Finance and Development: A Tale of Two Sectors

Francisco J. Buera* Joseph P. Kaboski† Yongseok Shin‡

October 25, 2009

Abstract

Income differences across countries primarily reflect differences in total factor productivity (TFP). More disaggregate data show that the TFP gap between rich and poor countries varies systematically across industrial sectors of the economy. For example, poor countries are particularly unproductive in manufacturing. We develop a quantitative framework to explain the relationship between aggregate/sector-level TFP and financial development across countries. Financial frictions distort the allocation of capital and talent across production units, adversely affecting measured productivity. In our model, sectors with larger scales of operation (e.g. manufacturing) have more financing needs, and are hence disproportionately vulnerable to financial frictions. Our quantitative analysis shows that financial frictions account for a large part of the observed cross-country differences in output per worker, aggregate TFP, sector-level relative productivity, and capital to output ratios.

*University of California at Los Angeles and NBER; fjbuera@econ.ucla.edu.
†Ohio State University and NBER; kaboski.1@osu.edu.
‡Washington University in St. Louis and Federal Reserve Bank of St. Louis; yshin@wustl.edu.

1 Introduction

Income per capita differences across countries are primarily accounted for by low total factor productivity (TFP) in poor countries (Klenow and Rodríguez-Clare, 1997; Hall and Jones, 1999). More disaggregate data show that the TFP gap between rich and poor countries varies systematically across industrial sectors of the economy. For instance, poor countries are particularly unproductive in producing manufactured goods, including investment goods (Hsieh and Klenow, 2007).[1] Here we attempt to explain these facts by building and using a quantitative model of financial markets and economic development.

There are stark differences in financial development across countries, and a large body of empirical and theoretical work has stressed the role of financial markets in economic development. Our model is intended to capture and quantify this role. We discover that, quantitatively, financial frictions explain a large part of the cross-country differences in aggregate TFP and sector-level relative productivity.

In our model, financial frictions distort the allocation of capital across heterogeneous establishments and also their entry/exit decisions, lowering aggregate and sector-level TFP as a consequence. Furthermore, the impacts of financial frictions are not equal across sectors. Consistent with the data, sectors that are particularly unproductive in poor countries (e.g. manufacturing) tend to have larger establishments on average than other sectors (e.g. services). Larger scales imply larger financing needs, and the manufacturing sector is disproportionately more vulnerable to financial frictions. Our model replicates the lower relative productivity of manufacturing to services in countries with underdeveloped financial markets.

We build an economy with two sectors (manufacturing and services) that differ in the per-period fixed costs of operating an establishment. This difference in fixed costs leads to a difference in the scale of establishments in the two sectors. In the model, individuals choose in each period whether to operate an establishment in either sector (entrepreneurship) or to supply labor for a wage. Individuals have different levels of entrepreneurial productivity and wealth, with the latter being endogenously determined by forward-looking saving decisions. The entrepreneurial productivity evolves stochastically, generating the need to reallocate capital and labor from previously-productive entrepreneurs to currently-productive ones. Financial frictions hinder this reallocation process.

We model financial frictions in the form of collateral constraints founded on imperfect enforceability of contracts. In an economy with perfect credit markets, occupational choices are based only on comparative advantage: high-productivity individuals become entrepreneurs,

[1] Jones (1994), Eaton and Kortum (2001), and Herrendorf and Valentinyi (2006) are recent contributions. Also, Balassa (1964) and Samuelson (1964) are the classic citations for tradables vs. non-tradables, which are closely related to the manufacturing vs. services distinction.

and the allocation of capital equalizes marginal products of capital across establishments. With financial frictions in the form of collateral constraints, individuals' occupational choices and entrepreneurs' production decisions are constrained by their available wealth (collateral), even though individuals do partly overcome the constraints over time with self-financing.

We use our model to quantitatively analyze the cross-country patterns in aggregate and sector-level TFP. We discipline our analysis by requiring that a benchmark model with well-functioning financial markets matches the US data on the establishment size distribution across and within sectors (e.g., average scales and thick right-tails of broadly-defined sectors), the dynamics of establishments, and income concentration. We then quantify the relationship between financial development (measured by the ratio of external financing to GDP) and economic development (output per worker, aggregate/sector-level TFP, and capital to output ratios), both in the cross-country data and in the model simulations. Finally, we use detailed data on the establishment size distribution in Mexico and the US to test additional implications of our model.

We find that financial frictions have sizable effects on output per worker, aggregate and sector-level TFP, and capital to output ratios.

The variation in financial development can explain a factor-of-two difference in output per worker across economies, or almost 80 per cent of the difference in output per worker between Mexico and the US. One thing to note is that the agricultural sector is not modeled or analyzed here. This factor-of-two difference goes a long way in explaining the factor-of-five difference in non-agricultural output per worker between the richest and the poorest fifth percentiles of countries. Consistent with the consensus view in the literature, most of the output per worker differences in our model are accounted for by the low TFP in economies with underdeveloped financial markets. Our model predicts that the aggregate TFP of the country with the least financial development will be 40 per cent below the US level.

The impact of financial frictions is particularly large in the large-scale, manufacturing sector. While the sector-level TFP declines by less than 30 per cent in services, it declines by more than 50 per cent in manufacturing, a result broadly in line with the available sector-level productivity data on 18 OECD countries. The differential impacts of financial frictions on sector-level productivity are reflected on the higher relative prices of manufactured goods to services in financially underdeveloped economies. The model accounts for a quarter of the observed relationship between the relative price and financial development.

Financial frictions also have a significant impact on the investment rate, when measured at common fixed prices across economies with different degrees of financial development. In our model simulations, the capital to output ratio (measured at common fixed prices) declines by 15 per cent with financial frictions. Consistent with the data, this decline is almost entirely driven by the higher relative prices of manufactured investment goods in financially underdeveloped economies (Parente and Prescott, 2000; Hsieh and Klenow, 2007): the investment rates measured at respective equilibrium prices are roughly constant across economies with varying degrees of financial development.

Our quantitative analysis provides a clear decomposition of the main margins distorted by financial frictions. First, for a given set of heterogeneous production units in operation, financial frictions distort the allocation of capital among them (misallocation of capital). Second, for a given number of production units in operation, financial frictions distort the selection into entrepreneurship, with productive-but-poor individuals delaying their entry and incompetent-but-rich entrepreneurs remaining in business (misallocation of talent). Third, financial frictions distort the number of production units for a given distribution of entrepreneurial productivity in an economy. In our simulations, whereas the misallocation of capital is responsible for 90 per cent of the effect of financial frictions on the service sector TFP, it is the misallocation of talent that accounts for more than 50 per cent of the effect on the manufacturing sector TFP. The distortion on the number of entrepreneurs per se has negligible impacts on the productivity of either sector.

Finally, the differential impacts of financial frictions across sectors in our model produce an interesting testable implication on the establishment size distribution within each sector. Financial frictions, together with the resulting higher relative price of manufactured goods and lower capital rental rates and wages in the equilibrium, lead to too few entrepreneurs and too large establishments in the sector with large fixed costs, and too many entrepreneurs and too small establishments in the sector with small fixed costs. To evaluate this implication, we perform a detailed, disaggregate-level case study of Mexico and the US, and do find empirical support.

To emphasize the essence, our purpose is to provide a robust quantitative assessment of the role of financial markets in economic development. To this end, we incorporate into our framework elements of microeconomic heterogeneity and dynamics that enable a transparent mapping between our model and the data, which in turn disciplines our quantitative exercises. The two main elements are as follows: the establishment-level production technology has sector-specific non-convexities (fixed costs), with the resulting scale of operation larger in manufacturing (including the investment goods sector) than in services; and individuals are given the opportunities to overcome financial frictions over time with internal funds (self-financing).

Empirical Underpinnings Our theory is built on two premises: cross-country differences in financial development and cross-sector differences in the scale of establishments. Both of these underlying premises have strong empirical support.

The first premise, cross-country differences in financial development—underdevelopment in poor countries in particular, has been well established in the literature. King and Levine (1993a) and Beck et al. (2000) show that aggregate measures of credit and financial development are closely correlated with output per capita across countries, while La Porta et al. (1998) document that these macro indicators are strongly related to underlying institutional differences such as the enforcement of contracts, creditor protection, and so on. Banerjee and Duflo (2005) review the literature documenting micro-level evidence for credit constraints in poor countries and for the resulting misallocation of capital. In his detailed analysis of Thailand, Townsend (2010) links observed misallocation to micro-level credit constraints and shows how their relaxation through financial development leads to faster economic growth.

One empirical contribution of our paper is to establish the second premise: cross-sector differences in scale, defined as workers per establishment. Using detailed sector-level data from the OECD countries (the US in particular), we document that the average size of establishments varies substantially across broadly-defined sectors. For example, the average establishment in manufacturing is more than three times as large as that in services. This sectoral difference in establishment size is robustly observed in a wide range of countries.

In addition, we carry out a detailed case study of Mexico and the US to evaluate our model prediction on the impact of financial frictions on the establishment size distribution. We use data from the economic censuses of the two countries (based on the common, and hence comparable, North American Industrial Classification System) and a survey of small businesses in Mexico (which provides data on small-scale, mobile, and informal entrepreneurs). We examine the data at a more disaggregate level of industry classification than the manufacturing vs. services dichotomy. The average establishments in Mexico are substantially smaller, especially in the service industries. However, industries with large-scale establishments in the US (mostly in manufacturing) tend to have even larger establishments in Mexico.[2]

One could define and explore sectoral heterogeneity along other characteristics than scale, but few are as easily-measured, ubiquitous, or robust as scale. We present two sets of empirical findings that further support our decision to focus on scale. Firstly, scale seems to be closely related to productivity: using sector-level TFP data from a subset of OECD countries and price data from a broader cross-section of countries, we show that, even at a more disaggregate level of industry classification, poor countries are particularly unproductive and have higher relative prices in industries with larger scales. Secondly, in our model, financial frictions have differential impacts on sectors with different scales because our

[2]To the best of our knowledge, we are the first to document this pattern, which does conform to the conventional wisdom of a "missing middle" in less developed countries (Tybout, 2000).

notion of scale (establishment size) translates directly into financing needs. The most widely-used empirical metric of financing needs is "external dependence" constructed by Rajan and Zingales (1998). We compute the external dependence for broadly-defined sectors in the US, and find that sectors with larger scales also have larger external dependence.

Related Literature This paper contributes to a vast literature relating financial frictions to economic development, including theoretical contributions by Banerjee and Newman (1993), King and Levine (1993b), Aghion and Bolton (1997), Piketty (1997), and Lloyd-Ellis and Bernhardt (2000).

There have been relatively fewer quantitatively-oriented studies. Giné and Townsend (2004) and Jeong and Townsend (2007, 2008) have pioneered quantitative analysis of the link between financial frictions and development. They estimate models of the theoretical literature using data on Thailand's growth experience. In a related vein, Buera and Shin (2008) study the transitional dynamics of economies with underdeveloped financial markets.

Amaral and Quintin (2009) and Greenwood et al. (2009) focus on quantifying the long-run impact of financial frictions.[3] They study models that allow for sharp analytical characterizations of the equilibrium. In obtaining such tractability, they limit the role of self-financing by entrepreneurs, either by assuming a two-period working life span or by directly ruling it out.[4] In our model, entrepreneurs can at least partly overcome financial constraints over time with self-financing, and the presence of such an avenue can materially affect the quantitative results (Section 4.3.3). In addition, we complement their work by building a multi-sector model with non-convex technologies at the establishment level. These features help better capture the aggregate impact of frictions (Section 4.3.2), and generate richer disaggregate-level predictions that can be evaluated empirically.

This paper is more closely related and complementary to three others in the literature that emphasize the differential effects of financial frictions on different industries. Rajan and Zingales (1998), an empirical paper, creates an index of dependence on external sources of financing for various manufacturing industries, and tests whether industries that are particularly dependent on financing grow relatively faster in countries with more developed financial markets. We reconstruct their measure of industry-specific financial dependence for our analysis, and show that our measure of sectoral scale (workers per establishment) is closely related to external dependence.[5] Erosa and Hidalgo Cabrillana (2008) is a theoretical

[3]Erosa (2001) is an earlier contribution that abstracts from microeconomic heterogeneity.

[4]Caselli and Gennaioli (2005) focus on relatively low frequency dynastic shocks requiring the reallocation of capital from incompetent heirs to productive-but-poor entrepreneurs. They assume an exogenous saving rule, consequently disallowing endogenous saving and self-financing decisions.

[5]Beck et al. (2008) examine the independent effects of scale and financing needs.

paper showing how financial frictions can have differential effects on the productivity of manufacturing industries with different fixed cost requirements. We study a broader set of sectors, and quantify the impacts of financial frictions by introducing scale as an empirical measure related to fixed costs and financing needs. Castro et al. (2009) start from the premise that different sectors (consumption vs. investment goods) are characterized by different volatility of underlying idiosyncratic productivity shocks. They show that in economies with less risk-sharing, sectors with more volatile shocks (investment goods) are particularly unproductive. Methodologically, they use a model with a two-period life span to obtain insightful analytical characterizations. We view our approach as complementary to theirs. Our model, with its rich dynamics and heterogeneity, eludes analytical tractability, but enables a clear mapping into data and a direct quantitative interpretation of the results. Our quantitative analysis emphasizes sectoral scale differences and cross-country differences in external financing—readily-measured and robust features in the data.

Finally, we add to the broader literature on the role of micro-level distortions (Hopenhayn and Rogerson, 1993; Guner et al., 2008; Restuccia and Rogerson, 2008). Our results complement the empirical findings of Hsieh and Klenow (2009) in particular, who find a factor-of-two difference in manufacturing TFP due to the misallocation of capital and labor. We explicitly model one source of such misallocation (financial frictions), and include the service sector in our analysis. Furthermore, we explore how distortions affect the entry and exit decisions of establishments.

2 Facts

This section documents the key empirical facts motivating our study. First, we show that the price of manufactured goods is high relative to services in less developed countries, and that such relative prices are closely linked to sector-level relative productivity. Second, we point out that there are large differences in scale (establishment size) across sectors, with the manufacturing sector having much larger establishments on average than the service sector. Next, we emphasize this sectoral patterns in relative scale and relative productivity by showing that this relationship holds at a more disaggregate level as well: goods and services that are produced with large-scale technologies tend to be relatively more expensive in less developed economies, reflecting their low relative productivity in these industries.

Relative Prices, Relative Productivity, and Economic Development In poor countries, the final prices of manufactured goods are high relative to services.[6] The left panel of Figure 1 shows this fact using the 1996 ICP benchmark by plotting the relative price of manufactured goods to services against PPP output per worker.[7] The log relative price of manufactured goods to services have a strong negative relationship with the log output per worker across countries. The slope coefficient of -0.45 is highly significant with a standard error of 0.06, and the regression has an R^2 of 0.40.

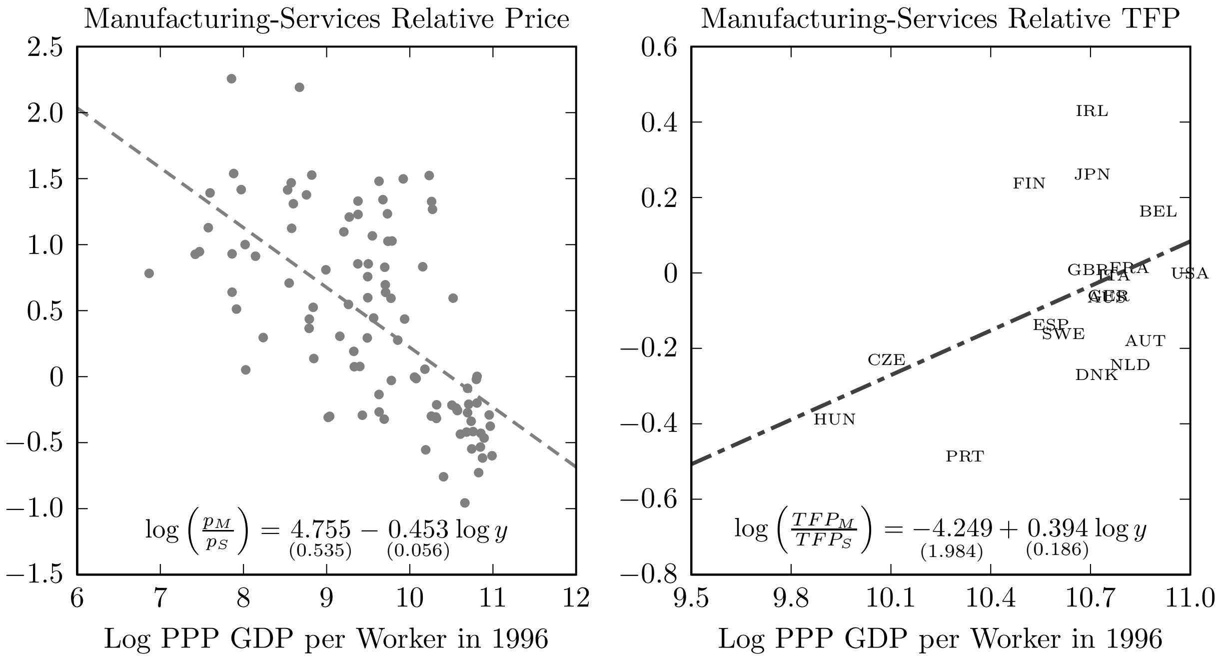

Fig. 1: Relative prices (left panel) and relative productivity (right panel) against output per worker. In the left panel, the log relative price of manufactured goods to services is plotted against the log of PPP output per worker (y) for each country. The price data are from the 1996 ICP benchmark, and the output data are from the Penn World Table 6.1. In the right panel, we construct sector-level TFP from the Productivity Level Database of the Groningen Growth and Development Centre. The log relative productivity of manufacturing to services is plotted against log output per worker. The standard errors of the regression coefficients are reported in parentheses.

---

[6] A related fact, first documented by Balassa (1964) and Samuelson (1964), is that the relative price of tradable goods is higher in poor countries.

[7] ICP stands for the International Comparison Programme of the United Nations. There are 115 ICP benchmark countries in 1996. To maintain a consistent sample, we present results based on the 102 countries for which Beck et al. (2000) constructed data on financial development. The results using all 115 countries are virtually identical. Here the relative price is compiled by creating Geary-Khamis aggregated prices for manufactured goods and services using 27 disaggregate product categories. The goods category consists of clothing, nine food and beverage categories, footwear, fuel, furniture/floor coverings, household appliances, household textiles/other household goods, machinery/equipment, tobacco, and transportation equipment. The services category consists of communication, education, medical/health, recreation/culture, rent/water, restaurants/hotels, and transportation services. We do not classify five final goods price categories: construction, changes in stocks, collective consumption by government, net foreign balance, and other goods and services.

Within many models, this relationship can be interpreted as lower relative TFP of manufacturing to services in less developed countries. Indeed, in models with constant-returns-to-scale production functions and equal factor shares across sectors, relative prices equal the inverse of relative TFP (Hsieh and Klenow, 2007). Differences in factor shares and the relative supply of factors (e.g., higher levels of capital per worker) could break this inverse relationship, but empirically factor shares do not vary much across sectors (Valentinyi and Herrendorf, 2008).

Furthermore, the available data on sector-level productivity across countries support the relative productivity interpretation of relative prices. Using the Productivity Level Database (Inklaar and Timmer, 2008) of the Groningen Growth and Development Centre (GGDC), we construct sector-level TFP measures for manufacturing and services value-added for 18 OECD countries.[8] The log relative sector-level TFP of manufacturing to services is plotted against log GDP per worker in the right panel of Figure 1. The regression coefficient is 0.39. Given the smaller sample size, the standard error is larger (0.19), but the estimate is still significant at the five-percent level, with an R^2 of 0.22. Thus, we find that the relative TFP of manufacturing to services is positively correlated with output per worker, just as relative prices are negatively correlated with output per worker, and the magnitudes are similar (0.39 and -0.45, respectively). This is comforting in two ways. Firstly, it confirms the predicted relationship between relative prices and relative productivity. Secondly, it gives evidence for value-added rather than final goods prices: we will explore these patterns in relative productivity, relative prices, and output per worker with a model that applies more transparently to value-added.

Relative Prices and Financial Development A common measure of a country's level of financial development is its ratio of external financing to GDP, where external financing is defined as the sum of private credit, private bond market capitalization, and stock market capitalization.[9] As is well documented in the literature, this measure is very closely correlated

[8]The database contains data for 19 of the 30 OECD countries (Australia, Austria, Belgium, Czech Republic, Denmark, Spain, Finland, France, Germany, Hungary, Ireland, Italy, Japan, Luxembourg, the Netherlands, Portugal, Sweden, the UK and the US) plus Slovenia. A double deflation method is used to construct real value-added by sector. While domestic intermediates are deflated by their source sector, imported intermediates are not. Imported intermediates are particularly important for two small countries with very high trade to GDP ratios, Luxembourg (112 per cent) and Slovenia (107 per cent). Trade to GDP ratios average 38 per cent for the remaining countries. These two countries are also strong outliers in terms of their sectoral TFP numbers (in opposite directions). We therefore exclude them in our analysis.

[9]We use the data from La Porta et al. (1998); Rajan and Zingales (1998); Beck et al. (2000), which report the market value of stock market capitalization. We discount the reported stock market capitalization for two reasons. First, some equity (roughly three per cent in the US) is internally held by the management (Clementi and Cooley, 2009). Second, the market value of equity overstates the book value which is conceptually closer to the financed capital investment in our model. Globally, the average of the book to market ratios is 0.45,

with output per worker: the regression of log GDP per worker on external financing to GDP yields a slope coefficient of 1.08, with a standard error of 0.13 and an R^2 of 0.48. Therefore it is unsurprising that the relationship between log relative prices and external financing to GDP ratios is similar to the one between log relative prices and log GDP per worker in the left panel of Figure 1. The estimated slope coefficient is -0.67, with a standard error of 0.10 and an R^2 of 0.38. For the 18 OECD countries with the sector-level TFP data, the regression coefficient of the log manufacturing-services relative TFP on external financing to GDP ratios is 0.08. While it has the same sign as the slope in the right panel of Figure 1, it is not significant with a standard error of 0.08.

The strength of the relationship suggests that financial development is potentially closely related to the patterns in relative prices, relative productivity, and output per worker across countries. In the model we develop, it is financial development that is the causal force behind these cross-country differences.

Scale Differences across Sectors Another key empirical fact that motivates our study is the clear differences in the scale of production units across broadly-defined sectors. Our interpretation is that the observed sectoral scale differences reflect differences in micro-level production technologies across these sectors. We will argue that these technological differences interact with financial development, so that financial development affects large-scale (e.g. manufacturing) and small-scale sectors (e.g. services) differently.

Here we use two measures of scale: workers per establishment and workers per enterprise. Establishments are locations of business, so that a single enterprise, Walmart, for example, may have multiple establishments. Table 1 presents measures of average scale across broadly-defined final goods sectors in the US and other OECD countries, along with other sector-level characteristics like financial dependence and factor intensities.[10]

The first column is based on data from the 2002 US Economic Census, which uses an establishment basis and the NAICS classification. The second column is based on the OECD Structural Statistics for Industry and Services (SSIS) data for 2002. These data follow

while for the US it is 0.24. We choose to multiply the reported stock market capitalization by 0.33. Our cross-country regression results are robust to this correction, because the stock market is important for only a small number of countries (notably the US) in the data. We also note that some private credit is used for consumer credit (e.g. credit cards). In the US flow of funds data, this amounts to roughly nine per cent of private credit during the 1990s. As we do not have data on consumer credit for most other countries—and this number is likely to be even smaller in those countries, we do not adjust for it.

[10]These sectors are constructed to be consistent with the final goods categories in the ICP data. Manufacturing consumption includes food/beverages, textiles, clothing, medicine, furniture, appliances, TVs and radios, cars, household items, and media. Equipment includes all manufactured equipment not included in consumption. Together, these two constitute manufacturing. Services include accommodation/food services, arts/entertainment, communication, education, FIRE, health, retail, sewage, transportation, and wholesale.

	Workers per Establishment	Workers per Enterprise	External Dependence	Capital Share
	US	OECD	US	US
Manufacturing	47	28	0.21	0.31
Services	14	8	0.09	0.27

Table 1: For the scale measures, we take simple averages in the US census data and the OECD SSIS data. The OECD data cover nine countries, and are complete at the enterprise level: Czech Republic, France, Germany, Hungary, Slovakia, Poland, Portugal, the UK, and the US. The external dependence ratio is calculated using the formula of Rajan and Zingales (1998), who measure the ratio of the difference between capital expenditures and cash flow to capital expenditures. Capital shares are from Valentinyi and Herrendorf (2008).

the common ISIC 3.2 four-digit classification, enabling comparison across OECD countries. Enterprise-level data permit comparison over the largest set of countries.[11]

Whether establishments or enterprises are the unit of measurement, the average scale varies considerably across sectors. In the US census data, establishments in manufacturing is on average 3.4 times as large as those in services (47 vs. 14). The relative scale of manufacturing to services is even larger with enterprises as the unit of observation. For the OECD average, the ratio is 3.5 (28 vs. 8). Establishments are our preferred unit of analysis because we think they embody production technologies, although we acknowledge that some technologies (e.g. the distribution system of Walmart) may be at the firm level. Data availability dictates which measure we use in certain cases.

In our model, financial frictions have differential impacts on sectors with different scales because our notion of scale (workers per establishment or enterprise) coincides with financing needs. Support for this interpretation can be obtained by comparing other measures of financing needs across sectors. The third column reports the measures of external dependence (Rajan and Zingales, 1998) that we construct using the US Compustat data for 1993–2003.[12] In the Compustat data, firms producing final manufacturing goods are substantially more financially-dependent than those producing services, with a median of 0.21 vs. 0.09.[13]

[11]There are subtle differences in the definition of workers between the two samples. In particular, the SSIS data measure "number of persons engaged," which can include some temporary or contract workers. The census data record "number of employees," which excludes proprietors. For some countries, SSIS data have both measures, and the two mirror each other closely.

[12]Rajan and Zingales measure the ratio of the difference between capital expenditures (Compustat #128) and cash flow (Compustat #110, or the sum of #123, 125, 126, 106, 213, and 217, for format code 7) to capital expenditures. In order to negate the influence of outliers in noisy firm-level data, they take the total capital expenditures and total cash flow over the sample period to compute firm-specific numbers, and then pick the median value within an industry as the industry-specific value. Note that Rajan and Zingales only included the manufacturing sector in their study.

[13]Alternative measures of firms' financing needs give a similar picture. For example, we have constructed a measure of setup costs. In particular, for each firm in Compustat, we have located the first period with

While external dependence may be a more direct measure of financing needs or investment requirements, we decide to focus on scale (employment), a measure that is available for all firms in the economy, as opposed to just the publicly-traded firms in Compustat.[14]

Finally, the fourth column reports the sectoral differences in factor intensities. These numbers are from Valentinyi and Herrendorf (2008), and correspond to the capital share of gross output in each of these broadly-defined sectors, calculated using input-output data from the Industry Accounts of the Bureau of Economic Analysis. Unlike differences in scale and external dependence, the variation in factor intensity is relatively small.[15] We follow the recent literature (Chari et al., 1997; Hsieh and Klenow, 2007) and build a model that abstracts from these differences. Our focus is instead on the large observed differences in scale, and the corresponding financing needs, across sectors.

We note that the distinctions in Table 1 hold even at a more disaggregate level. Broadly-defined sub-sectors within manufacturing tend to have larger scales and higher dependence on external financing than services and construction, but all have comparable factor shares.[16]

Relative Prices, Productivity, and Scale We have established that the large-scale, manufacturing sector has higher relative prices and lower relative productivity in less developed economies. A natural question to ask is whether relative prices, relative productivity, and the scale of production technologies are related at a more disaggregate level as well.

We examine this issue using disaggregate ICP price data from its 1996 benchmark. The scale of an industry is constructed by averaging across eight countries for which comparable data are available.[17] We then map ICP categories into closely related groups of industries and calculate the average scale for these industry groups.[18] Finally, we run a cross-country regression of 2,794 disaggregated ICP price data from 112 countries on log output per worker,

positive excess investment, and calculated the average excess investment over following consecutive periods with positive excess investment. For the firms in manufacturing, the ratio of this measure over annual sales has a median of 0.65, while for those in services, the median is 0.12.

[14]See Davis et al. (2006) for an example where conclusions drawn from the universe of firms and from the set of publicly-traded firms are very different.

[15]Chari et al. (1997) also find little difference in factor shares between investment and consumption goods sectors, with, if anything, investment goods producers being more labor-intensive.

[16]See the earlier version of this paper, Buera et al. (2009).

[17]UNIDO SSIS data—which cover all industries, not just manufacturing—are available for eight countries (Czech Republic, France, Germany, Hungary, Poland, Portugal, Slovakia, and the UK), but only at the enterprise level. We therefore use number of persons engaged per enterprise as our measure of scale. For any given industry at a disaggregate level, there is a strong correlation in scale across countries. Still, we average across countries to smooth out idiosyncratic variations that may arise from local market structures, government regulations, and so on.

[18]A reliable mapping could not be done for four of the 29 ICP categories: other household goods, operation of transportation equipment, other goods and services, and collective consumption by the government. Also, due to the lack of comparable data on agriculture, only the scale of food manufacturing establishments could be used. None of the food categories appears to be outliers.

log industry scale, and their interaction. The estimation result is (with standard errors in parentheses):

$$\log\left(\frac{p_{i,j}}{PPP_i}\right) = \underset{(0.56)}{-7.48} + \underset{(0.05)}{0.29}\log y_i + \underset{(0.15)}{1.01}\log\ \bar{l}_j - \underset{(0.02)}{0.10}\log y_i \log \bar{l}_j, \quad R^2 = 0.22,$$

where $p_{i,j}$ is the 1996 price of industry j output in country i, PPP_i is the average (Geary-Khamis) 1996 price level in country i, y_i is the output per worker of country i in 1996 international prices, $\bar{l}_j$ is the average number of workers engaged per enterprise for industry j. The negative coefficient on the interaction term indicates that prices of the output of industries with larger scales are relatively higher in low income countries. Given the log difference between manufacturing and service scales in Table 1—i.e., log(47/14) = 1.21, the coefficient of -0.10 implies a relative price elasticity with respect to output per worker of 0.12, about one-third of the full relationship in Figure 1.

We work out a parallel exercise using the disaggregate industry-level TFP data (29 industries) for the 18 OECD countries. We obtain a very similar result:

$$\log\left(\frac{TFP_{i,j}}{TFP_i}\right) = \underset{(0.28)}{12.18} - \underset{(0.32)}{1.10}\log y_i - \underset{(1.00)}{3.52}\log\ \bar{l}_j + \underset{(0.09)}{0.32}\log y_i \log \bar{l}_j, \quad R^2 = 0.08.$$

The magnitude of the interaction term is substantially higher. With the difference between manufacturing and service scales in Table 1, the coefficient of 0.32 implies a relative TFP elasticity with respect to output per worker of 0.4, almost the full estimated elasticity between the two in the data (right panel, Figure 1).

The general magnitude and significance of these two regression results are quite robust. The significance of the results at the five-percent level are robust to clustering standard errors by country or by ICP category. Also, we can use $p_{i,j}$ and $TFP_{i,j}$ without scaling by PPP_i and TFP_i respectively. The results are nearly identical, except that the coefficient on $\log y_i$ is substantially larger in the relative price regression and smaller in the relative TFP regression. More generally, alternative specifications that use country-specific fixed effects instead of controlling for $\log y_i$ or category-specific fixed effects instead of controlling for $\log \bar{l}_j$ yield very similar results. Finally, substituting countries' ratios of external finance to GDP for $\log y_i$ produces again very similar coefficients for the relative price regression, and smaller but still significant coefficients (0.13) for the relative TFP regression.

We conclude that, even at a more disaggregate level, less developed countries have relatively high prices and low productivity in industries with larger scales. This is further evidence supporting our emphasis on sectoral scale differences.

3 Model

We model an economy with two sectors, S (small scale, services) and M (large scale, manufacturing). In each sector, there are two occupations: workers and entrepreneurs. The output of the serivce sector is used for consumption only. The manufactured goods are used for consumption and investment, and are the numeraire.

There are a measure N of infinitely-lived individuals, who are heterogeneous in their wealth and the quality of their entrepreneurial ideas or talent, $z = (z_S, z_M)$. Individuals' wealth is determined endogenously by forward-looking behavior. The vector of entrepreneurial ideas is drawn from a distribution $\mu(z)$. Entrepreneurial ideas "die" with a constant hazard rate of $1-\gamma$, and a new vector of ideas is independently drawn from $\mu(z)$. The parameter γ therefore controls the persistence of the entrepreneurial idea or talent process.[19]

In each period, individuals choose their occupation: whether to work for a wage or to operate a business in sector S or M. Their occupational choices are based on their comparative advantage as an entrepreneur (z) and their access to capital. Access to capital is limited by agents' wealth through an endogenous collateral constraint, because capital rental contracts may not be perfectly enforceable in our model.

One entrepreneur can operate only one production unit (establishment) in a given period. Entrepreneurial ideas are inalienable, and there is no market for managers or entrepreneurial talent. The way we model an establishment draws upon the span of control of Lucas (1978) and per-period fixed costs as in Rossi-Hansberg and Wright (2007).

Preferences Individual preferences are described by the following expected utility function over sequences of consumption $c_t = (c_{S,t}, c_{M,t})$:

$$U(c) = \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t u(c_t)\right], \quad u(c_t) = \frac{1}{1-\sigma}\left(\psi c_{S,t}^{1-1/\varepsilon} + (1-\psi) c_{M,t}^{1-1/\varepsilon}\right)^{\frac{1-\sigma}{1-1/\varepsilon}}, \tag{1}$$

where β is the discount factor, σ the coefficient of relative risk aversion (and the reciprocal of the elasticity of intertemporal substitution), ε the intratemporal elasticity of substitution between services and manufactured goods, and ψ controls the share of services in overall consumption expenditure. The expectation is over the realizations of entrepreneurial ideas (z), which depend on stochastic death of ideas ($1-\gamma$) and on draws from $\mu(z)$.

[19]This shock can be interpreted as changes in market conditions that affect the profitability of individual skills.

Technology At the beginning of each period, an individual with vector of entrepreneurial ideas z and wealth a chooses whether to work for a wage w or operate a business in either sector $j = S, M$. To operate a business in a sector, individuals must pay a sector-specific per-period fixed cost of κ_j, in units of the sector's output. The crucial assumption is that the fixed cost to run an establishment in the manufacturing sector is higher than that in the service sector, $\kappa_M > \kappa_S$. This will generate the scale difference between the two sectors that we observe in the data. Note that κ_j is fixed costs that need to be paid in every period of operation. We discuss in Section 4.3.1 how the results will change if one were to introduce one-time fixed setup costs of starting a business.

After paying the fixed cost, an entrepreneur with talent z_j produces using capital (k) and labor (l) according to:

$$z_j f(k, l) = z_j k^\alpha l^\theta,$$

where α and θ are the elasticities of output with respect to capital and labor, and $\alpha + \theta < 1$, implying diminishing returns to scale in variable factors at the establishment level. Note that the factor elasticities are assumed to be the same in both sectors, consistent with the empirical findings in the literature (Chari et al., 1997; Valentinyi and Herrendorf, 2008).

Given factor prices w and R (rental rate of capital), the profit of an entrepreneur is:

$$\pi_j(k, l; R, w, p) = p_j z_j k^\alpha l^\theta - Rk - wl - (1 + r)p_j\kappa_j,$$

where r is the interest rate and p_j is the price of sector j output. Service sector output is only for consumption, and the manufacturing sector produces investment goods and manufactured consumption goods. We normalize p_M to one. For later use, we define the optimal level of capital and labor inputs when production is not subject to financial constraints:

$$\left(k_j^u(z_j), l_j^u(z_j)\right) = \arg\max_{k,l} \left\{p_j z_j k^\alpha l^\theta - Rk - wl\right\}.$$

The key feature of this technology is that the fixed costs introduce non-convexity. For any strictly positive fixed cost κ_j, the technology is feasible only if operated above the minimum scale; that is, $z_j k^\alpha l^\theta \geq (1 + r)\kappa_j$.[20]

[20] The model can be extended to allow for a choice of technologies within each sector: a technology with a small fixed cost and low productivity, $A_{1,j} z_j k^\alpha l^\theta - \kappa_1$, and a technology with a large fixed cost and high productivity, $A_{2,j} z_j k^\alpha l^\theta - \kappa_2$, with $A_{2,j} > A_{1,j}$ for $j = S, M$ and $\kappa_2 > \kappa_1$. Note that in this extension fixed costs are technology-specific but not sector-specific. One can think of our current setup as a case where the productivity gains associated with the κ_2-technology is substantially larger for the manufacturing sector; that is, $A_{2,M} \gg A_{2,S}$.

Credit and Rental Markets Individuals have access to competitive financial intermediaries, who receive deposits, rent capital k at rate R to entrepreneurs, and lend entrepreneurs the fixed cost $p_j \kappa_j$. In the benchmark model, we restrict the analysis to the case where both borrowing and capital rental are within a period—that is, $a \geq 0$. The zero-profit condition of the intermediaries implies $R = r + \delta$, where r is the deposit and lending rate and δ is the depreciation rate.

Borrowing and capital rental by entrepreneurs are limited by imperfect enforceability of contracts. In particular, we assume that, after production has taken place, entrepreneurs have the option to renege on the contracts. In such cases, the entrepreneurs can keep a fraction $(1-\phi)$ of the undepreciated capital and the revenue net of labor payments: $(1-\phi)\left[p_j z_j f(k,l) - wl + (1-\delta)k\right]$. The only punishment is the garnishment of their financial assets deposited with the financial intermediary, a. In the following period, this entrepreneur in default will regain access to financial markets, and will not be treated any differently despite the history of default.

Note that ϕ indexes the strength of an economy's legal institutions enforcing contractual obligations. This one-dimensional parameter captures the extent of frictions in the financial market due to imperfect enforcement of credit and rental contracts. This parsimonious specification allows for a flexible modeling of limited commitment that spans economies with no credit ($\phi = 0$) and those with perfect credit markets ($\phi = 1$).

We consider equilibria where the borrowing and capital rental contracts are incentive-compatible and are hence fulfilled. In particular, we study equilibria where the rental of capital is quantity-restricted by an upper bound $\bar{k}^j(a, z_j; \phi)$, which is a sector-specific function of the individual state (a, z). We choose the rental limits $\bar{k}^j(a, z_j; \phi)$ to be the largest limits that are consistent with entrepreneurs choosing to abide by their credit contracts. Without loss of generality, we assume $\bar{k}^j(a, z_j; \phi) \leq k_j^u(z)$, where k_j^u is the profit-maximizing capital inputs in an unconstrained (static) problem in sector j.

The following proposition provides a simple characterization of the set of enforceable contracts and the rental limits $\bar{k}^j(a, z_j; \phi)$ for $j = S, M$.

Proposition 1 *Capital rental k in sector j by an entrepreneur with wealth a and talent z_j is enforceable if and only if*

$$\max_l \{p_j z_j f(k,l) - wl\} - Rk - (1+r)\, p_j \kappa_j + (1+r)\, a$$
$$\geq (1-\phi)\left[\max_l \{p_j z_j f(k,l) - wl\} + (1-\delta)\, k\right]. \quad (2)$$

Therefore, the upper bound on capital rental that is consistent with entrepreneurs choosing to abide by their credit contracts can be represented with a function $\bar{k}^j(a, z_j; \phi)$, which is increasing in a, z_j and ϕ.

Condition (2) states that an entrepreneur must end up with (weakly) more economic resources when he fulfills his credit obligations (left-hand side) than when he defaults (right-hand side). This static condition is sufficient to characterize enforceable allocations because we assume that defaulting entrepreneurs regain access to financial markets in the following period.

This proposition also provides a convenient way to operationalize the enforceability constraint into a simple rental limit $\bar{k}^j(a, z_j; \phi)$.[21] As long as the unconstrained level of capital rental is not enforceable, the rental limit $\bar{k}^j(a, z_j; \phi)$ is implicitly defined as the larger root of the equation given by the equality in condition (2). Rental limits increase with the wealth of entrepreneurs, because the punishment for defaulting (loss of collateral) is larger. Similarly, rental limits increase with the talent of an entrepreneur because defaulting entrepreneurs keep only a fraction $1 - \phi$ of the output. In the rest of the paper, we restrict individuals' capital inputs to be less than or equal to the rental limit $\bar{k}^j(a, z_j; \phi)$.

While the enforceability of contracts as measured by ϕ is not sector-specific, the equilibrium enforceable rental contracts, as captured by the rental limits $\bar{k}^j(a, z_j; \phi)$, do vary across sectors because of the differences in technology and output prices across sectors.

Recursive Representation of Individuals' Problem Here we discuss the problem solved by individuals. In particular, we define the value for an individual before the occupational choice, $v(a, z)$, as well as the value of being a worker, $v^W(a, z)$, and being an entrepreneur in sector j, $v^j(a, z)$, for $j = S, M$.

Individuals maximize (1) by choosing sequences of consumption, financial wealth, occupations (including the sector), and capital/labor inputs if they choose to be entrepreneurs, subject to a sequence of period budget constraints and rental limits.

At the beginning of a period, an individual's state is summarized by his wealth a and the vector of abilities z. He then chooses whether to be a worker or to be an entrepreneur in sector S or M for the period. The value for him at this stage, $v(a, z)$, is the maximum over the value of being a worker, $v^W(a, z)$, and the value of being an entrepreneur in sector j, $v^j(a, z)$, for $j = S, M$:

$$v(a, z) = \max\left\{v^W(a, z), v^S(a, z), v^M(a, z)\right\}. \tag{3}$$

Note that the value of being a worker, $v^W(a, z)$, depends on his assets a, and on his entrepreneurial ideas z which may be implemented at a later date. Similarly, the value of being

[21] In general, the set of enforceable capital rentals dictated by (2) does not exactly correspond to $k \leq \bar{k}^j(a, z_j; \phi)$. For example, an entrepreneur who is only offered a very low (and unprofitable) level of capital will default if $p_j \kappa_j$ is large enough, $p_j \kappa_j > a$. Notwithstanding this, the solution to the individual's problem subject to (2) coincides with the solution to the individual's problem subject to the simpler rental limit $k \leq \bar{k}^j(a, z_j; \phi)$. See the proof in the appendix.

an entrepreneur in sector j, $v^j(a,z)$, depends on the entire vector of entrepreneurial ideas, as he may switch sectors at a later date. We denote the policy function associated with the optimal occupation choice by $o(a,z) \in \{W,S,M\}$.

As a worker, an individual chooses consumption bundle $c = (c_S, c_M)$ and the next period's assets a' to maximize his continuation value subject to the period budget constraint:

$$\begin{aligned} v^W(a,z) &= \max_{c,a' \geq 0} u(c) + \beta\{\gamma v(a',z) + (1-\gamma)\mathbb{E}_{z'}[v(a',z')]\} \\ &\text{s.t. } p \cdot c + a' \leq w + (1+r)a, \end{aligned} \tag{4}$$

where w is her labor income, and p denotes the vector of goods prices. The continuation value is a function of the end-of-period state (a',z'), where $z' = z$ with probability γ and $z' \sim \mu(z')$ with probability $1-\gamma$. In the next period, he will face an occupational choice again, and the function $v(a,z)$ appears in the continuation value.

Alternatively, individuals can choose to become an entrepreneur in sector j. The value function of being an entrepreneur in sector j is as follows.

$$\begin{aligned} v^j(a,z) &= \max_{c,a',k,l \geq 0} u(c) + \beta\{\gamma v(a',z) + (1-\gamma)\mathbb{E}_{z'}[v(a',z')]\} \\ &\text{s.t. } p \cdot c + a' \leq p_j z_j f(k,l) - Rk - wl - (1+r)p_j \kappa_j + (1+r)a \\ &\quad k \leq \bar{k}^j(a, z_j; \phi) \end{aligned} \tag{5}$$

Note that an entrepreneur's income is given by period profit $p_j z_j f(k,l) - Rk - wl$ net of fixed costs $(1+r)p_j\kappa_j$ plus the return to his initial wealth, and that his choices of capital inputs are constrained by the rental limit $\bar{k}^j(a, z_j; \phi)$.

Stationary Competitive Equilibrium A stationary competitive equilibrium is composed of: an invariant distribution of wealth and entrepreneurial ideas $G(a,z)$, with the marginal distribution of z denoted with $\mu(z)$; occupation choice function $o(a,z)$, and policy functions $c_S(a,z)$, $c_M(a,z)$, $a'(a,z)$, $l(a,z)$, and $k(a,z)$; rental limits $\bar{k}^j(a,z_j;\phi)$, $j = S, M$; and prices w, R, r, and p such that:

1. Given $\bar{k}^j(a,z_j;\phi)$, w, R, r and p, the individual policy functions $o(a,z)$, $c_S(a,z)$, $c_M(a,z)$, $a'(a,z)$, $l(a,z)$, $k(a,z)$ solve (3), (4) and (5);

2. Financial intermediaries make zero profit: $R = r + \delta$;

3. Rental limits $\bar{k}^j(a,z_j;\phi)$ are the most generous limits satisfying condition (2), with $\bar{k}^j(a,z_j;\phi) \leq k^u_j(z_j)$;

4. Capital rental, labor, services, and manufactured goods markets clear:

$$\frac{K}{N} \equiv \int k(a,z)\, G(da,dz) = \int aG(da,dz) \qquad \text{(Capital rental market)}$$

$$\int l(a,z)\, G(da,dz) = \int_{\{o(a,z)=W\}} G(da,dz) \qquad \text{(Labor market)}$$

$$\int c_S(a,z)\, G(da,dz) = \int_{\{o(a,z)=S\}} \left[zk(a,z)^{\alpha}\, l(a,z)^{\theta} - \kappa_S\right] G(da,dz)$$

$$\text{(Services market)}$$

$$\int c_M(a,z)\, G(da,dz) + \delta\frac{K}{N} = \int_{\{o(a,z)=M\}} \left[zk(a,z)^{\alpha}\, l(a,z)^{\theta} - \kappa_M\right] G(da,dz)$$

$$\text{(Manufactured goods market)}$$

5. The joint distribution of wealth and entrepreneurial ideas is a fixed point of the equilibrium mapping:

$$G(a,z) = \gamma \int_{\{(\tilde{a},\tilde{z})|\tilde{z}\leq z, a'(\tilde{a},\tilde{z})\leq a\}} G(d\tilde{a},d\tilde{z}) + (1-\gamma)\,\mu(z) \int_{\{(\tilde{a},\tilde{z})|a'(\tilde{a},\tilde{z})\leq a\}} G(d\tilde{a},d\tilde{z}).$$

Perfect-Credit Benchmark To clarify the basic mechanics of the model, we analyze the perfect credit benchmark, $\phi = 1$. This is an economy with unconstrained within-period borrowing and capital rental for production—that is, $\bar{k}^j(a, z_j, \phi) = k_j^u(z_j)$ for all a—but without between-periods borrowing or consumption insurance. We present two results characterizing the production side of the perfect-credit economy under the assumption that entrepreneurial talents for the two sectors follow mutually independent Pareto distributions with the same tail parameter η, $(z_S, z_M) \sim \eta^2 (z_S z_M)^{-(\eta+1)}$ for $z_j \geq 1$, $j = S, M$. This assumption permits approximate closed-form expressions for net sectoral production functions (sectoral output net of fixed costs), factor shares, and the establishment size distribution. The assumption of mutually-independent Pareto distributions also implies that the establishment size distribution within each sector exhibits a thick right tail, a salient feature of the data.[22] These characterizations will help us pin down the technological parameters of the model using data on establishment size distributions across and within sectors.[23]

[22]In the other extreme case, if entrepreneurial ideas were to be perfectly correlated across sectors, the model would predict that even the smallest manufacturing establishment will have more employees than the largest service establishment.

[23]We solve the perfect-credit benchmark in two steps. First, given an aggregate supply of capital and the intratemporal (homothetic) consumption decisions, we solve for optimal production decisions, occupation choices, and prices. We then use the wage and entrepreneurial profits coming from the production side of the economy to solve for the saving decisions of individuals facing idiosyncratic shocks to entrepreneurial ideas. By aggregating over individuals, we obtain the aggregate supply of capital. A stationary equilibrium with perfect credit markets is a nested fixed point of these two problems.

The first result is that the net output of a sector is given by a Cobb-Douglas, constant-returns-to-scale function of the population size (N), sectoral capital inputs (K_j) and labor inputs (L_j).

Proposition 2 *Assume that entrepreneurial talents for the two sectors follow mutually-independent Pareto distributions with the same tail parameter* η, $(z_S, z_M) \sim \eta^2 (z_S z_M)^{-(\eta+1)}$ *for* $z_j \geq 1$, $j = S, M$, *and that active entrepreneurs are a small fraction of the population. Then the output of a sector, net of fixed costs, equals:*

$$Y_j (K_j, L_j; N) = A_j N^{\frac{1/\eta}{\alpha+\theta+1/\eta}} K_j^{\frac{\alpha}{\alpha+\theta+1/\eta}} L_j^{\frac{\theta}{\alpha+\theta+1/\eta}},$$

where

$$A_j = \left[1 - \frac{p_j \kappa_j}{p_j \kappa_j + w}\left(1 - \alpha - \theta - \frac{1}{\eta}\right)\right] \left[\frac{\eta (1-\alpha-\theta)}{\eta (1-\alpha-\theta) - 1}\right]^{\frac{1}{1+\eta(\alpha+\theta)}} \left[\frac{p_j \kappa_j}{w} + 1\right]^{\frac{1-\eta(1-\alpha-\theta)}{1+\eta(\alpha+\theta)}}.$$

It follows that, as in the standard neoclassical sectoral growth model, the elasticities of output with respect to capital and labor are constant, $\frac{\alpha}{\alpha+\theta+1/\eta}$ and $\frac{\theta}{\alpha+\theta+1/\eta}$, respectively. Unlike in the standard model, however, the elasticities are not equal to the factor shares, because entrepreneurs earn rents. In particular, payments to capital as a share of income equals:

$$s_{K,j} = \frac{RK_j}{Y_j(K_j, L_j; N)} = \frac{\alpha}{1 - \frac{p_j \kappa_j}{p_j \kappa_j + w}\left(1 - \alpha - \theta - \frac{1}{\eta}\right)}.$$

For realistic parameterizations of the model, $1 - \alpha - \theta - \frac{1}{\eta}$ is close to zero, and hence factor shares in the two sectors are approximately equal.[24]

Our second result pertains to the establishment size distribution in the perfect-credit benchmark. In particular, we show that the establishment size in each sector follows a Pareto distribution with tail coefficient $\eta (1-\alpha-\theta)$, and that the overall establishment size distribution in the economy is a mixture of Pareto distributions. We also show that there is a direct mapping between the ratio of fixed costs to wage ($p_j \kappa_j / w$) and the ratio between the average establishment sizes ($\bar{l}_j$) of the two sectors.

Proposition 3 *Assume that entrepreneurial talents for the two sectors follow mutually-independent Pareto distributions with the same tail parameter* η, $(z_S, z_M) \sim \eta^2 (z_S z_M)^{-(\eta+1)}$ *for* $z_j \geq 1$, $j = S, M$, *and that active entrepreneurs are a small fraction of the population. Then the establishment size distribution in each sector follows the power law:*

$$\Pr\left[\tilde{l}_j > l\right] = \left(\frac{l(\hat{z}_j)}{l}\right)^{\eta(1-\alpha-\theta)}, \quad l \geq l(\hat{z}_j),$$

[24] One technical condition is $1 - \alpha - \theta - \frac{1}{\eta} > 0$. See the proof in the appendix.

where $l(\hat{z}_j)$ is the employment in the marginal establishment of sector j. Furthermore, the establishment size distribution in the aggregate economy is given by a mixture of Pareto distributions:

$$\Pr\left[\tilde{l} > l\right] = n_S \left(\frac{l(\hat{z}_S)}{l}\right)^{\eta(1-\alpha-\theta)} + n_M \left(\frac{l(\hat{z}_M)}{\max\{l, l(\hat{z}_M)\}}\right)^{\eta(1-\alpha-\theta)}, \quad l \geq l(\hat{z}_M),$$

where n_S and n_M are respectively the fraction of small and large-scale establishments in the economy, with $n_S + n_M = 1$. Also, the ratio of the average establishment sizes of the two sectors is:

$$\frac{\bar{l}_j}{\bar{l}_{j'}} = \frac{p_j \kappa_j + w}{p_{j'} \kappa_{j'} + w}.$$

This last result suggests a simple way of identifying the relative magnitude of sector-specific fixed costs from their relative scale. In our model, the large scale of a sector arises from the large fixed costs at the establishment level. In addition, the tail of the establishment size distribution identifies the parameter governing the distribution of entrepreneurial talents. In the next section, these observations will enable us to calibrate the model in a transparent manner. We then quantify the impact of financial frictions.

4 Quantitative Analysis

In this section, we first calibrate the perfect-credit benchmark of our model economy to the US economy. We then conduct experiments to assess the effect of financial frictions. In particular, we vary ϕ, the parameter governing the degree of financial frictions, to obtain variations in external finance to GDP ratios that are comparable to the range observed in a cross-section of countries. We evaluate our model predictions for aggregate/sector-level TFP, output per worker, and capital to output ratios.

In our quantitative analysis we hold fixed all technological parameters across countries, and vary only the parameter governing financial frictions (ϕ). In particular, we assume that countries are endowed with the same entrepreneurial talent distribution. We maintain this assumption because our goal is to isolate and quantify the direct impact of financial frictions. One of our main results is that, starting with the same potential pool of entrepreneurs, financial frictions distort the selection into active entrepreneurship. The productivity distribution of entrepreneurs in operation therefore differs across countries, with financial frictions lowering the mean and raising the dispersion of this distribution. The effect on the mean conforms to the conventional wisdom of aggregate TFP differences across countries. The increase in dispersion is consistent with the empirical findings of Hsieh and Klenow (2009), who show that less developed countries' establishment-level productivity (TFPQ in their terminology) dispersion is larger than that of the US. It would be straightforward to incorporate cross-country differences in the average productivity of potential entrepreneurs and workers by considering human capital and exogenous TFP differences. It is less obvious how one would discipline exogenous cross-country differences in higher moments of the entrepreneurial talent distribution.

4.1 Calibration

We calibrate preference and technology parameters so that the perfect-credit economy matches key aspects of the US, a relatively undistorted economy. In particular, our target moments include standard macroeconomic aggregates, features of the establishment size distribution within and across sectors, establishment dynamics, and the concentration of income in the population.

We need to specify values for eleven parameters: four technological parameters, α, θ, κ_S, κ_M, and the depreciation rate δ; two parameters describing the process for entrepreneurial talent, γ and η; the subjective discount factor β, the coefficient of relative risk aversion σ, the intratemporal elasticity of substitution ε, and the service share in consumption ψ.

Two preference parameters, σ and ε, and two technological parameters, $\frac{\alpha}{1/\eta+\alpha+\theta}$ and δ, can be set to standard values in the literature. We let $\sigma = 1.5$ and $\varepsilon = 1.0$.[25] The one-year depreciation rate is set at $\delta = 0.06$, and we choose $\frac{\alpha}{1/\eta+\alpha+\theta}$ to match the aggregate capital income share of 0.30.[26]

Target Moments	US Data	Model	Parameter
Top 10% employment share	0.69	0.69	$\eta = 4.84$
Top 5% income share	0.30	0.30	$\alpha + \theta = 0.79$
Average scale in services	17	17	$\kappa_S = 0.00$
Average scale in manufacturing	47	47	$\kappa_M = 4.68$
Exit rate	0.10	0.10	$\gamma = 0.89$
Manufacturing Share of GDP	0.25	0.25	$\psi = 0.91$
Interest rate	0.04	0.04	$\beta = 0.92$

Table 2: Calibration

We are thus left with the seven parameters that are more specific to our study.[27] We

[25] The estimate of Herrendorf et al. (2009) is close to one using data on final consumption of manufacturing and services. For value-added, Buera and Kaboski (2009) estimate the preference to be close to Leontieff: our choice of $\varepsilon = 1.0$ (Cobb-Douglas) will thus understate the quantitative effects of financial frictions.

[26] We are being conservative in choosing a relatively low capital share: the larger the share of capital, the bigger the role of capital misallocation. This way, we are also accommodating the fact that some of the measured payments to capital are actually payments to entrepreneurial input.

[27] As is common in heterogeneous-agent models with incomplete markets, the discount factor β must be

calibrate them to match seven relevant moments in the US data as shown in Table 2: the average size of establishments in services and in manufacturing; the employment share of the top decile of establishments; the share of income generated by the top five percentile of earners; the annual exit rate of establishments; the share of manufacturing value-added in the absorbed GDP; and the annual real interest rate.

The identification of these seven parameters follows the basic logic given in our discussion of the perfect-credit benchmark. We calibrate the sector-specific fixed costs, $\kappa_S = 0.0$ and $\kappa_M = 4.68$, to match the average establishment size in services and manufacturing (14 and 47, respectively). The per-period fixed cost in manufacturing sector, $\kappa_M = 4.68$, is tantamount to about three times the equilibrium wage in the perfect-credit benchmark. Given the returns to scale, $\alpha + \theta$, we choose the tail parameter of the entrepreneurial talent distribution, $\eta = 4.84$, to match the employment share of the largest ten percent of establishments, 0.63. We can then infer $\alpha + \theta = 0.79$ from the income share of the top five percent of earners. Top earners are mostly entrepreneurs (both in the US data and in the model), and $\alpha + \theta$ controls the fraction of output going to the entrepreneurial input. The parameter $\gamma = 0.89$ leads to an annual establishment exit rate of ten per cent in the model. This is consistent with the exit rate of establishments reported in the US Census.[28] We set $\psi = 0.91$ to match the share of manufacturing value added in absorbed GDP. Note that all investment goods are manufactured goods in our model. Finally, the model requires a discount factor of $\beta = 0.92$ to match the annual interest rate of four per cent.

Figure 2 shows the establishment size distribution from the calibrated perfect-credit benchmark (solid line for manufacturing, dashed line for services, dotted line for overall), and compares it with the US data (asterisk for manufacturing, triangle for services). The model is able to fit the tails of the empirical distribution, the distance between the two within-sector distributions, and the initial concavity in the overall (inclusive of both sectors) distribution of establishment size. The assumption that the entrepreneurial talents for the two sectors are drawn from the same Pareto distribution generates the identical slope for the right tails. The model cannot capture the initial concavity in the distribution of establishment size within a sector, however, presumably because we are abstracting from within-sector heterogeneity in fixed costs.

jointly calibrated with the parameters governing the stochastic income process.

[28]Note that $1 - \gamma$ is larger than ten per cent, because a fraction of those hit by this idea shock will choose to remain in business. Entrepreneurs exit only if their new idea is below the equilibrium cutoff level in either sector.

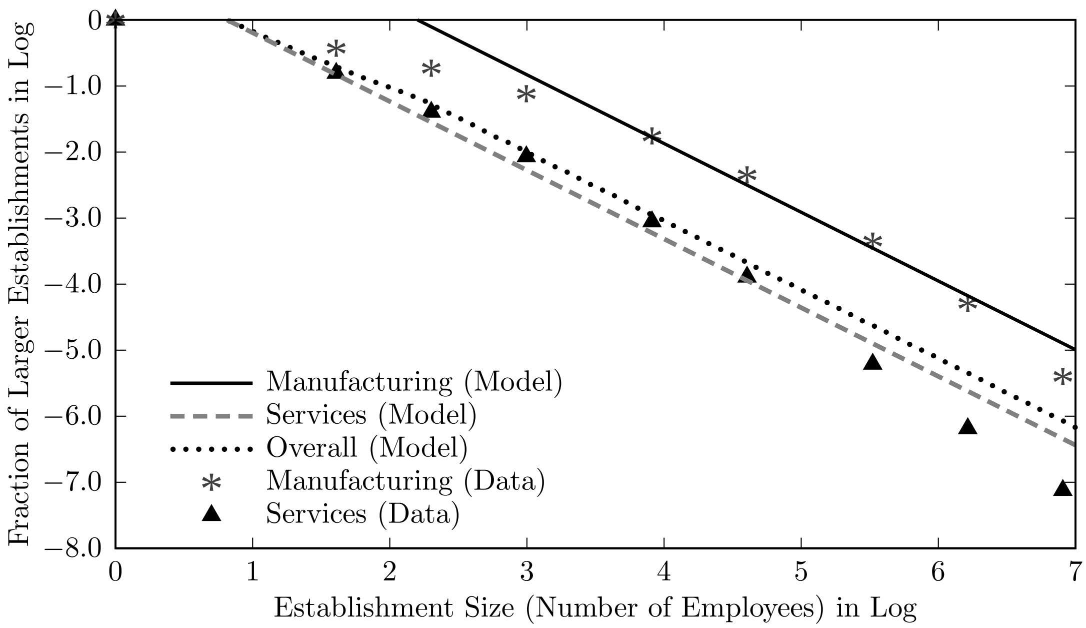

Fig. 2: Establishment size distribution in the data and in the model. The horizontal axis is the establishment size (number of employees, l) in log. For each l, we compute the fraction of establishments whose size is larger than or equal to l. This fraction in log is on the vertical axis. As we assume a Pareto distribution for entrepreneurial talent in each sector, the perfect-credit benchmark gives straight lines for services (dashed line) and manufacturing (solid line). One can construct a line using all the establishments in our perfect-credit benchmark (dotted line). We carry out the same calculation using the 2002 US Economic Census data. The establishment size distribution for the manufacturing sector is traced with asterisks, and the service sector with triangles.

4.2 Results

In this section, we quantify the effect of financial frictions on economic development. We first show that financial frictions have a substantial, adverse impact on output per worker. In our exercises, the lower per-capita income in economies with financial frictions is primarily explained by their low aggregate TFP, with particularly low productivity in manufacturing. These results are consistent with the empirical findings in Section 2.

We vary ϕ—the parameter governing the enforcement of contracts in condition (2)—to span a range of external finance to GDP ratios observed in the data. With quintiles of countries constructed in terms of GDP per worker at PPP, external finance to GDP ratios average 0.1 for the bottom quintile and 2.1 for the top quintile. The equilibrium external finance to GDP ratio is monotonically decreasing in ϕ in our exercises. We use 13 values of ϕ ranging from 0 to 1, which span variations in external finance to GDP ratios from 0 to 2.3. The parameter ϕ itself has no immediate empirical counterpart. Hence we plot the outcome of our model simulations against the (endogenous) ratio of external finance to GDP implied by a given ϕ.

The model quantities (e.g. output and TFP) are computed using a common fixed price, unless noted otherwise. In particular, we apply the output prices in the perfect-credit benchmark, with manufactured goods as numeraire. Similarly, the quantities in the data are measured at PPP.

4.2.1 Aggregate Impact of Financial Frictions

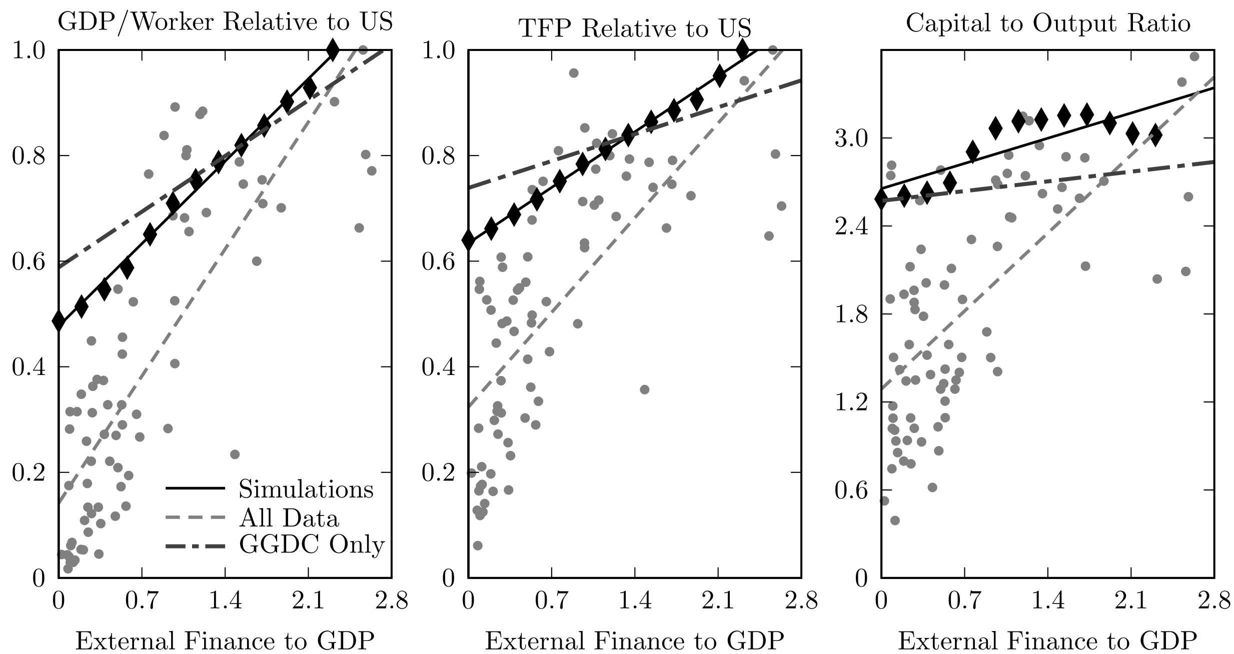

Fig. 3: GDP per worker at PPP (left panel), aggregate TFP (center panel), and capital to output ratio at PPP (right panel) against the ratio of external finance to GDP. For GDP per worker and TFP, the country data are relative to the US and model simulations are relative to the perfect-credit benchmark. The solid lines are regression lines for model simulations (♦), and the dashed lines are for all countries (•). For a subset of OECD countries, data on non-agricultural output, TFP and capital stock are available. These individual data points are not shown, but the corresponding regression lines are plotted with dash-dot lines.

Figure 3 plots the effect of financial frictions on output per worker at PPP, aggregate TFP, and capital to output ratios. The diamonds are from model simulations, and the dots represent country-level data.[29] The solid lines are regression lines for the model simulations, and the dashed lines are for the country data.

In our model, the variation in financial frictions can bring down output per worker to less than half of the US level (diamond, left panel). This roughly accounts for the difference between a country like Malaysia and the US, or about 80 per cent of the US-Mexico difference.

[29]The data are from the PWT 6.1. We use a perpetual inventory method and a depreciation rate of six per cent to construct the capital stock, using only the 79 countries with investment data series starting in 1980 or earlier. TFP in 1996 is constructed as $Y/(K^{1/3}L^{2/3})$, where Y is PPP GDP and L is the number of workers.

While this does not come close to the difference between the US and the poorest countries in sub-Saharan Africa, the magnitude is nevertheless sizable, considering that we are varying one single factor—financial markets—across countries. The regression coefficient of output per worker on external finance to GDP ratios is 0.22 for model simulations (solid line, left panel), and 0.34 for the data (dashed line, left panel).[30] Comparing the two coefficients, one may conclude that our model explains almost two thirds of the cross-country relationship between output per worker and financial development.

As in the data, the per-capita income differences in our model are primarily accounted for by differences in TFP (Figure 3, center panel). Financial frictions can reduce aggregate TFP by 40 per cent in our model (diamond, center panel). The regression coefficient of aggregate TFP on external finance to GDP ratios is 0.15 for model simulations (solid line, center panel), and 0.26 for the data (dashed line, center panel).[31] Comparing the two coefficients, one can conclude that our model explains about 60 per cent of the cross-country relationship between aggregate TFP and financial development.

Another effect of financial frictions is the impediment of capital accumulation. The right panel of Figure 3 shows the relationship between capital to output ratio and external finance to GDP. In our model, the capital to output ratios fall by 15 per cent as we move from the perfect-credit benchmark to financial autarky (diamond), when measured at common fixed prices across economies. The regression coefficients of capital to output ratios on external finance to GDP are 0.25 for model simulations (solid line) and 0.76 for the data measured at international prices (dashed line). As we discuss below (Sections 4.2.2 and 4.2.3), financial frictions affect the manufacturing sector and the service sector differentially, with the result that manufactured investment goods become relatively more expensive than services. Consistent with the data (Hsieh and Klenow, 2007), this higher relative price of investment explains almost all of the fall in capital to output ratios: the saving and investment rates measured at respective equilibrium prices are roughly constant across our model economies with varying degrees of financial development. This is because of two opposing forces at work. Firstly, financial frictions lead to lower equilibrium interest rates and hence returns to saving, exerting downward pressure on aggregate saving. On the other hand, the resulting lower wages and capital rental rates increase the returns to self-financing for constrained entrepreneurs. These forces seem to offset, leaving the saving and investment rates constant when measured at the respective equilibrium prices.

We also note that the factor-of-two difference in output per worker generated by our quantitative theory goes a long way if one were to focus on the factor-of-five difference in

[30] The univariate regression of output per worker on external finance in the data has an R^2 of 0.64.

[31] The univariate regression of aggregate TFP on external finance in the data has an R^2 of 0.51.

non-agricultural output per worker between the richest fifth percentile and the poorest fifth percentile of countries.[32] As we do not model the agricultural sector, ideally we would compare our model implications to data on non-agricultural output and productivity. The difficulty here is the dearth of data on sector-level capital stock that are comparable across countries.

We address this issue in two ways. Firstly, we use the data set of Restuccia et al. (2008) covering 82 countries to run a univariate regression of non-agricultural output per worker on external finance to GDP ratios. The regression coefficient is 0.24, and the R^2 is 0.45. This coefficient is about 30 per cent smaller than the one using the total output data (dashed line, left panel), and is very close to the slope, 0.22, generated by our model simulations (solid line, left panel): our model almost fully accounts for the empirical relationship between non-agricultural output per worker and financial development across countries. Secondly, because Restuccia et al. do not have data on capital input in the agricultural sector, we again turn to the Groningen Productivity Level Database and analyze the relationship between financial development and the performance of non-agricultural sectors.[33] In each of the three panels, we use a dash-dot line to show the regression line for these 18 OECD countries. We do not show individual data points. For these countries, the regression coefficients of non-agricultural output per worker, productivity, and capital to output ratios on financial development are smaller than those for the overall sample of countries. In fact, for this subset of countries, our model explains almost all of the empirical relationship between the dependent variables and external finance to GDP. The regression coefficients on external finance to GDP are 0.15, 0.07, and 0.09, respectively for output per worker, TFP, and capital to output ratios (dash-dot lines). Recall that the corresponding regression coefficients are 0.22, 0.15, and 0.25 with our model simulations (solid lines).

In summary, financial frictions in our model can explain a large part of the cross-country differences in economic development, measured by output per worker, aggregate TFP, and capital to output ratios. If one were to focus on the non-agricultural sector, available data even more strongly support the quantitative results of our model.

[32] See Restuccia et al. (2008) for a decomposition of cross-country per-capita income differences into agricultural and non-agricultural components.

[33] Among the 18 countries in the sample we use, Hungary is the poorest country, and also has the lowest external finance to GDP ratio (0.32). In the overall sample of the 79 countries analyzed here, Hungary ranks in the sixty-sixth percentile in terms of output per worker, and in the thirty-sixth percentile in terms of external finance to GDP.

4.2.2 Impact on Sectoral Productivity

Financial frictions have differential impacts on the two sectors. The solid lines in Figure 4 trace the effect of financial frictions on the measured TFP of the service sector (left panel) and the manufacturing sector (right panel). While TFP declines by 25 per cent in services, the manufacturing sector TFP declines by 55 per cent with financial frictions. This result is consistent with the empirical observations in Section 2 that productivity differences across countries are sharpest for the large-scale, manufacturing sector.

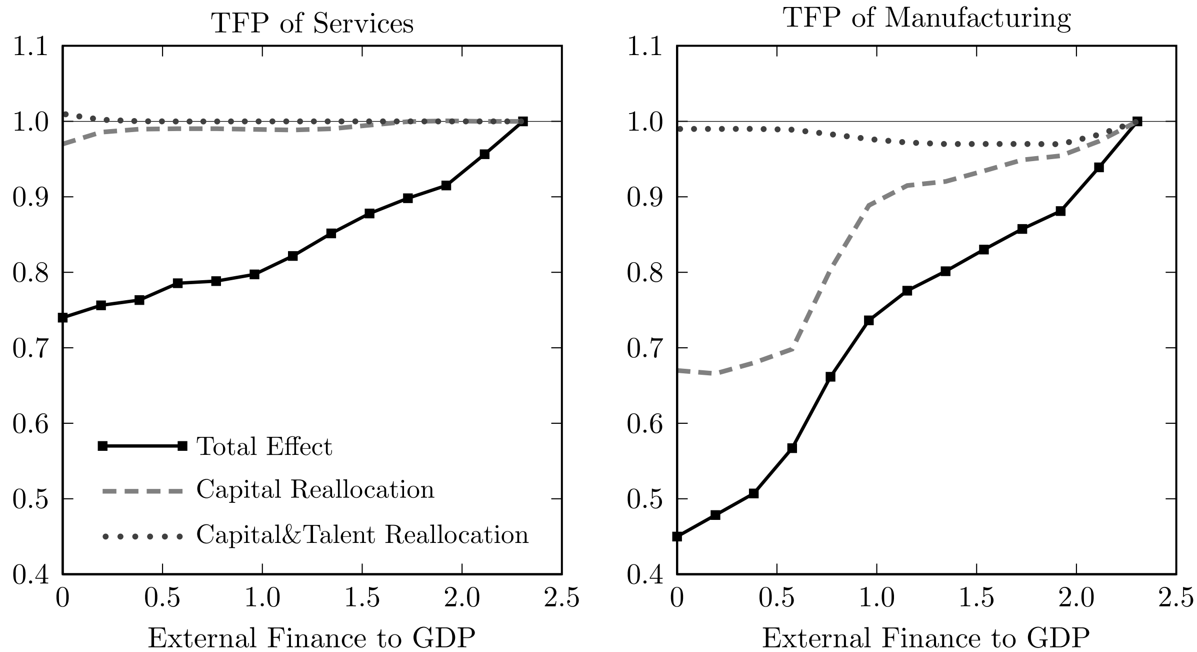

Fig. 4: Decomposing the impact of financial frictions on sector-level TFP in services (left panel) and in manufacturing (right panel). Solid lines are the sectoral TFP in the simulated economies. TFP is normalized by its level in the perfect-credit benchmark, and plotted against the corresponding external finance to GDP ratios. Efficient reallocation of capital among existing active entrepreneurs raises sectoral TFP (dashed lines). When distortions on selection into entrepreneurship are undone as well, sectoral TFP further increases (dotted lines).

Next, we ask what are the driving forces behind these effects on sector-level productivity (which in turn determine the aggregate productivity). Intuitively, financial frictions distort the allocation of productive capital among entrepreneurs in operation. Those with binding collateral constraints will have a marginal product of capital higher than the rental rate. In addition, financial frictions distort the entry and exit of entrepreneurs: Productive-but-poor entrepreneurs delay entry until they can overcome financing constraints, and incompetent-but-wealthy ones remain in business for too long. Here we quantitatively analyze how financial frictions affect sectoral and aggregate productivity by distorting the allocation of capital and the allocation of entrepreneurial talents.

In Figure 4, we decompose the effect of financial frictions on sectoral TFP into their effect on the allocation of capital across active entrepreneurs (intensive margin, or misallocation of capital), and their effect on the allocation of entrepreneurial talents (extensive margin). The extensive margin is further decomposed into the number of active entrepreneurs in each sector, and into the distribution of talent among active entrepreneurs (misallocation of talent). To quantify this decomposition, we perform three experiments on our simulated economies. First, we reallocate capital among active entrepreneurs within each sector to equalize the marginal product of capital across them. For this experiment, we hold fixed the number and the talent distribution of active entrepreneurs, as well as the total capital and labor employed in each sector. The sectoral TFP after this reallocation is the dashed lines in both panels. For the service sector, almost all of the lower TFP is explained by the misallocation of capital among active entrepreneurs. For the manufacturing sector, this intensive-margin distortion explains less than half of the lower TFP.

In the second experiment, while holding fixed the number of active entrepreneurs in each sector, we select the most talented individuals into entrepreneurship. We also allocate capital efficiently across the new set of active entrepreneurs, while holding the total capital and labor employed in each sector. The resulting sectoral TFP from this reallocation of talent and capital is the dotted lines. The misallocation of talent into entrepreneurship explains more than half of the lower TFP in the large-scale, manufacturing sector, and less than one tenth of the lower TFP in the small-scale, service sector.

Finally, in addition to the efficient reallocation of talent and capital above, we allow for the number of entrepreneurs to adjust in each sector at the perfect-credit equilibrium prices. This additional adjustment affects the sectoral TFP only slightly: The dotted lines from the second experiment above are already quite close to the horizontal lines going through one, which represent the sectoral TFP levels in the perfect-credit benchmark. This last experiment suggests that restrictions to entry per se may not have significant quantitative effects unless the distribution of entrants is distorted.

The following conclusions can be drawn from these sector-level exercises. Because of their larger scale and financing needs, establishments in the manufacturing sector are more susceptible to financial frictions. There are more misallocation of capital and more misallocation of entrepreneurial talent in manufacturing than in services. In particular, the distortions on the entry/exit decisions of entrepreneurs matter vastly more for manufacturing. This result suggests that modeling endogenous entry/exit of entrepreneurs is pivotal for capturing the large impact of financial frictions on productivity.

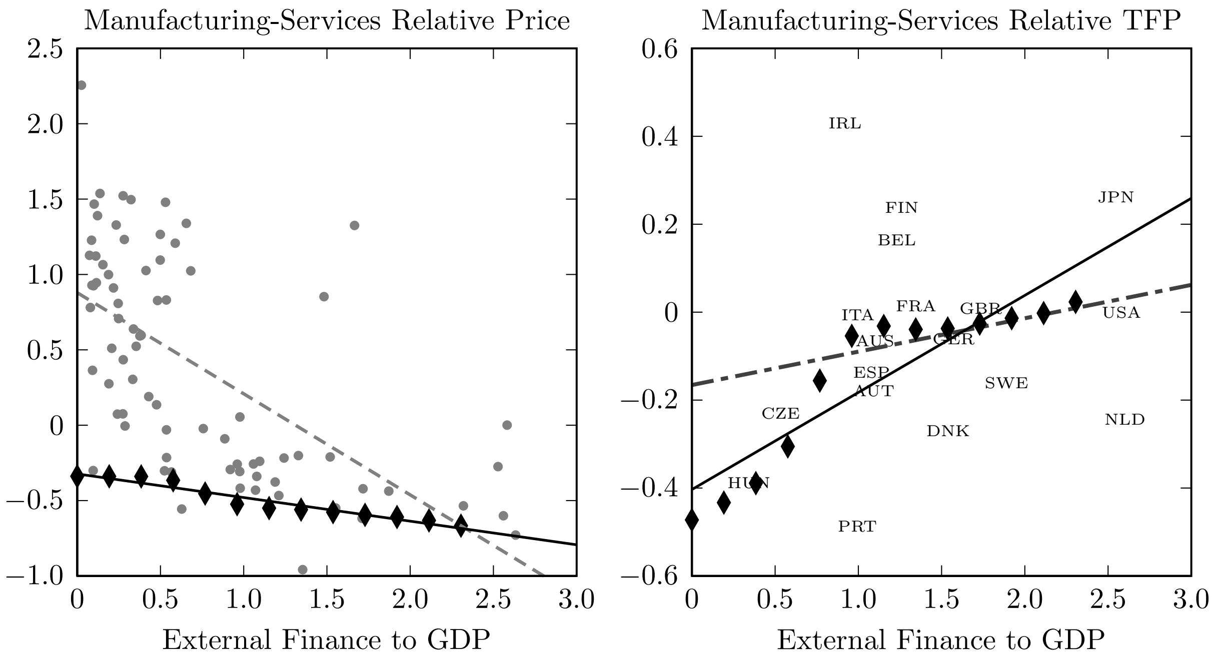

Fig. 5: Log relative price of manufacturing to services (left panel) and log relative productivity of manufacturing to services (right panel) against the ratio of external finance to GDP. Country data (• in the left panel and three-letter codes in the right panel) and model simulations (♦). The price data are from the 1996 ICP benchmark, and the sector-level TFP are from the GGDC. The dashed and dash-dot lines are regression lines for the data, and the solid ones are for the model simulation.

4.2.3 Relative Productivity and Relative Prices

The pattern of relative productivity between the two sectors leads to the price of manufactured goods relative to services being higher in countries with underdeveloped financial markets. The left panel of Figure 5 plots the relative price of manufactured goods to services in log ($\log(p_M/p_S)$) against external finance to GDP ratios. The diamonds represent the model simulations, and the dots represent country data. The regression coefficients are -0.67 for the data (dashed line) and -0.16 for model simulations (solid line): Our model accounts for about a quarter of the empirical cross-country relationship between the manufacturing-services relative price and financial development, leaving room for other explanations of relative prices that are correlated with financial frictions.

In the right panel of Figure 5, we plot the relative productivity of manufacturing to services in log against external finance to GDP ratios. The diamonds are the model simulations, with the corresponding regression line (solid line, right panel) that has a slope coefficient of 0.22. We again turn to the Groningen database for the sector-level relative productivity in the 18 OECD countries (three-letter codes, right panel). The regression line (dash-dot line, right panel) has a slope coefficient of 0.08: Our model generates a relationship between relative sector-level productivity and financial development that is starker than the empirical one found in the available data.

4.2.4 Impact on Distribution of Establishment-Level Productivity and Size

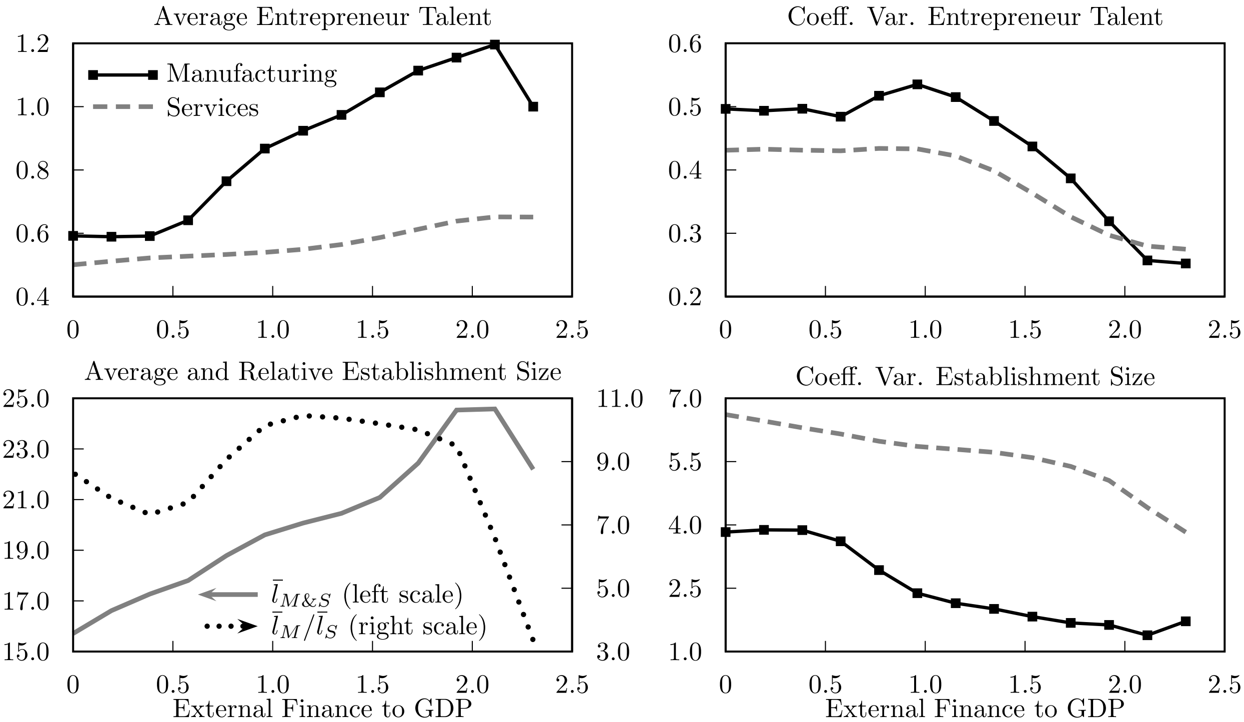

Fig. 6: Impact of financial frictions on the distributions of establishment-level productivity and size. Clockwise from top left, all plotted against the ratio of external finance to GDP (horizontal axis): Average productivity (z) of active entrepreneurs in the service sector (dashed line) and in the manufacturing sector (solid line), normalized by the average manufacturing entrepreneurial productivity in the perfect-credit benchmark; Coefficient of variation (coeff. var.) of the productivity of active entrepreneurs in each sector; Coefficient of variation of the establishment size (number of workers) in each sector; Average establishment size in the overall economy (solid gray line, left scale) and the ratio of average manufacturing establishment size to average service establishment size (dotted line, right scale).

In addition to their effects on sector-level productivity and relative prices, financial frictions have impact on the distribution of establishment-level productivity and size. Not surprisingly, the establishment-level distribution in each sector is differentially affected by financial frictions.

The top left panel of Figure 6 plots the average talent or productivity (z_j) of active entrepreneurs in the manufacturing sector (solid line) and the service sector (dashed line) of our simulated economies, against the ratio of external finance to GDP. The averages are normalized by the average manufacturing entrepreneurial productivity in the perfect-credit benchmark. With financial frictions, not only entrepreneurial talent but also an individual's wealth determine whether he will be an entrepreneur or not in any given period. As a result, incompetent-but-wealthy entrepreneurs remain in business, and talented-but-poor individuals do not operate business until they can self-finance the capital needed to operate at a profitable scale. With more financial frictions—and hence lower equilibrium ratio of external finance to GDP, the selection into entrepreneurship gets more distorted, and the average talent or productivity among active entrepreneurs falls. Figure 4 have shown that such misallocation of talent is more rampant in the manufacturing sector, which is confirmed here: The average entrepreneurial talent in the manufacturing sector drops by 40 per cent in response to financial frictions, while it goes down by 20 per cent in the small-scale, service sector.[34]

With more financial frictions, an individual's wealth becomes more influential upon their decision of entry into and exit from entrepreneurship. As a result, individuals with more diverse entrepreneurial talents will be operating business in the stationary equilibrium. With more financial frictions, the within-sector distribution of establishment-level productivity (z_j, $j = S, M$) becomes more disperse, and its coefficient of variation increases (top right panel, Figure 6.

Next, we explore how financial frictions affect the within-sector establishment size distribution. In the bottom left panel of Figure 6, we trace how the average establishment size changes in response to financial frictions. The solid line is the average establishment size, inclusive of both sectors (left-hand side vertical axis). With financial frictions, the average establishment size decreases by 30 per cent.[35] This is because of the low equilibrium wage: More individuals, including those who are not particularly talented, opt for entrepreneurship. In this sense, financial frictions generate too many establishments that are too small. The model prediction on the relationship between the average establishment size and financial development is consistent with available empirical evidence (Tybout, 2000). However, the magnitude of this effect in the model is not as large as in the available data, as relative prices leads the productive and wealthy entrepreneurs to run drastically larger establishments.

[34]Note that the average entrepreneurial talent in manufacturing is non-monotonic with respect to external finance to GDP ratios. With moderate degrees of financial frictions, where the external finance to GDP ratio runs from 2.3 down to 1.5, there seems to be a "positive" selection into entrepreneurship in the manufacturing sector. This can be explained with the fact that highly-productive entrepreneurs can overcome the collateral constraint more easily than low-productivity ones. Firstly, for a given level of output and factor prices, they make more profits. Secondly, our financing limit $\bar{k}(\cdot)$ is an increasing function of z. These forces seem to prevail with moderate degrees of financial frictions (high ϕ's), generating the hump in the average entrepreneurial talent in the manufacturing sector. We further note that, in the right panel of Figure 4, the misallocation of talent margin does not explain much of the lower sector-level TFP in manufacturing for moderate degrees of financial frictions, even less than the number of active entrepreneurs margin.

[35]The average establishment size first increases from 22.1 to 24.5, before it goes down to 15.4. The initial increase over the region where external finance to GDP is between 2.3 and 1.9 is the flip side of the fall in the number of establishments. It reflects the increased difficulty in starting a business with financial frictions, especially in the manufacturing sector.

Such marginal entrepreneurs in economies with financial frictions overwhelmingly choose to start businesses in the small-scale, service sector, where the establishment-level technology commands less financing needs. The larger scale and more financing needs in the manufacturing sector (with its larger fixed costs and operating scales) make it harder to set up an establishment there with financial frictions, reducing the number of manufacturing establishments. At the same time, the relative price of manufactured goods to services goes up with financial frictions. As a result, the service sector becomes even smaller relative to the manufacturing sector. The dotted line is the ratio of the average establishment size in manufacturing to that in services, $\bar{l}_M/\bar{l}_S$ (right-hand side vertical axis). We further pursue this prediction below (Section 4.2.5).

For given output and factor prices, the size of an establishment (measured by the number of employees) is determined by the entrepreneurial talent and the collateral constraint. Therefore, entrepreneurs with the same talent may operate at different scales, as they may have different levels of wealth (collateral). We have seen that financial frictions increase the dispersion of the entrepreneurial productivity, which is then compounded with the increased dispersion of scale for a given entrepreneurial talent. As a result, within-sector establishment size distributions become more disperse with financial frictions (bottom right panel, Figure 6). The overall (inclusive of both sectors) establishment size distribution also becomes more disperse with financial frictions, which is consistent with the "missing middle" in less developed economies (Tybout, 2000).

4.2.5 Relative Scale of Sectors: A Testable Implication

Our exercises confirm the intuition that financial frictions are more harmful to the sector with higher fixed costs, larger scales, and more financing needs (i.e. manufacturing). One interesting model implication is that financial frictions lead to greater disparity in the average establishment size between the manufacturing and the service sectors (dotted line, bottom left panel, Figure 6). As discussed above, this result stems not only from the direct effect of financial frictions on entrepreneurial entry decisions, but also from the general equilibrium effect on input and output prices. With financial frictions, wages—and hence the opportunity cost of entrepreneurship—go down, and entrepreneurship becomes more desirable, especially in the small-scale, service sector. On the other hand, reasonably talented entrepreneurs who have the means to meet the financing needs in the manufacturing sector face relatively higher output prices and lower input prices, and increase their scale of operation accordingly.

We evaluate this prediction using some new evidence from a case study of the US and Mexico. We compare establishment size data from the 2002 US Economic Census and the 2003 Mexican Economic Census, both of which follow the NAICS classification system.

For manufacturing, the classification schemes for Mexico and the US are identical, and allow for comparability at a very disaggregate level of industries (four digits). For services, however, the schemes differ and only allow for comparison at the two-digit industry level. The two censuses have more substantial differences in their coverage. We use Mexico's 1998 National Survey of Micro-Enterprises (ENAMIN) to impute corrections that make the US and the Mexican data fully comparable. These corrections include adjustments to remove non-employers—included in the Mexican Census but not for the US, and adjustments to add small-scale entrepreneurs without a fixed location—included in the US Census, though presumably unimportant, but not for Mexico, where they play an important role.

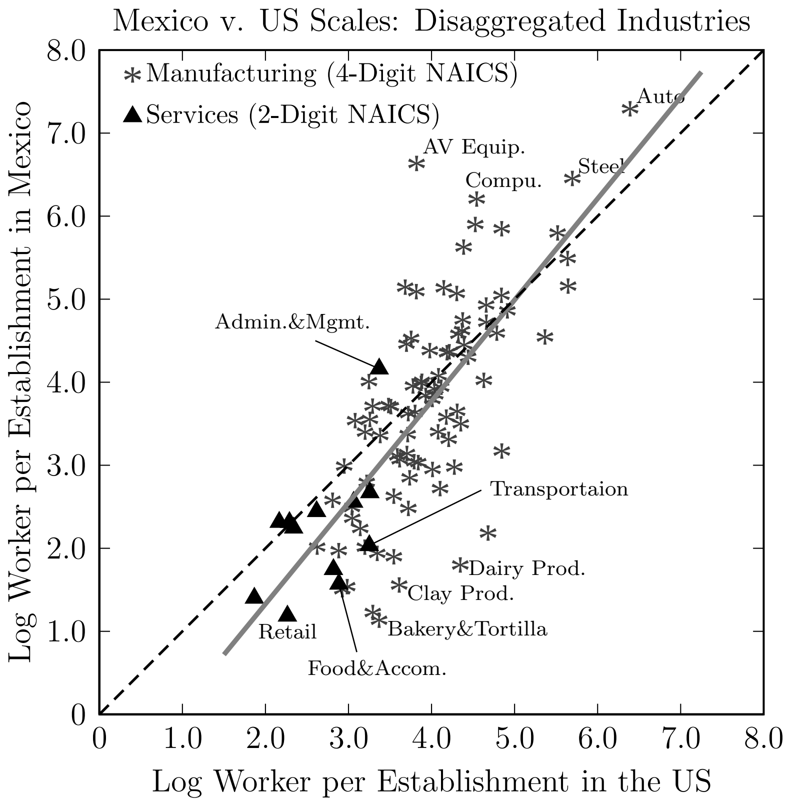

Fig. 7: Average establishment size of a given industry in Mexico (vertical axis) and in the US (horizontal axis): 86 industries in manufacturing and 12 in services. We use the data from the 2002 US Economic Census, the 2003 Mexican Economic Census, and Mexico's 1998 National Survey of Micro-Enterprises. The solid line is the regression line, and the dashed line is the 45-degree line.

Figure 7 plots the average establishment size in Mexico (in log, vertical axis) against the average establishment size in the US (in log, horizontal axis) for 86 four-digit manufacturing industries and 12 two-digit service industries. The overall average establishment size is substantially smaller in Mexico than in the US (6 vs. 17 workers per establishment). However, many industries (those lying above the 45-degree line) have an average establishment size that is larger in Mexico than in the US.[36] That is, the industries that are of a large scale in the US have an even larger scale in Mexico, while those that are of a small scale in the US have an even smaller scale in Mexico. With the exception of administration/management services, the ones above the 45-degree line are manufacturing industries. This finding supports our simulation results on the relative scale of sectors: The relative scale of manufacturing to services is larger with financial frictions.[37]

In addition, the fact that, at a more disaggregate industry level, Mexico has wider variation in scale than the US is consistent with our model prediction of a higher coefficient of variation of establishment size in financially less developed countries (bottom right panel, Figure 6). A more formal testing of this prediction, however, would require more detailed establishment-level data that are comparable across countries.

To summarize, one of our model predictions is that the the relative scale of manufacturing to services is even larger in economies with financial frictions. Using comparable establishment size data at a disaggregate industry level, we have shown that large-scale (mostly manufacturing) industries tend to have an even larger scale in Mexico than in the US, while small-scale industries such as services have an even smaller scale in Mexico than in the US. This finding, hitherto undocumented in the literature, supports our model implication on relative scale of sectors. Furthermore, we also conjecture that our mechanism, in which scale differences lead to differential impact of financial frictions across sectors, will also work at a more disaggregate level, not just at the level of the manufacturing vs. services dichotomy.

4.3 Discussions on Modeling Choices

4.3.1 Robustness

We briefly discuss the robustness of our results to alternative model specifications. We first consider two other ways of generating scale differences between the two sectors.

Setup Costs We have constructed numerical examples for a version of our model where the sectoral scale difference is driven by one-time setup costs, rather than by per-period fixed costs.[38] While these numerical examples were not precisely-calibrated exercises, we found

[36] Indeed, the data show a slope (solid line) that is clearly steeper than the 45-degree line (dashed line). The regression coefficient is 1.22 with a standard error of 0.11.

[37] There are other evidence consistent with our model prediction. Using the available OECD SSIS data, we find that financially less developed countries (Czech Republic, Hungary, Poland, Portugal, and Slovakia) have a larger relative scale of manufacturing to services (ranging from 3.4 to 4.1, with a median of 3.8) than the US (3.1), where the sector-level scale is defined to be the number of workers per enterprise. The OECD SSIS data are designed for comparability across countries, but suffer from coverage issues.

[38] In the recursive formulation of this specification, we need to carry an additional state variable indicating whether in the previous period an individual was a worker, an entrepreneur in services, or an entrepreneur in

that financial frictions have an even larger impact on aggregate and sectoral productivity, reflecting the stronger non-convexity that (front-loaded) setup costs impose. We chose our fixed-cost specification to avoid the risk of exaggerating the impact of financial frictions, especially when there are no reliable data on setup costs.[39]

Span of Control Differences between Sectors We have also considered the possibility that the sectoral scale difference hinges on the establishment-level span of control. In particular, we set α_j and θ_j, $j = S, M$, such that $\alpha_S + \theta_S < \alpha_M + \theta_M$, while the ratio α_j/θ_j is the same for $j = S, M$. The latter assumption reflects the empirical facts on factor shares reported in Table 1. In calibrated exercises with span-of-control differences and no fixed cost in either sector, we find that the effects of financial frictions are broadly consistent with our findings in Section 4.2, although not as large in magnitude. One important distinction is the implication on how relative scale of the two sectors change with financial frictions. With span-of-control differences, the relative scale of manufacturing to services decreases with financial frictions. Recall that our model with fixed cost differences predicts the opposite, consistent with our empirical findings in Section 4.2.5. It is mainly for this reason that we chose the fixed-cost model for our analysis.

Simpler Collateral Constraint We have also tried an alternative specification for the collateral constraint: a capital rental limit that does not depend on the entrepreneurs' talent z. More specifically, the sum of the capital rental and the fixed cost could not exceed λa, where λ is a parameter governing the degree of financial frictions, and a is the financial wealth of an entrepreneur at the beginning of a period. With this specification, $\lambda = 1$ corresponds to financial autarky where all capital needs to be self-financed, and $\lambda \to +\infty$ corresponds to the perfect-credit benchmark. In calibrated exercises with this simpler collateral constraint, the results are similar to our results in Section 4.2, both qualitatively and quantitatively. One nice feature of this specification is that the average establishment size in the overall economy (inclusive of both sectors) monotonically decreases with financial frictions, unlike the hump in the bottom left panel of Figure 6. We nevertheless decided to go with a more general

manufacturing. More important, we allow for between-periods borrowing, i.e., $\underline{a} < 0$. We assume that debt contracts are contingent on idea shocks so that the debt of those entrepreneurs hit with the idea shock (with probability $1 - \gamma$) at the end of a period is forgiven. The lower bound on asset $\underline{a}$ is defined to be the most generous debt limit that is enforceable when the borrower is not hit with the idea shock. The zero-profit condition of the financial intermediary will determine the spread between the deposit and the borrowing rates.

[39]See, for example, Paulson and Townsend (2004) and McKenzie and Woodruff (2006) for discussions on setup and fixed costs for small establishments in developing countries. Sutton (1991) explains the difficulties of estimating setup costs using data from oligopolistic industries in developed countries.

specification where collateral constraints depend endogenously on entrepreneurs' wealth and talent.

4.3.2 Comparison with One-Sector Models

One obvious advantage of our two-sector model over one-sector models is its richer implications on relative prices, productivity, and scale between the two sectors. At the same time, our two-sector model generates a larger impact of financial frictions on aggregate output and productivity than a one-sector model, when both models follow the same calibration strategy. In particular, we transformed the two-sector model into a one-sector model by imposing $\kappa_S = \kappa_M$ and $p_S = p_M$. We then follow our calibration strategy is Section 4.1. For comparability, we choose the fixed cost κ in the one-sector model in a way that the ratios of total fixed costs to GDP are the same in the perfect-credit benchmarks of the two models. Recall from Section 4.2 that financial frictions can bring down aggregate output by 52 per cent, aggregate TFP by 36 per cent, and the capital to output ratio by 15 per cent (Figure 3). In the calibrated one-sector model, the effect of financial frictions is 39 per cent on output, 30 per cent on TFP, and virtually none on the capital to output ratio.[40]

In our calibration of the two-sector model, we have larger fixed costs for manufacturing establishments than for service establishments, and these fixed costs are disciplined by the scale differences between the two sectors.[41]

We find that the effect of the non-convexities (i.e. fixed costs) is itself convex: The aggregate impact of the large fixed costs borne by one part of the economy (manufacturing) is larger than that of the small fixed costs spread over the whole economy, holding constant the ratio of total fixed costs to GDP: The impact of financial frictions is larger in our two-sector framework, and there is a qualitative difference in the predictions on capital to output ratios between the one-sector and the two-sector models.

4.3.3 Comparison with Two-Period Models: The Issue of Time Aggregation

Earlier attempts at quantifying the effect of financial frictions often relied on two-period overlapping-generations models. Given their tractability, two-period models are useful for qualitative characterizations of the economic mechanisms at work. However, there are difficulties in mapping the model into data and using it for quantitative purposes.[42] In particular,

[40]In an alternative one-sector exercise, we assumed that there is no fixed cost, while still following the same calibration strategy. The effect of financial frictions are marginally smaller without fixed costs: The effect is 35 per cent on output, 26 per cent on TFP, and again none on the capital to output ratio.

[41]We also discipline the relative importance of manufacturing vs. services in terms of GDP in our economies using data on (value-added) GDP.

[42]Amaral and Quintin (2009) provide a useful snapshot of the possible difficulties of quantitative work using two-period models, as well as some insights into such challenges can be met.

two-period models nearly preclude one important way of coping with financial frictions: self-financing. In our model with a realistic time horizon, a talented-but-poor individual saves up so that he can overcome the financial constraints over time and operate at the maximal-profit scale. In a two-period model, on the other hand, such a would-be entrepreneur is by construction condemned to the binding collateral constraint for half his lifetime. There is no real opportunity to overcome the constraint, and a larger fraction of individuals are constrained in the stationary equilibrium. As a result, the effect of financial frictions in a two-period model is made artificially larger than in a model where individuals are allowed to engage in self-financing. For example, in our own version of a two-period overlapping generation model, the aggregate effect of financial frictions (going from perfect credit to financial autraky) is about 50 per cent larger than in our baseline model with more realistic time horizon, when both models follow the same calibration strategy. Also, the lack of meaningful self-financing implies little growth in establishment size, and the average establishment size in the stationary equilibrium comes out too small, compared to the available data.

5 Concluding Remarks

We have developed a quantitative theory linking countries' level of financial development to their output per worker, aggregate TFP, sector-level relative productivity, and relative prices. Financial frictions distort the allocation of capital and entrepreneurial talents, and have sizable effects on a country's output per worker and aggregate productivity. Indeed, our model can explain almost all of the cross-country relationship between aggregate output or TFP and financial development, if one focuses on non-agricultural sectors. Establishments in sectors characterized by large fixed costs operate most efficiently at larger scales, and hence have bigger financing needs. For this reason, they are more vulnerable to financial frictions than those in small-scale sectors. We have shown that this mechanism is quantitatively important, almost fully explaining the relationship between financial development and relative productivity of sectors in the available data. We have also shown that our mechanism is also consistent with the larger differences in average establishment size across sectors observed in financially less developed countries.

Our analysis shows how micro-level (firm or establishment) technological differences across sectors help us better understand macroeconomic issues such as aggregate productivity and relative productivity across sectors. In this context, we view the study of other micro-level distortions—e.g. size-dependent policies of Guner et al. (2008) or entry barriers of Djankov et al. (2002)—and their interaction with financial frictions as promising avenues for future research.

Appendix

Proof of Proposition 1 The rental of capital k in sector j is enforceable iff:

$$\tilde{v}^j(k;a,z) \geq v^d\left(a^d,z\right),$$

where $\tilde{v}^j(k;a,z)$ is the value of a non-defaulting entrepreneur with wealth a and ability z that operates in sector j with rented capital k:

$$\tilde{v}^j(k;a,z) = \max_{c,a'}\{u(c)+\beta[\gamma v(a',z)+(1-\gamma)\mathbb{E}_{z'}v(a',z')]\}$$
$$\text{s.t. } pc+a' \leq \max_l\{p_j z_j f(k,l)-wl\}-Rk-(1+r)p_j\kappa_j+(1+r)a.$$

We define $v^d\left(a^d,z\right)$ to be the value of a defaulting entrepreneur with ability z who gets to keep $a^d=(1-\phi)[\max_l\{p_j z_j f(k,l)-wl\}+(1-\delta)k]$:

$$v^d\left(a^d,z\right) = \max_{c,a'}\{u(c)+\beta[\gamma v(a',z)+(1-\gamma)\mathbb{E}_{z'}v(a',z')]\}$$
$$\text{s.t. } pc+a' \leq (1-\phi)\left[\max_l\{p_j z_j f(k,l)-wl\}+(1-\delta)k\right].$$

It is straightforward to see that $\tilde{v}^j(k;a,z) \geq v^d\left(a^d,z\right)$ iff:

$$\max_l\{p_j z_j f(k,l)-wl\}-Rk-(1+r)p_j\kappa_j+(1+r)a$$
$$\geq (1-\phi)\left[\max_l\{p_j z_j f(k,l)-wl\}+(1-\delta)k\right],$$

which is equivalent to:

$$(1+r)(a-p_j\kappa_j) \geq ((1-\delta)(1-\phi)+r+\delta)k-\phi\max_l\{p_j z_j f(k,l)-wl\}$$
$$= -\phi\left[\max_l\{p_j z_j f(k,l)-wl\}-\frac{((1-\delta)(1-\phi)+r+\delta)}{\phi}k\right].$$

Notice that as long as $f_{kk}(k,l)<0$, $f_{ll}(k,l)<0$, $\lim_{k\to 0} f_k(k,l)=\infty$ and $\lim_{k\to\infty} f_k(k,l)=0$, the right hand side is minimized for some $\hat{k}(z_j;\phi)$ such that $0<\hat{k}(z_j;\phi)<k_j^u(z_j)$, as it equals $-\phi$ times the profits of an entrepreneur facing a rental price $((1-\delta)(1-\phi)+r+\delta)/\phi > r+\delta$. The set of enforceable levels of capital rental can be easily characterized by a simple set of rental limits. There are two cases to consider.

If $(1+r)(a-p_j\kappa_j) > -\phi\left[\max_l\left\{p_j z_j f\left(\hat{k}(z_j;\phi),l\right)-wl\right\}-\frac{((1-\delta)(1-\phi)+r+\delta)}{\phi}\hat{k}(z_j;\phi)\right]$, there exists a unique function $\bar{k}^j(a,z_j;\phi) \geq \hat{k}(z_j;\phi)$ given by the largest root of the equation:

$$(1+r)(a-p_j\kappa_j) = (1-\phi+r+\phi\delta)\bar{k}^j(a,z_j;\phi)-\phi\max_l\left\{p_j z_j f\left(\bar{k}^j(a,z_j;\phi),l\right)-wl\right\}.$$

Notice that if $a - p_j\kappa_j < 0$, there are two positive roots of the equation, with the smaller root satisfying $\underline{k}^j(a, z_j; \phi) \leq \hat{k}(z_j; \phi)$. In this case, the set of enforceable levels of capital rental is $\left[\underline{k}^j(a, z_j; \phi), \bar{k}^j(a, z_j; \phi)\right]$. If $a - p_j\kappa_j \geq 0$, the set of enforceable levels of capital rental is simply $\left[0, \bar{k}^j(a, z_j; \phi)\right]$. It is straightforward to see that $\bar{k}^j(a, z_j; \phi)$ is strictly increasing in a, z_j, and ϕ.

If $(1+r)(a - p_j\kappa_j) \geq -\phi\left[\max_l\left\{p_j z_j f\left(\hat{k}(z_j;\phi), l\right) - wl\right\} - \frac{((1-\delta)(1-\phi)+r+\delta)}{\phi}\hat{k}(z_j;\phi)\right]$, we set $\bar{k}^j(a, z_j; \phi) = 0$. $\blacksquare$

Proof of Proposition 2 In an economy with perfect credit markets, selection of individuals into entrepreneurship and sectors is determined by their entrepreneurial talents and relative prices. In particular, there exist two threshold ideas $\hat{z}_j$, $j = S, M$, and a function $\hat{z}_j(z_{-j})$, $(j, -j) = (S, M), (M, S)$, dividing the space of entrepreneurial ideas (z_S, z_M) into workers and entrepreneurs in the S and M sectors. These thresholds are defined by the following three indifference conditions:

$$(p_j\hat{z}_j)^{\frac{1}{1-\alpha-\theta}}\left(\frac{\alpha}{R}\right)^{\frac{\alpha}{1-\alpha-\theta}}\left(\frac{\theta}{w}\right)^{\frac{\theta}{1-\alpha-\theta}}(1-\alpha-\theta) = w + p_j\kappa_j(1+r), \quad s = S, M, \tag{6}$$

$$\begin{aligned}(p_j\hat{z}_j(z_{-j}))^{\frac{1}{1-\alpha-\theta}}\left(\frac{\alpha}{R}\right)^{\frac{\alpha}{1-\alpha-\theta}}\left(\frac{\theta}{w}\right)^{\frac{\theta}{1-\alpha-\theta}}&(1-\alpha-\theta) - p_j\kappa_j(1+r)\\ = (p_{-j}z_{-j})^{\frac{1}{1-\alpha-\theta}}\left(\frac{\alpha}{R}\right)^{\frac{\alpha}{1-\alpha-\theta}}\left(\frac{\theta}{w}\right)^{\frac{\theta}{1-\alpha-\theta}}&(1-\alpha-\theta) - p_{-j}\kappa_{-j}(1+r).\end{aligned} \tag{7}$$

Integrating over individual output of entrepreneurs in sector j net of fixed costs, we obtain an expression for the net output of sector j,

$$Y_j = N\int_{\hat{z}_j}^{\infty}\int_1^{\hat{z}_{-j}(z_j)} z_j k(z)^{\alpha} l(z)^{\theta}\mu(dz) - \kappa_j(1+r)N\int_{\hat{z}_j}^{\infty}\int_1^{\hat{z}_{-j}(z_j)}\mu(dz). \tag{8}$$

Using $k(z) = \frac{z^{\frac{1}{1-\alpha-\theta}}}{Z_j^{\frac{1}{1-\alpha-\theta}}}\frac{K_j}{N}$ and $l(z) = \frac{z^{\frac{1}{1-\alpha-\theta}}}{Z_j^{\frac{1}{1-\alpha-\theta}}}\frac{L_j}{N}$, which follow from the first order conditions of the entrepreneurs' problem, we can rewrite (8) as:

$$\begin{aligned}Y_j &= N\int_{\hat{z}_j}^{\infty}\int_1^{\hat{z}_{-j}(z_j)} z_j\left[\frac{z_j^{\frac{1}{1-\alpha-\theta}}}{Z_j^{\frac{1}{1-\alpha-\theta}}}\frac{K_j}{N_j}\right]^{\alpha}\left[\frac{z_j^{\frac{1}{1-\alpha-\theta}}}{Z_j^{\frac{1}{1-\alpha-\theta}}}\frac{L_j}{N_j}\right]^{\theta}\mu(dz) - \kappa_j(1+r)N\int_{\hat{z}_j}^{\infty}\int_1^{\hat{z}_{-j}(z_j)}\mu(dz)\\ &= \frac{N^{1-\alpha-\theta}K_j^{\alpha}L_j^{\theta}}{Z_j^{\frac{\alpha+\theta}{1-\alpha-\theta}}}\int_{\hat{z}_j}^{\infty}\int_1^{\hat{z}_{-j}(z_j)} z_j^{\frac{1}{1-\alpha-\theta}}\mu(dz) - \kappa_j(1+r)N\int_{\hat{z}_j}^{\infty}\int_1^{\hat{z}_{-j}(z_j)}\mu(dz).\end{aligned}$$

Then,

$$Y_j = N^{1-\alpha-\theta}K_j^{\alpha}L_j^{\theta}Z_j - \kappa_j(1+r)N\int_{\hat{z}_j}^{\infty}\int_1^{\hat{z}_{-j}(z_j)}\mu(dz), \tag{9}$$

where $Z_j = \left[\int_{\hat{z}_j}^{\infty}\int_1^{\hat{z}_{-j}(z_j)} z_j^{\frac{1}{1-\alpha-\theta}}\mu\left(dz\right)\right]^{1-\alpha-\theta}$.

Assuming $\mu\left(dz\right) = \eta^2\left(z_S z_L\right)^{-(\eta+1)}$ for $z_j \geq 1$ and that entrepreneurs are a small fractions of the population—i.e. $\hat{z}_j$ is large for $j = S, M$, we obtain

$$Z_j = \left[\int_{\hat{z}_j}^{\infty}\int_1^{\hat{z}_{-s}(z_j)} z_j^{\frac{1}{1-\alpha-\theta}}\mu\left(dz\right)\right]^{1-\alpha-\theta} \approx \left[\int_{\hat{z}_j}^{\infty} z_j^{\frac{1}{1-\alpha-\theta}}\eta z_j^{-(\eta+1)}dz_j\right]^{1-\alpha-\theta},$$

$$\int_{\hat{z}_j}^{\infty}\int_1^{\hat{z}_{-j}(z_j)}\mu\left(dz\right) \approx \hat{z}_j^{-\eta}.$$

Here, we further assume $\frac{1}{\eta} < (1-\alpha-\theta)$ to guarantee that the integral is finite.

Using (6), as well as $\frac{\alpha}{R} = \frac{K_j^{1-\alpha}}{L_j^{\theta}N^{1-\alpha-\theta}}\frac{1}{p_j Z_j}$ and $\frac{\theta}{w} = \frac{L_j^{1-\theta}}{K_j^{\alpha}N^{1-\alpha-\theta}}\frac{1}{p_j Z_j}$, we obtain

$$\hat{z}_j = \left\{\frac{\left(\kappa_j(1+r)+\frac{w}{p_j}\right)\left[\frac{(1-\alpha-\theta)}{\eta(1-\alpha-\theta)-1}\eta\right]^{\alpha+\theta}}{\left(1-\alpha-\theta\right)N^{-\alpha-\theta}K_j^{\alpha}L_j^{\theta}}\right\}^{\frac{1}{1+\eta(\alpha+\theta)}}. \tag{10}$$

Substituting into (9),

$$Y_j = A_j N^{\frac{1}{1+\eta(\alpha+\theta)}} K_j^{\frac{\alpha\eta}{1+\eta(\alpha+\theta)}} L_j^{\frac{\theta\eta}{1+\eta(\alpha+\theta)}},$$

where

$$A_j = \frac{\left[\frac{(1-\alpha-\theta)}{\eta(1-\alpha-\theta)-1}\eta\right]^{\frac{1}{1+\eta(\alpha+\theta)}}\left[1-\frac{p_j\kappa_j}{p_j\kappa_j+w}\left(1-\alpha-\theta-\frac{1}{\eta}\right)\right]}{\left[\left(\frac{p_j\kappa_j}{w}+1\right)\frac{w}{p(1-\alpha-\theta)}\right]^{\frac{\eta(1-\alpha-\theta)-1}{1+\eta(\alpha+\theta)}}}. \quad \blacksquare \tag{11}$$

Proof of Proposition 3 From the first order condition of an entrepreneur of productivity z and that of the marginal entrepreneur $(\hat{z}_j)$, we obtain

$$l\left(z\right) = \left(\frac{z}{\hat{z}_j}\right)^{\frac{1}{1-\alpha-\theta}} l\left(\hat{z}_j\right).$$

Thus,

$$\Pr\left[\tilde{l}_j > l\right] = \Pr\left[z > \left(\frac{l}{l\left(\hat{z}_j\right)}\right)^{1-\alpha-\theta}\hat{z}_j \,\middle|\, z \geq \hat{z}_j\right] = \left(\frac{l\left(\hat{z}_j\right)}{l}\right)^{\eta(1-\alpha-\theta)}.$$

The aggregate establishment size distribution in the economy is then given by a mixture of Pareto distributions:

$$\Pr\left[\tilde{l}_j > l\right] = n_S\left(\frac{l\left(\hat{z}_S\right)}{l}\right)^{\eta(1-\alpha-\theta)} + n_M\left(\frac{l\left(\hat{z}_M\right)}{\max\left\{l, l\left(\hat{z}_M\right)\right\}}\right)^{\eta(1-\alpha-\theta)}, \quad l \geq l\left(\hat{z}_S\right).$$

Finally, by integrating $l(z) = \frac{z^{\frac{1}{1-\alpha-\theta}}}{Z_j^{\frac{1}{1-\alpha-\theta}}} \frac{L_j}{N}$ over z, we calculate the average establishment size in sector j:

$$\bar{l}_j = \frac{L_j}{N(1-\mu(\hat{z}_j))}. \tag{12}$$

The optimal allocation of labor L_j and entrepreneurs $N(1-\mu(\hat{z}_j))$ to sector j implies:

$$\theta p_j N^{1-\alpha-\theta} Z_j K_j^{\alpha} L_j^{\theta-1} = w, \tag{13}$$

$$(1-\alpha-\theta) p_j N^{-\alpha-\theta} Z_j K_j^{\alpha} L_j^{\theta} = p_j \kappa_j (1+r)(1-\mu(\hat{z}_j)) + w\mu(\hat{z}_j). \tag{14}$$

Taking the ratio of these two conditions, we obtain:

$$\frac{1-\alpha-\theta}{\theta} \frac{L_j}{N(1-\mu(\hat{z}_j))} = \frac{p_j \kappa_j (1+r)}{w} + \frac{\mu(\hat{z}_j)}{1-\mu(\hat{z}_j)}. \tag{15}$$

Substituting (14) into (10), we obtain:

$$\hat{z}_j = \left\{ \frac{\left(\kappa_j (1+r) + \frac{w}{p_j}\right) \left[\frac{(1-\alpha-\theta)}{\eta(1-\alpha-\theta)-1}\eta\right]^{\alpha+\theta}}{\frac{p_j \kappa_j (1+r)(1-\mu(\hat{z}_j)) + \frac{w}{p_j}\mu(\hat{z}_j)}{Z_j}} \right\}^{\frac{1}{1+\eta(\alpha+\theta)}}, \tag{16}$$

$$\frac{p_j \kappa_j (1+r)}{w} + \frac{\mu(\hat{z}_j)}{1-\mu(\hat{z}_j)} = \left[\frac{p_j \kappa_j (1+r)}{w} + 1\right] \frac{\eta(1-\alpha-\theta)}{\eta(1-\alpha-\theta)-1}. \tag{17}$$

Combining (12), (15) and (17), we obtain the desired expression:

$$\frac{\bar{l}_j}{\bar{l}_{j'}} = \frac{p_j \kappa_j + w}{p_{j'} \kappa_{j'} + w}. \quad \blacksquare \tag{18}$$

References

Aghion, P. and P. Bolton (1997): "A Theory of Trickle-Down Growth and Development," *Review of Economic Studies*, 64, 151–172.

Amaral, P. and E. Quintin (2009): "Financial Intermediation and Economic Development: A Quantitative Assessment," *International Economic Review*, Forthcoming.

Balassa, B. (1964): "The Purchasing Power Parity Doctrine: A Reappraisal," *Journal of Political Economy*, 72, 584–596.

Banerjee, A. V. and E. Duflo (2005): "Growth Theory through the Lens of Development Economics," in *Handbook of Economic Growth*, ed. by P. Aghion and S. N. Durlauf, Amsterdam: Elsevier, vol. 1A, 473–552.

Banerjee, A. V. and A. F. Newman (1993): "Occupational Choice and the Process of Development," *Journal of Political Economy*, 101, 274–298.

Beck, T., A. Demirgüç-Kunt, L. Laeven, and R. Levine (2008): "Finance, Firm Size, and Growth," *Journal of Money, Credit and Banking*, 40, 1327–1554.

Beck, T., A. Demirgüç-Kunt, and R. Levine (2000): "A New Database on the Structure and Development of the Financial Sector," *World Bank Economic Review*, 14, 597–605.

Buera, F. J. and J. P. Kaboski (2009): "Can Traditional Theories of Structural Change Fit The Data?" *Journal of the European Economic Association*, 7, 469–477.

Buera, F. J., J. P. Kaboski, and Y. Shin (2009): "Finance and Development: A Tale of Two Sectors," Working Paper 14914, National Bureau of Economic Research.

Buera, F. J. and Y. Shin (2008): "Financial Frictions and the Persistence of History: A Quantitative Exploration," Manuscript, Washington University in St. Louis.

Caselli, F. and N. Gennaioli (2005): "Dynastic Management," Manuscript, London School of Economics.

Castro, R., G. L. Clementi, and G. MacDonald (2009): "Legal Institutions, Sectoral Heterogeneity, and Economic Development," *Review of Economic Studies*, 76, 529–561.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (1997): "The Poverty of Nations: A Quantitative Investigation," Staff Report 204, Federal Reserve Bank of Minneapolis.

Clementi, G. L. and T. Cooley (2009): "Executive Compensation: Facts," Manuscript, New York University.

Davis, S. J., J. C. Haltiwanger, R. Jarmin, and J. Miranda (2006): "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms," in *NBER Macroeconomics Annual 2006*, ed. by M. Gertler and K. Rogoff, Cambridge, MA: MIT Press, 107–156.

Djankov, S., R. La Prota, F. Lopez-de Silanes, and A. Shleifer (2002): "The Regulation of Entry," *Quarterly Journal of Economics*, 117, 1–37.

Eaton, J. and S. Kortum (2001): "Technology, Trade, and Growth: A Unified Framework," *European Economic Review*, 45, 742–755.

Erosa, A. (2001): "Financial Intermediation and Occupational Choice in Development," *Review of Economic Dynamics*, 4, 303–334.

Erosa, A. and A. Hidalgo Cabrillana (2008): "On Finance as a Theory of TFP, Cross-Industry Productivity Differences, and Economic Rents," *International Economic Review*, 49, 437–473.

Giné, X. and R. M. Townsend (2004): "Evaluation of Financial Liberalization: A General Equilibrium Model with Constrained Occupation Choice," *Journal of Development Economics*, 74, 269–307.

Greenwood, J., J. M. Sanchez, and C. Wang (2009): "Financing Development: The Role of Information Costs," *American Economic Review*, Forthcoming.

Guner, N., G. Ventura, and Y. Xu (2008): "Macroeconomic Implications of Size-Dependent Policies," *Review of Economic Dynamics*, 11, 721–744.

Hall, R. E. and C. I. Jones (1999): "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics*, 114, 83–116.

Herrendorf, B., R. Rogerson, and A. Valentinyi (2009): "Two Perspectives on Preferences and Structural Transformation," Manuscript, Arizona State University.

Herrendorf, B. and A. Valentinyi (2006): "Which Sectors Make the Poor Countries So Unproductive?" Manuscript, Arizona State University.

Hopenhayn, H. A. and R. Rogerson (1993): "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy*, 101, 915–938.

Hsieh, C.-T. and P. Klenow (2007): "Relative Prices and Relative Prosperity," *American Economic Review*, 97, 562–585.

——— (2009): "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, Forthcoming.

Inklaar, R. and M. P. Timmer (2008): "GGDC Productivity Level Database: International Comparisons of Output, Inputs and Productivity at the Industry Level," Working Paper GD-104, Groningen Growth and Development Centre.

Jeong, H. and R. M. Townsend (2007): "Sources of TFP Growth: Occupational Choice and Financial Deepening," *Economic Theory*, 32, 197–221.

——— (2008): "Growth and Inequality: Model Evaluation Based on an Estimation-Calibration Strategy," *Macroeconomic Dynamics*, 12, 231–284.

Jones, C. I. (1994): "Economic Growth and the Relative Price of Capital," *Journal of Monetary Economics*, 34, 359–382.

King, R. G. and R. Levine (1993a): "Finance and Growth: Schumpeter Might Be Right," *Quarterly Journal of Economics*, 108, 717–737.

——— (1993b): "Finance, Entrepreneurship and Growth: Theory and Evidence," *Journal of Monetary Economics*, 32, 513–542.

Klenow, P. J. and A. Rodríguez-Clare (1997): "The Neoclassical Revival in Growth Economics: Has it Gone Too Far?" in *Macroeconomics Annual 1997*, ed. by B. S. Bernanke and J. J. Rotemberg, Cambridge: MIT Press, 73–102.

La Porta, R., F. Lopez de Silanes, A. Shleifer, and R. W. Vishny (1998): "Law and Finance," *Journal of Political Economy*, 106, 1113–1155.

Lloyd-Ellis, H. and D. Bernhardt (2000): "Enterprise, Inequality and Economic Development," *Review of Economic Studies*, 67, 147–168.

Lucas, Jr., R. E. (1978): "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 9, 508–523.

McKenzie, D. J. and C. M. Woodruff (2006): "Do Entry Costs Provide an Empirical Basis for Poverty Traps? Evidence from Mexican Microenterprises," *Development and Cultural Change*, 55, 3–42.

Parente, S. L. and E. C. Prescott (2000): *Barriers to Riches*, Cambridge: MIT Press.

Paulson, A. L. and R. M. Townsend (2004): "Entrepreneurship and Financial Constraints in Thailand," *Journal of Corporate Finance*, 10, 229–262.

Piketty, T. (1997): "The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing," *Review of Economic Studies*, 64, 173–189.

Rajan, R. G. and L. Zingales (1998): "Financial Dependence and Growth," *American Economic Review*, 88, 559–586.

Restuccia, D. and R. Rogerson (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, 11, 707–720.

Restuccia, D., D. T. Yang, and X. Zhu (2008): "Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis," *Journal of Monetary Economics*, 55, 234–250.

Rossi-Hansberg, E. and M. L. J. Wright (2007): "Establishment Size Dynamics in the Aggregate Economy," *American Economic Review*, 97, 1639–1666.

Samuelson, P. A. (1964): "Theoretical Notes on Trade Problems," *Review of Economics and Statistics*, 46, 145–154.

Sutton, J. (1991): *Sunk Costs and Market Structure*, Cambridge, MA: MIT Press.

Townsend, R. M. (2010): *Financial Systems in Developing Economies: Growth, Inequality and Policy Evaluation in Thailand*, Oxford: Oxford University Press, first ed.

Tybout, J. R. (2000): "Manufacturing Frims in Developing Countries: How Well Do They Do, and Why?" *Journal of Economic Literature*, 38, 11–44.

Valentinyi, A. and B. Herrendorf (2008): "Measuring Factor Income Shares at the Sector Level," *Review of Economic Dynamics*, 11, 820–835.