

Why Do Inefficient Firms Survive? Management and Economic Development

Michael Peters*

January 2012

Abstract

There are large and persistent productivity differences across firms within narrowly defined industries. This is especially true in poor countries. Why do productivity differences decline as the economy develops? In this paper I propose a theory where productivity differences exist because different firms use different technologies. The negative correlation between economic development and productivity dispersion occurs because the set of economically viable techniques shrinks as the economy develops. My mechanism stresses the role of managerial inputs. If managers are essential to increase the scale of production, inefficient techniques survive in managerial-scarce economies as productive firms do not have the means to replace them. As the aggregate supply of managers increases, efficient firms expand, best-practice technologies dominate the industry and productivity differences decline. Using firm-level panel data from Chile, I test both cross-sectional and time-series implications of the theory and evaluate different approaches of how to introduce management in firms' production function.

1 Introduction

There are large and persistent differences in labor productivity across firms within narrowly defined industries. While this has also been established in developed economies, this seems to be particularly true in poor countries (Foster, Haltiwanger, and Syverson, 2008; Hsieh and Klenow, 2009; Bartelsman, Haltiwanger, and Scarpetta, 2009). As an example, consider Table 1, where I report the within-sector dispersion of revenue labor productivity in Indonesia, Colombia, Chile and France. It is not only seen that the dispersion is substantial but more interestingly, the dispersion is much lower in rich countries. The dispersion of labor productivity within an industry in Indonesia is more than 60% higher than in France. And even compared to a middle-income country like Chile, Indonesia has a 17% higher dispersion in labor revenue productivity. What is it about the growth process that reduces this heterogeneity across producers?

In this paper I argue that such differences in the average product of labor exist because different firms use different production technologies. Furthermore, the negative correlation between economic

*I especially thank my advisors Daron Acemoglu, Abhijit Banerjee and Rob Townsend for many helpful discussions and their ongoing support. I also thank Nick Bloom, Ricardo Caballero, V.V. Chari, Roberto Fattal Jaef, Daniel Keniston, Amir Kermani, Ben Moll, Michael Powell, Ivan Werning, participants of the MIT macro lunch and especially Joaquin Blaum and Felipe Iachan for their comments. I also thank Claire Lelarge for providing me with the data from France.

Labor Productivity				
	France (2005)	Chile (1992)	Colombia (1989)	Indonesia (1992)
sd	0.27	0.58	0.53	0.72
75-25	0.33	0.67	0.71	0.81
90-10	0.66	1.44	1.33	1.69
Prod diff 75-25	1.39	1.95	2.04	2.26
GDP capita	30973	6644	4549	2651

Notes: The table shows the within-industry dispersion of log labor productivity. An industry is a 4-digit sector. Labor productivity is calculated as $\frac{py}{wl}$, where py is the firm's real value added and wl denotes the firm's total wagebill. See appendix for details of the construction of the data.

Table 1: Differences in Labor Productivity Dispersion Within Industries Across Countries

development and productivity dispersion occurs because the set of economically viable techniques shrinks as the economy develops. My mechanism stresses the role of the availability of managerial inputs. In particular, I argue and provide empirical evidence that managers are essential to increase the scale of production. If managers are scarce in poor countries, inefficient techniques survive as productive firms do not expand sufficiently to replace them. This allows different technologies to coexist and causes the dispersion in labor productivity to be high. As the aggregate supply of managers grows, productive firms will have the means to expand and previously used technologies will cease to be economically viable. This reallocation process causes a concentration of production factors in “best-practice”-technologies and thereby reduces the within-sector dispersion of labor productivity.

I want to stress that it is the heterogeneity of technologies and not the heterogeneity of firms, which is at the heart of the mechanism. The data presented in Table 1 is about *revenue* measures of productivity and not about *quantity* measures.¹ In a canonical heterogeneous firms model a la Lucas (1978) or Melitz (2003), firms differ in their physical productivity but revenue-based productivity measures like labor productivity will be equalized across all firms. Precisely because all marginal products are equated, (physically) more productive firms will end up not having a higher labor productivity but they will be bigger. With heterogeneous technologies being active in the market however, there can still be ample heterogeneity in the average product of labor, despite all marginal products being equalized. It is only when the set of economically viable techniques shrink, that the dispersion of labor productivity declines.

To study this mechanism and to derive additional empirical predictions, I propose a highly tractable general equilibrium model, which formalizes this intuition. The model features within-sector heterogeneity in that each sectoral commodity can be produced using two technologies. While the (what I call) traditional technology uses only production workers, the modern technology has an efficiency advantage but also requires managerial inputs to produce. The only heterogeneity across sectors is the factor-neutral efficiency advantage of modern producers. In the competitive equilibrium, the available supply of managers is allocated according to the sectors' comparative advantage. While there are modern firms in all sectors of the economy, more managers will be allocated to those sectors, where the efficiency advantage is high and it is only in those sectors,

¹In the terminology of Foster, Haltiwanger, and Syverson (2008), this paper is about the dispersion of *TFPR* and *not* about the dispersion in *TFPQ*.

that modern firms can achieve a sufficiently large scale to fully replace traditional production techniques. In all other sectors, both technologies coexist and labor productivity differs across firms within the industry. Such productivity differences precisely exist, because managerial supplies are limited in the aggregate. With only few managers available to the economy as a whole, managerial demand of some sectors is low as they compete with other sectors where managerial productivity is comparatively high. As the aggregate supply of managerial inputs grows, within-sector productivity differences decline as the economy “can afford” to allocate managers to more and more sectors of the economy, which enables modern firms to replace traditional producers.

Besides this time-series comparative static result, the model also makes cross-sectional predictions. In particular, the model implies on the sectoral level a positive correlation between the average labor productivity and the average managerial intensity and a negative correlation between the within-sector employment share of low productivity units and the average managerial intensity. The key object generating those correlations are the equilibrium relative prices. Whenever different technologies coexist in a given sector, the equilibrium price is determined by the marginal agent within the sector. In the model, the marginal agent will always be a traditional producer so that relative prices do not depend on the efficiency of managerial firms. Managerial firms in more efficient sectors will therefore earn a premium as they benefit from relatively high prices. This induces them to increase their managerial intensity and expand their scale. By doing so, they increase the sector’s average productivity and reduce the employment share of low-productivity, traditional producers.

To test those predictions empirically, I analyze a comprehensive plant-level panel dataset of manufacturing firms in Chile. Crucially for this paper, this data has information on managerial inputs at the firm-level. This allows me to not only test the implications of the theory but to also provide some evidence on different ways to introduce management in the production function. Both in the time-series and the cross-section I show evidence, that is consistent with the model. In the cross-section, I show that the above mentioned predictions across sectors are borne out in the data. In the time-series, I show that the aggregate managerial intensity increased over time and that both the productivity dispersion and the employment share of low-productivity units within sectors declined as the model predicts.

Finally, I use the micro-data and the equilibrium implications of the model to evaluate different ways to introduce “management” in a neoclassical production function. The cross-sectional equilibrium implications are very informative to distinguish different approaches as they have strong implications for the correlation between firms’ inframarginal rents and the direction of reallocation - are resources moving from high to low labor productivity units or the other way round. In particular I show that thinking about management as an increase in a Lucas-type ‘span of control’ is consistent with the micro-data within industries, but has counterfactual implications for the cross-sectoral implications.

Related Literature The idea that managerial inputs might be central for firms’ ability to efficiently expand has a long tradition in economics. Of special importance for the particular mechanism of this paper is Edith Penrose’s landmark study “The Theory of the Growth of the Firm”(Penrose, 1959). Penrose not only argues that managerial resources “create a fundamental and inescapable limit to the amount of expansion a firm can undertake at any time” but also that it is precisely this scarcity of managerial inputs which provide so called interstices for small firms as “the bigger firms have not got around to mopping them up” (Penrose, 1959, p. 221). My model has exactly these two features: the scope of expansion is determined by the aggregate supply of manage-

rial resources and traditional producers survive in those sectors of the economy where bigger firms are limited in the amount of managerial resources they want to buy at at the current market price. However, the precise mechanism how managerial inputs are allocated across producer differs from Penrose's original work. While she argues that managerial inputs are to a large degree firm-specific, my model will only feature frictionless spot-market transactions. In her work the supply-bottleneck for managers is therefore at the level of the individual firm. I am focusing on a simpler problem - what are the consequences of managerial scarcity at the aggregate level. The view that managerial inputs were the crucial factor determining the efficiency with which firms could expand features also prominently in the work of Chandler (1977) and Robinson (1934). Especially the scarcity of managerial supplies in poor countries is discussed by the latter, who argues that "if we contrast the highly industrialized with the unindustrialized countries we may perhaps feel justified in hoping that industrial problems are very gradually producing men of the ability necessary to handle them." (Robinson, 1934, p. 255)

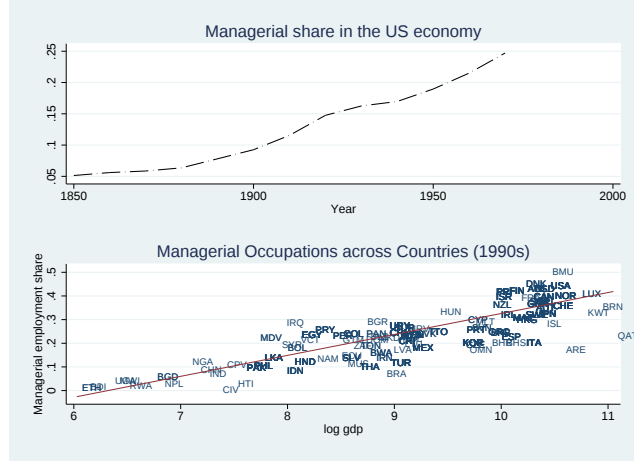
The interest in cross-country differences in managerial inputs has been revived recently. Bloom and Reenen (2007, 2010) provide evidence that managerial quality differs markedly across countries and argue that such differences might be important to account for the income differences across countries.² Bloom, Eifert, McKenzie, Mahajan, and Roberts (2010) report experimental evidence from India that providing managerial know-how increases productivity and argue that "one reason why better run firms do not dominate the market is constraints on growth through managerial span of control" as effective managers (outside the family) are in short supply given the contractual environment. This is also suggested by Hsieh and Klenow (2011) to explain differences in firms' life-cycle growth experience across countries. That managerial capital is missing in developing countries is also argued in Bruhn, Karlan, and Schoar (2010). These authors however focus on managerial inputs increasing productivity instead of being a necessary ingredient to expand scale.

Finally, there is a large and growing literature on productivity differences across firms (see Syverson (2011) for a recent survey). The majority of the literature follows Restuccia and Rogerson (2008) and Hsieh and Klenow (2009) to interpret differences in the average product of labor as dispersion in marginal products and hence concludes that poor countries are characterized by a larger degree of misallocation.³ There is a vast literature trying to provide a theoretical foundation for this pattern through financial frictions (recent contributions include Buera, Kaboski, and Shin (2011); Moll (2010); Banerjee and Moll (2010) and Midrigan and Xu (2010)), differences in mark-ups (Peters, 2010) or the presence of adjustment costs (Collard-Wexler, Asker, and De Loecker, 2011). These contributions are of course very different than this paper. There productivity differences are a sign of misallocation as some firms are not at their efficient size. In this paper, the economy is fully efficient and differences in productivity precisely stem from the fact that the efficient size is different for firms using different technologies.

The structure of the paper is as follows. In the next section I provide empirical evidence about the two crucial ingredients of the model, namely that economic development is correlated with an increase in the aggregate supply of managers and that managers are essential to expand the scale of production. Then I turn to the description of the model. After describing how I incorporate management in the production function, I discuss a very simple example to illustrate the main

²In fact, already 35 years Alfred Chandler argues that "administrative coordination helps to account for a significant segment of what economists have defined as a residual, that is the proportion of output that cannot be explained by the growth of input" (Chandler, 1977, p. 490)

³One recent paper that explicitly links measured productivity differences to the level of development is Ziebarth (2011) who applies the framework of Hsieh and Klenow (2009) to US plant level data from the 19th century and finds similar levels of productivity dispersion as in modern day India and China



Notes: The top panel depicts the time-series evolution of the managerial employment share of the US. The data is taken from the IPUMS files from the US Census. Managerial occupations are defined using the Census occupational category “Managers, Officials and Proprietors”. The bottom panel depicts the managerial employment share across countries. The GDP data comes from the Penn World Tables. Managerial employment is taken from the International Labor Organization (ILO). Managerial occupations are defined using the ILO occupational categories “Managerial, administrative, clerical and executive” occupations.

Figure 1: Managerial Suppliers and Economic Development

mechanism of this paper. Then I present the general multi-sector general equilibrium model and derive its empirical implications. In Section 4 I test those implications using the Chilean plant level data. Finally I go back to the management production function and use both the equilibrium implications of the model and the micro data to try to distinguish different specifications. Section 6 concludes.

2 Empirical Motivation for the Mechanism

This paper argues that managerial inputs into production are essential to expand firms’ scale of production. In such a world, the aggregate supply of managers determines the reallocation process across firms. If managers are scarce, the scope for expansion is limited and productivity differences across firms persist. Once managerial supplies grow, productive firms can expand and replace inefficient production techniques. While I will be presenting more empirical evidence about the particular mechanism suggested by the model below, I want to first present two pieces of evidence, which motivate the particular mechanism I am going to explore.

Consider first the variation in managerial supplies. For differences in the scarcity of managerial inputs to play an important part for the pattern of within-sector productivity differences across countries depicted in Table 1, it better be the case that the economywide supply of managers is positively correlated with economic development. That this seems to be the case is seen in Figure 1 below. In the top panel I depict the managerial employment share in the time series of the US. The data stems from the US census and I define workers in managerial occupations as all individuals that are assigned the occupational category “Managers, Officials and Proprietors” in the Census. There is a strong positive correlation in the time series in that the growth experience of the US

Dep. Variable: Managerial employment share				
$\ln(y)$	0.0878** (0.00496)	0.0823** (0.0212)	0.152** (0.0375)	0.136** (0.0410)
$\ln(k)$	0.0303** (0.0109)	-0.0283 (0.0181)	-0.0882** (0.0381)	-0.0612 (0.0423)
$\ln(h)$	0.246** (0.0597)	0.198** (0.0516)	0.293 (0.392)	0.0434 (0.431)
years of schooling			-0.00592 (0.0328)	0.0156 (0.0360)
labor share			-0.00512 (0.0983)	0.0612 (0.116)
Regional controls	No	No	No	Yes
N	71	67	67	45
R^2	0.775	0.761	0.813	0.759

Notes: Robust standard errors are shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. The dependent variable is the managerial employment share from the Internal Labor Organization. $\ln(y)$, $\ln(k)$ and $\ln(h)$ are the (log of) income per capita, the per-capita capital stock and a measure of human capital. "years of schooling" are the average years of schooling. The regional controls in column 5 contain dummy variables for the OECD, Africa and Asia. All the independent variables are taken from Caselli (2005).

Table 2: Determinants of managerial employment shares across countries

since the mid 19th century has been characterized by a trend of managerial deepening. While hardly 5% of the laborforce have been working in managerial occupations in the 19th century, in 1970, almost a quarter of the US laborforce has been employed in managerial occupations. The same pattern is apparent in the cross-country comparison. In the bottom panel of Figure 1 I show the simple correlation between income per capita and the managerial employment share in the 1990s. The employment data comes from the International Labour Organization (ILO) and I define as managerial workers all employees which work in managerial, executive, administrative and clerical occupations according to the ILO. Like for the US time series, there is a strong positive correlation between economic development and the reliance of managerial inputs.⁴ Furthermore, the cross-sectional relationships is not entirely driven by a simple correlation between managerial occupations and measures of human or physical capital. In Table 2 I take the cross-sectional "development accounting data" from Caselli (2005) and show that while measures of the stock of human and physical capital are highly correlated with the managerial employment share (column 2), they do not explain the correlation with income. In the remaining columns it is seen that this correlation is robust to other controls used in the traditional development accounting exercises.

In the preceding paragraph I took care in calling the relationship depicted in Figure 1 as an increase in the managerial employment share - and not an increase in managerial supplies. To

⁴The classification of the ILO and the US Census is somewhat different in that the classification of the ILO is more encompassing. I experimented with different measures of managerial employment shares (both for the Census and the ILO data) and they all have the same message: Economic development is positively correlated with employment in managerial occupations both in the cross-section of countries and the time series within countries.

distinguish demand and supply factors in the changing employment shares of skilled workers in general and managerial personnel in particular is subject of a large literature and this paper does not have much to add to this debate (Murphy and Welch, 1993; Autor and Katz, 1999; Goldin and Katz, 1999). In contrast, I take the evidence presented in Figure 1 as being consistent with a positive correlation between economic development and managerial supplies and the model will explore the economic implications of these differences in managerial supply on within-industry productivity differences across firms. The model is also silent on *why* managerial supplies differ across countries. Of first order importance are surely differences in aggregate human capital supplies. This is also suggested by the strongly positive correlation between the managerial employment share and human capital reported in Table 2. A second potential determinant of managerial supplies are differences in the institutional environment, for example the legal system. Bloom, Eifert, McKenzie, Mahajan, and Roberts (2010) for example present evidence that firms in underdeveloped countries do not hire managers as delegating authority seems to be infeasible given the contractual environment. Such a mechanism can be thought of a reduction in supply of managerial efficiency units. While the human capital explanation would imply that managerial skills are literally absent in poor countries, this mechanism would argue that *marketable* managerial skills are in short supply.

The second important ingredient I want to capture in the model is a strong form of complementarity between managerial inputs and firms' expansion decision. The plant-level data from Chile, which I will describe in more detail below, allows me to study if managers are indeed an essential input to expand. If such complementarities are important, changes in the non-managerial labor force should be predictive for changes in the number of managers a firms employs. To test this prediction, I look at a regression of the form

$$g_{i,t,s}^M = \alpha + \beta g_{i,t,s}^{BC} + \gamma l_{i,t-1,s}^{BC} + \delta_t + \delta_s + u_{i,t,s}, \quad (1)$$

where $g_{i,t,s}^M$ and $g_{i,t,s}^{BC}$ are the growth rates of the number of managers and blue-collar workers of firm i between $t - 1$ and t in sector s , $l_{i,t-1,s}^{BC}$ is the firm's lagged level of blue-collar employment and δ_s and δ_t are set of year and sector fixed effects. (1) is a regression akin to Doms, Dunne, and Troske (1997), who study complementarities between firms' skill demand and technology adoption. The results are contained in Table 3 below. In the first three columns I test if firms expand their managerial workers whenever they increase their blue-collar employment. In terms of (1), I regress an indicator if $g_{i,t,s}^M > 0$ on an indicator if $g_{i,t,s}^{BC} > 0$. Table 3 shows that there is a strong positive correlation. In column 4 I literally estimate (1), which confirms the earlier results: the two growth rates are strongly positively correlated. To put these effects into perspective, columns 5 and 6 repeat this exercise when I use the growth rates of non-managerial white collar workers as a dependent variable. While column 5 also shows a positive correlation between "times of expansion", the coefficient is only a third as big as for managerial personnel. This is even stronger in column 6, where I show that there is a negative correlation between the growth rate of blue and white-collar non-managerial employment. I interpret this as saying that managers and blue-collar workers are complements, while blue-collar workers and non-managerial personnel are more substitutable. In column 7 I show that this negative correlation is not driven by a different sample of firms. Even those firms that report a negative correlation between the blue-collar and non-managerial growth rates report a positive one with managerial employment. I view these results as generally supportive of the idea that managers are an important complementary factor for firms to grow.

Dep. Variable	$1[g^M > 0]$			g^M	$1[g^{WC} > 0]$	g^{WC}	g^M
$1[g^{BC} > 0]$	0.164** (0.00504)	0.155** (0.00510)	0.0605** (0.00540)		0.0194** (0.00501)		
$\ln(l_{t-1}^{BC})$			0.0352** (0.00284)	0.00762** (0.00328)	-0.0548** (0.00282)	0.0106** (0.00475)	0.00631 (0.00397)
g^{BC}				0.118** (0.0145)		-0.161** (0.0252)	0.186** (0.0187)
N	38335	38335	34326	30684	34326	17686	17686
R^2	0.027	0.038	0.020	0.011	0.111	0.026	0.023

Notes: Robust standard errors are shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. All regression except for column 1 contain year and sector fixed effects (4 digit). $g_{s,t}^j$ denotes the growth rate of managers, blue collar workers and white-collar non managerial workers respectively. $1[g_{s,t}^j > 0]$ is an indicator variable if $g_{s,t}^j > 0$. $\ln(l_{t-1}^{BC})$ is the log of lagged blue collar employment.

Table 3: Management as an essential input for expansion

3 The Model

Motivated by the evidence above, I will now study the economic implications of differences in managerial supplies if managers are an essential input for firms to expand.

3.1 How to model management?

This paper interprets the observed within-sector differences in labor productivity as stemming from differences of techniques used. It then interprets the negative correlation between productivity-dispersion and income per capita as being generated by a process of reallocation: as rich countries are manager-abundant, manager-intensive firms replace other technologies used and thereby reduce the dispersion of labor productivity within narrowly defined industries. To model this phenomenon, I not only have to specify how managers enter the production function, I also have to specify how the managerial-intensive technology itself differs from the non-managerial technology. It is this difference between technologies that determines which firms will earn the inframarginal rents.

For the most part of this paper I will assume that the “traditional”, non-managerial firms produce using labor only and that their technology is linear

$$y^T = l. \quad (2)$$

Managerial firms in contrast have access to a production function of the form

$$y^M = qf(l, m),$$

where q is a factor-neutral productivity term and l and m denote the number of workers and managers employed. I will now put more structure on f .

In accordance with results of Table 3, I want to capture that managerial inputs are essential for firms to increase their scale of production. Hence, f should have a low elasticity of substitution

so that managers and production workers are strong complements. Additionally, given that the traditional technology is assumed to have constant returns, f has to have decreasing returns - otherwise, only the technology with the lowest marginal costs (at given factor prices) will prevail and there were not any differences in labor productivity within sectors. In order to emphasize that it is the scarcity of managerial supplies that limit the expansion of management-intensive firms, I will take f to take the following form

$$f(l, m) = \left(\alpha l^{\frac{\sigma-1}{\sigma}} + (1-\alpha) (m^\gamma)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \quad (3)$$

where σ is small (in particular $\sigma < 1$) and $\gamma < 1$. Putting the decreasing returns directly on the managerial input “inside” the CES aggregator has the convenient property that f can also be expressed as

$$f(l, m) = \left(\alpha + (1-\alpha) \left(\frac{m^\gamma}{l} \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} l \equiv A \left(\frac{m^\gamma}{l} \right) l.$$

This stresses the interpretation that the total factor productivity of labor $A \left(\frac{m^\gamma}{l} \right)$ decreases in scale if managers are scarce. However, it will be clear in the model below, that a formulation of $f(l, m) = \left(\alpha l^{\frac{\sigma-1}{\sigma}} + (1-\alpha) m^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1} \gamma}$ will have almost identical results.

There is one special case of (3) which is particularly interesting. This is the case of $\sigma = 0$, where

$$f(l, m) = \min \{l, m^\gamma\}. \quad (4)$$

(4) is an interesting formulation, precisely because it focuses on the role of managers being necessary to expand scale. In contrast to (3), managers do not affect the marginal or average product of labor but if firms want to increase their scale, they require managerial inputs to do so. In the main section of the paper, I will analyze mostly the case of (4) in order to focus on that particular margin what managers are useful for. However, I also state the main results of the model for the general CES case of (3). It will be seen that the economy is “continuous in σ ” so that all the results are entirely analogous. Finally, consider the productivity term q . Throughout the paper I will assume that $q > 1$. For the case of (4) it is obvious that $q > 1$ is necessary for managerial firms to be willing to produce.⁵ I also think that $q > 1$ is the economically interesting case as I want to think of managerial scarcity preventing efficient firms to expand.

As mentioned above, it is not essential where I put the decreasing returns to scale *given* that it will be the managerial firms that earn inframarginal rents in this economy. *That* it is the managerial technology which has decreasing returns however is crucial. I will come back to this in Section 5 below, where I discuss the case of traditional producers having a decrease returns technology. In particular, I will show the general equilibrium implications of this case differ markedly from the setup described above and that these implications are at odds with the plant-level data.

3.2 A simple example

Before I present the general model, let me consider the simplest example that can illustrate the economic mechanism that is at the heart of this paper. Consider an economy, which consists of a

⁵For the general CES case, q could be smaller than one precisely because managerial firms earn inframarginal rents. In fact I will show below that the condition for managerial firms making positive profits is given by $q^{1-\sigma} \left(\frac{1}{\alpha} \right)^\sigma > 1$. For $\sigma = 0$, this implies that $q > 1$.

unique final good. The economy has access to two production techniques. In particular, there is a measure one of entrepreneurial firms that produce according to the production function

$$y = q \min \{l, m^\gamma\}, \quad q > 1, \quad \gamma < 1,$$

where as above l denotes the amount of production workers and m is the number of managers. The economy has also access to a traditional production technology, which transforms production labor into output, i.e.

$$y = l.$$

Workers and managers are in fixed supply L and M respectively and all markets are competitive. Characterizing the equilibrium is straight-forward. If both techniques are active, the equilibrium wage for production workers will be equal to the price of the final good, which is taken to be the numeraire. Given $w_L = 1$, profits of modern producers are given by

$$\pi = (q - 1) m^\gamma - w_M m,$$

as the Leontief structure of the production function implies that $l = m^\gamma$. Managerial demand is therefore determined from the first-order condition

$$\gamma (q - 1) m^{\gamma-1} = w_M. \quad (5)$$

Now consider the differences in labor productivity in this economy (which will of course trivially correspond to within-sector productivity differences). Following Foster, Haltiwanger, and Syverson (2008, 2011) I measure productivity as revenue per adjusted labor input. This involves weighing different types of labor by their relative wage. Formally, labor productivity is measured as

$$\xi \equiv \frac{py}{l + \frac{w_M}{w_L} m}. \quad (6)$$

Simply applying (6) to the different firms in this economy yields that

$$\xi^T = \frac{y}{l} = 1 \quad (7)$$

and

$$\xi^M = \frac{y}{l + \frac{w_M}{w_L} m} = \frac{qm^\gamma}{m^\gamma + w_M m} = \frac{q}{1 + \gamma(q - 1)} > 1, \quad (8)$$

where the last equality uses (5). Hence, entrepreneurial firms produce at a higher labor productivity precisely because they earn some inframarginal rents. Note that $\xi^M = 1$ if $\gamma = 1$, which again stresses that labor productivity is not primarily a statement about firm-specific factor neutral productivity (here q) but about the inframarginal rents different producers earn. The reason why I assumed that it is the managerial firms that earn inframarginal rents is seen in Table 4 below. There I report the results from running a regression of log labor productivity ($\ln(\xi)$) on firms' managerial intensity and a full set of industry fixed effects. Hence, the coefficient on the managerial intensity is the within-industry correlation between the managerial intensity and the inframarginal rents the firm earns. According to the model, the variation in the managerial share is generated

Dep. Variable	log labor productivity (ξ)					
Managerial intensity	0.625** (0.0349)	0.300** (0.0321)	0.241** (0.0300)			
$\ln(\frac{k}{l})$		0.249** (0.00365)	0.222** (0.00335)	0.243** (0.00365)	0.217** (0.00336)	
Managerial expenditure share				0.733** (0.0262)	0.450** (0.0244)	0.391** (0.0228)
N	41693	41693	40669	41693	41693	40669
R^2	0.289	0.390	0.419	0.299	0.394	0.423

Notes: Robust standard errors are shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. All regression contain year and sector fixed effects. A sector is a 4-digit industry. The dependent variable is log labor productivity (see (6)). “Managerial intensity” is the firm’s managerial employment share $\frac{m_i}{l_i}$, where m_i is the number of managers and l_i denotes total employment. $\ln(\frac{k}{l})$ denotes the log of firm’s capital intensity. “Managerial expenditure share” is given by firm’s managerial compensation over the total wagebill.

Table 4: Managerial intensity and inframarginal rents

by technological differences within the industry. The strong positive correlation suggests that inframarginal rents are earned by the - in the model - modern firms and not the traditional firms.⁶

Note also that measured productivity does not depend on any variables determined in equilibrium - in fact to derive (7) and (8) I did not even use the factor market clearing conditions. Once we impose market clearing, the equilibrium quantities are determined as

$$m = M, l_M = M^\gamma \text{ and } l_T = L - M^\gamma,$$

and the equilibrium wage for managers is simply given by $w_M = (q - 1)\gamma M^{-(1-\gamma)}$.

Now consider an increase in the aggregate supply of managers M . There are three main implications:

1. The dispersion in labor productivity vanishes once traditional technologies are fully replaced.
2. The employment share in low productivity establishments declines as more and more workers will be working in units that do earn inframarginal rents.
3. The economy’s average labor productivity increases as high-productivity units expand their scale.⁷

While admittedly simple, there are a couple of properties, which will also feature in the more general model. In particular, these three main implications will have their counterpart in the cross-sectoral dimension and the implications of managerial deepening are slightly more nuanced as the allocation of managers across sectors will respond. It is this model where I turn now.

⁶However, I will come back to this point in Section 5 when I present some other conflicting evidence.

⁷One way to see that this reallocation to high productivity units is mainly about increasing the inframarginal rents in the economy, is to note that the aggregate labor share $\frac{w_L L + w_M M}{Y} = \frac{L + (q-1)\gamma M^\gamma}{L + (q-1)M^\gamma}$ is decreasing in M .

3.3 The general model

I will now take the basic environment of this simple model and put it in a multi-sector environment. Doing so is useful for three reasons. First of all, the model generates additional economic insights, which are absent from the simple model. In particular, it allows me to study how different sectors in the economy “bear the burden” of managerial scarcity in the aggregate economy. Which sectors will absorb the few managers the economy has? How are different sectors affected by an increase in the aggregate supply of managers in the economy? And which sectors are characterized by within-sector productivity differences, precisely because they have to compete with other sectors for managers - the scarce resource in the economy? These are all questions, which the single-sector model above can not address. Secondly, the model generates additional empirical prediction on the sector level, which I can test using the plant level data. Finally, it turns out that it is these cross-sectoral predictions, which will be informative about the management production function and I will come back to this in Section 5 below.

Environment The environment is the following. The economy consists of a continuum of sectors in the unit interval. I will denote sectors by $i \in [0, 1]$. Individuals aggregate this unit mass of intermediate products according to a Cobb-Douglas aggregator

$$Y = \exp \left(\int_0^1 \ln (Y(i)) di \right), \quad (9)$$

where $y(i)$ is the total level of consumption of sector i 's production. The specific Cobb-Douglas form is convenient but not essential.⁸ There are three type of agents. There is a measure μ of entrepreneurs in each sector i (and hence also in the aggregate) that have access to a technology to produce variety i . Additionally there is a measure λ of managers and a measure $1 - \lambda$ of production workers. Each manager (worker) has $\frac{M}{\lambda}$ ($\frac{L}{\lambda}$) of efficiency units of labor so that the aggregate supply of managerial (worker) efficiency units is given by M and L respectively. Like in the example of Section 3.2, each sector's output $Y(i)$ can be produced by two technologies. Which of these technologies will prevail will be determined in equilibrium. More specifically, in each sector i production workers have access to a traditional technology, whose production function is given by

$$y_T(i) = A(i) l. \quad (10)$$

Here, l denotes the amount of production workers employed and $A(i)$ denotes the productivity in sector i . The distribution of $A(i)$ is unrestricted. Note that only production workers have access to this technology. Hence, I assume from the outset that managers will never want to run those technologies in equilibrium. Essentially this as an assumption on the parameters of the model that the relative wage of managers will be sufficiently high (in particular higher than the wage for production workers). I also assume that the technology in (10) is common knowledge among production workers and that there is free entry in all sectors.

Traditional producers in sector i compete with entrepreneurs. As in the single-sector example entrepreneurs have access to the technology

$$y_M(i) = q(i) A(i) \min \{l, m^\gamma\}. \quad (11)$$

⁸In the Appendix I derive the results for the more general CES specification with non-unitary demand elasticities.

Hence, $A(i)$ is a technology-neutral productivity term that applies to both the managerial and the traditional technology and $q(i)$ parametrizes the sector-specific comparative advantage of entrepreneurs. While the main analysis will focus on the specific Leontief case given in (11), I also state the results for the more general case of $y_M(i) = q(i) A(i) \left(\alpha l^{\frac{\sigma-1}{\sigma}} + (1-\alpha) m^{\gamma \frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$, whose analysis is relegated to the Appendix. It will be seen that all the results are “continuous in σ ” so that nothing is lost to focus on the notationally simpler case of $\sigma = 0$. Note also that I could allow entrepreneurs access to the traditional technology (10). However, precisely because entrepreneurs do earn inframarginal rents, they will strictly prefer to run the managerial technology.

Finally, I assume that all markets are perfectly competitive, i.e. all agents in the economy take wages (w_L, w_M) and sectoral prices $[p(i)]_{i=0}^1$ as given when making their decisions.

Equilibrium Consider now the equilibrium of this economy. An equilibrium has the obvious formal definition.

Definition 1. Consider the economy above. An equilibrium are wages for workers and managers (w_L, w_M) and sectoral prices $[p(i)]_{i=0}^1$ such that

- entrepreneurs and workers running the traditional technology maximize profits
- there is free entry in the traditional technology in all sectors
- factor markets for managers and production workers clear
- relative prices and quantities are consistent with demands generated from (9)

To characterize the equilibrium it is useful to take the productivity advantage q as the main state variable in this economy. In particular, as q is the main source of heterogeneity across sectors, let us index sectors directly by q and let $A(q)$ be the corresponding factor neutral productivity term which applies to all producers. I will denote the distribution of q by G_q and the support of q is given by $[q_L, \infty)$, where $q_L > 1$. Similarly, sectoral prices can be expressed directly as a function of q , i.e. the collection of sectoral prices is given by $[p(q)]_q$. Taking all prices as given, the entrepreneurial demand for production workers is given by

$$l_M(q) = m(q)^\gamma$$

and the scale of each firm is determined from

$$\begin{aligned} m(q) &= \operatorname{argmax}_m \{ (p(q) A(q) q - w_L) m^\gamma - w_M m \} \\ &= \operatorname{argmax}_m \left\{ \left(\frac{p(q) A(q) q}{w_L} - 1 \right) m^\gamma - \frac{w_M}{w_L} m \right\}. \end{aligned} \quad (12)$$

As clearly seen from (12), I can express managerial demands as a function of the wage premium $\frac{w_M}{w_L}$ and the statistic

$$z(q) \equiv \frac{p(q) A(q) q}{w_L} - 1, \quad (13)$$

which I will refer to as the *revenue productivity premium* in sector q . In particular, all the sectoral characteristics are fully contained in $z(q)$. (12) then directly implies that the managerial demand $m(z, \psi)$ is defined by

$$m(z(q), \psi)^{1-\gamma} = \gamma \psi z(q),$$

where $\psi \equiv \frac{w_L}{w_M}$ is the relative wage of production workers. This directly implies that $m(z, \psi)$ is increasing in both the revenue productivity premium z and in the relative wage of production workers.

The decision problem of traditional producers is straight-forward. Profits of traditional producers in sector q are given by

$$\pi^T = (p(q) A(q) - w_L) l,$$

which directly implies that equilibrium prices $[p(q)]_q$ and wages w_L have to satisfy the condition

$$p(q) \begin{cases} = \frac{w_L}{A(q)} & \text{if traditional producers are active in sector } q \\ < \frac{w_L}{A(q)} & \text{if traditional producers are not active in sector } q \end{cases}. \quad (14)$$

Hence, the main question is in which sectors traditional suppliers are active. In this economy, the answer is easy in that there will be a simple cutoff rule. In particular, let $\Omega^T(w_L, w_M)$ be the set of sectors where traditional suppliers produce at wages (w_L, w_M) and prices $[p(q)]_q$ satisfying (14). Then there is $\hat{q}(w_L, w_M)$ such that

$$q \in \Omega^T(w_L, w_M) \iff q \leq \hat{q}(w_L, w_M). \quad (15)$$

To see why (15) is true, note first that the Cobb-Douglas demand system in (9) implies that total sectoral revenues are equalized in all sectors, i.e.

$$p(q) Y(q) = x \text{ for all } q. \quad (16)$$

As total production in sector q is given by

$$Y(q) = A(q) l_T(q) + \mu q A(q) m(z(q), \psi)^\gamma,$$

(16) applied to $q, q' \in \Omega^T(w_L, w_M)$ implies that

$$p(q) A(q) [l_T(q) + \mu q m(z(q), \psi)^\gamma] = p(q') A(q') [l_T(q') + \mu q' m(z(q'), \psi)^\gamma].$$

As $p(q) A(q) = p(q') A(q') = w_L$, (13) implies that $z(q) = q - 1$ so that

$$l_T(q) = l_T(q') + \mu q' m(q' - 1, \psi)^\gamma - \mu q m(q - 1, \psi)^\gamma,$$

which implies that total employment in traditional technologies $l_T(q)$ is strictly decreasing in q . Hence, there is $\hat{q}$ such that in all sectors with $q > \hat{q}$ only entrepreneurial, management-intensive technologies will be active. This cutoff $\hat{q}$ is the key endogenous variable, which will be determined equilibrium. In particular, I can characterize the entire equilibrium allocations as a function of this cutoff $\hat{q}$. Consider first all sectors where only entrepreneurs are active, i.e. $q \geq \hat{q}$. The demand system (16) and the expression for $z(q)$ (see (13)) implies that

$$(1 + z(q)) m(z(q), \psi)^\gamma = (1 + z(\hat{q})) m(z(\hat{q}), \psi)^\gamma \text{ for all } q \geq \hat{q},$$

which directly shows that the revenue productivity premium $z(q)$ is equalized for all entrepreneurial firms, which do *not* compete against traditional suppliers within their sector. In terms of equilibrium prices, this together with (14) implies that

$$\begin{aligned} p(q, \hat{q}) &= \begin{cases} \frac{w_L}{A(q)} & \text{if } q < \hat{q} \\ \frac{w_L}{A(q)} \frac{\hat{q}}{q} & \text{if } q \geq \hat{q} \end{cases} \\ &= \frac{w_L}{A(q) q} \min\{q, \hat{q}\}, \end{aligned} \quad (17)$$

where the notation $p(q, \hat{q})$ stresses that the equilibrium price of sector q products depend on the aggregate statistic $\hat{q}$. Given the equilibrium pricing schedule in (17), the revenue productivity premium $z(q)$ of entrepreneurial firms is given by

$$z(q, \hat{q}) = \frac{p(q) A(q) q}{w_L} - 1 = \min\{q, \hat{q}\} - 1$$

and the allocation of managers across entrepreneurial firms of different sectors is given from (12) as

$$m(q, \hat{q}, \psi) = (\gamma \psi z(q, \hat{q}))^{\frac{1}{1-\gamma}} = (\gamma \psi)^{\frac{1}{1-\gamma}} (\min\{q, \hat{q}\} - 1)^{\frac{1}{1-\gamma}}. \quad (18)$$

Total production-worker employment in entrepreneurial firms is simply given by $l_M(q, \hat{q}, \psi) = m(q, \hat{q}, \psi)^\gamma$.

To finally determine aggregate employment in traditional technologies, use again the demand system (16) and the expression for equilibrium prices (17) to arrive at

$$l_T(q, \hat{q}, \psi) = \mu(\hat{q} l_M(\hat{q}, \hat{q}, \psi) - q l_M(q, \hat{q}, \psi)). \quad (19)$$

According to 3.2, traditional technologies will - in equilibrium - be used to account for the output not produced by large scale producers. Hence, the coexistence of heterogeneous technologies stems entirely from the demand side - less efficient producers survive because firms with a higher labor productivity are not willing to hire complementary factors (managers) to expand sufficiently. This keeps prices high and creates “interstices” for self-employed firms to survive (Penrose, 1959). Finally, consider the equilibrium wage for production workers in this economy. With the final good being the numeraire, the level of relative prices has to satisfy $\int_q \ln(p(q)) dG_q = 0$. Substituting (17) yields that

$$\ln(w_L) = \int_{q_L}^{\infty} \ln(A(q)) dG_q(q) + \int_{\hat{q}}^{\infty} \ln(q) dG_q(q) - \ln(\hat{q}) \int_{\hat{q}}^{\infty} dG_q(q), \quad (20)$$

which shows that wages are decreasing in $\hat{q}$. This is just the flip-side of traditional suppliers keeping prices high and is reminiscent of a long tradition of dual economy models in developing economics (Lewis, 1954). Given $(\psi, \hat{q})$ this fully characterizes all the allocations in the economy.

Now let us again measure labor productivity in this economy. In this economy we have that

$$\xi^T(q, \hat{q}) = \frac{p(q) A(q) l_T(q, \hat{q})}{l_T(q, \hat{q})} = w_L \quad (21)$$

and that

$$\begin{aligned} \xi^M(q, \hat{q}) &= \frac{p(q) A(q) q m(q, \hat{q}, \psi)^\gamma}{l_M(q, \hat{q}, \psi) + \frac{w_M}{w_L} m(q, \hat{q}, \psi)} = \frac{p(q) A(q) q}{1 + \gamma z(q, \hat{q})} = w_L + w_L \frac{(1 - \gamma) z(q, \hat{q})}{1 + \gamma z(q, \hat{q})} \\ &= w_L + w_L \frac{(1 - \gamma) (\min\{q, \hat{q}\} - 1)}{1 + \gamma (\min\{q, \hat{q}\} - 1)} > w_L \end{aligned} \quad (22)$$

Hence, akin to the simple single-sector model of Section 3.2, managerial firms earn a productivity premium precisely because they earn inframarginal rents ($\gamma < 1$). However, (21) and (22) now also contain results about the correlation of productivity across sectors. For traditional producers, labor productivity is equalized across sectors, precisely because the sectoral differences in productivity

$A(q)$ will shift relative prices equalize revenue per worker. This is different for management-intensive firms. As seen from (22), productivity is increasing in q as long as $q < \hat{q}$. This is of course precisely due to the fact that relative prices do *not* depend on q as long as traditional producers are present. Hence, managerial firms earn rents for two reasons: first of all they earn inframarginal rents simply due to their technology. But secondly, they earn rents because traditional firms (which are the marginal agents in this economy) put a floor on relative prices. This prevents relative prices from falling as sectoral productivity q increases and hence increases the revenue productivity premium, which entrepreneurial firms face. Once productivity differences *within* sectors vanish, i.e. for $q > \hat{q}$, productivity managerial productivity ξ^M is constant *across* sectors because prices then do depend on entrepreneurial productivity q which causes revenue per production unit to be constant. For future reference, I gather these results in the following Proposition.

Proposition 2. *Consider the economy above. Any equilibrium is characterized by a productivity cutoff $\hat{q}$, such that $l_T(q) = 0$ if and only if $q \geq \hat{q}$. Given this cutoff $\hat{q}$ and the relative wage $\psi = \frac{w_L}{w_M}$, equilibrium prices $[p(q)]_q$ are given by (17), the allocation of managers $m(q)$ is given in (18), employment in traditional technologies $l_T(q)$ is given in (19) and labor productivity ξ^T and ξ^M are given in (21) and (22).*

Proposition 2 characterizes the entire allocation as a function of the productivity cutoff and relative wage. To show that an equilibrium exists and that it is unique I only have to show that there exists a unique tuple $(\hat{q}, \psi)$ which clears the labor markets for both managers and production workers. Before doing so however, I want to reiterate that nothing in the simple characterization contained in Proposition 2 relied on the special Leontief structure. In particular, if I were to consider the general CES production function, the characterization of the equilibrium was essentially identical. It is contained in the following Proposition and the algebraic details can be found in the Appendix.

Proposition 3. *Consider the economy above and assume that the entrepreneurial technology is given as a general CES production function $f(l, m) = \left(\alpha l^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(m^\gamma)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$. Any equilibrium is characterized by a productivity cutoff $\hat{q}$, such that $l_T(q) = 0$ if and only if $q \geq \hat{q}$. Given this cutoff $\hat{q}$ and the relative wage $\psi = \frac{w_L}{w_M}$, equilibrium prices $[p(q)]_q$ are given by (17), the revenue productivity premium $z(q, \hat{q})$ is given by*

$$z^{CES}(q, \hat{q}) = \left(\frac{1}{\alpha} \right)^\sigma (\min\{q, \hat{q}\})^{1-\sigma} - 1,$$

the allocation of managers and production workers in entrepreneurial firms is given by

$$\begin{aligned} m^{CES}(q, \hat{q}, \psi) &= \left(\frac{1-\alpha}{\alpha} \right)^{\frac{\sigma}{\theta}} \gamma^{\frac{1}{1-\gamma}} \psi^{\frac{1}{1-\gamma}} z(q, \hat{q})^{\frac{1}{\theta}} \\ l_M^{CES}(q, \hat{q}, \psi) &= \left(\frac{1-\alpha}{\alpha} \right)^{\frac{\sigma}{\theta}} \gamma^{\frac{\gamma}{1-\gamma}} \psi^{\frac{\gamma}{1-\gamma}} z(q, \hat{q})^{\frac{1-\theta}{\theta}} \end{aligned}$$

where $\theta = (1-\gamma)(1-\sigma)$. Employment in traditional technologies is given by

$$l_T^{CES}(q, \hat{q}) = \left(\frac{1}{\alpha} \right)^\sigma (\hat{q}^{1-\sigma} l_M^{CES}(q, \hat{q}, \psi) - q^{1-\sigma} l_M^{CES}(q, \hat{q}, \psi))$$

and labor productivity ξ_T^{CES} and ξ_M^{CES} are given by

$$\begin{aligned}\xi_T^{CES}(q, \hat{q}) &= w_L \\ \xi_M^{CES}(q, \hat{q}) &= w_L + w_L \left(\frac{(1-\gamma) z^{CES}(q, \hat{q})}{1 + \gamma z^{CES}(q, \hat{q})} \right).\end{aligned}$$

Proof. See Appendix. $\square$

From Propositions 2 and 3 it is easy to see that the economy is essentially “continuous in σ ”. In particular, $\lim_{\sigma \rightarrow 0} z^{CES}(q, \hat{q}) = z(q, \hat{q}) = \min\{q, \hat{q}\} - 1$ and that the same holds true for all the allocations.

To fully characterize the equilibrium, I have to find the tuple $(\hat{q}, \psi)$, which ensures that the labor markets for both managers and production workers clear. Consider first the market for managers. Given that there are M units of managerial efficiency units to be hired in the market and that managerial demand only stems from entrepreneurial firms (given in (18)), the market clearing condition is given by

$$\frac{M}{\mu} = \int m(q, \hat{q}, \psi) dG_q(q) = \psi^{\frac{1}{1-\gamma}} \gamma^{\frac{1}{1-\gamma}} \int z(q, \hat{q})^{\frac{1}{1-\gamma}} dG_q(q).$$

Equilibrium on the market for production workers in turn requires that

$$L = \mu \int l_M(q, \hat{q}, \psi) dG_q(q) + \int_{q_L}^{\hat{q}} l_T(q, \hat{q}, \psi) dG_q(q).$$

Using (19) it can be shown (see Appendix) that the labor market clearing condition can be written as

$$\frac{L}{\mu} = \psi^{\frac{\gamma}{1-\gamma}} \gamma^{\frac{\gamma}{1-\gamma}} \left[\left(\frac{1 + z(\hat{q}, \hat{q})}{z(\hat{q}, \hat{q})} \right) z(q, \hat{q})^{\frac{1}{1-\gamma}} - \int z(q, \hat{q})^{\frac{1}{1-\gamma}} dG \right].$$

For future reference it is useful to rewrite these conditions as

$$\frac{M}{\mu} = \gamma^{\frac{1}{1-\gamma}} \psi^{\frac{1}{1-\gamma}} \Psi(\hat{q}) \tag{23}$$

$$\frac{L}{\mu} = \gamma^{\frac{\gamma}{1-\gamma}} \psi^{\frac{\gamma}{1-\gamma}} (D(\hat{q}) - \Psi(\hat{q})), \tag{24}$$

where

$$\begin{aligned}\Psi(\hat{q}) &= \int_q z(q, \hat{q})^{\frac{1}{1-\gamma}} dG_q(q) \\ D(\hat{q}) &= \left(\frac{1 + z(\hat{q}, \hat{q})}{z(\hat{q}, \hat{q})} \right) z(q, \hat{q})^{\frac{1}{1-\gamma}}.\end{aligned} \tag{25}$$

(23) and (24) are two equations in the two remaining unknowns $(\psi, \hat{q})$. That an equilibrium exists and it is unique is contained in the following Proposition.

Proposition 4. *Consider the economy above. There is a unique equilibrium $(\psi, \hat{q})$ as (23) and (24) have a unique intersection. The equilibrium is characterized by two functions*

$$\begin{aligned}\hat{q} &= \hat{q}\left(\frac{M}{\mu}, \frac{L}{\mu}, [q]_{q_L}^{\hat{q}}\right) \\ \frac{w_L}{w_M} &= \psi\left(\frac{M}{\mu}, \frac{L}{\mu}, [q]_{q_L}^{\hat{q}}\right)\end{aligned}$$

with

$$\begin{aligned}\frac{\partial \hat{q}}{\partial \frac{M}{\mu}} &< 0, & \frac{\partial \hat{q}}{\partial \frac{L}{\mu}} &> 0, & \frac{\partial \hat{q}}{\partial [q]_{q_L}^{\hat{q}}} &> 0 \\ \frac{\partial \frac{w_L}{w_M}}{\partial \frac{M}{\mu}} &> 0, & \frac{\partial \frac{w_L}{w_M}}{\partial \frac{L}{\mu}} &< 0, & \frac{\partial \frac{w_L}{w_M}}{\partial [q]_{q_L}^{\hat{q}}} &< 0.\end{aligned}$$

In particular, the equilibrium values $(\psi, \hat{q})$ do neither depend on $[A(q)]_{q_L}^{q_H}$ nor on $[q]_{\hat{q}}^{q_H}$.

Proof. See Appendix. □

The most important comparative static result for this paper is that productivity cutoff $\hat{q}$ is decreasing in the relative managerial supply of the economy. Hence, akin to the simple single sector model, as the aggregate supply of managerial resources grows, there will be more and more sectors in which traditional technologies are replaced. This comparative static result is contained in the following relationship, which determines the cutoff $\hat{q}$ as a function of relative managerial supplies. Using (23) and (24) it follows that

$$\frac{\left(\frac{M}{\mu}\right)^\gamma}{\left(\frac{L}{\mu}\right)} = \frac{\Psi(\hat{q})^\gamma}{D(\hat{q}) - \Psi(\hat{q})} \equiv Q(\hat{q}). \quad (26)$$

In the Appendix I show that $Q(\hat{q})$ is strictly decreasing (which is essentially the proof of Proposition (4)) and that

$$\begin{aligned}\lim_{\hat{q} \rightarrow q_L} Q(\hat{q}) &= 1 \\ \lim_{\hat{q} \rightarrow \infty} Q(\hat{q}) &= 0.\end{aligned}$$

$Q(\hat{q})$ is exactly the aggregate relative demand for managerial inputs. As entrepreneurial firms are management intensive, this relative demand is decreasing in $\hat{q}$ as a decline in $\hat{q}$ implies that more and more of the economy's output is produced using traditional techniques. It is in that sense that (26) exactly determines how much replacement of traditional technologies the economy can afford given that managerial supplies are limited. As with every scarce resource, managers are allocated according to their comparative advantage, which is parametrized by q . Hence, there will be managers in every sector of the economy but the relative importance of managerial firms is higher, the higher the productivity advantage of the technology they run. If managers are scarce, only a small pocket of the economy is characterized by technological homogeneity where traditional suppliers are fully replaced. As the managerial supply accumulates over time, this part of the economy grows, more and more traditional techniques will cease to be economically profitable and a smaller share of the economy will be characterized by a coexistence of different techniques and productivity differences.

With that intuition, the other comparative static results contained in Proposition (4) are also intuitive. First of all, note that the relative demand for managers $Q(\hat{q})$ does neither depend on

technology-neutral productivity $[A(q)]_q$ nor on the productivity advantage q in those parts of the economy where traditional suppliers have already been replaced, i.e. $[q]_{\hat{q}}^\infty$. This is again due to the general equilibrium adjustment of relative prices: factor neutral productivity differences, which apply to *all* producers within a sector equally, do not change their factor demands as relative prices adjust so that the price-weighted productivity is unaffected. Formally, this intuition gets its representation in the fact that the revenue productivity premium $z(q, \hat{q})$ does neither depend on $A(q)$ nor on $q > \hat{q}$. Secondly note that the cutoff $\hat{q}$ is *increasing* in the productivity of managerial firms whenever they compete with traditional firms. Hence, changes in productivity that make managerial techniques *more* productive lead to more sectors of the economy being populated by traditional techniques and to *lower* wages (see (20)). This seemingly counter-intuitive result is actually easy to understand. If productivity in a sector $q < \hat{q}$ increases, output of that particular commodity becomes relatively abundant. Its price however, cannot fall precisely because it is the productivity of the infra-marginal agent that increased. If reallocation was only taking place *within* sectors, production in that sector was now higher. Given that the existence of small-scale producers still put an upper bound on prices, this cannot be an equilibrium. Hence, production labor has to move *out* of this sector. To induce such reallocation, relative prices in sectors where self-employed firms have not been active previously *will increase*. This causes entry of traditional suppliers in sectors of the economy where their technology was not profitable prior to the productivity increase. It is of course exactly this increase in prices, which lowers the real wage as seen from (20). Propositions 2 and 4 fully characterize the equilibrium of this economy. Armed with those results I can now derive empirical predictions of this model.

3.4 Empirical Predictions

I will now use the model to test some of its testable implications using plant level data from Chile. For this paper it is of course essential to be able to observe the managerial intensity at which different firms produce. This is the case in the Chilean data as plants report their employment structure in detailed occupations categories, including their managerial staff. I will describe the data in more detail below. For the empirical application, I test both empirical predictions in a given cross-section (i.e. for a given value of $(\hat{q}, \psi)$) and in the time series. For the time-series implications I will mainly be thinking about the time-series variation being driven by changes in the aggregate managerial supply but I will also address what other aggregate changes could have generated the time-series variation in the data.

Cross-sectional predictions

The cross sectional predictions of the theory are essentially contained in Proposition 2. In particular, in line with the model, I will be thinking about the variation in the data as being generated by sectoral differences in q , i.e. in the comparative advantage of managerial-intensive technologies. To map the model to the firm-level data, I will weigh firms by their employment share. This is necessary, because the model does not have a clear mapping to the number of firms using the traditional technology given that the technology has constant returns. The employment share of

managerial firms in sector q is given by

$$\begin{aligned}
s_M(q, \hat{q}, \psi) &= \frac{\mu l_M(q, \hat{q}, \psi)}{\mu l_M(q, \hat{q}, \psi) + l_T(q, \hat{q}, \psi)} = \frac{\mu l_M(q, \hat{q}, \psi)}{\mu l_M(q, \hat{q}, \psi) + \mu(\hat{q} l_M(\hat{q}, \hat{q}, \psi) - q l_M(q, \hat{q}, \psi))} \\
&= \frac{l_M(q, \hat{q}, \psi)}{\hat{q} l_M(\hat{q}, \hat{q}, \psi) - (q-1) l_M(q, \hat{q}, \psi)} \\
&= \frac{(q-1)^{\frac{\gamma}{1-\gamma}}}{\hat{q}(\hat{q}-1)^{\frac{\gamma}{1-\gamma}} - (q-1)^{\frac{1}{1-\gamma}}},
\end{aligned}$$

which is clearly increasing in q (as long as $q < \hat{q}$). While q is not observable, the model suggests an obvious proxy for it - the average managerial intensity in the sector. In particular, let $\chi(q, \hat{q}, \psi)$ be the employment-weighted managerial share in the cross-section of firms within a sector. According to the model

$$\begin{aligned}
\chi(q, \hat{q}, \psi) &= \frac{m(q, \hat{q}, \psi)}{l_M(q, \hat{q}, \psi)} s_M(q, \hat{q}, \psi) = \left(\frac{m(q, \hat{q}, \psi)^\gamma}{l_M(q, \hat{q}, \psi)} \right) m(q, \hat{q}, \psi)^{1-\gamma} s_M(q, \hat{q}, \psi) \\
&= \left((\gamma\psi)^{\frac{1}{1-\gamma}} (\min\{q, \hat{q}\} - 1)^{\frac{1}{1-\gamma}} \right)^{1-\gamma} \frac{(q-1)^{\frac{\gamma}{1-\gamma}}}{\hat{q}(\hat{q}-1)^{\frac{\gamma}{1-\gamma}} - (q-1)^{\frac{1}{1-\gamma}}} \\
&= \frac{(\gamma\psi)(q-1)^{\frac{1}{1-\gamma}}}{\hat{q}(\hat{q}-1)^{\frac{\gamma}{1-\gamma}} - (q-1)^{\frac{1}{1-\gamma}}},
\end{aligned}$$

which is also increasing in q (as long as $q < \hat{q}$). Finally, the average productivity in sector q is given by

$$\begin{aligned}
ALP(q, \hat{q}, \psi) &= (1 - s_M(q, \hat{q}, \psi)) \xi^T(q, \hat{q}) + s_M(q, \hat{q}, \psi) \xi^M(q, \hat{q}) \\
&= (1 - s_M(q, \hat{q}, \psi)) w_L + s_M(q, \hat{q}, \psi) \left(w_L \left(1 + \frac{(1-\gamma)z(q, \hat{q})}{1 + \gamma z(q, \hat{q})} \right) \right) \\
&= w_L + w_L s_M(q, \hat{q}, \psi) \frac{(1-\gamma)z(q, \hat{q})}{1 + \gamma z(q, \hat{q})} \\
&= w_L + w_L s_M(q, \hat{q}, \psi) \frac{(1-\gamma)(q-1)}{1 + \gamma(q-1)},
\end{aligned}$$

which is also increasing in q both because $s_M(q, \hat{q}, \psi)$ and $\frac{q-1}{1+\gamma(q-1)}$ are increasing - not only have managerial firms high labor productivity in high q industries, they also account for a higher share of employment precisely because their high labor productivity induces them to invest in managerial resources which allows them to expand. Hence, in the cross-section of sectors, the model predicts

1. Positive correlation between the average managerial share and average productivity.
2. Positive correlation between the average managerial share and share of employment in management intensive firms within the respective industry.
3. Negative correlation between the average managerial share and the share of employment in low productivity production units within the respective industry.

Predictions in the Time-Series

Now think about the predictions in the time series. The model, in particular Proposition 4, makes tight predictions on which variables should have any effect on the firm-level allocations. Basically, the time series variation can either be induced by changes in the relative supply of managers or by changes in the comparative advantage of managerial intensive technologies in those parts of the economy, where different technologies still coexist. I will show below that the aggregate managerial intensity of the Chilean economy increased quite markedly between 1986 and 1995. So suppose that the time series variation in the data was induced by an increase in the relative aggregate supply of managers. According to the model, such an increase has the following observable implications

1. An decrease in the average of within-sector dispersion of log productivity. Note that it is important to consider the log of productivity because wages will respond to the increase in managerial supplies. In particular, they will increase. Note also that this implication holds true regardless of the weighting we give different sectors.
2. A decrease in the aggregate employment share in low productivity units within sectors, if sectors are weighted by their respective sectoral employment shares of non-managerial workers. To see this, note that

$$\begin{aligned}
 s_T(\hat{q}, \psi) &= \int_{q_L}^{\hat{q}} \frac{l_T(q, \hat{q}, \psi)}{L(q, \hat{q}, \psi)} \frac{L(q, \hat{q}, \psi)}{L} dG_q \\
 &= \frac{1}{L} \int_{q_L}^{\hat{q}} l_T(q, \hat{q}, \psi) dG_q \\
 &= \frac{1}{L} \int_{q_L}^{\hat{q}} \mu(\hat{q} l_M(\hat{q}, \hat{q}, \psi) - q l_M(q, \hat{q}, \psi)) dG_q,
 \end{aligned}$$

which is increasing in $\hat{q}$. Hence, an increase in managerial supplies decreases $s_T(\hat{q}, \psi)$ by reducing $\hat{q}$.

3. A decrease in the within-sector dispersion of log productivity of those sectors that have a high average product to begin with. This is due to the fact that the replacement of traditional technologies has a clear sectoral structure in this economy.

4 Empirical Application

I will now use plant-level panel data from Chile to test these empirical predictions in the data. The data has also been used in Moll (2010); Pavcnik (2002); Levinsohn and Petrin (2003).

Basic descriptive statistics about management While plant level data from developing economies have been widely used in the last couple of years, only few datasets contain information about managerial intensities at the firm level. Below I provide some basic descriptive statistics about firms' investment in managerial personnel.

Dep. Variable	Sectoral Average Productivity					
Managerial intensity	2.323** (0.173)	0.919** (0.118)	1.328** (0.218)		1.328** (0.421)	
$\ln\left(\frac{k}{l}\right)$		0.500** (0.0173)	0.488** (0.0297)	0.470** (0.0295)	0.488** (0.0298)	0.470** (0.0426)
Managerial expenditure share				1.016** (0.179)		1.016** (0.367)
N	462	462	462	462	462	462
R^2	0.303	0.787	0.789	0.788	0.789	0.788

Notes: Robust (columns 1-4) and clustered (on the sectoral level, columns 5,6) standard errors are shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. All regression contain year fixed effects. A sector is a 4-digit industry. The dependent variable the average (log) productivity in year t in sector s , i.e. $\frac{1}{N_{s,t}} \sum_{i=1}^{N_{s,t}} \ln(\xi_{s,t}^i)$, where $N_{s,t}$ is the number of firms in sector s at time t and $\xi_{s,t}^i$ is firm i 's revenue labor productivity (see (6)). "Managerial intensity" is the employment-weighted managerial share in sector s at time t , i.e. $\sum_{i=1}^{N_{s,t}} \left(\frac{m_i}{l_i}\right) \omega_i$, where m_i is the number of managers, l_i denotes total employment and ω_i is firm i 's employment share in industry s at time t , i.e. $\omega_i = l_i \left(\sum_{i=1}^{N_{s,t}} l_i\right)^{-1}$. "Managerial expenditure share" is the employment-weighted managerial share in sector s at time t , i.e. $\sum_{i=1}^{N_{s,t}} \chi_i \omega_i$, where χ_i is firm i 's expenditure share on managers (managerial compensation over the total wagebill). $\ln\left(\frac{k}{l}\right)$ is the average of log capital-intensity, i.e. $\ln\left(\frac{k}{l}\right) = \frac{1}{N_{s,t}} \sum_{i=1}^{N_{s,t}} \ln\left(\frac{k_i}{l_i}\right)$.

Table 5: Managerial Intensity and Average Productivity

4.1 Cross-sectional implications

I will first consider the cross-sectional implications derived above (see Section 3.4). To test those predictions I will consider regressions of the form

$$y_{s,t} = \alpha + \delta_t + \beta \chi_{s,t} + \gamma \ln\left(\frac{k}{l}\right)_{s,t} + u_{s,t}, \quad (27)$$

where $y_{s,t}$ is the respective dependent variable of interest, $\chi_{s,t}$ is the average managerial share in sector s at time t , $\ln\left(\frac{k}{l}\right)_{s,t}$ is the average log capital-intensity of sector s at time t and δ_t is a year fixed effect. Including the year fixed effects is conceptually important, because (27) is a cross-sectional relationship. In terms of the model, all the implications are conditional on a particular equilibrium, i.e. holding $\hat{q}$ fixed. Hence, δ_t in (27) should be thought of as controlling for $\hat{q}$. As derived above, I will consider three different dependent variables $y_{s,t}$. In particular, I will consider the mean productivity, the employment share in managerial-intensive production units and the employment share in low productivity firms in sector s at time t . I want to stress that for the latter two regressions, the respective employment shares are *within* sectors employment shares and (27) tests for a particular correlation how these within-sector statistics depend on a particular sectoral characteristics as implied by the model, namely the average managerial share. To control for the sector's capital-intensity is empirically important, because there is a strong correlation between the managerial intensity and the capital-intensity.

Consider first the average productivity in sector s as a dependent variable. The results are contained in Table 5 below. In column one, I report the simple correlation between the average

Dep. Variable	Within-industry employment share of managerial-intensive firms					
Managerial intensity	0.429** (0.0478)	0.495** (0.0517)	0.682** (0.0609)		0.682** (0.159)	
$\ln(\frac{k}{l})$		-0.0235** (0.00603)	-0.0356** (0.00536)	-0.0369** (0.00640)	-0.0356** (0.0112)	-0.0369** (0.0142)
Managerial expenditure share				0.409** (0.0495)		0.409** (0.121)
N	462	462	462	462	462	462
R^2	0.166	0.195	0.311	0.192	0.311	0.192

Notes: Robust (columns 1-4) and clustered (on the sectoral level, columns 5,6) standard errors are shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. All regression contain year fixed effects. A sector is a 4-digit industry. The dependent variable the share of employees within an industry-year cell that is employed at the 25% firms with the highest managerial intensity. "Managerial intensity" is the employment-weighted managerial share in sector s at time t , i.e. $\sum_{i=1}^{N_{s,t}} \left(\frac{m_i}{l_i}\right) \omega_i$, where m_i is the number of managers, l_i denotes total employment and ω_i is firm i 's employment share in industry s at time t , i.e. $\omega_i = l_i \left(\sum_{i=1}^{N_{s,t}} l_i\right)^{-1}$. "Managerial expenditure share" is the employment-weighted managerial share in sector s at time t , i.e. $\sum_{i=1}^{N_{s,t}} \chi_i \omega_i$, where χ_i is firm i 's expenditure share on managers (managerial compensation over the total wagebill). $\ln(\frac{k}{l})$ is the average of log capital-intensity, i.e. $\ln(\frac{k}{l}) = \frac{1}{N_{s,t}} \sum_{i=1}^{N_{s,t}} \ln\left(\frac{k_i}{l_i}\right)$.

Table 6: Managerial Intensity and Managerial Employment Share

managerial share and the sector's average productivity conditional on year fixed effects. The correlation is positive and strong. Column 2 shows why it is important to control for the sector's capital intensity. Putting this into the regression, reduces the coefficient on the managerial intensity by 60% but there is still a strong positive correlation. Column 3 reports the results when I weigh the regression by the number of firms used to calculate the sector-year averages. This increases both the coefficient and the standard error. Column 4 shows that these results do not depend on the particular measure of managerial intensity. Instead of the average number of managers per employee, column 4 uses the average expenditure share on managers as a measure of managerial intensity. According to the model, these two measures should be equivalent and indeed the coefficient is also positive and strongly significant. Finally, columns 5 and 6 present the results when I cluster the standard errors on the sector level. While the standard errors increase by 40%, both coefficients are still significant.

Consider now the second prediction, i.e. the positive correlation between the sectoral average managerial intensity and the within-sector employment share in managerial-intensive firms. The results are contained in Table 6. The structure of the Table is exactly the same as the one of Table 5. The first column shows that both variables are strongly positively correlated. That this correlation is neither driven by the sectoral capital-intensity, nor the weighting of firms is seen in columns 2 and 3. Column 4 shows again the robustness of this correlation with respect to the other measure of managerial intensity (i.e. the average managerial expenditure share). The last two columns finally show that these coefficients are significantly different from zero even when I cluster the standard errors on the sectoral level.⁹

⁹While the results in Table 6 are an implication of the theory, I do want to stress that simple forms of measurement errors will tend to introduce a mechanical correlation between the employment share of managerial intensive firms and the average managerial intensity. Suppose for example that all firms have the same managerial intensity $\frac{m}{l}$, but

Dep. Variable	Within-industry employment share of low productivity firms					
Managerial intensity	-0.177** (0.0550)	-0.255** (0.0622)	-0.266** (0.0476)	-0.266** (0.127)		
$\ln(\frac{k}{l})$		0.0278** (0.00623)	0.0153** (0.00504)	0.0230** (0.00516)	0.0153 (0.0138)	0.0230* (0.0131)
Managerial expenditure share				-0.263** (0.0406)	-0.263** (0.109)	
N	462	462	462	462	462	462
R^2	0.033	0.074	0.081	0.119	0.081	0.119

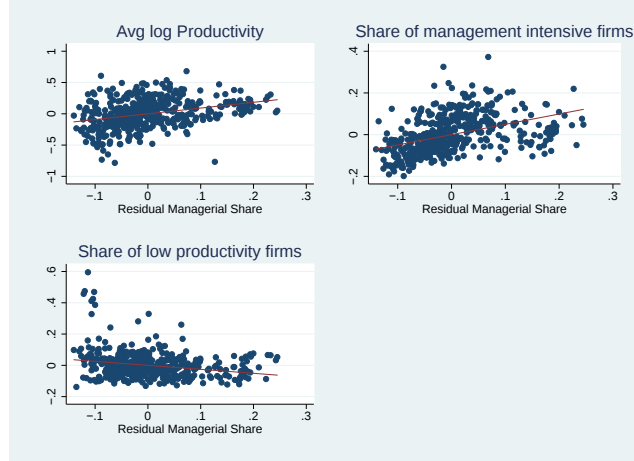
Notes: Robust (columns 1-4) and clustered (on the sectoral level, columns 5,6) standard errors are shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. All regression contain year fixed effects. A sector is a 4-digit industry. The dependent variable the share of employees within an industry-year cell that is employed at the 25% firms with the lowest labor productivity ξ (see (6)). “Managerial intensity” is the employment-weighted managerial share in sector s at time t , i.e. $\sum_{i=1}^{N_{s,t}} \left(\frac{m_i}{l_i}\right) \omega_i$, where m_i is the number of managers, l_i denotes total employment and ω_i is firm i ’s employment share in industry s at time t , i.e. $\omega_i = l_i \left(\sum_{i=1}^{N_{s,t}} l_i\right)^{-1}$. “Managerial expenditure share” is the employment-weighted managerial share in sector s at time t , i.e. $\sum_{i=1}^{N_{s,t}} \chi_i \omega_i$, where χ_i is firm i ’s expenditure share on managers (managerial compensation over the total wagebill). $\ln(\frac{k}{l})$ is the average of log capital-intensity, i.e. $\ln(\frac{k}{l}) = \frac{1}{N_{s,t}} \sum_{i=1}^{N_{s,t}} \ln\left(\frac{k_i}{l_i}\right)$.

Table 7: Managerial Intensity and Employment Share of Low Productivity Units

Finally, consider Table 7, where I report the results of (27) when I use the within-sector employment share of low productive units as the dependent variable. In particular, I report the results when I use the employment share of firms in the lowest quartile of the within-sector productivity distribution. The theory predicts a negative coefficient as sectors with a large managerial intensity are sectors with a large comparative advantage in managerial technologies (q) so that firms in those sectors are both more productive and have a large within-sector employment share. Table 7 shows that this prediction is borne out in the data as all the coefficient on the managerial intensity measures are negative and strongly significant. All in all, I interpret the results reported in Tables 5, 6 and 7 to be consistent with theory. A graphical representation of these results is also contained in Figure 2, where I depict the correlation between the sectoral managerial intensity and the three different dependent variables. Note that all the plots depict the residual correlation, i.e. after controlling for sectors capital intensity and a full set of year fixed effects. Hence, Figure 2 corresponds to column 3 in the respective Tables.

In the Appendix I report various robustness checks for these results. In particular, I consider the same implications, when I aggregate sectors at the three digit level. All results are qualitatively similar. I also report the results when I measure productivity conditional on the firm-level capital-intensity. Specifically, instead of using the raw data on labor productivity to calculate the sectoral mean productivity and employment share of low productivity units, I base all the calculations on a the residuals from a regression of log productivity on log capital-intensity on the firm level. This should at least address some worries that the capital-intensity on the sectoral level is too coarse a proxy to control for any correlation between managerial intensity, capital intensity and productivity

in the data I observe $\frac{m}{l} + \varepsilon$. Those sectors, where large firms happen to have a high ε will have both a high average managerial intensity and a large employment share in managerial intensive firms. This should be taken into account when interpreting the results in Table 6.



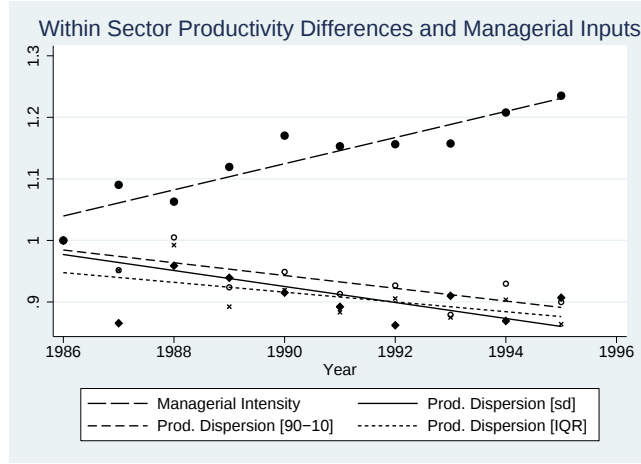
Notes: The figure depicts the scatter plot between an industry's average managerial intensity and its average productivity (top left), its within-industry employment share in managerial-intensive firms (top right) and its within-industry employment share in low-productivity firms. All plots are residual correlations after taking out a set of year fixed effects and the sector's average capital intensity. See Tables 5, 6 and 7 for details of the construction.

Figure 2: Cross-Sectional Predictions: Correlates of Managerial Intensity

at the firm level. As seen in the Appendix, all the results reported here hold true when I consider that productivity measure.

4.2 Time-Series Implications

I now turn to the time series prediction of the theory. In particular, I want to think of the time-series variation as being generated by a secular increase in managerial supplies in the Chilean economy between 1986 and 1996. When I look at the plant level data, I indeed find a steady increase of the aggregate managerial employment over the time period. The evolution of the managerial intensity is depicted in the Figure 3 below. The upward sloping line is the average managerial employment share relative to its level in 1986. The positive is apparent: In the decade following 1986, the managerial employment share increases by roughly 20%. According to the model, this increase in managerial supplies will reduce within-sector differences as less and less sectors of the economy will be characterized by a coexistence of heterogeneous technologies. In the lower part of Figure 3 I plot the evolution of three different measures of within-industry productivity differences. In particular, I consider the standard deviation, the inter-quartile range and the 90-10 difference in log labor productivity within industries. Each annual observation is a weighted average of the respective dispersion measure, where I weigh different industries by their respective value added share. Figure 3 shows that there is a strong negative relationship whereby each dispersion measure decreases by roughly 10% during the decade following 1986. In Table 8 I report the correlation between the aggregate managerial intensity and the respective dispersion measures. As can be expected from Figure 3 there is a strong negative correlation. Despite there being only 10 observations, I also include the average capital-intensity of the Chilean economy in the regression. During the time period studied, the firms in my sample not only increased their managerial-intensity but also experienced capital-deepening. At the annual data, this generates also a negative correlation



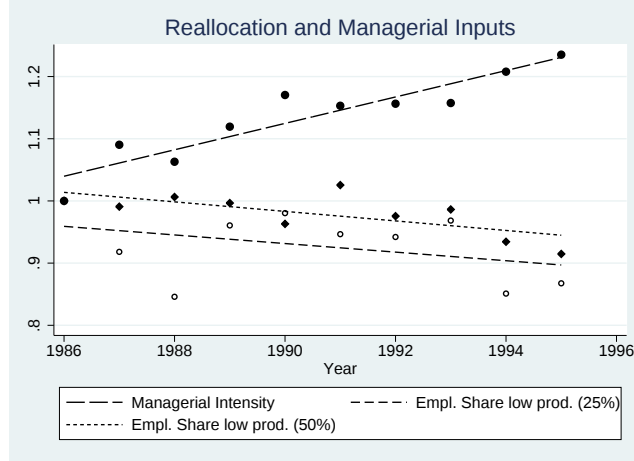
Notes: The figure depicts the time series evolution of the average managerial intensity and different measures of within-industry dispersion of log labor productivity. In particular, the employment-weighted managerial share in sector s is $\sum_{i=1}^{N_{s,t}} \left(\frac{m_i}{l_i} \right) \omega_i$, where m_i is the number of managers, l_i denotes total employment and ω_i is firm i 's employment share in industry s at time t , i.e. $\omega_i = l_i \left(\sum_{i=1}^{N_{s,t}} l_i \right)^{-1}$. To aggregate to the year level, I weigh industries by their share of value added.

Figure 3: Time-Series Predictions: Managerial deepening and productivity differences

Dep. Variable	Within-industry productivity differences							
	Standard deviation			IQR		90-10		
Managerial intensity	-3.433** (0.491)	-3.670** (1.226)	-3.106** (1.226)	-5.631** (2.126)	-6.346** (1.266)	-5.703 (3.165)		
Capital intensity	-0.153** (0.0455)	0.0141 (0.0759)	-0.106 (0.0819)	0.150 (0.136)	-0.298** (0.104)	-0.0382 (0.197)		
N	10	10	10	10	10	10	10	10
R^2	0.751	0.551	0.752	0.451	0.195	0.542	0.608	0.495

Notes: Robust standard errors are shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. A sector is a 4-digit industry. The dependent variables are three different measures of the within-industry dispersion of log labor productivity ξ (see (6)). “Managerial intensity” is the average managerial intensity in year t . “Capital intensity” is the average log capital-labor ratio in year t . In particular, the employment-weighted managerial share in sector s is $\sum_{i=1}^{N_{s,t}} \left(\frac{m_i}{l_i} \right) \omega_i$, where m_i is the number of managers, l_i denotes total employment and ω_i is firm i 's employment share in industry s at time t , i.e. $\omega_i = l_i \left(\sum_{i=1}^{N_{s,t}} l_i \right)^{-1}$. The capital intensity is calculated analogously with $\ln \left(\frac{k}{l} \right)$ instead of $\frac{m}{l}$. To aggregate to the year level, I weigh industries by their share of value added.

Table 8: Managerial deepening and productivity differences



Notes: The figure depicts the time series evolution of the average managerial intensity and two measures of the within-industry employment share in low-productivity firms. In particular, the employment-weighted managerial share in sector s is $\sum_{i=1}^{N_{s,t}} \left(\frac{m_i}{l_i} \right) \omega_i$, where m_i is the number of managers, l_i denotes total employment and ω_i is firm i 's employment share in industry s at time t , i.e. $\omega_i = l_i \left(\sum_{i=1}^{N_{s,t}} l_i \right)^{-1}$. The respective employment-share series refer to the within-industry employment share of firms with labor productivity below the 25% quantile and the median respectively. To aggregate to the year level, I weigh industries by their share of value added.

Figure 4: Time-Series Predictions: Managerial deepening and employment in low-productivity firms

between the dispersion measures and the capital-intensity of the Chilean economy. When I include both the managerial and the capital intensity in the regression, the coefficient on the managerial intensity remains significantly negative, while the capital-intensity loses its significance. However, I want to stress that each regression has only ten observations, so that I do not want to put too much weight on those results.

Figure 4 and Table 9 repeat this exercise with the time-series behavior of the within-industry aggregate employment share of low-productivity firms. In particular, I consider the employment share of 10%, 25% and 50% least productive firms within each industry. The theory predicts that an increase in managerial supplies in the aggregate should decrease all these employment shares as resources are reallocated towards managerial techniques, which have a relatively high labor productivity. The results reported in Figure 4 and Table 9 are not conclusive. While the trend is decreasing and all coefficients have the expected sign, the standard errors are large. In fact, Figure 4 shows that for the employment share in the lowest 25% of firms, the annual observations fluctuate widely around the estimated regression line. The pattern is more precise in case of the media.

I finally turn the more subtle implication of the theory. While the robust prediction is a negative correlation between managerial supplies and productivity dispersion, the model implies that the replacement process will have a clear “pecking-order”. As managerial supplies are allocated according to the industry’s comparative advantage, it will be the (initially) high average product industries where traditional supplies are replaced first. Hence, the model suggests, that within-sector productivity differences should vanish particularly fast in those industries that had a high labor productivity in 1986. To test for this pattern, I allow the evolution of productivity dispersion to differ by the industry’s initial labor productivity. In particular, I split the sample of sectors in

Dep. Variable	Within-industry employment share of low productivity firms							
	bottom 10%		bottom 25%		bottom 50%			
Managerial intensity	-0.115 (0.226)	-0.152 (0.440)	-0.336 (0.326)	-0.0333 (0.685)	-0.875** (0.324)			-0.204 (0.479)
Capital intensity	-0.00473 (0.0141)	0.00220 (0.0282)	-0.0195 (0.0160)	-0.0180 (0.0313)		-0.0491** (0.0115)		-0.0399 (0.0269)
N	10	10	10	10	10	10	10	10
R^2	0.023	0.015	0.024	0.123	0.153	0.500	0.584	0.591

Notes: Robust standard errors are shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. A sector is a 4-digit industry. The dependent variables are three different measures of the within-industry employment share of low-productivity firms. I consider the bottom 10%, 25% and 50% of firms as measured by their labor productivity within their industry. “Managerial intensity” is the average managerial intensity in year t . “Capital intensity” is the average log capital-labor ratio in year t . In particular, the employment-weighted managerial share in sector s is $\sum_{i=1}^{N_{s,t}} \left(\frac{m_i}{l_i} \right) \omega_i$, where m_i is the number of managers, l_i denotes total employment and ω_i is firm i ’s employment share in industry s at time t , i.e. $\omega_i = l_i \left(\sum_{i=1}^{N_{s,t}} l_i \right)^{-1}$. The capital intensity is calculated analogously with $\ln \left(\frac{k}{l} \right)$ instead of $\frac{m}{l}$. To aggregate to the year level, I weigh industries by their share of value added.

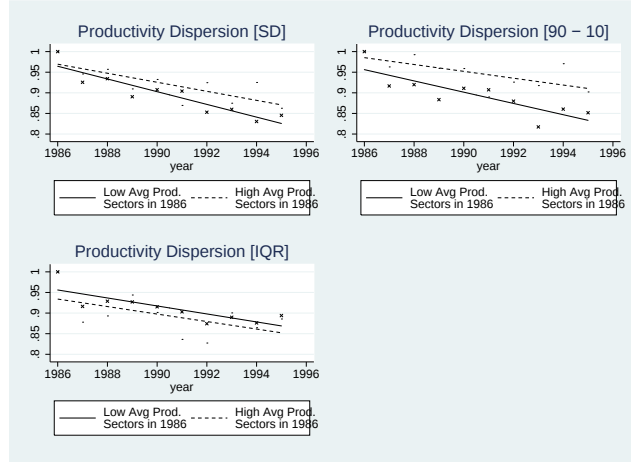
Table 9: Managerial deepening and employment share of low-productivity firms

1986 in high (above media) and low (below median) productivity industries and track the evolution of within-sector productivity differences in those two groups over time. The results are depicted in Figure 5. It is clearly seen that this implication of the theory is not borne out. While I am not able to reject that the pattern is exactly identical across the two groups, it is if anything the initially low productivity sectors, where productivity dispersion falls the most in the time period after 1986.

5 Revisiting the management production function

Let me now briefly come back to the basic setup of the heterogeneous techniques in this economy. As explained in Section 3.1, the crucial modelling decision is which technology to “endow” with inframarginal rents. In the main model of this paper, I assumed that these rents accrue to the managerial technology. In fact the positive correlation between managerial intensity and revenue labor productivity within industries reported in Table 4 suggested that this is the empirically relevant case.

This conclusion however depends on the particular measure of labor productivity used. In the main analysis of this paper, I followed the literature to measure labor productivity as value added per adjusted labor input. In the model with two types of workers, this measure was given by $\xi = \frac{py}{l + \frac{w_L}{w_M} m}$. In the model where all firms face the same factor prices, this measure is of course equivalent to measuring labor productivity as value added per dollar spent on labor inputs, namely $\xi^W = \frac{py}{w_L l + w_M m} = w_L \xi$. Now suppose we were to take ξ^W instead of ξ and ask the question which technology earns inframarginal rents. The results of running exactly the same regression as in Table 4 using $\ln(\xi^W)$ as the dependent variable are contained in Table 10. It is clearly seen that the results are very different. As soon as I control for firms’ capital-intensity in the regression, the coefficient on the managerial intensity is *negative*. Hence, from Table 10 we could conclude that it is the managerial technology that earns lower inframarginal rents.



Notes: The figure depicts the time series evolution of three different measures of within-industry dispersion of log labor productivity for two groups of firms respectively. High productivity firms are firms with a labor-productivity exceeding the median in labor productivity in their industry in 1986. Low productivity firms are the complement. An industry is a 4-digit sector. To aggregate to the year level, I weigh industries by their share of value added.

Figure 5: Time-Series Predictions: Managerial deepening and productivity dispersion in high and low productivity industries

Dep. Variable	log of Value added over wagebill [$\ln(\frac{py}{w_L l + w_M m})$]					
Managerial intensity	0.143** (0.0288)	-0.00423 (0.0285)	-0.0552** (0.0257)			
$\ln(\frac{k}{l})$		0.112** (0.00301)	0.0997** (0.00271)	0.114** (0.00304)	0.101** (0.00274)	
Managerial expenditure share				0.0447** (0.0219)	-0.0878** (0.0218)	-0.120** (0.0198)
N	41693	41693	40669	41693	41693	40669
R^2	0.145	0.181	0.208	0.145	0.181	0.209

Notes: Robust standard errors are shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. All regression contain year and sector fixed effects. A sector is a 4-digit industry. The dependent variable is log of firms' value added over their wagebill. In the model this refers to $\frac{py}{w_L l + w_M m}$. "Managerial intensity" is the firm's managerial employment share $\frac{m_i}{l_i}$, where m_i is the number of managers and l_i denotes total employment. $\ln(\frac{k}{l})$ denotes the log of firm's capital intensity. "Managerial expenditure share" is given by firm's managerial compensation over the total wagebill.

Table 10: Managerial intensity and inframarginal rents revisited

Starting from this evidence, we might have assumed production technologies of the form

$$\begin{aligned} y_T(q) &= A(q) l^\theta \\ y_M(q) &= A(q) q \min\{l, m^\gamma\}, \end{aligned}$$

were $\theta < \gamma$. This specification has intuitive appeal as it stresses that the managerial technology has a higher span-of-control (compared to traditional producers). This specification however, will have very different *equilibrium* implications when viewed through the mechanism of the model. Suppose we were to take that production structure and impose a free entry condition of the form $\pi_T(q) = w_L$. This is the standard indifference condition in the original Lucas (1978) span of control model. This model will have a very similar structure as the model presented above. In particular, there will again be a cutoff productivity $\hat{q}$ such that traditional producers will only be active in industries with $q < \hat{q}$. Furthermore, this cutoff is again decreasing in the aggregate managerial supply. Hence, this alternative model has the same implication that managerial deepening will reduce the within-sector productivity dispersion. However, the cross-sectional implications are very different. In particular, this model will imply a *negative* correlation between the average managerial intensity and the average labor productivity and a *positive* correlation between the managerial intensity and the employment share in the low-productivity units. The reason is of course that it will now be the managerial firms that have low inframarginal rents so that their labor productivity is low (despite the physical productivity q being high). These implications are at odds with the results reported above: when I perform the cross-sectional regressions of Section 4 using the measure $\xi^W = \frac{\text{value added}}{\text{wagebill}}$, all the results reported in Tables 5, 6 and 7 are confirmed.

6 Conclusion

Even within narrowly defined industries, labor productivity is not equalized across different producers. While this is also true in developed economies, this pattern is much more severe in poor countries. In this paper I try to give an explanation for this empirical regularity. In particular, I interpret the observed differences in labor productivity across producers as stemming from firms using different technologies. If different techniques are profitable in the marketplace, the average product of factors can vary across producers despite all marginal products being equalized. I then argue that there is less productivity dispersion in rich countries because the set of economically viable technologies shrinks as the economy develops. In particular, efficient firms expand and replace small-scale producers. This reduces the within-sector productivity dispersion as industries become more and more dominated by a single best-practice technology. My mechanism stresses the role of managerial inputs. In particular, I consider a model where managers are essential to increase firms' scale of production. In managerial-scarce economies, inefficient producers survive because their competitors do not have the necessary inputs to produce at large enough a scale to replace them. As managerial inputs accumulate, these interstices of surviving small-scale producers vanish as managerial-intensive firms increase their scale of production. I formalize this intuition in a very tractable multi-sector general equilibrium and test its cross-sectional and time-series implications using firm-level panel data from Chile. While the main implications of the model are borne out in the data, the evidence on how to best introduce management in firms' production function is less decisive.

References

- AUTOR, D., AND L. F. KATZ (1999): “Changes in the Wage Structure and Earnings Inequality,” in *Handbook of Labor Economics*, vol. 3A. Elsevier.
- BANERJEE, A., AND B. MOLL (2010): “Why Does Misallocation Persist?,” *American Economic Journal: Macroeconomics*, 2(1), p. 189–206.
- BARTELSMAN, E. J., J. C. HALTIWANGER, AND S. SCARPETTA (2009): “Cross-Country Differences in Productivity: The Role of Allocation and Selection,” Working paper 15490, NBER.
- BLOOM, N., B. EIFERT, D. MCKENZIE, A. MAHAJAN, AND J. ROBERTS (2010): “Does management matter: Evidence from India,” .
- BLOOM, N., AND J. V. REENEN (2007): “Measuring and Explaining Management Practices Across Firms and Countries,” *Quarterly Journal of Economics*.
- (2010): “Why Do Management Practices Differ across Firms and Countries?,” *Journal of Economic Perspectives*, 24(1), 203–224.
- BRUHN, M., D. KARLAN, AND A. SCHOAR (2010): “What Capital is Missing in Developing Countries?,” *American Economic Review (P&P)*.
- BUERA, F., J. KABOSKI, AND Y. SHIN (2011): “Finance and Development: A Tale of Two Sectors,” *American Economic Review*, 101(5), 1964–2002.
- CASELLI, F. (2005): “Accounting for Cross-Country Income Differences,” in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf, vol. 1A, chap. 9, pp. 679–741. Elsevier B.V.
- CHANDLER, A. D. J. (1977): *The Visible Hand. The Managerial Revolution in American Business*. The Belknap Press of Harvard University Press, Cambridge, MA.
- COLLARD-WEXLER, A., J. ASKER, AND J. DE LOECKER (2011): “Productivity Volatility and the Misallocation of Resources in Developing Economies,” Discussion paper, NBER Working Paper 17175.
- DOMS, M., T. DUNNE, AND K. R. TROSKE (1997): “Workers, Wages, and Technology,” *Quarterly Journal of Economics*, 112, 253–290.
- FOSTER, J., J. C. HALTIWANGER, AND C. SYVERSON (2008): “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?,” *American Economic Review*, 98(1), 394–425.
- (2011): “The Slow Growth of New Plants: Learning about Demand,” Discussion paper, Working Paper.
- GOLDIN, C., AND L. KATZ (1999): “The Returns to Skill in the United States Across the Twentieth Century,” Discussion paper, NBER Working Paper 7126.
- HSIEH, C.-T., AND P. KLENOW (2009): “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics*, 124(2), 771–807.

- (2011): “The Life-Cycle of Plants in India and Mexico,” Discussion paper, Working Paper.
- LEVINSOHN, J., AND A. PETRIN (2003): “Estimating Production Functions Using Inputs to Control for Unobservables,” *Review of Economic Studies*, 70, 317–341.
- LEWIS, A. (1954): “Economic Development with Unlimited Supplies of Labor,” *The Manchester School*, 22(2), 131–191.
- LUCAS, R. E. (1978): “On the Size Distribution of Business Firms,” *The Bell Journal of Economics*, 9(2), 508–523.
- MELITZ, M. (2003): “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 71(6), 1695–1725.
- MIDRIGAN, V., AND D. Y. XU (2010): “Finance and Misallocation: Evidence From Plant-Level Data,” NBER Working Paper 15647.
- MOLL, B. (2010): “Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?,” Working Paper.
- MURPHY, K. M., AND F. WELCH (1993): “Occupational Change and the Demand for Skill, 1940 - 1990,” *American Economic Review (P&P)*, 83(2), 122–126.
- PAVCNIK, N. (2002): “Trade Liberalization, Exit and Productivity Improvements: Evidence from Chilean Plants,” *Review of Economic Studies*, 69(1), 245–276.
- PENROSE, E. T. (1959): *Theory of the Growth of Firms*. J. Wiley & Sons, New York.
- PETERS, M. (2010): “Heterogeneous Mark-Ups and Endogenous Misallocation,” .
- RESTUCCIA, D., AND R. ROGERSON (2008): “Policy Distortions and Aggregate Productivity with Heterogenous Establishments,” *Review of Economic Dynamics*, 11(4), 707–720.
- ROBINSON, A. (1934): “The Problem of Management and the Size of Firms,” *The Economic Journal*, 44(174), 242–257.
- SYVERSON, C. (2011): “What determines productivity?,” *Journal of Economic Literature*.
- ZIEBARTH, N. L. (2011): “Are China and India Backward? Evidence from the 19th century US Census of Manufacturers,” Discussion paper, Working Paper.

7 Appendix

7.1 The Demand for Production Workers

To derive (24), note that

$$\begin{aligned}
\frac{L}{\mu} &= \int [l_M(q, \hat{q}, \psi) + l_T(q, \hat{q}, \psi)] dG_q(q) \\
&= \int l_M(q, \hat{q}, \psi) dG(q) + \int_{q_L}^{\hat{q}} (\hat{q} l_M(\hat{q}, \hat{q}, \psi) - q l_M(q, \hat{q}, \psi)) dG(q) \\
&= \psi^{\frac{\gamma}{1-\gamma}} \gamma^{\frac{\gamma}{1-\gamma}} \left[\int z(q, \hat{q})^{\frac{\gamma}{1-\gamma}} dG(q) + G(\hat{q}) \hat{q} z(\hat{q}, \hat{q})^{\frac{\gamma}{1-\gamma}} - \int_{q_L}^{\hat{q}} q z(q, \hat{q})^{\frac{\gamma}{1-\gamma}} dG(q) \right] \\
&= \psi^{\frac{\gamma}{1-\gamma}} \gamma^{\frac{\gamma}{1-\gamma}} \left[(1 - G(\hat{q})) z(\hat{q}, \hat{q})^{\frac{\gamma}{1-\gamma}} + G(\hat{q}) \hat{q} z(\hat{q}, \hat{q})^{\frac{\gamma}{1-\gamma}} - \int_{q_L}^{\hat{q}} (q - 1) z(q, \hat{q})^{\frac{\gamma}{1-\gamma}} dG(q) \right] \\
&= \psi^{\frac{\gamma}{1-\gamma}} \gamma^{\frac{\gamma}{1-\gamma}} \left[(1 - G(\hat{q})) z(\hat{q}, \hat{q})^{\frac{\gamma}{1-\gamma}} + G(\hat{q}) \hat{q} z(\hat{q}, \hat{q})^{\frac{\gamma}{1-\gamma}} - \int_{q_L}^{\hat{q}} z(q, \hat{q})^{\frac{1}{1-\gamma}} dG(q) \right] \\
&= \psi^{\frac{\gamma}{1-\gamma}} \gamma^{\frac{\gamma}{1-\gamma}} \left[(1 - G(\hat{q})) z(\hat{q}, \hat{q})^{\frac{\gamma}{1-\gamma}} + G(\hat{q}) \hat{q} z(\hat{q}, \hat{q})^{\frac{\gamma}{1-\gamma}} + (1 - G(\hat{q})) z(\hat{q}, \hat{q})^{\frac{1}{1-\gamma}} - \int z(q, \hat{q})^{\frac{1}{1-\gamma}} dG(q) \right] \\
&= \psi^{\frac{\gamma}{1-\gamma}} \gamma^{\frac{\gamma}{1-\gamma}} \left[z(\hat{q}, \hat{q})^{\frac{\gamma}{1-\gamma}} ((1 - G(\hat{q})) + G(\hat{q}) \hat{q} + (1 - G(\hat{q})) (\hat{q} - 1)) - \int z(q, \hat{q})^{\frac{1}{1-\gamma}} dG(q) \right] \\
&= \psi^{\frac{\gamma}{1-\gamma}} \gamma^{\frac{\gamma}{1-\gamma}} \left[(\hat{q} - 1)^{\frac{\gamma}{1-\gamma}} \hat{q} - \int z(q, \hat{q})^{\frac{1}{1-\gamma}} dG(q) \right] \\
&= \psi^{\frac{\gamma}{1-\gamma}} \gamma^{\frac{\gamma}{1-\gamma}} \left[\frac{z(\hat{q}, \hat{q}) + 1}{z(\hat{q}, \hat{q})} z(\hat{q}, \hat{q})^{\frac{1}{1-\gamma}} - \int z(q, \hat{q})^{\frac{1}{1-\gamma}} dG(q) \right].
\end{aligned}$$

7.2 Proof of Proposition 4

To show that there is a unique equilibrium, we have to establish that there is a unique tuple $(\hat{q}, \psi)$ that solves

$$\begin{aligned}
\frac{M}{\mu} &= \gamma^{\frac{1}{1-\gamma}} \psi^{\frac{1}{1-\gamma}} \Psi(\hat{q}) \\
\frac{L}{\mu} &= \gamma^{\frac{\gamma}{1-\gamma}} \psi^{\frac{\gamma}{1-\gamma}} (D(\hat{q}) - \Psi(\hat{q})),
\end{aligned}$$

where

$$\begin{aligned}
\Psi(\hat{q}) &= \int_q z(q, \hat{q})^{\frac{1}{1-\gamma}} dG_q(q) \\
D(\hat{q}) &= \left(\frac{1 + z(\hat{q}, \hat{q})}{z(\hat{q}, \hat{q})} \right) z(\hat{q}, \hat{q})^{\frac{1}{1-\gamma}},
\end{aligned}$$

with $z(q, \hat{q}) = \min\{q, \hat{q}\} - 1$. To do so, I will show that there is a unique $\hat{q}$ that solves

$$\frac{\left(\frac{M}{\mu}\right)^\gamma}{\left(\frac{L}{\mu}\right)} = \frac{\Psi(\hat{q})^\gamma}{D(\hat{q}) - \Psi(\hat{q})} \equiv Q(\hat{q}).$$

First of all note that

$$\begin{aligned} Q(\hat{q} = q_L) &= \frac{(q_L - 1)^{\frac{\gamma}{1-\gamma}}}{q_L (q_L - 1)^{\frac{\gamma}{1-\gamma}} - (q_L - 1)^{\frac{1}{1-\gamma}}} = 1 \\ \lim_{\hat{q} \rightarrow \infty} Q(\hat{q}) &= \lim_{\hat{q} \rightarrow \infty} \frac{\left(\int (q - 1)^{\frac{1}{1-\gamma}} dG_q\right)^\gamma}{\hat{q}(\hat{q} - 1)^{\frac{\gamma}{1-\gamma}} - \left(\int (q - 1)^{\frac{1}{1-\gamma}} dG_q\right)} \\ &= \lim_{\hat{q} \rightarrow \infty} \frac{\left(E\left[(q - 1)^{\frac{1}{1-\gamma}}\right]\right)^\gamma}{\hat{q}(\hat{q} - 1)^{\frac{\gamma}{1-\gamma}} - E\left[(q - 1)^{\frac{1}{1-\gamma}}\right]} = 0 \end{aligned}$$

To see that $Q(\hat{q})$ is strictly decreasing, note that

$$\begin{aligned} \frac{\partial Q(\hat{q})}{\partial \hat{q}} &= \frac{\gamma \Psi(\hat{q})^{\gamma-1} \Psi'(\hat{q}) (D(\hat{q}) - \Psi(\hat{q})) - \Psi(\hat{q})^\gamma (D'(\hat{q}) - \Psi'(\hat{q}))}{(D(\hat{q}) - \Psi(\hat{q}))^2} \\ &= \frac{\Psi(\hat{q})^{\gamma-1}}{(D(\hat{q}) - \Psi(\hat{q}))^2} (\gamma \Psi'(\hat{q}) (D(\hat{q}) - \Psi(\hat{q})) - \Psi(\hat{q}) (D'(\hat{q}) - \Psi'(\hat{q}))) \\ &= \frac{\Psi(\hat{q})^{\gamma-1}}{(D(\hat{q}) - \Psi(\hat{q}))^2} (\gamma \Psi'(\hat{q}) D(\hat{q}) - \Psi(\hat{q}) D'(\hat{q}) + (1 - \gamma) \Psi(\hat{q}) \Psi'(\hat{q})). \end{aligned}$$

Hence, it is sufficient to show that

$$\Psi(\hat{q}) D'(\hat{q}) - (1 - \gamma) \Psi(\hat{q}) \Psi'(\hat{q}) - \gamma \Psi'(\hat{q}) D(\hat{q}) > 0. \quad (28)$$

Now note that

$$\begin{aligned} D'(\hat{q}) &= (\hat{q} - 1)^{\frac{\gamma}{1-\gamma}} + \frac{\gamma}{1-\gamma} \hat{q} (\hat{q} - 1)^{\frac{\gamma}{1-\gamma} - 1} = (\hat{q} - 1)^{\frac{\gamma}{1-\gamma}} \left(1 + \frac{\gamma}{1-\gamma} \frac{\hat{q}}{\hat{q} - 1}\right) \\ \Psi'(\hat{q}) &= (1 - G_q(\hat{q})) \frac{1}{1-\gamma} (\hat{q} - 1)^{\frac{\gamma}{1-\gamma}} \end{aligned}$$

so that

$$D'(\hat{q}) = \frac{(1 - \gamma) \Psi'(\hat{q})}{1 - G_q(\hat{q})} \left(1 + \frac{\gamma}{1-\gamma} \frac{\hat{q}}{\hat{q} - 1}\right).$$

Substituting this in (28) yields

$$\begin{aligned}
& \Psi(\hat{q}) \frac{(1-\gamma)\Psi'(\hat{q})}{1-G_q(\hat{q})} \left(1 + \frac{\gamma}{1-\gamma} \frac{\hat{q}}{\hat{q}-1}\right) - (1-\gamma) \Psi(\hat{q}) \Psi'(\hat{q}) - \gamma \Psi'(\hat{q}) D(\hat{q}) \\
= & \frac{\Psi(\hat{q})\Psi'(\hat{q})}{1-G_q(\hat{q})} \left[(1-\gamma) \left(1 + \frac{\gamma}{1-\gamma} \frac{\hat{q}}{\hat{q}-1}\right) - (1-\gamma) (1-G_q(\hat{q})) - \gamma \frac{D(\hat{q})(1-G_q(\hat{q}))}{\Psi(\hat{q})} \right] \\
= & \frac{\Psi(\hat{q})\Psi'(\hat{q})}{1-G_q(\hat{q})} \left[(1-\gamma) \left(1 + \frac{\gamma}{1-\gamma} \frac{\hat{q}}{\hat{q}-1}\right) - (1-\gamma) (1-G_q(\hat{q})) - \gamma \frac{\hat{q}}{\hat{q}-1} \frac{(1-\gamma)^{\frac{1}{1-\gamma}} (1-G_q(\hat{q}))}{\int_{q_L}^{\hat{q}} (q-1)^{\frac{1}{1-\gamma}} dG_q(q) + (\hat{q}-1)^{\frac{1}{1-\gamma}} (1-G_q(\hat{q}))} \right] \\
> & \frac{\Psi(\hat{q})\Psi'(\hat{q})}{1-G_q(\hat{q})} \left[(1-\gamma) + \gamma \frac{\hat{q}}{\hat{q}-1} - (1-\gamma) (1-G_q(\hat{q})) - \gamma \frac{\hat{q}}{\hat{q}-1} \right] \\
= & \frac{\Psi(\hat{q})\Psi'(\hat{q})}{1-G_q(\hat{q})} (1-\gamma) G_q(\hat{q}) \\
> & 0,
\end{aligned}$$

which proves the proposition.

7.3 Proof of Proposition 3

If the managerial technology is given by $f(l, m) = \left(\alpha l^{\frac{\sigma-1}{\sigma}} + (1-\alpha) m^{\gamma \frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$, the two optimality conditions are given by

$$\begin{aligned}
w_L &= p(q) A(q) q \left(\alpha l^{\frac{\sigma-1}{\sigma}} + (1-\alpha) m^{\gamma \frac{\sigma-1}{\sigma}} \right)^{\frac{1}{\sigma-1}} \alpha l^{\frac{-1}{\sigma}} \\
w_M &= p(q) A(q) q \left(\alpha l^{\frac{\sigma-1}{\sigma}} + (1-\alpha) m^{\gamma \frac{\sigma-1}{\sigma}} \right)^{\frac{1}{\sigma-1}} \gamma (1-\alpha) m^{\frac{\sigma-1}{\sigma} \gamma - 1}.
\end{aligned} \tag{29}$$

These imply that the managerial intensity is given by

$$\frac{m}{l} = \left(\gamma \frac{1-\alpha}{\alpha} \psi \right)^{\frac{\sigma}{1-(1-\sigma)(1-\gamma)}} l^{\frac{(1-\sigma)(1-\gamma)}{1-(1-\sigma)(1-\gamma)}}, \tag{30}$$

and that

$$\left(\frac{m^\gamma}{l} \right)^{\frac{\sigma-1}{\sigma}} = \left(\frac{\psi^{\frac{\sigma\gamma}{\sigma+(1-\sigma)\gamma}} l^{\frac{\gamma}{\sigma+(1-\sigma)\gamma}}}{l} \right)^{\frac{\sigma-1}{\sigma}} = \frac{\alpha}{\gamma(1-\alpha)} \frac{1}{\psi} \frac{m}{l}.$$

(29) can therefore be written as

$$\begin{aligned}
1 &= \frac{p(q) A(q) q}{w_L} \alpha \left(\alpha + (1-\alpha) \left(\frac{m^\gamma}{l} \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{1}{\sigma-1}} \\
&= \min\{q, \hat{q}\} \alpha^{\frac{\sigma}{\sigma-1}} \left(1 + \frac{1}{\gamma \psi} \frac{m}{l} \right)^{\frac{1}{\sigma-1}},
\end{aligned}$$

as prices are still given by (17). This implies that

$$\frac{m}{l} = \gamma \psi \left(\left(\frac{1}{\alpha} \right)^\sigma (\min\{q, \hat{q}\})^{1-\sigma} - 1 \right) = \gamma \psi z^{CES}(q, \hat{q}).$$

Using (30), we get that

$$\gamma \psi z^{CES}(q, \hat{q}) = \left(\gamma \frac{1-\alpha}{\alpha} \psi \right)^{\frac{\sigma}{1-(1-\sigma)(1-\gamma)}} l^{\frac{(1-\sigma)(1-\gamma)}{1-(1-\sigma)(1-\gamma)}},$$

which implies that

$$l^{CES}(q, \hat{q}, \psi) = \left(\frac{\alpha}{1-\alpha} \right)^{\frac{\sigma}{(1-\sigma)(1-\gamma)}} (\gamma \psi)^{\frac{\gamma}{1-\gamma}} z^{CES}(q, \hat{q})^{\frac{1-(1-\sigma)(1-\gamma)}{(1-\sigma)(1-\gamma)}}$$

and

$$m^{CES}(q, \hat{q}, \psi) = \left(\frac{\alpha}{1-\alpha} \right)^{\frac{\sigma}{(1-\sigma)(1-\gamma)}} (\gamma \psi)^{\frac{1}{1-\gamma}} z^{CES}(q, \hat{q})^{\frac{1}{(1-\sigma)(1-\gamma)}}.$$

To derive the expression for labor productivity, note that

$$\begin{aligned} \xi_M^{CES}(q, \hat{q}) &= \frac{p(q) A(q) q f(l, m)}{l + \frac{w_M}{w_L} m} = w_L \frac{\frac{p(q) A(q) q}{w_L} \frac{f(l, m)}{l}}{1 + \frac{w_M}{w_L} \frac{m}{l}} \\ &= w_L \frac{\min\{q, \hat{q}\} \left(\alpha + (1-\alpha) \left(\frac{m}{l} \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}}{1 + \frac{1}{\psi} \frac{m}{l}} \\ &= w_L \frac{\min\{q, \hat{q}\} \alpha^{\frac{\sigma}{\sigma-1}} \left(1 + \frac{1}{\gamma} \frac{1}{\psi} \frac{m}{l} \right)^{\frac{\sigma}{\sigma-1}}}{1 + \frac{1}{\psi} \frac{m}{l}} \\ &= w_L \frac{\min\{q, \hat{q}\} \alpha^{\frac{\sigma}{\sigma-1}} (1 + z^{CES}(q, \hat{q}))^{\frac{\sigma}{\sigma-1}}}{1 + \gamma z^{CES}(q, \hat{q})} \\ &= w_L \frac{\min\{q, \hat{q}\} \alpha^{\frac{\sigma}{\sigma-1}} \left(\frac{1}{\alpha} \right)^{\sigma \frac{\sigma}{\sigma-1}} (\min\{q, \hat{q}\})^{-\sigma}}{1 + \gamma z^{CES}(q, \hat{q})} \\ &= w_L \frac{\left(\frac{1}{\alpha} \right)^{\sigma} (\min\{q, \hat{q}\})^{1-\sigma}}{1 + \gamma z^{CES}(q, \hat{q})} \\ &= w_L \frac{z^{CES}(q, \hat{q}) + 1}{1 + \gamma z^{CES}(q, \hat{q})} \\ &= w_L + w_L \left(\frac{(1-\gamma) z^{CES}(q, \hat{q})}{1 + \gamma z^{CES}(q, \hat{q})} \right). \end{aligned}$$

7.4 Robustness of empirical results

Dep Variable:	Sectoral Average Productivity					
Managerial intensity	2.653** (0.335)	0.413* (0.225)	1.022** (0.311)		1.022 (0.611)	
$\ln(\frac{k}{l})$		0.597** (0.0260)	0.533** (0.0338)	0.549** (0.0322)	0.533** (0.0650)	0.549** (0.0648)
Managerial expenditure share				0.657** (0.238)		0.657 (0.513)
N	254	254	254	254	254	254
R^2	0.199	0.824	0.768	0.763	0.768	0.763

Table 11: Managerial Intensity and Average Productivity (3digit)

Dep. Variable	Within-industry employment share of managerial-intensive firms					
Managerial intensity	0.637** (0.0713)	0.686** (0.0826)	0.693** (0.0976)		0.693** (0.285)	
$\ln(\frac{k}{l})$		-0.0132* (0.00746)	-0.0330** (0.00844)	-0.0195** (0.00900)	-0.0330 (0.0209)	-0.0195 (0.0225)
Managerial expenditure share				0.401** (0.0704)		0.401* (0.197)
N	254	254	254	254	254	254
R^2	0.311	0.324	0.359	0.212	0.359	0.212

Table 12: Managerial Intensity and Managerial Employment Share (3digit)

Dep. Variable	Within-industry employment share of low-productivity units					
Managerial intensity	-0.107** (0.0444)	-0.0734 (0.0511)	-0.197** (0.0550)		-0.197 (0.141)	
$\ln(\frac{k}{l})$		-0.00891 (0.00651)	0.00682 (0.00665)	0.00551 (0.00680)	0.00682 (0.0157)	0.00551 (0.0161)
Managerial expenditure share				-0.156** (0.0503)		-0.156 (0.134)
N	254	254	254	254	254	254
R^2	0.021	0.031	0.063	0.065	0.063	0.065

Table 13: Managerial Intensity and Employment Share of Low Productivity Units (3digit)

Dep. Variable	Average Residual Labor Productivity					
Managerial intensity	1.379** (0.115)	0.919** (0.118)	1.328** (0.218)		1.328** (0.421)	
$\ln(\frac{k}{l})$		0.164** (0.0173)	0.152** (0.0297)	0.134** (0.0295)	0.152** (0.0298)	0.134** (0.0426)
Managerial expenditure share				1.016** (0.179)		1.016** (0.367)
N	462	462	462	462	462	462
R^2	0.280	0.422	0.457	0.454	0.457	0.454

Table 14: Managerial Intensity and Average Productivity (Residual Productivity)

Dep. Variable	Within-industry employment share of low-residual-productivity firms					
Managerial intensity	-0.159** (0.0542)	-0.230** (0.0617)	-0.264** (0.0478)		-0.264** (0.120)	
$\ln(\frac{k}{l})$		0.0253** (0.00599)	0.0191** (0.00561)	0.0248** (0.00582)	0.0191 (0.0125)	0.0248* (0.0126)
Managerial expenditure share				-0.233** (0.0391)		-0.233** (0.0998)
N	462	462	462	462	462	462
R^2	0.029	0.064	0.086	0.103	0.086	0.103

Table 15: Managerial Intensity and Employment Share of Low Productivity Units (Residual Productivity)