital or throughout capitalized knowledge capital.

2.A.1.4 Additional Validations Firm-Level Intangible Capital

Here we compare some additional trends, related to intangible capital investment, between aggregate data from BEA and our measure from Compustat. Figure 2.A.5 compares the share of tangible capital investment into total investment and the share of intangible capital investment into total investment both in the BEA data and in the Compustat data between 1980 and 2015. We can see that both data sources tell a similar story: In 1980 most of the investment was in tangible capital, whereas by 2015 tangible investment is roughly 70% of total investment in BEA and is 50% in Compustat. Despite the two data sources tell similar stories they also show some discrepancy. In Compustat the decline in the share of tangible capital investment of total investment is more pronounced; this could be due to for instance to (i) undercapitalization of true IPP capital in BEA or to (ii) selection of intangible-intensive firms in Compustat.

FIGURE 2.A.5: Investment Components Share



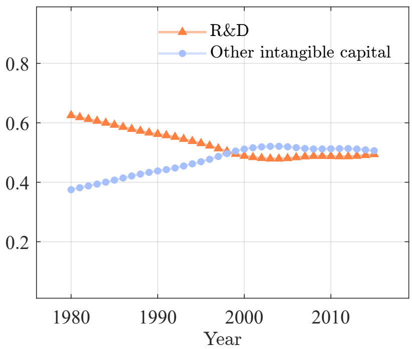
(a) BEA

(b) Compustat

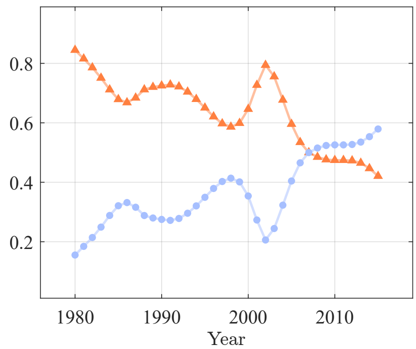
Note. The figures show the evolution of the share of tangible capital investment and of intangible capital investment over total investment both in BEA data and in Compustat data for the period 1980–2015. The data are detrended with an Hpfilter with $\lambda = 6.25$.

Figure 2.A.6 shows the evolution of the different components of intangible capital investment both in BEA and in Compustat for the pe-

FIGURE 2.A.6: Intangible Capital Components Share



(a) BEA



(b) Compustat

Note. The figures show the evolution of the share of knowledge capital investment (R&D) and of others intangible capital investment (intangible capital investment different from R&D) over total intangible capital investment both in BEA data and in Compustat data for the period 1980–2015. The data are detrended with an Hpfilter with $\lambda = 6.25$.

riod 1980-2015. Again the two data sources show a similar tendency: In 1980 most of intangible capital investment was investment in research and development whereas by 2015 investment in research and development accounts for less than 50% of total intangible capital investment.

Finally, in Figure 2.A.7 we compare the evolution of the intangible capital investment share across different sectors for both BEA data and Compustat data for the period 1998–2015. The sector-level intangible capital investment shares emerging from the Compustat data show similar trends with the one computed with the BEA data. However we see some difference in the level within some sectors. It is difficult to know what are the sources of these discrepancies; overall we conclude that our firm-level measure of intangible capital does a reasonable good job in capturing the tendencies that are present in the aggregate data.

2.A.2 *Production Function Estimation*

To estimate the firm-level production function, we follow De Loecker, Eeckhout, and Unger (2020) and use two main approaches: (i) The control function approach and (ii) the cost shares approach. Both of them are popular methods used to estimate firm-level production functions. We review here the two methodologies emphasizing their virtues and their limitations.

2.A.2.1 Ackerberg-Caves-Frazer

The control function approach has been pioneered by Olley and Pakes (1996), and developed further by Levinsohn and Petrin (2003) and Ackerberg, Caves, and Frazer (2015). The main insight from this literature is that firm-level unobservable productivity can be proxied by some variable expenditure.

To overcome some of the criticism emphasized in Gandhi, Navarro, and Rivers (2020) we work with a structural value added specification as in Ackerberg, Caves, and Frazer (2015) and De Loecker and Scott (2016), given by:

$$\mathcal{Q}_{ft} = \min\Big\{K^{\alpha}_{T,ft}K^{\nu}_{I,ft}L^{1-\alpha-\nu}_{ft}\exp(\omega_{ft}+\varepsilon_{ft}),\ \beta M_{ft}\Big\}, \tag{2.20}$$

where $\mathcal{Q}_{ft}$ is output, $K_{T,ft}$ is tangible capital, $K_{I,ft}$ is intangible capital, L_{ft} is labor, ω_{ft} is log-productivity, ε_{ft} is the error term, and M_{ft} is material. This structural value added production function yields the following first order condition:

$$\mathcal{Q}_{ft} = K^{\alpha}_{T,ft}K^{\nu}_{I,ft}L^{1-\alpha-\nu}_{ft}\exp(\omega_{ft}+\varepsilon_{ft}), \tag{2.21}$$

justifying the regression of $\mathcal{Q}_{ft}$ on tangible capital, intangible capital, and labor while ignoring materials. A caveat is that, in theory, equation 2.21 may not be satisfied in certain situations. If both capitals and labor are quasi-fixed, and materials are a flexible input, then when output prices are sufficiently low relative to the price of materials, it will be better to set $M_{ft} = 0$ and not produce at all. However, given that our data only includes actively producing firms, we assume that equation

2.21 always holds.[40] Therefore, under the specification in equation 2.20 the estimation of the firm-level production function reduces to:

$$q_{ft} = \alpha k_{T,ft} + \nu k_{I,ft} + (1 - \alpha - \nu)\ell_{ft} + \omega_{ft} + \varepsilon_{ft}, \tag{2.22}$$

where $q_{ft} = \log(\mathcal{Q}_{ft})$, $k_{T,ft} = \log(K_{T,ft})$, $k_{I,ft} = \log(K_{I,ft})$, and $\ell_{ft} = \log(L_{ft})$. As usual, the main identification challenge to the production function estimation is the simultaneity bias induced by the unobserved time-varying firm-level productivity, ω_{ft}. We follow the control function literature, and in particular Ackerberg, Caves, and Frazer (2015) and De Loecker, Eeckhout, and Unger (2020), to estimate the production function in 2.22 using a two step approach based on the use of a control function for the productivity process. The identification relies on the observation that tangible capital investment demand of the firm is given by a policy function of the form $x_{T,ft} = x_T(k_{T,ft}, k_{I,ft}, \omega_{ft})$. Then, providing that the policy function is invertible, the productivity process can be proxied by a control function given by $\omega_{ft} = \omega(k_{T,ft}, k_{I,ft}, \omega_{ft})$ where $\omega(\cdot) = x_T^{-1}(\cdot)$.[41]

Therefore, in the first stage of this estimation procedure, we can clean firm-level output value from measurement errors and unanticipated productivity shocks regressing output on a polinomial of tangible capital, intangible capital, labor, and potential demand shifters given by:

$$q_{ft} = \mathcal{P}_t(k_{T,ft}, k_{I,ft}, \ell_{ft}, \mathbf{d}_{ft}) + \varepsilon_{ft}. \tag{2.23}$$

Then, in the second stage, using the estimate $\widehat{\mathcal{P}}_t$ from the previous stage, we can construct a measure of productivity that does not depend

[40] For a more detailed discussion on this issue see Ackerberg, Caves, and Frazer (2015).

[41] The assumptions needed to ensure the invertibility of the policy functions associated with a wide class of production functions have been discussed extensively by Pakes (1991), Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg, Caves, and Frazer (2015).

on the measurement error ε_{ft} given by:

$$\omega_{ft}(\alpha,\nu) = \widehat{\mathcal{P}}_t(k_{T,ft}, k_{I,ft}, \ell_{ft}, \mathbf{d}_{ft}) - \alpha k_{T,ft} - \nu k_{I,ft} - (1-\alpha-\nu)\ell_{ft}. \tag{2.24}$$

Finally, taking advantage of the assumption that productivity follows an AR(1) process, it is possible to construct a measure of productivity innovations given by:

$$\xi(\alpha,\nu,\rho) = \omega_{ft}(\alpha,\nu) - \rho\omega_{ft-1}(\alpha,\nu). \tag{2.25}$$

Therefore, using the productivity innovations we can construct a set of moment conditions to estimated the parameters of the production function given by:

$$\mathbb{E}(\xi(\alpha,\nu,\rho) \times \mathbf{z}_{ft}) = \mathbf{0}_{Z\times 1}, \tag{2.26}$$

where Z ≥ 3 and, under the assumption that firms react to unanticipated productivity shocks contemporaneously and that capital is predetermined, the set of admissible instruments is
$\mathbf{z}_{ft} \in \{\ell_{ft}, k_{T,ft}, k_{I,ft}, \ell_{it-1}, k_{T,ft-1}, k_{I,ft-1}, \dots\}$.

Units. It is well known that most of the time standard production data, such as Compustat, record revenues and expenditures, rather than physical production and input used. In the presence of product differentiation (be it through physical attributes or location) an additional source of endogeneity presents itself through unobserved output and input prices.[42] This implies that, when bringing the model to the data, the structural value added production function takes the following form:

$$q_{ft}+p_{ft} = \alpha(k_{T,ft}+p_t^T)+\nu(k_{I,ft}+p_t^I)+(1-\alpha-\nu)(\ell_{ft}+p_{ft}^\ell)+\omega_{ft}+\varepsilon_{ft}, \tag{2.27}$$

[42]See De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) for a recent treatment of these issues.

where p_{ft} is the output price, p_t^T is the common user cost of tangible capital, p_t^I is the common user cost of intangible capital, and p_{ft}^{ℓ} is the price of labor. This empirical specification produces the following structural error term:

$$\omega_{ft} + p_{ft} - \alpha p_t^T - \nu p_t^I - (1 - \alpha - \nu) p_{ft}^{\ell}. \tag{2.28}$$

We follow De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) and let the wedge between the output and the input price (scaled by the output elasticity) be a function of the demand shifters and the productivity difference.[43] The inclusion in the control function of demand shifters $\mathbf{d}_{ft}$, constructed using measures of market shares as in De Loecker, Eeckhout, and Unger (2020), should therefore capture the relevant output and input market forces that generate differences in output and input price. As discussed in De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) this is an exact control when output prices, conditional on productivity, reflect input price variation, and when demand is of the (nested) logit form.

This is a second-best solution to address the aforementioned challenge in the estimation of the production function, however, without more detailed data on output quantities, it is not possible to go beyond this second-best solution to the problem.

2.A.2.2 Cost Shares

The cost shares approach has been prominently adopted in Foster, Haltiwanger, and Syverson (2008) and it exploits the first order conditions of the firm. To make fruitful use of the first order conditions of the firm two assumptions are needed, namely: (i) Constant returns to scale in production and (ii) that all inputs are variable. Under these assumptions the output elasticities can be calculated from cost shares. The cost

[43] De Loecker, Eeckhout, and Unger (2020) note that not observing output prices has the perhaps unexpected benefit that output price variation absorbs input price variation, thus eliminating part of the variation in the error term.

shares of both inputs are defined as:

$$\alpha = \text{med}\left\{\frac{r_t^T k_{T,ft}}{w_{ft}\ell_{ft} + r_t^T k_{T,ft} + r_t^I k_{I,ft}}\right\} \quad \text{and} \tag{2.29}$$

$$\nu = \text{med}\left\{\frac{r_t^I k_{I,ft}}{w_{ft}\ell_{ft} + r_t^T k_{T,ft} + r_t^I k_{I,ft}}\right\}, \tag{2.30}$$

where $w_{ft}\ell_{ft}$ is the wage bill, $r_t^T k_{T,ft}$ is the rental cost of tangible capital, and $r_t^I k_{I,ft}$ is the rental cost of intangible capital. Therefore, an extra requirement to apply this method is the possibility to calculate the return on both capitals, r_t^T and r_t^I.

2.A.3 *Robustness Production Function Estimation*

In this subsection of the appendix we explain the alternative specifications that we use to test the robustness of the IBTC. Results are presented in Figure 2.3.

2.A.3.1 Unconstrained Returns to Scale

To test the robustness of our results to a more flexible specification of returns to scale we estimate with the ACF approach the following production function:

$$q_{ft} = \alpha k_{T,ft} + \nu k_{I,ft} + \beta \ell_{ft} + \omega_{ft} + \varepsilon_{ft}, \tag{2.31}$$

where the only difference with equation 2.22 is that now returns to scale are unconstrained. Therefore, with this alternative specification, the set of moment conditions becomes:

$$\mathbb{E}(\xi(\alpha, \nu, \beta, \rho) \times \mathbf{z}_{ft}) = \mathbf{0}_{Z\times 1}, \tag{2.32}$$

where Z ≥ 4.

2.A.3.2 Sector-Level Production Technology

One restrictive assumption of our benchmark specification is that the production technology is the same across all sectors. We relax this assumption by allowing the production technology to be sector-specific. Effectively, this means that we estimate the following production function:

$$q_{ft} = \alpha_s k_{T,ft} + \nu_s k_{I,ft} + (1 - \alpha_s - \nu_s)\ell_{ft} + \omega_{ft} + \varepsilon_{ft}, \tag{2.33}$$

which is identical to the benchmark one except that now output elasticity are sector-specific. Finally, with this specification, the average output elasticities are going to be computed using a sales-weighted average.

2.A.3.3 Translog Production Function

We also test the robustness of our results to a more flexible production function: The Translog production. This production function approximates up to a second-order a CES production function. We choose a specification with constant returns to scale given by:

$$\begin{aligned} q_{ft} = {} & \alpha k_{T,ft} + \nu k_{I,ft} + (1 - \alpha - \nu)\ell_{ft} \\ & - \beta k_{T,ft} k_{I,ft} - \beta k_{T,ft}\ell_{ft} - \beta k_{I,ft}\ell_{ft} + \beta k_{T,ft}^2 + \beta k_{I,ft}^2 + \beta \ell_{ft}^2 + \omega_{ft} + \varepsilon_{ft}, \end{aligned} \tag{2.34}$$

Therefore, with this alternative specification, the set of moment conditions becomes:

$$\mathbb{E}(\xi(\alpha, \nu, \beta, \rho) \times \mathbf{z}_{ft}) = \mathbf{0}_{Z\times 1}, \tag{2.35}$$

where Z ≥ 4. Finally, the endogenous output elasticities are going to be given by:

$$\theta^T = \text{med}\big(\alpha - \beta k_{I,ft} - \beta \ell_{ft} + 2\beta k_{T,ft}\big), \tag{2.36}$$

$$\theta^I = \text{med}\big(\nu - \beta k_{T,ft} - \beta \ell_{ft} + 2\beta k_{I,ft}\big), \tag{2.37}$$

$$\theta^\ell = \text{med}\big(1 - \alpha - \nu - \beta k_{T,ft} - \beta k_{T,ft} + 2\beta \ell_{ft}\big). \tag{2.38}$$

Table 2.A.2: Lumpiness by Period

Investment rates	1980-1999	2000-2015
Average	0.36	0.34
Positive fraction, $i > 1$	0.87	0.90
Negative fraction, $i < -1$	0.02	0.03
Inaction rate	0.11	0.07
Spike rate, $\lvert i \rvert > 20$	0.76	0.75
Positive spikes, $i > 20$	0.75	0.74
Negative spikes, $i < -20$	0.01	0.01
Standard Deviation	0.29	0.30
Serial correlation, Corr(i_t, i_{t-1})	0.38	0.26

Note. This table shows the moments of the investment rate distribution of intangible and tangible capital. The statistics are computed for a balance panel of 3860 firm year observations between 1980 and 1999 and for a balance panel of 2992 firms year observations between 2000 and 2015.

2.A.4 *Robustness Lumpiness*

In this section of the appendix, we present some robustness analyses regarding the patterns of the investment rate distribution of intangible capital. In particular, we look at two additional dimensions that we neglected in the main analysis: The time dimension and the sector-level dimension.

Table 2.A.2 shows the moments of the investment rate distribution of intangible capital for different time frames: The period between 1980 and 1999 and the period between 2000 and 2015. The investment rate distribution of intangible capital does not show any qualitative difference over time. Overall, it seems that the most salient features of the distribution are stable over time, hence they are not an artifact of the fact that intangible capital is rising over time and could be subject

Table 2.A.3: Lumpiness by Sector

Investment rates	MIN	CON	MAN	TCU	WHO	RET	SRV
Average	0.22	0.15	0.38	0.35	0.14	0.16	0.30
Positive fraction, $i > 1$	0.65	0.79	0.91	0.93	0.69	0.81	0.87
Negative fraction, $i < -1$	0.06	0.14	0.02	0.04	0.14	0.09	0.06
Inaction rate	0.29	0.07	0.07	0.03	0.17	0.10	0.07
Spike rate, $\lvert i \rvert > 20$	0.49	0.50	0.81	0.67	0.43	0.40	0.67
Positive spikes, $i > 20$	0.48	0.38	0.80	0.64	0.38	0.35	0.64
Negative spikes, $i < -20$	0.01	0.12	0.01	0.03	0.05	0.05	0.03
Standard Deviation	0.30	0.30	0.29	0.33	0.29	0.29	0.32
Serial correlation, Corr(i_t, i_{t-1})	0.41	0.14	0.32	0.21	0.18	0.14	0.29

Note. This table shows the moments of the investment rate distribution of intangible across different sectors. The statistics are computed for a balance panel between 1980 and 1990. MIN is the Mining sector and has 209 firm year observations. CON is the Construction sector and has 99 firm year observations. MAN is the Manufacturing sector and has 4455 observations. TCU is the Transportation and Public Utilities sector and has 88 firm year observations. WHO is the Wholesale sector and has 176 firm year observations. RET is the Retail sector and has 176 firm year observation. SRV is Services sector and has 352 firm year observations.

to an initial adoption phase.

Table 2.A.3 shows the moments of the investment rate distribution of intangible capital for different sectors. Here the analysis is complicated by the fact that at the sector-level the construction of a balanced panel sacrifices a lot of observation, leaving us with relatively small samples. Nonetheless, most of the salient characteristics of the investment rate distribution of intangible capital seem to emerge anyway from this analysis. This suggests that the investment rate distribution of intangible capital exhibits a fractal behavior overall.

Concluding, the high level of the lumpiness of the investment rate distribution of intangible capital documented in the main analysis seems robust to different time frameworks and to different sectors. This suggests that this lumpiness has to come from intrinsic properties of the investment process in intangible capital.

2.A.5 *Aggregate Trends*

In this section, we present the evolution between 1980 and 2015 of the main trends of interest for the quantitative analysis. For the trends constructed with the Compustat data, we explain the measurement procedure, for the others instead, we just refer to main papers that document them. In particular, we look at: (i) The rise in concentration, (ii) the decline in the labor share, (iii) the rise in intangible capital investment share, (iv) the rise in tangible capital investment share, (v) the rise in the tangible capital investment share, (vi) the rise in the profit rate, (vii) the rise in the average firm size, (viii) the decline in the allocative efficiency of the economy, that is, the rise in the standard deviation of $TFPR$.

We measure concentration using the HHI index as in Grullon, Larkin, and Michaely (2019). In Compustat the HHI index of a sector s is constructed as:

$$HHI_{st} = \sum_{f} \left(\frac{\texttt{SALE}_{ft}}{\sum_f \texttt{SALE}_{ft}} \right). \tag{2.39}$$

Then, the aggregate concentration is simply the sales-weighted average of the sector-level concentrations.[44]

The firm-level profit rate, adjusted for intangible capital, is defined as:

$$\pi_{ft} = \frac{\texttt{SALE}_{ft} - \texttt{COGS}_{ft} - (\texttt{XSGA}_{ft} - \texttt{XRD}_{ft}) - r_{T,t}k_{T,ft} - r_{I,t}k_{I,ft}}{\texttt{SALE}_{ft}}, \tag{2.40}$$

the construction of the user cost of both capitals is described in Appendix 2.A.1.2, we drop `XRD` from `XSGA` to not double-count research and development costs as the are both part of our measured intangible capital and of selling general and administrative costs. The standard

[44]We follow Grullon, Larkin, and Michaely (2019) and use 2-digit NAICS level as the definition of sector-level.

un-adjusted profit rate is instead defined as:

$$\pi_{ft} = \frac{\texttt{SALE}_{ft} - \texttt{COGS}_{ft} - \texttt{XSGA}_{ft} - r_{T,t}k_{T,ft}}{\texttt{SALE}_{ft}}. \tag{2.41}$$

To obtain the aggregate profit rate we use a sales-weighted average of both measures of the firm-level profit rate.

Finally, to measure the allocative efficiency of the US economy we measure the standard deviation of $TFPR$. We compute $TFPR$ as:

$$TFPR_{ft} = \log \texttt{SALE}_{ft} - \alpha k_{T,ft} - \nu k_{I,ft} - (1-\alpha-\nu)\log \texttt{EMP}_{ft}, \tag{2.42}$$

where α and ν are the estimates from Section 2.3.1. Then our measure of allocative efficiency is just dispersion in $TFPR$ over the different years.

2.B QUANTITATIVE APPENDIX

2.B.1 *Additional Comparisons between Model and Data in 1980*

Here we compare the distribution of the average product of tangible capital, $ARPK_T$, and the distribution of the average product of intangible capital, $ARPK_I$, from the model with the ones form the data for the 1980s.

Figure 2.B.1 shows the distributions. The distributions implied from the model capture well the main features of the distributions in the data. In particular, the model is able to capture the excess dispersion in the distribution of $ARPK_I$ relative to the distribution of $ARPK_T$. This is due to the fact that indeed intangible capital faces higher distortions due to the presence of higher adjustment costs.

2.B.2 *Additional Comparisons between Model and Data over Time*

In this section, we document two additional implications of the model over time and compare them with the data. In particular, we look at the distribution of intangible intensity, defined as the ratio of intangible capital to labor bill, and at the distribution of $TFPR$.

Figure 2.B.2 shows the evolution over time of the distribution of intangible intensity both in the model and in the data. Overall, despite some qualitative differences, both the model and the data show a shift towards the right in the distribution of intangible intensities, highlighting the fact that firms are using on average more intangible capital relative to labor.

Figure 2.B.3a shows the evolution, both in the model and in the data, of the distribution of $TFPR$. Both in the model and in the data the distribution of $TFPR$ is more dispersed in 2015, highlighting a decline in allocative efficiency. This, as emphasized in the main text, is due to the fact that firms rely more on an input that is highly dispersed due to technological constraints and this hinders a fast reallocation of inputs towards high marginal product firms.

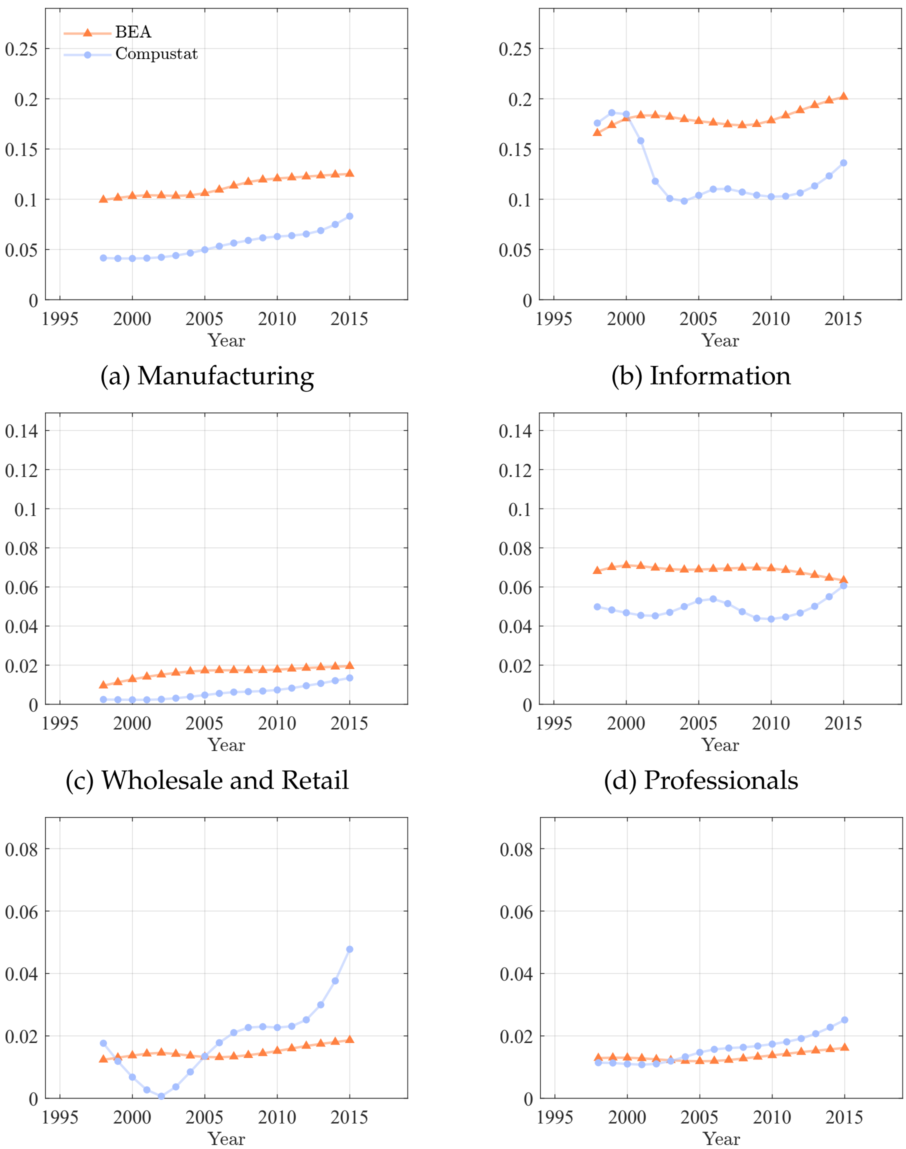
2.B.3 *Additional Robustness*

Table 2.B.1 shows the parameters from the robustness exercise where we re-estimate the adjustment costs as presented in Section 2.7.2. two things are changed relative to the calibration for the 1980 steady state. First, the intangible capital share has increased to 0.12 as we are estimating the economy in 2015. Second, now all the adjustment costa associated with the investment process of both capitals have been changed to match the moments from the later part of our sample. Remaining parameters not associated with the production technology or with the adjustment costs are left the same as in the 1980 steady state to facilitate the comparison across steady state and to pin down what are the fundamental forces underlying the results.

Table 2.B.1: Parameters and Moments

Fixed	Value	Description			
R	1.05	Annual interest rate			
δ_T	0.07	Annual depreciation rate tangible capital			
δ_I	0.29	Annual depreciation rate intangible capital			
α	0.28	Tangible capital share			
ν	0.12	Intangible capital share			
ω	0.90	Returns to scale			
ρ_z	0.90	Autocorrelation idiosyncratic productivity			
σ_z	0.20	Standard deviation idiosyncratic productivity			
c_e	3·e-4	Fixed to 1980 SS			
c_f	2.540	Fixed to 1980 SS			
η	2.025	Fixed to 1980 SS			
m	6.2·e-3	Fixed to 1980 SS			
Fitted	**Value**	**Description**	**Moments**	**Model**	**Data**
γ_T	0.012	Convex adj. cost k_T	corr($x_{T,ft}, x_{T,ft-1}$)	0.16	0.16
γ_I	0.060	Convex adj. cost k_I	corr($x_{I,ft}, x_{I,ft-1}$)	0.26	0.27
f_p	2.3·e-3	Fixed adj. cost k_T	Inaction rate: x_T	0.03	0.03
f_i	0.017	Fixed adj. cost k_I	Inaction rate: x_I	0.07	0.07

FIGURE 2.A.7: Intangible Capital Components Share



(a) Manufacturing

(b) Information

(c) Wholesale and Retail

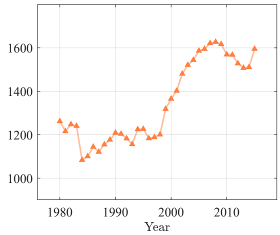
(d) Professionals

(e) Healthcare

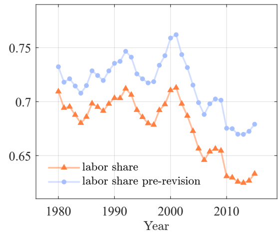
(f) Others

Note. The figures show the evolution of intangible capital investment share across different sectors of the US economy for both BEA−KLEMS data and for Compustat data between 1998−2015. The data are detrended with an Hpfilter with $\lambda = 6.25$.

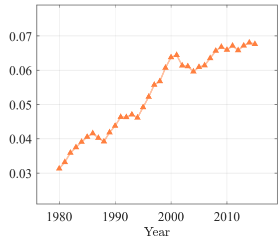
FIGURE 2.A.8: Aggregate Trends



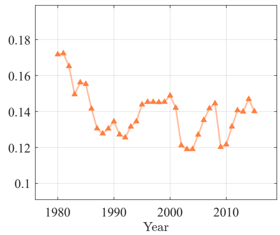
(a) Concentration (HHI)



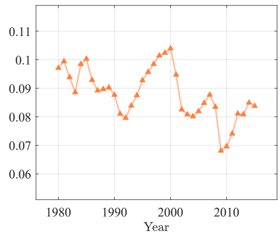
(b) Labor Share



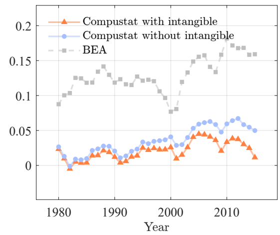
(c) Intangible Capital Share



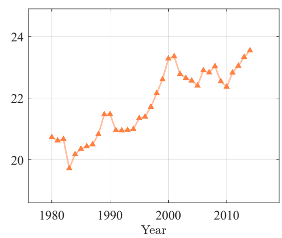
(d) Tangible Capital Share



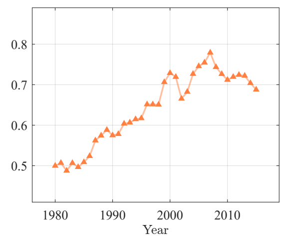
(e) Tangible Investment Rate



(f) Profit Rate



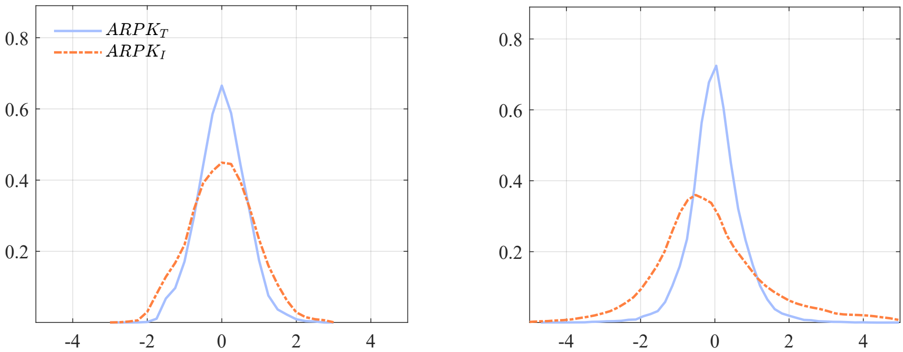
(g) Avg. Firm Size



(h) sd($TFPR$)

Note. Figure 2.A.8a replicate the evolution of the HHI index in Compustat as documented by Grullon, Larkin, and Michaely (2019). Figure 2.A.8b shows the evolution of the labor share, pre end post−revision, in the corporate non−financial sector as reported in Koh, Santaeulàlia-Llopis, and Zheng (2020). Figure 2.A.8c shows the evolution of the intangible capital investment share in the corporate non−financial sector as reported in Koh, Santaeulàlia-Llopis, and Zheng (2020). Figure 2.A.8d shows the evolution of the tangible capital investment share in the corporate non−financial sector as reported in Koh, Santaeulàlia-Llopis, and Zheng (2020). Figure 2.A.8e shows the evolution of the tangible capital investment rate as reported by Crouzet and Eberly (2019). Figure 2.A.8f shows the evolution of the profit rate as reported in De Loecker et al. (2020) and the profit rate adjusted for intangible capital. Figure 2.A.8g shows the evolution of the average firm size measured in number of employees from BDS data. Figure 2.A.8h shows the evolution of the standard deviation of $TFPR$ in Compustat.

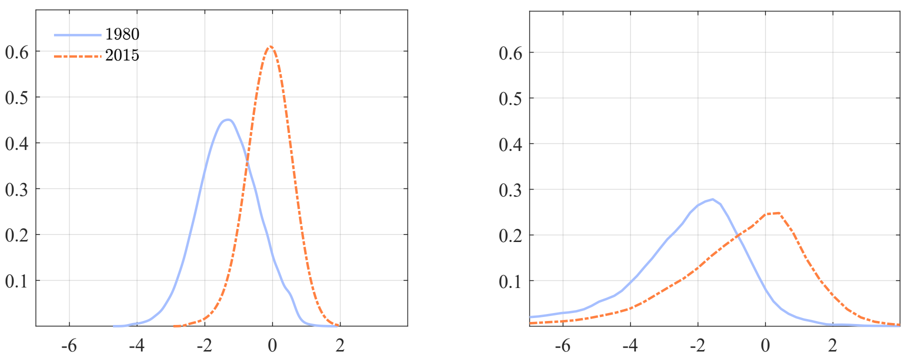
FIGURE 2.B.1: Average Product of Tangible and Intangible Capital



(a) Model (b) Data

Note. Figure 2.B.1a shows the distribution of $ARPK_T$ (solid light blue line) and of $ARPK_I$ (dashed orange line) from the model. Figure 2.B.3a shows the same distributions from the data. All distributions are demeaned.

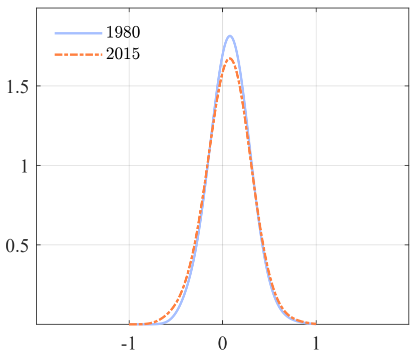
FIGURE 2.B.2: Log Intangible Intensity



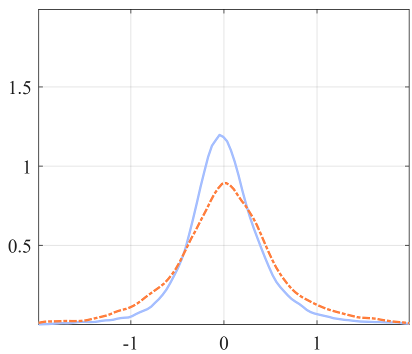
(a) Model (b) Data

Note. Figure 2.B.2a shows the distribution of log intangible intensity both in the 1980 (solid light blue line) and in 2015 (dashed orange line) from the model. Figure 2.B.2b shows the same distributions form the data.

Figure 2.B.3: Total Factor Productivity Revenue



(a) Model



(b) Data

Note. Figure 2.B.3a shows the distribution of $TFPR$ in 1980 (solid light blue line) and in 2015 (dashed orange line) from the model. Figure 2.B.3b shows the same distributions from the data. All distributions are demeaned.

3

PROCUREMENT, COMPETITION AND MARKET POWER

Joint with Jurica Zrnc

3.1 INTRODUCTION

Every year around 250,000 government authorities indulge into public procurement that represents 14% of the Eurozone's GDP.[1] Despite its large size, there is a growing evidence that only handful of firms take part in the procurement process which raises the concerns about low competition and high market power of firms in this sector.[2] Thriving competition among firms helps the government agencies to carry out their mission by procuring best-quality of goods and services at the

[1]See, for instance, European Commission (2017b).

[2]Kang and Miller (2017), for instance, document that 44 percent of the procurement budget was paid to contracts that attracted a single bid during fiscal year 2015 in United States. In EU co-funded projects, European Commission (2019) also documents a large share of single bidder contracts in European countries.

lowest price possible.[3] In the absence of competition, firms gain market power and command high prices. This has implications for public budget and for people's economic interests at large. In addition to lowering government surplus, market power in procurement sector may spillover to private sector and this may leads to a decline in the labor demand, business dynamics and innovation.

Despite the importance of procurement sector in the overall economy, there is lack of systematic evidence on the degree of competition in the procurement process, spanning all the sectors of the economy and over time.[4] In this article, our main goal is to document the market power in the procurement sector for the Croatian economy. First, we analyze markups, the ability of sellers to price their goods above marginal cost. This measure is of particular importance because it is informative about the underlying production technology and efficiency relative to other measures of market power such as Herfindahl-Hirschman Index. Based on firm-level data, we find that firms that are active in the procurement sector charge 9% percent higher markups than firms that only sell goods in the private sector. This markup premium for procurement firms increases with the share of sales coming for procurement sector (procurement intensity). Furthermore, we find strong positive association between markups, our proxies for prices in public procurement and the number of bids for a contract. We show that single bid procurement procedures consistently incur higher prices than expected by the procuring body relative to a multi bid setting, even for standardized products.

In principle, markups for procurement firms may be high because overhead costs or fixed costs of producing for government are high. These costs may stem from bureaucratic/administrative costs of bid-

[3] Competitive bidding can be a powerful tool to reduce procurement prices, Bulow and Klemperer (1996).

[4] There is a rich, both theoretical and empirical literature in Industrial organization that documents level of competition in some specific cases.

ding in the procurement process. In that case, the firm charges prices well above marginal costs to cover these excess costs. Therefore, markups alone can not be associated directly with market power. We thus also analyze measures of profitability (earning before taxes as share of total sales) that take into account not only the marginal cost but total costs for firms. We find that procurement firms tend to have much higher profits relative to non-procurement firms.

In general, measuring market power is hard. It is even harder in our setting as firms produce in two different sectors: private and procurement, where they face different demand functions. Moreover, both production decision and amount of quantity sold in procurement sector is driven non-market forces such as political connections, networks, and government policies which are hard to observe in a normal firm-level datasets. Therefore, in order to rigorously study markups for procurement firms, we need a detailed, large-scale micro dataset on firms, their procurement history, and their political affiliations.

To this end, we construct a new dataset for the Croatian economy, spanning the entire period of 2013-2019, wherein we merge: (i) firm-level balance sheet data; (ii) administrative data on procurement contracts above 25,000 euros; (iii) court registry data on firms and manager of firms; (iv) registry of politicians that contain information of ownership and employment histories of local and central government officials. The nature of this data allows us to exploit rich heterogeneity in the type of procurement contract across different sectors of the economy to understand the degree of competition and its implications for the markups.

We use the above mentioned dataset to measure markups at the firm-level. To do so, we follow recent advancements in the literature on markup estimation by De Loecker and Warzynski (2012) and De Loecker et al. (2020), and rely on individual firm output and input data. Together with the assumption of cost minimization at firm-level, a measure of the markup is obtained for each firm at a given point in time as the wedge between a variable input's expenditure share in revenue, i.e., directly observed in the data, and that output elasticity of variable input. The latter is obtained by estimating the associated production function. The advantage of this approach is that the production approach does not require to model demand for many heterogeneous markets, such as private and procurement, over a period of time.

The procurement process process is a complicated endeavour where rules and regulation poses many challenges for the economists to infer demand functions. Furthermore, each auction has different properties that ties ones' hands and limit the analyses to a specific set of auctions or a certain procurement agency. This is the reason that we adopt the methodology as in De Loecker and Warzynski (2012) because it helps to ignore the microcosm of the underlying allocation mechanism and measure market power for broad range of procurement activity in the economy. Needless to say that this convenience comes at a cost of assumptions that help us identify the markups using the production data. Furthermore, we only estimate one markup for each firm in a given year rather than a separate markups for private and procurement markets. However, even with these limitations, this methodology helps us to understand the conditions under which these firms sell goods and services to government.

We use our empirical model to estimate markups for Croatian firms jointly with the production technology, and test whether procurement firms, on average, have different markups relative to firms that only produce in the private sector.[5] We document a significant and positive markup premium for procurement firms in all the major sectors of the economy. For instance, the markup premium stands at 4% in manufacturing, 24% in construction sector and 6% in wholesale and retail sector.[6] We also find that large proportion of firms only have a minority

[5]We follow a two-stage estimation procedure following Ackerberg et al. (2015).

[6]Together these three sector, constitutes around 85% of the sales and procurement values within in our sample.

of their sales coming from procurement sector whereas there are few firms that have a majority of sales coming from procurement activities. The mean yearly procurement intensity, defined as the share of procurement sales over total sales is 32% and markup premium increases sharply with procurement intensity. Further, we analyze to what extent this markup premium is driven by higher prices or lower marginal cost. We document that the markups increase upon entry into procurement sectors even after controlling for a proxy for changes in marginal cost. This provide suggestive evidence that the markup premium is driven by higher prices.[7]

Furthermore, we exploit substantial entry of new firms into procurement in our data to disentangle the effect of entry into procurement from the incumbent procurement firms. We find that an entry of a new firm (that has not won a procurement contract before) into procurement is linked to a 10% percent increase in markup on average. Moreover, we find that the entry markup premium increases substantially with procurement intensity. For instance, a firm with 1% of sales coming from procurement sector only observe a 0.1% increase in markup, whereas a firm with 50% of sales coming from procurement sector witness an increase of 5%.

In what follows, we explore to what extent high markups in the procurement sector are associated with low competition. To do so, we exploit another dimension of our data that provides the number of bids for individual procurement auctions. Firms that enter into procurement with contracts for which there was only one bid (single bidder) charge even higher markups relative to multi-bid contracts. Furthermore, we corroborate our findings of positive procurement markup premium with the data on the estimated value of the procurement projects by the government and actual values at which the contract was awarded. The

[7]We use total factor productivity revenue $tfpr$ as a proxy for marginal cost. This is not a perfect proxy as it contain output variation but given the data constraints, this is the best available option.

estimate and actual contract value differential is 6% higher for single-bid relative to multi-bid setting, even after controlling for firm-year-8 digit product-county fixed effects. The estimates are also similar when we restrict the sample to standardized products such as printing paper following Bandiera et al. (2009).

We then analyze firm profitability. The objective is to analyze whether markups have not increased exclusively due to a rise in overhead/administrative costs that are associated with procurement process. To address this issue, we calculate the profit rate, which is total sales minus all cost (including overhead and the expenditure on capital) as a share of sales. We find that firms in procurement sector do have higher profits relative to firms that only operate in the private sector in all sectors of the economy. For instance, procurement firms in manufacturing sector have 3 percentage points higher profits, whereas in construction sector they have 6 percentage points higher profits relative to non-procurement firms. These magnitudes are large as the average profit rate is close to 1.5 percent (of sales). We also find in a staggered difference in differences framework that after winning procurement contracts firm profits increase.

We analyse firm-level labor share that has gained a lot of attention in the recent times.[8] In particular, we analyze how the rise in markups is associated with the labor share after firms enter procurement sector. We find a negative relation between procurement activity and labor share in all the sectors. We also find this result when comparing the dynamics of labor share after winning the procurement contract. This is in line with the theoretical literature, for instance, Atkeson and Burstein (2008), where high markups firms also have labor share. In conjunction, this also highlights the concerns for monopsony power and subsequent wage suppression in the labor markets.

Finally, we estimate dynamic effects of public procurement. We

[8]See, for instance, Autor, Dorn, Katz, Patterson, and Van Reenen (2020) and Kehrig and Vincent (2021).

use the staggered difference in differences research design to show how the evolution of our variables interest changed prior and after winning the procurement contracts. We find very similar results as in the cross-sectional analysis. Before winning the public procurement contract, markups of treated and not treated firms were statistically similar, while at the year of winning the procurement contract markups increase by 10%. In the following years there is no effect of procurement on markups, suggesting that price increases are driving our results and not shocks to productivity, as these are persistent (see e.g. Foster et al., 2008). We find that firms increase profits, employment and decrease labor share after public procurement. We also check the robustness of our results to heterogeneous treatment effects by using the estimator from De Chaisemartin and d'Haultfoeuille (2020a).

This paper is linked to a comprehensive literature in Industrial Organization on market structure in procurement sector. Laffont and Tirole (1987), Laffont and Tirole (1990), Laffont and Tirole (1993), Goldberg (1977), and Bajari and Tadelis (2001) propose a theoretical farmeworks of procurement contracts that emphasizes the choice of contract terms as a means for influencing the ex post performance of the underlying project. Subsequently, there are empirical papers studying the interplay between competitive mechanisms and contract outcomes (Spulber 1990, Bajari, McMillan, and Tadelis 2009, Warren 2014, and Decarolis 2014).

There are few papers that have directly studied empirically the relation between competition and rents in the procurement sector. For instance, Kang and Miller (2017) looks into ICT industry and highlight the role of costs in determining the number of bidders in government auctions, whereas Kroft, Luo, Mogstad, and Setzler (2020) focus on construction sector in US and estimates labor both labor and product market power together. Carril and Duggan (2020) studies the effects of large mergers on competition in US defense sector. Gugler, Weichselbaumer, and Zulehner (2015) documents the effects of crisis on competition in procurement in Austrian construction sector. Relative to all these papers, we estimated the market power for all the firms active in procurement sector and link that to other firm-level outcomes such as firm-level labor share.

This paper is also related to the literature that documents the effects of procurement on firm dynamics, for instance, Colonnelli and Prem (2017) highlights the negative impact of corruption in procurement on local entrepreneurial activity, di Giovanni, Garica-Santana, Jeenas, Moral-Benito, and Pijoan-Mas focus on the high collateralizability of cash-flows emerging from procurement sectors, and Ferraz, Finan, and Szerman (2015) and Gugler, Weichselbaumer, and Zulehner (2020) documents the growth effects of winning a procurement contract. Relative to all these papers, we focus on how procurement activity affects markups at the firm-level.

The organization of the paper is as follows. In the next section, we describe the construction of our main dataset and provide descriptive statistics. Section 3.3 documents the institutional setting, highlighting the main features of procurement in Croatia. In Section 3.4, we present structural framework to estimate markups and in Section 3.5 we discuss main results. In Section 3.6, we conclude.

3.2 DATA AND MEASUREMENT

We tap on data from multiple administrative sources to build a comprehensive dataset on firms, procurement contracts, workers, and local politicians in Croatia for the period of 2013-2019. The core of this data construction is the firm level financial statements that are collected by the Financial agency (FINA). We combine it with the administrative dataset, collected by the Ministry of Economy and Sustainable Development, on public procurement contracts allocated to firms. Further, we merge this dataset with court registry data on firm owners and managers and then link it with registry of politicians to construct our proxy for political networks. In this section, we give details about the multiple administrative sources used to build our dataset and further, provide additional definitions relevant to the analysis.

Firm level financial statements are collected by the Financial agency (FINA). All non-financial companies in Croatia are obliged to provide their financial reports to FINA, only sole proprietors are excluded. This data is used for statistical purposes by the Croatian Bureau of Statistics and Croatian National Bank. We exclude predominantly non-market sectors and sectors dominated by the government such as agriculture, public utilities, education, health and defense. We also remove the financial sector. After dropping these sectors, our dataset contains firm level data for more than 400 sectors in industry and services according to NACE Rev. 2. four digit classification. Similarly, we drop all firms with zero capital, labor and sales. After all of this data cleaning our sample capture captures 64% of the GVA. We have access to the data from 2002 to 2019.

Public procurement data is also an administrative dataset, collected by the Ministry of Economy and Sustainable Development. It contains all public procurement that was published in the Official Gazette from 2008 until 2019. In our analysis, we use the data starting from 2013 because it includes information on the unique firm identifier and the published data corresponds very closely to official statistics on total public procurement. After 2013 the threshold for required publishing of the tender in the Official Gazette is 200,000 Kuna (approx. 25,000 EUR) for goods and services, and 500,000 Kuna (approx. 70,000 EUR) for construction. Although very small contracts are not published and hence not in our data, we capture the majority of public procurement. For example in 2019, public procurement that was above the thresholds and published in the Official Gazette is 43 billion Kuna, while the below threshold and unpublished procurement was 11 billion Kuna (see Table 3.A.1). Overall procurement is around 13% of GDP. After cleaning the financial statements dataset and merging it with procurement data, we capture around half of total procurement $\approx 6\%$ of GDP. The data contains rich set of information about the final value of contracted procurement, estimated value of the tender by the procuring body, identity of the procuring body, location, number of bidders, type of contract, type of procedure, Common Procurement Vocabulary (CPV) code for the goods and services procured. In this data, we do not have information on the duration of the procurement contracts. We use a separate web-scrapped dataset acquired from GONG, a leading non-governmental organization in Croatia, and has information on the timeline for the delivery of goods and services for procurement in 2010 and 2011. The dataset shows that a large majority of goods and services are due in one year, but it also has many missing observations. Ferraz et al. (2015) also find that most of public procurement is due within a year.

The court registry is a dataset acquired from GONG. It contains all managers and owners of firms in Croatia from 2005 onwards. The registry of politicians also comes from GONG, and is web-scraped from web-pages of high ranking public officials that need to disclose their work and ownership histories. This includes parliament members, ministers, deputy ministers, mayors and deputy mayors. The dataset contains their ownership and employment histories. We use these datasets to build political connections indicators. First, we flag all firms that are owned or managed by former or current politicians. Second, we identify former managers of all procurement bodies (e.g. public utilities and regulatory agencies) and flag all firms that are owned or managed by them. Using these indicators we flag on average 3,721 unique firms as politically connected, out of 15,551 total unique firms in our sample (see Table 3.A.2).

In our analysis we define a procurement firm as a firm that has at least one procurement contract in a given year. At the individual procurement contract level a firm is a single bidder if it was the only bidder. At the firm-year level a firm is a single bidder if it has at least one single bidder contract in a given year. Approximately 60% of total

Table 3.1: Data Sources

Data Name	Data Source	Sample	Years	Variables & Content
1) Firm-level Data		Universe of firms	2002-2019	Balance Sheet Information and Income statement.
2) Procurement Data	Ministry of Economy	Administrative data on published public procurement contracts value $\geq$ 25k EUR in goods and 65k EUR for construction	2008-2019	Estimated value of the contract, Final amount paid for the contract (after completion), Winner of the contract and number of bids received.
3) Court registry				Court registry that has info on all managers and owners of firms and government entities.
4) Registry of Politicians				Ownership and employment histories of local and central government officials.
5) Procurement Bidders	Web-Scrapped	Food and Construction	2016-2019	Tenders

procurement contracts are awarded in procedures with only one bidder. This translates to $\approx 9,000$ firm-year single bidder observations out of $\approx 16,000$ firm-year procurement observations (see Table 3.A.3).

In Table 3.2, we present the descriptive of our merged dataset. In Panel A, we report the sectoral allocation of output and procurement contracts in the economy. Manufacturing, Constructions and Wholesale & Retail sector constitutes almost 85% of the total sales and procurement value in the economy. All of three sectors contain substantial number of procurement contract that attracted only one bidder. For instance, 40% of the firms in the manufacturing sector have atleast one single bidder contract, that points towards the low degree of competition in the procurement sector.

In Panel B of Table 3.2, we provide descriptive statistics on the firms. In our dataset, around 16% firms are involved in the procurement sector. The procurement firms are substantially larger than that of non-procurement firms. Furthermore, they also have revenue labor productivity, lower average revenue product of capital and higher profit rate.

Next, we document how firms characteristics varies over procurement intensity (procurement sales over total sales). First, we find that as firms size is non-monotonic in procurement intensity. For instance, firms with procurement intensity less than 0.1 is are much larger than that the firms that have majority of sales coming from procurement sector (procurement dependent firms). These procurement dependent firms gets much larger procurement contracts on average and they also make much higher profits relative to other procurement firms. All these findings hint towards a higher markups in the procurement sector, where procurement dependent firms charge even higher markups.

3.3 INSTITUTIONAL FRAMEWORK

All public procurement above the thresholds mentioned in Section 3.2 is subject to the Public Procurement Law, which is harmonized with European Union Public Procurement directives. According to the Law, public procurement needs to be done through the Electronic Public Procurement Notice (Elektronicki Oglasnik Javne Nabave). The majority of public procurement is done through an open procedure in which anybody can submit a bid - 85%, while in the rest the buyer invites selected sellers to submit a bid. Firms submit sealed bids until a known deadline. The winner is determined as the most economically advantageous bid, taking into account the price and other factors such as quality. Before the change in the Law in 2017, there was a bigger focus on prices, while after the Law changed the procuring bodies are instructed to take other factors such as quality and reliability in consideration. The procuring authority can assign how much weight it gives to each factor. In most cases, the lowest price is still a dominant decision factor. The government can sign various types of contracts with suppliers. Most of contracts are standard individual purchase contracts where the government pays for the delivery of previously specified goods or services. However, soliciting bids and selecting best offers for recurring needs of the public authority might entail large administrative costs. Framework agreements allow more flexibility for the procuring authority. They can be signed with multiple suppliers to a duration of up to 4 years and allow the procurer to choose among the pre-specified set of suppliers. While this procedure allows the procuring body to reduce administrative costs it decreases competition due to long term relationships between the procurer and a set of suppliers. Around 40% of procurement is done under framework agreements.

Although Croatia is fully compliant with the EU directives on public procurement, corruption is still a major concern. According to the Eurobarameter Survey of Business Attitudes in 2019, 77% respondents think that corruption is widespread in public procurement, which is well above the EU average (European Commission, 2020). The main concerns that discouraged participation in public procurement according to respondents were "The criteria seemed to be tailor-made for certain participants", "The deal seemed to be done before the call to tender" and "The procedure seemed too bureaucratic or burdensome". Corruption, however, is not the only concern about the procurement process. The European Commission expressed concerns about the limited administrative capacity of contracting authorities in Croatia (European Commission, 2017a). A report on single bidding in the EU (European Commission, 2019) suggests that these political and administrative issues result in the high share of single bidder contracts. The report finds that Croatia has the highest share of single bid contracts in analyzed countries, but it also uses a non-representative sample of public procurement.

Table 3.2: Summary Statistics on the Matched Firm-level Data: 2013-2019

PANEL A: Output and procurement allocation across sectors

Sector	Sales Share	Procurement Share	Single Bidder share of Procurement	Single Bidder Firms	Procurement Firms	Single Bidder Share (N)
Manufacturing	0.26	0.14	0.40	1688	2976	0.57
Construction	0.08	0.39	0.26	1665	3371	0.49
Wholesale & Retail	0.46	0.29	0.37	2903	4825	0.60
Transportation & Storage	0.04	0.02	0.54	322	544	0.59
Accommodation and Food	0.04	0.00	0.60	154	201	0.77
Information & Communication	0.04	0.06	0.58	957	1184	0.81
Real Estate	0.01	0.00	0.48	18	32	0.56
Professional, S & T	0.05	0.07	0.41	1571	3006	0.52
Administrative & Support	0.02	0.02	0.29	304	658	0.46

PANEL B: Firms characteristics and procurement intensity

$\frac{\text{Procurement}}{\text{Sale}}$	$\frac{\text{Procurement}}{\text{Total Proc.}}$	Labor	Obs.	$\frac{\text{Sale}}{\text{Employees}}$	$\frac{\text{Sale}}{\text{Capital}}$	$\frac{\text{Profits}}{\text{Sale}}$	$\frac{\text{Wagebill}}{\text{Sale}}$	$\frac{\text{Loans}}{\text{Capital}}$	Mean Proc.(10^3)
0	0.00	9	361572	5984	140	0.01	0.28	0.02	0
$< 10\%$	0.09	88	6902	13329	80	0.06	0.18	0.02	1766
$[10\%, 20\%)$	0.09	42	2663	8994	63	0.08	0.21	0.02	4819
$[20\%, 30\%)$	0.10	36	1737	8957	59	0.07	0.22	0.02	8048
$[30\%, 50\%)$	0.16	37	2133	9066	119	0.08	0.21	0.02	10521
$[50\%, 100\%)$	0.23	32	2106	8705	114	0.07	0.22	0.02	15035
$\geq 100\%$	0.32	31	1234	7385	107	0.06	0.24	0.02	35802

Note. Panel A presents the descriptive statistics at the sector-level. The variable Sales share is defined as the sector sales over total sales, Procurement share is sector procurement value over total procurement value, Single Bidder share of procurement is the value of single bidder contracts over total procurement, Single bidder firm is number of firm-year observation of firms with atleast one single bidder contract, Procurement firm is number of firm-year observation of firms active in procurement. Panel B presents firm-level statistics.

3.4 A STRUCTURAL FRAMEWORK TO ESTIMATE MARKUPS

In this section, we present the structural framework to estimate the firm-level markups. It relies on the production approach as used in De Loecker and Warzynski (2012) and De Loecker, Eeckhout, and Unger (2020). In particular, we rely on optimal input demand conditions obtained from cost minimization and the ability to identify the output elasticity of a variable input free of adjustment costs. This approach does not require us to make any assumption on the demand side or the market structure. Therefore, this is helpful in estimating the markups of firms that are active in procurement sector without taking a particular stand on the underlying contract allocation mechanism. However, we need to estimate a production function at the firm-level.

To obtain output elasticities, we need to estimate a firm-level production function and we rely on control function approach as developed by Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg, Caves, and Frazer (2015). To do so, we use material demand as a proxy for productivity and control both simultaneity and selection bias due to unobservable productivity shocks. This requires a two-stage estimation procedure that is explained in the next section.

3.4.1 *Deriving an Expression for Markups*

In order to derive markups, we need to define the cost minimization problem of the firms. In particular, we assume that the production function is defined as:

$$Y_{ft} = min[\gamma_m M_{ft}, F(L_{ft}, K_{ft})exp(\omega_{ft})]exp(\varepsilon_{ft}). \tag{3.1}$$

where Y_{ft} is output of firm f at time t, M_{ft} is material input, K_{ft} is capital stock, L_{ft} is labor input, ω_{ft} is productivity, and ε_{ft} is measurement error in output.

Assumption I: The production technology is sector-specific.

We assume that all the firms use the same technology within a sector, however, they can have different productivity ω_{ft} and that makes Y_{ft} to differ across firms. This implies that input shares are same for all firms within a sector.

Assumption II: Materials are a perfect complement to the combination of labor and capital.

This is a structural value added specification, the dependent variable is gross output. Here, labor and capital may be substitutable, but material are perfect complement to the combination of labor and capital. This is specification of the production function is not subject to the critique of Gandhi, Navarro, and Rivers (2020).

Assumption III: $F(.)$ is continuous and twice differentiable in labor and materials (variable input).

This assumption implies that the firms can adjust their output quanity by changing a particular variable input. Moreover, this implies that firms cost minimization involves two first-order conditions with respect to variable inputs of production.

Assumption IV: ω_{ft} is hicks neutral.

The productivity ω_{ft} affects both labor and capital in the same way such that capital-labor ratio would equalize across firms in the absence of capital adjustment costs (capital is a free input).

Assumption V: State Variables $S_{ft} = \{K_{ft}, \omega_{ft}, Proc_{ft}, G_{ft}\}$.

The state variables of the firms include the stock of capital K_{ft} (determined in period $t-1$), productivity ω_{ft}, firm active in procurement status $Proc_{ft}$ (decision to produce in procurement sector at time t is taken in period $t-1$), and the exogenous factors such as location.[9]

Assumption VIa: Firms minimize short-run costs taking input prices as given.

Firms face a vector of input prices and it does not depend on the quantity of input, ruling out any sort of monopsony power. We consider the firms' cost minimization problem conditioning on the state variables. From previous assumptions, firms minimize costs with respect to variable inputs (labor and material).

[9] $Proc_{ft}$ is a dummy variable that takes the value one if the firm is active.

The associated Lagrangian function for any firm f at time t is

$$\begin{aligned}\mathcal{L}_{ft}(L_{ft}, M_{ft}, K_{ft}, \lambda_{1,ft}, \lambda_{2,ft}) = {} & \mathcal{F} + rK_{ft} + wL_{ft} + p^m M_{ft} \\ & - \lambda_{1,ft}(F(L_{ft}, K_{ft}) exp(\omega_{ft} + \varepsilon_{ft}) - Y_{ft}) \\ & - \lambda_{2,ft}(\gamma_m M_{ft} exp(\varepsilon_{ft}) - Y_{ft}).\end{aligned} \tag{3.2}$$

The firm's first-order condition for variable inputs would pin down $\lambda_{1,ft}$ and $\lambda_{2,ft}$ and thus, the marginal cost of increasing output by one unit. In particular, the marginal cost is defined as:

$$MC_{ft} = \lambda_{1,ft} + \lambda_{2,ft} = \frac{wL_{ft}}{\beta_\ell Y_{ft}} + \frac{p^m M_{ft}}{Y_{ft}},$$

where, β_ℓ is the output elasticity of labor, wL_{ft} is the wage bill and $p^m M_{ft}$ is the cost of materials.

Assumption VIb: Prices are flexible. They are set annually.

Finally, we assume that the prices are flexible and set annually that helps us to pin down the markups at the firm-level. This means that we can write expression for markups as;

$$\mu_{ft} := \frac{P_{ft}}{MC_{ft}} = \beta_\ell \frac{P_{ft} Y_{ft}}{wL_{ft} + \beta_\ell \, p^m M_{ft}}.$$

This expression only requires us to estimate the output elasticity of labor and all others items required are directly observed in the data (namely sales, material cost and wagebill). In order to estimate the output elasticity of labor, we assume $F(L_{ft}, K_{ft})$ is Cobb-Douglas.[10] This implies that the structural value added in logs is;

$$y_{ft} = \beta_0 + \beta_k k_{ft} + \beta_l \ell_{ft} + \omega_{ft} + \varepsilon_{ft}. \tag{3.3}$$

[10] One can allow for translog production function and this does not change results substantially.

Here, the left-hand side is given directly in the data, and in the right-hand side the parameter vector β must be estimated. Estimating Equation 3.3 directly by OLS could potentially suffer from simultaneity bias (as unobserved productivity shocks in ω_{ft} are correlated with input choices), serial correlation bias (if the observed productivity ω_{ft} has correlated effects), and selection bias (if, over time, sample selection occurs among exiting low-productivity firms). To deal with these issues, we proceed using the control function approach as in Olley and Pakes (1996) that entails two-stage estimation procedure.

Assumption VIII: $m_{ft} = m_t(k_{ft}, \omega_{ft}, \mathbf{proc}_{ft}, \mathbf{ms}_{ft}, \mathbf{z}_{ft})$.

The vector $\mathbf{z}_{ft}$ includes wages, procurement agency dummies, etc. In this context, $\mathbf{z}_{ft}$ may also include political network proxies, and dummy for privatized firms. In particular, we will include proxy for political connections as it can change material demand via directly influencing input prices or indirectly by increasing the probability of winning a procurement contract. For instance, in appendix, we show in Table 3.A.2 that politically connected firms do on average have larger procurement contracts relative to relative unconnected firms. Further, we also show in Table 3.A.4 that connected firms are more likely to be involved in procurement sector (overall 4% of firms in our sample are involved with public procurement, however, when it comes to politically connected firm, this number increases to 13%).[11]

Assumption IX: Material demand increasing in ω_{ft}.

[11]These findings are in line with large literature that find positive link between political connection and procurement activity at firm-level, such as Cingano and Pinotti (2013), Gerardino, Litschig, and Pomeranz (2017), Decarolis, Fisman, Pinotti, and Vannutelli (2020a), Brogaard, Denes, and Duchin (2021), Baltrunaite, Giorgiantonio, Mocetti, and Orlando (2021) , Decarolis, Giuffrida, Iossa, Mollisi, and Spagnolo (2020b), and Decarolis (2014).

$$\omega_{ft} = h_t^{-1}(k_{ft}, m_{ft}, \mathbf{proc}_{ft}, \mathbf{ms}_{ft}, \mathbf{z}_{ft}), \tag{3.4}$$

where h_t is a non-parametric function (e.g. a polynomial) of variable inputs, capital, labour, possibly other fixed inputs, and time dummies.

Now, we focus on the selection bias due to non-random entry and exit of firm. If the current state variables indicate continuing in operation is not worthwhile, the firm closes down the plant. If this is not the case the firm chooses an optimal investment level (constrained to be non-negative). The solution to this control problem generates an exit rule and an investment demand function. If we define the indicator function χ_{ft} to be equal to zero if the firm exits, then the exit rule and the investment demand equation are written, respectively, as

$$\chi_{ft} = \left\{ \begin{array}{cc} 1 & \omega_{ft} \geq \overline{\omega}_{ft}(\mathbf{s}_{ft}) \\ 0 & \text{otherwise} \end{array} \right\}$$

Consider next the problem self-selection into production that may bias the production function estimation. Assuming, temporarily, that there are no variable factors, the conditional expectation of y, (conditional on current inputs, survival, and information available at $t-1$), includes the term $E[\omega_{ft-1}, S_{ft}, \chi_{t-1} = 1]$.

Recall that $\chi_{t-1} = 1$ if and only if $\omega_{ft-1} > \overline{\omega}_{ft-1}$. For instance, if the profit function is increasing in k_{ft}, the value function must be increasing and $\overline{\omega}_{ft}$ decreasing in k_{ft}. Firms with larger capital stocks can expect larger future returns for any given level of current productivity, and hence will continue in operation at lower co realizations. Furthermore, access to procurement sector may delay the exit of unproductive firms as they remain afloat with profits from public sector contracts. Hence, the self-selection generated by exit behavior implies that $E[\omega_{ft-1}, S_{ft}, \chi_{t-1} = 1]$ will be decreasing in k_{ft}, leading to negative basis in the capital coefficients.

The survival probabilities can written as;

$$\begin{aligned} Pr(\chi_{ft} = 1) &= Pr[\omega_{ft} \geq \overline{\omega}_{ft} | \overline{\omega}_{ft}(s_{ft}), \omega_{ft-1}] \\ &= \kappa_{t-1}(\overline{\omega}_{ft}(s_{ft}), \omega_{ft-1}) \\ &= P_{ft} \end{aligned} \tag{3.5}$$

To control for the impact of the unobservable productivity on selection, we need a measure of ω_{ft} and a measure of the value of $\overline{\omega}_{ft}(s_{ft})$ which makes the firm just indifferent between continuing in operation and selling off. As in Olley and Pakes (1996), we have two different indexes of firm heterogeneity, the productivity and the productivity cutoff point. Note that, using equation 3.5, we can write $\overline{\omega}_{ft}(s_{ft}) = \kappa_{t-1}^{-1}(\omega_{ft}, P_{ft})$.

To estimate the parameter vector β, we follow Ackerberg, Caves, and Frazer (2015) and form moments based on the innovation in the productivity ξ_{ft}. We consider the following law of motion for productivity,

Assumption X: ω_{ft} follows Markov process.

$$\omega_{ft} = g(\omega_{ft-1}, P_{ft}) + \xi_{ft}. \tag{3.6}$$

To form moment based on the innovation in the productivity shock in Equation 3.6, we need to express productivity in forms of data and parameters. In particular, we follow a two-stage procedure to estimate the production function. In the first-stage, we run the following regression,

$$y_{ft} = \phi_t(k_{ft}, \ell_{ft}, m_{ft}, proc_{ft}, z_{ft}, G_{ft}) + \varepsilon_{ft},$$

where we estimate $\widehat{\phi_t}$. Next, we obtain the estimates for productivity $\omega_{ft}(\beta) = \widehat{\phi_t} - \beta_k k_{ft} - \beta_\ell \ell_{ft}$. By regressing $\omega_{ft}(\beta)$ on a function of $\omega_{ft-1}(\beta)$ and P_{ft} (nonparametrically), we recover the innovation to productivity giver β, $\xi(\beta) = \omega_{ft}(\beta) - g(\omega_{ft-1}(\beta), P_{ft})$.[12]

[12]In principle, one may also include other state variables that may affect productivity in period t.

In the second-stage, to obtain our estimated parameters $\widehat{\beta}$, we use the following moment conditions,

$$E\left[\xi_{ft}(\beta)\begin{pmatrix} k_{ft} \\ \ell_{ft-1} \\ m_{ft-1} \\ \phi_{ft-1} \end{pmatrix}\right] = 0 \tag{3.7}$$

and we use standard GMM technique to obtain the estimate of the production function and rely on the bootstrapping for the standard errors.

3.5 MARKUPS AND PROCUREMENT

We rely on our empirical framework to analyze how markups, size, labor share and profits differ between procurement firms and firms without procurement contracts. Moreover, we are interested in how the procurement intensity (share of procurement in total sales) is correlated with our main outcome variables. Further, we are also interested in how entry into procurement impact markups.

There is rich literature in IO on public procurement auctions and how different auctions designs changes firm-level surplus. A recent paper by Kang and Miller (2017) shows that more competition in the procurement can increase government surplus (reducing profits for firms). Other papers that are more directly linked to the firm dynamics such as Ferraz, Finan, and Szerman (2015) document an increase in firm size when it gets an procurement contract and di Giovanni, Garica-Santana, Jeenas, Moral-Benito, and Pijoan-Mas argue that the procurement revenues are more collateralizable and therefore, relaxes borrowing constraints for the firms. However, there is no systematic evidence on the the impact of procurement on the firm-level markups.

Theoretically, the firm-level markups are driven by physical productivity in model of oligopolistic competition such as in Atkeson and Burstein (2008). More productive firms are larger in size and also charge a higher markup; supply-side sources of markup of heterogeneity. Meanwhile, the demand-side factors (profitability) is also important for firm growth and size (Foster, Haltiwanger, and Syverson 2008) and thus, they can also explain markup heterogeneity in the cross-section.

Access to procurement sector acts as a demand shifter for firms and therefore, it may increase firm size and thus, can increase their markup as well. From the supply-side, the firms in the procurement sector may have higher physical productivity. On the one hand, entry into the procurement process is costly as firms may face bureaucratic costs (Krasnokutskaya and Seim (2011) discuss the participation cost in public auctions in detail), which would imply a entry of productive firms in the sector. On the other hand, firms' networks and connections with bureaucrats and politicians may allow unproductive firms to enter, which would imply a lower productivity of procurement firms.[13] Finally, procurement agencies may demand higher quality goods that allows firms to charge higher markups.[14] Given the fact, there are multitude of counter-balancing forces involved, it is not ex-ante clear how markups change as firms enter procurement sector. In this regard, we see this paper as providing a holistic evidence on procurement affects firm characteristics such as, size, markups, profits and labor share.

Taking stock of the above, we therefore expect higher markups for procurement firms. As we discussed before, markups differences are related to both cost and price side factors, our methodology, to a certain extent, will be able to disentangle these two forces. Using the method-

[13]Colonnelli and Prem (2017) shows that corruption allows less productive firms to enter and produce into procurement.

[14]The intuition is line with the literature on firm-level exporting, quality and markups. If exporters produce higher-quality goods, while relying on higher-quality inputs, all things equal, they can charge higher markups. See Kugler and Verhoogen (2012) and Hallak and Sivadasan (2013).

ology presented in the previous section, we can measure markups and revenue productivity at the firm-level. We use revenue productivity $tfpr$ as a crude measure of the physical productivity tfp and disentangle demand and supply side forces. Needless to say that $tfpr$ is an imperfect measure and contain unobserved price heterogeneity. In the future, we will combine product price data to provide more precise evidence on the different role of supply and demand side forces.

3.5.1 *Results*

In this section, we use our empirical model to estimate markups for Croatian firms, and test whether procurement firms, on average, have different markups relative to firms that only produce in the private sector. Furthermore, we exploit substantial entry of new firms into procurement in our data to disentangle the effect of entry into procurement from the incumbent procurement firms.

We plot the distribution of the estimated markups in Figure 3.1. Similar to the findings in the literature, we document a fat tailed distribution of markups, where a lot of mass is concentrated around one and some firms have high markups above 2. Further, we find that the markup distribution of procurement firms is systematically skewed towards right, suggesting a higher markups for them relative to others.

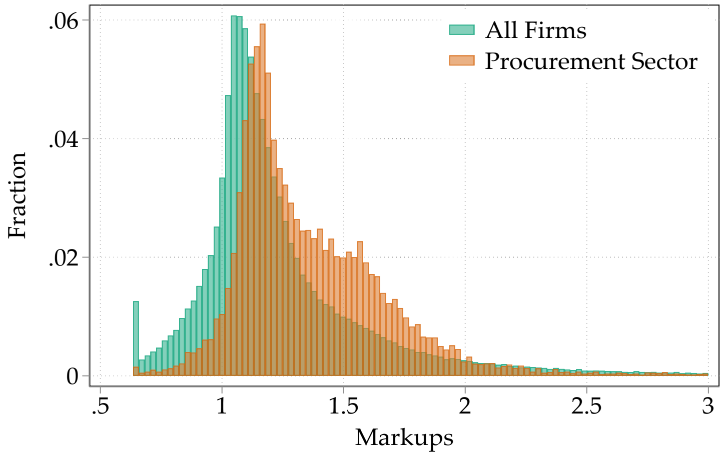
3.5.1.1 Markup Premium

More specifically, we run the following regression for each 2 digit industry separately,

$$ln\,\mu_{ft} = \delta_0 + \delta_1\,Procurement_{ft} + b'_{ft}\sigma + \nu_{ft}, \tag{3.8}$$

where $Procurement_{ft}$ is a dummy that indicates if a firms f is active in the procurement sector in a given year t and δ_1 measures the percentage markup premium for procurement firms. We control for labor and capital use in order to capture the differences in size and

FIGURE 3.1: Markups



Note. The distribution of estimated markups between 2013-2019. The x-axis is markups and y-axis is the fraction of firms for a certain bin of the histogram. We winsorize the variable at 1% and 99% to remove outliers. Green represents the distribution for all firms in the economy whereas red represents the firms active in the procurement sector. The mean and median for non-procurement firms are 1.27 and 1.14 respectively, and in case of procurement firms they stand at 1.39 and 1.29 respectively.

factor intensity, as we as full year-industry (four digit level) interaction to take out industry specific trends in markups. Needless to say that, we are not interpreting δ_1 as the causal impact of procurement on markups. We collect all the controls in a vector b_{ft} with σ the corresponding coefficients.

The results are presented in Table 3.3. We find that firms that are active in procurement have 6% higher markup relative to those firms that are not active in procurement in the Manufacturing sector. Similarly, markup premium for procurement firms is 25% in the Construction sector and 4% in the Retail & Wholesale sector. These sectors constitute almost 85% of the revenues and procurement value in the economy. Apart from them, we also find that markup premium for procurement firm is 14% in Professional Services & Technical Services sector. In the rest of the sectors, we do not find any statistically significant difference in the markups across procurement and non-procurement firms.

Table 3.3: Markups and Procurement I: Cross-Sectional Analysis

	(1) Manufacturing	(2) Construction	(3) Wholesale & Retail	(4) Transportation & Storage	(5) Accomm. & Food	(6) Information & Comm.	(7) Financial & Insurance	(8) Professional, S & T Services	(9) Admin. & Support
PANEL A. Baseline Specification in Equation 3.8									
Procurement	0.06***	0.25***	0.04***	0.01	0.03	-0.03	0.06	0.14***	-0.03
	(0.01)	(0.02)	(0.01)	(0.02)	(0.04)	(0.03)	(0.08)	(0.01)	(0.02)
PANEL B. Alternative Specification in Equation 3.9									
Procurement	0.02***	0.15***	0.01***	-0.06**	0.01	0.03	0.10	0.19***	-0.05***
	(0.01)	(0.02)	(0.00)	(0.02)	(0.03)	(0.02)	(0.09)	(0.01)	(0.02)
Observations	58,139	43,092	102,582	18,704	28,840	17,842	6,514	60,845	14,229
Sector#year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note. This table presents the estimated coefficients from regression specification Equation 3.8 in Panel A and Equation 3.9 in Panel B. We always use capital and lagged labor as controls. Manufacturing, Construction and Wholesale & Retail capture 85% of the sales and procurement activity in the economy. Furthermore, Professional, Scientific & Technical services sector capture another 7% of the procurement contracts in value. The regression standard errors in parentheses below the coefficient values are clustered at 4 digit NACE rev.2 sector level $\times$ year, where significance level is defined as *** p$<$0.01, ** p$<$0.05, * p$<$0.1.

In many models of imperfect competition (Atkeson and Burstein 2008), firms with same productivity will charge the same markups, making productivity differences the only source of markup differences. However, there could be other forces apart from technical efficiency that may generate markups dispersion across procurement and non-procurement firms. Our procedure generates estimates for both markups and revenue productivity and, to a certain extent, we can disentangle these forces by including both. Before moving to our next specification, we stress the fact that our measure of productivity is revenue based ω_{ft} that potentially picks up price differences and therefore, we expect is to absorb additional variation, that is not solely to linked to technical efficiency of production, in markups as well. To this end, we run the following specification

$$ln\, P_{ft} - ln\, MC_{ft} = \delta_0 + \delta_1\, Procurement_{ft} + \delta_2\, \omega_{ft} + b'_{ft}\sigma + \nu_{ft}, \quad (3.9)$$

which shows that δ_1 will measure the average price difference in percentages if ω_{ft} capture $ln\, MC_{ft}$ completely. The results are presented in Panel B of Table 3.3. We find that procurement dummy still explain a substantial difference in markups in the cross-section even after controlling for productivity. This implies that the factor that are not related to productivity, such as demand-side forces play an important role in explaining markup differences between procurement and non-procurement firms. In order to dip deeper into the exact mechanism, we would require data on product-level prices and this is left for future research.

These results are important in order to interpret the differences between procurement and non-procurement firms. As documented above, supply-side factors, such as more productive firms enter procurement as it entails fixed costs of entry, can not be whole story of markup differences. Moreover, revenue based measures of productivity to differentiate procurement firms from others may bias our inferences that revenue productivity contains prices and thus markups as well. Our results, however, cautions a more careful interpretation of procurement markup premium as truly a productivity premium.

3.5.1.2 Procurement at the Intensive Margin

So far, we have documented the procurement markup premium at the extensive margin, however, there is a lot of heterogeneity among the procurement firms themselves. As document in Section 3.2, there is rich heterogeneity in the size of procurement contract that a firm gets relative to their overall size. A large proportions of firms only have a minority of their sales coming from procurement sector whereas there are few firms that have a majority of sales coming from procurement activities. The mean yearly procurement intensity, defined as the share of procurement sales over total sales, is 32%.[15] We exploit this heterogeneity to estimate the procurement markup premium at the intensive margin. To see this, we run the following regression

$$ln\,\mu_{ft} = \delta_0 + \delta_1 \left(\frac{Procurement}{Sale}\right)_{ft} + b'_{ft}\sigma + \nu_{ft}, \tag{3.10}$$

where, δ_1 will measure the markup premium at the intensive margin. The results are presented in Table 3.4. At the extensive margin, procurement firms charge a 9% markup premium in the cross-section. Meanwhile, at the intensive margin, we find that higher procurement intensity also implies higher markup premium in the cross-section. The estimates δ_1 is 0.13 and it is statistically significant. An average firm has a procurement intensity of 32% that leads to a markup premium of 4%. However, it may be the case that the firms that are active in the procurement sector are inherently different from the non-procurement firms and therefore, the markup heterogeneity is a results of some unobserved heterogeneity. To this end, we provide estimates from the same regression including firm fixed-effects. In this case, we find that δ_1 declines from 0.13 in the baseline case (sector$\times$year fixed-effects) to 0.09 but still substantially and statistically significant.

3.5.1.3 Markups at Entry into Procurement

So far, we have just estimated difference in average markups for procurement and non-procurement producers. For procurement firms, however, we rely on a markup across the procurement and private market. In principle, our methodology can generate markup by markets. Applying first-order condition of labor and material by market $\mathcal{M}$, where $\mathcal{M} = \{procurement, private\}$, we can compute markup as

[15] This measure potentially suffer from the bias that is generated due to the fact that some procurement contract are multi-year and we do not observe this directly.

Table 3.4: Markups and Procurement II: Intensive Margin

VARIABLES	(1) μ_{ft}	(2) μ_{ft}	(3) μ_{ft}	(4) μ_{ft}
$Procurement_{ft}$	0.09***		0.11***	
	(0.01)		(0.01)	
$(\frac{Procurement}{Sale})_{ft}$		0.13***		0.09***
		(0.02)		(0.01)
Observations	309,948	309,948	293,115	293,115
R-squared	0.59	0.58	0.85	0.85
Sector#year FE	✓	✓	✓	✓
Firm FE	-	-	✓	✓

Note. This table presents the estimated coefficients from regression specification Equation 3.10. Dependent variable is log markups. $Procurement_{ft}$ is dummy variable. Clustered standard errors at sector×years are in brackets below the coefficient values. The average procurement over sales value is 32%. The regression standard errors in parentheses clustered at 4 digit NACE rev.2 sector level × year, where significance level is defined as *** p<0.01, ** p<0.05, * p<0.1.

before. However, in our data, we do not observe wagebill and material used for each market. We observe only total wagebill and material cost. Therefore, We estimate a weighted average in two markets: $\mu_{ft} = \rho_{ft}^{Priv}\mu_{ft}^{priv} + \rho_{ft}^{proc}\mu_{ft}^{proc}$. ρ_{ft}^{Proc} is the cost share in private sector (unknown to us), where, markups in the procurement sector is given by

$$\mu_{ft}^{Proc} = \beta_\ell \frac{[P_{ft}Y_{ft}]^{Proc}}{\rho_{\ell,ft}^{Proc}[wL_{ft}] + \beta_\ell\ \rho_{m,ft}^{Proc}\,[p^m M_{ft}]},$$

where $\rho_{m,ft}^{Proc}$ measures the share of material cost and $\rho_{\ell,ft}^{Proc}$ measures the share of wagebill used in production sold in procurement market.

We adopt the following strategy to verify whether the private sector markup of firms changes in the period when they enter the procurement market. First, we define the a dummy variable $Entry_{ft}$ for firms that are new entrants into the procurement sector. Second, we interact $Entry_{ft}$ with the procurement intensity. This specification allows us to inspect whether the increase in the firm's average markups (across procurement and private sectors) due to procurement entry depends on the intensity of procurement. We can look at firms with a very small fraction if sales coming from procurement, say less than 1 percent, when they enter the procurement market, which can be informative about what happens to their private sector markup. In particular, we run the specification

$$ln\,\mu_{ft} = \delta_0 + \delta_1\, Entry_{ft} + \delta_2\, Incumbents_{ft} + b'_{ft}\sigma + \nu_{ft}, \qquad (3.11)$$

where $Entry_{ft}$ is the dummy variable for entrants in the procurement sector. The coefficient of interest here is δ_1 that captures the average effect of entry. The results are presented in Table 3.5. We find that the entrants have a markup premium of 10%. Further, when we interact entry dummy with procurement intensity, we estimate δ_1 to be 0.15. To get the total effect of export entry, however, we need to multiply this estimate with the procurement intensity, and this implies that the markup entry effect is very small for firms selling a small share of their production to procurement. For procurement firms selling less than 1 percent to government, markups only increase by 0.002 percent, suggesting that that markup in the private market did not change. This approach to make an inference is not without problems as the procurement share increases over time and the separation between markups in procurement and private markets. In addition, this approach does not necessarily use the optimal weights, which will depend on how we aggregate inputs across production by markets within a firm. The procurement sales weight assumes implicitly that inputs are used in proportion to final sales (this assumption is made in Foster, Haltiwanger, and Syverson 2008).

Finally, in order to stress the decomposition of markups across two markets, we can write the change in a markup before and after entry into procurement share as $\Delta\mu_{ft} = \Delta\,\rho_{ft}^{Priv}\mu_{ft}^{priv} + \rho_{ft}^{proc}\mu_{ft}^{proc}$. Using this decomposition, our results suggest that for firms with very small procurement sales, markups do not change, suggesting that the private market markups are unaffected, under the assumption that the input cost share ρ_{ft}^{proc} will be small as well.

Table 3.5: Markups and Procurement III: Entry Margin

Var.	(1) μ_{ft}	(2) μ_{ft}
$Entry$	0.10***	
	(0.01)	
$Incumbents$	0.08***	
	(0.01)	
$\left(\frac{Procurement}{Sale}\right)_{ft} \times Entry$		0.15***
		(0.01)
$\left(\frac{Procurement}{Sale}\right)_{ft} \times Incumbents$		0.11***
		(0.02)
Observations	308,970	308,970
R-squared	0.59	0.58
Sector#year FE	YES	YES

Note. This table presents the estimated coefficients from regression specification Equation 3.10. Dependent variable is log markups. $Procurement_{ft}$ is dummy variable. Clustered standard errors at sector×years are in brackets below the coefficient values. The average procurement over sales value is 32%. The regression standard errors in parentheses clustered at 4 digit NACE rev.2 sector level × year, where significance level is defined as *** p<0.01, ** p<0.05, * p<0.1.

3.5.2 *Procurement, Profits and Labor share*

In this section, we provide further evidence on firm-level effects of procurement activities. Previously, we have documented that the procurement firms do charge a substantial markup premium. In a simple framework of imperfect competition, markups may be associated with higher profits rate (total profits over sales) and lower labor share.

For instance, Atkeson and Burstein (2008) show that larger firms (with higher market share) face lower demand elasticity and therefore, can charge a higher price-cost markups. However, higher markups does not necessarily imply that firms have more market power and therefore higher economic profits. In fact, increasing markups in procurement sector can come from a variety of reasons that are not associated with a decline in competition, such as ex-ante high entry costs, fixed bureaucratic costs, bribery costs, product customization costs, among others. I order to cover these costs, firms may charge a higher markups in the procurement sector. However, if firms do make excess profits in the procurement sector then that would be a stronger evidence in the support of the hypothesis that procurement sector is crippled with low competition and firms have market power.

In what follows, we measure firm's profit with EBT (earnings before taxes). Our measure of profitability is EBT over sales.[16] Further, we run the same specification as in 3.8, however, in this case the dependent variable in profit rate. The results are presented in Panel A of Table 3.6. We find that firms in procurement sector do have higher profits relative to firms that only operate in the private sector in all sectors of the economy. For instance, procurement firms in manufacturing sector have 3 percentage points higher profits, whereas in construction sector they have 6 percentage points higher profits relative to non-procurement firms. These magnitudes are large as the average profit rate is close to 1.5 percent (of sales). These results highlights the fact the high markups in the procurement sector reflects lack of competition and not only high administrative costs.

Next, we move towards analysing labor share of procurement firms. In simple framework of imperfect competition, labor share is inversely proportional to markups and profits. In order to build intuition of this result, let us suppose production function with only one factor, labor. If firms are making more profits then the share of production that goes to labors declines.[17]

[16]We also compute accounting profits as $Sales - rK - WL - p^m M - Overhead\, Costs$, and results do not change by a lot.

[17]Recent work by Autor, Dorn, Katz, Patterson, and Van Reenen (2020) summarize these results. A firm will have a lower labor share if its markup is higher. Moreover, larger firms that have lower demand elasticity would have higher price-cost markups and lower labor share. The reason is because markups are generally falling in the absolute value of the elasticity of demand, and according to Marshall's "Second Law of Demand," consumers will be more price-inelastic at higher levels of consumption and lower levels of price.

We define labor share at firm-level as the ratio of wagebill and sales. Further, we run the same specification as in 3.8, however, in this case the dependent variable in labor share. The results are presented in Panel B of Table 3.6. We find that firms in procurement sector do have lower labor share relative to firms that only operate in the private sector in all sectors of the economy. For instance, procurement firms in manufacturing sector have 8 percentage points higher profits, whereas in construction sector they have 13 percentage points lower labor share relative to non-procurement firms. These results highlights that procurement affects firms in multiple ways apart from the standard size/growth effects that has been discussed in the literature.

Table 3.6: Profits and Procurement I: Cross-Sectional Analysis

	(1) Manufacturing	(2) Construction	(3) Wholesale & Retail	(4) Transportation & Storage	(5) Accommodation & Food	(6) Information & Comm.	(7) Financial & Insurance	(8) Real Estate	(9) Admin. & Support
				PANEL B. Profit Share					
Procurement	0.03***	0.06***	0.05***	0.06***	0.07***	0.03***	0.03**	0.06***	0.04***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.02)	(0.01)	(0.01)	(0.00)	(0.01)
Observations	50,936	37,512	88,341	16,517	25,993	15,876	5,541	53,468	12,645
R-squared	0.05	0.02	0.02	0.03	0.02	0.05	0.02	0.04	0.03
				PANEL B. Labor Share					
Procurement	-0.08***	-0.13***	-0.10***	-0.05***	-0.03***	-0.04***	-0.08***	-0.07***	0.01
	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.02)	(0.00)	(0.02)
Observations	51,390	38,086	89,133	16,636	26,234	16,018	5,807	54,039	12,774
R-squared	0.11	0.03	0.06	0.08	0.01	0.06	0.04	0.04	0.16
Sector#year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note. This table presents the estimated coefficients from regression specification 3.8 in Panel A and 3.9 in Panel B. We always use capital and lagged labor as controls. Clustered standard errors at sector×years are in brackets below the coefficient values. Manufacturing, Construction and Wholesale & Retail capture 85% of the sales and procurement activity in the economy. Furthermore, Professional, Scientific & Technical services sector capture another 7% of the procurement contracts in value. The regression standard errors in parentheses below the coefficient values are clustered at 4 digit NACE rev.2 sector level, where significance level, where significance level is defined as *** p<0.01, ** p<0.05, * p<0.1.

3.5.3 *Dynamics Effects of Procurement*

In order to highlight the dynamic effects of public procurement on firms, we use the staggered difference-in-difference design. In particular, our regression model is;

$$y_{it} = \alpha + \sum_{l \in \Lambda} \beta_{t+l} \text{Procurement}_{it+l} + \gamma X_{it-1} \phi_i + \delta_{st} + \varepsilon_{it}, \qquad (3.12)$$

where ϕ_i is the individual fixed effect, and δ_{st} are year and sector FE. $\Lambda = \{-2, 0, 1, 2\}$ are coefficients of interest that estimate the evolution of the outcome relative to the year prior to getting the contract. X_{it-1} is the vector of pre-treatment controls. The identification assumption is that that conditional on controlling for time invariant heterogeneity and time variant characteristics in the vector X_{it-1} there are no unobservables driving both procurement and our outcomes of interest. Firm TFP is one such unobservable, given that a positive TFP shock can both drive procurement and markups. We will discuss the plausibility of TFP shocks explaining our results.

Additionally, a recent literature on staggered difference in differences shows that estimating Equation 3.12 with OLS in the presence of heterogeneous treatment effects for different treatment cohorts may lead to biased estimates.[18] For instance, the bias appears when units treated later have lower treatment effects than units treated earlier. While it is not obvious why this would happen in the context of procurement, for robustness we use the estimator from De Chaisemartin and d'Haultfoeuille (2020b) which provides unbiased estimates in this setting. In particular, they propose comparing units that switch from untreated to treated from $t-1$ to t with never treated units at both dates and comparing units that switch from treated to untreated from $t-1$ to t with units treated at both dates. This approach allows units

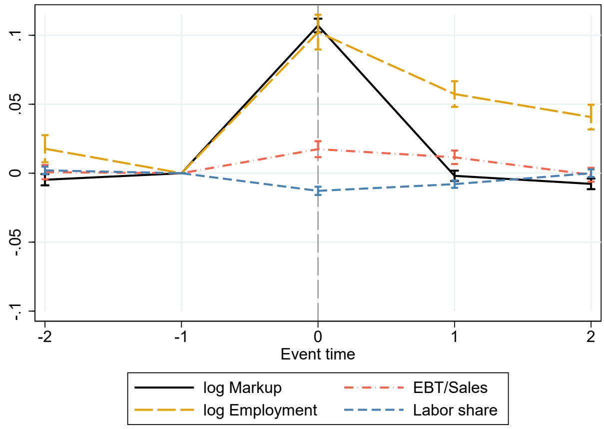
[18] See Goodman-Bacon 2018; Abraham and Sun 2018; Callaway and Sant Anna 2020; De Chaisemartin and d'Haultfoeuille 2020b.

entering and exiting from treatment, which is crucial in our case given that procurement is not necessarily an absorbing state. De Chaisemartin and d'Haultfoeuille (2020a) extend this framework to identify dynamic treatment effects across multiple periods. Intuitively, they rely on comparing switchers l periods after the switch to units whose treatment status remained constant.

We find that, at the year of getting their procurement contract, firms experience a sharp increase (10 percentage points) in their markups. The effect completely dissipates after one year, which is consistent with firms charging higher prices for public procurement contracts. Another potential explanation for this effect is a transitory productivity shock without persistence that reduced the marginal cost only in the year when the firm won public procurement and later it reverted to the mean. This is inconsistent with a large body of research that finds productivity shocks to be very persistent (Foster, Haltiwanger, and Syverson 2008; İmrohoroğlu and Tüzel 2014; Pozzi and Schivardi 2016). Similarly to our cross-sectional evidence, we find that firms increase employment, profit share and decrease labor share. These effects are more persistent, especially the effect on employment, which is in accordance with stickiness of employment contracts and other research on employment effects of procurement (Ferraz, Finan, and Szerman 2015; Gugler, Weichselbaumer, and Zulehner 2020). We do not see any pre-trends in markups, labor share and profit share. A pre-trend in employment is visible, but it is very small relative to the effect at the time of getting the public procurement. The reason behind this pre-trend might be differences in size growth rates between procurement and non-procurement firms, which we proceed to explore below.

In Figure 3.A.1 we present results using the De Chaisemartin and d'Haultfoeuille (2020b) estimator, which doesn't suffer from the heterogeneous treatment effect bias. We get very similar instantaneous effects on markups, profit share and labor share relative to the OLS approach. The dynamic effects differ, because of the way De Chaise-

FIGURE 3.2: Dynamic effects of procurement - OLS



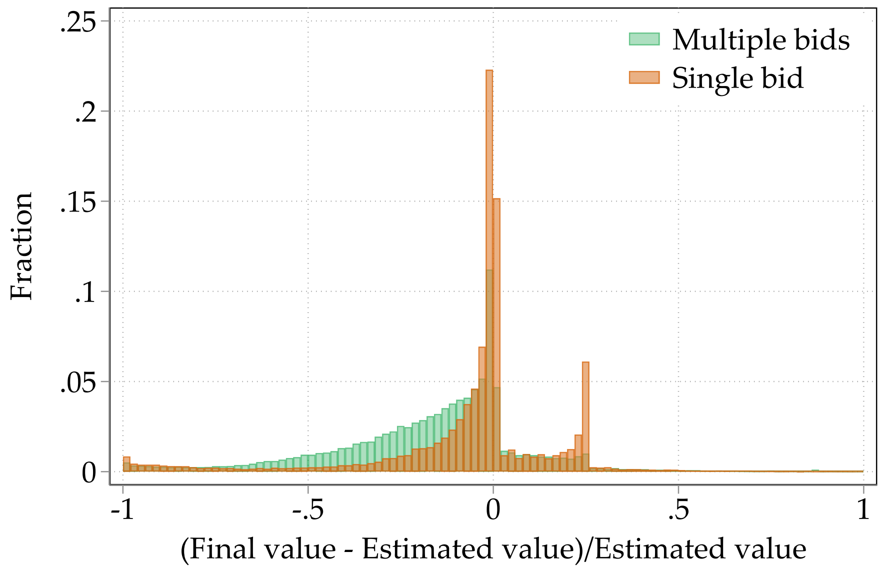
Note: This graph present the results from event study in Equation 3.12. The base period is -1 to which all other times periods are normalized and the treatment period is 0 (when procurement contract is allocated). EBT is earnings before taxes. Labor share is labor cost over sales

martin and d'Haultfoeuille (2020a) estimator constructs treatment and control groups. Their approach doesn't take into account that firms which got procurement at time t did get procurement again at $t+l$, so in Figure 3.A.1 the effects of procurement are more persistent on all variables. This happens because winning a procurement contract at time t increases the probability of winning contracts in the future. In the OLS approach outlined in Equation 3.12 future procurement is included as a variable in the equation, so the persistence in procurement contracts doesn't affect OLS results. The De Chaisemartin and d'Haultfoeuille (2020a) estimator doesn't show any pre-trends for markups, labor share and profit share. It shows, however, significant pre-trends in employment, indicating that firms that get a new procurement contract might be growing faster than other firms. To further alleviate concerns about possible differing trends across treated and non-treated firms, we also run a specification with firm specific time trends, thus controlling for possibly differing growth rates across firms. In Figure 3.A.2, we show that include unit time trends shrinks the pre-trends in employment towards zero, while the coefficients on markups, profit share and labor share stay almost the same. This robustness check suggests that we are estimating the true effect of procurement on these variables. The estimated effect of procurement on employment is 3.6%. This is similar to estimates from close auctions in Ferraz, Finan, and Szerman (2015); Gugler, Weichselbaumer, and Zulehner (2020), which arrive at estimates ranging between 3%-5% after one year. Ferraz, Finan, and Szerman (2015), and Gugler, Weichselbaumer, and Zulehner (2020) also find the effects of procurement on employment to be persistent.

3.5.4 *Markups, Competition and Number of Bids*

Until now, we have documented both static and dynamic effects of entry into the procurement sector on various firm-level characteristics. In this section, we provide corroborating evidence on the underlying mechanism that is driving high markups in public procurement. In what follows, we use the information on the number of bids in each procurement procedure. We will use the data on estimated costs of procurement projects, which are provided by the contracting authority. The contracting authority is mandated by law to estimate the cost of goods and services it wants to procure. Following European Commission (2019) and **?** we use the percent difference between actual cost of the public procurement and the estimated cost as a proxy for prices. We explore how the difference between actual and estimated cost systematically vary with number of bids. We use number of bids as proxy of competition for a given procurement auction. In Figure 3.3 we show the distribution of this cost differential across procurement contracts. There is a much larger mass around zero for single bid auctions, while for multiple bid auctions there is more mass below zero. This is consistent with competitive bidding driving prices down and reducing firm markups.

FIGURE 3.3: Actual vs estimated costs of procurement



Note. The Figure plots the distribution of contract value difference (Final value - Estimated values) for public procurement projects. Green bars show the contracts with multiple bids and red bars show contracts with single bid.

Given that we have very granular data, we can analyze how this price differential varies across single bid and multiple bid procurement procedures for the same firm, selling the same 8-digit CPV good or service, within the same county (approx. 150,000 people). In other words we can control for firm-year-product-county fixed effects. We find that the firm which wins a procurement contract within the same county for the same product wins contracts with consistently higher price differentials in single bid relative to multiple bid auctions (Table 3.7). This suggests that lower competition in public procurement increases the prices that firms charge. Given that Croatia is a small country there might be a single firm that supplies the good for the whole country and thus drives our results. In column 2 of Table 3.7 we check that our estimates hold if we exclude firms that sell to more than three municipalities (Croatia has 21 municipalities). Furthermore, the variation in prices might come from the fact that we do not observe the quality of the goods and services. We also perform the same exercise, but restrict the sample to standardized goods such as printing paper as in Bandiera, Prat, and Valletti (2009). We find very similar estimates, but the number of procurement contracts shrinks drastically, so standard errors increase correspondingly. This suggest that differences in quality are not driving our results.

Next we relate firm level markups to the number of bidders in the procurement auction. Given that we only have estimates of markups at the firm level we separate the firm FE from the Year-Product-County FE. This allows us to follow the change in firm level markups related to single bid auctions for the same product in the same county and year. We find that an average single bid auction for the same good is associated with higher firm level markups by 0.3%. The estimates increase if we consider firms that sell only in a few municipalities. The estimate is similar in size if we focus only on standardized products, but the amount of observations shrinks considerably (there are only 111 unique firms in column 6).

Next, we move to evidence at the firm-level by aggregating contract level data to firm-level. We define a firm as a single bidder if it has at least one single bidder contract in a given year. We find that single bidder firms are larger (see Table 3.A.3) on average in relative to other procurement firms both in terms of sales and employment. Furthermore, the average single bidder contract value is smaller relative to multi-bidder contract. Although the contract value is smaller, the estimated markups are even higher for single bidders relative to other procurement firms Table 3.A.5 . This is in line with the hypothesis

Table 3.7: Procurement price differentials, competition and markups

	(1)	(2)	(3)	(4)	(5)	(6)
	$\frac{\text{Final value-Estimated value}}{\text{Estimated value}}$				log Markup	
Single bidder	0.064***	0.079***	0.054*	0.003**	0.007***	0.006
	(0.005)	(0.006)	(0.028)	(0.001)	(0.002)	(0.009)
Observations	61,637	25,949	658	72,346	31,919	798
R-squared	0.62	0.70	0.59	0.93	0.95	0.87
Firm × Year× Product× County FE	✓	✓	✓			
Selling to max 3 counties		✓			✓	
Standardized products			✓			✓
Year×Product× County & Firm FE				✓	✓	✓

In columns (1)-(3) we present the estimates for δ_1 from regression:$Diff_{fit} = \delta_0 + \delta_1\, Single\ Bidder_{fit} + \phi_{fpct} + \nu_{fit}$, where $Diff_{fit} = \frac{\text{Final value}_{fit} - \text{Estimated value}_{fit}}{\text{Estimated value}_{fit}}$. f, m, c, i, p, t are the firm, municipality, county, contract, product and time indices, respectively. ϕ_{ftpc} is the firm-year-product-county fixed effect. In columns (4)-(6) we present estimates for δ_1 from regression:log Markup$ft = \delta_0 + \delta_1\, Single\ Bidder_{fit} + \phi_{pct} + \tau_f + \nu_{ft}$. The number of observations is smaller in columns (1)-(3) because there are missing values on estimated value of procurement project. The regression standard errors in parentheses clustered at firm level, where significance level is defined as *** p<0.01, ** p<0.05, * p<0.1.

that low competition is driving high markups in this sector. Finally, we perform an event study as in refeq:diff to disentangle impact of winning a single bidder contract relative to multiple bid contracts and find that all effects on markups, profits, labor share and employment are more pronounced for these firms (Table 3.A.3). For example, single bidder firms have 3% higher markups after winning a single bid contract than firms winning a procurement contract in a multi-bid environment.

3.6 CONCLUSION

In this paper, we find evidence that firms command higher market power in public procurement relative to the private market and this may be driven by low competition. Employing the recent advancement in the estimation techniques of firm-level markups, we show that firms in the procurement sector charge much higher markups relative to firm that only sell in the private sector. We provide suggestive evidence that these markups difference emerges due to high pricing rather than lower marginal cost of firms producing in the procurement sector. As further test of this finding, we show that procurement sector firms also earn high profits, suggesting that the high markups are not an artifact of bureaucratic or administrative costs of doing business with government. Furthermore, we show that the markup premium of producing in government sector increase with the share of sales coming from that sector. Finally, we exploit the data on number of bids for a select group of contracts and document that number of bids are negatively correlated with markup premium and positively correlated with the associated cost of the project. All these evidence taken together highlights that the low competition in the public procurement may rationalize high markups of firms that produce for the public sector.

Given the suggestive evidence documented in this paper, a natural next step is to understand the causes of the low competition in the government sector and highlighting the macroeconomic implications of these distortions. In the future, we want to combine these empirical facts together with a theoretical framework to quantify the impact of various forces that are responsible for low competition in the procurement sector.

3

APPENDIX

3.A FIGURES AND TABLES

3.A.1 *Tables*

Table 3.A.1: Procurement dataset aggregates

Year	Total published procurement	Total procurement in our dataset	Unpublished procurement
2013	33	25	6
2014	33	31	9
2015	31	26	9
2016	35	35	10
2017	31	31	9
2018	37	37	10
2019	43	43	11

Note: Procurement figures in our dataset are drawn from Ministry of economy and sustainable development. The overall procurement activity in the economy is takes from Official Gazette. Numbers are billions of Kuna. 1 EUR $\approx$ 7.4 Kuna.

Table 3.A.2: Political connections, procurement and firm descriptives

	All Firms	Connected	Procurement	Connected and Procurement
Revenue (Mean)	8,498,811	53,481,001	58,515,672	221,306,285
Capital (Mean)	3,439,618	27,547,091	15,586,897	68,129,141
Employees (Mean)	11	58	59	217
Procurement dummy	0.04	0.13		
Procurement value (Mean)	341,625	2,780,636	7,789,598	22,107,196
Wage bill (Mean)	1,071,563	6,556,182	6,478,207	24,583,862
Wages (Mean)	75,747	98,868	109,299	125,175
Loans (Mean)	1,624,731	14,144,206	7,733,551	31,133,466
% of total procurement value				40%
Observations	384,111	15,551	13,993	1,956
Unique firms	94,702	3,721	5,177	587

Note: All Numbers are in Kuna. 1 EUR $\approx$ 7.4 Kuna. The political connection are defined as in Section 3.2. First column presents statistics for all firms, second is for firms with political connections, third column is for firms active in public procurement, fourth column is for firms that have political connections and are active in public procurement. Capital is mean stock of fixed assets at firm-level. Wages are average wage paid by firm to all of its employees. Loans are outstanding value of borrowing in a given year. Procurement dummy captures the percentage of firms active in procurement sector. Unique firms represents the number of firms in each category. The time period is 2013-2019.

Table 3.A.3: Descriptive Statistics: Firm Level

	All Firms	Procurement	Single - Bidders
Revenue (000s)	8442	58033	74817
Capital (000s)	3374	15280	20401
Employees	11.20	57.71	70.25
Procurement Contract Value (000s)		8178	11883
Single-Bidder Contract Value (000s)			4990
Wagebill (000s)	1070	6410	8204
Observations	384,111	16,797	9,582
Unique firms	94,702	5,775	3,810

Note: All Numbers are in Kuna. 1 EUR $\approx$ 7.4 Kuna. The single bidder contracts are defined as contracts that received only one bid and firms those have atleast one single-bidder contract in a given year would be classified as single-bidder firms in our data. First column presents statistics for all firms, second is for firms with procurement contracts, third column is for firms that are classified as single-bidders. Procurement Contract Value is the average value of contract awarded and Single-Bidder Contract Value is the average value of single-bidder contract awarded in the sample period. Capital is mean stock of fixed assets at firm-level. Unique firms represents the number of firms in each category. The time period is 2013-2019.

Table 3.A.4: Political Connection and Entry into Procurement

	(1)	(2)	(3)	(4)	(5)	(6)
	Procurement dummy$_{it}$			Single bidder$_{it}$		
Politically connected$_{ft}$	0.12***	0.11***	0.11***	0.12***	0.09**	0.09***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
4 digit sector and year FE		✓	✓		✓	✓
Municipality and year FE			✓			✓
Observations	319,063	318,948	318,866	13,997	13,549	12,893
R-squared	0.00	0.06	0.07	0.00	0.16	0.23

Note: Formally, the first three columns show the coefficient α_0 from the estimation of the following regression model:

$$Procurement_{ft} = \alpha_0 + \alpha_0\, Politically\, connected_{ft} + b'_{ft}\sigma + \nu_{1,ft},$$

and the column 4,5 and 6 show the coefficient δ_0 from the estimation of the following regression model:

$$Single\, bidder_{ft} = \delta_0 + \delta_1\, Politically\, connected_{ft} + c'_{ft}\sigma + \nu_{2,ft},$$

The regression standard errors in parentheses clustered at 4 digit NACE rev.2 sector level × year. *** p<0.01, ** p<0.05, * p<0.1. Regressions in columns (4)-(6) are done for the subsample of firms that have public procurement.

Table 3.A.5: Markups and Procurement I: Cross-Sectional Analysis

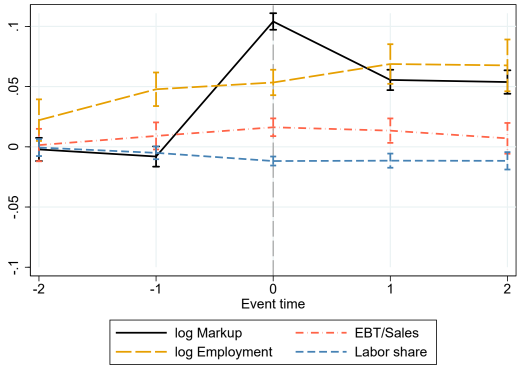
	(1) Manufacturing	(2) Construction	(3) Wholesale & Retail	(4) Transportation & Storage	(5) Accommodation & Food	(6) Information & Comm.	(7) Financial & Insurance	(8) Professional, S & T Services	(9) Admin. & Support
Procurement	0.06***	0.25***	0.04***	0.01	0.03	-0.03	0.06	0.14***	-0.03
	(0.01)	(0.02)	(0.01)	(0.02)	(0.04)	(0.03)	(0.08)	(0.01)	(0.02)
Single bidder	0.02***	0.03***	0.02***	-0.01	-0.01	-0.01	0.11	0.00	0.04*
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.02)	(0.09)	(0.01)	(0.02)
Observations	58,139	43,092	102,582	18,704	28,840	17,842	6,514	60,845	14,229
R-squared	0.23	0.21	0.09	0.35	0.18	0.20	0.17	0.46	0.55
Sector#year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note. This table presents the estimated coefficientsδ_1 and δ_2 from regression specification:

$$ln\, \mu_{ft} = \delta_0 + \delta_1\, Procurement_{ft} + \delta_2 Single\, bidder_{ft} + b'_{ft}\sigma + \nu_{ft},$$

We always use capital and lagged labor as controls. Clustered standard errors at sector×years are in brackets below the coefficient values. Manufacturing, Construction and Wholesale & Retail capture 85% of the sales and procurement activity in the economy. Furthermore, Professional, Scientific & Technical services sector capture another 7% of the procurement contracts in value. The regression standard errors in parentheses clustered at 4 digit NACE rev.2 sector level × year where significance level is defined as *** p<0.01, ** p<0.05, * p<0.1.

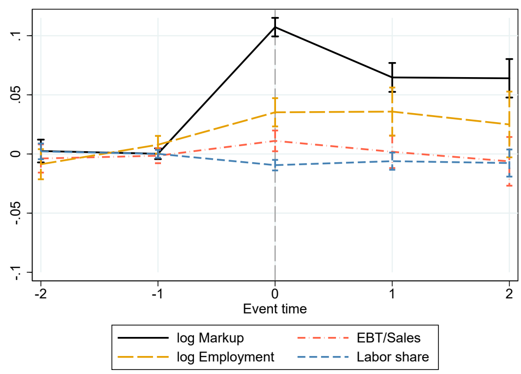
FIGURE 3.A.1: De Chaisemartin and d'Haultfoeuille (2020b) Staggered Difference-in-Differences



Note: This figure shows graphically the estimates using De Chaisemartin and d'Haultfoeuille (2020b) staggered difference-in-differences approach. EBT is earnings before taxes. Labor share is labor cost over sales. We control for lagged log capital, log sales and the share of exports in sales. In De Chaisemartin and d'Haultfoeuille (2020b) there is no base period, e.g. at event time -1 the comparison is between switchers at time 0 and the never treated.

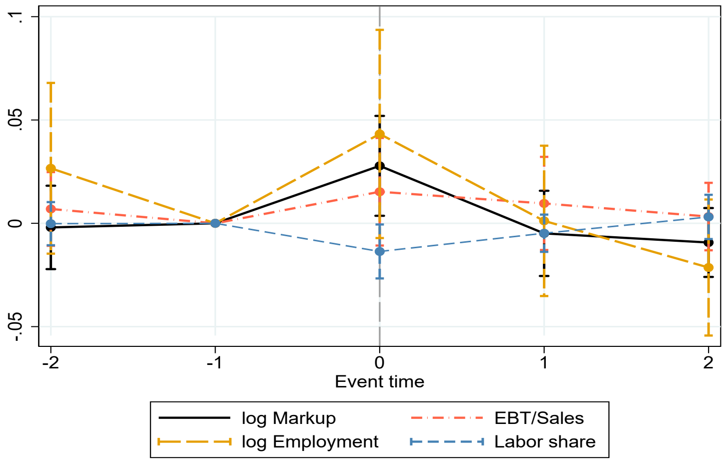
3.A.2 *Figures*

FIGURE 3.A.2: De Chaisemartin and d'Haultfoeuille (2020b) Staggered Difference-in-Differences with unit specific time trends



Note: This figure shows graphically the estimates using De Chaisemartin and d'Haultfoeuille (2020b) staggered difference-in-differences approach with unit specific time trends. EBT is earnings before taxes. Labor share is wagebill over sales. We control for lagged log capital, log sales and the share of exports in sales. In De Chaisemartin and d'Haultfoeuille (2020b) there is no base period, e.g. at event time -1 the comparison is between switchers at time 0 and the never treated.

FIGURE 3.A.3: Dynamic Effects Procurement on Single Bidder Firms-OLS



Note: This graph presents the results from event study in 3.12. The dependent variable in this case if single bidder dummy and all effects are relative to procurement firms. The base period is -1 to which all other times periods are normalized and the treatment period is 0 (when procurement contract is allocated). EBT is earnings before taxes. Labor share is labor cost over sales

BIBLIOGRAPHY

Andrew B Abel and Janice C Eberly. A unified model of investment under uncertainty. *American Economic Review*, 84:1369–1384, 1994.

Andrew B Abel and Janice C Eberly. Optimal investment with costly reversibility. *The Review of Economic Studies*, 63(4):581–593, 1996.

Sarah Abraham and Liyang Sun. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *arXiv preprint arXiv:1804.05785*, 2018.

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

Philippe Aghion, Antonin Bergeaud, Timo Boppart, Peter J Klenow, and Huiyu Li. A theory of falling growth and rising rents. 2019.

Alberto Alesina and Eliana La Ferrara. Ethnic diversity and economic performance. *Journal of Economic Literature*, 43(3):762–800, 2005.

John Asker, Allan Collard-Wexler, and Jan De Loecker. Dynamic inputs and resource (mis) allocation. *Journal of Political Economy*, 122(5): 1013–1063, 2014.

Jose Asturias, Manuel García-Santana, and Roberto Ramos. Competition and the welfare gains from transportation infrastructure: Evidence from the golden quadrilateral of india. *Journal of the European Economic Association*, 17(6):1881–1940, 2019.

Andrew Atkeson. Alternative facts regarding the labor share. *Review of Economic Dynamics*, 37:S167–S180, 2020.

Andrew Atkeson and Ariel Burstein. Pricing-to-market, trade costs, and international relative prices. *American Economic Review*, 98(5): 1998–2031, 2008.

David Autor, 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*, 135(2):645–709, 2020.

Meghana Ayyagari, Asli Demirguc-Kunt, and Vojislav Maksimovic. Does local financial development matter for firm lifecycle in india? *The World Bank*, 2014.

Patrick Bajari and Steven Tadelis. Incentives versus transaction costs: A theory of procurement contracts. *Rand journal of Economics*, pages 387–407, 2001.

Patrick Bajari, Robert McMillan, and Steven Tadelis. Auctions versus negotiations in procurement: an empirical analysis. *The Journal of Law, Economics, & Organization*, 25(2):372–399, 2009.

Audinga Baltrunaite, Cristina Giorgiantonio, Sauro Mocetti, and Tommaso Orlando. Discretion and supplier selection in public procurement. *The Journal of Law, Economics, and Organization*, 37(1):134–166, 2021.

Oriana Bandiera, Andrea Prat, and Tommaso Valletti. Active and passive waste in government spending: evidence from a policy experiment. *American Economic Review*, 99(4):1278–1308, 2009.

Abhijit Banerjee and Kaivan Munshi. How efficiently is capital allocated? evidence from the knitted garment industry in tirupur. *The Review of Economic Studies*, 71(1):19–42, 2004.

Abhijit V Banerjee and Esther Duflo. Growth theory through the lens of development economics. *Handbook of Economic Growth*, 1:473–552, 2005.

Abhijit V Banerjee and Benjamin Moll. Why does misallocation persist? *American Economic Journal: Macroeconomics*, 2(1):189–206, 2010.

Simcha Barkai. Declining labor and capital shares. *Stigler Center for the Study of the Economy and the State New Working Paper Series*, 2, 2016.

Robert J Barro. Double-counting of investment. Technical report, National Bureau of Economic Research, 2019.

Marc F Bellemare, Christopher B Barrett, and David R Just. The welfare impacts of commodity price volatility: evidence from rural ethiopia. *American Journal of Agricultural Economics*, 95(4):877–899, 2013.

Chris Bidner and Mukesh Eswaran. A gender-based theory of the origin of the caste system of india. *Journal of Development Economics*, 114: 142–158, 2015.

Mark Bils, Peter J Klenow, and Cian Ruane. Misallocation or mismeasurement? Technical report, National Bureau of Economic Research, 2020.

Jonathan Brogaard, Matthew Denes, and Ran Duchin. Political influence and the renegotiation of government contracts. *The Review of Financial Studies*, 34(6):3095–3137, 2021.

Francisco J. Buera and Yongseok Shin. Financial frictions and the persistence of history: A quantitative exploration. *Journal of Political Economy*, 121(2):221–272, 2013. ISSN 00223808, 1537534X. URL http://www.jstor.org/stable/10.1086/670271.

Francisco J Buera, Joseph P Kaboski, and Yongseok Shin. Finance and development: A tale of two sectors. *The American Economic Review*, 101(5):1964–2002, 2011.

Francisco J Buera, Joseph P Kaboski, and Yongseok Shin. Entrepreneurship and financial frictions: A macrodevelopment perspective. *Annual Reviews of Economics*, 7(1):409–436, 2015.

Jeremy Bulow and Paul Klemperer. Auctions vs. negotiations. Technical report, American Economic Review, 1996.

Marco Cagetti and Mariacristina De Nardi. Entrepreneurship, frictions, and wealth. *Journal of Political Economy*, 114(5):835–870, 2006.

Brantly Callaway and Pedro HC Sant Anna. Difference-in-differences with multiple time periods. *Journal of Econometrics*, 2020.

Rodrigo Carril and Mark Duggan. The impact of industry consolidation on government procurement: Evidence from department of defense contracting. *Journal of Public Economics*, 184:104141, 2020.

Francesco Caselli and James Feyrer. The marginal product of capital. *The Quarterly Journal of Economics*, 122(2):535–568, 2007.

Judith A Chevalier and David S Scharfstein. Capital market imperfections and countercyclical markups: Theory and evidence. Technical Report 4, 1996.

Andrea Chiavari. The macroeconomics of rising returns to scale: Costumer acquisition, markups, and dynamism. 2021.

Federico Cingano and Paolo Pinotti. Politicians at work: The private returns and social costs of political connections. *Journal of the European Economic Association*, 11(2):433–465, 2013.

Gian Luca Clementi and Berardino Palazzo. On the calibration of competitive industry dynamics models. *Unpublished working paper*, 2016a.

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

Gian Luca Clementi and Berardino Palazzo. Investment and the cross-section of equity returns. *The Journal of Finance*, 74(1):281–321, 2019.

Emanuele Colonnelli and Mounu Prem. Corruption and firms: evidence from randomized audits in brazil. *Available at SSRN*, 2931602, 2017.

Russell W Cooper and John C Haltiwanger. On the nature of capital adjustment costs. *The Review of Economic Studies*, 73(3):611–633, 2006.

Carol Corrado, Charles Hulten, and Daniel Sichel. Intangible capital and us economic growth. *Review of income and wealth*, 55(3):661–685, 2009.

Carol A Corrado and Charles R Hulten. How do you measure a" technological revolution"? *American Economic Review*, 100(2):99–104, 2010.

Nicolas Crouzet and Janice C Eberly. Understanding weak capital investment: The role of market concentration and intangibles. *National Bureau of Economic Research*, 2019.

Harish Damodaran. *India's new capitalists: caste, business, and industry in a modern nation*. Springer, 2008.

Joel M David and Venky Venkateswaran. The sources of capital misallocation. *American Economic Review*, 109(7):2531–67, 2019.

Steven J Davis, John Haltiwanger, Ron Jarmin, Javier Miranda, Christopher Foote, and Eva Nagypal. Volatility and dispersion in business growth rates: Publicly traded versus privately held firms [with comments and discussion]. *NBER macroeconomics annual*, 21:107–179, 2006.

Clément De Chaisemartin and Xavier d'Haultfoeuille. Difference-in-differences estimators of intertemporal treatment effects. *Available at SSRN 3731856*, 2020a.

Clément De Chaisemartin and Xavier d'Haultfoeuille. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96, 2020b.

Eeckhout Jan De Loecker, Jan and Simon Mongey. Quantifying market power. 2019.

Jan De Loecker and Paul T Scott. Estimating market power evidence from the us brewing industry. Technical report, National Bureau of Economic Research, 2016.

Jan De Loecker and Frederic Warzynski. Markups and firm-level export status. *American economic review*, 102(6):2437–71, 2012.

Jan De Loecker, Pinelopi K Goldberg, Amit K Khandelwal, and Nina Pavcnik. Prices, markups, and trade reform. *Econometrica*, 84(2): 445–510, 2016.

Jan De Loecker, Jan Eeckhout, and Gabriel Unger. The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics*, 135(2):561–644, 2020.

Suresh De Mel, David McKenzie, and Christopher Woodruff. Returns to capital in microenterprises: evidence from a field experiment. *The Quarterly Journal of Economics*, 123(4):1329–1372, 2008.

Maarten De Ridder. Market power and innovation in the intangible economy. 2019.

Francesco Decarolis. Awarding price, contract performance, and bids screening: Evidence from procurement auctions. *American Economic Journal: Applied Economics*, 6(1):108–32, 2014.

Francesco Decarolis, Raymond Fisman, Paolo Pinotti, and Silvia Vannutelli. Rules, discretion, and corruption in procurement: Evidence from italian government contracting. Technical report, National Bureau of Economic Research, 2020a.

Francesco Decarolis, Leonardo M Giuffrida, Elisabetta Iossa, Vincenzo Mollisi, and Giancarlo Spagnolo. Bureaucratic competence and procurement outcomes. *The Journal of Law, Economics, and Organization*, 36(3):537–597, 2020b.

Sonalde Desai, Reeve Vanneman, and New Delhi National Council of Applied Economic Research. *India Human Development Survey (IHDS), 2005*. Inter-university Consortium for Political and Social Research [distributor], 2018-08-08. https://doi.org/10.3886/ICPSR22626.v12, 2018.

Ashwini Deshpande, Smriti Sharma, et al. Entrepreneurship or survival? caste and gender of small business in india. Technical report, Centre for Development Economics Department of Economics, Delhi School of Economics, 2013.

Manali S Deshpande. History of the indian caste system and its impact on india today. 2010.

Julian di Giovanni, Manuel Garica-Santana, Priit Jeenas, Enrique Moral-Benito, and Josep Pijoan-Mas. Government procurement and credit growth: Firm-level evidence and macro consequences.

Mark Doms and Timothy Dunne. Capital adjustment patterns in manufacturing plants. *Review of economic dynamics*, 1(2):409–429, 1998.

William Easterly and Ross Levine. Africa's growth tragedy: policies and ethnic divisions. *The Quarterly Journal of Economics*, 112(4):1203–1250, 1997.

Chris Edmond, Virgiliu Midrigan, and Daniel Yi Xu. How costly are markups? *NBER working papers*, 2018.

Andrea L Eisfeldt and Dimitris Papanikolaou. Organization capital and the cross-section of expected returns. *The Journal of Finance*, 68(4): 1365–1406, 2013.

Michael WL Elsby, Bart Hobijn, and Ayşegül Şahin. The decline of the us labor share. *Brookings Papers on Economic Activity*, 2013(2):1–63, 2013.

Andrés Erosa, Luisa Fuster, Gueorgui Kambourov, and Richard Rogerson. Hours, occupations, and gender differences in labor market outcomes. *NBER Working Paper*, (w23636), 2017.

European Commission. Policy measure fact sheet croatia public procurement act thematic objective 4. Technical report, European Commission, 2017a.

European Commission. Making public procurement work in and for europe. Technical report, European Commission, 2017b.

European Commission. Single bidding and noncompetitive tendering procedures in eu co-funded projects. Technical report, European Commission, 2019.

European Commission. Business' attitudes toward corruption in the eu. Technical report, European Commission, 2020.

Michael Ewens, Ryan H Peters, and Sean Wang. Acquisition prices and the measurement of intangible capital. Technical report, National Bureau of Economic Research, 2019.

Ernst Fehr and Karla Hoff. Tastes, castes and culture: The influence of society on preferences. *The Economic Journal*, 121(556):F396–F412, 2011.

Claudio Ferraz, Frederico Finan, and Dimitri Szerman. Procuring firm growth: the effects of government purchases on firm dynamics. Technical report, National Bureau of Economic Research, 2015.

Raymond Fisman, Daniel Paravisini, and Vikrant Vig. Cultural proximity and loan outcomes. *American Economic Review*, 107(2):457–92, 2017.

Lucia Foster, John Haltiwanger, and Chad Syverson. Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, 98(1):394–425, 2008.

Rosemary R Fullerton, Cheryl S McWatters, and Chris Fawson. An examination of the relationships between jit and financial performance. *Journal of Operations Management*, 21(4):383–404, 2003.

Amit Gandhi, Salvador Navarro, and David A Rivers. On the identification of gross output production functions. *Journal of Political Economy*, 128(8):2973–3016, 2020.

Manuel Garcia-Santana and Josep Pijoan-Mas. The reservation laws in india and the misallocation of production factors. *Journal of Monetary Economics*, 66:193–209, 2014.

Maria Paula Gerardino, Stephan Litschig, and Dina Pomeranz. Can audits backfire? evidence from public procurement in chile. *NBER Working Papers*, (23978), 2017.

Simon Gilchrist, Raphael Schoenle, Jae Sim, and Egon Zakrajšek. Inflation dynamics during the financial crisis. *American Economic Review*, 107(3):785–823, 2017.

Victor P Goldberg. Competitive bidding and the production of precontract information. *The Bell Journal of Economics*, pages 250–261, 1977.

Andrew Goodman-Bacon. Difference-in-differences with variation in treatment timing. Technical report, National Bureau of Economic Research, 2018.

Sampreet Goraya and Akhil Ilango. Demand segmentation and markups. *Mimeo*, 2020.

Gustavo Grullon, Yelena Larkin, and Roni Michaely. Are us industries becoming more concentrated? *Review of Finance*, 23(4):697–743, 2019.

Klaus Gugler, Michael Weichselbaumer, and Christine Zulehner. Competition in the economic crisis: Analysis of procurement auctions. *European economic review*, 73:35–57, 2015.

Klaus Gugler, Michael Weichselbaumer, and Christine Zulehner. Employment behavior and the economic crisis: Evidence from winners and runners-up in procurement auctions. *Journal of Public Economics*, 182:104112, 2020.

Nezih Guner, Gustavo Ventura, and Yi Xu. Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics*, 11(4):721–744, 2008.

Germán Gutiérrez and Thomas Philippon. Investment-less growth: An empirical investigation. Technical report, National Bureau of Economic Research, 2016.

Charles J Hadlock and Joshua R Pierce. New evidence on measuring financial constraints: Moving beyond the kz index. *The Review of Financial Studies*, 23(5):1909–1940, 2010.

Robert E Hall. Quantifying the lasting harm to the us economy from the financial crisis. *NBER Macroeconomics Annual*, 29(1):71–128, 2015.

Robert E Hall and Charles I Jones. Why do some countries produce so much more output per worker than others? *The quarterly journal of economics*, 114(1):83–116, 1999.

Juan Carlos Hallak and Jagadeesh Sivadasan. Product and process productivity: Implications for quality choice and conditional exporter premia. *Journal of International Economics*, 91(1):53–67, 2013.

Jonathan Haskel and Stian Westlake. *Capitalism without capital: the rise of the intangible economy*. Princeton University Press, 2018.

Jonas Hjort. Ethnic divisions and production in firms. *The Quarterly Journal of Economics*, 129(4):1899–1946, 2014.

Karla Hoff and Priyanka Pandey. Discrimination, social identity, and durable inequalities. *The American Economic Review*, 96(2):206–211, 2006.

Hugo Hopenhayn, Julian Neira, and Rish Singhania. The rise and fall of labor force growth: Implications for firm demographics and aggregate trends. *NBER Working Paper*, pages 1–28, 2018.

Hugo A Hopenhayn. Firms, misallocation, and aggregate productivity: a review. *Annual Review of Economics*, 6(1):735–770, 2014.

Chang-Tai Hsieh and Peter J. Klenow. Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics*, CXXIV (November), 2009a.

Chang-Tai Hsieh and Peter J Klenow. Misallocation and manufacturing tfp in china and india. *The Quarterly journal of economics*, 124(4): 1403–1448, 2009b.

Chang-Tai Hsieh and Peter J Klenow. The life cycle of plants in india and mexico. *The Quarterly Journal of Economics*, 129(3):1035–1084, 2014.

Chang-Tai Hsieh and Esteban Rossi-Hansberg. The industrial revolution in services. 2019.

Chang-Tai Hsieh, Erik Hurst, Charles I Jones, and Peter J Klenow. The allocation of talent and us economic growth. *Econometrica*, 87(5): 1439–1474, 2019.

Ayşe İmrohoroğlu and Şelale Tüzel. Firm-level productivity, risk, and return. *Management Science*, 60(8):2073–2090, 2014.

Lakshmi Iyer, Tarun Khanna, and Ashutosh Varshney. Caste and entrepreneurship in india. *Economic and Political Weekly*, 48(6):52–60, 2013.

Surinder S Jodhka. Sikhism and the caste question: Dalits and their politics in contemporary punjab. *Contributions to Indian sociology*, 38 (1-2):165–192, 2004.

Surinder S Jodhka. Dalits in business: Self-employed scheduled castes in north-west india. *Economic and Political Weekly*, pages 41–48, 2010.

Karam Kang and Robert A Miller. Winning by default: Why is there so little competition in government procurement? *Unpublished working paper*, 2017.

Loukas Karabarbounis and Brent Neiman. The global decline of the labor share. *The Quarterly journal of economics*, 129(1):61–103, 2013.

Matthias Kehrig and Nicolas Vincent. The micro-level anatomy of the labor share decline. *The Quarterly Journal of Economics*, 136(2): 1031–1087, 2021.

Aubhik Khan and Julia K Thomas. Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics. *Econometrica*, 76(2):395–436, 2008.

Dongya Koh, Raül Santaeulàlia-Llopis, and Yu Zheng. Labor share decline and intellectual property products capital. *Econometrica*, 88 (6):2609–2628, 2020.

Elena Krasnokutskaya and Katja Seim. Bid preference programs and participation in highway procurement auctions. *American Economic Review*, 101(6):2653–86, 2011.

Kory Kroft, Yao Luo, Magne Mogstad, and Bradley Setzler. Imperfect competition and rents in labor and product markets: The case of the construction industry. Technical report, National Bureau of Economic Research, 2020.

Maurice Kugler and Eric Verhoogen. Prices, plant size, and product quality. *The Review of Economic Studies*, 79(1):307–339, 2012.

Jean-Jacques Laffont and Jean Tirole. Auctioning incentive contracts. *Journal of Political Economy*, 95(5):921–937, 1987.

Jean-Jacques Laffont and Jean Tirole. Adverse selection and renegotiation in procurement. *The Review of Economic Studies*, 57(4):597–625, 1990.

Jean-Jacques Laffont and Jean Tirole. *A theory of incentives in procurement and regulation*. MIT press, 1993.

Danial Lashkari, Arthur Bauer, and Jocelyn Boussard. Information technology and returns to scale. *Available at SSRN 3458604*, 2018.

Yoonsoo Lee and Toshihiko Mukoyama. Productivity and employment dynamics of us manufacturing plants. *Economics Letters*, 136:190–193, 2015.

Baruch Lev and Feng Gu. *The end of accounting and the path forward for investors and managers*. John Wiley & Sons, 2016.

James Levinsohn and Amil Petrin. Estimating production functions using inputs to control for unobservables. *The review of economic studies*, 70(2):317–341, 2003.

Ellen R McGrattan and Edward C Prescott. Technology capital and the us current account. *American Economic Review*, 100(4):1493–1522, 2010a.

Ellen R McGrattan and Edward C Prescott. Unmeasured investment and the puzzling us boom in the 1990s. *American Economic Journal: Macroeconomics*, 2(4):88–123, 2010b.

Ellen R McGrattan and Edward C Prescott. A reassessment of real business cycle theory. *American Economic Review*, 104(5):177–82, 2014.

Virgiliu Midrigan and Daniel Yi Xu. Finance and misallocation: Evidence from plant-level data. *The American Economic Review*, 104(2): 422–458, 2014.

Benjamin Moll. Productivity losses from financial frictions: can self-financing undo capital misallocation? *American Economic Review*, 104 (10):3186–3221, 2014.

Jose G Montalvo and Marta Reynal-Querol. Ethnic diversity and economic development. *Journal of Development Economics*, 76(2):293–323, 2005.

Kaivan Munshi. Caste networks in the modern indian economy. In *Development in India*, pages 13–37. Springer, 2016.

Masao Nakamura, Sadao Sakakibara, and Roger Schroeder. Adoption of just-in-time manufacturing methods at us-and japanese-owned plants: some empirical evidence. *IEEE transactions on engineering management*, 45(3):230–240, 1998.

Daniel M Neuman. Caste and social stratification among muslims in india. *The Journal of Asian Studies*, 40(2):400–402, 1981.

G Steven Olley and Ariel Pakes. The dynamics of productivity in the telecommunications equipment industry. Technical report, 1996.

Ariel Pakes. Dynamic structural models: Problems and prospects. mixed continuous discrete controls and market interactions. Technical report, Cowles Foundation for Research in Economics, Yale University, 1991.

Ryan H Peters and Lucian A Taylor. Intangible capital and the investment-q relation. *Journal of Financial Economics*, 123(2):251–272, 2017.

Andrea Pozzi and Fabiano Schivardi. Demand or productivity: What determines firm growth? *The RAND Journal of Economics*, 47(3): 608–630, 2016.

Harish K Puri. Scheduled castes in sikh community: A historical perspective. *Economic and political weekly*, pages 2693–2701, 2003.

Diego Restuccia and Richard Rogerson. Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic dynamics*, 11(4):707–720, 2008.

Daniel F Spulber. Auctions and contract enforcement. *Journal of Law, Economics, & Organization*, 6(2):325–344, 1990.

Sukhadeo Thorat and Nidhi Sadana. Caste and ownership of private enterprises. *Economic and Political Weekly*, pages 13–16, 2009.

Espen Villanger. Entrepreneurial abilities and barriers to microenterprise growth: A case study in nepal. *The Journal of Entrepreneurship*, 24(2):115–147, 2015.

Patrick L Warren. Contracting officer workload, incomplete contracting, and contractual terms. *The RAND Journal of Economics*, 45(2):395–421, 2014.

Lichen Zhang. Intangibles, concentration, and the labor share. *Mimeo*, 2019.