The London School of Economics and Political Science

ESSAYS ON INSTITUTIONS AND ECONOMIC PERFORMANCE

YAN LIANG

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy, London, June 2018

DECLARATION

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without my prior written consent.

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 33,230 words.

STATEMENT OF INCLUSION OF PREVIOUS WORK

I can confirm that chapter 2 is extended from the previous study (for a Master of Research award) I undertook at the London School of Economics.

STATEMENT OF USE OF THIRD PARTY FOR EDITORIAL HELP

I can confirm that chapter 1 of my thesis was copy edited for conventions of language, spelling and grammar by the LSE Language Center, and chapter 2 and 3 by Cambridge Proofreading LLC.

ABSTRACT

This dissertation consists of three essays in the intersection of macroeconomics and international trade.

The first essay studies the causes and consequences of the differences in the use of outsourcing across countries. I start by observing that there is considerable variation in outsourcing intensity across countries. I then show that this pattern can be rationalized in a theoretical framework that combines a Coase-Williamson view of the firm with Kiyotaki-Moore-Manova view of financial friction. The model pins down the intensity of outsourcing and shows how it varies with the financial characteristics of the suppliers. Econometric evidence reveals that the model is consistent with the features of both sectoral-level and firm-level data. The model also clarifies two conflicting mechanisms of outsourcing on productivity. Quantitative analysis reveals that both mechanisms are quantitatively significant so that the net effect on aggregate productivity is modest. My study implies that outsourcing is unlikely a significant source of cross-country differences in productivity.

In the second essay, I examine how heterogeneous market power affects the quantification of resources misallocation within sector. I extend the Hsieh-Klenow framework of misallocation to allow for heterogeneous market power and use the model to study the impact of resources misallocation on India's aggregate productivity. Quantitative results show that heterogeneous market power has a large impact on the quantification of misallocation. In particular, in the presence of heterogeneous market power, the impact of tax-related distortions on aggregate productivity is about one-seventh of the effect found by previous literature. My study implies that increasing market competition is an effective way to reduce market power and enhance aggregate productivity.

The third essay studies how factors of different quality are allocated to the production chain. The essay starts by unveiling two systematic patterns in factor inputs and factor rewards along production chains. I then show that these patterns can be rationalized in a theoretical framework with heterogeneous factors. In the model, products become increasingly complex as they move along the production chain. Downstream firms hire skilled workers to process complex products. To the extent that skill is strongly complementary to the quality of physical capital, downstream firms also employ high quality capital goods. The analysis sheds light on the organization of factors along production chains.

ACKNOWLEDGMENTS

I am grateful to Francesco Caselli and Gianmarco Ottaviano for their invaluable guidance and support, and to Swati Dhingra for detailed discussions and advice. My gratitude extends to Karun Adusumilli, Philippe Aghion, David Baqaee, Johannes Boehm, Paola Conconi, Pedro Franco De Campos Pinto, Wouter Den Haan, Thomas Drechsel, Jiajia Gu, Chao He, Hanwei Huang, Ethan Ilzetzki, Sevim Kosem, Per Krusell, Rachel Ngai, Kieu-Trang Nguyen, Michael Peters, Jonathan Pinder, Frank Pisch, Veronica Rappoport, Ricardo Reis, Filip Rozsypal, Thomas Sampson, Silvana Tenreyro, Catherine Thomas, Shengxing Zhang, and seminar participants at LSE, Deakin, LBS TADC, LKYSPP, Sussex, and Tokyo for helpful comments. Financial support from the Centre for Macroeconomics is gratefully acknowledged. All errors are my own.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1

THE IMPACT OF FINANCIAL DEVELOPMENT ON OUTSOURCING AND AGGREGATE PRODUCTIVITY

This paper highlights a novel channel through which financial development affects aggregate productivity, namely endogenous changes in the extent to which firms outsource production of intermediate inputs. I present a multi-sector general equilibrium model of outsourcing with heterogeneous firms, and clarify the role of financial development in shaping outsourcing, entry, prices, and productivity. I then empirically test the key implication that financial development increases outsourcing, especially when suppliers are highly dependent on external finance and have fewer tangible assets. This implication is strongly supported by both cross-country sectoral input-output data and firm-level data. Finally, I structurally estimate the model and perform counterfactual experiments. Quantitative analysis shows that financial development has a sizable impact on outsourcing, but its effect on aggregate productivity is relatively modest. Outsourcing implies a more efficient use of resources, but transaction costs generate a powerful negative effect through prices, offsetting the majority of productivity gains.

1.1 INTRODUCTION

An important question in macroeconomics concerns how financial development affects real activity. Most research focuses on how financial development affects the production decisions of the firm.[1] While this channel is clearly important, it does not exhaust all the available channels for financial development to influence real activity. In this paper, I show that financial development is an important determinant of any firm's sourcing decisions.

To get a sense of how these two concepts are related, Figure 1.1 provides some suggestive evidence at the country level. It plots the share of intermediate purchases in gross output against an index of financial development, the ratio of private credit to GDP.[2] The diagram clearly shows that firms in countries with higher levels of financial development tend to purchase more products and services from outside suppliers (outsourcing) rather than produce them in-house (vertical integration).

Outsourcing has become an important business strategy, whereby productive efficiency may be gained when products and services are produced at lower cost by outside suppliers.[3] From a social perspective, further allocative efficiency may be gained when the resources conserved by outsourcing are devoted to the creation of new or more products and services, leading to higher aggregate productivity.[4] In this paper, I present a simple model to formalize these arguments, confront the model with data, and quantitatively assess the impact of financial development on aggregate productivity.

The central finding is that financial development has a large impact on outsourcing, but net productivity gains are relatively modest. Quantitative analysis shows that setting financial development to the U.S. level would increase outsourcing by more than 10 percentage points, if all countries are weighted equally. Yet despite the ubiquitous reorganization of production, aggregate productivity gains are relatively modest: only about 0.3 percent on average. Outsourcing implies a more efficient use of resources, but it also entails costs of market transactions. The transaction costs generate a powerful negative effect through prices, which offsets the majority of productivity gains. If

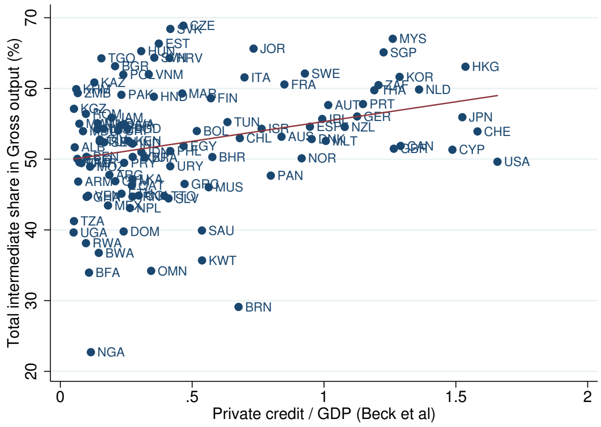
1 See, for example, B. S. Bernanke, Gertler, and Gilchrist (1999), B. Bernanke and Gertler (1989), Kiyotaki and Moore (1997), and Townsend (1979), and subsequent works by their followers.

2 The share of intermediate purchases in gross output is calculated from the cross-country sectoral input-output tables from GTAP 9 Database. The ratio of private credit to GDP is a proxy for financial development initially developed by Beck, Demirg, and Levine (2010). The slope (0.054) is precisely estimated at less than 1% level.

3 The important benefit of outsourcing is consistent with the transaction cost economics pioneered by Coase (1937) and Williamson (1971, 1979, 1985). By contrast, the benefit of vertical integration is that it reduces the costs of transactions, adaptation, and opportunism.

4 Since the seminal work of R.Krugman (1979), product variety has played an important role in theoretical models of growth (G. M. Grossman and Helpman, 1991; Romer, 1990) and trade (Melitz, 2003; Melitz and G. I. P. Ottaviano, 2008). New varieties are important sources of economic growth (Bils and Klenow, 2001) and welfare gains from trade (Broda and Weinstein, 2006).

Figure 1.1: Total intermediate purchases and financial development



there were no such negative price effects, the net productivity gains would be fivefold (about 1.5 percent on average). Hence, in evaluating the impact of outsourcing related policies, it is not sufficient to measure changes in the prevalence of outsourcing; one must take into account the effect through prices.

My contribution to the literature is three-fold. First, I highlight a channel for financial development to affect aggregate productivity that has received relatively little attention in the literature, namely endogenous changes in outsourcing. Financial development can induce a reorganization of production that enhances aggregate productivity. My theoretical work clarifies the role of financial development in shaping outsourcing, entry, price and productivity. I further show how the effects on these individual components together create the impact on aggregate productivity. My analysis also sheds light on how resources are being reallocated both within and across sectors as financial markets develop. The reallocation of resources further elevates aggregate productivity.

Second, I test the main implication of the model with both sector-level and firm-level data. My model predicts that financial development increases outsourcing, especially when suppliers are highly dependent on external finance and have fewer tangible assets. I first confront the model with cross-country sectoral input-output data. The empirical results strongly support the view that financial development disproportionately increases a sector's outsourcing from sectors characterized by high dependence on external finance and fewer tangible assets. To further inspect the mechanism, I complement the sector-level regressions with firm-level regressions. Specifically, I examine the impact of the contraction in credit supply on the intensity of outsourcing of UK firms using data from Orbis. To establish causal effect, I exploit the 2008 financial crisis as a source of exogenous variation in credit supply of banks.[5] My results confirm that a large contraction in

5 There is also anecdotal evidence that credit shortages reduce demand because firms are discouraged from applying for finance or become cautious about business prospects in an uncertain economic environment (Schans, 2012). However, overwhelming evidence suggests that there is continued tightening in the supply of credit

credit reduces outsourcing disproportionately in firms which rely on suppliers that are highly dependent on external finance or having few tangible assets. Furthermore, the results show that credit shocks affect both the intensive and extensive margins of outsourcing, consistent with the prediction of my model.

Third, I structurally estimate the model and perform counterfactuals to quantify the impact of financial development on aggregate productivity. The quantitative analysis shows that financial development has a sizable impact on outsourcing. Setting financial development to the U.S. level for all countries would increase the prevalence of outsourcing from an average of 60 percent to more than 70 percent. Consistent with the arguments set out at the beginning of the paper, outsourcing implies a more efficient use of resources, thereby elevating aggregate productivity by about 1.5 percent on average. However, outsourcing also incurs transaction costs, which generate a powerful negative effect through prices. Taken together, the net productivity gains are relatively modest (about 0.3 percent on average). The results point to product price as an important factor for aggregate productivity.

How does financial development affect the sourcing decisions of the firm? The proposed explanation is best motivated by a simple example. Suppose a firm has two options to acquire an intermediate input: either manufacturing it in-house, or outsourcing it to a specialized supplier. Outsourcing is more cost efficient, but the supplier may hold up investment when contracts are incomplete. The firm, therefore, chooses sourcing modes by balancing the cost and benefit. When there are financial frictions, additional complications may arise. Outsourcing not only purchases inputs from the supplier, but also transfers the responsibility of production and the associated technology to the supplier, in exchange for an upfront payment.[6] The supplier's lack of access to finance may hinder its ability to make upfront payments, and induce the firm to inefficiently retain production in-house.[7]

I build a multi-sector, general equilibrium, transaction cost model of outsourcing to generalize this example. My model incorporates firm heterogeneity and financial frictions into the industry equilibrium model of G. M. Grossman and Helpman (2002). To provide a close link between theory and empirics, I model financial frictions à la Manova (2013) on the supplier side.[8] Specifically, the supplier has to borrow a

since the 2008 crisis (Armstrong et al., 2013). Moreover, the presence of demand factors means that I underestimate the effect of credit supply.

6 The empirical counterpart to the upfront payment is royalties and licences. Royalties and licences is the second most important category in UK trade in services, accounting for 22.6% of service imports and 26.1% of service exports (Breinlich and Criscuolo, 2011).

7 Having the downstream firm supply the technology on credit does not change the results. The firm, acting the same as banks, would require its lending break even. The only difference may be that the firm can monitor its supplier more efficiently and recover more from collaterals in the event of supplier default. In my model, this means a change of values in two parameters, monitoring efficiency and asset tangibility. See section 1.2.1 for more details.

8 My model differs from Manova's in three important ways. First, I focus on financial frictions on the supplier rather than the firm. Second, in Manova's model, financial frictions have a limited impact on the intensive margin of exports, as productive

fraction of the upfront payment (external-finance dependence) against its collateral. The supplier repays when the bank monitors it effectively (financial development), and defaults otherwise. In the event of default, the bank can recover a fraction of the collateral value (asset tangibility). My model predicts that financial development increases outsourcing, especially when the supplier is highly dependent on external finance and has fewer tangible assets.

I further analyse the role of financial development in shaping sectoral entry, price, and productivity. My model predicts that financial development fosters new firm formation, reduces sector price, and enhances sector productivity.[9] Furthermore, financial development improves allocative efficiency across sectors by reallocating resources to sectors that are highly dependent on external finance or have few tangible assets.[10] Hence, financial development elevates aggregate productivity by improving allocative efficiency both within and across sectors.

Having characterized the model, I confront it with data. First, I test the model implication with cross-country sectoral input-output tables from the GTAP 9 database. My baseline specification is a differences-in-differences regression with fixed effects, which focuses on the interaction of financial development and the financial characteristics (external-finance dependence and asset tangibility) of upstream (supplier) sectors. Following Rajan and Zingales (1998), I use the ratio of private credit to GDP as a proxy for financial development, and calculate external-finance dependence and asset tangibility from the U.S. Compustat database. The results strongly support the prediction of the model. The baseline results are robust with regard to (1) excluding intermediates purchased from within the sector, (2) restricting the sample to domestic intermediates, (3) restricting the sample to service intermediate inputs, (4) exploiting variations of intermediate purchases over time, and (5) using credit information index as an alternative measure of financial development.[11]

---

exporters are unaffected by credit constraints. My model emphasizes both intensive and extensive margin of outsourcing. Third, Manova analysed the effect of financial frictions on exports in a partial equilibrium setting. I characterize how financial development induces the reorganization of production both within and across sectors in general equilibrium.

9 More precisely, my model predicts that financial development always increases the mass of entrants. When adjustment at the extensive margin dominates the intensive margin, financial development reduces sector price and elevates sector productivity unambiguously.

10 One way to understand why resources are reallocated across sectors is to compare my model with a competitive model. In a competitive model with Cobb-Douglas preference, a change in sector productivity does not induce resource reallocation across sectors, because the effect of higher productivity is offset by the effect of lower product price. For this to be true, revenue productivity must be the same across sectors. In my model, financial development disproportionately reduces revenue productivity in financial vulnerable sectors, inducing resources reallocation to those sectors.

11 The credit information index is from the World Bank. This index measures rules affecting the scope, accessibility, and quality of credit information available through public and private credit registries. Higher values indicate the availability of more credit information to facilitate lending decisions.

Second, I confront the model with firm-level data from the Orbis database. My baseline regression is based on the differences-in-differences framework. I test the hypothesis that after the crisis, when credit was scarcer, firms that source from more external-finance dependent industries outsource less, while firms that source from industries with more tangible assets outsource more. To measure the financial characteristics of suppliers, I link the primary industry of the firm to industries in the U.S. input-output table, and calculate the average financial characteristics of firms' upstream industries. My results strongly support the view that a large contraction in credit reduces outsourcing disproportionately in firms which rely on suppliers that are highly dependent on external finance or having few tangible assets. The baseline results are robust with regard to (1) including additional controls, and (2) restricting the sample to the balanced panel.

Having empirically tested the model, I analyse the quantitative importance of the sourcing channel. I ask: How much do countries benefit from further development of their financial markets to the U.S. level? To provide an answer, I first estimate the structural parameters of the model. My estimation strategy proceeds in two stages, with each stage featuring a non-linear Poisson pseudo-maximum-likelihood (PML) estimator with high-dimensional fixed effects.[12] I then use the model to perform counterfactuals. The experiment shows that financial development has a sizable impact on outsourcing, but its effect on aggregate productivity is relatively modest. The majority of productivity gain comes from the reallocation of resources within sectors. Financial development also induces resource reallocation across sectors, but with a limited impact on aggregate productivity.

1.1.1 *Related Literature*

There is a large body of theoretical literature that studies the boundaries of the firm. Many theories of the firm suggest that integration, while costly, reduces transaction costs and enhances profitability.[13] Consistent with this view, I build firm boundaries on the transaction cost theory of the firm pioneered by Coase (1937) and Williamson (1985). However, the novel channel for financial development to improve allocative efficiency applies well to the property rights approach of the firm (Antràs, 2003; S. J. Grossman and O. D. Hart, 1986; O. D. Hart, 1995; O. Hart and Moore, 1990).[14] I embed this channel in a multi-sector general equilibrium model with heterogeneous firms, and fully characterize the impact of financial development on outsourcing and aggregate productivity.

My paper is closely related to a large body of empirical literature on the determinant of vertical integration. The propensity for firms to

12 Poisson PML estimators are particularly suitable for non-linear conditional means with multiplicative forms (see Silva and Tenreyro, 2006).

13 See Gibbons (2005) for a survey.

14 See Appendix A.2.2 for an alternative exposition based on the property rights approach. It is worth noting that the two approaches have distinct predictions regarding differences in contracting institutions.

integrate varies systematically with technology intensity (Acemoglu, Aghion, et al., 2010), factor intensity (Antràs, 2003), product market competition (Aghion, Griffith, and Howitt, 2006; Bloom, Sadun, and Reenen, 2010), product price (Alfaro, Conconi, et al., 2016), and institutions (Acemoglu, S. Johnson, and Mitton, 2009; Boehm, 2015; Macchiavello, 2012).[15] I contribute to this literature by highlighting a novel channel for financial development to affect a firm's vertical integration decisions, namely supplier access to finance. Empirical results strongly support the view that supplier access to finance is a key determinant of vertical integration.

There is also a large body of empirical literature that studies the role of financial market imperfections in economic development. Early contributions are by Demirgüç-Kunt and Maksimovic (1998), Jayaratne and Strahan (1996), King and Levine (1993), and Rajan and Zingales (1998) and Braun (2003). See Banerjee and Duflo (2005) and Levine (2005) for recent surveys. I contribute to this literature by proposing a new channel for financial development to affect economic development, namely firms' outsourcing decisions.

My paper also adds to the growing body of work on financial market imperfections on trade. Previous work has shown that credit constraints distort trade flows by impeding a firm's export operations (Carluccio and Fally, 2012; Chan and Manova, 2015; Chor and Manova, 2012; Feenstra, Li, and Yu, 2014; Manova, 2013). More recently, scholars have explored exogenous shocks to a firm's access to finance to establish a causal effect of credit constraints on trade (Amiti and Weinstein, 2011; Bricongne et al., 2012; Manova, Wei, and Zhang, 2015; Paravisini et al., 2015). My contribution to this literature is using a novel source of identification: credit contraction during the 2008 financial crisis combined with the variation in financial characteristics of input suppliers.

My paper is also related to the branch of literature that studies the impact of resource misallocation on aggregate productivity. Misallocation of resources can occur both across sectors (Jones, 2011, 2013) and within sectors (Banerjee and Duflo, 2005; Foster, Haltiwanger, and Syverson, 2008), potentially reducing aggregate productivity (Hsieh and Klenow, 2009; Restuccia and Rogerson, 2008). While various frictions can generate misallocation, the most relevant source of misallocation to my study is financial frictions (Caselli and Gennaioli, 2013; Midrigan and Xu, 2014; Moll, 2014).[16] I contribute to this literature by highlighting a novel channel for financial frictions to affect resource misallocation both within and across sectors, namely endogenous changes in outsourcing. Empirical evidence provides strong support for this channel.

15 See Lafontaine and Slade (2007) for a survey of early contributions.

16 The potential sources of misallocation include tax and regulations (Hsieh and Klenow, 2009; Restuccia and Rogerson, 2008), capital adjustment costs (Asker, Collard-Wexler, and Loecker, 2014), information frictions (David, H. A. Hopenhayn, and Venkateswaran, 2016), heterogeneous markups (Dhingra and Morrow, 2016; Peters, 2013), and financial frictions (as cited in the main text).

In terms of focus, my paper is closely related and complementary to Boehm (2015). Boehm studied how contract enforcement costs affect a firm's sourcing decisions and aggregate productivity. My paper differs from his in three dimensions. First, while Boehm emphasized the role of contract enforcement costs in shaping buyer-seller relationships, I focus on the role of financial development. Second, Boehm's theory builds on the Ricardian trade model, in which contract enforcement costs lower a supplier's prospect of supplying inputs, but do not affect the prices charged by the supplier.[17] In my model, the hold-up problem induces the supplier to lower the scale of production and raise the product price, offsetting most of the productivity gains. Third, while Boehm investigated the impact of institutional changes that reduce transaction costs, I focus on institutional changes that improve allocative efficiency holding fixed the transaction costs.

The rest of the paper is organized as follows. Section 1.2 presents the model and derives the main propositions of the paper. Section 1.3 tests the main predictions of the model with cross-country sectoral input-output table data. Section 1.4 further examines the proposed mechanism with firm-level data. Section 1.5 estimates the structural parameters and performs counterfactual experiments. Section 1.6 concludes. Proofs of the propositions are provided in Appendix A.1.

1.2 A MODEL OF FINANCIAL FRICTIONS AND OUTSOURCING

In this section, I develop a model of outsourcing under financial frictions. Two features characterize the model. The first feature is that investments are relationship specific. Parties are partially locked into the bilateral relationship, and are likely to withhold investments. Firms choose to outsource or internalize production to minimize the transaction costs the hold-up problem generates. The second feature is that financial markets are imperfect. Financial frictions interfere with firms' choices of ownership structures, which reduce the efficiency at which production is carried out. I first introduce the basic environment, and then explore various aggregate implications of the model.

1.2.1 *Basic Environment*

Consider an economy with L consumers, each supplying one unit of labor. Preferences are defined over final goods from N sectors, $U = \prod_{n=1}^{N} Y_n^{\theta_n}$, where $\sum_{n=1}^{N} \theta_n = 1$. The final good Y_n is produced by combining intermediates Y_{ni} with a Cobb-Douglas production technology:

$$Y_n = \prod_{i=1}^{N} Y_{ni}^{\gamma_{ni}}, \text{ where } \sum_{i=1}^{N} \gamma_{ni} = 1.$$

17 Put differently, external suppliers (outsourcing) overcome their cost disadvantage by selling a smaller range of inputs, exactly to the point at which the distribution of prices for what they sell to the firm is the same as the distribution of prices offered by internal suppliers (integration).

Here, Y_{ni} represents the intermediate goods from upstream sector i. The intermediate Y_{ni} in turn is a CES aggregate of a continuum of differentiated products,[18]

$$Y_{ni} = \left(\int_0^1 y_{ni}(j)^{\alpha} \, dj \right)^{1/\alpha}.$$

Inputs are imperfect substitutes, with an elasticity of substitution $1/(1-\alpha)$. Firms supplying differentiated inputs therefore face a demand $y_{ni}(j) = A_{ni} p_{ni}(j)^{-1/(1-\alpha)}$, where $A_{ni} = P_{ni}^{\alpha/(1-\alpha)} \theta_n \gamma_{ni} PY$ is an aggregate demand shifter. Here, $Y = \prod_{n,i=1}^{N} Y_{ni}^{\theta_n \gamma_{ni}}$ refers to the aggregate output and $P = \prod_{n,i=1}^{N} (P_{ni}/\theta_n \gamma_{ni})^{\theta_n \gamma_{ni}}$ represents the ideal aggregate price index. In what follows, whenever there is no ambiguity, I shall focus on a particular product and omit the index n, i, and j.

Firm F can transform input x into product y with productivity z, $y = zx$. The firm can build the input in-house with constant marginal (labor) cost $c > 1$, or buy it from a specialised supplier S, who can produce it at constant unity marginal cost. The former organizational form is vertical integration and the latter is outsourcing. Setting up an integrating firm also requires a fixed (labor) cost f^V, which is assumed to be higher than the fixed cost for an outsourcing firm, f. Assume firms can freely enter the market by paying an entry (labor) cost f^e. After the entry cost is sunk, firms draw their productivities from a common distribution $G(z)$.

An integrating firm chooses input x to maximize its profit,

$$\pi_V(z) = \max_x A^{1-\alpha} z^{\alpha} x^{\alpha} - cx - f_V.$$

This delivers a profit function $\pi_V(z) = \psi_V z^{\frac{\alpha}{1-\alpha}} - f_V$, where $\psi_V = (1-\alpha) A (c/\alpha)^{-\frac{\alpha}{1-\alpha}}$. An outsourcing firm F can be paired with a specialised supplier S who builds the input. The firm supplies technological information z to the supplier, in exchange for a lump-sum upfront payment $T(z)$. Assume *ex-ante* the firm faces a perfect elastic supply of suppliers. It would demand the upfront payment to ensure the supplier participates at minimum cost. Once the relationship is formed, however, the supplier has an incentive to renegotiate the division of sales revenue. Assume the two parties engage in Nash bargaining. The firm, with a bargaining power $0 < \beta < 1$, obtains a share β of the revenue; the supplier obtains the rest. The supplier, anticipating only a share $1-\beta$ of the sales revenue, would withhold the investment in input x. This can be clearly seen from the supplier's problem:

$$\pi_S(z) = \max_x (1-\beta) A^{1-\alpha} z^{\alpha} x^{\alpha} - x - T(z).$$

---

18 An example of this is the production of computers. One can think of Y_n as the final good computers, Y_{ni} as the components of a computer, power supply, motherboard, microprocessor, memory, storage devices etc. Focus on one of the components, say, the microprocessor. The product $y_{ni}(j)$ can be thought of as semiconductors, gold, copper, aluminum, silicon and various plastics. Alternative, one can view $y_{ni}(j)$ as tasks, such as exposure, washing, etching, planting, wiring, slicing, packaging etc.

Under-investment occurs as the marginal product of input exceeds the unity marginal cost.

The under-investment problem can be mitigated by allowing the firm to choose the amount of transfer, and to switch to integration altogether. The only additional complication is that financial markets are imperfect; therefore, the suppliers may not be able to borrow enough funds to make the transfer. Specifically, assume that the supplier must borrow a fraction $0 \leq \delta \leq 1$ of the upfront payment from an external financier, and can finance the rest by internal funds. The supplier, however, cannot promise to repay with certainty. With probability $0 \leq \lambda \leq 1$ the supplier repays F in full, otherwise it defaults. In the event of default, the financier can seize a fraction $0 \leq \mu \leq 1$ of the collateral. The supplier would need to replenish the collateral in order to continue production in future. To maintain the tractability of the model, assume the amount of collateral equals the entry cost f^e.[19] With financial frictions, the supplier's problem becomes:

$$
\begin{aligned}
\pi_S(z) = \max_{x, F(z)} & \ (1-\beta) A^{1-\alpha} z^\alpha x^\alpha - x - (1-\delta) T(z) - \lambda F(z) - (1-\lambda) \mu f^e \\
s.t. \quad & (1-\beta) A^{1-\alpha} z^\alpha x^\alpha - x - (1-\delta) T(z) \geq F(z), \\
& \lambda F(z) + (1-\lambda) \mu f^e \geq dT(z).
\end{aligned}
$$

The first constraint states that the amount of repayment $F(z)$ must be feasible. The second requires that external financiers break even in expectation. The most important parameters in this model are the probability of repayment (financial development) λ, and supplier's external-finance dependence δ and collateralizability of assets (asset tangibility) μ.

An implication of financial friction is that it imposes an upper bound on the amount of upfront payment. In addition to ensuring the supplier would participate at minimum cost, the amount of upfront payment must also ensure the supplier can borrow funds from an external financier. Hence the firm faces an additional credit constraint:

$$
T(z) \leq \frac{(1-\alpha)(1-\beta) A \left(\frac{1}{\alpha(1-\beta)}\right)^{-\frac{\alpha}{1-\alpha}} z^{\frac{\alpha}{1-\alpha}}}{1+\delta\left(\frac{1-\lambda}{\lambda}\right)} + \frac{\left(\frac{1-\lambda}{\lambda}\right)\mu f^e}{1+\delta\left(\frac{1-\lambda}{\lambda}\right)}
$$

If there were no financial frictions ($\lambda = 1$), the firm would demand an amount of transfer equal to the supplier's *ex-post* operating profit, which is the numerator in the first term of the credit upper bound. When there are financial frictions ($\lambda < 1$), external finance is more costly to obtain than internal funds. The price of internal funds is unity, while the price of external funds is $1/\lambda$. Hence the average price of funds is $1 + \delta\left(\frac{1-\lambda}{\lambda}\right)$, which is the denominator of the credit upper bound. Financial frictions require lowering the amount of upfront payment to respect the price of funds. Nonetheless, financial frictions are not all bad news for the firm. The fact that the supplier needs to replenish the collateral after default effectively lowers its outside option below zero, allowing the

19 The settings of financial frictions are similar to those in Manova (2013).

firm to extract more rents from the relationship. This is reflected in the second term of the credit upper bound. When the credit constraint binds, the firm's profit function is $\pi_O(z) = \psi_O z^{\frac{\alpha}{1-\alpha}} - f_O$, where $\psi_O = \left(\beta + \frac{(1-\alpha)(1-\beta)}{1+\delta\left(\frac{1-\lambda}{\lambda}\right)}\right) A \left(\frac{1}{\alpha(1-\beta)}\right)^{-\alpha/(1-\alpha)}$, and $f_O = f - \frac{\left(\frac{1-\lambda}{\lambda}\right)\mu f^e}{1+\delta\left(\frac{1-\lambda}{\lambda}\right)}$.

The assumptions of the model can now be stated. The first assumption is motivated by two empirical facts. First, outsourcing and vertical integration coexist even in the most financially developed country, the United States.[20] Second, vertically integrated firms tend to be larger and more productive (Atalay, Hortaçsu, and Syverson, 2014). To account for these facts, I assume that integration is more profitable than outsourcing at the expense of higher fixed cost.

Assumption 1.1. *Integration is more profitable than outsourcing at the expense of higher overhead costs. That is,* $\frac{f_V}{f} > \frac{1-\alpha}{1-\alpha(1-\beta)} \left(c\left(1-\beta\right)\right)^{-\frac{\alpha}{1-\alpha}} > 1$.

The second assumption concerns the credit constraints. Depending on parameter values, the participation constraint can be more binding than the credit constraint. Since the purpose of the paper is to study the effect of financial development, I focus on a more interesting case in which the credit constraint is more stringent than the participation constraint.

Assumption 1.2. *Credit constraint is binding for all producing firms,* $\frac{f^e}{f} < \frac{\delta}{\mu}\frac{(1-\alpha)(1-\beta)}{1-\alpha(1-\beta)}$.

Intuitively, this condition says that in the worst case scenario, in which the level of financial development is zero ($\lambda = 0$), the supplier's share of the revenue $\frac{(1-\alpha)(1-\beta)}{1-\alpha(1-\beta)} f$, exceeds the amount of upfront payment to the firm $\frac{\mu}{\delta} f^e$. The second assumption ensures that all suppliers, given the chance, would participate in production.

The last assumption bounds firm's bargaining power from below. It requires that the firm's share of revenue exceeds that of the supplier under outsourcing.

Assumption 1.3. *The firm's bargaining power is sufficiently high,* $\beta > (1-\alpha)(1-\beta)$.

Bargaining power governs the outsourcing firm's trade-off between under-investment and rent extraction. A high bargaining power exacerbates the supplier's hold-up problem, by lowering its perceived revenue. This is clearly seen from the marginal product of input, $1/(1-\beta)$, which exceeds the unity marginal cost. A low bargaining power aggravates the rent extraction problem by placing the firm in a weak position. In the worst case scenario, in which the level of financial development is zero ($\lambda = 0$), the firm is left with a fraction β of the revenue. The last assumption ensures that financial frictions do not dominate contracting frictions in all circumstances. That is, the marginal product of input (hold-up problem) exceeds the relative

20 Data on value added reveal that, in the United States, transactions that occur in the firm are roughly equal in value to those that occur in markets. See also figure 1.1.

revenue shares between integration and outsourcing (rent extraction), $\frac{1}{1-\beta} > \frac{1-\alpha}{\beta}$. The last assumption therefore maintains the most essential feature of the transaction cost model at all times. This completes the description of the model.

1.2.2 *Aggregate Implications*

It is now possible to determine the equilibrium of the model. I proceed as follows. First, I derive an expression for the prevalence of different organizational forms. I then combine the free entry condition with labor market equilibrium to solve for the mass of entrants. Finally, I combine the aggregate demand for labor in each sector with the allocation of aggregate expenditure across sectors to determine the allocation of labor across sectors. To simplify the analysis, I use a specific parameterization of the productivity distribution. I assume that the productivity z follows a Pareto distribution with lower bound b and shape parameter $\theta \geq 1$,

$$G(z) = 1 - \left(\frac{b}{z}\right)^{\theta}, \text{ where } z \in [b, \infty).$$

How does the prevalence of different organizational forms respond to changes in financial developments? First, I need to define what I mean by the "prevalence of organizational forms". I use the fraction of firms that choose a specific organizational form as the measure of prevalence. By assumption 1.1, high productivity firms vertically integrate while low productivity firms outsource. It follows that the prevalence of integrating firms is $\sigma = (1 - G(z_V)) / (1 - G(z_O))$, where z_O is the cutoff productivity for the marginal entrant, and $z_V > z_O$ is the cutoff productivity for the marginal organizational switcher. Note that the prevalence measure is independent of the mass of entrants. Under Pareto distribution, the prevalence of vertical integration admits a close-form solution:[21]

$$\sigma = \left[\frac{\psi_V - \psi_O}{\psi_O} \frac{f_O}{f_V - f_O}\right]^{\theta(1-\alpha)/\alpha}. \tag{1.1}$$

After totally differentiating this equation, I derive an expression for the percentage change in the prevalence of integration,

$$\hat{\sigma} = -\vartheta_1 \hat{\lambda} + \vartheta_2 \hat{\delta} - \vartheta_3 \hat{\mu},$$

where the ϑ's collect the terms that multiply $d\lambda/\lambda$, $d\delta/\delta$ and $d\mu/\mu$ respectively. Hence an improvement in financial markets would induce integrating firms to switch to outsourcing. Intuitively, better financial institutions lower the cost of external finance, making the organizational form that relies on external finance more attractive. The following proposition summarizes the results.

21 An alternative measure of prevalence is the fraction of sales captured by each organizational form, $\sigma_R = \left[\frac{\psi_V - \psi_O}{\psi_O} \frac{f_O}{f_V - f_O}\right]^{\theta(1-\alpha)/\alpha - 1}$. All results concerning σ go through if swiching to σ_R instead.

Proposition 1.1. *The prevalence of vertical integration σ decreases in financial development λ, increases in external financial dependence δ, and decreases in asset tangibility μ. Furthermore, the prevalence of vertical integration σ is log-supermodular in λ and μ, $\frac{\partial^2 \ln \sigma}{\partial \mu \partial \lambda} > 0$, and is log-submodular in λ and d, $\frac{\partial^2 \ln \sigma}{\partial \delta \partial \lambda} < 0$.*

The proposition also summarizes the interaction effects between financial development and the financial characteristics of the supplier. In particular, it states that financial development increases outsourcing, especially when supplier is highly dependent on external finance and has fewer tangible assets. The intuition is straightforward. Financial development disproportionately benefits those suppliers that are highly dependent on external finance or have fewer tangible assets to serve as collateral. These testable implications will be closely examined in subsequent sections.

Next, I solve for the mass of entrants M^e. First, using the zero cutoff condition for marginal entrants and the free entry condition, I derive an expression for the equilibrium cutoff productivity z_O:

$$\frac{f^e}{f_O} = \left(\frac{b}{z_O}\right)^{\theta} \left[\left(\frac{\overline{z}}{z_O}\right)^{\frac{\alpha}{1-\alpha}} - \frac{\overline{f}}{f_O}\right],$$

where $\overline{z}$ and $\overline{f}$ are proportional to the average profit and fixed cost for of producing firms.[22] Further combining with labor market equilibrium gives an expression for the mass of entrants:

$$M^e = \frac{L}{f^e\left(\frac{\alpha(1-\beta)}{\phi_O}\Omega_p/\overline{\Omega}+1\right)\left(1+\frac{\overline{f}}{\overline{\Omega}f_O-\overline{f}}\right)}. \tag{1.2}$$

Here, Ω_p and $\overline{\Omega}$ are two multiplying factors summarizing the composition effect of different organizations on average labor demand and profit.[23] By assumption, integrating firms are more profitable; thus, an increase in the number of integrating firms would raise average profit. Nonetheless, integrating firms are less efficient in production cost; therefore, the composition effect on average labor demand is greater than that on average profit, $\Omega_p > \overline{\Omega}$. Together with the composition effect on fixed cost, $\overline{f}$, these multiplying factors determine the effect of financial development on the number of entrants in equilibrium. The following proposition summarizes the effects.

Proposition 1.2. *The mass of entrants M^e increases in financial development λ, decreases in external financial dependence δ, and increases in asset tangibility μ.*

22 The expression for average profit fixed cost is $\overline{f} = (1-\sigma) f_O + \sigma f_V$, and for average profit productivity is $\overline{z} = \left(V(z_O) + \left(\frac{\phi_V}{\phi_O}(c(1-\beta))^{-\frac{\alpha}{1-\alpha}} - 1\right)\sigma V(z_V)\right)^{\frac{1-\alpha}{\alpha}}$, where $\phi_V = 1-\alpha$, $\phi_O = \beta + \frac{(1-\alpha)(1-\beta)}{1+\delta\left(\frac{1-\lambda}{\lambda}\right)}$, and $V(y) = \int_y^{\infty} z^{\frac{\alpha}{1-\alpha}} \frac{g(z)}{1-G(z)} dz$.

23 The expression for the two multiplying factors are $\overline{\Omega} = \frac{\theta}{\theta - \frac{\alpha}{1-\alpha}}\left[1 + \left(\frac{\phi_V}{\phi_O}(c(1-\beta))^{-\frac{\alpha}{1-\alpha}} - 1\right)\sigma_R\right]$, and $\Omega_p = \frac{\theta}{\theta-\frac{\alpha}{1-\alpha}}\left[1+\left(\frac{\alpha}{\alpha(1-\beta)}(c(1-\beta))^{-\frac{\alpha}{1-\alpha}} - 1\right)\sigma_R\right]$.

Intuitively, integrating firms generate more profits at the expense of higher production costs. Financial development induces fewer integrating firms, freeing up more labor resources for the creation of new firms. Across sectors, those sectors relying more on external finance attract smaller cohorts of entrants. The reason is that external finance is costlier to obtain than internal funds. Sectors more reliant on external finance divert more labor resources away from firm creations. Conversely, sectors with ample tangible assets tend to attract a large cohorts of entrants. Tangible assets, by serving as collateral, could reduce the cost of external finance. This, in turn, allows more resources to be devoted to the creation of firms.

Having solved for the mass of entrants, I now derive an expression for the aggregate sector price:

$$P = \left(\frac{1}{\alpha\,(1-\beta)}\right) M^{-\frac{1-\alpha}{\alpha}} \tilde{z}^{-1}. \tag{1.3}$$

Here, $M = M^e \left(\frac{b}{z_O}\right)^{\theta}$ is the mass of producing firms, and $\tilde{z}$ is proportional to their average productivity.[24] The expression for sector price is intuitive. It is the price charged by an average firm as if it is operating under outsourcing. Unlike the mass of entrants, the effect of financial development on sector price can be ambiguous. The reason is that, while financial development induces more producing firms (extensive margin), by so doing, it also reduces the average productivity of those firms (intensive margin). When adjustment at the extensive margin dominates intensive margin, sector price unambiguously declines with financial development. The following proposition summarizes the findings.

Proposition 1.3. *When $\phi_V > \phi_O$, aggregate sector price P decreases in financial development λ, increases in external financial dependence d, and decreases in asset tangibility μ, where $\phi_V = 1-\alpha$, $\phi_O = \beta + \frac{(1-\alpha)(1-\beta)}{1+\delta\left(\frac{1-\lambda}{\lambda}\right)}$.*

The condition requires that the *ex-post* revenue share of an integrating firm is greater than that of an outsourcing firm. When this happens, the integrating firm's profitability is much greater than that of the outsourcing firm. As financial development takes place, adjustment of the firms can occur at two margins. First, marginal low productivity firms enter the market and operate under outsourcing. Second, marginal integrating firms switch to outsourcing as it becomes more profitable. When integration is much more profitable than outsourcing, adjustment along the second margin may be limited. Most adjustment occurs at the first margin. This means the quantity effect (the mass of producing firms) on sector price dominates the composition effect (average productivity of the firms). Hence sector price unambiguously falls with financial development. It should be noted that this condition is a sufficient condition. In practice, the quantity effect is likely to dominate the composition effect, as shown in later sections.

24 The expression for average productivity is $\tilde{z} = \left(V(z_O) + \left((c\,(1-\beta))^{-\frac{\alpha}{1-\alpha}} - 1\right)\sigma V(z_V)\right)^{\frac{1-\alpha}{\alpha}}$.

Now an expression for sector TFP can be derived. First, I aggregate the revenue and labor demand from individual firms, and obtain an expression for sector revenue TFP:

$$TFPR = \frac{R}{L} = \frac{\tilde{\Omega}}{\alpha\left(1-\beta\right)\Omega_p + \phi_O \overline{\Omega}}. \tag{1.4}$$

where $\tilde{\Omega}$ is a multiplying factor that captures the composition effect of different organizations on average productivity.[25] Note that if there were no financial frictions and there was only one type of organizational form, ϕ_O would become $1-\alpha\left(1-\beta\right)$, all of these Ω's, and hence TFPR, would equal to one. The direct effect of financial frictions can be seen by setting all Ω's to unity. TFPR is above one because ϕ_O falls below $1-\alpha\left(1-\beta\right)$. This implies that the sector as a whole is underinvesting in labor resources. The indirect effect can be seen by setting ϕ_O in the denominator to $1-\alpha\left(1-\beta\right)$, while leaving Ω's unchanged. It is easy to verify that TFPR is once again above unity, implying under-investment for the sector as a whole. Furthermore, both effects increase in the severity of financial frictions.

Armed with the TFPR expression, I now obtain an expression for sector physical productivity:

$$TFP = \frac{Y}{L} = \frac{\alpha\left(1-\beta\right)\tilde{\Omega}}{\alpha\left(1-\beta\right)\Omega_p + \phi_O \overline{\Omega}} M^{\frac{1-\alpha}{\alpha}} \tilde{z}. \tag{1.5}$$

Of particular interest is how financial development affects sector TFP. The following proposition summarizes the effects.

Proposition 1.4. *When $\phi_V > \phi_O$, aggregate sector TFP increases in financial development λ, decreases in external financial dependence d, and increases in asset tangibility μ.*

Under the sufficient condition, sector TFPR rises with financial development while sector price falls with it. Hence sector TFP, defined as the ratio of sector TFPR to sector price, unambiguously rises with financial development. Why does sector TFPR comove positively with financial development? The reason is that financial development reduces all three multiplying factors representing the composition effects on productivity ($\tilde{\Omega}$), labor demand (Ω_p), and profit ($\overline{\Omega}$). Nonetheless, the composition effects on labor demand and profit are more responsive than the composition effect on productivity. Hence the overall effect of financial development on TFPR is positive. This, together with the negative effect on sector price, implies that financial development enhances sector TFP.

I have discussed various aggregate implications at the sector (pair) level. To close the model, it remains to determine the allocation of labor across sectors. I now resume the indices for downstream sector n and upstream sector i. I combine the aggregate demand for labor in

25 The expression for the multiplying factor is $\tilde{\Omega} = \frac{\theta}{\theta - \frac{\alpha}{1-\alpha}} \left[1 + \left(\left(c\left(1-\beta\right)\right)^{-\frac{\alpha}{1-\alpha}} - 1\right)\sigma_R\right]$.

each sector with the allocation of aggregate expenditure across sectors to obtain an expression for labor allocation:

$$L_{ni} = L \frac{\theta_n \gamma_{ni} / TFPR_{ni}}{\sum_{m,k=1}^{N} \theta_m \gamma_{mk} / TFPR_{nk}}.$$

Note that if there were no financial frictions, the share of labor in each sector would be equal to its share of aggregate expenditure. Under financial frictions, the market allocates labor away from sectors more exposed to financial problems to sector sectors less exposed to financial problems. Since the mass of entrants is proportional to labor, financially vulnerable sectors attract fewer entrants and therefore have lower sector TFP. This is the general equilibrium effect of financial frictions.

The central question of this paper is: how does financial development affect aggregate productivity? Using the Cobb-Douglas aggregator and the allocation of labor across sectors, I derive an expression for aggregate productivity (TFP):

$$TFP = \prod_{n,i=1}^{N} \left(TFP_{ni} \frac{\theta_n \gamma_{ni} / TFPR_{ni}}{\sum_{m,k=1}^{N} \theta_m \gamma_{mk} / TFPR_{mk}} \right)^{\theta_n \gamma_{ni}}. \tag{1.6}$$

This expression makes it clear that financial development has dual effects on aggregate productivity. Not only does it directly enhance the productivity in financially vulnerable sectors, but in the process, it also allocates more labor resources to those sectors. Both effects tend to enhance aggregate productivity. Finally, notice that preferences are defined over final goods; hence, social welfare W is equal to aggregate productivity, $W = U = Y/L = TFP$.

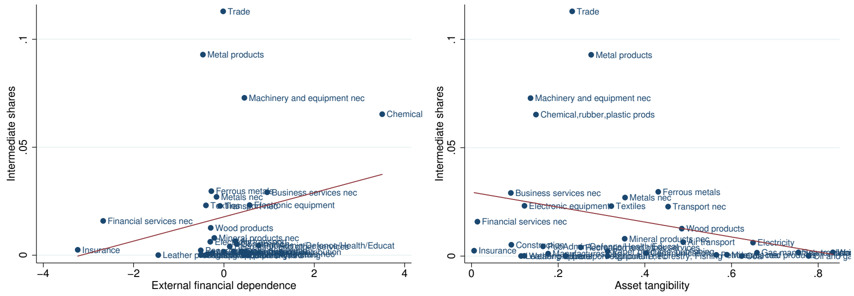
We have completed the discussion of the model. The key finding is that financial development affects organizational forms, through which it also affects aggregate productivity. How plausible is the model and its assumptions? One testable implication is given by Proposition 1.1. Namely, financial development increases outsourcing, especially when supplier is highly dependent on external finance and has fewer tangible assets. The following two sections test this implication with both sector-level and firm-level data.

1.3 SECTOR-LEVEL EVIDENCE

1.3.1 *Data*

The primary source of data is the GTAP 9 Data Base coordinated by the Center for Global Trade Analysis at Purdue University. GTAP provides a consistent global input-output table that covers 57 sectors in 129 countries and regions. The database describes bilateral trade flows, production, consumption and intermediate use of commodities and services. GTAP sectors are defined by reference to Central Product Classification (CPC) and the International Standard Industry Classification (ISIC). Using the GTAP data, I can compute the share of intermediate purchases in gross output at both the sector-pair level

Figure 1.2: External financial dependence and asset tangibility



(a) External financial dependence (b) Tangibility

and the sector level, which are used as the dependent variables in the reduced-form regressions.

To proceed with the regressions, I need a country measure of financial development and two industry measures of external financial dependence and asset tangibility. The measure of financial development is the share of domestic credit to private sector to GDP, compiled by the World Bank. The construction of industry measures of external financial dependence and tangibility follows Rajan and Zingales (1998) and Braun (2003). Each industry's external financial dependence and tangibility is calculated as the median of all U.S. based firms in the industry, as contained in Compustat's annual Fundamental Annual files for the 1997 – 2006 period. External financial dependence is defined as the investment needs that cannot be met by cash flows from operations. Cash flow from operations is broadly defined as cash flow from operations plus decreases in inventory, decreases in account receivables, and increases in account payables. Investment need is broadly defined as capital expenditure plus increases in investment and acquisitions. Tangibility is defined as the share of net property, plant, and equipment in the book value of assets. Since the industry segments are different in Compustat and GTAP, the former is mapped to the latter using a correspondence between 1987 US SIC to GTAP sectoral classification (GSC2). The summary statistics of these measures are given in Table 1.1.

Before proceeding with regressions, it is instructive to inspect the level effects of these industrial measures. Figure 1.2a plots the input expenditure share of the U.S. motor vehicle and parts industry against the external financial dependence of upstream industries. The diagram illustrates that U.S. motor vehicle producers tend to source more inputs from financially vulnerable sectors. A fixed effect regression using input expenditure share for all U.S. downstream sectors confirms that the positive correlation is statistically significant (coefficient is 0.003 and is significant at less than 1% level). By contrast, Figure 1.2b reveals that U.S. motor vehicle producers tend to source fewer inputs from those sectors endowed with more tangible assets. A similar fixed effect regression using data from U.S. input-output tables confirms that the negative correlation is statistically significant (coefficient is -0.012 and is significant at less than 10% level).

Table 1.1: Summary statistics for regression variables

	Mean	Median	S.D.	Min	Max	N
Country Characteristics						
Private credit / GDP	0.594	0.416	0.476	0.025	2.063	117
Contract enforcement cost	0.371	0.321	0.245	0.114	1.508	114
Log human capital index	0.912	0.976	0.276	0.128	1.297	114
Log capital per worker	11.166	11.372	1.353	7.659	13.023	119
Resource abundance	0.077	0.028	0.104	0.000	0.498	118
Industry Characteristics						
External financial dependence	-0.055	-0.057	1.080	-3.237	3.503	34
Asset tangibility	0.347	0.314	0.229	0.006	0.834	34
Contract intensity	0.225	0.185	0.227	0.019	1.099	34
Skill intensity	0.345	0.308	0.152	0.141	0.628	34
Capital intensity	0.143	0.103	0.150	0.010	0.860	35
Resource intensity	0.114	0.000	0.323	0.000	1.000	35
Dependent Variables						
Intermediate purchases	0.017	0.002	0.057	0.000	0.999	171,500
Outside intermediates	0.014	0.001	0.049	0.000	0.999	166,600
Domestic intermediates	0.011	0.001	0.041	0.000	0.999	171,500
Service intermediates	0.018	0.003	0.051	0.000	0.999	73,500

Notes: Private credit to GDP ratio and resource abundance are obtained from the World Development Indicators as compiled by the World Bank (Beck, Demirg, and Levine, 2010). Private credit to GDP ratio is defined as the domestic credit provided to the private sector as a share of GDP; resource abundance is defined as the total natural resources rents as a share of GDP. Contract enforcement cost is calculated using the Doing Business survey from the World Bank. It is defined as the monetary cost plus the interest foregone during the proceedings (assuming a three percent annual interest rate). Log human capital index and log capital per worker are obtained from the Penn World Table (PWT) 9.0. Each sector's financial characteristics variables are calculated as the median of all U.S. based firms in the sector, as contained in Compustat's annual Fundamental Annual files for the period from 1996 to 2005. External financial dependence is defined as the investment needs that could not be met by cash flows from operations. Cash flow from operations is broadly defined as cash flow from operations, plus decreases in inventory, decreases in account receivables, and increases in account payables. Investment need is broadly defined capital expenditures, plus increases in investment and acquisitions. Tangibility is defined as the share of net property, plant, and equipment in the book value of assets. Each sector's contract intensity is obtained from Boehm (2015), defined as the average enforcement intensity (the number of court cases between a pair of sectors) across all downstream sectors. Factor intensities (skill, capital, and resource intensity) and dependent variables are obtained from the GTAP 9 database (with reference year 2007). Factor intensity is defined as the share of factor rent in gross output. Intermediate purchase refers to the share of expenditure on intermediate inputs in gross output. Outside intermediate refers to the expenditure on inputs purchased from all upstream sectors except for the sector itself. Domestic intermediate is the expenditure on inputs sourced from domestic suppliers rather than from abroad. Service intermediate refers to the services (including the distributional service) purchased from upstream sectors. All values are evaluated at market prices.

Overall, these diagrams confirm that the input expenditure share varies systematically, and it comoves with the two industry measures in the expected directions. Since theory emphasizes the differential effects of financial development on sectors with different financial characteristics, I will now turn to the main specification of regressions.

1.3.2 *Empirical Strategy*

The main prediction is that financial development increases outsourcing, especially when supplier is highly dependent on external finance and has fewer tangible assets. This prediction can be examined with the following specification:

$$\frac{X_{nic}}{X_{nc}} = \alpha_{ni} + \alpha_{nc} + \beta_1 PC_c ExtFinDep_i + \beta_2 PC_c Tang_i + Z_{ic}\gamma + \epsilon_{nic}.$$

Here, X_{nic}/X_{nc} is the share of intermediate purchases in gross output specific to sector-pair ni in country c. PC_c is the measure of private credit to GDP in country c, $ExtFinDep_i$ and $Tang_i$ are the measures of external financial dependence and asset tangibility in upstream sector i. Z_{ic} is a vector of the additional interaction terms of country and industry characteristics. It includes the interactions of contract enforcement cost and contract intensity, skill abundance and skill intensity, capital abundance and capital intensity, resource abundance and resource intensity. The sector-pair fixed effect α_{ni} controls for the technological requirements of the usage of intermediates in the production of final goods. The downstream-country fixed effect α_{nc} captures unobservable supply and demand shocks hitting the downstream sector in a particular country. Standard errors are clustered at the country level to account for potential correlation of errors at various levels within the country.

Table 1.2 reports the results of the baseline specifications. Column (1) is the basic regression with dual fixed effects. It shows that firms tend to source more inputs from financially dependent sectors as financial development takes place. Since the regression relies on a generalized difference-in-difference approach, one can get a sense of the economic magnitude of this effect as follows. In the dataset, the country at the 25th percentile of financial development is Nigeria, and at the 75th percentile is Singapore. The industry at the 25th percentile of external financial dependence is textiles, and at the 75th percentile is coal. A firm would purchase 0.1 percentage points more inputs from coal suppliers than from textile suppliers if it moved from Nigeria to Singapore. To put the number in perspective, the average intermediate share is 1.7 percentage points. This means that the firm would increase outsourcing by about 6 percent. Meanwhile, the regression shows that the firm would purchase 0.3 percentage points fewer inputs from the 75th percentile tangible suppliers (air transport) than from the 25th percentile tangible suppliers (plastic) should it move from Nigeria to Singapore. Equivalently, the firm would increase outsourcing by about 12 percent. The pattern of coefficients is consistent with the prediction in Proposition 1.1.

Table 1.2: Baseline results: Interaction effects at sector-pair level

Dependent variable: $\ln\left(X_{nic}/X_{nc}\right)$

	(1)	(2)
$PC_c ExtFinDep_i$	0.00184***	0.00204***
	(0.000406)	(0.000391)
$PC_c Tang_i$	-0.0132***	-0.00932***
	(0.00228)	(0.00289)
Contract interaction		0.0102
		(0.00848)
Skill interaction		0.0353***
		(0.00835)
Capital interaction		-0.00315
		(0.00389)
Resource interaction		0.0192***
		(0.00604)
Downstream-country FE α_{nc}	Yes	Yes
Sector-pair FE α_{ni}	Yes	Yes
N	139230	128520
R^2	0.444	0.452

Notes: The sample includes the intermediates purchased by all downstream sectors from all upstream sectors in all countries for the year 2007. The dependent variable is the share of intermediate purchases in gross output specific to sector-pair ni in country c. PC_c is an index for financial development, referring to the private credit to GDP ratio in country c. $ExtFinDep_i$ is the intensity of external-finance dependence of upstream sector i. $Tang_i$ is a proxy for the collateralizability of upstream sector i, defined as the share of tangible assets in total assets. In column 2, contract (skill/capital/resource) interaction is the interaction between contract enforcement cost (skill/capital/resource intensity) of country c and contract intensity of upstream sector i. Robust standard errors are in parentheses, allowing for correlation at the country level. ***p<0.01, **p<0.05, and *p<0.1.

The difference-in-difference approach was designed to mitigate the omitted variable bias problem. Nonetheless, country-dependent variables measured at the upstream sector level may still be a source of omitted variable bias. To address this concern, I add a vector of covariates to the basic regression in column (2). The first covariate is the interaction of national contract enforcement cost and contract intensity of the upstream sector.[26] It controls for the comparative advantage arising from the quality of contracting institutions. That is, countries with better contracting institutions may foster input markets that are particularly reliant on contracts. Analogously, I add further interaction terms to the basic regression: skill abundance and skill intensity, capital abundance and capital intensity, resource abundance and resource intensity.[27] The summary statistics of these covariates are given in Table 1.1. After adding all these covariates, the interaction terms of interest preserve the correct signs and remain highly statistically significant. Hence, the concern about omitted variables is well-placed, but does not fundamentally change the results of the basic regression.

1.3.3 *Robustness*

Having examined the baseline results, I look at five robustness checks. First, I exclude intermediates purchased from within the same industry. Second, I change the dependent variable to shares of domestic intermediate purchases in gross outputs. Third, I restrict the sample to service intermediates. Fourth, I exploit time variation with panel regressions. Finally, I use credit information index as an alternative measure for financial development.

1.3.3.1 *Outside Intermediates*

In the baseline regressions, I include the intermediate shares for all pairs of sectors. However, it is well known that firms tend to purchase more inputs from inside the sector. Put differently, the diagonal elements tend to be greater than other elements in a typical input output table. This raises the question of whether the baseline results are affected by these diagonal elements. In this robustness check, I exclude these diagonal elements from the sample and rerun the baseline regressions. The results are then reported in Table 1.3. The coefficients of the main interaction terms preserve their correct signs and

26 Contract enforcement cost is constructed from the World Bank Doing Business survey. Following Boehm (2015), I measure enforcement cost as the sum of monetary cost (the fraction of value of the claim) and opportunity cost (interest forgone during the proceedings). The measure of contract intensity is taken from Boehm (2015). It is the enforcement intensity of an upstream sector, averaged across downstream sectors. The interaction term is not statistically significant. However, it is worth noting that the channel here is different from the channel in his paper.

27 Skill and capital abundance are construct from the Penn World Table 9.0 (PWT). Skill intensity is defined as the log of human capital intex in the PWT. Capital intensity is defined as the log of capital stock per worker. Resource abundance is the total resource rents as a share of GDP in the World Development Indicators from World Bank. Skill intensity, capital intensity, and resource intensity are constructed from the GTAP data, defined as their respective shares in the gross output.

Table 1.3: Robustness: Intermediates purchased from outside the sector

Dependent variable: $\ln\left(X_{nic}/X_{nc}\right)$

	(1)	(2)
$PC_cExtFinDep_i$	0.00167***	0.00177***
	(0.000366)	(0.000351)
PC_cTang_i	-0.0120***	-0.00939***
	(0.00227)	(0.00288)
Contract interaction		0.00842
		(0.00690)
Skill interaction		0.0341***
		(0.00855)
Capital interaction		-0.000298
		(0.00364)
Resource interaction		0.0132**
		(0.00566)
Downstream-country FE α_{nc}	Yes	Yes
Sector-pair FE α_{ni}	Yes	Yes
N	135252	124848
R^2	0.403	0.410

Notes: The sample includes the intermediates purchased by all downstream sectors from all upstream sectors except for the downstream sector itself for the year 2007. The dependent variable is the share of intermediate purchases in gross output specific to sector-pair ni in country c. PC_c is an index for financial development, referring to the private credit to GDP ratio in country c. $ExtFinDep_i$ is the intensity of external-finance dependence of upstream sector i. $Tang_i$ is a proxy for the collateralizability of upstream sector i, defined as the share of tangible assets in total assets. In column 2, contract (skill/capital/resource) interaction is the interaction between contract enforcement cost (skill/capital/resource intensity) of country c and contract intensity of upstream sector i. Robust standard errors are in parentheses, allowing for correlation at the country level. ***p<0.01, **p<0.05, and *p<0.1.

remain highly statistically significant. Furthermore, the magnitudes of these coefficients are similar to the baseline results. This reassures me that the baseline results are not driven by the inclusion of diagonal elements.

1.3.3.2 *Domestic Intermediates*

Another concern might be that, while theory focuses on domestic financial development, the baseline regressions use all intermediates regardless of whether they are purchased domestically or abroad. In light of the recent development of global production, one might worry that the baseline results are affected by foreign intermediates. Fortunately, the GTAP data report domestic and overseas purchases separately. This allows me to examine the interaction effects on domestic intermediates only. Table 1.4 reports the results. It is clear that all results in the baseline regressions go through. The magnitudes of the coefficients are also similar to those in the previous regressions.

Table 1.4: Robustness: Domestic intermediates only

Dependent variable: $\ln(X_{nic}/X_{nc})$

	(1)	(2)
$PC_c ExtFinDep_i$	0.00191***	0.00208***
	(0.000368)	(0.000353)
$PC_c Tang_i$	-0.0134***	-0.00603**
	(0.00202)	(0.00276)
Contract interaction		0.0116
		(0.00809)
Skill interaction		0.0393***
		(0.00728)
Capital interaction		-0.0108***
		(0.00358)
Resource interaction		0.0294***
		(0.00554)
Downstream-country FE α_{nc}	Yes	Yes
Sector-pair FE α_{ni}	Yes	Yes
N	139230	128520
R^2	0.327	0.336

Notes: The sample includes the domestic intermediates purchased by all downstream sectors in all countries for the year 2007. The dependent variable is the share of intermediate purchases in gross output specific to sector-pair ni in country c. PC_c is an index for financial development, referring to the private credit to GDP ratio in country c. $ExtFinDep_i$ is the intensity of external-finance dependence of upstream sector i. $Tang_i$ is a proxy for the collateralizability of upstream sector i, defined as the share of tangible assets in total assets. In column 2, contract (skill/capital/resource) interaction is the interaction between contract enforcement cost (skill/capital/resource intensity) of country c and contract intensity of upstream sector i. Robust standard errors are in parentheses, allowing for correlation at the country level. ***p<0.01, **p<0.05, and *p<0.1.

The baseline results, therefore, are robust to the exclusion of foreign intermediates.

1.3.3.3 *Service Intermediates*

Following the IO literature, I measure outsourcing as the share of intermediate purchases in gross outputs. One might worry that there could be potential mismeasurement problems. Large firms, for example, may have subsidiaries in the upstream sectors. Transactions between these firms and their suppliers would then show up as outsourcing in my measure. Could this fundamentally affect the baseline results?[28] To address this concern, I re-run the baseline regressions using intermediate purchases from service sectors only. Unlike intermediate goods, services performed within the boundaries of the firm are less likely to be priced, and therefore, they would not appear in the manufacturing surveys on which input-output tables are constructed. Table 1.5 reports the results. It is clear that all baseline results still go through when restricted to service intermediates. This reassures me that the potential mismeasurement problem does not fundamentally change the results.

1.3.3.4 *Time Variation*

One of the advantages of the GTAP 9 database is that it provides data with three reference years: 2004, 2007, and 2011. For each reference year, the input-output tables are updated and reconciled using macroeconomic data set and a variety of international data sets on trade, protection, and energy. The regional data bases are then assembled to construct the panel of globally consistent input-output tables. This short panel spans the period of the recent 2008 financial crisis, a rare event in which financial markets are interrupted on a global scale. If the proposed mechanism is at work, the variation of input-output data should pick up the effect of the credit crunch on the sourcing decisions of the firm. In this section, I exploit the time variation with the following specification:

$$\frac{X_{nict}}{X_{nct}} = \alpha_{nit} + \alpha_{nct} + \alpha_{nic} + \beta_1 PC_{ct} ExtFinDep_i + \beta_2 PC_{ct} Tang_i + \epsilon_{nict}.$$

The regression includes sector-pair-time fixed effects, α_{nit}, which control for all unobserved (time-varying) heterogeneity in technological requirements across pairs of sectors. It also includes a full set of downstream-sector-country-time dummies, α_{nct}, which control for all unobserved (time-varying) heterogeneity in the demand of intermediate inputs. Finally, it includes the sector-pair-country fixed effects, α_{nic}, which account for all unobserved (time-invariant) heterogeneity in the supply and demand of intermediate goods. I am interesting in

28 Recent researches may help alleviate this concern. Atalay, Hortaçsu, and Syverson (2014) find that roughly one-half of upstream establishments report no shipments to downstream establishments within the same firm. Ramondoa, Rappoport, and Ruhl (2016) find that the input-output coefficient linking the parent's and affiliate's industry of operation is unrelated to a corresponding intrafirm flow of goods.

Table 1.5: Robustness: Service intermediates only

Dependent variable: $\ln(X_{nic}/X_{nc})$

	(1)	(2)
$PC_c ExtFinDep_i$	0.00305^{***}	0.00362^{***}
	(0.000841)	(0.000839)
$PC_c Tang_i$	-0.0189^{***}	-0.0108^{**}
	(0.00374)	(0.00458)
Contract interaction		0.0116
		(0.00763)
Skill interaction		0.0558^{***}
		(0.0145)
Capital interaction		-0.0115^{**}
		(0.00484)
Downstream-country FE α_{nc}	Yes	Yes
Sector-pair FE α_{ni}	Yes	Yes
N	57330	52920
R^2	0.297	0.304

Notes: The sample includes the service intermediates purchased by all downstream sectors in all countries for the year 2007. The dependent variable is the share of intermediate purchases in gross output specific to sector-pair ni in country c. PC_c is an index for financial development, referring to the private credit to GDP ratio in country c. $ExtFinDep_i$ is the intensity of external-finance dependence of upstream sector i. $Tang_i$ is a proxy for the collateralizability of upstream sector i, defined as the share of tangible assets in total assets. In column 2, contract (skill/capital) interaction is the interaction between contract enforcement cost (skill/capital intensity) of country c and contract intensity of upstream sector i. Resource interaction is omitted due to lack of variation in resource intensity across upstream service sectors. Robust standard errors are in parentheses, allowing for correlation at the country level. ***p<0.01, **p<0.05, and *p<0.1.

Table 1.6: Robustness: Time Variation
Dependent variable: $\ln (X_{nict}/X_{nct})$

	(1)	(2)	(3)
$PC_{ct}ExtFinDep_i$	0.000368***	0.000405***	0.000607***
	(0.000113)	(0.000106)	(0.000211)
$PC_{ct}Tang_i$	-0.00380*	-0.00375*	-0.00593*
	(0.00223)	(0.00221)	(0.00316)
Downstream-country-time FE α_{nct}	Yes	Yes	Yes
Sector-pair-time FE α_{nit}	Yes	Yes	Yes
Sector-pair-country FE α_{nic}	Yes	Yes	Yes
N	412930	401132	170030
R^2	0.964	0.961	0.962

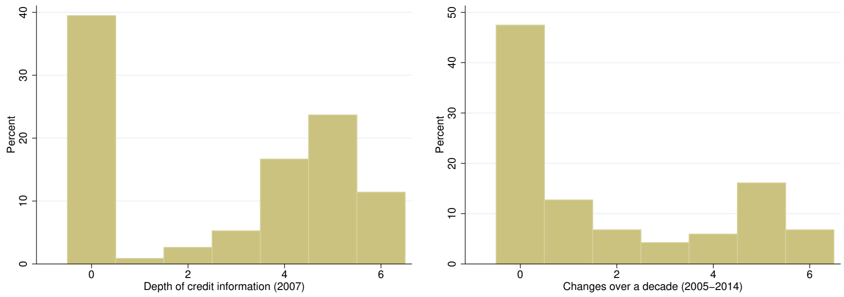
Notes: The sample includes the intermediates purchased by all downstream sectors from all upstream sectors in all countries for the year 2004, 2007, and 2011. The dependent variable is the share of intermediate purchases in gross output specific to sector-pair ni in country c. PC_c is an index for financial development, referring to the private credit to GDP ratio in country c. $ExtFinDep_i$ is the intensity of external-finance dependence of upstream sector i. $Tang_i$ is a proxy for the collateralizability of upstream sector i, defined as the share of tangible assets in total assets. Robust standard errors are in parentheses, allowing for correlation at the upstream sector and country level. ***p<0.01, **p<0.05, and *p<0.1.

the residual variation that captures sector-pair-specific supply heterogeneity over time. The interaction terms measure the differentiated response to changes in credit supply across upstream sectors. In all the regressions, I adjust the standard errors for clustering at the upstream sector and country level to account for the fact that the intermediate purchase X_{nict}/X_{nct} is constant for an upstream-sector-country-time triplet.

Table 1.6 spells out the results. Column (1) presents the baseline results based on the full sample. As in the cross-section regressions, firms tend to source more inputs from financially dependent sectors as credit become more abundant; they tend to source less inputs from sectors with more tangible assets at the same time. The difference is that, the cross-section regressions control for observable characteristic at the upstream sector and country level. One may concern that the set of covariates do not exhaust all sources of variation that may be relevant to firms' sourcing decisions. Here, the panel regressions alleviate the concern by accounting for all unobservable characteristics at the upstream sector and country level. The next two columns show that the baseline results are robust with regard to (a) excluding intermediate purchases from within the sector (column 2), and (b) restricting the sample to service intermediate inputs (column 3).[29]

29 Due to data constraints, I do not exploit the variation in domestic intermediate purchases. The reason is that the underlying domestic input-output tables are the same across years. These input-output tables are updated to incorporate information mainly from the international trade data sets. Hence, variation in domestic input purchases over time most likely reflects the adjustment processes based on trade changes.

Figure 1.3: Distribution of Depth of credit information



(a) Index values in 2007 (b) Index changes 2005-2014

The pattern of coefficients is also similar, with the coefficients for asset tangibility greater than those for external financial dependence. Moreover, restricting to service inputs almost doubles the interaction effects in the baseline results. The results reassure me that focusing on cross-section regressions is without loss of generality.

1.3.3.5 *Alternative measure of financial development*

The measure of financial development has been used extensively in the previous literature, but one may (rightly) worry whether private credit captures the aspect of financial development appropriately. The prosperity of credit markets certainly depends on the quality of other aspects of institutions, such as the quality of laws, and the effectiveness of corporate governance. As an imperfect attempt to address this concern, I use an index of the depth of credit information from the World Bank Doing Business survey as an alternative measure for financial development. This index measures the scope and accessibility of credit information available through credit reporting service providers such as credit bureaus or credit registries. It ranges from 0 to 6, with 6 indicating the greatest accessibility. Unlike private credit, the accessibility of credit information is a direct measure of the strength of financial markets. It therefore helps relieve the concern that private credit may pick up the quality of other institutions.

Figure 1.3a shows the distribution of this index in 2007 for the subset of countries in the sample. It shows that apart from a group of underdeveloped countries, the distribution is slightly skewed to the right. The depth of credit information is also statistically highly correlated with the depth of private credit (coefficient is 0.103 and is significant at less than 1% level). Meanwhile, Figure 1.3b shows the distribution of (absolute) changes in the index over the decade 2005-2014. The figure shows that the depth of credit information is relatively stable over time. About 45 percent of countries have no changes at all over a decade. Furthermore, about 70 percent of countries maintain their values within a band of 3 levels. This is more stable than the private credit which changes frequently year by year. The stability of this index means that it is likely to capture the long run development of financial markets.

Table 1.7: Robustness: Credit Information

Dependent variable: $\ln\left(X_{nic}/X_{nc}\right)$

	(1)	(2)
$HighCI_c ExtFinDep_i$	0.00245***	0.00241***
	(0.000450)	(0.000455)
$HighCI_c Tang_i$	-0.0139***	-0.0100***
	(0.00313)	(0.00378)
Contract interaction		0.0102
		(0.00850)
Skill interaction		0.0359***
		(0.00829)
Capital interaction		-0.00360
		(0.00367)
Resource interaction		0.0207***
		(0.00628)
Downstream-country FE α_{nc}	Yes	Yes
Sector-pair FE α_{ni}	Yes	Yes
N	135660	129710
R^2	0.445	0.450

Notes: The sample includes the intermediates purchased by all downstream sectors from all upstream sectors in all countries for the year 2007. The dependent variable is the share of intermediate purchases in gross output specific to sector-pair ni in country c. $HighCI_c$ is a dummy for financial development, indicating that there is a high level of credit information available in country c. $ExtFinDep_i$ is the intensity of external-finance dependence of upstream sector i. $Tang_i$ is a proxy for the collateralizability of upstream sector i, defined as the share of tangible assets in total assets. In column 2, contract (skill/capital/resource) interaction is the interaction between contract enforcement cost (skill/capital/resource intensity) of country c and contract intensity of upstream sector i. Robust standard errors are in parentheses, allowing for correlation at the country level. ***p<0.01, **p<0.05, and *p<0.1.

To exploit the variation in credit information, I divide countries into two groups by the median of credit information index: a group with high credit information and the other with low credit information. I then examine the main prediction with the following specification:

$$\frac{X_{nic}}{X_{nc}} = \alpha_{ni} + \alpha_{nc} + \beta_1 HighCI_c ExtFinDep_i + \beta_2 HighCI_c Tang_i + Z_{ic}\gamma + \epsilon_{nic}.$$

Here, $HighCI_c$ is a dummy indicating that there is a high level of credit information available in country c. As in the baseline regressions, α_{ni} is a sector-pair fixed effect, α_{nc} is a downstream-country fixed effect, and Z_{ic} is a vector of additional interaction terms of country and industry characteristics. Standard errors are clustered at the country level to account for potential correlation of errors at various levels within the country.

Table 1.7 reports the results. The results confirm that more credit information encourages more outsourcing, especially when the supplier sector is highly dependent on external finance or have fewer tangible assets. Moreover, the magnitudes of these coefficients are greater than their counterparts in the baseline regressions. This is consistent with the view that private credit is an imperfect proxy for financial development, and the associated measurement errors lead to attenuation bias in the baseline regressions. Nonetheless, the consistent results reassure me that private credit, while imperfect, is a reasonably good proxy for financial development.

In summary, I have showed that the prediction of the model is consistent with the sector-level data. Specifically, financial development has differential effects on outsourcing depending on the characteristics of upstream sectors. When upstream suppliers are more dependent on external finance, financial development induces more firms to engage in outsourcing. Conversely, when suppliers have more collateralizable assets, financial development induces more firms to switch to integration. These patterns are consistent with the reduced-form evidence.

1.4 FIRM-LEVEL EVIDENCE

The main testable implication of the model is that financial development increases outsourcing, especially when the supplier is highly dependent on external finance and has fewer tangible assets. The interaction effects have been confirmed by cross-country sectoral input-output data. In this section, I further provide firm-level evidence on the impact of credit supply shocks on firms' sourcing decisions in the United Kingdom. My aim is to examine the impact of the contraction in credit supply following the 2008 financial crisis on the intensity of outsourcing of UK firms. For this purpose, I employ the difference-in-difference approach and exploit the financial crisis as a source of exogenous variation in credit supply of banks.

1.4.1 *Data*

The core firm-level dataset is the Orbis database provided by Bureau van Dijk Electronic Publishing (BvD). The Orbis database provides information on firms' financial and productive activities from balance sheets and income statements for over 200 million companies across more than 100 countries and regions. One of the advantages of Orbis is therefore the inclusion of a wide set of countries at different levels of development. Nonetheless, the purpose here is not to conduct a full cross-country firm-level study, but to assess the validity of the novel channel proposed in this paper. For this purpose, I focus on UK firms for the period of 2007-2011. I exploit the recent financial crisis as a source of exogenous credit variation, and examine how firms' input sourcing decisions respond to changes in credit supply.

The unit of observation in Orbis is the firm, for which Orbis reports one or more financial statements. Most of the large firms report either consolidated accounts or unconsolidated accounts, although some firms report both. To avoid double counting, I restrict the sample to firms reporting unconsolidated accounts. I first calculate a firm's value added as the sum of net income, taxation, cost of employees, depreciation, and interest paid. Following the IO literature, I then calculate the measure of outsourced inputs as one minus the share of value-added in operating revenue. Finally, I exclude (outlier) firms whose measure of outsourcing fall outside the unit interval. This leaves me with 231,965 distinct firms in the sample. The summary statistics for all variables are reported in Table 1.8.

The Orbis database provides information on the 4-digit NAICS code of the firm's primary industry. To determine which upstream industries are important for the firm, I combine the above with information from U.S. Input-Output (IO) Tables. The IO data are from the Bureau of Economic Analysis (BEA). I use the Use of Commodities by Industries after Redefinitions 2007 (Purchasers' Prices) tables. To the extent that U.S. financial markets are highly developed, the U.S. IO tables should be informative about input flows across industries driven by technological requirements. Using the U.S. IO tables also helps alleviate the concern that input-output tables are endogenously determined by financial development, as predicted in my model. The 2007 IO accounts provide a level of industry detail at the 6-digit IO industry codes, with a total of 389 IO industries. The BEA provides a concordance from IO industries to 2007 NAICS codes. Readers familiar with these tables will be aware that the concordance is not one-to-one. However, by aggregating the IO industries to 4-digit level, I was able to develop a many-to-one concordance from 4-digit NAICS codes to 4-digit IO codes.[30] This concordance maps 305 NAICS into 199 IO industries. It will be used in subsequent analysis.

To measure the financial characteristics of the industries, I use the Compustat database as explained in the main text. Besides the 4-digit SIC codes, Compustat also provides information on the 4-digit NAICS code of the firm's primary industry, which will be used in the current analysis. As before, each industry's external financial dependence and tangibility is calculated as the median of all U.S. based firms in the industry, as contained in Compustat's annual Fundamental Annual files for the period from 1997 – 2006. External financial dependence is defined as the investment needs that could not be met by cash flows from operations. Tangibility is defined as the share of net property, plant, and equipment in the book value of assets. With the data explained above, I can identify the primary NAICS industry for each firm in Orbis, along with its upstream NAICS industries, and the financial characteristics of these upstream industries. This information provides the basis for the subsequent analysis.

30 The only exception is two IO industries (23 Construction, 531 Real Estates), which cannot be disaggregated to 4-digit level. Hence I keep these two IO industries at 2-digit and 3-digit level respectively. The correspondence is available upon request.

Table 1.8: Summary statistics for firm-level regression variables

	Mean	Median	S.D.	Min	Max	N
Firm-level Variables						
Log Net income	10.240	10.246	2.450	0.000	22.800	536,005
Log Taxation	9.165	9.070	2.438	0.000	22.065	424,478
Log Costs of employees	12.333	12.502	2.741	0.000	22.335	260,087
Log Depreciation	8.815	8.469	2.771	0.000	20.950	448,784
Log Interest paid	8.775	8.815	3.331	0.000	21.090	251,041
Log Value added	11.143	10.889	2.825	0.000	22.929	625,755
Log Operating revenue	12.581	12.114	2.638	0.000	25.289	626,941
Outsourcing	0.618	0.674	0.283	0.000	1.000	626,941
Industry-level Variables						
External financial dependence	0.139	-0.074	1.345	-3.254	9.061	172
Asset tangibility	0.312	0.271	0.198	0.003	0.866	172
Ave External financial dependence	-0.321	-0.238	0.439	-2.004	1.158	178
Ave Tangibility	0.268	0.281	0.114	0.046	0.532	178
Ave Number of firms	1,894	1,909	1,444	189	5,943	890

Notes: All firm-level variables are obtained from the Orbis UK database by Bureau van Dijk (BvD). Net income, taxation, costs of employees, depreciation, interest paid, and operating revenue are reported in the profit and loss statement of the firm. Value added is defined as the sum of net income, taxation, costs of employees, depreciation, and interest paid. Outsourcing is defined as unity minus the share of value-added in operating revenue. Industry level variables are obtained from the Compustat's annual Fundamental Annual files for the period from 1997 – 2006. External financial dependence is defined as the investment needs that could not be met by cash flows from operations. Cash flow from operations is broadly defined as cash flow from operations, plus decreases in inventory, decreases in account receivables, and increases in account payables. Investment need is broadly defined capital expenditures, plus increases in investment and acquisitions. Tangibility is defined as the share of net property, plant, and equipment in the book value of assets. Average external financial dependence is defined as a simple average of the external financial dependence of the top 10 upstream industries, as ranked by input flows according to the 2007 U.S. Input-Output (IO) tables. Similarly, average tangibility and average number of suppliers are defined as simple averages of the asset tangibility and the number of suppliers of the top 10 upstream industries. All industry-level measures except for the average number of suppliers are time-invariant.

1.4.2 *Empirical Strategy*

Having explained the data, I turn now to the empirical strategy. The aim is to examine whether UK firms actively adjust sourcing decisions in response to changes in credit supply. For this purpose, I use the difference-in-difference framework and exploit the latest financial crisis as a source of exogenous variation in the credit supply of banks.

The UK is particularly suitable for the study, because it is heavily dependent on banks and it suffered a relatively large credit shock. Unlike the United States, UK firms are highly dependent on banks as a source of external finance. Bank loans account for about two-thirds of the corporate debt in the UK, as compared to about one-quarter in the US. The reason is that the vast majority of UK firms are small and medium-sized enterprises (SMEs).[31] Lending to SMEs is generally riskier as they are often young businesses and are less transparent due to low reporting requirements. Banks, therefore, have a comparative advantage over equity markets in screening and monitoring the borrowers. It is typical that only large firms have access to equity markets while SMEs primarily rely on bank financing.[32] Since the model of this paper focuses on debt financing, it provides a more appropriate account for an economy that principally relies on bank financing.

Furthermore, UK banks were hit hard following the financial crisis. Following the 2008 financial crisis, there was intense pressure on UK banks to recapitalise. The major four banks – Barclays, HSBC, Lloyds Banking Group (LBG), and the Royal Bank of Scotland (RBS) – had started to issue equity in order to improve their capital positions, which they did not need to do before and after the crisis. As a result of the financial stress, banks started to cut back on lending. The corporate lending fell accordingly, by 20 percent between 2007 and 2008, after it had been growing at an average rate of roughly 10% a year in the previous decade. The contraction of credit supply is also apparent at the aggregate level. Figure 1.4 plots the stock of private credit (as a share of GDP) in the UK during the last six decades. The stock of private credit had been growing steadily until 2008, reaching a peak in 2009, and has been declining at an average rate of about 10% a year ever since.[33] The credit supply shock should be large enough to incentivise the firms to revise their sourcing strategies.

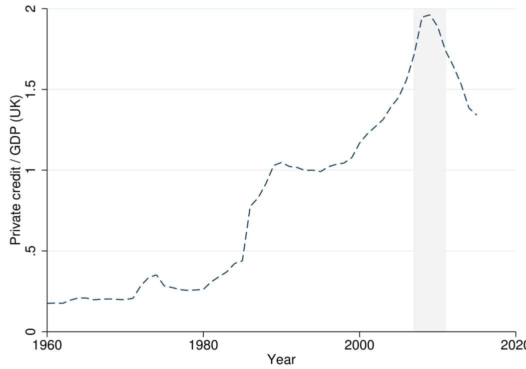
In Appendix A.2.3, I develop a variant of the baseline model, in which the firm chooses the intensity of outsourcing. One testable implication of the model is that firms adjust their sourcing decisions according to the abundance of credit supply. After the financial crisis, when credit became scarcer, firms that source from industries which are more reliant on external finance should lower the inten-

31 According to the Business Population Estimates 2013, at the start of 2013 there were 4.9 million SMEs forming 99.9 percent of all private sector firms, accounting for about 60 percent of private sector employment and nearly a half of private sector turnover.

32 According to a 2012 Department of Business Innovation and Skills (BIS) report, half of all SMEs have used financial institutions to obtain finance, while less than one percent of all SMEs have used equity finance (Schans 2012).

33 The delay of the peak is because private credit is a stock, which reflects both the repayment of old debts and the issuance of new debts.

Figure 1.4: Time series of private credit as a share of GDP for the UK



sity of outsourcing, whilst firms that source from industries which are equipped with more collateralizable assets should increase the intensity of outsourcing. I test these predictions with the following specification:

$$Outsourcing_{jnt} = \beta_1 Post_t AveExtFinDep_n + \beta_2 Post_t AveTang_n + X_{jnt}\gamma + \alpha_j + \alpha_t + \epsilon_{jnt}$$

Here, $Outsourcing_{jnt}$ is the outsourcing intensity of firm j from industry n in year t, $Post_t$ is a dummy that switches on after the 2008 financial crisis. $AveExtFinDep_n$ is a simple average of the external financial dependence of the top 10 upstream industries excluding the industry in which the firm operates. This covariate measures, on average, how much the suppliers of the firm rely on external finance. I use a simple average instead of a weighted mean to avoid this measure being cofounded by technological requirements. I further exclude the industry of the firm to ensure this measure reflects the financial characteristics of the suppliers but not those of the firm. The firm's financial characteristics are included separately in the vector of covariates instead. Analogously I construct the measure $AveTang_n$ as a simple average of the asset tangibility of the top 10 upstream industries excluding the industry of the firm. The covariate vector X_{jnt} includes the average number of suppliers in the top 10 upstream industries, and the financial characteristics of the firm. The average number of suppliers is a simple average of the numbers of suppliers in the top 10 upstream industries excluding the industry of the firm. The financial characteristics of the firm includes its external financial dependence and asset tangibility. The summary statistics for these measures are reported in Table 1.8.

As in the classic difference-in-difference framework, the baseline specification includes a firm fixed effect and a time fixed effect. The former fixed effect accounts for any time-invariant, unobservable characteristics of the firm. The latter fixed effect controls any supply and demand shocks that are common to all firms. To account for potential correlation of the shocks within the same industry, standard errors are clustered at the industry level.

Table 1.9: Baseline firm-level results

Dependent variable: $Outsourcing_{jnt}$

	(1) All firms	(2) Balanced panel
$Post_t AveExtFinDep_n$	-0.00990**	-0.0109**
	(0.00459)	(0.00454)
$Post_t AveTang_n$	0.0753**	0.0789**
	(0.0296)	(0.0307)
Firm FE α_j	Yes	Yes
Year FE α_t	Yes	Yes
N	609,019	222,545
R^2	0.861	0.849

Notes: In column 1, the sample includes all active UK firms as reported in the Orbis database for the period of 2007 – 2011. In column 2, the sample is restricted to firms who stay active in the whole period from 2007 to 2011. The dependent variable is the measure of outsourcing, as defined as unity minus the share of value-added in operating revenue. $Post_t$ is a dummy indicating the post-crisis period 2009 – 2011. $AveExtFinDep_n$ is a simple average of the external financial dependence of the top 10 upstream industries of industry n. $AveTang_n$ is a simple average of the external financial dependence of the top 10 upstream industries of industry n. Robust standard errors are in parentheses, allowing for correlation at the industry level. ***p<0.01, **p<0.05, and *p<0.1.

1.4.3 *Results*

Table 1.9 sets out the baseline results, for both the unbalanced and balanced sample. Column (1) reports the results for the full sample. The OLS estimate of the first interaction term (between post and the average external financial dependence) is negative and statistically significant. This implies that as credit became scarcer following the financial crisis, firms with more external-finance dependent suppliers engage less in outsourcing. The OLS coefficient of the second interaction term (between post and the average asset tangibility) is positive and statistically significant, telling us that firms whose suppliers are equipped with more tangible assets engage more in outsourcing when credit supply is tightened. The coefficients of interest (the two interaction terms) have the correct signs as predicted by the model, and both are estimated precisely. This reassures me that the proposed channel is indeed at work. An outstanding question is whether the results are driven by firm entry and exit, or by incumbent firms who actively adjust their sourcing strategy in response to a change in credit supply. To examine this, I construct a balanced panel which consists of the firms that report the variables of interests in all years. Column (2) reports the results for the balanced sample. The coefficients have the same signs and similar magnitudes as those for the unbalanced sample. This implies that incumbent firms have indeed actively adjusted their outsourcing intensity in response to the contraction of credit supply following the 2008 financial crisis.

Table 1.10 reports the results with additional covariates. The covariates are motivated as follows. In the model, I have made two simplifying assumptions. First, financial frictions apply only to the suppliers but not to the firms. This assumption allows me to focus on the main mechanism of the model while remaining agnostic about the effect of financial frictions on the firm. Would the contraction of credit supply directly affect firms' sourcing capability? Column (1) examines this possibility by examining the interaction term between post and the financial characteristics of the firm. The coefficients are statistically indistinguishable from zero. At first sight, this may seem surprising; why do financial frictions have no effect on firms' sourcing capability? Upon reflection, however, there are at least two reasons as to why this is the case. First, the dependent variable is the share of outsourcing instead of the level of outsourced activities. As long as the firms do not systematically drop products whose inputs are more suitable for outsourcing, there is no obvious reason to believe that the credit shortage directly affects the intensity of outsourcing. Second, previous literature has found that credit supply has an insignificant effect on outsourcing.[34] The reason, as the authors argued, is that financial development has opposite effects on outsourcing. More credit supply may allow for more suppliers to enter the market, thereby increasing the probability of the firm to engage in outsourcing. However, to the extent that mergers and acquisitions require external finance, more credit supply also makes vertical integration more possible. The insignificance of coefficients here echoes the findings of the previous literature.

Another simplifying assumption is that there is a large pool of potential suppliers. The firm can always match with one supplier should it want to do so. In reality, however, the number of suppliers may affect the outcome of outsourcing as well. All else equal, the probability of finding a suitable match should be higher when the number of potential suppliers increases, and so does the intensity of outsourcing. Column (2) explores this possibility. Since I do not observe the number of potential suppliers, I use the actual number of suppliers as a proxy for it.[35] To arrive at a metric at the firm level, I calculate the average number of suppliers in the top 10 industries of the firm. The regression then examines the impact of the average number of suppliers on the intensity of outsourcing. As expected, more suppliers are indeed associated with a higher probability of outsourcing.

Column (3) incorporates the additional covariates into the baseline regression. It shows that the interaction terms of interest preserve the right signs and remain statistically significant. The magnitudes of the coefficients are similar to those in the baseline results. This reassures me that the baseline results are not affected by the inclusion of additional covariates. Interestingly, the coefficient of the average number of

34 See, for instance, Acemoglu, S. Johnson, and Mitton (2009) found that the interaction between financial development and firms' external-finance dependence is not robustly significant across different specifications. Macchiavello (2012) found that the interaction term is not significant unless industry heterogeneity in firm size distribution is taken into account.

35 In theory, the actual number of suppliers is proportional to the number of potential suppliers (entrants).

Table 1.10: Additional firm-level results
Dependent variable: $Outsourcing_{jnt}$

	All firms			Balanced panel
	(1)	(2)	(3)	(4)
$Post_t AveExtFinDep_n$			-0.0112***	-0.0111***
			(0.00404)	(0.00394)
$Post_t AveTang_n$			0.0680**	0.0742**
			(0.0326)	(0.0331)
$Post_t ExtFinDep_n$	-0.00207		-0.000785	-0.00101
	(0.00258)		(0.00195)	(0.00182)
$Post_t Tang_n$	0.0154		0.00424	0.00703
	(0.0117)		(0.00824)	(0.00669)
$Ave\#Suppliers_{nt}$		0.0000169***	0.00000673	0.00000162
		(0.00000640)	(0.00000804)	(0.00000738)
Firm FE α_j	Yes	Yes	Yes	Yes
Year FE α_t	Yes	Yes	Yes	Yes
N	570,937	609,019	570,937	215,650
R^2	0.863	0.861	0.864	0.850

Notes: In column 1-3, the sample includes all active UK firms as reported in the Orbis database for the period of 2007 – 2011. In column 4, the sample is restricted to firms who stay active in the whole period from 2007 to 2011. The dependent variable is the measure of outsourcing, as defined as unity minus the share of value-added in operating revenue. $Post_t$ is a dummy indicating the post-crisis period 2009 – 2011. $AveExtFinDep_n$ is a simple average of the external financial dependence of the top 10 upstream industries of industry n. $AveTang_n$ is a simple average of the external financial dependence of the top 10 upstream industries of industry n. $ExtFinDep_n$ is the measure of external financial dependence for the industry to which the firm belong. $Tang_n$ is the measure of asset tangibility for the industry of the firm. $Ave\#Suppliers_{nt}$ is the average number of suppliers in the top 10 upstream industries of industry n. All industry-level measures except for the average number of suppliers are time-invariant. Robust standard errors are in parentheses, allowing for correlation at the industry level. ***p<0.01, **p<0.05, and *p<0.1.

suppliers becomes insignificant. The loss of significance suggests that it may be appropriate to maintain the assumption of a perfectly elastic supply of potential sub-contractors. The reason is as follows. The contraction of credit supply may have shifted the (horizontal) supply curve upwards, leading to a smaller number of suppliers observed in equilibrium. However, once the factors shifting the supply curve (the interaction terms of interest) have been accounted for, the equilibrium number of suppliers yields no additional information, and therefore becomes insignificant. Column (4) shows the OLS estimates for the balanced sample. The results are broadly similar. The coefficients of interest are of similar magnitudes as those for the unbalanced sample, and they remain highly statistically significant. When used on its own, the number of average suppliers is positively and statistically significant (not shown); however, it becomes insignificant when a full set of covariates is included. Finally, the interaction terms between post and the financial characteristics of the firm remain statistically insignificant.

In brief, I have confronted the input-level model with firm-level data from Orbis. I have employed the difference-in-difference approach and exploited the recent financial crisis as a source of exogenous variation in credit supply. The results of reduced-form regressions broadly confirmed the predictions of my model. Taken together, the reduced-form evidence supports the hypothesis that firms actively revise their sourcing strategies in response to a change in credit supply. We can therefore rely on the model to answer policy questions. Since theory predicts that organizational changes have an impact on aggregate productivity, a natural question then is: how big is the impact? To answer this question, I turn to counterfactual experiments in the following section.

1.5 QUANTITATIVE ANALYSIS

Financial development has an impact on aggregate productivity through outsourcing. This section utilizes the model to answer the question: how much? Specifically, how much do countries benefit from setting their financial development to the U.S. level? The purpose of this exercise is to quantify the impact of financial development on aggregate productivity while focusing exclusively on the outsourcing channel. I proceed as follows. First, I outline a procedure to estimate model parameters. I then discuss the design of a counterfactual experiment. Next, I present the results from the estimation and the experiment. Finally, I discuss the results of the sensitivity analysis.

1.5.1 *Estimation*

Before explaining the estimation procedure, I need to specify how the parameters vary according to country and sector. I assume that the preference parameter α, the firm's bargaining power β, (the inverse of) productivity dispersion θ, and cost disadvantages of integration c and

f_V/f are country and sector independent. The intermediate demand parameter γ_{ni} is independent of the country. The overhead cost f_n and fixed entry cost f_n^e depend on the sector but not on the country. The sector composition parameter θ_{nc}, and the location parameter of productivity distribution b_{nc} are both country and sector dependent.

The estimation procedure can now be described. The share of intermediate purchases conveys valuable information on outsourcing. It is therefore the natural starting point for estimation. With the nested CES preference, the model admits a simple expression for the share of intermediate purchases:

$$\frac{X_{nic}}{X_{nc}} = \gamma_{ni}\left(1 - \sigma_{Rnic}\right).$$

Here, γ_{ni} is the aggregate demand for sector-pair ni intermediate goods, and σ_{Rnic} is the share of demand captured by integrating firms. The expression for σ_{Rnic} is given by:[36]

$$\sigma_{Rnic} = \left[\frac{\psi_V - \psi_{Oic}}{\psi_{Oic}}\frac{f_{Onic}}{f_{Vn} - f_{Onic}}\right]^{\theta(1-\alpha)/\alpha - 1}.$$

These closed-form expressions serve as the basis for estimation.

My estimation strategy involves two stages. First, I need to obtain estimates of the demand parameter γ_{ni} and the revenue share of integrating firms σ_{Rnic}. Second, using the estimates of revenue share, I obtain estimates of the structural parameters of the model.

I now turn to the first stage estimation. In light of the multiplicative form, I use the Poisson pseudo-maximum-likelihood (PML) estimator as proposed by Silva and Tenreyro (2006). To ensure the revenue shares fall into the unit interval, I assume it can be approximated by an inverse logit function. The inverse logit function takes two arguments: the upstream-sector-country fixed effect α_{ic} and the downstream-sector fixed effect α_n. Taking these together gives the following conditional mean:

$$E\left(X_{nic}/X_{nc}|\text{data}\right) = \gamma_{ni}\frac{1}{1 + \exp\left(\alpha_{ic} + \alpha_n\right)}.$$

To maintain the Cobb-Douglas assumption, I estimate this conditional mean function with restrictions $\sum_i \gamma_{ni} = 1$ for all n.

In the second stage, I estimate the structural parameters of the model from the first-step estimates $\hat{\sigma}_{Rnic}$. To keep the estimation work transparent, I calibrate some of the parameters and estimate the rest of them. I set the elasticity of substitution between firm products to $1/\left(1-\alpha\right) = 3.5$. Estimates of the substitutability of competing products typically range from three to ten. These estimates are based on industry level products. Since the model concerns sector level products, an elasticity of substitution of 3.5 seems appropriate. Later, I consider higher values for $1/\left(1-\alpha\right)$ as robustness checks. I set the

---

36 The expression for the profitability of outsourcing is $\psi_{Oic} = \beta + \frac{(1-\alpha)(1-\beta)}{1+\delta_i\left(\frac{1-\lambda_c}{\lambda_c}\right)}$, and for the fixed cost of outsourcing is $f_{Onic} = f_n - \frac{\left(\frac{1-\lambda_c}{\lambda_c}\right)\mu_i}{1+\delta_i\left(\frac{1-\lambda_c}{\lambda_c}\right)} f_n^e$.

bargaining power of the firm to $\beta = 0.5$. I have in mind the null hypothesis that neither party has an advantage in negotiations under outsourcing. I also set the fixed cost disadvantage for integration to $f_V/f = 1.2$. This means that integration requires 20 percent more of overhead costs. Since existing literature provides little guidance on how these two parameters should be calibrated, I choose their values conservatively. Later, I experiment with different values for these two parameters as sensitivity checks.

The calibration leaves me with two main parameters: the productivity dispersion θ and the marginal cost disadvantage of integration c. I obtain estimates of these parameters using another Poisson PML estimator with the conditional mean:

$$E\left(\hat{\sigma}_{Rnic}|\text{data}\right) = \left[\left(\frac{b_1}{1 + x_{1,ic}} - 1\right)\left(\frac{b_2}{1 - x_{2,ic}\alpha_n} - 1\right)^{-1}\right]^{b_3}.$$

Here, the constant b_1 contains parameter c, b_2 equals the fixed cost disadvantage, and b_3 contains parameter θ. The downstream-sector fixed effect α_n also requires estimation. The independent variables $x_{1,ic}$ and $x_{2,ic}$ are constructed from the share of private credit to GDP, and the two industry measures of external financial dependence and asset tangibility.[37] This completes the description of the estimation procedure.

1.5.2 *Counterfactual Experiment*

The purpose of this experiment is to compare the aggregate productivity of a country c at the current level of financial development, TFP_c, and at the hypothetical level of financial development, $T\check{F}P_c$. Aggregate productivity gain is defined as $\left(T\check{F}P_c/TFP_c - 1\right) \times 100$. The hypothetical level of financial development is the current U.S. level of financial development.

To calculate productivity gains, I proceed as follows. For each sector pair, I calculate the ratio of hypothetical sector TFP (1.5) to actual sector TFP, and the levels of hypothetical sector TFPR (1.4) and actual sector TFPR, and then aggregate these values across sector pairs using the Cobb-Douglas aggregators:

$$\frac{T\check{F}P_c}{TFP_c} = \prod_{n,i=1}^{N}\left(\frac{T\check{F}P_{nic}}{TFP_{nic}} \cdot \frac{T\check{F}PR_c/T\check{F}PR_{nic}}{TFPR_c/TFPR_{nic}}\right)^{\theta_{nc}\gamma_{ni}},$$

where $TFPR_c = \left[\sum_{m,k=1}^{N}\theta_{mc}\gamma_{mk}/TFPR_{mkc}\right]^{-1}$ is the aggregate TFPR.

37 The expressions for the constants are $b_1 = \left(\frac{1-\alpha}{\beta}\right)\left(c\left(1-\beta\right)\right)^{-\frac{\alpha}{1-\alpha}}$, $b_2 = f_{Vn}/f_n$, and $b_3 = \frac{\theta(1-\alpha)}{\alpha} - 1$. The expression for the fixed cost is $\alpha_n = f_n^e/f_n$. The expressions for the two independent variables are $x_{1,ic} = \frac{(1-\alpha)(1-\beta)}{\beta}\left(1+\delta_i\left(\frac{1-\lambda_c}{\lambda_c}\right)\right)^{-1}$ and $x_{2,ic} = \mu_i\left(\frac{1-\lambda_c}{\lambda_c}\right)\left(1+\delta_i\left(\frac{1-\lambda_c}{\lambda_c}\right)\right)^{-1}$. The variable λ_c corresponds to the share of domestic credit to private sector to GDP. The variables δ_i and μ_i correspond to the industry measures of external financial dependence and asset tangibility. I rescale λ_c and δ_i to unit interval range by logit transformations.

The productivity gains can be decomposed as follows. Recall that financial development has dual effects on aggregate productivity. The first term of the preceding equation refers to the direct effect of financial development on sector TFP. The second term represents the indirect effect of financial development through the reallocation of resources across sectors. Using the identify $TFP_{nic} = TFPR_{nic}/P_{nic}$, I further decompose the direct TFP effect into two components: the effect on sector revenue productivity and the effect on sector price. For the impact on sector price, the most important component is the effect of financial development on the entry of firms, which reflects the reallocation of resources within sectors. I will report the counterfactual productivity gain, along with its decomposition as outlined above.[38]

1.5.3 *Results*

Having explained the estimation procedure and experimental design, I now present the results in order. The first step estimation recovers the share of outsourcing from data using a Poisson PML estimator. Since this step involves only fixed effects, I therefore report the predicted values for the share of outsourcing. The summary statistics of the predicted shares are given in Table 1.11. It shows that the average sector purchases about 60 percent of inputs from upstream suppliers. The distribution of the shares is slightly skewed to the right, with the median slightly higher than the mean. Furthermore, it is worth mentioning that the R-squared (0.568) is higher than the reduced-form regressions. This means the Poisson estimator performs better than OLS in capturing systematic variations.

The second step estimation backs out key parameters (productivity dispersion, cost disadvantage of integration) from the predicted values for the share of outsourcing. Table 1.12 reports the results. The point estimate for the productivity dispersion θ is 4.678. This parameter is comparable to the trade elasticity in the international trade literature. Estimates of elasticities of trade in this literature typically range from three to eight. The estimate of θ falls well within this range. The point estimate for the cost disadvantage of integration is 1.427, implying that integration is about 40 percent less cost efficient than outsourcing. Both point estimates are statistically highly significant.

With these point estimates, I can now perform counterfactual calculations. The summary statistics for the counterfactual variables are given in Table 1.11. Figure 1.5a shows the percentage change in the prevalence of outsourcing. From the diagram, it is clear that the prevalence of outsourcing increases in the vast majority of sectors. The average increase of outsourcing is 13 percentage points, bringing up

38 The direct TFP effect is defined as $\left(D\check{TFP}_c/DTFP_c - 1\right) \times 100$, where $\frac{D\check{TFP}_c}{DTFP_c} = \prod_{n,i=1}^{N}\left(\frac{\check{TFP}_{nic}}{TFP_{nic}}\right)^{\theta_{nc}\gamma_{ni}}$. The indirect effect through resource reallocation across sectors is defined as $\left(ID\check{TFP}_c/IDTFP_c - 1\right) \times 100$, where $\frac{ID\check{TFP}_c}{IDTFP_c} = \prod_{n,i=1}^{N}\left(\frac{\check{TFPR}_c/\check{TFPR}_{nic}}{TFPR_c/TFPR_{nic}}\right)^{\theta_{nc}\gamma_{ni}}$. The revenue productivity and price effect can be defined analogously by noting that $TFP_{nic} = TFPR_{nic}/P_{nic}$. The entry effect is defined similarly by noting $P_{nic} \propto \left(M^e_{nic}\right)^{-\alpha/(1-\alpha)}$.

Table 1.11: Summary statistics for estimation and counterfactuals

	Mean	Median	S.D.	Min	Max	N
First stage						
Prevalence of outsourcing	0.613	0.639	0.297	0.000	1.000	135,252
Second Stage						
Change in outsourcing (p.p.)	13.079	11.836	8.737	-8.191	46.712	135,252
Change in mass of entrants (%)	9.060	9.420	4.423	-0.100	18.891	135,252
Change in sector price (%)	-0.585	-0.632	1.069	-2.159	17.460	135,252
Change in sector TFP (%)	0.274	0.306	1.134	-16.464	2.255	135,252
Aggregate productivity gain (%)	0.276	0.317	0.273	-1.163	0.789	117
Decomposition of productivity gains						
TFP effect (%)	0.276	0.317	0.273	-1.164	0.788	117
Reallocation effect (%)	0.000	0.000	0.000	0.000	0.001	117
Price effect (%)	0.592	0.607	0.304	-0.569	1.162	117
TFPR effect (%)	-0.314	-0.350	0.110	-0.598	0.000	117
Entry effect (%)	1.832	2.086	0.653	0.000	2.633	117

Notes: All variables are estimated/inferred using the GTAP 9 database with reference year 2007. Prevalence of outsourcing refers to predicted value for the revenue share of integrating firms from the first stage estimation. Values are reported for each pair of sectors and each country. Change in outsourcing, mass of entrant/sector price/sector TFP, and aggregate productivity gains are inferred from the second stage estimation. Change in mass of entrants is the percentage change of the mass of entrants for each pair of sector after setting financial development to the U.S. level for all countries. Analogously, change in sector price (TFP) is the percentage change of the sector price index (TFP) for each pair of sectors after setting financial development to the U.S. level. Aggregate productivity gain is defined as the percentage change of productivity for each country when switching to the U.S. level of financial development. Aggregate productivity gains can be decomposed into two components: the direct TFP effect and the indirect effect through resource reallocation across sectors. The direct TFP effect can be further decomposed into two effects: the effect on revenue productivity and the effect on sector price. The effect on sector price can be decomposed into three components: entry effect, the effect on surviving probability, and the effect on the composition of organizational forms. The most important component (the entry effect) is reported here.

Table 1.12: Baseline result of second step estimation

Dependent variable: $\hat{\sigma}_{Rnic}$	
	(1)
Productivity dispersion θ	4.678***
	(0.0567)
Cost disadvantage c	1.427***
	(.000977)
Downstream FE α_n	Yes
N	135252
R^2	0.0242

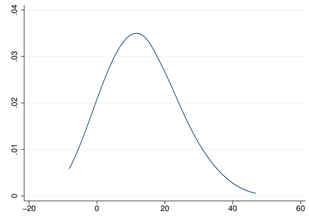
Notes: The sample includes the prevalence of integration for all pair of sectors in all countries for the year 2007. The dependent variable is the prevalence of integration specific to sector pair *ni* and country *c*, as obtained from the first stage estimation. Productivity dispersion θ refers to (the inverse of) the dispersion of productivity drawn from the Pareto distribution. Cost disadvantage c refers to the higher marginal cost for integration as compared to the unity marginal cost for outsourcing. Robust standard errors are in parentheses, allowing for correlation at the country level. ***p<0.01, **p<0.5, and *p<0.1.

the mean of outsourcing from about 60 to more than 70 percentage points. This is consistent with the prediction in Proposition 1.1. Of special interest is the change in the mass of entrants. Figure 1.5b shows that the mass of entrants increases in all but a few sectors (pairs). The average sector attracts 9 percent more entrants. The reason is that outsourcing requires fewer labor resources, freeing up labor for the creation of new firms. This is consistent with the prediction in Proposition 1.2.

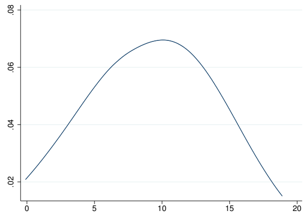
How do sector prices respond to the increase in financial development? Figure 1.5c plots the distribution of percentage changes in sector price. As expected, price rises and falls coexist. The assumed reason for price rises is that integration is more profitable. Integrating firms can generate more profits by charging lower markups and lower prices. Financial development induces integrating firms to switch to outsourcing, thereby causing sector prices to rise. The reason for price falls is that outsourcing is more cost efficient. It allows more resources to be devoted to the creation of new products. More products mean lower sector prices. Although price rises and falls occur at the same time, they are not evenly distributed. Prices fall in about 85% of all sectors (pairs), implying that aggregate prices are likely to fall. Meanwhile, Figure 1.5d plots the distribution of percentage changes in sector (pair) TFP. The changes in sector TFP mirror the changes in sector prices. TFP rises and falls coexist, with TFP rises occurring more often than TFP falls. The average increase of TFP of all sectors is about 0.3%. This is an important channel of productivity gains.

Aggregate productivity gains can now be discussed. Table 1.11 shows the summary statistics for the productivity gains of 107 countries in the sample. Setting financial development to the U.S. level would increase aggregate productivity by about 0.3% on average. However, countries do not equally benefit from further financial devel-

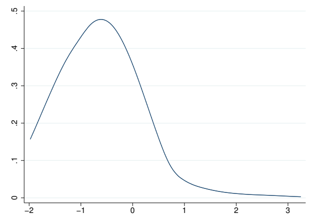
Figure 1.5: Counterfactual: The response of key variables in counterfactuals



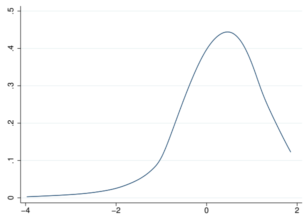
(a) Change in outsourcing (p.p.)



(b) Change in mass of entrants (%)



(c) Change in sector price (%)



(d) Change in sector TFP (%)

opment. Figure 1.6 shows the cross-country differences in productivity gains. In this diagram, I plot productivity gains against the current level of financial development. The diagram shows that, as expected, countries with a low level of financial development are more likely to benefit from setting their financial development to the U.S. level. However, countries with the same level of financial development do not necessarily enjoy the same productivity gains. There are a few economies that actually lose from further financial development. The reason is that price rises in more sectors than price falls in those economies.[39] The productivity gains for an average country are about 0.3%, which is roughly the same as the average sectoral TFP gains. This suggests that the direct effect of financial development on sector TFP is the primary source of productivity gains. The indirect effect of resource reallocation across sectors has a limited impact on aggregate productivity.

This intuition is confirmed by the decomposition of productivity gains. The direct TFP effect is almost the same as the net productivity gains; the indirect effect from resource reallocation across sectors is virtually zero. Furthermore, the productivity gains from falling sector prices (about 0.6 percent) are about twofold of the net productivity gains. The average productivity gain from entry (about 1.8 percent), which represents a more efficient use of resources, is nearly sevenfold of the net productivity gain. The missing productivity gains (1.8-0.6=1.2 percent) are due to the changes in composition of organizations. By assumption, integrating firms are more profitable as they charge lower product prices and sell more products. Financial development induces more outsourcing firms, and thereby (ceteris

39 Since by assumption neither organizational form dominates the other, productivity loses are theoretically possible in this model.

Figure 1.6: Counterfactual: Aggregate productivity gains from financial development

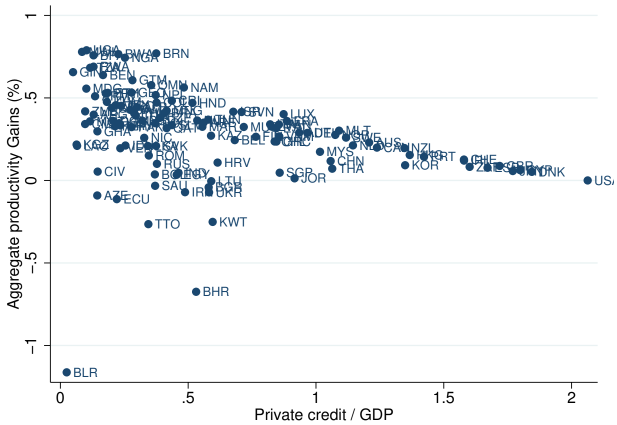


Table 1.13: Sensitivity analysis: The role of bargaining power and fixed cost disadvantage

	Fixed cost f_V/f			
Bargaining power β	1.2	1.4	1.6	1.8
0.3	-0.080	0.998	1.192	1.230
0.5	0.276	0.484	0.562	0.630
0.7	0.009	0.043	0.298	0.456

Notes: The table shows the mean value of cross-country productivity gains in counterfactuals when fixed production cost and bargaining power are calibrated to different values. Aggregate productivity gain is calculated as the percentage change in aggregate productivity when moving from the current equilibrium to the hypothetical equilibrium. In the hypothetical equilibrium, the level of financial development was set to the U.S. level for all countries.

paribus) raises sector price index. This assumed price effect offsets most of the productivity gains from improved allocative efficiency.

1.5.4 *Sensitivity Analysis*

Having presented the results, I discuss how sensitive the results are to the calibrated parameters. In the previous section, I set the bargaining power of the firm to $\beta = 0.5$. This amounts to the null hypothesis that neither party has an advantage during the negotiation. The alternative hypothesis is that one party has greater bargaining power over the other. I therefore consider a lower value 0.3 and a higher value 0.7 as robustness checks. Meanwhile, I set the fixed cost disadvantage of integration to $f_V/f = 1.2$. This means integration requires 20 percent more of overhead cost than outsourcing. This is a conservative number; hence, I entertain three higher values of 1.4, 1.6, and 1.8. Of course, these are still arbitrary numbers, so my purpose is to ascertain whether the productivity gains are sensitive to changes in parameter values.

The results are reported in Table 1.13. The results show that (the mean) productivity gains fall with bargaining power. Intuitively, a higher bargaining power means the firm can retain a bigger fraction of revenue; the concern for rent extraction is therefore less important. Hence, the potential scope of productivity gains from financial development is limited. On a technical level, a high bargaining power aggravates the underinvestment problem, making outsourcing less attractive. To generate a pattern of outsourcing that is consistent with real data, the estimator therefore infers that the cost disadvantage for integration should also increase. A higher cost disadvantage encourages more marginal integrating firms to switch to outsourcing as financial development takes place. Adjustments along the intensive margin (organizational switch) tend to suppress productivity gains, because outsourcing firms are more likely to charge higher prices. Meanwhile, the estimator infers that the (inverse of) productivity dispersion θ should also increase. This means the relative number of low productivity firms increases, and the productivity distribution is more concentrated at these low productivity levels. As financial development occurs, adjustment along the extensive margin (entry and the surviving probability post entry) tends to enhance productivity gains. Productivity gains fall with bargaining power since the effect at the intensive margin dominates the effect at the extensive margin.

By contrast, productivity gains rise with the fixed cost disadvantage f_V/f. The intuition is straightforward. A higher fixed cost disadvantage implies a greater scope for productivity gains from improving allocative efficiency. Hence as financial development takes place, productivity gains become bigger. On a technical level, as the fixed cost disadvantage increases, integration becomes less attractive. To maintain the relative attractiveness of integration, the estimator infers that the marginal cost disadvantage for integration should be lower. The lower marginal cost disadvantage discourages marginal integrating firms from switching to outsourcing. This tends to suppress adjustments along the extensive margin. The estimator further infers that the (inverse of) productivity dispersion should be higher as well. As the productivity distribution becomes more concentrated at the lower productivity levels, productivity gains become higher, for adjustments along the extensive margin become more important. Unlike bargaining power, however, the two effects tend to work in the same direction. It follows that productivity gains rise with the fixed cost disadvantage. Overall, the estimated productivity gains remain within the narrow interval between -0.1 and 1.2. Hence the benchmark productivity gains are not very sensitive to these two parameters.

I also set the elasticity of substitution between firm products to $1/(1-\alpha) = 3.5$. Estimates of substitutability of competing products in the trade literature typically range from three to ten, depending on the level of disaggregation. Since I have in mind products at the sector level, the elasticity of substitution is likely to fall into the interval between three and five. I therefore experiment with a variety of values within this range. The role of elasticity of substitution is relatively well known. As the substitutability between competing products becomes higher, productivity gains from variety become smaller. Furthermore, as consumer demands become less elastic, the profitabilities of dif-

Table 1.14: Sensitivity analysis: The role of elasticity of substitution

Elasticity of substitution $1/(1-\alpha)$	3	3.5	4	5
Aggregate productivity gains (%)	0.310	0.276	0.224	0.121

Notes: The table shows the mean value of cross-country productivity gains in counterfactuals when the elasticity of substitution is calibrated to different values. Aggregate productivity gain is calculated as the percentage change in aggregate productivity when moving from the current equilibrium to the hypothetical equilibrium. In the hypothetical equilibrium, the level of financial development was set to the U.S. level for all countries.

ferent organizational forms become more similar. This means that productivity gains from choosing efficient organizational forms become smaller. Hence productivity gains should fall with the elasticity of substitution. Table 1.14 reports the results. As expected, productivity gains decline as competing products become more substitutable. However, the changes in productivity gains remain small.

I have experimented with a variety of values for the calibrated parameters. Aggregate productivity gains are more responsive to changes in the bargaining power and the elasticity of substitution. They are less responsive to changes in the fixed cost disadvantage. Nonetheless, the absolute changes in aggregate productivity gains remain small. They remain within the narrow interval between 0.12 and 0.34. The baseline productivity gains (0.3 percent) fall in the middle of this interval. I conclude that the baseline results are not sensitive to the calibrated parameters.

1.6 CONCLUSION

There is a large body of theoretical and empirical literature that studies the organizational choice of the firm.[40] Nonetheless, its effect on aggregate productivity and the underlying mechanisms are not well understood. A recent paper of Boehm (2015) argued that an improvement of contracting institutions reduces the costs of market transactions. Firms, driven by transaction cost rewards, alter the boundary of an organization and outsource internal activities to outside suppliers. Through this channel, contracting institutions have been shown to have a large impact on outsourcing and aggregate productivity. How outsourcing

40 On the theory side, numerous explanations have been advanced to understand the boundary of the firm. Firms alter the boundary of organization to reduce transaction costs (Klein, Crawford, and Alchian, 1978; Williamson, 1971, 1979, 1985), to align control and incentives (S. J. Grossman and O. D. Hart, 1986; O. D. Hart, 1995; O. Hart and Moore, 1990), to provide multi-tasking incentives (Holmström, 1999; Holmström and Milgrom, 1991, 1994; Holmström and Tirole, 1991), and to facilitate adaptive decision-making (Klein, 1996, 2000; Klein and Murphy, 1988, 1997; Williamson, 1971, 1973, 1975, 1991). See Gibbons (2005) for a survey. On the empirical side, numerous studies have studied the determinants of vertical integration. The propensity for firms to integrate varies systematically with technology intensity (Acemoglu, Aghion, et al., 2010), factor intensity (Antràs, 2003), product market competition (Aghion, Griffith, and Howitt, 2006; Bloom, Sadun, and Reenen, 2010), product price (Alfaro, Conconi, et al., 2016), contracting institutions (Boehm, 2015), and financial development (Acemoglu, S. Johnson, and Mitton, 2009; Macchiavello, 2012). See Lafontaine and Slade (2007) for a survey of early contributions.

per se affects aggregate productivity is less clear. Distinguishing the two aspects is important. The former (reduction of transaction cost) resembles technological progress, while the latter (reorganization of production) is related to productive and allocative efficiency.

In this paper, I highlight a channel for financial development to affect outsourcing without affecting the underlying transaction costs. I show that even in the absence of cost reduction, financial development may still induce a reorganization of production that enhances aggregate productivity. With outsourcing, productive efficiency may be gained when products and services are produced more efficiently by outside suppliers. From the social perspective, further allocative efficiency may be gained when the resources conserved by outsourcing are devoted to the creation of new products and services. I structurally estimate the model and perform counterfactuals to quantify the impact of this channel. Perhaps surprisingly, quantitative analysis shows that financial development has a sizable impact on outsourcing, but its effect on aggregate productivity is relatively modest. Outsourcing implies a more efficient use of resources, but it also entails costs of market transactions. The transaction costs generate a powerful negative effect through prices, which offsets the majority of productivity gains.

What policy implications can we draw from this study? First, in light of the preceding discussion, policymakers should take into account the negative price effect when evaluating the impact of outsourcing-related policies on aggregate productivity. Second, financial development improves the allocative efficiency of input markets. A better functioning input market enhances aggregate productivity and social welfare. Developing financial markets are especially beneficial for less-developed and developing countries, as the costs are likely to be lower. Financial development can be achieved by a variety of policies. For instance, policies that facilitate the access of credit information and strengthen the legal rights of the borrowing will improve the financial market and eventually the allocative efficiency. Third, while it is important to enhance universal credit access, it is especially important to ensure access to finance for small and medium-sized enterprises. A lack of access to finance is more likely to cause small businesses to exit and medium-sized businesses to switch organization inefficiently. This issue is particularly acute following the 2008 financial crisis.[41]

A number of areas are left for future research. Many extensions of the baseline model can be considered to assess the full impact of financial development on aggregate productivity. First, the model assumes that financial development does not affect the underlying transaction costs. It would be interesting to extend the model to allow for financial development to reduce transaction costs (Boehm, 2015). Second, the model abstracts from input-output linkages. It would be interesting to extend the model by incorporating network production (Acemoglu, Carvalho, et al., 2012; Baqaee, 2017), roundabout produc-

41 A recent report from the Department for Business Innovation & Skills confirmed that small and medium-sized businesses in the UK are facing ongoing tight credit since 2008 (Armstrong et al 2013).

tion (Eaton and Kortum, 2002; Jones, 2013; Krugman and Venables, 1995), or sequential production (Antràs and Chor, 2013; Melitz and Redding, 2014). Introducing inter-sector linkages may substantially increase productivity gains, as efficiency gain in one sector may have implications for productivity in other sectors. Third, the model focuses on static gains from resource reallocation. It would be interesting to bring in a dynamic dimension. It has been shown that financial development fosters entry. If entry is related to innovation and technology adoption, financial development may have a persistent effect on growth. The potential growth effect is likely to be greater than the static effect as reported here (Peters, 2013). These extensions will bring together a quantitative framework for a full evaluation of the productivity impact of financial development.

Another area for future research is to investigate the relationship between outsourcing and product prices. This subject has long been a source of controversy among economists and policy makers.[42] In this paper, I have shown that transaction costs generate a powerful negative effect through prices, offsetting the majority of productivity gains from outsourcing. This result hinges on a key model property that integration leads to lower product prices. Several studies have found that vertical integration is associated with higher product prices (Alfaro, Conconi, et al., 2016; Ford and Jackson, 1997). If this is the case, efficiency loss from the transaction costs of outsourcing may be lower than the high product price of vertical integration. Financial development may induce more outsourcing, reducing the product prices and reinforcing the productivity gains. Extending the model by relaxing the CES assumption (Melitz and G. I. P. Ottaviano, 2008) or incorporating richer contractual arrangements (Bernard and Dhingra, 2016) can generate a positive relationship between integration and product prices.

42 Early evidence for the view that integration reduces prices includes McBride (1983), Muris, Scheffman, and Spiller (1992), and Shepard (1993), whilst evidence against it includes Ford and Jackson (1997). See Lafontaine and Slade (2007) for further discussion. More recently, Hortaçsu and Syverson (2007) investigated the U.S. cement industry and concluded that more integration leads to lower product prices. Alfaro, Conconi, et al. (2016) addressed the opposite direction of causality and found that higher product prices generate more integration.

2 IMPERFECT COMPETITION AND MISALLOCATIONS

This paper studies the effect of variable markups on misallocation and aggregate productivity using the manufacturing data from India. Three main findings are presented. First, markup dispersion is an important source of misallocation. Such dispersion accounts for approximately one third of the misallocation observed in the economy. Due to the endogenous response of markups, removing non-markup distortions has only a modest impact on misallocation and aggregate productivity. Second, models with CES preferences and constant markups substantially overstate productivity gains. TFP gain implied by a CES model is nearly six times the TFP gain when variable markups are taken into account. Third, competition policies that reduce barriers of entry can reduce markup distortions and enhance aggregate productivity.

2.1 INTRODUCTION

Resource misallocation can lower aggregate total factor productivity (TFP). Understanding the sources of misallocation has becoming an important research agenda for macroeconomists. Research on this agenda has largely focused on the Dixit-Stiglitz CES preferences, implying that all firms have the same market power and charge the same markup. This is a strong simplifying assumption. This paper studies the impact of variable markups by relaxing the CES assumption.

My main contribution is to demonstrate that variable markups are an important source of misallocation. When firms charge differing markups, resources will be reallocated away from high- to low-markup firms. If additionally, markups systematically correlate with firm productivity, variable markups would lead to systematic resource misallocation, causing substantial losses in aggregate productivity.

The study of India's manufacturing data reveals that this channel is quantitatively significant. First, the raw data suggest that variable markup, in its static form, accounts for about one third of the observed misallocation (as measured by the dispersion in revenue TFP), while non-markup exogenous distortions (government restrictions, taxes and subsidies) account for the remaining two thirds. Second, removing non-markup distortions has only a modest effect on misallocation and aggregate productivity, despite the fact that these distortions are substantial. The reason is the endogenous response of markups: removal of distortions, which tax productive firms and subsidize unproductive ones, leads to more dispersion in effective firm productivity and hence greater dispersion in markups, offsetting the direct effect of removing distortions. Third, I show that increasing market competition by reducing entry barriers has a significant effect on resource allocation and aggregate productivity because tougher competition reduces markups across the board and hence the dispersion of revenue TFP.

My second contribution is to caution against the use of CES preferences (with constant markups) in the context of studying misallocation. The reason is that a significant portion of variation in the data will be misinterpreted as non-markup distortions. In addition, models based on CES preferences ignore the endogenous response of markups to changes in distortions. Taken together, CES models would significantly overstate the TFP gains. The study shows that TFP gain predicted by a CES model is nearly six times the effect when variable markups are taken into account. Hence, studies based on CES preferences may provide misleading results.

The first step of the analysis is to separately identify markups from other types of distortions. In principle, it is difficult to distinguish markups from output taxes and subsidies given standard data for wage bills, capital, and value added. This approach leads to previous studies either interpreting the relevant variation as output distortions (Hsieh and Klenow, 2009; Restuccia and Rogerson, 2008) or as markups (De Loecker and Warzynski, 2012; Edmond, Midrigan, and Xu, 2015; Peters, 2016). To make progress, I use a profit measure to back out markups from data in a separate stage. Plausibility checks show that the recovered markups systematically correlate with other firm-level measures in ways that are consistent with the theory.

Since markups endogenously respond to changes in the economy, I need a model to fully assess the impact of variable markups on misallocation and aggregate productivity. I therefore build a quantitative model on the quasi-linear demand model by Melitz and G. I. P. Ottaviano (2008). An important feature of the model is that productive firms produce more but also charge higher markups, which implies that variable markups would cause systematic resource misallocation across firms. Another important feature of the model is that demand-side heterogeneity across sectors is summarized in a sufficient statistic, namely the sector-specific choke price at which consumer demand vanishes. This approach allows me to focus on the estimation of supply-side parameters, which significantly simplifies the policy experiments.

In the first policy experiment, I ask: How much gains in aggregate productivity can be achieved by removing exogenous distortions? I find that removing distortions has only a modest effect on resource misallocation and aggregate productivity. In particular, after removing non-markup distortions, approximately 80% of the dispersion in revenue TFP (a measure for the degree of misallocation) still remains in the market. To the extent that productivity gains come from improvements in misallocation, removing non-markup distortions has a small effect on aggregate productivity (about 14%). By contrast, a CES model significantly overstates the productivity gain (more than 80%).

A second policy experiment quantifies the impact of increasing market competition on misallocation and aggregate productivity. In this experiment, in addition to removing distortions, I further reduce the entry cost to a half for each sector. Tougher competition reduces markups across the board and further increases aggregate productivity by 8%. This is a large effect that is comparable with the effect of eliminating non-markup distortions. The policy experiment suggests a way to improve market allocative efficiency: increasing market competition either by lowering barriers of entry or by opening to international trade.

To explore whether the main results are driven by model misspecification, I perform two robustness checks. In the benchmark analysis, I have considered two types of non-markup distortions: a tax on capital and a subsidy on output. Naturally, one may wonder whether the small effect of removing distortions has to do with the fact that distortions on labor costs have been omitted. In the first robustness check, I include a labor distortion to capture these firm-specific taxes on labor costs. The results show that including a labor distortion only modestly increases the productivity gain (to approximately 19%).

The second robustness check concerns the production elasticity with respect to capital. In the benchmark analysis, I have followed Hsieh and Klenow (2009) to set sector capital elasticities to those in the corresponding U.S. sector. This approach avoids potential mismeasurements of capital elasticities when there are sector-specific capital distortions. However, it also raises concern about whether using the

U.S. capital elasticities is appropriate because production technologies in India may differ significantly from the United States. To address this concern, I estimate capital elasticities directly from the India data, using the proxy method introduced by Olley and Pakes (1996) and refined by Levinsohn and Petrin (2003). The results show that using native capital elasticities does not change the main results: the productivity gain from removing distortions is almost the same as that in the benchmark analysis (at about 14%).

This paper is closely related to the literature on misallocation and aggregate productivity, pioneered by Restuccia and Rogerson (2008) and Hsieh and Klenow (2009). In particular, Hsieh and Klenow (2009) find that improving misallocation to the U.S. level would increase aggregate TFP by about 30–50% in China and 40–60% in India. The sizable TFP gains have inspired many researchers to study the specific factors driving misallocation. Potential sources of misallocation have been suggested, including financial frictions (Banerjee and Duflo, 2005; Caselli and Gennaioli, 2013; Midrigan and Xu, 2014; Moll, 2014), capital adjustment cost (Asker, Collard-Wexler, and Loecker, 2014; Gopinath et al., 2017), information friction (David, H. A. Hopenhayn, and Venkateswaran, 2016), and markup dispersion (Peters, 2016). My paper is closely related to (Peters, 2016), who argues that variable markups have dynamic consequences on growth. My contribution to this literature is demonstrating that variable markups are a significant source of static misallocation.

This paper is also related to a branch of trade literature that theoretically models and empirically estimates variable markups. On the theory front, Melitz and G. I. P. Ottaviano (2008) propose a general equilibrium model with quasi-linear demands, and Nocco, G. I. P. Ottaviano, and Salto (2014) analyze the optimal allocations in this model. Dhingra and Morrow (2016) study the optimality of market allocation in models with addictively separable preferences. On the empirical front, De Loecker and Warzynski (2012) develop a method to estimate markups using production data, Edmond, Midrigan, and Xu (2015) use this method to study the effect of trade on markup distortions using Taiwanese data, and Peters (2016) applies this method to study the dynamic consequences of markup distortions using Malaysian data.

The remainder of the paper is structured as follows. Section 2.2 outlines the quasi-linear demand model and the accounting framework. Section 2.3 discusses the data and measurement issues. Section 2.4 presents the main results regarding the effect of variable markups on misallocation and aggregate productivity. Section 2.5 performs robustness checks. Section 2.6 concludes.

2.2 LINEAR DEMAND MODEL AND ACCOUNTING FRAMEWORK

Consider an economy with one homogenous good sector and S differential variety sectors. The representative consumer has the preferences:

$$U = Y_0 + \beta \sum_{s=1}^{S} \sum_{i=1}^{N_s} Y_{si} - \frac{1}{2}\gamma \sum_{s=1}^{S} \sum_{i=1}^{N_s} Y_{si}^2 - \frac{1}{2}\delta \sum_{s=1}^{S} \left(\sum_{i=1}^{N_s} Y_{si} \right)^2 - \frac{1}{2}\eta \left(\sum_{s=1}^{S} \sum_{i=1}^{N_s} Y_{si} \right)^2$$

Here, Y_0 is the homogenous good, which is used as numeraire. Y_{si} is the differential variety i in sector s, and N_s is the number of varieties in sector s. Think of the homogenous good sector as agriculture and the differential good sectors as manufacturing and services. Having a homogenous good sector ensures that the marginal utility of income is constant, hence the allocation of expenditure among differential goods are not affected by income.

The demand parameters β, γ, δ, and η are all positive. β and η govern the substitution pattern between the homogenous good and the differentiated varieties. An increase of β (or a decrease of η) shifts out the demand for varieties relative to the homogenous good. The parameter gamma governs the elasticity of substitution between the varieties within a sector, and delta governs the elasticity of substitution between differentiated variety sectors.

The consumer preferences imply a downward-sloping linear demand for variety i:

$$Y_{si} = \frac{1}{\gamma} \left(P_{\max,s} - P_{si} \right)$$

where $P_{max,s}$ is the choke price for sector s and P_{si} is the price of the variety i in sector s. The expression of the choke price is given by:

$$P_{\max,s} = (1 - \omega_s) \left[\frac{\delta}{\delta + \eta W} \beta + \frac{\eta W}{\delta + \eta W} \overline{P} \right] + \omega_s \overline{P}_s$$

where $\overline{P}_s = N_s^{-1} \sum_i P_{si}$ is the price index for sector s and $\overline{P} = W^{-1} \sum_s \omega_s \overline{P}_s$ is the aggregate price index for differentiated goods. The sector weights are given by $\omega_s = \delta N_s / (\gamma + \delta N_s)$ and $W = \sum_s \omega_s$.

Two features of the linear demand function are particularly noteworthy. First, the price elasticity $\epsilon_{si} = -(\partial Y_{si} / \partial P_{si})(P_{si} / Y_{si}) = (P_{\max,s} / P_{si} - 1)^{-1}$ varies across varieties. In particular, the price elasticity increases in the level of price P_{si}, and thus decreases in the level of output Y_{si}. Larger firms produce more outputs and also face a less elastic consumer demand. Hence, they charge higher markups and higher prices on their products.

Second, the choke price captures the information on consumer preferences (β, γ, δ, and η) and market conditions (N_s, $\overline{P}_s$, $\overline{P}$). Since the slope of the demand function ($1/\gamma$) is common for all differentiated variety sectors, the choke price is the only source of cross-sector heterogeneity on the demand side. This observation motivates a sufficient statistic approach to quantify the impact of variable markups on aggregate productivity in the later section.

There are two factors of production, capital, and labor. Producing the variety requires combining capital and labor with a Cobb-Douglas production function:

$$Y_{si} = A_{si} K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}$$

where A_{si} is the variety-specific technology, and α_s is the capital share of output which is specific to sector s.

Each variety is produced by a single firm. Firms differ in two dimensions. First, each firm is endowed with a variety-specific technology A_{si}. Second, each firm faces idiosyncratic distortions. Specifically, denote the distortion that affect both capital and labor as output distortion τ_Y, and the distortion that changes the marginal product of capital relative to labor as capital distortion τ_K. For example, think of τ_Y as the government restrictions or taxes to the firm, and τ_K as the differing costs of financing capital expenditure due to imperfect credit markets.

Profit of the firm is given by:

$$\pi_{si} = (1 - \tau_{Ysi}) P_{si} Y_{si} - w L_{si} - (1 + \tau_{Ksi}) R K_{si}$$

where w is wage to worker and R is rent to capital, both of which are common across sectors.

Markets are monopolistically competitive, implying that each firm will maximize profit by setting input quantities and output price while taking the consumer demand as given. Cost minimization implies that capital intensity is determined by relative factor price:

$$\frac{K_{si}}{L_{si}} = \frac{\alpha_s}{1-\alpha_s} \frac{w}{R} \frac{1}{1+\tau_{Ksi}}$$

Profit maximization yields the standard condition that the firm's output price is a markup over its marginal cost:

$$P_{si} = \mu_{si} \left(\frac{R}{\alpha_s}\right)^{\alpha_s} \left(\frac{w}{1-\alpha_s}\right)^{1-\alpha_s} \frac{(1+\tau_{Ksi})^{\alpha_s}}{A_{si}(1-\tau_{Ysi})}$$

Here, $\mu_{si} = \epsilon_{si} / (\epsilon_{si} - 1)$ is the markup charged by the firm, and the rest of the terms is the marginal cost C_{si} for the firm.

The solution for optimal markup is given by:

$$\mu_{si} \propto \frac{A_{si}(1-\tau_{Ysi})}{(1+\tau_{Ksi})^{\alpha_s}} - \text{constant}$$

The markup charged by firms depends not only on firm productivity levels, but also on the output and capital distortions they face. To the extent that markup is a function of distortions, changes in distortions would lead to endogenous response of markups. One of the key insights in this paper is that the endogenous response of markups may offset the direct effect of eliminating distortions.

The allocation of resources across firms depends on markups, firm productivity levels, and distortions.

$$\begin{aligned} L_{si} &\propto \frac{1}{\mu_{si}} \left(\text{constant} - \left(\frac{(1+\tau_{Ksi})^{\alpha_s}}{A_{si}(1-\tau_{Ysi})} \right)^2 \right) (1-\tau_{Ysi}) \\ K_{si} &\propto L_{si} \frac{1}{1+\tau_{Ksi}} \\ Y_{si} &\propto \text{constant} - \frac{(1+\tau_{Ksi})^{\alpha_s}}{A_{si}(1-\tau_{Ysi})} \end{aligned}$$

Conditional on firm productivity level and distortions, a higher markup implies that fewer resources will be allocated to the firm. In an ideal world without any distortions, there is still within-sector misallocation due to endogenous markups. More productive firms charge higher markups, implying that resources are allocated away from these firms to less productive firms, leading to inefficient outcomes. If more productivity firms also face more distortions, then eliminating distortions would allow them to charge higher markups. The endogenous response of markups exacerbates the within-sector misallocation, offsetting the direct effect of eliminating distortions. I will come back to this point later.

The revenue marginal products of labor and capital are given by:

$$\begin{aligned} MRPL_{si} &= \frac{1}{\mu_{si}} (1-\alpha_s) \frac{P_{si}Y_{si}}{L_{si}} = w \frac{1}{1-\tau_{Ysi}} \\ MRPK_{si} &= \frac{1}{\mu_{si}} \alpha_s \frac{P_{si}Y_{si}}{K_{si}} = R \frac{1+\tau_{Ksi}}{1-\tau_{Ysi}} \end{aligned}$$

The intuition here is that the after-tax revenue marginal products of labor and capital are equalized across firms. Hence, the before-tax revenue marginal products must be higher in firms that face taxes and lower in firms that face subsidies.

The revenue productivity is given by:

$$TFPR_{si} = \frac{P_{si}Y_{si}}{K_{si}^{\alpha_{si}} L_{si}^{1-\alpha_{si}}} = \mu_{si} \left(\frac{MRPK_{si}}{\alpha_s} \right)^{\alpha_s} \left(\frac{MRPL_{si}}{1-\alpha_s} \right)^{1-\alpha_s}$$

Intuitively, high revenue productivity can arise either because the firm charges a high markup or because the firm has high revenue marginal products. In the linear demand model, firms charge differing markups, hence TFPR would not be equalized even if firms face no capital or output distortions. However, regardless of the reasons, high firm TFPR is a sign that the firm uses fewer resources and hence is smaller than optimal (in the sense of maximizing output in the economy).

It is now ready to consider the aggregate economy. The sectoral revenue is $P_s Y_s$, and sectoral inputs are K_s and L_s. The sectoral markup and revenue marginal products are the weighted harmonic means of their firm-level measures:

$$\mu_s = \left(\sum_{i=1}^{N_s} \frac{P_{si} Y_{si}}{\sum_{j=1}^{N_s} P_{sj} Y_{sj}} \mu_{si}^{-1} \right)^{-1}$$
$$MRPL_s = \left(\sum_{i=1}^{N_s} \frac{\mu_{si}^{-1} P_{si} Y_{si}}{\sum_{j=1}^{N_s} \mu_{sj}^{-1} P_{sj} Y_{sj}} MRPL_{si}^{-1} \right)^{-1}$$
$$MRPK_s = \left(\sum_{i=1}^{N_s} \frac{\mu_{si}^{-1} P_{si} Y_{si}}{\sum_{j=1}^{N_s} \mu_{sj}^{-1} P_{sj} Y_{sj}} MRPK_{si}^{-1} \right)^{-1}$$

The sectoral revenue TFP is a geometric mean of the sectoral markup and revenue products:

$$TFPR_s = \frac{P_s Y_s}{K_s^{\alpha_s} L_s^{1-\alpha_s}} = \mu_s \left(\frac{MRPK_s}{\alpha_s} \right)^{\alpha_s} \left(\frac{MRPL_s}{1-\alpha_s} \right)^{1-\alpha_s}$$

Finally, the sectoral TFP can be expressed as:

$$TFP_s = \sum_{i=1}^{N_s} \frac{TFPR_{si}^{-1} P_{si} Y_{si}}{\sum_{j=1}^{N_s} TFPR_{sj}^{-1} P_{sj} Y_{sj}} A_{si}$$

This equation illustrates how endogenous markups and exogenous distortions affect sectoral productivity. If there were no capital and output distortions, marginal revenue products would be equalized and the sectoral TFP would be:

$$TFP_s = \sum_{i=1}^{N_s} \frac{\mu_{si}^{-1} P_{si} Y_{si}}{\sum_{j=1}^{N_s} \mu_{sj}^{-1} P_{sj} Y_{sj}} A_{si}$$

If more productive firms face more disincentives, then eliminating distortions induces more resources to move toward more productive firms, resulting in a higher sectoral productivity (since weights of more productive firms increase). However, these high productive firms also increase their markups accordingly, resulting in a lower sectoral productivity than optimal (since weights of more productive firms decrease). Hence, the endogenous response of markups partially offsets the direct effect of eliminating distortions.

2.3 DATA AND MEASUREMENT

The primary source of data comes from India's Annual Survey of Industries (ASI). India's ASI is conducted annually by its Central Statistical Organization. The basic survey unit is plant, also known as establishment. The plant here corresponds to the firm in the model. I have access to the cross-section data from 2001–2008. In the interest of space, I only report results for 2001, 2004, and 2007. Results for other years are available upon request. The raw data for 2001 contains approximately 26 thousand plants, and the data for 2004 and 2007 each contains about 36 thousand plants.

The variables used in the analysis include plants' industry (4-digit NIC), labor compensation, value added, book value of the fixed capital stock, and profit. Specifically, the ASI reports the plant's total wage payments, bonus payments, and benefit payments. Labor compensation is the sum of wages, bonuses, and benefits. The novel element here is that I also measure profit from the ASI data. The ASI reports the plant's depreciation of the fixed capital stock, rents paid for property, plant, and equipment (PP&E), and interest payments. The measure of profit is value added less the sum of depreciation, rents, interests, and labor compensation. In addition, the ASI reports the book value of fixed capital at the beginning and end of the fiscal year. I use the average book value of fixed capital at the beginning and end of the fiscal year as the plant's capital.

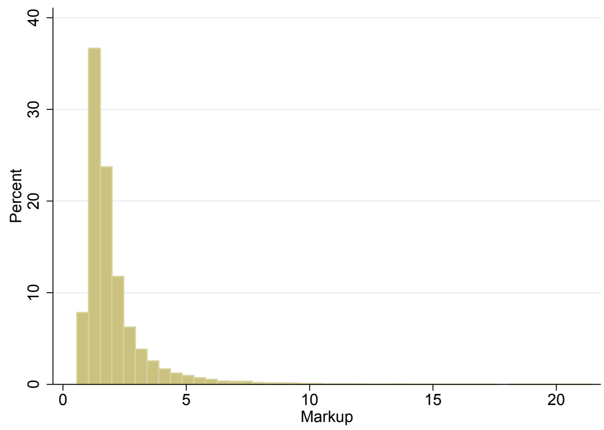
The capital share of output alphas are taken from the NBER Productivity Database (E. J. Bartelsman and Gray, 1996). This means that I set the capital share of output to the value in the corresponding industry in the United States. As argued by Hsieh and Klenow (2009), it is difficult to separately identify the average capital distortion and the capital production elasticity from the Indian data, and hence they use the corresponding values in the United States. However, using the capital share from the United States also amounts to assuming that India has access to the same production technology. This assumption may be too strong. In the robustness check, I estimate the capital shares from the Indian data using the method proposed by Olley and Pakes (1996) and subsequently refined by Levinsohn and Petrin (2003). The main results in this paper do not depend on the choice of capital shares.

To infer markups and distortions from the Indian data, I follow the method of Hsieh and Klenow (2009). The key insight is that distortions can be inferred from the differences between observed values and the valued implied by theory. For example, theory predicts that if there were no capital distortions, the marginal rate of technological substitution must equal the relative factor price. Any deviation from this observation, therefore, can be inferred as capital distortion. Following the same idea, I infer markups, distortions, and productivity for each plant as follows:

$$
\begin{aligned}
\mu_{si} &= \frac{(1-\alpha_s)\,\pi_{si}}{wL_{si}} + 1 \\
1+\tau_{Ksi} &= \frac{\alpha_s}{1-\alpha_s}\frac{wL_{si}}{RK_{si}} \\
1-\tau_{Ysi} &= \mu_{si}\frac{wL_{si}}{(1-\alpha_s)\,P_{si}Y_{si}} \\
A_{si} &= \kappa_s\frac{\left(2-\mu_{si}^{-1}\right)P_{si}Y_{si}}{K_{si}^{\alpha_s}L_{si}^{1-\alpha_s}}
\end{aligned}
$$

The first equation infers markups on the basis of definition. The third equation bases on the observation that if there were no output distortion, revenue (value added) should be equal to the product of markup and total production cost ($wL_{si}/(1-\alpha_s)$). Deviations from

Figure 2.1: Distribution of markup



this observation imply there are output distortions. In the last equation, the sector-specific constant $\kappa_s = 1/P_{\max,s}$. Although I do not observe κ_s, productivity gains are unaffected by setting $\kappa_s = 1$ for each industry s. The summary statistics of these variables are provided in Table 2.1.

Three points are particularly noteworthy. First, for labor input I use the plant's labor compensation rather than its employment to measure L_{si}. This means that labor input is measured in efficiency units rather than raw counts. Second, here it is possible to separately infer markup and output distortion because I also measure profit from the Indian data. Third, it is useful to compare the approach here with that in De Loecker and Warzynski (2012), who developed a general method to estimate markups. The key insight there is that if there were no markups, labor production elasticity should be equal to its expenditure share. Hence, any deviation from this observation can be inferred as markups. This is equivalent to setting output distortion to one in the second equation, and attributing all the variation in $1/(1-\alpha_s) \times (wL_{si}/P_{si}Y_{si})$ to markups. By contrast, Hsieh and Klenow (2009) set a constant markup and attribute all the variation in $1/(1-\alpha_s) \times (wL_{si}/P_{si}Y_{si})$ to output distortion. The approach in this paper differs in that it attributes the variation to both markup and output distortion. Since this is a key departure from the two papers mentioned above, it is useful to perform plausibility checks on the key variables.

One way to check whether the inference makes good sense is to look at the distribution of markup. Figure 2.1 plots the distribution of markup for the year 2004. Except for a few firms, most firms charge a gross markup that is greater than one. The median firm charges a net markup of 60% over its marginal cost. The distribution is highly skewed to the left, with only a few firms charging very high markups.

Another way is to examine the distributions of revenue and physical TFP. Foster, Haltiwanger, and Syverson (2008) find that physical productivity is more dispersed than revenue productivity. Figure 2.2 plots the distribution of the log of revenue and physical TFP. There is clearly more dispersion of revenue TFP than physical TFP, which is consistent with findings in the previous literature.

Table 2.1: Summary statistics for relevant variables

	Mean	Median	S.D.	Min	Max	#obs
2001						
Value added	15.710	15.490	2.103	8.293	24.146	19,559
Capital	15.200	14.996	2.447	5.814	24.893	19,559
Labor compensation	14.445	14.184	1.997	7.090	22.960	19,559
Profit	14.785	14.598	2.317	5.919	23.684	17,887
Employment	3.905	3.689	1.508	0.000	10.631	19,559
Materials	16.235	16.292	2.374	5.371	25.317	19,558
Markup	1.970	1.547	1.357	0.542	17.123	19,559
Capital distortion	1.818	1.748	1.548	-4.098	9.709	19,559
Output distortion	0.207	0.152	0.467	-1.306	3.339	19,559
Productivity	1.082	1.092	0.812	-2.022	4.378	19,559
2004						
Value added	15.748	15.511	1.938	9.141	24.645	29,224
Capital	15.240	15.075	2.288	6.163	25.019	29,224
Labor compensation	14.469	14.232	1.821	8.666	23.315	29,224
Profit	14.859	14.660	2.181	5.159	24.256	27,609
Employment	3.783	3.611	1.371	0.000	10.708	29,224
Materials	16.323	16.385	2.299	5.704	25.489	29,221
Markup	2.036	1.597	1.487	0.566	21.402	29,224
Capital distortion	1.790	1.732	1.537	-4.449	9.947	29,224
Output distortion	0.213	0.155	0.437	-1.167	3.151	29,224
Productivity	1.122	1.112	0.760	-1.955	4.580	29,224
2007						
Value added	15.978	15.600	2.023	8.287	24.714	25,324
Capital	15.467	15.240	2.360	5.886	25.757	25,324
Labor compensation	14.665	14.290	1.913	7.535	22.091	25,324
Profit	15.141	14.810	2.220	4.920	24.367	24,475
Employment	3.814	3.526	1.451	0.000	10.303	25,317
Materials	16.502	16.531	2.459	5.521	26.343	25,322
Markup	2.070	1.669	1.314	0.551	15.761	25,324
Capital distortion	1.753	1.696	1.505	-4.477	9.150	25,324
Output distortion	0.207	0.158	0.418	-1.155	3.104	25,324
Productivity	1.160	1.135	0.724	-1.571	3.840	25,324

Notes: All variables except markup are in log terms.

Figure 2.2: Dispersion of revenue and physical TFP

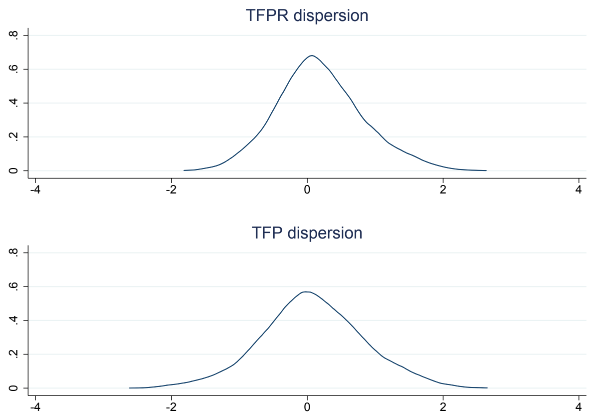


Table 2.2: Plausibility check

Year: 2004

	(1) Markup	(2) Revenue	(3) Profit
Productivity	1.609***	0.612***	1.426***
	(0.172)	(0.060)	(0.096)
Capital distortion	-1.009***	-0.502***	-0.857***
	(0.138)	(0.042)	(0.086)
Output distortion	1.800***	0.348*	1.172***
	(0.374)	(0.178)	(0.361)
Industry FE	Yes	Yes	Yes
N	29224	29224	27609
R^2	0.699	0.246	0.343

Notes: All variables except markup are in log terms. ** 5%, *** 1% levels of significance. Standard errors are clustered at the industry level.

An additional plausibility check is to ask whether the inferred variables are consistent with the predictions of the model. For example, the model implies that more productive firms charge higher markups, earn higher revenues and profits. A firm is more productive either because it has higher physical productivity, or because it receives more output subsidy and pays less capital tax. Table 2.2 reports the results of OLS regressions of markups, revenues, and profits on the inferred productivity, and output and capital distortions for the year 2004. Column (1) shows that higher markups are indeed associated with firms that have higher physical productivity, receive more output subsidies and pay less capital taxes. Column (2) and (3) show that such firms are also likely to earn higher revenues (value added) and profits. Of course, there is no way to claim that these results are causal: these relationships may simply hold by construction. What the results shows is that the way of inferring variables is consistent with theory.

To summarize, the inferred variables do seem plausible. With these variables in hand, I can now perform counterfactual experiments to quantify the impact of variable markups on resource misallocation and aggregate productivity.

2.4 COUNTERFACTUAL EXPERIMENTS

The central question in the paper is: How much of the observed resource misallocation (as measured by dispersion of TFPR) is driven by variable markups rather than exogenous distortions? To answer this question, I perform two counterfactual experiments. In each of the experiments, I first measure the dispersion of TFPR in the data as a benchmark. I then remove capital and output distortions from the economy, and let the firms adjust their markups accordingly. After that, I measure the dispersion of TFPR in the counterfactual data and compare it with the benchmark value. High dispersion of TFPR in the counterfactual data is a sign that variable markups are important for resource misallocation.

The counterfactual experiments rely on a sufficient statistic of the model: the industry-specific choke price for which a firm is indifferent between producing and exiting the industry. As explained in the theory section, the choke price captures the information on consumer preferences and market conditions. It is the only source of demand-side heterogeneity across industries. All the firm performance variables, including price, quantity, revenue, and profit are functions of the choke price and the marginal cost of the firm.

The choke price for each industry is determined by the corresponding free-entry condition. Specifically, if the distribution of marginal cost in sector s is denoted by $G_s\left(C_{si}\right)$, then the free-entry condition for sector s is given by:

$$\int_0^{P_{\max,s}} \pi_{si}\left(C_{si}, P_{\max,s}\right) dG_s(C_{si}) = f_{E,s}$$

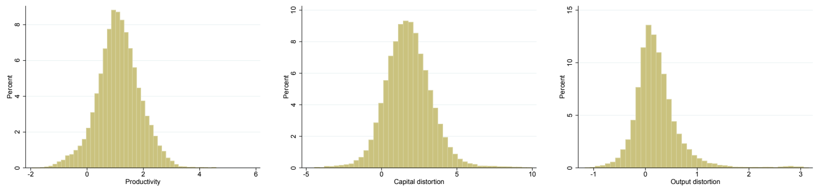
where $f_{E,s}$ is the fixed entry cost as measured by the numeraire good. The intuition here is that potential firms will enter the industry to the point that the cost of entry equals the expected benefit of entry. The latter is the average profit of all the firms except those whose marginal costs are too high to survive the industry ($C_{si} > P_{\max,s}$).

Suppose that consumer preferences and entry cost ($f_{E,s}$) do not change following an removal of capital and output distortions, then the free-entry condition implies:

$$\int_0^{P_{\max,s}} \pi_{si}\left(C_{si}, P_{\max,s}\right) dG_s(C_{si}) = \int_0^{\tilde{P}_{\max,s}} \pi_{si}\left(C_{si}, \tilde{P}_{\max,s}\right) d\tilde{G}_s(C_{si})$$

where tilde represents counterfactual outcomes. This equation provides a way to estimate the choke price in the counterfactual data. Specifically, first estimate the choke price as the highest price of all firms that have non-negative profit. Then estimate the counterfactual choke price by solving the equation above. With the counterfactual choke price in hand, I can calculate the dispersion of TFPR in the counterfactual data and compare it with the benchmark value.

Figure 2.3: Distribution of productivity and distortions



(a) Productivity (b) capital distortion (c) output distortion

To the extent that marginal cost is affected by productivity, and capital and output distortions, the distribution of marginal cost is determined by the joint distribution of these variables. Hence, I will first discuss the estimation of the joint distribution of productivity, and capital and output distortions, and then discuss the counterfactual experiments in detail.

2.4.1 *Estimation of Distributional Parameters*

What distribution is appropriate to model the joint relationship of productivity, and capital and output distortions? To motivate the distributional assumption, it is useful to examine the data. Figure 2.3 plots firms' productivity, and capital and output distortions in log terms. The data seem to suggest that a joint lognormal distribution is appropriate. Hence, I assume that firm productivity, and capital and output distortions are jointly lognormally distributed with mean μ_A, μ_K, μ_Y, and variance covariance matrix:

$$\begin{pmatrix} \sigma_A^2 & 0 & 0 \\ 0 & \sigma_K^2 & \sigma_{KY} \\ 0 & \sigma_{KY} & \sigma_Y^2 \end{pmatrix}$$

Note that I have assumed that firm productivity and distortions are ex-ante uncorrelated. The reasons are twofold. First, the purpose of the counterfactual experiment is to quantify the effect of variable markups on aggregate productivity. It would be useful to compare the counterfactual outcome of the linear demand economy versus that of the CES economy. As shown by Hsieh and Klenow (2009), when consumer preference is CES, the covariances of firm productivity and distortions do not affect the aggregate productivity. However, these covariances do affect the aggregate productivity in the linear demand economy. Had I assumed that productivity and distortions are ex-ante correlated, it would be difficult to assess whether the differences in outcomes are due to variable markups or due to the correlations between productivity and distortions.

Second, correlated distortions have larger effects on the endogenous response of markups than unrelated ones (as will be explained later). Assuming these away implies that the effect of variable markups will be underestimated. Hence, the results should provide a conservative estimate of the impact of variable markups on aggregate productivity.

Besides the distributional assumption, there is one more complication in estimation. In particular, the counterfactual experiments require knowing the full distribution of firm productivity and distortions, whereas in the data I only observe a selected sample of firms that choose to operate. Hence, I need to estimate the distributional parameters while explicitly accounting for the selected nature of the sample. The data selection rule is:

$$P_{\max,s} \geq C_{si} \propto \frac{(1+\tau_{Ksi})^{\alpha_s}}{A_{si}(1-\tau_{Ysi})}$$

Intuitively, only low marginal cost firms would choose to operate in the economy.

The set of parameters that await estimation is $(\mu_A, \mu_K, \mu_Y, \sigma_A^2, \sigma_K^2, \sigma_Y^2, \sigma_{KY})$. I estimate these parameters using a truncated maximum likelihood estimator with the conditional density function:

$$\begin{aligned} & f\left(\log A_{si}, \log\left(1+\tau_{Ksi}\right), \log\left(1-\tau_{Ysi}\right) | C_{si} \leq P_{\max,s}\right) \qquad (2.1) \\ & = \frac{1}{\sigma_A}\phi\left(\frac{\ln A_{si}-\mu_A}{\sigma_A}\right)\frac{1}{\sigma_{\tau s}}\phi\left(\frac{\ln \tau_{si}-\mu_{\tau s}}{\sigma_{\tau s}}\right)\left[\Phi\left(\frac{\mu_{cs}-\ln P_{\max,s}}{\sigma_{cs}}\right)\right]^{-1} \end{aligned}$$

where Φ is the distribution function of $N(0,1)$ and ϕ is the corresponding density function. For notational convenience, I define $\tau_{si} = (1-\tau_{Ysi}) / (1+\tau_{Ksi})^{\alpha_s}$ and denote its mean and variance by $\mu_{\tau s}$ and $\sigma_{\tau s}^2$. Think of τ_{si} as the effective subsidy received by the firm. I also denote the mean and variance of the marginal cost by μ_{cs} and σ_{cs}.

Before presenting the results, it is useful to discuss the intuitions behind the estimation. The estimation is constructed in a parsimonious and effective manner. The conditional density is a function of firm productivity and distortions. The estimation of productivity parameters (μ_A, σ_A^2) is straightforward. The estimation of parameters associated with distortions is subtler. To estimate σ_K^2, σ_Y^2, and σ_{KY} from a linear combination of capital and output distortions, I reply on the fact that 1, α_s, and α_s^2 are independent polynomials. To see how this is relevant, notice that the estimator can determine $\sigma_{\tau s}^2 = \sigma_Y^2 - 2\alpha_s\sigma_{KY} + \alpha_s^2\sigma_K^2$ with precision. Once this is done, $\sigma_{\tau s}$ can be decomposed into three components $(\sigma_K^2, \sigma_Y^2, \sigma_{KY})$ when 1, α_s, and α_s^2 are independent polynomials.

This decomposition strategy also depends on the variability of industry-specific capital share of output. Fortunately, there is a great deal of variation in the data. Table 2.3 reports the summary statistics of the capital shares for those industries that have a counterpart in the United States (NBER Productivity Database). There are more than 80 matched industries in each year. Capital shares tend to center on 0.57 and have a standard deviation of about 0.17. The minimum capital share is 0.031 and the maximum is 0.95. The large variation in capital share reassures me that the decomposition strategy would be effective.

Table 2.4 reports the estimates of the distributional parameters. The estimated parameters are broadly consistent with the data patterns. In particular, both capital and output distortions (in log terms) have a positive mean. Capital distortion is more dispersed than output distortion. Interestingly, capital distortion is positively correlated with

Table 2.3: Summary statistics for matched capital elasticity

Year	Mean	Median	S.D.	Min	Max	#obs
2001	0.571	0.548	0.174	0.031	0.950	82
2004	0.572	0.550	0.173	0.031	0.950	85
2007	0.569	0.546	0.177	0.031	0.950	81

Table 2.4: Maximum likelihood estimation of parameters

Year	μ_A	μ_K	μ_Y	σ_A^2	σ_K^2	σ_Y^2	σ_{KY}	#obs
2001	-0.412***	3.038***	0.539***	0.689***	3.237***	0.361***	0.899***	17887
	(0.0119)	(0.0448)	(0.0226)	(0.0112)	(0.137)	(0.0238)	(0.0580)	
2004	-0.344***	3.212***	0.671***	0.613***	3.901***	0.450***	1.171***	27609
	(0.00818)	(0.0333)	(0.0156)	(0.00765)	(0.109)	(0.0190)	(0.0456)	
2007	-0.321***	3.314***	0.740***	0.593***	3.982***	0.438***	1.172***	24475
	(0.00827)	(0.0381)	(0.0175)	(0.00773)	(0.123)	(0.0192)	(0.0486)	

Notes: ** 5%, *** 1% levels of significance.

output distortion, implying that firms that receive more output subsidy also face higher capital tax. From previous regressions, we also know that these firms tend to be low in terms of productivity and small. Hence, the result is consistent with the observation that small firms in India tend to receive government subsidies, but also face high cost of financing capital expenditure due to credit frictions. Meanwhile, productivity (in log terms) has a slightly negative mean and is more dispersed than output distortion but less than capital distortion.

2.4.2 *Experiment 1: Elimination of Distortions*

With the distributional parameters in hand, I can now perform the first counterfactual experiment. The goal of the experiment is to assess how much of the dispersion in revenue TFP would remain after eliminating capital and output distortions from the economy. The graphs in this section refer to the year 2004 unless otherwise stated.

The key step of the counterfactual experiment is to estimate the counterfactual choke prices. Figure 2.4 shows the relationship between the actual and counterfactual choke prices for a total of 107 industries.

Clearly, choke prices tend to decrease after removing capital and output distortions. Intuitively, firms become more productive after the removal of distortions, and hence the industries become more competitive.

One important statistic in the counterfactual experiment is the dispersion in revenue TFP. Table 2.5 reports the dispersion of revenue TFP before and after removing distortions. It also reports the ratio of revenue TFP dispersion before and after removing distortions. This is reported as the fraction of revenue TFP dispersion remaining.

For the year 2004, removing distortions reduces the dispersion in revenue TFP from 0.338 to 0.284. That is, approximately 84% of the

Figure 2.4: Counterfactual: Choke price response to removal of non-markup distortions

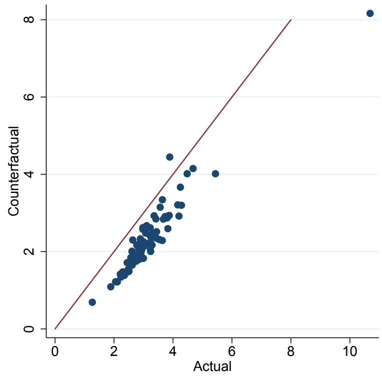


Table 2.5: Counterfactual: TFPR dispersion upon removing non-markup distortions

TFP Dispersion	2001	2004	2007
Actual	0.362	0.338	0.330
Counterfactual	0.301	0.284	0.281
Remaining (%)	83.1	83.9	85.2

dispersion in revenue TFP remains in the economy. Hence, removing distortions has only a modest effect in improving resource allocation.

Interestingly, although markups account for about one third of the dispersion in revenue TFP in the data (see section 2.4.4), they alone can generate substantial dispersion in revenue TFP. This result suggests that distortions somehow reduce the variation in markups.

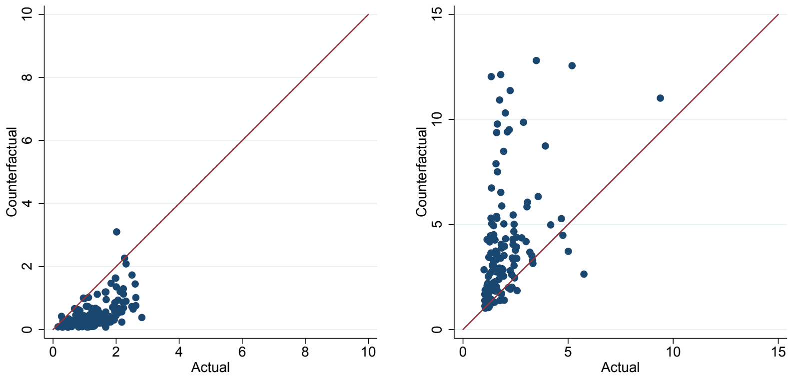
To understand the reason behind this, it is useful to examine the raw data. Table 2.6 reports the results from OLS regressions of firm productivity on capital and output distortions. The results show that more productive firms tend to receive less output subsidies and face more capital taxes.

Table 2.6: OLS regression of productivity on distortions

Dependent variable: Productivity			
	(1) 2001	(2) 2004	(3) 2007
Capital distortion	0.423***	0.420***	0.444***
	(0.055)	(0.058)	(0.051)
Output distortion	-1.385***	-1.493***	-1.648***
	(0.140)	(0.150)	(0.146)
Industry FE	Yes	Yes	Yes
N	17887	27609	24475
R^2	0.307	0.310	0.340

Notes: All variables are in log terms. ** 5%, *** 1% levels of significance. Standard errors are clustered at the industry level.

Figure 2.5: Counterfactual: Marginal cost and markup response to removal of non-markup distortions



(a) Marginal cost (b) Markup

This observation is consistent with the predictions of the model. In particular, the selection rule predicts that a firm chooses to operate only if its marginal cost is lower than the choke price. The firm has a low marginal cost either because it has high physical productivity (but unfavorable distortions such as capital tax), or because it receives favorable distortions such as output subsidy (but low physical productivity). When distortions are removed, productive firms are free of unfavorable distortions and therefore they charge even higher markups.

This phenomenon can be seen clearly from Figure 2.5. Figure 2.5 (a) plots the marginal costs before and after removing distortions for firms in industry 1513 (processing and preserving of fruit, vegetables, and edible nuts). After removing distortions, marginal cost tends to be lower for most of the firms in this industry. Figure 2.5 (b) plots the markups before and after removing distortions for firms in the same industry. Most firms increase their markups. This effect is particularly strong for productive firms which exhibit high markups before removing distortions. Hence, removing distortions tends to increase the dispersion of markups.

Another important statistic is the gains in aggregate productivity (TFP gain). TFP gain is defined as the percentage change in aggregate productivity after removing distortions, 100 times $(\tilde{TFP}/TFP - 1)$.

Table 2.7 reports the TFP gains for the linear demand model (VES) versus the CES model. It is useful to compare the results because the CES model has no variation in markups. The data for 2004 reveals that removing distortions enhances aggregate productivity by approximately 14% (VES). The TFP gain is substantially higher (about 87%) in the CES model because there is no resource misallocation from variable markups. This pattern is robust across different sampling years.

To summarize, three are two points that are particularly noteworthy. First, variable markups are an important source of resource misallocation. It induces resources to be allocated away from high- to

Table 2.7: Counterfactual: TFP gain from removing non-markup distortions

TFP Gain	2001	2004	2007
VES model	16.6	14.4	13.7
CES model	88.7	87.3	86.1

Table 2.8: Counterfactual: TFPR dispersion upon removing non-markup distortions and lowering entry barriers

TFP Dispersion	2001	2004	2007
Actual	0.362	0.338	0.330
Counterfactual	0.266	0.253	0.249
Remaining (%)	73.5	74.8	75.6

low-markup firms. If productive firms charge high markups, then variable markups imply a systematic misallocation of resources which could substantially lower aggregate productivity. Second, variable markups are endogenously determined by both productivity and distortions. Removing distortions may not have the intended effect because markups may respond accordingly. The counterfactual experiment in this section shows that the effect of removing distortions is very modest.

2.4.3 *Experiment 2: Remove Distortions and Lower Entry Barriers*

If removing distortions does not have a large effect as expected, what could be done to enhance aggregate productivity? Since variable markups affect resource misallocation, it is possible to enhance aggregate productivity by reducing variation in markups. One way to achieve this is to increase market competition by lowering entry barriers.

I therefore perform a second counterfactual experiment, in which entry costs are halved in addition to removing capital and output distortions. The counterfactual choke prices can be estimated using the following free-entry condition:

$$\int_0^{\tilde{P}_{\max,s}} \pi_{si}\left(C_{si}, \tilde{P}_{\max,s}\right) d\tilde{G}_s(C_{si}) = \frac{1}{2} f_{E,s}$$

Table 2.8 reports the dispersion in revenue TFP before and after the entry barriers are lowered. In 2004, the dispersion of revenue TFP changes from 0.342 to 0.220. There is still a significant portion of dispersion in revenue TFP remains (about 64%) but it is lower than before (about 74%). This suggests that increasing market competition is effective in improving resource allocation.

Table 2.9 reports the TFP gains from reducing entry barriers by a half and removing distortions. In 2004, the TFP gain is about 22%, which is substantially higher than before (about 14%). However, the TFP gain is still far lower than what a CES model would predict

Table 2.9: Counterfactual: TFP gain from removing non-markup distortions and lowering entry barriers

TFP Gain	2001	2004	2007
VES model	25.1	22.1	22.3
CES model	88.7	87.3	86.1

Table 2.10: Variance decomposition of TFPR dispersions

	Variance Decomposition				Contribution (%)	
Year	TFPR	Markup	MRP	Covariance	Markup	MRP
2001	0.475	0.257	0.349	-0.065	40.3	59.7
2004	0.419	0.231	0.331	-0.071	38.1	61.9
2007	0.392	0.209	0.324	-0.070	35.3	64.7

(approximately 87%), implying that the CES assumption may lead to a significant over-estimation of TFP gains.

2.4.4 *Naive TFPR Decomposition*

In this section, I describe naive decompositions of the dispersion in revenue TFP. The goal is to show that naive decompositions yield misleading results, hence providing justifications for the use of counterfactual experiments.

Let $T\hat{FPR}_{si}$ denote the log deviation of revenue TFP from its industry mean value. The log deviation of TFPR can be expressed as the sum of the log deviations of markup and revenue marginal products (MRP).

$$T\hat{FPR}_{si} = \hat{\mu}_{si} + M\hat{R}P_{si}$$

The dispersion in revenue TFP can be measured by the variance of $T\hat{FPR}_{si}$. The dispersion of revenue TFP can be decomposed into three components: dispersion in markup, dispersion in revenue marginal products, and two covariance terms. Revenue marginal products capture the direct effect of exogenous distortions on resource misallocation.

$$Var\left(TF\hat{P}R_{si}\right) = Var\left(\hat{\mu}_{si}\right) + Var\left(M\hat{R}P_{si}\right) + 2Cov\left(\hat{\mu}_{si}, M\hat{R}P_{si}\right)$$

The decomposition results are reported in Table 2.10. The results seem to suggest that deviation in markup is half as important as deviation in revenue marginal products.

Another way to quantify the relative importance of the two sources of TFPR dispersion is the following. Assign one of the covariance terms to markup and the other to revenue marginal product. Then, it is possible to decompose the TFPR dispersion in two components: contribution by markup dispersion and contribution by MRP dispersion. Using this decomposition, Table 2.10 suggests that markup accounts for about one third of the TFPR dispersion, while revenue marginal products accounts for about two thirds of the TFPR dispersion. Once again, the decomposition suggests that markup is less important than revenue marginal products.

To summarize, naive decompositions tend to conclude that variable markups are less important than exogenous distortions in explaining resource misallocations.

Why do naive decompositions provide misleading results? The reason is that they ignore the endogenous response of markups to changes in exogenous distortions. When distortions are correlated with productivity, endogenous response of markups could offset the direct effect from removing distortions, as demonstrated in the previous sections.

2.5 ROBUSTNESS ANALYSIS

The results are startling. Due to the endogenous response of markups, removing non-markup distortions has only a modest impact on misallocation and aggregate productivity, despite the fact that these distortions are substantial. A natural question then is: How robust are these findings? This section performs two robustness checks: (1) including a labor distortion and (2) using native capital production elasticities.

2.5.1 *Labor Distortions*

The benchmark analysis includes two types of exogenous distortions: a tax on capital and a subsidy on output. Is it possible that the large role of markup is due to the omission of labor distortions? To address this concern, I include a labor distortion 1+tau-l to the benchmark model. The profit of the firm is now given by:

$$\pi_{si} = (1 - \tau_{Ysi}) P_{si} Y_{si} - (1 + \tau_{Lsi}) w L_{si} - (1 + \tau_{Ksi}) R K_{si}$$

The methodology spelled out in the previous section broadly goes through, except for three amendments. First, labor distortion needs to be added to the following expressions: optimal capital intensity, marginal cost, revenue marginal product of capital, and labor. Second, labor distortions must be inferred from the data. Third, the distributional assumption and the estimation procedure need to be modified. The first modification is straightforward. I will explain the next two modifications in what follows.

To infer labor distortions, I use the following variables from the ASI data: labor compensation and employment. I use employment as a measure for labor input, and labor compensation as a measure for distortion inclusive wage payment $(1 + \tau_{Lsi}) w L_{si}$. Labor distortions can be recovered by dividing labor compensation by employment.

Measuring labor input by employment means that any differences in hours worked and human capital per worker across plants are treated as labor distortions. Hence, the labor distortion model is likely to overstate the effect of labor distortions.

With labor distortions in hand, I can back out other variables using the same formula as in the benchmark analysis, except that the term

Table 2.11: Robustness: TFPR dispersion with additional labor distortions

TFP Dispersion	2001	2004	2007
Actual	0.421	0.399	0.395
Counterfactual	0.327	0.315	0.316
Remaining (%)	77.6	78.8	80.1

Table 2.12: Robustness: TFP gain with additional labor distortions

TFP Gain	2001	2004	2007
VES model	22.5	18.8	18.4
CES model	106.3	102.4	106.2

wL_{si} needs to be replaced with distortion-inclusive wage payment $(1+\tau_{Lsi})wL_{si}$.

The modification of the estimation procedure is the following. Assume that productivity, capital, labor, and output distortions are jointly log-normally distributed. As before, assume that productivity is ex-ante uncorrelated with distortions. Denote the means of the distortions by μ_K, μ_L, and μ_Y. Denote the variances of the distortions by σ_K^2, σ_L^2, and σ_Y^2, and their covariances by σ_{KL}, σ_{KY}, and σ_{LY}. Then, the set of distributional parameters to be estimated are given by $(\mu_A, \mu_K, \mu_L, \mu_Y, \sigma_A^2, \sigma_K^2, \sigma_L^2, \sigma_Y^2, \sigma_{KL}, \sigma_{KY}, \sigma_{LY})$.

The conditional density function is still given by (2.1), except that the second term is replaced by a tri-variate normal density for $\log(1+\tau_{Ksi})$, $\log(1+\tau_{Lsi})$, and $\log(1-\tau_{Ysi})$. The rest of the procedure is the same as in the benchmark analysis. In the interest of space, I will not repeat the procedure. Instead, I turn now to reporting the results.

Table (2.11) reports the dispersion of revenue TFP before and after removing exogenous distortions. It shows that after removing non-markup distortions, approximately 80% of the TFPR dispersion still remains in the market. Table (2.12) further reports the TFP gains from the removal of non-markup distortions. If there were no variable markups (CES), aggregate productivity would increase by more than 100%. With variable markups, however, the TFP gains reduce to about 18.8%. Once again, CES models with constant markups significantly overestimate productivity gains. Also, the modest effects from removing non-markup distortions on misallocation and aggregate productivity are consistent with the findings in the baseline analysis.

2.5.2 *Capital Production Elasticities*

Now I turn to the second robustness analysis. In the benchmark analysis, I have followed Hsieh and Klenow (2009) to use capital production elasticities in the corresponding U.S. industry. This approach avoids potential mismeasurement of capital elasticities when there are industry-specific capital distortions. However, it also raises the concern about the appropriateness of applying the U.S. elasticities

Table 2.13: Summary statistics for estimated capital elasticity

Year	Mean	Median	S.D.	Min	Max	#obs
2001	0.525	0.530	0.118	0.231	0.806	100
2004	0.500	0.510	0.120	0.165	0.781	107
2007	0.547	0.523	0.120	0.310	0.877	112

to India. If production technology in India differs systematically to the United States, then the benchmark analysis would overestimate capital distortions. To address this concern, I now estimate the capital production elasticities using the India ASI data and use these native elasticities to perform counterfactual experiments.

To estimate capital production elasticities, I use the proxy method developed by Olley and Pakes (1996) and subsequently refined by Levinsohn and Petrin (2003).

The starting point is the revenue function:

$$R_{si} = P_s\left(Z_{si}\right) Z_{si} K_{si}^{\beta_{Ks}} L_{si}^{\beta_{Ls}} e^{u_{si}}$$

Where $Z_{si} = A_{si}\left(1 - \tau_{Ysi}\right) / \left(1 + \tau_{Ksi}\right)^{\alpha_s}$ is the effective productivity and u_{si} is measurement errors. For the purpose of exposition, it is more convenient to index firms by their effective productivity rather than marginal cost. Taking logs gives:

$$r_{si} = \beta_{Ks} k_{si} + \beta_{Ls} l_{si} + z_{si} + p_s(z_{si}) + u_{si}$$

where lower case means that the variable is in log terms.

Assume that labor is a variable input but capital is predetermined. This means that labor is likely to be correlated with a firm's effective productivity. The key insight by Olley and Pakes (1996) and Levinsohn and Petrin (2003) is that if more productive firms also use more materials, then it is possible to use a control function of capital and materials as a proxy for productivity (and price). To the extent that material costs are undistorted, the condition that more productive firms use more materials is likely to hold. Hence, I estimate the capital production elasticities with the following Olley-Pakes estimator:

$$r_{si} = \beta_{Ls} l_{si} + \phi(m_{si}, k_{si}) + u_{si}$$

where the control function ϕ is a 4th-order polynomial of capital k_{si} and materials m_{si}. The summary statistics of the estimated capital elasticities are given in Table (2.13).

With the capital elasticities in hand, I now perform the first policy experiment again. Table (2.14) reports the dispersion of revenue TFP before and after removing exogenous distortions. More than 85% of the TFPR dispersion remains in the market after the removal of non-markup distortions. This result is even stronger than the benchmark case. Table (2.15) further reports the resulting TFP gains. The TFP gain with variable markups (VES) is approximately 14%, which is slightly lower than the benchmark result. Once again, a CES model with constant markups significantly overestimates the TFP gain, implying

Table 2.14: Robustness: TFPR dispersion using estimated capital elasticities

TFP Dispersion	2001	2004	2007
Actual	0.306	0.284	0.297
Counterfactual	0.260	0.246	0.253
Remaining (%)	85.0	86.1	86.0

Table 2.15: Robustness: TFP gain using estimated capital elasticities

TFP Gain	2001	2004	2007
VES model	15.4	14.4	14.8
CES model	77.5	75.5	79.4

that removing distortions would enhance aggregate productivity by 76%.

To summarize, the two robustness checks reassure me that the main findings of the paper are unlikely to be driven by model specifications.

2.6 Conclusions

In this paper, I have shown that variable markups are a significant source of resource misallocation. Not only do variable markups account for one third of the observed misallocation, but they also endogenously respond to changes in the economy, offsetting gains from removing other types of non-markup distortions. Competition policies can reduce markup dispersion and improve market allocative efficiency. Such policies are more effective and implementable than removing firm-specific taxes and subsidies to improve misallocation and enhance aggregate productivity.

The study also cautions against the use of CES preferences (with constant markups) in the context of studying misallocation. Studies based on CES preferences may provide misleading results. In particular, models with CES preferences could significantly overstate the gains from removing distortions. The reason is that CES models lead to mismeasurement of non-markup distortions, and also ignore the endogenous response of markups to changes in the economy. Instead, models that explicitly account for variable markups should be used.

There are also limitations in the current study. Two aspects are particularly noteworthy. First, as with many studies, this paper focuses on a specific source of misallocation, namely markup dispersion. There are many other potential sources too. The natural question then is: which sources of misallocation are relatively important in explaining variation in the data? In this regard, studies along the line of David and Venkateswaran (2017) would be particularly useful.

Second, quantitative studies on markup distortions typically reply on a particular form of market structure. For example, in Peters (2016) this was Bertrand competition, in Edmond, Midrigan, and Xu (2015) this was oligopolistic competition, and in the current paper this is quasi-linear demand with monopolistic competition. This variation raises the question regarding how much the quantitative findings depend on the specific assumption of the market structures. A more ambitious question is whether there is a general way to quantify the impact of markup distortions on aggregate productivity without specific details on the market structure. These questions are left for future research.

3

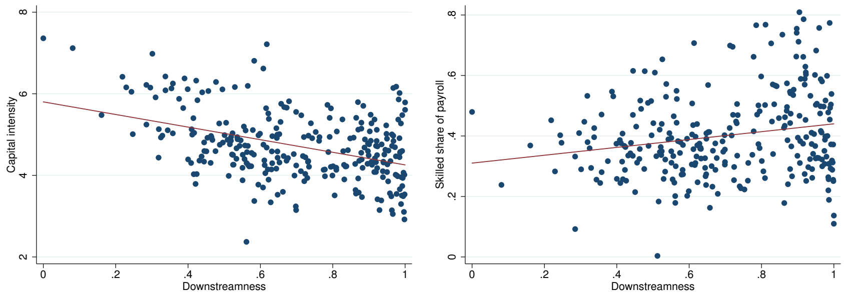
HETEROGENEOUS CAPITAL ALONG PRODUCTION CHAINS

Using industry-level data from the NBER-CES Manufacturing Database, this study documents three stylized facts about production chains that substantiate the need for measuring quality differences in capital. Firstly, downstream industries are less capital intensive. Secondly, downstream industries are more skill intensive. Thirdly, capital share of output does not vary systematically according to production line position. This analysis presents a stylized model of production chains with heterogeneous capital that can account for all of these three facts. The developed model sheds light on the role of capital quality in determining the pattern of and the reward to factors along production chains. The findings have important implications for policy design regarding how to remain competitive in the global economy.

3.1 INTRODUCTION

How are capital goods of differing quality allocated to production chains? Traditionally, production chains have been studied in the domain of international trade. Macroeconomic literature tends to treat industries symmetrically and as abstract from the interconnection of production. More generally, this literature places increased emphasis on differences in human capital while treating physical capital as homogenous goods, with only a few exceptions (Caselli and D. Wilson, 2004; Eaton and Kortum, 2001). This paper documents new stylized facts that substantiate the need for incorporating heterogeneous capital into production chain models.

Figure 3.1: Factor intensity and production line position



(a) Capital intensity (b) Skill intensity

Recent advances in measurement of production line positions make it possible to characterize production chains. Using industry-level data from the NBER-CES Manufacturing Database (E. J. Bartelsman and Gray, 1996), the patterns of factor allocation and factor prices along the production chains are documented. The unit of observation is an industry. Figure 3.1 (a) shows the relationship between capital intensity and production line position ("downstreamness"). Here, downstreamness is measured by the average position in the production chain at which an industry's output is used (Antràs, Chor, et al., 2012). Capital intensity is the log of capital per worker. The diagram shows that capital intensity declines towards the end of the production chain. Figure 3.1 (b) examines how skill intensity varies with production line position. Here, skill intensity is defined as the share of skilled labor income (non-production workers) in total labor income (production and non-production workers). The diagram illustrates that downstream industries are more skill intensive. These diagrams raise the first question: Why would capital and skill intensity vary systematically with production line position?

Figure 3.2: Factor rewards and production line position

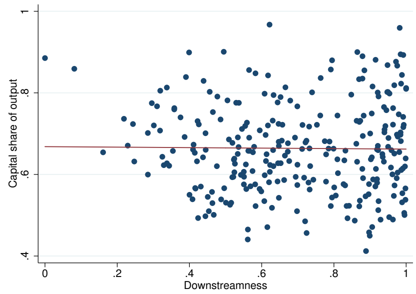


Figure 3.2 examines the relationship between capital share of output and production line position. Capital share of output is defined as payment to capital as a share of total industry value added. Surprisingly, capital share of output does not vary systematically with production line position. How is it possible that downstream industries use more labor and also more skilled labor without compensating workers proportionately? The only plausible explanation seems to be that capital goods have differing quality and that downstream industries tend to use high-quality capital goods. Here, variation in capital quality can be interpreted as the differences in productivity or costs of capital goods across industries.

This paper incorporates heterogeneous capital into a model of production chains. The model builds on the previous literature on organization of firms (Garicano, 2000) and matching between heterogeneous factors and tasks (Costinot and Vogel, 2010). The basic idea is the following. As moving towards downstream, products become more and more complex. This change requires workers to solve more challenging problems. Hence, downstream industries tend to hire more skilled workers to solve problems. Moreover, skill is strongly complementary to the quality of capital. Thus, downstream industries tend to deploy high-quality capital goods. My model can account for the stylized facts discussed at the beginning of the paper.

This work's first contribution is to substantiate the need to document differences in capital goods. Capital can be different for many reasons. First, there are different types of capital goods. Caselli and D. Wilson (2004) find that there is enormous cross-country difference in the composition of capital goods. Second, the cost of capital goods can differ. Eaton and Kortum (2001) show that trade cost generates substantial variation in the cost of capital goods across countries. To the extent that different industries have different composition of capital goods, there is substantial variation too in the cost of capital goods across industries. Third, capital goods can have different productivity across countries and industries. For example, in Australia, land is more productive for producing wine than rice, while in Thailand the reverse is true.

The second contribution of this analysis is to highlight the role of capital quality in shaping the pattern of and the reward to factors along production chains. Specifically, it is shown that downstream industries tend to use more skilled labor and high-quality capital goods. This observation is particularly relevant today as countries are increasingly specializing in global production chains. The proposed model predicts that countries with more high-quality capital goods tend to specialize in downstream industries. Since capital is more mobile than labor across countries, my model also predicts that high-skilled countries tend to attract more investments in high-quality capital goods (while specializing in downstream industries).

This paper builds upon a rapidly growing body of literature using assignment models in various contexts. The central theme of these papers is to offer specific comparative static predictions under strong functional form assumptions on the pattern of substitution across goods. Assignment models have been used in the context of international trade (Costinot, 2009; Costinot and Vogel, 2010; Sampson, 2014, 2016; Yeaple, 2005), offshoring (Antràs, Garicano, and Rossi-Hansberg, 2006; G. M. Grossman and Rossi-Hansberg, 2008), and management (Garicano, 2000; Garicano and Rossi-Hansberg, 2006). Among previous papers, this one is most closely related to the work by Costinot and Vogel (2010), who analyze the determinants of factor allocation and factor prices in economics with a large number of goods and factors. This paper's contribution to the literature is to study a new assignment problem: how capital goods of differing quality are allocated to different positions of production chains. The results are useful for understanding how countries specialize in global production chains.

This paper is at the cross-roads of several other works. It is clearly related to the literature on global production chains. A first strand of this literature has developed methodologies for credibly estimating the volume of cross-border production activity by combining data on trade and input output tables (R. C. Johnson and Noguera, 2012; Koopman, Z. Wang, and Wei, 2014), as well as for measuring the positions of countries and industries in the production chains (Antràs, Chor, et al., 2012; Fally, 2012). A second strand this literature has developed is the theoretical models of production chains to account for firms' location decisions (Antràs, Fort, and Tintelnot, 2017; Antràs and Gortari, 2017) and internalization decisions (Alfaro, Antràs, et al., 2017; Antràs and Chor, 2013; Fally and Hillberry, 2015), volume of trade (Yi, 2003), and pattern of trade (Costinot, Vogel, and S. Wang, 2013). This paper's contribution to this literature is to develop a new model of production chain based on assignments of heterogeneous factors to production line positions.

This work is related to the tradition on embodied technology, which emphasize differences in R&D contents across different types of capital (Caselli and D. Wilson, 2004; Eaton and Kortum, 2001; Sampson, 2016; D. J. Wilson, 2002), or differences in the efficiency units delivered by different vintages of capital (Greenwood, Hercowitz, and Krusell, 1997; Jovanovic and Rob, 1997; Solow, 1960). The analysis substantiates the need to document differences in the quality of capital goods. It is shown that the homogenous capital assumption, which is used extensively in the literature, is unlikely to be consistent with new stylized facts about production chains.

This work is also related, though less closely, to recent literature on macroeconomic implications of intersectoral input-output linkages. The presence of multi-sector linkages tends to amplify the effect of resource misallocation (Jones, 2011, 2013), generate aggregate fluctuations from microeconomic idiosyncratic shocks (Acemoglu, Carvalho, et al., 2012; Baqaee, 2017; Gabaix, 2011), and transmit financial shocks across the economy (Altinoglu, 2017; Bigio and La'O, 2016). While the presented work also emphasizes the interconnection of industries, it takes a different approach by visualizing the production process as a chain rather than a network. Both approaches are useful for understanding the nature of production processes.

The rest of the paper is organized as follows. Section 3.2 documents the stylized facts in details. Section 3.3 presents a basic model of production chains with homogenous factors and discusses its usefulness and limitations. Section 3.4 extends the basic model to allow for heterogeneous capital goods to account for the stylized facts. Section 3.5 concludes. All proofs are provided in Appendix B.1.

3.2 EMPIRICAL MOTIVATION

The main data source is the NBER CES Manufacturing Industry Database (E. J. Bartelsman and Gray, 1996; Becker, Gray, and Marvakov, 2016). The data set provides information on total value added, capital stock, and labor income at the industry level. It also breaks capital stock down into plant and equipment, and labor income down into skilled labor income and unskilled labor income. The data is highly disaggregated at the 6-digit NAICS level, with a total of 473 industries. Data from 2002 are used, for which a measure of production line position at the same industry level is obtained.

Recent advances in measurement of production line positions make it possible to characterize production chains. This paper follows Antràs, Chor, et al. (2012) to construct a measure of industry "upstreamness" as the average distance of an industry's output from final use. It then normalizes the upstreamness measure to unit interval and computes the industry "downstreamness" as one minus the normalized upstreamness. This measure is constructed using the 2002 US Input-Output Tables provided by the Bureau of Economic Analysis (BEA). The 2002 I-O Tables report information on production linkages at a highly disaggregate level, namely 6-digit I-O industry codes. There are altogether 426 industries in the I-O Tables, of which 279 are manufacturing industries.

BEA provides a concordance between the codes used in the I-O tables and the Census NAICS industry classification. This enables merging of the two data sets. The merged data set has a total of 279 manufacturing industries. The summary statistics of key variables are given in Table 3.1. Using the merged data set, three stylized facts regarding production chains are documented.

Fact 3.1. *Downstream industries are less capital intensive.*

Table 3.1: Summary statistics

	Mean	Median	S.D.	Min	Max	N
Capital intensity	4.721	4.607	0.828	2.370	7.360	279
log Plant per worker	3.602	3.518	0.779	1.573	6.199	279
log Equipment per worker	4.298	4.244	0.896	1.771	6.993	279
Skill intensity	0.401	0.375	0.139	0.003	0.809	279
Capital share of output	0.664	0.658	0.111	0.412	0.967	279
Rent to capital	0.845	0.744	0.518	0.133	4.966	279
Downstreamness	0.699	0.699	0.223	0.000	1.000	279

Notes: Data for all variables except downstreamness are from the NBER CES Manufacturing Industry Database (E. J. Bartelsman and Gray, 1996). Capital intensity is the log of capital per worker. Skill intensity is the share of skilled (non-production) labor income in total labor income. Capital share of output is defined as payment to capital as a share of total industry value added. Rent to capital is payment to capital divided by capital stock. Downstreamness is measured by the average position in the production chain at which an industry's output is used (Antràs, Chor, et al., 2012).

Table 3.2: Capital intensity

	(1)	(2)	(3)
	Capital intensity	log Plant per worker	log Equipment per worker
Downstreamness	-1.545***	-1.220***	-1.752***
	(0.223)	(0.223)	(0.234)
Constant	5.802***	4.455***	5.524***
	(0.166)	(0.168)	(0.170)
N	279	279	279
R^2	0.173	0.122	0.190

Notes: Robust standard errors are in parentheses. ***p<0.01, **p<0.05, and *p<0.1.

Capital intensity here is defined as log of capital per worker. This is a widely used proxy for how intensively capital is used in production. Within manufacturing, downstreamness is negatively correlated with physical capital intensity (Table 3.2). This observation is also true when capital is broken down into plant and equipment. Both log plant per worker and equipment per worker are negatively correlated with downstreamness.

Fact 3.2. *Downstream industries are more skill intensive.*

Following the previous literature, skill intensity is defined as the share of skilled (non-production) labor income in total labor income. Note that skilled and unskilled labor is measured by labor income rather than employment, because hours worked per worker may vary systematically across production line positions. Think of labor income as the efficiency unit of labor. Table 3.3 shows that within manufacturing, downstreamness is positively correlated with skill intensity.

Fact 3.3. *Capital share of output does not vary systematically according to production line position.*

Table 3.3: Skill intensity

	(1)
	Skill intensity
Downstreamness	0.130***
	(0.0358)
Constant	0.310***
	(0.0240)
N	279
R^2	0.043

Notes: Robust standard errors are in parentheses. ***p<0.01, **p<0.05, and *p<0.1.

Table 3.4: Capital share of output

	(1)	(2)
	Capital share of output	Rent to capital
Downstreamness	-0.00603	1.009***
	(0.0314)	(0.112)
Constant	0.668***	0.139*
	(0.0222)	(0.0739)
N	279	279
R^2	0.000	0.188

Notes: Robust standard errors are in parentheses. ***p<0.01, **p<0.05, and *p<0.1.

Since downstream industries employ less capital and more labor, particularly more skilled labor, one would expect that capital share of output (value added) decreases with the downstreamness of industry. Surprisingly, as shown in Table 3.4, capital share of output does not vary systematically according to production line position. To understand the reason behind this, the rent paid to capital is computed and correlated with downstreamness. Table 3.4 shows that the rent is increasing in the downstreamness of industry. This result suggests that the efficiency embodied in capital is different across industries. In particular, the efficiency of capital goods is increasing in the downstreamness of industry.

In the next two sections, two stylized models will be presented to rationalize these stylized facts. In section 3.3, a model with homogenous factors is presented and discussed regarding its usefulness and limitations. In section 3.4, a model with heterogeneous capital is presented and shown to be consistent with the stylized facts provided here.

3.3 Homogenous factor model

All markets are perfectly competitive. There is only one final good, which is used as numeraire. Producing this final good requires a continuum of stage inputs indexed by m, m belongs to $[\underline{m}, \overline{m}]$. The final good producer assembles these inputs using a CES technology:

$$Y = \left(\int_{\underline{m}}^{\overline{m}} y_m^{\rho} dm \right)^{\frac{1}{\rho}}$$

Here, y_m is input at stage m, and $1/(1-\rho) > 0$ is the elasticity of substitution across inputs.

Producing stage goods requires capital and labor. The economy is endowed with a mass K of homogenous capital and a mass L of homogenous labor. Output of stage m is given by:

$$y_m = \left[k_m^{\sigma} + (A_m l_m)^{\sigma} \right]^{\frac{1}{\sigma}}$$

where k_m and l_m are the capital and labor employed at stage m, $1/(1-\sigma) > 0$ is the elasticity of substitution between capital and labor, and A_m is the labor productivity at stage m. Labor productivity is determined as follows.

Producing stage goods requires solving problems indexed by z_m. Problems are ordered so that the easiest problems correspond to the lowest value of z_m. The cumulative distribution of problems is given by $F(z_m)$, which is twice continuously differentiable and has a positive density $f(z_m)$ on the support $[0, \overline{z}]$. Production requires that problems z_m be drawn and solved. Workers can solve the problem at a time cost that is proportional to the interval of know-how required $[0, z_m]$. For example, the producer can have a fraction z_m of the workers focus on solving problems and the rest of $1 - z_m$ workers produce the output.

The frequency of problems that may arise is proportional to the number of stage, m. A worker at stage m confronts a mass m of problems. All of these problems must be solved to produce an input that is compatible with the inputs in the previous stages. The expected labor productivity at stage m is given by:

$$F(z_m)^m (1 - z_m)$$

The stage producer chooses know-how z_m to maximize labor productivity. The first order condition with respect to z_m is:

$$m \frac{f(z_m)}{F(z_m)} = \frac{1}{1 - z_m}$$

The key assumption here is that the cumulative distribution of problems is log-concave. Under this assumption, the left-hand side of the equation above decreases in z_m while the right-hand side increases in z_m. Hence, there is a unique solution for optimal know-how z_m. Furthermore, an increase of stage m shifts the left-hand side upward, which results in a greater value of z_m. The intuition is that downstream producers are likely to confront more problems and therefore allocate more workers to solve those problems. Since downstream producers allocate fewer workers to production, the labor productivity A_m is declining in stage m.

Cost minimization of the stage m producer gives the optimal capital to labor ratio:

$$\frac{k_m}{l_m} = \left(\frac{r}{w}\right)^{\frac{1}{\sigma-1}} A_m^{\frac{\sigma}{\sigma-1}}$$

where r is the rent paid to capital and w is the wage paid to labor. When $\sigma < 0$, that is the elasticity of substitution between capital and labor is less than one, capital intensity decreases in stage m. Hence, capital and labor are strong complements. Since downstream producers have lower labor productivity, they find it optimal to employ more labor to compensate for the productivity loss. It is evident that the capital share of output also declines in stage m.

The limitations of the homogenous factor model are apparent. If $\sigma < 0$, so that capital and labor are strong complements, both capital intensity and capital share of output decline in stage m. If $\sigma > 0$, both capital intensity and capital share of output increase in stage m. To account for the stylized facts, it is necessary to have $\sigma = 0$. However, if $\sigma = 0$, so that the elasticity of substitution between capital and labor is one, then neither capital intensity nor capital share of output varies with stage m because factor prices are constant across production line positions. Since labor is measured by efficiency units, capital used in different production stages must be different. The next section presents a model with heterogonous capital goods.

3.4 heterogeneous capital model

The economy is endowed with a mass K of capital goods with differing quality, θ. Let the $K(\theta)$ be the mass of capital goods with quality less than θ and suppose K has support on $\left[\underline{\theta}, \overline{\theta}\right]$. Think of θ as the input augmenting productivity term which captures the technology variation across different types of capital goods. Alternatively, consider $\theta K'(\theta)$ to be the quantity of capital goods measured in efficiency units and $1/\theta$ as the cost per efficiency unit of capital.

The most important insight from the basic model is that production know-how z increases in stage m. The final good will be rewritten as an aggregate in terms of z instead of m, while maintaining the key insight that z increases with m. In what follows, production line position will be indexed by z instead of m. z will also be referred to as both know-how and production line position.

$$Y = \left(\int_{\underline{\theta}}^{\overline{\theta}} y(z)^{\rho} \mu(z)\, dz\right)^{\frac{1}{\rho}} \tag{3.1}$$

Here, $y(z)$ is the output of stage z, and $\mu(z)$ is the measure of stage z.

Producing stage goods requires physical inputs and production know-how. The production function of stage z good is given by:

$$y(z) = \left[\int_{\underline{\theta}}^{\overline{\theta}} A(\theta, z)\, k(\theta, z)\, d\theta\right]^{\alpha} l(z)^{1-\alpha} \tag{3.2}$$

Here, $k(\theta, z)$ is the mass of quality theta capital, and $l(z)$ is the amount of labor employed at stage z. $A(\theta, z)$ is the Solow-neutral technology term, which is twice differentiable, strictly increasing and concave in both its arguments, and strictly log supermodular.

The key assumption here is that A is strictly log supermodular in θ and z:

$$\frac{\partial^2 \log A(\theta, z)}{\partial\theta\partial z} > 0$$

The intuition here is that capital quality and production know-how are strong complements. Higher quality capital goods can make better use of otherwise the same production know-how. For a CES function, log supermodularity requires that the elasticity of substitution between two inputs is less than one.

Strong complementarity between quality and know-how implies that in equilibrium, there is a strictly increasing matching function T mapping know-how to quality such that $k(\theta, z) > 0$ if and only if $z = T(\theta)$. In other words, there is positive assortative matching between capital quality and production know-how. High quality capital goods are associated with better production know-how. This result is formally stated in the following proposition.

Proposition 3.1. *In the competitive equilibrium, there is positive assortative matching between capital quality and production know-how.*

Given the matching function, the output of stage m is given by:

$$y(z) = \left[A\left(T^{-1}(z), z\right) k\left(T^{-1}(z), z\right)\right]^{\alpha} l(z)^{1-\alpha} \tag{3.3}$$

The matching function must ensure that capital markets clear. Obviously, it satisfies $T(\underline{\theta}) = \underline{z}$ and $T(\overline{\theta}) = \overline{z}$. The market for quality theta capital clears when:

$$K(\theta) = \int_{\underline{z}}^{T(\theta)} k\left(T^{-1}(z), z\right) \mu(z)\, dz$$

Differentiating the equation above with respect to theta and using equations (1) and (3) gives a differential equation for the matching function:

$$T'(\theta) = \frac{r(\theta) K'(\theta)}{\alpha\mu(z)\, p(z)^{\frac{\rho}{\rho-1}} Y} \tag{3.4}$$

where $r(\theta)$ is the rent paid to capital with quality θ, and $p(z)$ is the price for output at stage z.

Two features of the matching function are particularly noteworthy. First, the matching function depends on supply of capital goods $K'(\theta)$ and demand of capital goods $\mu(\theta)$. Second, the matching function affects the relative factor price $w/r(\theta)$. In later part of the section, it will be shown how changes in capital supply and demand affect the matching function, which in turn affects equilibrium relative factor price and factor intensity.

How is the equilibrium rent schedule determined? Since there is a one-to-one matching between capital quality and production know-how, it is possible to characterize the rent schedule by studying producers' choices of capital goods. In particular, the producer at stage z chooses capital with quality θ to minimize the unit production cost:

$$\left(\frac{1}{A(\theta,z)}\frac{r(\theta)}{\alpha}\right)^{\alpha}\left(\frac{w}{1-\alpha}\right)^{1-\alpha}$$

The first order condition is given by:

$$\theta\frac{r'(\theta)}{r(\theta)}=\theta\frac{A_\theta(\theta,z)}{A(\theta,z)} \tag{3.5}$$

The intuition here is that return to quality $\theta r'(\theta)/r(\theta)$ reflects how much quality has contributed to the increase of productivity.

Strong complementarity between quality and know-how implies that return to quality increases when workers have more production know-how:

$$\frac{\partial}{\partial z}\left(\theta\frac{r'(\theta)}{r(\theta)}\right)=\theta\frac{\partial^2\log A(\theta,z)}{\partial\theta\partial z}>0$$

Proposition 3.2. *In a competitive equilibrium, the matching function and rent schedule satisfy* (3.4) *and* (3.5).

The Cobb-Douglas production function implies that capital intensity $k(\theta,z)/l(z)$ is proportional to the relative factor price $w/r(\theta)$, which decreases in know-how z. Since z is also an index of the production line position, it follows that capital intensity is lower in downstream industries.

To summarize, the basic model has been extended to incorporate heterogeneous capital goods. The new model is consistent with all stylized facts discussed at the beginning of the paper. In particular, the model predicts that downstream industries hire more and pay more to skilled workers (z), operate at lower capital intensity ($k(\theta,z)/l(z)$), and that capital share of output does not vary across production line positions.

Let us now investigate how exogenous changes in factor supply and demand, captured by K and μ, affect factor allocation and factor prices. In each case, this work first analyzes how exogenous affects the matching function. The rent schedule is then used to analyze how changes in the matching function affect the variation of factor intensity along the production chain.

First, consider an exogenous change that increases the relative supply of high quality capital goods. Specifically, consider a shift in the distribution of capital quality K to $\tilde{K}$ such that:

$$\frac{\tilde{K}'(\theta')}{\tilde{K}'(\theta)}\geq\frac{K'(\theta')}{K'(\theta)},\quad\forall\theta'>\theta$$

The intuition here is that the proportion of capital goods with quality above theta is always higher under $\tilde{K}$ than under K, and therefore, the proportion of capital goods with quality below theta is always lower under $\tilde{K}$ than under K.

Proposition 3.3. *An increase in the relative supply of high-quality capital goods shifts the matching function downwards and implies less variation in capital intensity.*

Moving from K to $\tilde{K}$ shifts the matching function downwards. This flattens the matching function at low quality levels and makes it steeper at high quality levels. The intuition is that the increase in high quality capital goods must be absorbed in the production chain, and hence there is quality upgrading at each production line position.

The downward shift in the matching function implies that return to quality increases at a slower pace. Intuitively, holding fixed the demand for quality, an increase of the relative supply of high-quality goods lowers the relative return to high-quality capital goods. Hence, there is less variation in return to quality along the production chain. Since capital intensity monotonically depends on relative factor price $w/r\,(\theta)$, there is also less variation in capital intensity across production line positions.

Second, consider an exogenous change that increases the relative demand for high-quality capital goods. In particular, consider a shift in the distribution of know-how μ to $\tilde{\mu}$ such that:

$$\frac{\tilde{\mu}\,(\theta')}{\tilde{\mu}\,(\theta)} \geq \frac{\mu\,(\theta')}{\mu\,(\theta)}, \quad \forall \theta' > \theta$$

Think of this as a technological change that increases the demand for downstream output and hence for high-quality capital goods.

Proposition 3.4. *An increase in the relative demand of downstream goods shifts the matching function upwards and implies more variation in capital intensity.*

Moving from μ to $\tilde{\mu}$ shifts the matching function upwards. This shift flattens the matching function at high quality levels and makes it steeper at low quality levels. Intuitively, an increase in the relative demand of downstream goods means that more capital goods must be allocated to produce downstream goods. Hence, there is quality downgrading at each production line position.

The upward shift in the matching function implies that return to quality increases at a faster speed. Therefore, there is more variation in return to quality along the production chain. Once again, the Cobb-Douglas production function implies that capital intensity is proportional to relative factor price $w/r\,(\theta)$. It follows that there is more variation in capital intensity across production line positions.

3.5 conclusions

Accounting for quality differences in capital goods could be important for understanding production chains. This work has shown that incorporating heterogeneous capital into production chain models is necessary to rationalize the stylized facts. The developed model also implies that countries with more human capital tend to attract more high-quality capital goods and specialize in downstream industries.

This implication is particularly relevant to policy because countries are becoming increasingly specialized in global production chains.

APPENDIX TO CHAPTER 1

A.1 PROOFS OF PROPOSITIONS

A.1.1 *Proof of Proposition 1.1*

Total differentiating the expression of σ gives

$$d\ln\sigma = -\frac{\psi_V}{\psi_V-\psi_O}Bd\ln\psi_O + \frac{f_V}{f_V-f_O}Bd\ln f_O,$$

where $B = \theta(1-\alpha)/\alpha$. Total differentiating the expression of ψ_O yields

$$d\ln\psi_O = \frac{\psi_S}{\psi_O}\frac{d}{\lambda\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2}d\ln\lambda - \frac{\psi_S}{\psi_O}\frac{d\left(\frac{1-\lambda}{\lambda}\right)}{\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2}d\ln d,$$

where $\psi_S = (1-\alpha)(1-\beta)$. Total differentiating the expression of f_O gives

$$d\ln f_O = \frac{1}{f_O}\frac{\mu f^e}{\lambda\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2}d\ln\lambda + \frac{1}{f_O}\frac{d\left(\frac{1-\lambda}{\lambda}\right)^2\mu f^e}{\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2}d\ln d - \frac{1}{f_O}\frac{\left(\frac{1-\lambda}{\lambda}\right)\mu f^e}{1+d\left(\frac{1-\lambda}{\lambda}\right)}d\ln\mu$$

Substituting and collecting terms,

$$d\ln\sigma = -\vartheta_1 d\ln\lambda + \vartheta_2 d\ln d - \vartheta_3 d\ln\mu$$

The coefficients ϑ's are $\vartheta_1 = \frac{\psi_V}{\psi_V-\psi_O}\frac{\psi_S}{\psi_O}\frac{Bd}{\lambda\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2} - \frac{f_V}{f_V-f_O}\frac{1}{f_O}\frac{B\mu f^e}{\lambda\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2}$, $\vartheta_2 = \frac{\psi_V}{\psi_V-\psi_O}\frac{\psi_S}{\psi_O}\frac{Bd\left(\frac{1-\lambda}{\lambda}\right)}{\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2} + \frac{f_V}{f_V-f_O}\frac{1}{f_O}\frac{Bd\left(\frac{1-\lambda}{\lambda}\right)^2\mu f^e}{\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2}$, and $\vartheta_3 = \frac{f_V}{f_V-f_O}\frac{1}{f_O}\frac{B\mu\left(\frac{1-\lambda}{\lambda}\right)f^e}{1+d\left(\frac{1-\lambda}{\lambda}\right)}$. It is evident that ϑ_2 and ϑ_3 are strictly positive. It remains to show that ϑ_1 is also strictly positive. Note that $\sigma\in(0,1)$ implies $\frac{\psi_V}{\psi_V-\psi_O} > \frac{f_V}{f_V-f_O}$. Furthermore, Assumption 1.2 implies that equilibrium cutoff exceeds the ideal cutoff for which financial market is perfect. This in turn implies $\frac{\psi}{f} > \frac{\psi_O}{f_O}$, where $\psi = 1-\alpha(1-\beta)$. Applying these two inequalities gives

$$\begin{aligned}\vartheta_1 &> \frac{\psi_V}{\psi_V-\psi_O}\frac{Bd}{\lambda\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2}\left\{\frac{\psi_S}{\psi_O} - \frac{1}{f_O}\frac{\mu}{d}f^e\right\}\\ &> \frac{\psi_V}{\psi_V-\psi_O}\frac{Bd}{\lambda\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2}\left\{\frac{\psi_S}{\psi} - \frac{1}{f}\frac{\mu}{d}f^e\right\} > 0\end{aligned}$$

where the last inequality follows from Assumption 1.2.

I then show that σ is log-supermodular in λ and μ, $\frac{\partial^2\ln\sigma}{\partial\mu\partial\lambda} > 0$. Since $\frac{\partial^2\ln\sigma}{\partial\mu\partial\lambda} = \frac{\partial}{\partial\lambda}\left(\frac{\partial\ln\sigma}{\partial\lambda}\right) = -\frac{1}{\lambda}\frac{\partial\vartheta_3}{\partial\lambda}$, it suffices to show that $\frac{\partial\vartheta_3}{\partial\lambda} < 0$. Differentiating the expression of ϑ_3 gives

$$\frac{d\vartheta_3}{d\lambda} = \frac{f_V}{(f_V-f_O)^2}\frac{1}{f_O}\frac{B\mu\left(\frac{1-\lambda}{\lambda}\right)f^e}{1+d\left(\frac{1-\lambda}{\lambda}\right)}\frac{df_O}{d\lambda} - \frac{f_V}{f_V-f_O}\frac{1}{f_O}\frac{B\mu f^e}{\lambda^2\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2} - \frac{1}{f_O}\vartheta_3\frac{df_O}{d\lambda}$$

Since the third term is negative, it suffices to show that the sum of the first two terms is negative. Substituting the expression of ϑ_3 yields

$$\begin{aligned}\frac{d\vartheta_3}{d\lambda} &< \frac{f_V}{(f_V-f_O)^2}\frac{1}{f_O}\frac{B\mu\left(\frac{1-\lambda}{\lambda}\right)f^e}{1+d\left(\frac{1-\lambda}{\lambda}\right)}\frac{df_O}{d\lambda} - \frac{f_V}{f_V-f_O}\frac{1}{f_O}\frac{B\mu f^e}{\lambda^2\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2} \\ &= \frac{f_V}{(f_V-f_O)^2}\frac{1}{f_O}\frac{B\mu\left(\frac{1-\lambda}{\lambda}\right)f^e}{1+d\left(\frac{1-\lambda}{\lambda}\right)}\frac{\mu f^e}{\lambda^2\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2} - \frac{f_V}{f_V-f_O}\frac{1}{f_O}\frac{B\mu f^e}{\lambda^2\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2} \\ &= \frac{f_V}{f_V-f_O}\frac{1}{f_O}\frac{B\mu f^e}{\lambda^2\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2}\left\{\frac{d\left(\frac{1-\lambda}{\lambda}\right)}{1+d\left(\frac{1-\lambda}{\lambda}\right)}\frac{1}{f_V-f_O}\frac{\mu}{d}f^e - 1\right\}\end{aligned}$$

Denote by z_S the cutoff productivity for the supplier. By Assumption 1.2, $z_S < z_V$, which in turn implies $\frac{\psi_S}{\psi_V-\psi_O} - \frac{1}{f_V-f_O}\frac{\mu}{d}f^e > 0$. Hence

$$\frac{d\vartheta_3}{d\lambda} < \frac{f_V}{f_V-f_O}\frac{1}{f_O}\frac{B\mu f^e}{\lambda^2\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2}\left\{\frac{d\left(\frac{1-\lambda}{\lambda}\right)}{1+d\left(\frac{1-\lambda}{\lambda}\right)}\frac{\psi_S}{\psi_V-\psi_O} - 1\right\}$$

Since $\psi_V > \psi \geq \psi_O$, we have

$$\frac{\psi_S}{\psi_V-\psi_O} < \frac{\psi_S}{\psi-\psi_O} = \frac{\psi_S}{\psi_S - \psi_S/\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]} = \frac{1+d\left(\frac{1-\lambda}{\lambda}\right)}{d\left(\frac{1-\lambda}{\lambda}\right)}$$

It follows that $\frac{d\vartheta_3}{d\lambda} < 0$.

Finally, it remains to show that σ is log-submodular in λ and d, $\frac{\partial^2 \ln\sigma}{\partial d\partial\lambda} < 0$. It suffices to show that $\frac{\partial\vartheta_1}{\partial d} > 0$. Proving this inequality is difficult; therefore, I provide sufficient conditions under which it holds. Differentiate ϑ_1 with respect to d gives

$$\begin{aligned}\frac{d\vartheta_1}{dd} &= \frac{\psi_V}{\psi_V-\psi_O}\frac{\psi_S}{\psi_O}\frac{B\left[1-d\left(\frac{1-\lambda}{\lambda}\right)\right]}{\lambda\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^3} + \frac{f_V}{f_V-f_O}\frac{1}{f_O}\frac{2B\left(\frac{1-\lambda}{\lambda}\right)\mu f^e}{\lambda\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^3} \\ &+ \left(1-\frac{\psi_O}{\psi_V-\psi_O}\right)\frac{\psi_V}{\psi_V-\psi_O}\left(\frac{\psi_S}{\psi_O}\right)^2\frac{Bd\left(\frac{1-\lambda}{\lambda}\right)}{\lambda\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^4} \\ &+ \left(1-\frac{f_O}{f_V-f_O}\right)\frac{f_V}{f_V-f_O}\left(\frac{1}{f_O}\right)^2\frac{B\left[\left(\frac{1-\lambda}{\lambda}\right)\mu f^e\right]^2}{\lambda\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^4}\end{aligned}$$

Assume $\lambda > \frac{1}{2}$ so that $d\left(\frac{1-\lambda}{\lambda}\right) \in (0,1)$. Applying the inequalities $\frac{\psi_V}{\psi_V-\psi_O} > \frac{f_V}{f_V-f_O}$ and $\frac{f^e}{f} < \frac{d}{\mu}\frac{\psi_S}{\psi}$ gives a lower bound for the sum of first two terms, $\frac{f_V}{f_V-f_O}\frac{1}{f_O}\frac{B\frac{\mu}{d}f^e}{\lambda\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2}$. Assume also that $\psi_V > 2\psi_O$ and $f_V > 2f_O$. If follows that $\frac{\psi_O}{\psi_V-\psi_O} < 1$ and $\frac{f_O}{f_V-f_O} < 1$, and the last two terms are strictly positive. Hence the set of sufficient conditions for $\frac{\partial^2\ln\sigma}{\partial d\partial\lambda} < 0$ are $\lambda > \frac{1}{2}$, $\psi_V > 2\psi_O$, and $f_V > 2f_O$. This completes the proof of Proposition 1.1.

A.1.2 *Proof of Proposition 1.2*

Totally differentiating the expression of M^e gives

$$d\ln M^e = \frac{L_p}{L}\left(d\ln\phi_O - d\ln\left(\frac{\Omega_p}{\overline{\Omega}}\right)\right) + \frac{\overline{f}}{\overline{\Omega}f_O - \overline{f}}\left(d\ln\overline{\Omega} - d\ln\left(\frac{\overline{f}}{f_O}\right)\right)$$

First consider the terms in the second parentheses,

$$\begin{aligned}
d\ln\overline{\Omega} - d\ln\left(\frac{\overline{f}}{f_O}\right) &= \overline{\Delta}\left(\frac{B-1}{B}d\ln\sigma - \frac{\overline{D}}{\overline{D}-1}d\ln\phi_O\right) - \overline{\Theta}\left(d\ln\sigma - \frac{f_V/f_O}{f_V/f_O-1}d\ln f_O\right) \\
&= \overline{\Delta}\left(\frac{B-1}{B}d\ln\sigma - \frac{\psi_V}{\psi_V-\psi_O}d\ln\phi_O\right) - \overline{\Theta}\left(d\ln\sigma - \frac{f_V}{f_V-f_O}d\ln f_O\right)
\end{aligned}$$

where $\overline{\Delta} = \frac{(\overline{D}-1)\sigma_R}{1+(\overline{D}-1)\sigma_R}$, $\overline{D} = \frac{\phi_V}{\phi_O}(q(1-\beta))^{-\frac{\alpha}{1-\alpha}}$, and $\overline{\Theta} = \frac{(f_V/f_O-1)\sigma}{1+(f_V/f_O-1)\sigma}$ for notational convenience. Note that $(\overline{D}-1)\sigma_R = \left(\frac{\psi^V}{\psi^O}-1\right)\sigma^{1-\frac{1}{B}} = \left(\frac{f^V}{f^O}-1\right)\sigma$ implies $\overline{\Delta} = \overline{\Theta}$. We can now collect terms

$$\begin{aligned}
d\ln\overline{\Omega} - d\ln\left(\frac{\overline{f}}{f_O}\right) &= \overline{\Delta}\left(\frac{B-1}{B}d\ln\sigma - d\ln\sigma - \frac{\psi_V}{\psi_V-\psi_O}d\ln\phi_O + \frac{f_V}{f_V-f_O}d\ln f_O\right) \\
&= \overline{\Delta}\left(\frac{B-1}{B}d\ln\sigma - d\ln\sigma + \frac{1}{B}d\ln\sigma\right) = 0
\end{aligned}$$

Next consider the terms in the first parentheses,

$$\begin{aligned}
d\ln\phi_O - d\ln\left(\frac{\Omega_p}{\overline{\Omega}}\right) &= d\ln\phi_O - \Delta_p d\ln\sigma_R - \overline{\Delta}\left(d\ln\sigma_R - \frac{\overline{D}}{\overline{D}-1}d\ln\phi_O\right) \\
&= \left(1-\overline{\Delta}\frac{\overline{D}}{\overline{D}-1}\right)d\ln\phi_O - \left(\Delta_p - \overline{\Delta}\right)\frac{B-1}{B}d\ln\sigma
\end{aligned}$$

where $\Delta_p = \frac{(D_p-1)\sigma_R}{1+(D_p-1)\sigma_R}$, and $D_p = \frac{\alpha}{\alpha(1-\beta)}(q(1-\beta))^{-\frac{\alpha}{1-\alpha}}$. Note that $1-\overline{\Delta}\frac{\overline{D}}{\overline{D}-1} = \frac{1-\sigma_R}{(1-\sigma_R)+\overline{D}\sigma_R} > 0$. By assumption 1.3, $\Delta_p > \overline{\Delta}$ always holds. It follows from the proof of Proposition 1.1 that all statements in Proposition 1.2 are true.

A.1.3 *Proof of Proposition 1.3*

Totally differentiating the expression of sector price gives

$$d\ln P = -(B/\theta)\,d\ln M^e + d\ln\left(z_O^B\tilde{z}^{-1}\right)$$

By Proposition 1.2 the first term has the expected signs. It remains to show the second term also has the expected signs. Totally differentiating the second term gives

$$d\ln\left(z_O^B\tilde{z}^{-1}\right) = (B-1)\,d\ln z_O + \left(\tilde{\Delta}B/\theta\right)d\ln\sigma_R$$

where $\tilde{\Delta} = \frac{(D-1)\sigma_R}{1+(D-1)\sigma_R}$, $\tilde{D} = (q(1-\beta))^{-\frac{\alpha}{1-\alpha}}$. Totally differentiating the expression of z_O gives

$$d\ln z_O = -\frac{B}{\theta}d\ln\psi_O + \frac{B}{\theta}d\ln f_O$$

whereas totally differentiating the expression of σ_R yields,

$$d\ln\sigma_R = -\frac{\psi_V}{\psi_V-\psi_O}(B-1)\,d\ln\psi_O + \frac{f_V}{f_V-f_O}(B-1)\,d\ln f_O$$

Substituting and collecting terms yields,

$$d\ln\left(z_O^B\tilde{z}^{-1}\right) = (B-1)\,\frac{B}{\theta}\left\{-\left(1-\frac{\tilde{\Delta}\psi_V}{\psi_V-\psi_O}\right)d\ln\psi_O + \left(1-\frac{\tilde{\Delta}f_V}{f_V-f_O}\right)d\ln f_O\right\}$$

Substituting and using the expressions of $d\ln\psi_O$ and $d\ln f_O$ gives,

$$d\ln\left(z_O^B\tilde{z}^{-1}\right) = -\xi_1 d\ln\lambda + \xi_2 d\ln d - \xi_3 d\ln\mu$$

where the coefficients are given as follows: $\xi_1 = \left(1-\frac{\tilde{\Delta}\psi_V}{\psi_V-\psi_O}\right)\frac{\psi_S}{\psi_O}\frac{Bd}{\lambda\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2} - \left(1-\frac{\tilde{\Delta}f_V}{f_V-f_O}\right)\frac{1}{f_O}\frac{B\mu f^e}{\lambda\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2}$, $\xi_2 = \left(1-\frac{\tilde{\Delta}\psi_V}{\psi_V-\psi_O}\right)\frac{\psi_S}{\psi_O}\frac{Bd\left(\frac{1-\lambda}{\lambda}\right)}{\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2} + \left(1-\frac{\tilde{\Delta}f_V}{f_V-f_O}\right)\frac{1}{f_O}\frac{B\left(\frac{1-\lambda}{\lambda}\right)^2\mu f^e}{\left[1+d\left(\frac{1-\lambda}{\lambda}\right)\right]^2}$, and $\xi_3 = \frac{\tilde{\Delta}f_V}{f_V-f_O}\frac{1}{f_O}\frac{B\left(\frac{1-\lambda}{\lambda}\right)f^e}{1+d\left(\frac{1-\lambda}{\lambda}\right)}$. Note that $\phi_V > \phi_O$ implies $\tilde{D} < \overline{D}$, which in turn implies $\tilde{\Delta} < \overline{\Delta}$. It follows that $1-\frac{\tilde{\Delta}f_V}{f_V-f_O} > 1-\frac{\tilde{\Delta}\psi_V}{\psi_V-\psi_O} > 1-\frac{\overline{\Delta}\psi_V}{\psi_V-\psi_O} > 0$. It follows from the proof of Proposition 1.1 the all three coefficients ξ_i's are positive numbers. This completes the proof of Proposition 1.3.

A.1.4 *Proof of Proposition 1.4*

Totally differentiating the expression of sector TFP gives

$$d\ln TFP = d\ln TFPR - d\ln P$$

By Proposition 1.2 the second term has the expected signs. It remains to show the first term also has the expected signs. Totally differentiating the first term gives

$$d\ln TFPR = -\frac{\alpha(1-\beta)\,\Omega_p}{\alpha(1-\beta)\,\Omega_p+\phi_O\overline{\Omega}}\left(\Delta_p-\tilde{\Delta}\right)d\ln\sigma_R - \frac{\phi_O\overline{\Omega}}{\alpha(1-\beta)\,\Omega_p+\phi_O\overline{\Omega}}\left(\overline{\Delta}-\tilde{\Delta}\right)d\ln\sigma_R$$

By assumption 1.3, $\Delta_p > \overline{\Delta}$ always holds. Note $\phi_V > \phi_O$ implies $\tilde{D} < \overline{D}$, which in turn implies $\tilde{\Delta} < \overline{\Delta}$. It follows from the proof of Proposition 1.1 the first term has the expected signs. This completes the proof of Proposition 1.3.

A.2 EXTENSIONS

In this section, I develop three extensions of the benchmark model in the main text. The purpose here is not to developed a full-fledged model, but to check whether the mechanism of the baseline model continue to hold in different settings. The first extension considers the case when investments in inputs are partially contractible. The second extension goes beyond the transaction costs approach and adopts the more sophisticated property rights approach. The third extension allows the firm to outsource a fraction of its inputs while integrating the rest of them.

A.2.1 *Partial Contractibility*

In the baseline model, we have assumed that investments in inputs are fully noncontractible under outsourcing. Here, we consider the case when investments in inputs are partially contractible as in Acemoglu, Antràs, and Helpman (2007). Firm faces the demand $y = Ap^{-1/(1-\alpha)}$ which implies the revenue function $R = A^{1-\alpha}y^{\alpha} = A^{1-\alpha}z^{\alpha}x^{\alpha}$. The intermediate input x is produced by assembling a continuum of differentiated inputs using the Cobb-Douglas production function,

$$x = \exp\left[\int_0^1 \ln x\left(k\right) dk\right].$$

Assume that input $k \in [0, \eta)$ is fully contractible, while input $k \in [\eta, 1]$ is non-contractible. An integrating firm solves the same problem as in the benchmark model. Hence we focus on the case of outsourcing. Under outsourcing, the supplier solves the problem:

$$\begin{aligned}
\pi_S\left(z\right) = &\max_{\{x(k)\}_{k=\eta}^{1}, F(z)} \left(1-\beta\right) A^{1-\alpha}z^{\alpha}x^{\alpha} - \int_{\eta}^{1} x\left(k\right) dk - \left(1-d\right) T\left(z\right) - \lambda F\left(z\right) - \left(1-\lambda\right)\mu f^{e}. \\
s.t. \quad &\left(1-\beta\right) A^{1-\alpha}z^{\alpha}x^{\alpha} - \int_{\eta}^{1} x\left(k\right) dk - \left(1-d\right) T\left(z\right) \geq F\left(z\right), \\
&\lambda F\left(z\right) + \left(1-\lambda\right)\mu f^{e} \geq dT\left(z\right).
\end{aligned}$$

The firm offers a contract that specifies the amount of transfer $T\left(z\right)$ and the investment levels in contractible inputs $\{x\left(k\right)\}_{k=0}^{\eta}$. The transfer ensures the supplier can obtain external funds at minimum cost. Investments in contractible inputs maximize the profit of the firm while taking into account the under-investment problem of noncontractible inputs:

$$\pi_F\left(z\right) = \max_{\{x(k)\}_{k=0}^{\eta}} \beta A^{1-\alpha}z^{\alpha}x^{\alpha} - \int_0^{\eta} x\left(k\right) dk - f + T\left(z\right).$$

Solving this program yields the profit function of the firm under outsourcing $\pi_O\left(z\right) = \psi_O z^{\frac{\alpha}{1-\alpha}} - f_O$, where $\psi_O = \left(1-\alpha\right) A \left[\alpha^{\alpha}\left(1-\beta\right)^{\alpha(1-\eta)}\left(\frac{\phi_O}{1-\alpha(1-\eta)}\right)^{1-\alpha(1-\eta)}\right]^{\frac{1}{1-\alpha}}$, $\phi_O = \beta + \frac{(1-\alpha(1-\eta))(1-\beta)}{1+d\left(\frac{1-\lambda}{\lambda}\right)}$, and $f_O = f - \frac{\left(\frac{1-\lambda}{\lambda}\right)\mu f^{e}}{1+d\left(\frac{1-\lambda}{\lambda}\right)}$. Since $\psi_O \propto \phi_O^{\frac{1-\alpha(1-\eta)}{1-\alpha}}$ and $\frac{1-\alpha(1-\eta)}{1-\alpha} > 1$, we conclude that Proposition 1.1 continue to hold in the case of partial contractibility.

A.2.2 *Property Rights Approach*

In the benchmark model, suppliers are the only agents undertaking relationship-specific investments. Here, we follow the property rights approach which allows the firm to also make relationship-specific investments. Since neither the firm nor the supplier is the full residual reclaimant of the output gains derived from their investments, both agents tend to withhold investments. Ownership is assigned to the agent that undertakes the most important investment to mitigate the overall under-investment problem.

Firm faces the demand $y = Ap^{-1/(1-\alpha)}$ which implies the revenue function $R = A^{1-\alpha}y^{\alpha} = A^{1-\alpha}z^{\alpha}x^{\alpha}$. The intermediate input x is produced with a Cobb-Douglas technology,

$$x = \left(\frac{h}{\eta}\right)^{\eta}\left(\frac{m}{1-\eta}\right)^{1-\eta}.$$

The firm is responsible for undertaking investment in headquarter service h, and the supplier is responsible for making investment in component m. Investment are relationship-specific in the sense that agents cannot costlessly switch to alternative trading partners and are partially locked in the bilateral relationship. Hence the two agents negotiate the division of output gains derived from the investments, with the firm receives a fraction β_k of the gains. The division of gains depend on the organizational forms $k \in \{V, O\}$ where V refers to integration and O refers to outsourcing. For our purpose, it suffices to assume $\beta_V > \beta_O$.

The firm's problem with organizatinal form k is:

$$\pi_F(z) = \max_h \beta_k A^{1-\alpha}z^{\alpha}x^{\alpha} - h - f + T(z).$$

The supplier's problem is:

$$\begin{aligned}
\pi_S(z) = \max_{m,F(z)} & \left(1-\beta_k\right)A^{1-\alpha}z^{\alpha}x^{\alpha} - m - (1-d)\,T(z) - \lambda F(z) - (1-\lambda)\,\mu f^e. \\
s.t. \quad & \left(1-\beta_k\right)A^{1-\alpha}z^{\alpha}x^{\alpha} - m - (1-d)\,T(z) \geq F(z), \\
& \lambda F(z) + (1-\lambda)\,\mu f^e \geq dT(z).
\end{aligned}$$

The firm and the supplier choose investments in h and m simutaneously. Solving the intersection of the their best response functions and using the optimal investments gives the profit function under organizational mode k, $\pi_k(z) = \psi_k z^{\frac{\alpha}{1-\alpha}} - f_k$, where $\psi_k = \phi_k A\left(\frac{1}{z\beta_k^{\eta}(1-\beta_k)^{1-\eta}}\right)^{-\frac{\alpha}{1-\alpha}}$, $\phi_k = \beta_k(1-\alpha\eta) + \frac{(1-\beta_k)(1-\alpha(1-\eta))}{1+d\left(\frac{1-\lambda}{\lambda}\right)}$ and $f_k = f - \frac{\mu\left(\frac{1-\lambda}{\lambda}\right)}{1+d\left(\frac{1-\lambda}{\lambda}\right)}f^e$. It is straightforward to verify that financial development has a greater effect on ψ_V than ψ_O, and has the same effect on f_V and f_O. The same statement is true for external financial dependence and asset tangibility. It follows that Proposition 1.1 continue to hold with the property rights approach.

A.2.3 *Input-level Model*

In the baseline model, firms' sourcing decisions are a binary choice: either outsource all of inputs or integrate all of them. This section extends the baseline model to allow the firm to outsource a fraction of the inputs while integrating the rest of them. The purpose here is to derive predictions on the prevalence of outsourcing at the firm level, which will be tested with firm-level data in the next section.

Firm faces the demand $y = Ap^{-1/(1-\alpha)}$ which implies the revenue function $R = A^{1-\alpha}y^{\alpha} = A^{1-\alpha}z^{\alpha}x^{\alpha}$. The intermediate input x is produced by assembling a continuum of differentiated inputs using the CES production function,

$$x = \left(\int_0^1 x\left(k\right)^{\rho}dk\right)^{1/\rho}, \quad \text{where } 0 < \alpha < \rho < 1.$$

Hence competiting inputs are more substitutable than differentiated products. Each input k is produced by a distinct supplier who faces a downward-sloping demand $x\left(k\right) = Bp\left(k\right)^{-1/(1-\rho)}$, where $B = Ap^{\frac{\rho-\alpha}{(1-\rho)(1-\alpha)}}$. The supplier can transform raw input l into input x with productivity φ, $x\left(k\right) = \varphi\left(k\right)l\left(k\right)$. Assume that suppliers draw productivity from a common distribution $G_x\left(\varphi\right)$. As all suppliers with the same producitivity behave symmetrically, I will index suppliers by productivity φ from now on.

The firm can integrate the supplier φ, solving the problem:

$$\pi_{xV}\left(\varphi\right) = \max_{l} B^{1-\rho}\varphi^{\rho}l^{\rho} - ql - f_{xV}.$$

Here, f_{xV} is the fixed overhead cost for integrating supplier φ. This program delievers a profit function $\pi_{xV}\left(\varphi\right) = \left(1-\rho\right)B\left(q/\rho\right)^{-\frac{\rho}{1-\rho}}\varphi^{\frac{\rho}{1-\rho}} - f_{xV} = \psi_{xV}\varphi^{\frac{\rho}{1-\rho}} - f_{xV}$. The firm can also outsource the production of input to the supplier, in which case the supplier solves the problem:

$$\begin{aligned}
\pi_S\left(\varphi\right) = \max_{l,F(\varphi)} & \left(1-\beta\right)B^{1-\rho}\varphi^{\rho}l^{\rho} - l - \left(1-d\right)T_x\left(\varphi\right) - \lambda F\left(\varphi\right) - \left(1-\lambda\right)\mu f^{e}. \\
s.t. \quad & \left(1-\beta\right)B^{1-\rho}\varphi^{\rho}l^{\rho} - l - \left(1-d\right)T_x\left(\varphi\right) \geq F\left(\varphi\right), \\
& \lambda F\left(\varphi\right) + \left(1-\lambda\right)\mu f^{e} \geq dT\left(\varphi\right).
\end{aligned}$$

Solving this program yields the profit function of the firm from input φ, $\pi_{xO}\left(\varphi\right) = \psi_{xO}\varphi^{\frac{\rho}{1-\rho}} - f_{xO}$, where $\varphi_{xO} = \phi_{xO}B\left(\frac{1}{\rho(1-\beta)}\right)^{-\frac{\rho}{1-\rho}}$, $\phi_{xO} = \beta + \frac{(1-\rho)(1-\beta)}{1+d\left(\frac{1-\lambda}{\lambda}\right)}$, and $f_{xO} = f_x - \frac{\left(\frac{1-\lambda}{\lambda}\right)\mu f^{e}}{1+d\left(\frac{1-\lambda}{\lambda}\right)}$.

Assume the distribution of input productivity is $G_x\left(\varphi\right) = 1 - \left(b_x/\varphi\right)^{\theta_x}$, then the measure of integrating inputs is:

$$\sigma_x = \left[\frac{\psi_{xV}-\psi_{xO}}{\psi_{xO}}\frac{f_{xO}}{f_{xV}-f_{xO}}\right]^{\theta_x(1-\rho)/\rho}.$$

It is straightforward to verify that Proposition 1.1 continue to hold at the input level. In particular, firms are more likely to outsource inputs when financial development is high, when the supplier less rely on external finance and has more collateralizable assets. Furthermore, the first order effects of external financial dependence and asset collateralizability fall with financial development. These predictions can be tested with firm-level data.

B

APPENDIX TO CHAPTER 3

B.1 PROOFS

B.1.1 *Proof of Proposition 3.1*

Define the following correspondence

$$T \;:\; \left[\underline{\theta}, \overline{\theta}\right] \rightarrow \left[\underline{z}, \overline{z}\right], \quad T\left(\theta\right) = \left\{z \;:\; k\left(\theta, z\right) > 0\right\}$$

Goods market clearing implies that final good producer has non-zero demand for all stage goods. Capital market clearing implies that all stage goods have non-zero demand for capital. It follows that T is surjective.

The next step is to show that T is non-decreasing. I proceed by contradiction. Start with producer z's problem:

$$\max_{\{k(\theta,z)\},l(z)} p\left(z\right) \left[\int_{\underline{\theta}}^{\overline{\theta}} A\left(\theta, z\right) k\left(\theta, z\right) d\theta\right]^{\alpha} l\left(z\right)^{1-\alpha} - \int_{\underline{\theta}}^{\overline{\theta}} r\left(\theta\right) k\left(\theta, z\right) d\theta - wl\left(z\right)$$

The first order condition for $l\left(z\right)$ is

$$\left(1-\alpha\right) p\left(z\right) y\left(z\right) = wl\left(z\right)$$

And for $k\left(\theta, z\right)$ is

$$\alpha p\left(z\right) \left[\int_{\underline{\theta}}^{\overline{\theta}} A\left(\theta, z\right) k\left(\theta, z\right) d\theta\right]^{\alpha-1} l\left(z\right)^{1-\alpha} A\left(\theta, z\right) \leq r\left(\theta\right), \quad \text{with equality iff } k\left(\theta, z\right) > 0$$

Combining the two conditions implies

$$\alpha \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} p\left(z\right)^{\frac{1}{\alpha}} A\left(\theta, z\right) \leq r\left(\theta\right), \quad \text{with equality iff } k\left(\theta, z\right) > 0$$

Now suppose T is not non-decreasing, implying that there exists $\theta_1 < \theta_2$ and $z_1 > z_2$ such that $z_1 = T\left(\theta_1\right)$ and $z_2 = T\left(\theta_2\right)$. The condition above implies

$$\alpha \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} p\left(z_1\right)^{\frac{1}{\alpha}} A\left(\theta_1, z_1\right) = r\left(\theta_1\right)$$

$$\alpha \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} p\left(z_1\right)^{\frac{1}{\alpha}} A\left(\theta_2, z_1\right) \leq r\left(\theta_2\right)$$

$$\alpha \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} p\left(z_2\right)^{\frac{1}{\alpha}} A\left(\theta_2, z_2\right) = r\left(\theta_2\right)$$

$$\alpha \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} p\left(z_2\right)^{\frac{1}{\alpha}} A\left(\theta_1, z_2\right) \leq r\left(\theta_1\right)$$

It follows that

$$\frac{A(\theta_2, z_1)}{A(\theta_1, z_1)} \leq \frac{r(\theta_2)}{r(\theta_1)} \leq \frac{A(\theta_2, z_2)}{A(\theta_1, z_2)}$$

But this contradicts A is strict log-supermodular in θ and z. Hence, T is non-decreasing.

The final step is to show that T is a single-valued, strictly increasing and continuous function. Suppose T is multi-valued, implying that for some θ, there exist z_1, z_2 in $T(\theta)$ with $z_1 < z_2$. Since T is non-decreasing, for any z in $[z_1, z_2]$, z must be in $T(\theta)$. But this contradicts there is no mass point in the quality distribution. It follows that T is single-valued, well-defined function. Similar reasoning can be combined with the assumption that there is no mass point in the know-how distribution to show that T is strictly increasing. Combining this result with the fact that T is surjective implies that T is continuous.

B.1.2 *Proof of Proposition 3.2*

The first step is to derive the differential equation for matching function. Capital market clearing implies

$$K(\theta) = \int_{\underline{z}}^{T(\theta)} k\left(T^{-1}(z), z\right) \mu(z)\, dz$$

Differentiating this equation with respect to theta implies

$$K'(\theta) = k(\theta, z)\, \mu(z)\, T'(\theta)$$

Cobb-Douglas production of stage good implies capital demand

$$k(\theta, z) = \frac{\alpha p(z)\, y(z)}{r(\theta)}$$

CES production of final good implies demand for stage good z

$$y(z) = p(z)^{\frac{1}{\rho-1}} Y$$

Combining the four equations above implies

$$T'(\theta) = \frac{r(\theta)\, K'(\theta)}{\alpha \mu(z)\, p(z)^{\frac{\rho}{\rho-1}} Y}$$

The next step is to show the differential equation for rent schedule. Reproduce the combined first order condition for producer z's problem below:

$$\alpha \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} p(z)^{\frac{1}{\alpha}} A(\theta, z) \leq r(\theta), \quad \text{with equality iff } k(\theta, z) > 0$$

It follows that

$$\alpha\left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} p\left(T\left(\theta\right)\right)^{\frac{1}{\alpha}} A\left(\theta, T\left(\theta\right)\right) = r\left(\theta\right)$$

$$\alpha\left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} p\left(T\left(\theta\right)\right)^{\frac{1}{\alpha}} A\left(\theta+d\theta, T\left(\theta\right)\right) \leq r\left(\theta+d\theta\right)$$

$$\alpha\left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} p\left(T\left(\theta+d\theta\right)\right)^{\frac{1}{\alpha}} A\left(\theta+d\theta, T\left(\theta+d\theta\right)\right) = r\left(\theta+d\theta\right)$$

$$\alpha\left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} p\left(T\left(\theta+d\theta\right)\right)^{\frac{1}{\alpha}} A\left(\theta, T\left(\theta+d\theta\right)\right) \leq r\left(\theta\right)$$

Hence,

$$\alpha\left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} p\left(T\left(\theta\right)\right)^{\frac{1}{\alpha}} \frac{A\left(\theta+d\theta, T\left(\theta\right)\right) - A\left(\theta+d\theta, T\left(\theta\right)\right)}{d\theta} \leq \frac{r\left(\theta+d\theta\right) - r\left(\theta\right)}{d\theta}$$
$$\leq \alpha\left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} p\left(T\left(\theta+d\theta\right)\right)^{\frac{1}{\alpha}} \frac{A\left(\theta+d\theta, T\left(\theta+d\theta\right)\right) - A\left(\theta+d\theta, T\left(\theta\right)\right)}{d\theta}$$

Now taking the limit as $d\theta \to 0$ implies that the rent schedule is differentiable and satisfies the differential equation (5).

B.1.3 *Proof of Proposition 3.3*

The first step is to show that an increase of quality abundance shifts the matching function downwards. That is,

$$\frac{\tilde{K}'\left(\theta'\right)}{\tilde{K}'\left(\theta\right)} \geq \frac{K'\left(\theta'\right)}{K'\left(\theta\right)}, \quad \forall \theta' > \theta$$

implies $\tilde{T}\left(\theta\right) \leq T\left(\theta\right)$ for all θ in $\left[\underline{\theta}, \overline{\theta}\right] \cap \left[\underline{\tilde{\theta}}, \overline{\tilde{\theta}}\right]$.

Suppose for some θ in $\left[\underline{\theta}, \overline{\theta}\right] \cap \left[\underline{\tilde{\theta}}, \overline{\tilde{\theta}}\right]$, $\tilde{T}\left(\theta\right) > T\left(\theta\right)$. Obviously, $\underline{z} = \tilde{T}\left(\underline{\tilde{\theta}}\right) \leq T\left(\underline{\tilde{\theta}}\right)$ and $\tilde{T}\left(\overline{\theta}\right) \leq T\left(\overline{\theta}\right) = \overline{z}$. Since both T and $\tilde{T}$ are continuous, $\tilde{T}$ must move above T and then move below T at least once. Hence, there exists $\underline{\tilde{\theta}} \leq \theta_1 \leq \theta_2 \leq \overline{\theta}$ and $\underline{z} \leq z_1 \leq z_2 \leq \overline{z}$ such that (i) $\tilde{T}\left(\theta_1\right) = T\left(\theta_1\right) = z_1$, $\tilde{T}\left(\theta_2\right) = T\left(\theta_2\right) = z_2$, (ii) $\tilde{T}'\left(\theta_1\right) \geq T'\left(\theta_1\right)$, $\tilde{T}'\left(\theta_2\right) \leq T'\left(\theta_2\right)$, (iii) $\tilde{T}\left(\theta\right) > T\left(\theta\right)$, for all θ in $\left(\theta_1, \theta_2\right)$.

Condition (ii) implies

$$\frac{\tilde{T}'\left(\theta_2\right)}{\tilde{T}'\left(\theta_1\right)} \leq \frac{T'\left(\theta_2\right)}{T'\left(\theta_1\right)}$$

Condition (i), the differential equation for matching function and the unit production cost together imply

$$\left(\frac{\tilde{r}\left(\theta_2\right)}{\tilde{r}\left(\theta_1\right)}\right)^{1+\alpha\frac{\rho}{\rho-1}} \frac{\tilde{K}'\left(\theta_2\right)}{\tilde{K}'\left(\theta_1\right)} \leq \left(\frac{r\left(\theta_2\right)}{r\left(\theta_1\right)}\right)^{1+\alpha\frac{\rho}{\rho-1}} \frac{K'\left(\theta_2\right)}{K'\left(\theta_1\right)}$$

By assumption,

$$\frac{\tilde{K}'\left(\theta_2\right)}{\tilde{K}'\left(\theta_1\right)} \geq \frac{K'\left(\theta_2\right)}{K'\left(\theta_1\right)}$$

It follows that

$$\frac{\tilde{r}(\theta_2)}{\tilde{r}(\theta_1)} \leq \frac{r(\theta_2)}{r(\theta_1)}$$

This equation cannot hold because of condition (iii), the strict log-supermodularity of A, and the differential equation for rent schedule.

The next step is to show that downward shift of the matching function implies less variation in capital intensity. Combining condition $\tilde{T}(\theta) < T(\theta)$ for all θ, the strict log-supermodularity of A, and the differential equation for rent schedule, I obtain

$$\frac{d \log \tilde{r}(\theta)}{d\theta} = \frac{\partial \log A\left(\theta, \tilde{T}(\theta)\right)}{\partial \theta} \leq \frac{\partial \log A(\theta, T(\theta))}{\partial \theta} = \frac{d \log r(\theta)}{d\theta}$$

Integrating the above inequality implies

$$\frac{\tilde{r}(\theta')}{\tilde{r}(\theta)} \leq \frac{r(\theta')}{r(\theta)}, \quad \forall \theta' > \theta$$

Let R and $\tilde{R}$ be the distributions induced by $r(\theta)$ and $\tilde{r}(\theta)$ respectively. That is,

$$R(x) = \Pr\{r : r \leq x\} = \Pr\left\{\theta : \theta \leq r^{-1}(x)\right\} = K\left(r^{-1}(x)\right)$$

$$R'(x)\, dx = K'(\theta) \frac{1}{r'(\theta)} d\theta$$

Similarly,

$$\tilde{R}(x) = \Pr\{\tilde{r} : \tilde{r} \leq x\} = \Pr\left\{\theta : \theta \leq \tilde{r}^{-1}(x)\right\} = K\left(\tilde{r}^{-1}(x)\right)$$

$$\tilde{R}'(x)\, dx = \tilde{K}'(\theta) \frac{1}{\tilde{r}'(\theta)} d\theta$$

To show there is less variation in capital intensity, it suffices to show that $\tilde{R}$ second-order stochastically dominates R. Combining the above three equations implies

$$\tilde{R}'(x')\, R'(x) \geq \tilde{R}'(x)\, R'(x')$$

Integrating both sides over x from $\underline{r}$ to x' implies

$$R(x')\, \tilde{R}'(x') \geq \tilde{R}(x')\, R'(x')$$

Integrating both sides over x' from x to $\bar{r}$ implies

$$\left[1 - \tilde{R}(x)\right] R'(x) \geq [1 - R(x)]\, \tilde{R}'(x)$$

The x and x' in the above two equations are arbitrary. Combine these equations by letting $x = x' = r$, I obtain

$$\frac{1 - \tilde{R}(r)}{1 - R(r)} \geq \frac{\tilde{R}'(r)}{R'(r)} \geq \frac{\tilde{R}(r)}{R(r)}$$

This implies

$$\tilde{R}(r) \leq R(r), \quad \forall r$$

Hence, $\tilde{R}$ first-order stochastically dominates R. The conclusion that there is less variation in capital intensity follows from the fact that first order stochastic dominance implies second order stochastic dominance.

B.1.4 *Proof of Proposition 3.4*

The first step is to show that an increase in the relative demand of downstream goods shifts the matching function upwards. That is,

$$\frac{\tilde{\mu}(z')}{\tilde{\mu}(z)} \geq \frac{\mu(z')}{\mu(z)}, \quad \forall z' > z$$

implies $\tilde{T}(\theta) \geq T(\theta)$ for all θ in $[\underline{\theta}, \overline{\theta}]$.

Suppose for some θ in $[\underline{\theta}, \overline{\theta}]$, $\tilde{T}(\theta) < T(\theta)$. Obviously, $\underline{z} = \tilde{T}(\underline{\theta}) = T(\underline{\theta})$ and $\tilde{T}(\overline{\theta}) = T(\overline{\theta}) = \overline{z}$. Since both T and $\tilde{T}$ are continuous, $\tilde{T}$ must move below T and then move above T at least once. Hence, there exists $\underline{\theta} \leq \theta_1 \leq \theta_2 \leq \overline{\theta}$ and $\underline{z} \leq z_1 \leq z_2 \leq \overline{z}$ such that (i) $\tilde{T}(\theta_1) = T(\theta_1) = z_1$, $\tilde{T}(\theta_2) = T(\theta_2) = z_2$, (ii) $\tilde{T}'(\theta_1) \leq T'(\theta_1)$, $\tilde{T}'(\theta_2) \geq T'(\theta_2)$, (iii) $\tilde{T}(\theta) < T(\theta)$, for all θ in (θ_1, θ_2).

Condition (ii) implies

$$\frac{\tilde{T}'(\theta_2)}{\tilde{T}'(\theta_1)} \geq \frac{T'(\theta_2)}{T'(\theta_1)}$$

Condition (i), the differential equation for matching function and the unit production cost together imply

$$\left(\frac{\tilde{r}(\theta_2)}{\tilde{r}(\theta_1)}\right)^{1+\alpha\frac{\rho}{\rho-1}} \frac{\tilde{\mu}'(z_1)}{\tilde{\mu}'(z_2)} \geq \left(\frac{r(\theta_2)}{r(\theta_1)}\right)^{1+\alpha\frac{\rho}{\rho-1}} \frac{\mu'(z_1)}{\mu'(z_2)}$$

By assumption,

$$\frac{\tilde{\mu}(z_1)}{\tilde{\mu}(z_2)} \leq \frac{\mu(z_1)}{\mu(z_2)}$$

It follows that

$$\frac{\tilde{r}(\theta_2)}{\tilde{r}(\theta_1)} \geq \frac{r(\theta_2)}{r(\theta_1)}$$

This equation cannot hold because of condition (iii), the strict log-supermodularity of A, and the differential equation for rent schedule.

The next step is to show that upward shift of the matching function implies more variation in capital intensity. The result follows logically from the proof of the second part of Proposition 3.3.

BIBLIOGRAPHY

Acemoglu, Daron, Philippe Aghion, Rachel Griffith, and Fabrizio Zilibotti (2010). "Vertical integration and technology: theory and evidence." In: *Journal of the European Economic Association* 8.5, pp. 989–1033 (cit. on pp. 7, 46).

Acemoglu, Daron, Pol Antràs, and Elhanan Helpman (2007). "Contracts and Technology Adoption." In: *American Economic Review* 97.3, pp. 916–943 (cit. on p. 91).

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi (2012). "The Network Origins of Aggregate Fluctuations." In: *Econometrica* 80.5, pp. 1977–2016 (cit. on pp. 47, 78).

Acemoglu, Daron, Simon Johnson, and Todd Mitton (2009). "Determinants of Vertical Integration: Financial Development and Contracting Costs." In: *Journal of Finance* 64.3, pp. 1251–1290 (cit. on pp. 7, 35, 46).

Aghion, Philippe and Patrick Bolton (1992). "An Incomplete Contracts Approach to Financial Contracting." In: *Review of Economic Studies* 59.3, pp. 473–494.

Aghion, Philippe, Rachel Griffith, and Peter Howitt (2006). "Vertical Integration and Competition." In: *American Economic Review: Papers and Proceedings* 96.2, pp. 97–102 (cit. on pp. 7, 46).

Aghion, Philippe and Jean Tirole (1994). "The Management of Innovation." In: *Quarterly Journal of Economics* 109.4, pp. 1185–209.

Alfaro, Laura, Pol Antràs, Davin Chor, and Paola Conconi (2017). "Internalizing Global Value Chains: A Firm-Level Analysis." In: *Journal of Political Economy* (cit. on p. 77).

Alfaro, Laura, Paola Conconi, Harald Fadinger, and Andrew F. Newman (2016). "Do Prices Determine Vertical Integration?" In: *Review of Economic Studies* 0.0, pp. 1–34 (cit. on pp. 7, 46, 48).

Altinoglu, Levent (2017). "The Origins of Aggregate Fluctuations in a Credit Network Economy." Working paper (cit. on p. 78).

Amiti, Mary and David E. Weinstein (2011). "Exports and Financial Shocks." In: *Quarterly Journal of Economics* 4.1, pp. 1841–1877 (cit. on p. 7).

Antràs, Pol (2003). "Firms, Contracts, and Trade Structure." In: *Quarterly Journal of Economics* 118.4, pp. 1375–1418 (cit. on pp. 6, 7, 46).

– (2013). "Grossman–Hart (1986) Goes Global: Incomplete Contracts, Property Rights, and the International Organization of Production." In: *The Journal of Law, Economics, and Organization* 0.0, pp. 1–58.

Antràs, Pol and Ricardo J. Caballero (2009). "Trade and Capital Flows: A Financial Frictions Perspective." In: *Journal of Political Economy* 117.4, pp. 701–744.

Antràs, Pol and Davin Chor (2013). "Organizing the Global Value Chain." In: *Econometrica* 81.6, pp. 2127–2204 (cit. on pp. 48, 77).

Antràs, Pol, Davin Chor, Thibault Fally, and Russell Hillberry (2012). "Measuring the Upstreamness of Production and Trade Flows." In: *American Economic Review* 102.3, pp. 412–416 (cit. on pp. 75, 77–79).

Antràs, Pol, Mihir A. Desai, and C. Fritz Foley (2009). "Multinational Firms, FDI Flows, and Imperfect Capital Markets." In: *Quarterly Journal of Economics* 124.3.

Antràs, Pol, Teresa C. Fort, and Felix Tintelnot (2017). "The Margins of Global Sourcing: Theory and Evidence from US Firms." In: *American Economic Review* 107.9, pp. 2514–64 (cit. on p. 77).

Antràs, Pol, Luis Garicano, and Esteban Rossi-Hansberg (2006). "Offshoring in a Knowledge Economy." In: *Quarterly Journal of Economics* 121.1, pp. 31–77 (cit. on p. 77).

Antràs, Pol and Alonso de Gortari (2017). "On the Geography of Global Value Chains." NBER Working Paper No. 23456 (cit. on p. 77).

Antràs, Pol and Elhanan Helpman (2004). "Global Sourcing." In: *Journal of Political Economy* 112.3, pp. 552–580.

– (2008). "Contractual Frictions and Global Sourcing." In: *The Organization of Firms in a Global Economy*. Ed. by Elhanan Helpman, Dalia Marin, and Thierry Verdier. 9-54. Harvard University Press.

Armstrong, Angus, E Philip Davis, Iana Liadze, and Cinzia Rienzo (2013). "Evaluating Changes in Bank Lending to UK SMEs over 2001-12 – Ongoing Tight Credit." In: *Department for Business, Innovation and Skills* (cit. on p. 4).

Asker, John, Allan Collard-Wexler, and Jan De Loecker (2014). "Dynamic Inputs and Resource (Mis)Allocation." In: *Journal of Political Economy* 112.5, pp. 1013–1063 (cit. on pp. 7, 52).

Atalay, Enghin, Ali Hortaçsu, and Chad Syverson (2014). "Vertical Integration and Input Flows." In: *American Economic Review* 104.4, pp. 1120–1148 (cit. on pp. 11, 24).

Banerjee, Abhijit and Esther Duflo (2005). "Growth Theory through the Lens of Development Economics." In: *Handbook of Economic Growth*. Ed. by Philippe Aghion and Steven N. Durlauf. Vol. 1, Part A. Elsevier. Chap. 7, pp. 473–552 (cit. on pp. 7, 52).

Baqaee, David Rezza (2017). "Cascading Failures in Production Networks." Working paper (cit. on pp. 47, 78).

Barro, Robert and Jong Wha Lee (2010). "A New Data Set of Educational Attainment in the World, 1950-2010." NBER Working Paper No. 15902.

Bartelme, Dominick and Yuriy Gorodnichenko (2015). "Linkages and Economic Development." Working paper.

Bartelsman, Eric J. and Wayne Gray (1996). "The NBER Manufacturing Productivity Database." NBER Technical Working Paper No. 205 (cit. on pp. 57, 75, 78, 79).

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta (2013). "Cross-Country Differences in Productivity: The Role of Allocation and Selection." In: *American Economic Review* 103.1, pp. 305–34.

Beck, Thorsten, Á-Kunt Asli Demirg, and Ross Levine (2010). "A New Database on Financial Development and Structure." World Bank dataset (cit. on pp. 2, 18).

Becker, Randy, Wayne Gray, and Jordan Marvakov (2016). “NBER-CES Manufacturing Industry Database: Technical Notes.” mimeo (cit. on p. 78).

Bernanke, Ben S., Mark Gertler, and Simon Gilchrist (1999). “The financial accelerator in a quantitative business cycle framework.” In: *Handbook of Macroeconomics*. Ed. by John B. Taylor and Michael Woodford. 1st. Vol. 1. C. North Holland. Chap. 21, pp. 1341–1393 (cit. on p. 2).

Bernanke, Ben and Mark Gertler (1989). “Agency Costs, Net Worth, and Business Fluctuations.” In: *American Economic Review* 79.1, pp. 14–31 (cit. on p. 2).

Bernard, Andrew B. and Swati Dhingra (2016). “Importers, Exporters and the Division of the Gains from Trade.” Working paper (cit. on p. 48).

Bigio, Saki and Jennifer La’O (2016). “Financial Frictions in Production Networks.” Working paper (cit. on p. 78).

Bils, Mark and Peter J. Klenow (2001). “The Acceleration in Variety Growth.” In: *American Economic Review: Papers and Proceedings* 91.2, pp. 274–280 (cit. on p. 2).

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen (2010). “Does Product Market Competition Lead Firms to Decentralize?” In: *American Economic Review: Papers and Proceedings* 100.2, pp. 434–438 (cit. on pp. 7, 46).

Boehm, Johannes (2015). “The Impact of Contract Enforcement Costs on Outsourcing and Aggregate Productivity.” Working paper (cit. on pp. 7, 8, 18, 21, 46, 47).

Braun, Matias (2003). “Financial Contractibility and Asset Hardness.” Working paper (cit. on pp. 7, 17).

Breinlich, Holger and Chiara Criscuolo (2011). “International trade in services: A portrait of importers and exporters.” In: *Journal of International Economics* 84.2, pp. 188–206 (cit. on p. 4).

Bricongne, Jean-Charles, Lionel Fontagné, Guillaume Gaulier, Daria Taglioni, and Vincent Vicard (2012). “Firms and the global crisis: French exports in the turmoil.” In: *Journal of International Economics* 87.1, pp. 134–146 (cit. on p. 7).

Broda, Christian and David E. Weinstein (2006). “Globalization and the Gains From Variety.” In: *Quarterly Journal of Economics* 121.2, pp. 541–585 (cit. on p. 2).

Carluccio, Juan and Thibault Fally (2012). “Global Sourcing under Imperfect Capital Markets.” In: *Review of Economics and Statistics* 94.3, pp. 740–763 (cit. on p. 7).

Caselli, Francesco (2005). “Accounting for Cross-Country Income Differences.” In: *Handbook of Economic Growth*. Ed. by Philippe Aghion and Steven N. Durlauf. Vol. 1, Part A.A. Elsevier. Chap. 9, pp. 679–741.

Caselli, Francesco and James Feyrer (2007). “The Marginal Product of Capital.” In: *Quarterly Journal of Economics* 122.2, pp. 535–568.

Caselli, Francesco and Nicola Gennaioli (2013). “Dynastic Management.” In: *Economic Inquiry* 51.1, pp. 971–996 (cit. on pp. 7, 52).

Caselli, Francesco and Wilbur John Coleman II (2006). “The World Technology Frontier.” In: *American Economic Review* 96.3, pp. 499–522.

Caselli, Francesco and Daniel Wilson (2004). “Importing technology.” In: *Journal of Monetary Economics* 51.1, pp. 1–32 (cit. on pp. 75–77).

Chan, Jackie M.L. and Kalina Manova (2015). “Financial development and the choice of trade partners.” In: *Journal of Development Economics* 116, pp. 122–145 (cit. on p. 7).

Chor, Davin and Kalina Manova (2012). “Off the cliff and back? Credit conditions and international trade during the global financial crisis.” In: *Journal of International Economics* 87.1, pp. 117–13 (cit. on p. 7).

Coase, Ronald Harry (1937). “The Nature of the Firm.” In: *Economica* 4.16, pp. 386–405 (cit. on pp. 2, 6).

Costinot, Arnaud (2009). “An Elementary Theory of Comparative Advantage.” In: *Econometrica* 77.4, pp. 1165–1192 (cit. on p. 77).

Costinot, Arnaud and Jonathan Vogel (2010). “Matching and Inequality in the World Economy.” In: *Journal of Political Economy* 118.4, pp. 747–786 (cit. on pp. 76, 77).

Costinot, Arnaud, Jonathan Vogel, and Su Wang (2013). “An Elementary Theory of Global Supply Chains.” In: *Review of Economic Studies* 80.1, pp. 109–144 (cit. on p. 77).

David, Joel M., Hugo A. Hopenhayn, and Venky Venkateswaran (2016). “Information, Misallocation and Aggregate Productivity.” In: *Quarterly Journal of Economics* 131.2, pp. 943–1005 (cit. on pp. 7, 52).

David, Joel M. and Venky Venkateswaran (2017). “The Sources of Capital Misallocation.” NBER Working Paper No. 23129 (cit. on p. 72).

De Loecker, Jan and Jan Eeckhout (2017). “The Rise of Market Power and the Macroeconomic Implications.” NBER Working Paper No. 23687.

De Loecker, Jan and Frederic Warzynski (2012). “Markups and Firm-Level Export Status.” In: *American Economic Review* 102.6, pp. 2437–71 (cit. on pp. 50, 52, 58).

Demirgüç-Kunt, Asli and Vojislav Maksimovic (1998). “Law, Finance, and Firm Growth.” In: *Journal of Finance* 53.6, pp. 2107–2137 (cit. on p. 7).

Dhingra, Swati and John Morrow (2016). “Monopolistic Competition and Optimum Product Diversity Under Firm Heterogeneity.” In: *Journal of Political Economy* (cit. on pp. 7, 52).

Dixit, Avinash K. and Joseph E. Stiglitz (1977). “Monopolistic Competition and Optimum Product Diversity.” In: *American Economic Review* 67.3, pp. 297–308.

Eaton, Jonathan and Samuel Kortum (2001). “Trade in capital goods.” In: *European Economic Review* 45.7, pp. 1195–1235 (cit. on pp. 75–77).

– (2002). “Technology, Geography, and Trade.” In: *Econometrica* 70.5, pp. 1741–1779 (cit. on p. 48).

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu (2015). “Competition, Markups, and the Gains from International Trade.” In: *American Economic Review* 105.10, pp. 3183–3221 (cit. on pp. 50, 52, 72).

Fally, Thibault (2012). "Production staging: measurement and facts." mimeo (cit. on p. 77).

Fally, Thibault and Russell Hillberry (2015). "A Coasian Model of International Production Chains." NBER Working Paper No. 21520 (cit. on p. 77).

Feenstra, Robert C., Zhiyuan Li, and Miaojie Yu (2014). "Exports and Credit Constraints under Incomplete Information: Theory and Evidence from China." In: *Review of Economics and Statistics* 96.4, pp. 729–744 (cit. on p. 7).

Ford, George S. and John D. Jackson (1997). "Horizontal Concentration and Vertical Integration in the Cable Television Industry." In: *Review of Industrial Organization* 12.4, pp. 501–518 (cit. on p. 48).

Foster, Lucia, John Haltiwanger, and Chad Syverson (2008). "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" In: *American Economic Review* 98.1, pp. 394–425 (cit. on pp. 7, 58).

Franklin, Jeremy, May Rostom, and Gregory Thwaites (2015). "The banks that said no: banking relationships, credit supply and productivity in the United Kingdom." In: *Bank of England Working Paper No. 557*.

Gabaix, Xavier (2011). "The Granular Origins of Aggregate Fluctuations." In: *Econometrica* 79.3, pp. 733–772 (cit. on p. 78).

Garicano, Luis (2000). "Hierarchies and the Organization of Knowledge in Production." In: *Journal of Political Economy* 108.5, pp. 874–904 (cit. on pp. 76, 77).

Garicano, Luis and Esteban Rossi-Hansberg (2006). "Organization and Inequality in a Knowledge Economy." In: *Quarterly Journal of Economics* 121.4, pp. 1383–1435 (cit. on p. 77).

Gibbons, Robert (2005). "Four formal(izable) theories of the firm?" In: *Journal of Economic Behavior and Organization* 58.2, pp. 200–245 (cit. on pp. 6, 46).

Gopinath, Gita, Şebnem Kalemli-Özcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez (2017). "Capital Allocation and Productivity in South Europe." In: *Quarterly Journal of Economics* 132.4, pp. 1915–1967 (cit. on p. 52).

Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell (1997). "Long-Run Implications of Investment-Specific Technological Change." In: *American Economic Review* 87.3, pp. 342–362 (cit. on p. 77).

Grossman, Gene M. and Elhanan Helpman (1991). *Innovation and growth in the global economy*. Cambridge, MA: The MIT Press (cit. on p. 2).

– (2002). "Integration versus Outsourcing in Industry Equilibriumersus Outsourcing in Industry Equilibrium." In: *Quarterly Journal of Economics* 117.1, pp. 85–120 (cit. on p. 4).

Grossman, Gene M. and Esteban Rossi-Hansberg (2008). "Trading Tasks: A Simple Theory of Offshoring." In: *American Economic Review* 98.5, pp. 1978–1997 (cit. on p. 77).

Grossman, Sanford J. and Oliver D. Hart (1986). "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." In: *Journal of Political Economy* 94.4, pp. 691–719 (cit. on pp. 6, 46).

Hart, Oliver D. (1995). *Firms, Contracts, and Financial Structure*. New York: Oxford University Press (cit. on pp. 6, 46).

Hart, Oliver and John Moore (1990). "Property Rights and the Nature of the Firm." In: *Journal of Political Economy* 98.6, pp. 1119–1158 (cit. on pp. 6, 46).

Head, Keith and Thierry Mayer (2014). "Gravity Equations: Workhorse, Toolkit, and Cookbook." In: *Handbook of International Economics*. Ed. by Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff. Vol. 4. Elsevier. Chap. 3, pp. 131–195.

Hirschman, Albert O (1958). *The Strategy of Economic development*. New Haven, CT: Yale University Press.

Holmström, Bengt (1999). "The Firm as a Subeconomy." In: *Journal of Law, Economics, and Organization* 15.1, pp. 74–102 (cit. on p. 46).

Holmström, Bengt and Paul Milgrom (1991). "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." In: *Journal of Law, Economics, and Organization* 7.1, pp. 24–52 (cit. on p. 46).

– (1994). "The Firm as an Incentive System." In: *American Economic Review* 84.4, pp. 972–991 (cit. on p. 46).

Holmström, Bengt and Jean Tirole (1991). "Transfer Pricing and Organizational Form." In: *Journal of Law, Economics, and Organization* 7.2, pp. 201–228 (cit. on p. 46).

Hopenhayn, Hugo A. (2014). "On the Measure of Distortions." NBER Working Paper No. 20404.

Hopenhayn, Hugo and Richard Rogerson (1993). "Job Turnover and Policy Evaluation: A General Equilibrium Analysis." In: *Journal of Political Economy* 101.5, pp. 915–938.

Hortaçsu, Ali and Chad Syverson (2007). "Cementing Relationships: Vertical Integration, Foreclosure, Productivity, and Prices." In: *Journal of Political Economy* 115.2, pp. 250–301 (cit. on p. 48).

Hsieh, Chang-Tai and Peter J. Klenow (2009). "Misallocation and Manufacturing TFP in China and India." In: *Quarterly Journal of Economics* 124.4, pp. 1403–1148 (cit. on pp. 7, 50–52, 57, 58, 62, 70).

Hulten, Charles R. (1978). "Growth Accounting with Intermediate Inputs." In: *Review of Economic Studies* 45.3, pp. 511–518.

Jayaratne, Jith and Philip E. Strahan (1996). "The Finance-Growth Nexus: Evidence from Bank Branch Deregulation." In: *Quarterly Journal of Economics* 111.3, pp. 639–670 (cit. on p. 7).

Johnson, Robert C. and Guillermo Noguera (2012). "Accounting for intermediates: Production sharing and trade in value added." In: *Journal of International Economics* 86.2, pp. 224–236 (cit. on p. 77).

Jones, Charles I. (2011). "Intermediate Goods and Weak Links in the Theory of Economic Development." In: *American Economic Journal: Macroeconomics* 3.2, pp. 1–28 (cit. on pp. 7, 78).

– (2013). "Misallocation, Economic Growth, and Input-Output Economics." In: *Advances in Economics and Econometrics: Tenth World Congress*. Ed. by Daron Acemoglu, Manuel Arellano, and Eddie Dekel. Vol. II. Cambridge University Press, pp. 419–456 (cit. on pp. 7, 48, 78).

Jovanovic, Boyan and Rafael Rob (1997). "Solow vs. Solow: Machine Prices and Development." NBER Working Paper No. 5871 (cit. on p. 77).

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi (2011). "World Governance Indicators, 1996-2010." World Bank dataset.

King, Robert G. and Ross Levine (1993). "Finance and Growth: Schumpeter Might Be Right." In: *Quarterly Journal of Economics* 108.3, pp. 717–737 (cit. on p. 7).

Kiyotaki, Nobuhiro and John Moore (1997). "Credit Cycles." In: *Journal of Political Economy* 105.2, pp. 211–248 (cit. on p. 2).

Klein, Benjamin (1996). "Why hold-ups occur: The self-enforcing range of contractual relationships." In: *Economic inquiry* 34.3, pp. 444–463 (cit. on p. 46).

– (2000). "The Role of Incomplete Contracts in Self-Enforcing Relationships." In: *Revue d'Economie Industrielle* 92.1, pp. 67–80 (cit. on p. 46).

Klein, Benjamin, Robert G. Crawford, and Armen A. Alchian (1978). "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process." In: *Journal of Law, Economics, and Organization* 21.2, pp. 297–326 (cit. on p. 46).

Klein, Benjamin and Kevin M. Murphy (1988). "Vertical Restraints as Contract Enforcement Mechanisms." In: *The Journal of Law and Economics* 31.2, pp. 265–297 (cit. on p. 46).

– (1997). "Vertical Integration as a Self-Enforcing Contractual Arrangement." In: *American Economic Review: Papers and Proceedings* 87.2, pp. 415–420 (cit. on p. 46).

Koopman, Robert, Zhi Wang, and Shang-Jin Wei (2014). "Tracing Value-Added and Double Counting in Gross Exports." In: *American Economic Review* 104.2, pp. 459–494 (cit. on p. 77).

Krugman, Paul and Anthony J. Venables (1995). "Globalization and the Inequality of Nations." In: *Quarterly Journal of Economics* 110.4, pp. 857–880 (cit. on p. 48).

Lafontaine, Francine and Margaret Slade (2007). "Vertical Integration and Firm Boundaries: The Evidence." In: *Journal of Economic Literature* 45.3, pp. 629–685 (cit. on pp. 7, 46, 48).

Legros, Patrick and Andrew F. Newman (2008). "Competing for Ownership." In: *Journal of the European Economic Association* 6.6, pp. 1279–1308.

Leontief, Wassily W. (1936). "Quantitative Input and Output Relations in the Economic Systems of the United States." In: *Review of Economics and Statistics* 18.3, pp. 105–125.

Levine, Ross (2005). "Finance and Growth: Theory and Evidence." In: *Handbook of Economic Growth*. Ed. by Philippe Aghion and Steven N. Durlauf. Vol. 1A. Chap. 12, pp. 865–934 (cit. on p. 7).

Levinsohn, James and Amil Petrin (2003). "Estimating Production Functions Using Inputs to Control for Unobservables." In: *Review of Economic Studies* 70.2, pp. 317–341 (cit. on pp. 52, 57, 71).

Long, John B. and Charles I. Plosser (1983). "Real Business Cycles." In: *Journal of Political Economy* 91.1, pp. 39–69.

Macchiavello, Rocco (2012). "Financial Development and Vertical Integration: Theory and Evidence." In: *Journal of the European Economic Association* 10.2, pp. 255–289 (cit. on pp. 7, 35, 46).

Manova, Kalina (2013). "Credit Constraints, Heterogeneous Firms, and International Trade." In: *Review of Economic Studies* 80, pp. 711–744 (cit. on pp. 4, 7, 10).

Manova, Kalina, Shang-Jin Wei, and Zhiwei Zhang (2015). "Firm Exports and Multinational Activity Under Credit Constraints." In: *Review of Economics and Statistics* 97.3, pp. 574–588 (cit. on p. 7).

Mayer, Thierry, Marc J. Melitz, and Gianmarco Ottaviano (2014). "Market Size, Competition, and the Product Mix of Exporters." In: *American Economic Review* 104.2, pp. 495–536.

McBride, Mark E. (1983). "Spatial Competition and Vertical Integration: Cement and Concrete Revisited." In: *American Economic Review* 73.5, pp. 1011–1022 (cit. on p. 48).

Melitz, Marc J. (2003). "The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity." In: *Econometrica* 71.6, pp. 1695–1725 (cit. on p. 2).

Melitz, Marc J. and Gianmarco I. P. Ottaviano (2008). "Market Size, Trade, and Productivity." In: *Review of Economic Studies* 75, pp. 295–316 (cit. on pp. 2, 48, 51, 52).

Melitz, Marc J. and Stephen J. Redding (2014). "Missing Gains from Trade?" In: *American Economic Review: Papers and Proceedings* 104.5, pp. 317–321 (cit. on p. 48).

Midrigan, Virgiliu and Daniel Yi Xu (2014). "Finance and Misallocation: Evidence from Plant-Level Data." In: *American Economic Review* 104.2, pp. 422–458 (cit. on pp. 7, 52).

Moll, Benjamin (2014). "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?" In: *American Economic Review* 104.10, pp. 3186–3221 (cit. on pp. 7, 52).

Muris, Timothy J., David T. Scheffman, and Pablo T. Spiller (1992). "Strategy and Transaction Costs: The Organization of Distribution in the Carbonated Soft Drink Industry." In: *Journal of Economics and Management Strategy* 1.1, pp. 83–128 (cit. on p. 48).

Nocco, Antonella, Gianmarco I. P. Ottaviano, and Matteo Salto (2014). "Monopolistic Competition and Optimum Product Selection." In: *American Economic Review* 104.5, pp. 304–309 (cit. on p. 52).

Nunn, Nathan (2007). "Relationship-specificity, incomplete contracts, and the pattern of trade." In: *Quarterly Journal of Economics* 122.2, pp. 569–600.

Nunn, Nathan and Daniel Trefler (2014). "Domestic Institutions as a Source of Comparative Advantage." In: *Handbook of International Economics*. Ed. by Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff. Vol. 4. Chap. 5, pp. 263–315.

Oberfield, Ezra (2013). "Productivity and Misallocation During a Crisis: Evidence from the Chilean Crisis of 1982." In: *Review of Economic Dynamics* 16.1, pp. 100–119.

– (2017). "A Theory of Input-Output Architecture." Working paper.

Olley, G. Steven and Ariel Pakes (1996). "The Dynamics of Productivity in the Telecommunications Equipment Industry." In: *Econometrica* 64.6, pp. 1263–1297 (cit. on pp. 52, 57, 71).

Paravisini, Daniel, Veronica Rappoport, Philipp Schnabl, and Daniel Wolfenzon (2015). "Dissecting the Effect of Credit Supply on Trade: Evidence from Matched Credit-Export Data." In: *Review of Economic Studies* 82.1, pp. 333–359 (cit. on p. 7).

Peters, Michael (2013). "Heterogeneous mark-Ups and endogenous misallocation." In: *Working paper* (cit. on pp. 7, 48).

– (2016). "Heterogeneous Mark-Ups, Growth and Endogenous Misallocation." Working Paper (cit. on pp. 50, 52, 72).

Rajan, Raghuram G. and Luigi Zingales (1998). "Financial Dependence and Growth." In: *American Economic Review* 88.3, pp. 559–586 (cit. on pp. 5, 7, 17).

Ramondoa, Natalia, Veronica Rappoport, and Kim J. Ruhl (2016). "Intrafirm trade and vertical fragmentation in U.S. multinational corporations." In: *Journal of International Economics* 98, pp. 51–59 (cit. on p. 24).

Restuccia, Diego and Richard Rogerson (2008). "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments." In: *Review of Economic Dynamics* 11.4, pp. 707–720 (cit. on pp. 7, 50, 52).

R.Krugman, Paul (1979). "Increasing returns, monopolistic competition, and international trade." In: *Journal of International Economics* 9.4, pp. 469–479 (cit. on p. 2).

Romer, Paul M. (1990). "Endogenous Technological Change." In: *Journal of Political Economy* 98.5, pp. 71–102 (cit. on p. 2).

Sampson, Thomas (2014). "Selection into Trade and Wage Inequality." In: *American Economic Journal: Microeconomics* 6.3, pp. 157–202 (cit. on p. 77).

– (2016). "Assignment reversals: Trade, skill allocation and wage inequality." In: *Journal of Economic Theory* 163, pp. 365–409 (cit. on p. 77).

Schans, Daniel van der (2012). "SME Access to External Finance." In: *BIS Economics Paper No. 16* Department for Business, Innovation and Skills (cit. on pp. 3, 32).

Shepard, Andrea (1993). "Contractual Form, Retail Price, and Asset Characteristics in Gasoline Retailing." In: *RAND Journal of Economics* 24.1, pp. 58–77 (cit. on p. 48).

Silva, Santos J.M.C. and Silvana Tenreyro (2006). "The Log of Gravity." In: *Review of Economics and Statistics* 88.4, pp. 641–658 (cit. on pp. 6, 38).

– (2011). "Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator." In: *Economics Letters* 112.2, pp. 220–222.

Solow, Robert M. (1960). "Investment and Technical Progress." In: *Mathematical Methods in the Social Science*. Ed. by Kenneth Arrow, Samuel Karlin, and Paul Suppes. Stanford University Press (cit. on p. 77).

Townsend, Robert M (1979). "Optimal contracts and competitive markets with costly state verification." In: *Journal of Economic Theory* 21.2, pp. 265–293 (cit. on p. 2).

Williamson, Oliver E. (1971). "The Vertical Integration of Production: Market Failure Considerations." In: *American Economic Review* 61.2, pp. 112–123 (cit. on pp. 2, 46).

– (1973). "Markets and Hierarchies: Some Elementary Considerations." In: *American Economic Review* 63.2, pp. 316–325 (cit. on p. 46).

– (1975). *Markets and Hierarchies: Analysis and Antitrust Implications*. New York, NY: Free Press (cit. on p. 46).

– (1979). "Transaction-Cost Economics: The Governance of Contractual Relations." In: *The Journal of Law and Economics* 22.2, pp. 233–261 (cit. on pp. 2, 46).

– (1985). *The Economic Institutions of Capitalism*. New York, NY: Free Press (cit. on pp. 2, 6, 46).

– (1991). "Comparative Economic Organization: The Analysis of Discrete Structural Alternatives." In: *Administrative Science Quarterly* 36.2, pp. 269–296 (cit. on p. 46).

Wilson, Daniel J. (2002). "Is Embodied Technology the Result of Upstream R&D? Industry-Level Evidence." In: *Review of Economic Dynamics* 5.2, pp. 285–317 (cit. on p. 77).

Yeaple, Stephen Ross (2005). "A simple model of firm heterogeneity, international trade, and wages." In: *Journal of International Economics* 65.1, pp. 1–20 (cit. on p. 77).

Yi, Kei-Mu (2003). "Can Vertical Specialization Explain the Growth of World Trade?" In: *Journal of Political Economy* 111.1, pp. 52–102 (cit. on p. 77).