Investment in skills, managerial quality and economic development

Mehrdad Esfahani*

Arizona State University

For the latest version, click here

August 13, 2018.

Abstract

I document that for a group of 38 countries ranging from low to high income: (1) the share of skilled managers is higher in richer countries, (2) the relative income of managers to non-managers is lower in richer countries, and (3) the relative income of skilled to unskilled individuals is lower in richer countries. In addition, the share of managers is lower in richer countries while the mean plant size is larger in richer countries. I explore these facts through the lens of a general equilibrium model of investment in skills and occupational choice. Countries differ in productivity level in production and the level of size-dependent distortions. I find that exogenous productivity differences alone can produce the above facts qualitatively, but size-dependent distortions are needed to account for these facts quantitatively and the output elasticity of productivity is 2.6.

JEL classification codes: E23, E24, J24, M11, O43, O47.

Keywords: Economic development, Cross-Country Income Differences, Managers, Management Practices, Skill Investment

*I am grateful to Gustavo Ventura, Bart Hobijn and Domenico Ferraro. The views expressed here are those of the author's and do not necessarily reflect the views of EuroStat, the European Commission or any national statistical authorities whose data has been used.

1 Introduction

It is well-established that the observed cross-country income differences are large. Development accounting methodology reveals that differences in total factor productivity (TFP) are crucial in understanding why some countries produce more than others.[1] The question then becomes, what determines productivity differences across countries?

There are multiple proposals for the roots of productivity differences such as measurement of the physical capital and its composition, measurement of the quality of human capital, and monopolistic barriers to technology adoption. There is a relatively recent empirical literature that highlights the role of management practices as a root of productivity differences.[2] Management practices vary considerably across countries and across firms within a country and they manifest themselves on the aggregate level by higher total factor productivity and output.

In this paper, I present novel evidence on educational attainment of managers, which I define to be managers based on occupational classification and self-employed, and its relation to output per worker. I first document that the share of managers with more than a high school degree among all managers is higher in richer countries. I refer to managers with more than a high school degree as skilled. For example, the share of skilled managers in the U.S. is 71%, while it is 17% in Mexico and 14% in Brazil. The correlation between the share of skilled managers and GDP per worker is 0.18. I do this using micro data for a set of low-, middle- and high-income countries.

I subsequently document that the relative income of managers to non-managers, which I call managerial premium, is smaller in richer countries. For example, the managerial premium in the U.S. is 1.4 while it is 1.5 in Mexico and 1.7 in South Africa. The correlation between managerial premium and GDP per worker is -0.4. To my knowledge, this is the first paper to document such a fact about relative income of managers across countries.[3] I also document that the share of managers among the working age population (age 15-64), which I refer to as managerial rate, is lower in richer countries. For instance, the share of managers in the U.S. is 22% while it is 42% in Mexico and 37% in Greece and the correlation between the managerial rate and GDP per worker is -0.2.

There is the well-established fact that the mean plant size is larger in richer countries. Bento and Restuccia (2017) document this fact using the survey data for the manufacturing sector and find that the correlation between the mean plant size and GDP per worker is 0.33. So richer countries have less managers, more skilled managers and those managers operate larger plants. Finally, I document that the relative income of skilled versus unskilled individuals in the working-age population, which I refer to as skill premium, is lower in richer countries. For example, the skill premium in the U.S. is 1.8 while

[1] Klenow and Rodriguez-Clare (1997), Prescott (1998), Hall and Jones (1999), Caselli (2005).

[2] See Bloom and Reenen (2011) and Bloom *et al.* (2011) among others.

[3] Guner *et al.* (2018) document that the relative income of managers versus non-managers grow faster over the life-cycle in richer countries. My study abstracts from this growth and document a different fact about the income levels rather than growth.

it is 4.0 in Mexico and 4.3 in Brazil. This finding is not new and Fernandez *et al.* (2005) documented a similar trend. What is new here is that the correlation between managerial premium and skill premium is 0.6. This shows that the above facts about skill segregation and relative income should be interpreted in a unified framework to better understand the income differences across countries. I develop such a framework in this paper.

I will interpret the differences in management practices, documented by the empirical literature, in my framework as th differences in "managerial quality" coming from differences in *selection* into management occupation, along the lines of Lucas (1978), and differences in *investment in skills*. Hence in my framework, the incentives for investment in skills and occupational choice and the resulting endogenous skill distribution are at the core of cross-country income differences.

To study the effects of selection and investment as a source of income differences across countries, I develop a span-of-control model with heterogeneous agents with respect to two generic notions of talent: schooling talent and managerial talent. There is a stand-in household who maximizes the lifetime utility of each household member and every period, a large number of new members are born. The household decides about investment in skills (based on schooling talent), how much to invest in skills, and occupational choice (based on managerial talent), i.e. who becomes a manager within the pool of skilled and unskilled members. The rest of the members become skilled and unskilled workers and their schooling talent is their efficiency units of labor. A key feature of the model is that investment in skills augments both schooling and managerial talents and I will refer to this "augmentation" as becoming skilled. So the initial heterogeneity is amplified after selection for skill investment which will in turn translate into more heterogeneity among managers. Managers use capital, skilled and unskilled workers along with their managerial talent to operate a plant[4] in order to produce output and collect managerial profits. There is an economy-wide productivity term in the production function which is the same for all managers, but it is a source of variation across countries. The equilibrium sum of the managerial talents of all managers (skilled and unskilled) divided by the number of managers is the average managerial quality in this framework.

Investment in skills in the model indicates the costs (investing resources rather than consuming them) and the benefits (the future reward because of having higher schooling and managerial talents). Since the input for the investment in skills is consumption goods, a lower level of aggregate productivity lowers the incentive for the household to invest in its members' skills. The lower investment in skills results in lower average managerial quality.

A fundamental element of the model is the *complementarity* between schooling talent and investment in skills. The members that have higher schooling talent are selected for skill investment, which means an investment in production of both schooling and managerial talent. This selection amplifies and propagates the initial differences in talents and in a stationary equilibrium along a balanced growth

[4] In this paper, I refer to a production unit as a plant. I use the word "establishment" and "plant" interchangeably.

path, skilled managers operate larger and more productive production plants.

I calibrate the parameters in the model to match a set of cross-sectional and aggregate facts of the U.S. economy: macroeconomic statistics, composition of working-age population with regard to education (skilled vs unskilled) and occupation (managers and non-managers), and also the composition of managers regarding their educational attainment (skilled vs unskilled managers). Remarkably, the model can reproduce the central features of the U.S. working-age population based on education and occupation. The model also does an excellent job of generating the skill premium and managerial premium.

The role of size-dependent distortions as a source of income differences across countries has been emphasized in the literature of misallocation.[5] To study this role, I proceed to introduce these distortions. I model them as progressive taxes on the output of a plant and use a simple parametric function, which was originally proposed by Benabou (2002). These distortions have two main effects in my model. First, a *"reallocation effect"* as the presence of distortions suggests that capital and labor services move from more distorted (large) to less distorted (small) plants. Second, an "*incentive effect*" as distortions affect the incentive of investment in skills and thus the average plant level productivity. Overall, the model has a natural framework to study how changes in the level of distortions can account for differences in output per worker. It also helps to understand the differences in managerial quality, educational attainment of the working-age population, skill premium and managerial premium. For my purposes, I assume the U.S. economy is free of distortions.

I find that the model can generate the documented facts when the source of variation is productivity differences or size-dependent distortions. In order to find the elasticity of output per worker with respect to productivity, I calibrate the productivity and distortion parameters to match two fact in my sample of countries: GDP per worker and the share of unskilled working-age population. I find that the model can generate almost all of the disparities in managerial rate, share of skilled managers, managerial premium and skill premium. I then calculate the elasticity of output with respect to productivity which is 2.6. This elasticity is higher than most previous papers[6] and the presence of investment in skills coupled with size-dependent distortions are the reasons for higher elasticity.

Consistent with the facts presented above, my model suggests that lower levels of economy-wide productivity result in both lower average quality of managers as well as lower share of skilled managers. I find that a reduction in productivity by 20% decreases mean plant size by 10% and the share of managers increases by 5%. Output per worker declines by 34% which implies an elasticity of output with respect to productivity of 2.6. This elasticity is higher than the standard neoclassical growth model because by lowering the productivity, the incentives for occupational and educational choices change. The added amplification in the model is due to the investment in talents.

[5]See Hopenhayn (2014), Restuccia and Rogerson (2013) and Restuccia (2013) for recent reviews.

[6]To my knowledge, only Manuelli and Seshadri (2014) finds higher elasticity which is 5.7. Their model is a life-cycle model that features investment in skills for childhood, adulthood and training on the job. So productivity differences affect more margins that I abstracted from in this work.

Lower productivity also changes the composition of managers. The share of skilled managers in the workforce declines by 6%, while the share of skilled managers among the pool of managers decreases by 15%. This shows a high effect of productivity on the selection into managerial occupation. Lower productivity also means that on average, the quality of managers declines by 46%. The decline in quality comes from two channels: lower share of skilled manages and lower investment in skill accumulation by the household.

Size-dependent distortions have similar qualitative effects to lower productivity, but the quantitative effects are more dramatic. A level of distortions that generates a similar output per capita to the case of lower productivity, increases the share of managers by 42%. It also reduces the mean plant size by 84% and the share of skilled managers drops by 80%. Moreover, the average quality of managers declines by 80%.

I then fit the model to data using two parameters (productivity and distortions parameter) and assess the performance of the model regarding other statistics. The model successfully generates skill premium, managerial premium, managerial rate and share of skilled managers observed in the data. I take it from this exercise that the important margins of occupational choice and investment in skills along with the presence of size-dependent distortions are necessary to account for the observed differences in the relevant statistics across countries.

2 Background

My paper builds on the research on how micro level misallocation of resources can emerge as aggregate income and productivity differences on the macro level. I focus on implicit size distortions as a source of misallocation following Guner *et al.* (2008) and Restuccia and Rogerson (2008). What is new here, compared to those papers, is that I explicitly model the investment in education and selection into different occupations which results in endogenous distribution of skills.

My emphasis on education and income of managers as well as investment in their skill links my paper to empirical literature on differences in management practices.[7] It also links to the trade and development literature that study how investment in skills and R&D amplify the effects of productivity differences and distortions.[8] Caselli and Gennaioli (2013) emphasizes the importance of management and quality of managers for cross-country income differences. More recent works show how managers and their incentives matter for aggregate productivity and the size distributions of plants.[9] Guner *et al.* (2018) consider a life-cycle model of heterogeneous managers based on managerial abilities with investment in skills, and study their wage growth relative to non-managers across countries. In contrast, I provide a tractable model to study multiple facts about skill and occupational segregation

[7] see Bloom and Reenen (2011) and Bloom *et al.* (2014)

[8] See Erosa *et al.* (2010), Rubini (2014), Atkeson and Burstein (2011) and Gabler and Poschke (2013) among others.

[9] See Bhattacharya *et al.* (2013), Roys and Seshadri (2017) and Akcigit *et al.* (2016).

of the working-age population along with relative income based on skill level and occupation.

My paper is also connected to work that documents plant and firm-level productivity and size.[10] Poschke (2017) studies a model of occupational choice where the managerial technology improvement is biased towards the skill level of a manager, i.e. managers with higher skills benefit more from improvements in technology. He uses the skill-biased technology change as the driver of the differences in firm size distribution across countries. Bento and Restuccia (2017) present evidence on manufacturing plant size distribution using country level data and develop a model where distortions affect the investment decision in plant-level productivity. In both their models and mine, distortions are amplified by endogenous investment decisions.

3 Data

In this section, I document a set of facts about the number of managers, their skill composition (educational attainment) and relative income across countries using multiple individual level datasets: IPUMS-USA (Ruggles *et al.* (2017)), IPUMS-International (Minnesota Population Center (2018)) and the European Union Statistics on Income and Living Conditions (EU-SILC) (2004-2016). IPUMS-International provides harmonized Census data for a large set of countries. Only few international censuses, however, contain information both on incomes and occupations. The EU-SILC contains both cross-sectional and longitudinal micro-data data for European countries on income, work, poverty, social exclusion and living conditions.

My final sample consists of 38 countries: Austria, Belgium, Bulgaria, Switzerland, Cyprus, Czech Republic, Germany, Denmark, Estonia, Greece, Spain, Finland, France, Croatia, Hungary, Ireland, Italy, Lithuania, Latvia, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Serbia, Sweden, Slovenia, Slovakia, United Kingdom, United States, Canada, South Africa, Brazil, Puerto Rico, Panama and Mexico.

Table (11) in appendix 6 shows the data source for each country as well as the year of the survey and the number of observations. Since I document facts about income of individuals, the sample is limited to countries that has reliable data on income and occupation. I should also emphasize that there are factors other than the initial abilities are important in less developed countries for determining the educational attainment, and occupational choice such as borrowing constraints. I abstracted from these channels and postpone them for future research.

The definition of managers is the individuals working for wage with occupational code 11 to 13 based on ISCO-88 and individuals who identify themselves as self-employed. The occupational codes cover legislators, senior officials and managers. These include occupations whose main tasks consist of determining and formulating government policies, as well as laws and public regulations, overseeing

[10]See Hsieh and Klenow (2009), Bartelsman *et al.* (2013), Hsieh and Klenow (2014) and García-Santana and Ramos (2015) among others.

their implementation, representing governments and acting on their behalf, or planning, directing and coordinating the policies and activities of enterprises and organizations, or departments. The reason to consider self-employed as managers is that self-employed have supervisory and coordinating roles in their own business and do a lot of tasks that managers who work for wage do. Individuals are classified as skill if they have strictly more than a high school degree.

Specifically, I run the following regression in each country. The dependent variable is the log of income. This income is either wage income (for workers and managers with occupational code 11-13), or self-employed income or both. I should emphasize that when an individual has more than one occupation (both wage and self-employed income), he/she is classified as self-employed or working for wage based on whichever occupation he/she spends more hours in.

$$\begin{aligned}\log(\text{labor income}) = \alpha + \beta X + \delta_1 \mathbb{1}\{\text{skilled manager}\} + \delta_2 \mathbb{1}\{\text{unskilled manager}\} \\ + \delta_3 \mathbb{1}\{\text{skilled worker}\} + \epsilon.\end{aligned} \tag{1}$$

The vector X contains individual-specific controls such as log of hours, age, age-squared, gender, race, marital status, veteran status. The inclusion of a variable for years of education is redundant since the skilled vs unskilled classification controls partially for years of education. The coefficients δ_1, δ_2 and δ_3 are used to calculate skill premium and managerial premium.

I present three stylized facts about managers based on the sample of countries that I have and two complementary facts that was documented before in the literature, and I intend to explain these facts through the lens of my model.

3.1 Stylized Facts

1. The share of skilled managers increases with GDP per worker.

It is a well established fact that in less developed countries, the educational attainment of the workforce is generally lower than developed countries and these differences can potentially explain a large part of the differences in income per capita across countries.[11] The GEM survey reveals that similar systematic differences are present even when I focus on managers. In other words, as income per capita increases, so does the educational attainment of managers. This shows an interplay between education and selection into managerial occupation.

Figure (2) shows the educational attainment of adults aged 18-64 in each country (dots) along with the educational attainment of managers (stars). Clearly the managers are more educated than the average population. This shows that simple models of education cannot account for the differences in education of managers across countries. If less developed countries have low levels of education, it seems that managers in those countries have less education as well. The figure indicates that this

[11] Caselli (2005), Cordoba and Ripoll (2013), Erosa *et al.* (2010), Hendricks (2002), Schoellman (2012), Cubas *et al.* (2016).

explanation is not enough. The managers are clearly more educated than the general population, so there is a selection for managerial occupation, but there are still systematic differences across countries, though the differences are smaller as indicated by the correlation coefficient.

2. The managerial premium decreases with GDP per worker.

The second fact about the managers is about their income. I define the relative income of managers to non-managers as the managerial premium. Figure (3) illustrates the relationship between managerial premium and GDP per worker. The correlation between the two is -0.33. This is a new finding that relate the relative income of managers to non-managers and economic development. The negative relationship means that in as one moves from poorer to richer countries, the income gap between the managers and non-managers shrinks. To calculate this premium, I control for observable characteristics of individuals such as age, race, marital status and education. These observables cannot account for the income gap between managers and non-managers. In appendix 4, I explain in more detail how the managerial premium in each country is calculated.

3. The managerial rate decreases with GDP per worker.

The third fact is about the relationship between the number of managers and economic development. Figure (4) shows that in richer countries, the share of managers among all the working individuals (age 25-64), which I call the managerial rate, is lower than poorer countries. In other words, as countries become more developed, the number of managers decline and the correlation between managerial rate and GDP per worker is -0.11. This fact is to my knowledge new. I should emphasize that the number of managers come from individual level micro data and it is based on occupation. Due to data limitations, I could not document this fact from the business side and count the number of managers in establishments. I assume that the same pattern will emerge from analyzing such data.

4. Mean establishment size increases with income per capita.

The forth fact is about the average plant size and its relation to economic development. There are several datasets that provide information on the size of establishments measured by the number of workers. GEM provides data only on relatively small plants which are basically the left tail of the size distribution of establishments. Bento and Restuccia (2017) provide an internationally comparable dataset for manufacturing plants using country-level data and administrative surveys. Poschke (2017) provides data on establishment size for the right tail of the distribution, namely the large establishments. Since all these data sets show that the mean establishment size increases with income per capita, I replicate the evidence by Bento and Restuccia (2017) since it covers most of the size distribution in an internationally comparable fashion. The correlation between log average plant size and GDP per worker is 0.4.

5. The skill premium decreases with GDP per worker.

The fifth fact is about the relative income of skilled versus unskilled individuals which I call skill premium. I define a skilled individual as one with strictly more than a high school degree and unskilled otherwise. Figure (6) shows the relationship between skill premium and economic development. As economies become more developed the income gap between skilled and unskilled individuals shrinks and the correlation between skill premium and GDP per worker is -0.56. I should emphasize that this income gap is only about skill level, regardless of occupation. In other words, I did not take into account the occupation of the individuals to look at the relative income of skilled versus unskilled. The details of calculating the skill premium after controling for observable characteristics of the individuals are presented in appendix 4.

If I put all the facts together, the following picture emerges: In less developed countries, the share of the working individuals who are skilled is lower, the share of managers is higher while the share of skilled managers is lower. On the income side, the relative income of managers to non-managers is higher and the relative income of skilled versus unskilled individuals is also higher. Finally, managers in less developed countries operate smaller plant. In the next section, I provide a theoretical framework to study all of these facts together and I will also give an intuitive interpretation of managerial quality based on investment in skills. This unifying framework will help us understand economic development better and also works a theoretical background for the empirical literature that connects the managerial quality with economic development.

4 Theoretical Framework

There is a single representative household in the economy. The household has a continuum of measure L_t members at time t, who only value consumption. The size of the household grows at the constant rate of (g_L).[12] The household is infinitely lived and maximizes

$$\sum_{t=0}^{\infty} \beta^t L_t \log(C_t/L_t) \tag{2}$$

where $\beta \in (0,1)$ and C_t denotes total household consumption at time t.

4.1 Endowments

Each household member is born with two innate talents: schooling talent (a) and managerial talent (z). Talents are distributed with support in $[0,\bar{a}] \times [0,\bar{z}]$, with CDF $F(a,z)$ and density $f(a,z)$. Household members have one unit of time that is supplied inelastically. Each household member can be a *skilled*

[12] I introduce population growth so the model has standard balanced-growth properties, and thus can be better mapped to data.

manager, *unskilled manager*, *skilled worker* or *unskilled worker*.[13] I describe the related decisions and income of each type of member below. The household is also endowed with an initial capital stock of $K_0 > 0$.

4.2 Technology

There are two types of production plants in this economy. One that is operated by a skilled manager which is decreasing returns to scale and requires four inputs: capital (k),two types of labor services: skilled labor (s) and unskilled labor (u) and managerial talent if the manager who is operating this technology z. Output is given by

$$\boldsymbol{y = f(k,u,s;A,z) = y = Az^{1-\gamma}\Big(k^{\alpha}(u^{\theta}s^{1-\theta})^{1-\alpha}\Big)^{\gamma}} \tag{3}$$

where A denotes an exogenous measure of "economy-wide" productivity which I will refer to as TFP from this point forward, $\gamma \in (0,1)$ is the span of control parameter and $\theta, \alpha \in (0,1)$. The elasticity of substitution between skilled and unskilled labor is one.

An unskilled manager operates the same production technology and the only difference between the two types of managers is that the skilled managers had their managerial talent augmented due to obtaining education. How this augmentation occurs is similar to the skilled workers and I will elaborate on it when discussing the effect of education on the skilled workers in section (4.3).

Both types of managers face competitive prices. The price of one unit of physical capital in period t is R_t. The price for the use of skilled labor in period t is $W_{s,t}$ per efficiency units of skilled labor. $W_{u,t}$ is a similar price for unskilled labor. From the stand point of a manager, all efficiency units of skilled labor are perfect substitutes. The same is true for unskilled labor. So a plant is a manager that operates a technology which requires physical capital and efficiency units of skilled and unskilled labor. Capital depreciates at the rate δ.

4.2.1 Managers

Managers maximize their profit, taking prices of inputs as given. Since their problem is a static one, I omit the subscript t. The problem of a manager with managerial talent z is:

$$\max_{\{k,u,s\}} \; Az^{1-\gamma}\Big(k^{\alpha}(u^{\theta}s^{1-\theta})^{1-\alpha}\Big)^{\gamma} - Rk - W_u u - W_s s \tag{4}$$

The solution to the above optimization for an unskilled manager gives the demands for the unskilled labor $u^u(z, W_u, W_s, R)$, skilled labor $s^u(z, W_u, W_s, R)$ and capital $k^u(z, W_u, W_s, R)$. The profits of an unskilled manager is $\pi^u(z, W_u, W_s, R)$.
Similarly, for a skilled manger, the demands for unskilled labor, skilled labor, capital and profits are

[13] When I refer to efficiency units of workers, I refer to them as "labor". The context should eliminate any confusion.

$u^s(z, W_u, W_s, R)$, $s^s(z, W_u, W_s, R)$, $k^s(z, W_u, W_s, R)$ and $\pi^s(z, W_u, W_s, R)$. The following proposition illustrates two characteristics of the above functions:

Proposition 1 *The demand and profit functions of both types of managers satisfy the following conditions:*

1. *They are strictly increasing in managerial talent and strictly decreasing in prices.*
2. *They are linear functions of the managerial talent. Specifically, the demands for input factors and profits per managerial talent are equal for both types of managers:*

$$\frac{i^u(z, W_u, W_s, R)}{z} = \frac{i^s(z, W_u, W_s, R)}{z}, \quad i \in \{u, s, k, \pi\} \tag{5}$$

The proof is in (9). I will refer to the demands for factors and profits per managerial talent as $u(W_u, W_s, R)$, $s(W_u, W_s, R)$, $k(W_u, W_s, R)$ and $\pi(W_u, W_s, R)$.

4.3 The Household Problem

The household has to decide which newborn member becomes skilled and within each pool of skilled and unskilled, which occupation each member is assigned to. Specifically, the household observes the schooling talent of each member and assigns the member to one of the two pools of skilled and unskilled. Turning a newborn into skilled is costly; it requires time and goods. After this segregation, the household must decide about the occupation of the newborns in each pool: worker or manager. So the members of the household are categorized into four categories: *skilled manager*, *unskilled manager*, *skilled worker* or *unskilled worker*.

If a newborn household member is selected for the unskilled labor pool at time t, her schooling talent is transformed into efficiency units of unskilled labor in the same period and her income is given by $W_{u,t}a$. If she is selected to be an unskilled manager, her schooling talent is foregone and she becomes an unskilled managers and earns profits.

If she instead is selected for the skilled pool, it takes one period to turn her into a skilled member and the household has to forgo her earnings for one period either as an unskilled worker or an unskilled manager. The household invests x_t units of consumption goods to augment her talents. Investing x_t implies that her talents are augmented by the factor h_{t+1}, where

$$\boldsymbol{h_{t+1} = B x_t^{\phi}} \tag{6}$$

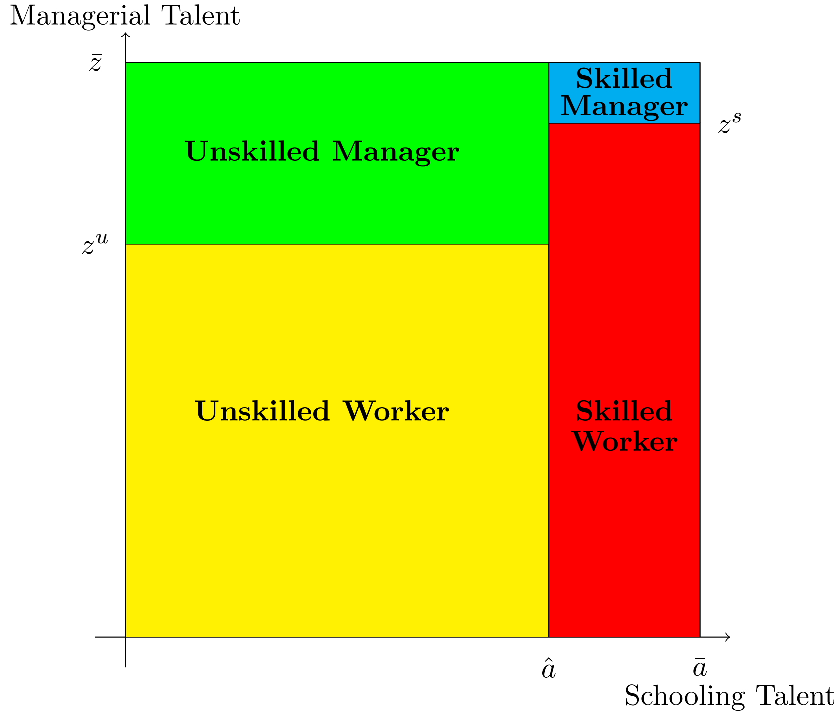
with $\phi \in (0, 1)$ and B is a parameter determining the relative efficiency of consumption goods that are invested in augmenting talents. Her income then is given at $t + 1$ by $W_{s,t+1}ah_{t+1}$ if she becomes a skilled worker. If she is selected to become a skilled manager, her efficiency units as a skilled worker are foregone and she earns profits. Her managerial talent is augmented by the same factor as her schooling talent, so her managerial talent at time $t+1$ is zh_t. This augmentation of managerial talent is a result of education and the skilled and unskilled managers differ in this regard.

It is worth emphasizing that the segregation of household members is based on the schooling talent (a). The fate of a member is going to be very different when she is selected for one of the skilled or unskilled pools. The members who were selected for the skilled pool obtain education and this investment in education will change their occupational outcomes dramatically. As we will see in section (6), the choice of schooling is a dominant margin when the economic environment such as economy-wide productivity changes. Most of the differences across countries can be attributed to the choice of schooling based on the model.

It follows that only members with sufficiently high levels of schooling talent become skilled and within each pool of skilled and unskilled, members with sufficiently high managerial talent become managers. Given rental prices, there exists a unique threshold $\hat{a}_t$ such that newborns with schooling talent below this threshold become unskilled at time t, and those with schooling talent above this become skilled from $t+1$ on. There are also two unique thresholds for becoming a manager. Unskilled members with managerial talent higher than z_t^u become unskilled managers at time t and the ones below this threshold become unskilled workers at time t. Skilled members at time $t+1$ with managerial talent higher than z_{t+1}^s become skilled managers and the ones below this threshold become skilled workers.

Figure (1) shows how the household members are segregated into the four categories based on education and occupation.

Figure 1: **Segregation of household members into four categories, based on education and occupation**



The household's problem is to choose (i) the sequence of consumption, (ii) the fraction of household members who are skilled and unskilled, (iii) the fraction of unskilled managers, (iv) the fraction of skilled managers, (v) the amount of consumption goods invested in augmenting the talents of new skilled members, and (vi) the capital for the next period. Formally, the household's problem is to choose $\{C_t, I_t, x_t, \hat{a}_t, z_t^u, z_t^s\}_0^\infty$ to maximize (2) subject to (6) and

$$C_t + I_t + N_t x_t \int_0^{\bar{z}} \int_{\hat{a}_t}^{\bar{a}} f(a, z) da dz \leq (W_{u,t} U_t + W_{s,t} S_t) + R_t K_t + \Pi_t^u + \Pi_t^s \tag{7}$$

$$K_{t+1} = (1 - \delta) K_t + I_t \tag{8}$$

and

$$N_0, S_0, U_{-1}, \Pi_0^s, \Pi_{-1}^u, K_0 > 0 \tag{9}$$

where U_t is the stock of the efficiency units of unskilled labor, S_t is the stock of the efficiency units of skilled labor, Π_t^u is the flow of profits of all the unskilled managers and Π_t^s is the flow of profits of all the skilled managers at time t. The laws of motion for these stocks and flows are as follows.

Each period, the number of newborns is $\boldsymbol{N_t = g_L L_{t-1}}$. Based on the above assignment, the laws of motion for the stocks of the efficiency units of unskilled and skilled labor are:

$$U_t = U_{t-1} + N_t \int_0^{z_t^u} \int_0^{\hat{a}_t} af(a,z)dadz \tag{10}$$

$$S_t = S_{t-1} + N_{t-1}h_t \int_0^{z_t^s} \int_{\hat{a}_{t-1}}^{\bar{a}} af(a,z)dadz \tag{11}$$

Note that at any time t, $U_t + S_t$ is not equal to L_t, since U_t is the total efficiency units of unskilled labor and S_t is the total efficiency units of skilled labor, whereas L_t is the size of the household at time t. Household also needs to know about the profits that managers are earning. The laws of motion for profits of unskilled and skilled managers are:

$$\Pi_t^u = \Pi_{t-1}^u + N_t \int_{z_t^u}^{\bar{z}} \int_0^{\hat{a}_t} z\pi(W_{u,t}, W_{s,t}, R_t)f(a,z)dadz \tag{12}$$

$$\Pi_t^s = \Pi_{t-1}^s + N_{t-1}h_t \int_{z_t^s}^{\bar{z}} \int_{\hat{a}_{t-1}}^{\bar{a}} z\pi(W_{u,t}, W_{s,t}, R_t)f(a,z)dadz \tag{13}$$

where Π_{t-1}^u is the flow of profits from unskilled managers who were managers from period $t-1$ and before. Equation (12) states that the flow of profits to household from unskilled managers at time t equals to the flow of profits from all the previously assigned unskilled managers and the flow of profits of the unskilled managers that are going to be assigned in period t. The law of motion for the flow of profits from skilled managers has a similar interpretation. Note that I used the result of proposition (1) for the profits of both types of managers. They are both linear in managerial talent with the augmentation for skilled managers due to education. The solution to the household's problem is described by the following first order conditions:

- Euler Equation

$$\frac{1}{\frac{C_t}{L_t}} = \beta \frac{[R_{t+1} + (1-\delta)]}{\frac{C_{t+1}}{L_{t+1}}} \tag{14}$$

  Condition (14) is the standard Euler equation.

- Selection for Education ($\hat{a}$)

$$\frac{W_{u,t}\int_0^{z_t^u} \hat{a}_t f(\hat{a}_t,z)dz + \int_{z_t^u}^{\bar{z}} z\pi f(\hat{a}_t,z)dz + x_t \int_0^{\bar{z}} f(\hat{a}_t,z)dz}{\frac{C_t}{L_t}} = \beta \frac{\Big(W_{s,t+1}(\int_0^{z_{t+1}^s} \hat{a}_t f(\hat{a}_t,z)dz) + \int_{z_{t+1}^s}^{\bar{z}} z\pi f(\hat{a}_t,z)dz\Big)Bx_t^{\phi}}{\frac{C_{t+1}}{L_{t+1}}} \tag{15}$$

  Condition (15) states that the optimal decision for education threshold should be in such a way that the compensation of marginal unskilled workers plus profits of marginal unskilled managers plus the cost of skill augmentation on the margin (marginal cost of education) in this period, is equal to the discounted compensation of marginal skilled workers plus the profits of marginal skilled managers (marginal benefits of education) in the next period.

- Investment in Education (x)

$$\frac{\int_{\hat{a}_t}^{\bar{a}}\int_0^{\bar{z}} f(a,z)dadz}{\frac{C_t}{L_t}} = \beta\frac{\Big(W_{s,t+1}(\int_{\hat{a}_t}^{\bar{a}}\int_0^{z_{t+1}^s} af(a,z)dadz) + \int_{\hat{a}_t}^{\bar{a}}\int_{z_{t+1}^s}^{\bar{z}} z\pi f(a,z)dadz\Big)B\phi x_t^{\phi-1}}{\frac{C_{t+1}}{L_{t+1}}} \tag{16}$$

Condition (16) states that the marginal cost of investing one unit of consumption good in the talents of marginal skilled members must be equal to the marginal benefit which is the marginal increase in efficiency units times its rental price times "raw" addition to its pool plus the profits of marginal managers whose managerial talent is augmented.

- Occupation Choice for Unskilled Members (z^u)

$$W_{u,t}\int_0^{\hat{a}_t} af(a,z_t^u)da = \int_0^{\hat{a}_t} z_t^u\pi f(a,z_t^u)da \tag{17}$$

Condition (17) states that the compensation of marginal unskilled workers must be equal to the profits of marginal unskilled managers. I can rewrite this condition to get an expression for the threshold of choosing managerial occupation for the unskilled members:

$$z_t^u = \frac{\frac{W_{u,t}\int_0^{\hat{a}_t} af(a,z_t^u)da}{\int_0^{\hat{a}_t} f(a,z_t^u)da}}{\pi} \tag{18}$$

The numerator is the marginal wage earnings per marginal unskilled worker. The denominator is the profits per unit of managerial talent. The optimality condition dictates that for a chosen threshold of education ($\hat{a}$), if the wage per unskilled efficiency units increases relative to the profits per managerial talent, household decides to choose less unskilled members to become unskilled managers and hence the threshold increases.

- Occupation Choice for Skilled Members (z^s)

$$W_{s,t}\int_{\hat{a}_{t-1}}^{\bar{a}} af(a,z_t^s)da = \int_{\hat{a}_{t-1}}^{\bar{a}} z_t^s\pi f(a,z_t^s)da \tag{19}$$

Condition (19) states that the compensation of marginal skilled workers must be equal to the profits of marginal skilled managers. I can rewrite this condition as well:

$$z_t^s = \frac{\frac{W_{s,t}\int_{\hat{a}_{t-1}}^{\bar{a}} af(a,z_t^s)da}{\int_{\hat{a}_{t-1}}^{\bar{a}} f(a,z_t^s)da}}{\pi} \tag{20}$$

The numerator is the marginal wage earnings per marginal skilled worker. The denominator is the profits per unit of managerial talent. The optimality condition dictates that for a chosen threshold of education ($\hat{a}$), if the wage per skilled efficiency units increases relative to profits the per managerial talent, household decides to choose less skilled members to become skilled managers and hence the threshold increases.

4.4 Equilibrium

A Competitive Equilibrium is a collection of sequences $\{C_t^*, K_{t+1}^*, \hat{a}_t^*, z_t^{u*}, z_t^{s*}, x_t^*, W_{u,t}^*, W_{s,t}^*, R_t^*\}_{t=0}^{\infty}$ such that given prices, managers maximize their profits, $\{C_t^*, K_{t+1}^*, \hat{a}_t^*, z_t^{u*}, z_t^{s*}, x_t^*\}_{t=0}^{\infty}$ solves the household's problem and labor markets for skilled and unskilled workers as well as capital and goods markets clear.

Balanced Growth

Along a balanced growth path, aggregate consumption, investment, production, capital, profits of unskilled and skilled managers, and the pools of skilled and unskilled workers are growing at the rate of population growth. Also, the thresholds for education and occupation choices are fixed as well as the investment in education per member. I omit the subscript t and superscript $*$ since the equilibrium quantities are either constant or growing with constant population growth rate.

A feature of the balanced growth equilibrium in the model is that the rental rate of capital is constant and equals $(1/\beta - (1 - \delta))$. This helps with computing the equilibrium. Specifically, starting with a guess for the pair (W_u, W_s), one can solve for $\{\hat{a}, z^u, z^s, x\}$ simultaneously using equations (15), (16), (17) and (19). Then market clearing for skilled and unskilled labor can determine the new guess for the initial pair. Iterating on this pair, the equilibrium objects can be calculated. Consumption and capital can be calculated using aggregate feasibility and capital market clearing conditions.[14]

5 Parameter Values

I start by setting the model period to *four* years to reflect a more realistic time period to become a skilled member. I use the U.S. as a benchmark to calibrate the parameter values and choose these values to match the observed characteristics of the U.S. economy in the steady state equilibrium of the model. Since the model does not have any implications over the life-cycle, I focus on prime age individuals (age 40-54) and years 2014-2016. The details of calculating moment conditions from the data is described in Appendix (9).

[14]See Appendix 2 for more details.

5.1 Technology

I follow Guner *et al.* (2018) for choosing the value for the span of control parameter (γ) and set it to 0.77. In order to match the capital to output ratio of 0.33 based on the findings of Gollin (2002) on labor shares across countries, I set the share of capital in the production function (α) to 0.43. I choose (θ) so that in the steady state, the model reproduces the fraction of unskilled labor (worker+managers) in the U.S. This share is 40% and it is the fraction of the population aged 15 years and older that completed secondary education or less.

I set the depreciation rate to 26.6% (7.4% at the annual rate), so that given the capital-to-output ratio,the model is consistent with the observed investment-to-output ratio. Cubas *et al.* (2016) used NIPA data for the period 1960-2010 to calculate an average investment-to-output ratio of 0.27 and a capital-to-output ratio of 0.8 at the four-year frequency (3.2 at the annual rate). The productivity level (A) is set to 1.0 as a normalization.

5.2 Preferences and Demographics

I set the discount factor so that in the four-year steady-state, the capital-to-output ratio is 0.8. This implies that (β) equals to 0.869 (0.966 at the annual rate). The value for β means an annual interest rate of 3.6%. Based on the Penn World Tables 7.0, I set the growth rate of population to the annual rate of 0.9%.

5.3 Distributions of Talents

I calibrate the distribution of talents to match several moments. I assume exponential distributions for talents with parameters λ_a fo schooling talent and λ_z for managerial talent. There is also a level of correlation between talents which I will indicate with a parameter ρ. I construct a joint distribution of talents with CDF $F(a, z)$ and PDF $f(a, z)$ using the notion of copula.[15] Specifically, I create a bivariate exponential distribution using the method developed by Gumbel (1960). Suppose the marginal distributions of talent are univariate exponential distributions as follows:

$$g(a;\lambda_a) = \begin{cases} \frac{1}{\lambda_a} e^{-\frac{a}{\lambda_a}} & a \geq 0, \\ 0 & a < 0. \end{cases} \qquad G(a;\lambda_a) = \begin{cases} 1 - e^{-\frac{a}{\lambda_a}} & a \geq 0, \\ 0 & a < 0. \end{cases} \tag{21}$$

and

$$h(z;\lambda_z) = \begin{cases} \frac{1}{\lambda_z} e^{-\frac{z}{\lambda_z}} & z \geq 0, \\ 0 & z < 0. \end{cases} \qquad H(z;\lambda_z) = \begin{cases} 1 - e^{-\frac{z}{\lambda_z}} & z \geq 0, \\ 0 & z < 0. \end{cases} \tag{22}$$

[15] A copula is a function that produces multivariate distributions out of any set of arbitrary univariate distributions; see Nelsen (2006) for more details.

where the upper case letters denote CDF and the lower case letters denote pdf of the distributions. Then the bivariate distribution can be characterized by the following probability distribution function:

$$\boldsymbol{f(a,z;\lambda_a,\lambda_z,\rho)} = \begin{cases} \boldsymbol{g(a)h(z)\Big(1+\rho(2G(a)-1)(2H(z)-1)\Big)} & \boldsymbol{a,z \geq 0,} \\ \boldsymbol{0} & \boldsymbol{a<0} \text{ or } \boldsymbol{z<0.} \end{cases} \tag{23}$$

The parameter ρ governs the correlation between the random variables with $\rho \in [-1,1]$. The correlation coefficient between the two random variables is $\frac{\rho}{4}$. Several remarks about the choices of distributions and copula are in order.

First, they allow for richness and flexibility in matching data and at the same time, retaining a parsimonious set of parameters. One dimension of richness is that household members' talents are dependent which a priori seems a reasonable assumption. This allows for the feature that a household member who is talented in one talent is more talented in the other talent as well. Another dimension of richness is that dispersion in household members' talent is not the same for each talent. Specifically, schooling talent is a broad notion for relative advantage in obtaining education which constitutes a variety of activities related to education such as studying hard, managing personal finances, balancing between school and social life and so on. Managerial talent is a different notion for relative advantage in managing a business and performing tasks such as hiring/firing decisions, forecasting the needs of the business, motivating and monitoring employees, organizing, negotiating contracts and so on. So these notions of talent are capturing different aspects of heterogeneity in the household members. Since managerial talent is a stand-in for arguably more diverse types of activities than those related to education, one might expect that the dispersion of managerial talent of the household members is larger than schooling talent. The choice of distributions allows for this possibility.

The second reason for the above choices is that they allow for closed-form solutions to double integrations needed to solve the model. Given that the equilibrium objects consist of threshold levels in integration, the solution is very sensitive to the error level if I want to compute the integrals numerically. The closed-form solutions eliminate this difficulty.

The third reason for choosing these distributions is that they allow my theory to fail. In particular, there is nothing inherent in these distributions that assures that the model should behave in a certain way to reproduce the empirical targets and generate the cross country patterns observed in the data. The success of the model in this regard will be dictated by the data and the calibration procedure.

In order to identify the parameters of the distributions, I calibrate the model to reproduce three observed moments in U.S. data in equilibrium. The first moment is managerial rate which is the percentage of prime-age working population who are either owner-manager of a business or employed to be the manager of a established business. So the managers in the data are those who are employed as managers and can be labeled as managers based on the Census Bureau occupation codes as well as self-employed individuals. Using data from BLS, this rate is about 17%. The second moment is the share of skilled managers in the U.S. where I defined the level of education to be considered *"skilled"* in section (5.1). Using BLS data, this amounts to 77%. The third moment is the managerial premium which is defined as the relative earnings of a manager to a worker (regardless of skill level).

The curvature (ϕ) and level parameter (B) in the production of skills are set so that the model reproduces two empirical targets in a steady state equilibrium. The first target is the expenditure per tertiary student as a fraction of GDP per worker (at PPP values) in the U.S. According to OECD (2017) report, this fraction was 25% in U.S. in 2014. The second target is the skill premium. It is defined as the relative earnings of a skilled person to an unskilled one regardless of occupation which is 1.71 in U.S. The details of calculation the skill premium and managerial premium are outlined in Appendix (9)

Table (5) shows the calibrated parameters based on matching the above moments. The interesting result is the parameter values for talent distributions.

As I mentioned before, the notion of managerial talent is capturing a wider variety of activities and skills than the notion of schooling talent. The calibrated values for parameters of the distribution indeed approves this a priori expectation. The schooling distribution has mean and standard deviation of 9.7, while the managerial talent has mean and standard deviation of 10. The correlation between the two talents is positive but small 0.005. It seems that this is reasonable since a negative correlation does not have a good economic meaning and a positive but large correlation means that individuals who are talented in one dimension are more likely to be talented in the other and it sounds like natural selection.

Table(6) shows the reproduced moments from the model and compares them with the empirical moments. The model does an excellent job of matching the moments in the data. The equilibrium for the benchmark model has an interesting feature about relative quality of managers and their steady state distribution. The share of skilled managers in the U.S. is 77% which model does a good job of reproducing. As a measure of external validity of the model, I reported two additional moments. One is the relative income of skilled managers to unskilled workers, which is 2.2 in the U.S. data. The other is the income of unskilled managers to skilled workers, which is 0.67 in the U.S. data. The model reproduces these two moments fairly well. Note that since there are four types of members in the model, skilled and unskilled managers along with skilled and unskilled workers, I can only match at most three moments related to relative income. The forth moment will be mechanically matched. I chose to match only two income moments and use the other two to test for external validity.

But the important distinction between skilled and unskilled managers in the model is the fact that skilled manager obtained education and augmented their managerial talent which amplifies the starting heterogeneity between skilled and unskilled managers. This feature reveals itself in the relative income of skilled vs unskilled managers. Specifically, the average income of a skilled managers in the model and data is more than twice as much as the average income of an unskilled worker. The selection for schooling along with selecting into higher income occupation based on having a higher managerial talent are the channels contributing to this income gap.

It is worth emphasizing that the mean plant size in the current setup is about 5, whereas in the data, the mean plant size for the U.S. is 17, as calculated by Guner *et al.* (2008). The reason being that in the real world, plants can have multiple managers, while in the model, every plant has only one manager. So, unless I deviate from the simple span-of-control setup here to allow for multiple managers per plant, I cannot reproduce simultaneously the share of managers as well as the mean plant size. Guner *et al.* (2008) focuses on the size distribution to study the aggregate implications of size-dependent distortions without investment in skills, while I focus on the share of managers and the composition of managers with respect to skills. Therefore, the current setup matches the share of managers while failing to match the mean plant size. Since my main concern is not the size distribution per se and the relative change in the mean plant size of skilled vs unskilled managers is a result of the framework under study, I chose to match the share of managers rather than the mean plant size.

Figure (7) shows the equilibrium distribution of managers in the model economy. One can see that the cutoff values for selection into managerial occupation for skilled and unskilled managers are different. Specifically, the cutoff value for unskilled managers is higher. The reason is that the outside options for skilled and unskilled manager are different. The unskilled manger can become an unskilled worker and earn w_u for her efficiency units while the skilled manger gains w_s if she works as a skilled worker. Since in equilibrium, the share of unskilled managers is low, the selection is higher for them in contrast with the skilled managers.

What is important here is the average managerial talent of skilled and unskilled managers in the equilibrium which eliminates the relative share of each type of managers and gives us a better statistic to gauge the relative quality of skilled and unskilled managers. This quality translates one to one into the average plant level productivity run by different types of managers for a fixed level of A. It turns out that the average managerial talent of the unskilled managers is 46% lower than the average managerial talent of the skilled managers. So the investment in education outweighs the higher share of skilled managers and the average plant level productivity for skilled mangers is higher than the unskilled mangers.

6 Model Mechanics

I present a set of results pertaining to economy-wide notion of productivity differences across countries along with an experiment of introducing size-dependent distortions in the model and assessing their effects on the model equilibrium. The two channels, productivity and distortions, have different implications for the model, both qualitatively and quantitatively. The productivity channel has a uniform impact on all firms; the generic productivity term (A) reduces by 20%. The reaction of managers to this productivity loss is heterogeneous due to the fact that managers have different levels of managerial talent. The distortion channel has a heterogeneous effect from the beginning since it hits each plant based on its size (value-added in the model). I will elaborate on the exact modeling strategy for size distortions in section (6.2).

6.1 Productivity Differences

The results from a 20% reduction in productivity is presented in table (7). A reduction in productivity results in a lower level of output per capita,[16] share of skilled labor, skilled manager, investment in skills and mean size of establishments. It also results in an increase in the shares of unskilled labor, unskilled manger, share of managers in the population, managerial premium and skill premium.

Before continuing with the intuition behind these results and how they connect to the stylized facts in section (3), a note on the notion of size in the model is in order. The size of establishments are usually measured by the number of workers in the establishment. In my theory, managers demand efficiency units and workers supply them. So the size of a plant is not well-defined. But based on the segregation of the population based on occupation and the fact that each manager represents one and only one establishment and all the workers are hired by managers, the notion of the mean size of establishments can be defined as:

$$\text{mean size} = \frac{\#\text{ of workers}}{\#\text{ of managers}} \tag{24}$$

So the notion of size is a by-product of the model and changes in this variable come from the relative changes in the number of managers and workers. I can also define the mean plant size for skilled and unskilled managers. The trick is to determine the share of skilled and unskilled workers that are employed by each type of manager. I define these shares as follows: the share of efficiency units of skilled workers that is demanded by the skilled managers is the same as the share of the number of skilled workers that is demanded by skilled managers. The rest of the skilled workers are employed by unskilled managers. The share of unskilled workers for each type of manager is defined analogously. Based on the model setup, this is a natural way of mapping the efficiency units demanded to the number of each type of worker that is demanded by managers. When I know the total number of workers demanded by each type of managers, then the mean plant size can be calculated using (24).

Panel A in table (7) shows how the model conforms to the patterns found in the data. Fact 1 states that the share of skilled managers and workers increase with income per capita. Fact 2 states that the managerial premium declines with income per capita, while Fact 3 asserts that the managerial

[16] Since in the model, every household member is either a worker or a manager and there is no unemployment, the notion of output per capita which is the total output divided by the size of the household seems more appropriate than output per worker. The latter may confuse the reader into thinking that the total output is divided only by the size of workers in the household.

rate declines with income per capita. Fact 4 reveals that the mean size of establishments increases with income per capita. Finally, fact 5 states that the skill premium declines with income per capita.

A 20% reduction in productivity in the model results in a decline in the share of skilled labor 34% while the share of skilled managers among managers only decreases by 15% which is almost half of the decline in skilled population share. Based on the calibration of the benchmark economy, the share of skilled managers among managers is still higher than the share of skilled working-age population in the new equilibrium. All of these changes in shares conforms to the fact (1). In line with fact (2), the managerial premium increases by 8%. The managerial rate increases by 5% which is in line with fact (3) and the mean size of establishments declines by 10%, conforming to fact (4). Finally, skill premium increases by 12% which is what fact (5) shows.

Panel B gives some additional information about the output and segregation of the household members into different occupations and skill levels. The output per capita falls by 12% which is intuitive based on lower productivity that translates into lower levels of output. The share of unskilled workers increases by 6% while the share of skilled workers declines by 8%. The share of unskilled managers increases by 22%. Putting panels A and B together, the total share of managers in the population increases along with the share of unskilled managers. Within the pool of managers, the share of skilled managers declines. This shows that in the model, an economy with 20% lower productivity exhibits more managers who on average, runs smaller plants and the share of skilled managers among them is lower.

Panel C elaborates more on the notion of quality and size among managers. With 20% lower productivity in the model economy, the investment in skills which provides higher quality declines by 23%, the average quality of all managers drops by 28% and the average quality of skilled managers declines by 12% while the average quality of unskilled managers decreases by 42%. This means that in an economy with lower productivity, the quality of all managers is lower. The relative quality of a skilled manager to an unskilled manager increases by 56%. This shows that the skilled managers in the economy with lower productivity are more selected: they have less investment in them, but the gap between their talents relative to the unskilled managers is higher.

There is an asymmetry between the two types of managers in terms of quality. Specifically, the average quality of the unskilled managers declines more due to two reasons. First, the lower productivity induces the household to select more unskilled members to become managers, which results in lower average productivity among unskilled managers. I call this the *"selection effect"*. The other reason is that the skilled managers have investment in their managerial talents. The investment is lower due to lower productivity, and it is nonexistent for unskilled managers. This investment increases the average quality of skilled managers relative to unskilled managers and I call it the *"skill effect"*. Together, these effects cause a wider gap between the average quality of skilled versus unskilled managers in economies with lower levels of productivity.

The average plant size of skilled managers almost stays the same when productivity is lower while the average plant size of unskilled managers declines by 53%. The relative plant size increases by 114%. This change in plant size is consistent with studies showing that in less developed countries, the distribution of plant size consists of a lot of small plants along with a few big ones and the middle part of the distribution is *"missing"*.[17] Although I am not trying to explain the size distribution of plants across countries, it is interesting that even with my notion of size in this paper, the relative size gap is higher in economies with lower productivity.

The intuition behind the results is as follows. As the productivity declines, it is relatively more expensive for the household to invest in skill accumulation of its members, so the household selects fewer members for skill investment and invest in them less. The choice of occupation will change based on three factors: profits per managerial talent, prices per efficiency units of labor (whether skilled or unskilled) and investment in skills. In a simple span-of-control model without skill accumulation such as Lucas (1978), the change in productivity will not affect occupational choice if the technology has a Hicks neutral productivity term.[18] The profits and wages would adjust in a way that the choice of occupation is unchanged (i.e. the threshold for becoming a manager does not change). The effects of productivity change are more subtle here because of the skill accumulation and selection for skill investment.

For the choice of unskilled managers, the profits per managerial talent declines as a result of the drop in productivity, so it becomes less profitable to be an unskilled manager. On the other hand, the wage per efficiency units of unskilled workers will adjust and it may go up or down depending on the supply and demand for unskilled workers in the new equilibrium. This wage may weaken or strengthen the desire of the household to select unskilled members to become unskilled managers. In the new steady state equilibrium, the share of unskilled managers increases.[19] The interesting feature of the composition of the unskilled managers in the new equilibrium is that the increase is because of the fact that the household selects some members with high initial schooling and managerial talent to become unskilled (i.e. not selecting them for skill investment), but because their managerial talent is high, they become unskilled managers and some of the unskilled managers in the previous equilibrium with low managerial talent are unskilled workers in the new equilibrium.[20] This means that the choice of the schooling has a very important effect on the choice of occupation.

The effects on the profits per managerial talent for the skilled managers is the same. The other determinant of this choice for the household is the wage per efficiency units of a skilled worker and the investment in skills which augments also managerial talent. With lower investment in skills, lower share of skilled household members and lower profits per managerial talent, one expects that the

[17]See Hsieh and Olken (2014) for a discussion of the literature of the firm size distribution in developing countries.

[18]This requires a Cobb-Douglas technology which is the case in my model.

[19]Table (7), Panel B.

[20]For the mathematically inclined reader, it means that z^u increases, but because a also increases, the net effect on the share of unskilled managers is an increase rather than a decreases. As a visual aid, take a look at figure (1).

share of skilled managers declines in the new equilibrium which indeed happens.[21] Since in the new equilibrium, there are more unskilled managers and less skilled managers, the share of skilled managers among the pool of managers declines.

What happens to the total number of managers? The above explanation shows lower share of skilled managers and higher share of unskilled managers in the equilibrium. The total effect is an increase in the share of managers in the workforce and based on the definition for size of a plant in the model, the mean size declines with a decrease in productivity.

What happens to relative income of based on skills and occupation? The relative income of managers to workers increases since in the new equilibrium the wages decline more than the profits. The fact that there are more managers in the new equilibrium does not offset the relative income of managers to workers and the managerial premium increases. The relative income of skilled versus unskilled members increases because of two effects. The first is an adjustment of wages where in the new equilibrium, wages are lower. The other effect is the lower share of skilled members in the new equilibrium. Although the investment in the skills of skilled members decline, the relative income of skilled members increases which causes the skill premium to increase.

6.2 Size Dependent Distortions

I investigate the role of size-dependent distortions in my model here. The environment is the same as before, but the managers now face distortions when operating plants and these distortions depend on the size of the operation. My model has a delicate notion of size; the size of a plant is only defined in a steady-state equilibrium via the relative share of managers to workers. So, instead of introducing distortions on the size of the operations, I model them to be output taxes that act based on the value-added of the plant. In particular, I assume an establishment with output y faces an average tax rate $T(y) = 1 - \nu y^{-\tau}$. This tax function was proposed originally by Benabou (2002) and has an intuitive interpretation: if $\tau = 0$, all establishments face the same constant tax rate $(1-\nu)$. For $\tau > 0$, the distortions are dependent on size; larger establishments face higher distortions than smaller ones. Hence, τ controls the level of the dependence of distortions on size. In (9), the profit and demand functions for a manager with managerial talent z and a given level of τ are presented. I am going to assume that the economy-wide productivity is constant and normalize it to $A = 1$. An interesting feature of the distortions is that the demand and profit functions per managerial talent are no longer the same for all managers, specifically:

$$\frac{k(z', W_u, W_s, R, A)}{k(z, W_u, W_s, R, A)} = \frac{u(z', W_u, W_s, R, A)}{u(z, W_u, W_s, R, A)} = \frac{s(z', W_u, W_s, R, A)}{s(z, W_u, W_s, R, A)} = \frac{\pi(z', W_u, W_s, R, A)}{\pi(z, W_u, W_s, R, A)} = \left(\frac{z'}{z}\right)^{\frac{(1-\gamma)(1-\tau)}{1-\gamma(1-\tau)}} \tag{25}$$

[21] Table (7), Panel A, fact 1.

where

$$\frac{(1-\gamma)(1-\tau)}{1-\gamma(1-\tau)} < 1 \tag{26}$$

as long as $\tau > 0$. This means that for a *given* distribution of managerial talents, the distribution of plant size becomes more compressed when size-dependent distortions are present.

6.2.1 The Case Without Investment in Skills

It is worth considering a simpler version of the model where the choice of investment in skills is absent. In this case, the dynamics of occupational choice disappears and the choice of occupation between worker and manager is contemporaneous for each cohort of newborns in the same period that they are born. Therefore, the economy contains two types of members: workers and managers. The managers employ capital and efficiency units of workers. This setup is essentially the one studied in Guner *et al.* (2008).

The distortions reduces the overall demand for labor in the model economy. This reduction in the labor demand drives down the equilibrium wage rate, while the rental rate of capital stays constant. The reduction in the wage rate changes the incentives of the marginal worker in favor of becoming a manager since the reduction in the wage rate translates one-to-one in reduction in income as a worker, while the reduction in profits for that marginal worker if she was a manager is less than proportional. As a result, the marginal worker becomes manager. The aggregate output declines despite the fact that more members are now managers since the distortion affects the managers with high managerial talent harder than the managers with low managerial talent. The share of output accounted for by the managers with high managerial talent declines and the share of the output accounted for by the bottom managers increases. So, the restrictions on the size increases the number of plants, decreases the mean size of plants, and redistribute production from high talent to low talent managers.

6.2.2 The Case With Investment in Skills

I now study the effects of size-dependent distortions on various variables in the model and present some numerical results. I first set $\nu = 1$ and compare steady state equilibrium of the model under two values for τ: 0 and 0.05. The reason for choosing $\tau = 0.05$ is that the output per worker in the distorted economy is close to an undistorted economy with $A = 0.8$ and so the results in this section can be compared to the results in section (6.1).

The results are qualitatively similar to the ones in table (7). Once again the model conforms to the patterns observed in the data as shown in ***Panel A***. The main differences between the two tables is quantitative. Specifically, size-dependent distortions have dramatic impact on some statistics. The number of managers increases by 42%. As a result, the mean size of the establishments drop by 84%.

The effect on share of skilled managers is about 34% and the share of skilled managers among all managers drops by 80%.

Panel B shows the decline in output per capita by 30% which is expected since the distortions are modeled as output taxes. The skilled labor share declines by 85%, the managerial premium increases by 49% and the skill premium goes up by 76%. The effects on the quality of managers and the mean size of the establishments managed by different types of managers are summarized in ***Panel C***. The investment in skills declines by 44%. As a result, the average quality of all managers goes down by 80% and the average quality of skilled managers drops by 66%. The increase in the number of managers is almost solely due to the addition of unskilled managers which means that the average quality of unskilled managers decreases by 85%. The relative quality of a skilled manager to unskilled manager increases by 175%. The mean size of the plants run by skilled managers declines by 78% while the drop in the mean size is about 92% for the unskilled managers. The relative average establishment size of skilled to unskilled managers increases by 130% which means a very skewed size distribution for the model economy.

To understand the intuition behind these numbers, one should focus on the margins that are affected by the size distortions. Since some fraction of the output is taxed away, the household decides to select fewer members for schooling and invest in them less. This means that the average quality of the skilled managers would decline along with the share of skilled workforce. The choice of occupation would change in favor of household members with lower managerial talent since the distortions for members with higher managerial talents are higher. So the relative advantage of managers with high managerial talent (either intrinsic talent or augmented talent as a result of investment in skills) is diminished because the distortions tax away a higher share of their output relative to managers with lower levels of managerial talent. This is a reallocation of resources from larger to smaller plants in the model. As a result, the share of unskilled managers increases.

The share of skilled managers decreases since the outside option for them is to work for other managers and with more managers in the economy as a result of an influx of unskilled mangers, it is optimal for the household to select some of the previous skilled managers to become skilled workers. This margin turns out to be not so large and the main reason that there are less skilled managers is that, some of them are not selected for investment in skills in the distorted economy and become unskilled, but since their managerial talent is high, the household still selects them to become unskilled managers. This again emphasizes the importance of the choice of schooling for the choice of occupation.

Size distortions can provide an interpretation for the overall productivity in the model economy. Recall that for the results of this section, I set $A = 1$. This means that the managerial talent of a manager can be interpreted as the productivity level of the plant she is running. The productivity of the whole economy in the model can be defined as the average talent (quality) of all the managers, whether skilled or unskilled. This way of looking at the managerial talent provides a new source for cross-country productivity differences. In the model economy in this section, the average quality of managers is 75% lower when distortions are present. This is due to two margins. On the intensive margin, the skilled managers have less investment in them which decreases the average quality of the skilled managers. On the extensive margin, the addition of unskilled managers with lower managerial talent contributes to the overall decline in the average quality of managers. The two margins working together provide the sharp decline in the average managerial quality in the economy which can be interpreted as the lower productivity in the model economy.

The segregation of household members based on skill level and occupation provides more margins of change that is worth noting. In a simple setting without investment ins skills such as Guner *et al.* (2008), the introduction of size-dependent distortions results in a reallocation of resources from large to small plants through addition of more managers with low managerial talent. Therefore, the mean size of the establishments declines and the number of managers go up. The same pattern is present here, but selection and skill effects amplify the effect of distortions through lowering the share of skilled managers and investment in skills.

The results here can be contrasted to Guner *et al.* (2018) as well. In their model, the population consists of workers and managers. The workers are all the same and managers are heterogeneous based on their managerial talent[22] and they invest in their skills. The effect of distortions in my model is larger than theirs. Specifically, distortions increase the number of managers in both models, but their effects on the incentive of workers and managers are different in each model. The workers and managers in the current setting are heterogeneous based on skill level, so the outside option for a skilled worker, which is a skilled manager, is different from an unskilled worker, which is an unskilled managers. In Guner *et al.* (2018), the outside option for all managers is becoming a worker and having a fixed level of efficiency units. The fact that my model provides more heterogeneity in terms of skill level and occupation amplifies the effects of size distortion since these distortions are impacting different margins. These margins reveal themselves in selection and skill effects and result in different aggregate implications in my model compared to theirs. For example, the number of managers increase more in my model with $\tau = 0.05$ (it is less than 100 percent increase in their model) and the average quality of managers is lower in mine (80% lower versus 50%).

7 Matching Data

In this section, I use the model to replicate two statistics in every country in my sample: the GDP per worker and the share of unskilled working individuals. Then I study the model's implications for other statistics associated with the facts that I presented in section (3), namely managerial rate, share of skilled managers, managerial premium and skill premium. In particular, I calibrate the productivity

[22] They call it managerial ability.

term (A) and distortions parameter (τ) to match those moments in the data. The model generates other statistics and I can assess how the model performs regarding the unmatched moments.

Table (9) shows the results of this experiment. I compare three sections of the distribution of GDP per worker: top 10% to bottom 10%, top 20% to bottom 20% and top 35% to bottom 35%. The model match GDP per worker and the share of unskilled working individuals exactly and it also produces other statistics related to the facts that I documented,namely managerial rate, share of skilled managers among all managers, managerial premium and skill premium. Since the notion of size in the model is not the same as the data, I refrain from reporting this statistic, but I explained earlier that qualitatively the model generates the fact about mean establishment size and its relation to GDP per worker. I also presented three statistics that the model generates but I do not have reliable data to compare: average managerial quality, investment in skills as a share of GDP per worker and capital per worker.

The model reproduces the other moments in the data quite well. For example, the managerial rate in the model and the data are sufficiently close and the variation in managerial rate coming from the model is less than what I observe in the data. The same pattern is true for the share of skilled managers. The model can also produce the managerial premium quite well, although the variation generated by the model is more than what I find in the data. The most impressive statistic generated from the model is skill premium. This statistic in the model match the data quite well.

I should emphasize that the two moments related to relative income, i.e. managerial premium and skill premium, are quite complicated statistics. They contains all the decision rules about investment in skills and occupational choice as well as equilibrium prices and wages. There is nothing in the model parameters that insures that these statistics can match data closely. In other words, I only calibrate the model parameters to generate skill premium and managerial premium in the U.S. and there is no guarantee that if I match two moments other than these two, the model can match them sufficiently well. In fact, Cubas *et al.* (2016) tried to match the skill premium by matching the GDP per worker, the division of labor between skilled and unskilled and distribution of talents matching the PISA distribution in each country. The skill premium from their model is far from the data. This comparison shows that since investment in skills are affecting occupational choice in the current setup, coupled with the presence of size-dependent distortions are necessary to produce the massive variation in the skill premium observed in the data.

I take the results of this exercise as a success story for my current setup. It shows that in order to explain the five facts about economic development, one needs a rich, yet parsimonious setup such as the one I developed. The relevant margins of investment in skills and occupational choice with the presence of size-dependent distortions are enough to generate the large observed differences in several dimensions of economic development. In this regard, this paper adds a significant understanding to the economic choices that individuals in developing countries are making.

8 Discussion

In this section, I provide a discussion about the importance of different margins in the model. One margin is occupational choice and the other is investment in skills.

8.1 Importance of occupational choice

How important is the occupational choice? For one thing, if I shutdown the margin of occupational choice, I cannot relate to the facts presented about the managerial rate, the skilled managers and managerial premium. I also cannot study the effect of size-dependent distortions. The model is going to reduce to the model presented by Cubas *et al.* (2016). Their choice of parameters is different from mine and they identify the distribution of talents, which is a gamma distribution in their setup, by the mean and variance of the PISA test scores in each country. However, the investment in skills and other modeling assumptions are fairly similar. In the case with no occupational choice, the productivity term can be interpreted as total factor productivity (TFP).

Figure (8) illustrates the elasticity of output with respect to productivity term (A) when the curvature parameter ϕ in the technology of skill investments change from 0 to 0.6. There are three scenarios presented in the figure: no occupational choice, no size-dependent distortions and the model with occupational choice and size-dependent distortions. In the figure, the elasticity of output with respect to productivity term (A) increases from 1.5 to around 3 which is similar to what Cubas *et al.* (2016) found. This case can be contrasted to the scenario with occupational choice and the figure demonstrates that the elasticity is higher for every value of ϕ. The reason is that the investment in skills affecting the margin of occupation and each unit of investment in the skills of managers and workers is going to increase output from the production side in two ways: (1) the efficiency units of skilled workers will be higher and (2) the managerial talent of skilled managers will be higher as well. Together, they contribute to more production. Occupational choice has another obvious advantage that it helps me to understand the facts about managers and how the productivity differences can account for observed variations in the data through differences in productivity.

8.2 Importance of investment in skills

To understand the importance of investment in skills, I did the following experiment. I increased the value of curvature parameter in the production of skills (ϕ) from 0 to 0.25 and for each value of ϕ studied the effects of changes in productivity and size-dependent distortions on output per worker, skill premium and managerial premium. Table (10) presents the result of such experiment. When $\phi = 0$, the channel of skill investment is shutdown and $\phi = 0.17$ is the benchmark case. In the case of no skill investments, the output drops by 68%, skill premium by 44% and managerial premium by 42% when $A = 1$ and $\tau = 0$. This is the case where there is no size-dependent distortions.

Decline in productivity term and increase in size-dependent distortions is going to decrease the output further and increase in skill premium and managerial premium. But the changes in all these cases are less dramatic than the benchmark case with $\phi = 0.17$. As the value of ϕ increases the magnitude of changes in the three statistics increases. This illustrates that fixed changes in productivity and size-dependent distortions are magnified in the model when investment in skills is present and the marginal rate of return on investment in skills (ϕ) is higher. I previously showed that quantitatively, skill investments are necessary to account for the observed differences in GDP per worker in the data along with the statistics on managerial rate, share of skilled managers, skill premium and managerial premium. I thereby emphasize that to account for the observed differences in these statistics across countries, the channel of investment in skills and a reasonable value for the rate of return of this investment are needed.

I further calculate the elasticity of output with respect to exogenous productivity term A in the model when size-dependent distortions are present. I set the distortion parameter $\tau = 0.015$ since this is the average value for this parameter when I match the data in section (7). Figure (8) shows that as the marginal rate of investment in skills increases, the elasticity of output with respect to A increases, when the channel of occupational choice and size-dependent distortions are present. It also shows the output elasticity in case without distortions and without occupational choice. The output elasticity in the model is always higher when distortions are present and diverge from the other two cases as ϕ increases. Since the model matched data quite well, I take this evidence as showing the necessary role of investment in skills for accounting for observed differences in GDP per worker in the data.

Finally, I highlight the role that size-dependent distortions are playing in affecting the margins of skill investment and occupational choice. These distortions change the incentives for becoming a manager. As a result, the household does not invest in the skills of its members enough. So the distortions on the production side ultimately changes the incentive for skill investment. The resulting elasticity of output with respect to productivity shows that a distortion-free environment is not always enough to account for observed differences in income across countries.

9 Final remarks

The importance of investment in skills as a source of variation in the observed GDP per worker across countries was emphasized before in the macro-development literature. In this paper, I developed a parsimonious model to study investment in skills and occupational choice in the presence of size-dependent distortions as the central component of observed differences in managerial rate, share of skilled managers, managerial premium and skill premium. I showed that the model can successfully generate the statistics in the data and emphasized the important role of size-dependent distortions for accounting for the facts in the data. I abstracted from channels such as borrowing constraints, which are very important in developing countries, and organizational considerations in those countries about promotions and occupational choice which are important for the managerial occupations and left them for future research.

References

Akcigit, U., Alp, H. and Peters, M. (2016). *Lack of Selection and Limits to Delegation: Firm Dynamics in Developing Countries*. NBER Working Papers 21905, National Bureau of Economic Research, Inc.

Atkeson, A. and Burstein, A. T. (2011). *Aggregate Implications of Innovation Policy*. Working Paper 17493, National Bureau of Economic Research.

Bartelsman, E., Haltiwanger, J. and Scarpetta, S. (2013). Cross-country differences in productivity: The role of allocation and selection. *American Economic Review*, **103** (1), 305–34.

Benabou, R. (2002). Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica*, **70** (2), 481–517.

Bento, P. and Restuccia, D. (2017). Misallocation, establishment size, and productivity. *American Economic Journal: Macroeconomics*, **9** (3), 267–303.

Bhattacharya, D., Guner, N. and Ventura, G. (2013). Distortions, endogenous managerial skills and productivity differences. *Review of Economic Dynamics*, **16** (1), 11–25.

Bloom, N., Genakos, C., Sadun, R. and Reenen, J. V. (2011). *Management Practices Across Firms and Countries*. CEP Discussion Papers dp1109, Centre for Economic Performance, LSE.

—, Lemos, R., Sadun, R., Scur, D. and Van Reenen, J. (2014). Jeea-fbbva lecture 2013: The new empirical economics of management. *Journal of the European Economic Association*, **12** (4), 835–876.

— and Reenen, J. V. (2011). Human resource management and productivity. *Handbook of Labor Economics*, **4**, 1697 – 1767.

Caselli, F. (2005). Chapter 9-accounting for cross-country income differences. In P. Aghion and S. N. Durlauf (eds.), *Handbook of Economic Growth*, vol. 1, Part A, Elsevier, pp. 679 – 741.

— and Gennaioli, N. (2013). Dynastic management. *Economic Inquiry*, **51** (1), 971–996.

Cordoba, J. C. and Ripoll, M. (2013). What explains schooling differences across countries? *Journal of Monetary Economics*, **60** (2), 184 – 202.

Cubas, G., Ravikumar, B. and Ventura, G. (2016). Talent, labor quality, and economic development. *Review of Economic Dynamics*, **21**, 160 – 181.

Erosa, A., Koreshkova, T. and Restuccia, D. (2010). How important is human capital? a quantitative theory assessment of world income inequality. *The Review of Economic Studies*, **77** (4), 1421–1449.

European Union Statistics on Income and Living Conditions (EU-SILC) (2004-2016). European union statistics on income and living conditions (eu-silc).

Fernandez, R., Guner, N. and Knowles, J. (2005). Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality. *The Quarterly Journal of Economics*, **120** (1), 273–344.

Gabler, A. and Poschke, M. (2013). Experimentation by firms, distortions, and aggregate productivity. *Review of Economic Dynamics*, **16** (1), 26 – 38, special issue: Misallocation and Productivity.

García-Santana, M. and Ramos, R. (2015). Distortions and the size distribution of plants: evidence from cross-country data. *SERIEs*, **6** (3), 279–312.

Gollin, D. (2002). Getting income shares right. *Journal of Political Economy*, **110** (2), 458–474.

Gumbel, E. J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association*, **55** (292), 698–707.

Guner, N., Parkhomenko, A. and Ventura, G. (2018). Managers and productivity differences. *Review of Economic Dynamics*, **29**, 256 – 282.

—, Ventura, G. and Xu, Y. (2008). Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics*, **11** (4), 721 – 744.

Hall, R. E. and Jones, C. I. (1999). Why do some countries produce so much more output per worker than others? *The Quarterly Journal of Economics*, **114** (1), 83.

Heathcote, J., Storesletten, K. and Violante, G. L. (2010). The macroeconomic implications of rising wage inequality in the united states. *Journal of Political Economy*, **118** (4), 681–722.

—, — and — (2014). Consumption and labor supply with partial insurance: An analytical framework. *American Economic Review*, **104** (7), 2075–2126.

Hendricks, L. (2002). How important is human capital for development? evidence from immigrant earnings. *American Economic Review*, **92** (1), 198–219.

Hopenhayn, H. A. (2014). Firms, misallocation, and aggregate productivity: A review. *Annual Review of Economics*, **6** (1), 735–770.

Hsieh, C.-T. and Klenow, P. (2014). The life cycle of plants in india and mexico. *The Quarterly Journal of Economics*, **129** (3), 1035–1084.

— and Klenow, P. J. (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly Journal of Economics*, **124** (4), 1403.

— and Olken, B. A. (2014). The missing "missing middle". *Journal of Economic Perspectives*, **28** (3), 89–108.

Klenow, P. and Rodriguez-Clare, A. (1997). The Neoclassical Revival in Growth Economics: Has It Gone Too Far? In *NBER Macroeconomics Annual 1997, Volume 12*, NBER Chapters, National Bureau of Economic Research, Inc, pp. 73–114.

Lucas, R. E. (1978). On the size distribution of business firms. *The Bell Journal of Economics*, **9** (2), 508–523.

Manuelli, R. E. and Seshadri, A. (2014). Human capital and the wealth of nations. *American Economic Review*, **104** (9), 2736–62.

Minnesota Population Center (2018). Integrated public use microdata series, international: Version 7.0 [dataset].

Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer Series in Statistics, Springer.

OECD (2017). *Education at a Glance 2017*. The Organisation for Economic Co-operation and Development.

Poschke, M. (2017). The firm size distribution across countries and skill-biased change in entrepreneurial technology. *unpublished*.

Prescott, E. (1998). Needed: A theory of total factor productivity. *International Economic Review*, **39** (3), 525–51.

Restuccia, D. (2013). factor misallocation and development. In S. N. Durlauf and L. E. Blume (eds.), *The New Palgrave Dictionary of Economics*, Basingstoke: Palgrave Macmillan.

— and Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics*, **11** (4), 707 – 720.

— and — (2013). Misallocation and productivity. *Review of Economic Dynamics*, **16** (1), 1 – 10.

Roys, N. and Seshadri, A. (2017). Economic development and the organization of production.

Rubini, L. (2014). Innovation and the trade elasticity. *Journal of Monetary Economics*, **66**, 32 – 46.

Ruggles, S., Genadek, K., Goeken, R., Grover, J. and Sobek, M. (2017). Integrated public use microdata series: Version 7.0 [dataset]. *Minneapolis: University of Minnesota*.

Schoellman, T. (2012). Education quality and development accounting. *The Review of Economic Studies*, **79** (1), 388.

Appendix 1

Proof of Proposition 1. First, I solve the maximization problem of a manager with managerial talent z.

$$\max_{\{k,u,s\}} Az^{1-\gamma}\Big(k^{\alpha}(u^{\theta}s^{1-\theta})^{1-\alpha}\Big)^{\gamma} - Rk - W_u u - W_s s \tag{27}$$

The optimality conditions for input factors are:

- Capital

$$(\alpha\gamma)Az^{1-\gamma}k^{\alpha\gamma-1}u^{\theta(1-\alpha)\gamma}s^{(1-\theta)(1-\alpha)\gamma} = R$$

- Unskilled labor

$$\theta(1-\alpha)\gamma Az^{1-\gamma}k^{\alpha\gamma}u^{\theta(1-\alpha)\gamma-1}s^{(1-\theta)(1-\alpha)\gamma} = W_u \tag{28}$$

- Skilled labor

$$(1-\theta)(1-\alpha)\gamma Az^{1-\gamma}k^{\alpha\gamma}u^{\theta(1-\alpha)\gamma}s^{(1-\theta)(1-\alpha)\gamma-1} = W_s$$

Solving the above system of equations gives the following factor demand and profit functions:

- Capital

$$k(z,W_u,W_s,R) = z(A\theta(1-\alpha)\gamma)^{\frac{1}{1-\gamma}}\Big(\frac{\alpha}{\theta(1-\alpha)}\Big)^{\frac{1-\gamma(1-\alpha)}{1-\gamma}}\Big(\frac{1-\theta}{\theta}\Big)^{\frac{(1-\theta)(1-\alpha)\gamma}{1-\gamma}}\Big(\frac{1}{R}\Big)^{\frac{1-\gamma(1-\alpha)}{1-\gamma}}\Big(\frac{1}{W_u}\Big)^{\frac{\theta\gamma(1-\alpha)}{1-\gamma}}\Big(\frac{1}{W_s}\Big)^{\frac{(1-\theta)(1-\alpha)\gamma}{1-\gamma}} \tag{29}$$

- Unskilled labor

$$u(z,W_u,W_s,R) = z(A\theta(1-\alpha)\gamma)^{\frac{1}{1-\gamma}}\Big(\frac{\alpha}{\theta(1-\alpha)}\Big)^{\frac{\alpha\gamma}{1-\gamma}}\Big(\frac{1-\theta}{\theta}\Big)^{\frac{(1-\theta)(1-\alpha)\gamma}{1-\gamma}}\Big(\frac{1}{R}\Big)^{\frac{\alpha\gamma}{1-\gamma}}\Big(\frac{1}{W_u}\Big)^{\frac{1-\gamma+(1-\alpha)\gamma\theta}{1-\gamma}}\Big(\frac{1}{W_s}\Big)^{\frac{(1-\theta)(1-\alpha)\gamma}{1-\gamma}} \tag{30}$$

- Skilled labor

$$s(z,W_u,W_s,R) = z(A\theta(1-\alpha)\gamma)^{\frac{1}{1-\gamma}}\Big(\frac{\alpha}{\theta(1-\alpha)}\Big)^{\frac{\alpha\gamma}{1-\gamma}}\Big(\frac{1-\theta}{\theta}\Big)^{\frac{1-\gamma(\alpha+\theta-\alpha\theta)}{1-\gamma}}\Big(\frac{1}{R}\Big)^{\frac{\alpha\gamma}{1-\gamma}}\Big(\frac{1}{W_u}\Big)^{\frac{\theta\gamma(1-\alpha)}{1-\gamma}}\Big(\frac{1}{W_s}\Big)^{\frac{1-\gamma(\alpha+\theta-\alpha\theta)}{1-\gamma}} \tag{31}$$

- Profit

$$\pi(z,W_u,W_s,R) = \Omega\Big(\frac{1}{R}\Big)^{\frac{\alpha\gamma}{1-\gamma}}\Big(\frac{1}{W_u}\Big)^{\frac{\theta\gamma(1-\alpha)}{1-\gamma}}\Big(\frac{1}{W_s}\Big)^{\frac{(1-\alpha)(1-\theta)\gamma}{1-\gamma}} \tag{32}$$

where $\Omega = (1-\gamma)Az(A\theta(1-\alpha)\gamma)^{\frac{\gamma}{1-\gamma}}\Big(\frac{\alpha}{\theta(1-\alpha)}\Big)^{\frac{\alpha\gamma}{1-\gamma}}\Big(\frac{1-\theta}{\theta}\Big)^{\frac{\gamma(1-\alpha)(1-\theta)(1-\alpha\gamma(1-\theta))}{1-\gamma}}$

Based on the assumptions about the production technology parameters, all the demand and profit functions are strictly increasing in managerial talent and strictly decreasing in factor prices. The second part of the proposition is straightforward given that the above functions are linear in managerial talent z.

■

Appendix 2

The Lagrangian for the household problem is:

$$
\begin{aligned}
\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \Big\{ & L_t \ln(\frac{C_t}{L_t}) + \mu_t \Big[w_{u,t} \Big(U_{t-1} + N_t \int_0^{z_t^u} \int_0^{\hat{a}_t} a f(a,z) da dz \Big) + \\
& w_{s,t} \Big(S_{t-1} + N_{t-1} h_t \int_0^{z_t^s} \int_{\hat{a}_{t-1}}^{\bar{a}} a f(a,z) da dz \Big) + \\
& \Pi_{t-1}^u + N_t \int_{z_t^u}^{\bar{z}} \int_0^{\hat{a}_t} z \pi(W_{u,t}, W_{s,t}, R_t) f(a,z) da dz + \\
& \Pi_{t-1}^s + N_{t-1} h_t \int_{z_t^s}^{\bar{z}} \int_{\hat{a}_{t-1}}^{\bar{a}} z \pi(W_{u,t}, W_{s,t}, R_t) f(a,z) da dz + \\
& r_t K_t + (1-\delta) K_t - K_{t+1} - C_t - N_t x_t \int_0^{\bar{z}} \int_{\hat{a}_t}^{\bar{a}} f(a,z) da dz \Big] \Big\}
\end{aligned} \tag{33}
$$

The first order conditions with respect to $\{C_t, K_{t+1}, x_t, \hat{a}_t, z_t^u, z_t^s\}$ are equations (14), (15), (16), (17) and (19) with $\mu_t = \frac{1}{\frac{C_t}{L_t}}$. The market clearing conditions and aggregate feasibility are[23]:

- Unskilled labor market

$$
\frac{U}{L} = \int_0^{\hat{a}} \int_{z^u}^{\bar{z}} z u(W_u, W_s, R) f(a,z) da dz + (B x^{\phi}) \int_{\hat{a}}^{\bar{a}} \int_{z^s}^{\bar{z}} z u(W_u, W_s, R) f(a,z) da dz \tag{34}
$$

- Skilled labor market

$$
\frac{S}{L} = \int_0^{\hat{a}} \int_{z^u}^{\bar{z}} z s(W_u, W_s, R) f(a,z) da dz + (B x^{\phi}) \int_{\hat{a}}^{\bar{a}} \int_{z^s}^{\bar{z}} z s(W_u, W_s, R) f(a,z) da dz \tag{35}
$$

- Capital Market

$$
\frac{K}{L} = \int_0^{\hat{a}} \int_{z^u}^{\bar{z}} z k(W_u, W_s, R) f(a,z) da dz + (B x^{\phi}) \int_{\hat{a}}^{\bar{a}} \int_{z^s}^{\bar{z}} z k(W_u, W_s, R) f(a,z) da dz \tag{36}
$$

- Aggregate feasibility

$$
\frac{C}{L} + \delta \frac{K}{L} + \frac{g_L}{1+g_L} x \int_{\hat{a}}^{\bar{a}} \int_0^{\bar{z}} f(a,z) da dz = \frac{Y}{L} \tag{37}
$$

From the laws of motion for the stocks of unskilled and skilled labor, we can further simplify the market clearing conditions in the labor market:

$$
\begin{aligned}
& \frac{U}{L} \equiv \bar{u} = \int_0^{\hat{a}} \int_0^{z^u} a f(a,z) da dz = \\
& \int_0^{\hat{a}} \int_{z^u}^{\bar{z}} z u(W_u, W_s, R) f(a,z) da dz + (B x^{\phi}) \int_{\hat{a}}^{\bar{a}} \int_{z^s}^{\bar{z}} z u(W_u, W_s, R) f(a,z) da dz
\end{aligned} \tag{38}
$$

[23] I omitted the market clearing condition for the goods market since it will clear by Walras' law.

$$\frac{S}{L} \equiv \bar{s} = \frac{Bx^{\phi}}{1+g_L} \int_{\hat{a}}^{\bar{a}} \int_{0}^{z^s} af(a,z)dadz = \\ \int_{0}^{\hat{a}} \int_{z^u}^{\bar{z}} zs(W_u, W_s, R) f(a,z)dadz + (Bx^{\phi}) \int_{\hat{a}}^{\bar{a}} \int_{z^s}^{\bar{z}} zs(W_u, W_s, R) f(a,z)dadz \tag{39}$$

Equations (38) and (39) can be used to refine the initial guess for (W_u, W_s) pair. Equation (36) gives the capital per worker while equation (37) can be used to calculate consumption per worker. Aggregate production is calculated based on the equilibrium production of managers:

$$\frac{Y}{L} = \int_{0}^{\hat{a}} \int_{z^u}^{\bar{z}} zy(W_u, W_s, R) f(a,z)dadz + (Bx^{\phi}) \int_{\hat{a}}^{\bar{a}} \int_{z^s}^{\bar{z}} zy(W_u, W_s, R) f(a,z)dadz \tag{40}$$

where $y(W_u, W_s, R)$ is the production per managerial talent of a manager.

Appendix 3

The problem of a manager with managerial talent z facing size distortions is as follows:

$$\max_{\{k,u,s\}} \nu A^{1-\tau} z^{(1-\gamma)(1-\tau)} \Big(k^{\alpha}(u^{\theta} s^{1-\theta})^{1-\alpha}\Big)^{\gamma(1-\tau)} - Rk - W_u u - W_s s \tag{41}$$

The optimality conditions for input factors are:

- Capital

$$(\alpha\gamma(1-\tau))\nu A^{1-\tau} z^{(1-\gamma)(1-\tau)} k^{(\alpha\gamma)(1-\tau)-1} u^{\theta(1-\alpha)\gamma(1-\tau)} s^{(1-\theta)(1-\alpha)\gamma(1-\tau)} = R$$

- Unskilled labor

$$\theta(1-\alpha)\gamma(1-\tau)\nu A^{1-\tau} z^{(1-\gamma)(1-\tau)} k^{\alpha\gamma(1-\tau)} u^{\theta(1-\alpha)\gamma(1-\tau)-1} s^{(1-\theta)(1-\alpha)\gamma(1-\tau)} = W_u \tag{42}$$

- Skilled labor

$$(1-\theta)(1-\alpha)\gamma(1-\tau)\nu A^{1-\tau} z^{(1-\gamma)(1-\tau)} k^{\alpha\gamma(1-\tau)} u^{\theta(1-\alpha)\gamma(1-\tau)} s^{(1-\theta)(1-\alpha)\gamma(1-\tau)-1} = W_s$$

Solving the above system of equations gives the following factor demand and profit functions:

- Capital

$$\begin{aligned} &k(z, W_u, W_s, R) = \\ &\Phi z^{\frac{(1-\gamma)(1-\tau)}{1-\gamma(1-\tau)}} \left(\frac{\alpha}{\theta(1-\alpha)}\right)^{\frac{1-\gamma(1-\alpha)(1-\tau)}{1-\gamma(1-\tau)}} \left(\frac{1-\theta}{\theta}\right)^{\frac{(1-\theta)(1-\alpha)\gamma(1-\tau)}{1-\gamma(1-\tau)}} \left(\frac{1}{R}\right)^{\frac{1-\gamma(1-\alpha)(1-\tau)}{1-\gamma(1-\tau)}} \\ &\left(\frac{1}{W_u}\right)^{\frac{\theta\gamma(1-\alpha)(1-\tau)}{1-\gamma(1-\tau)}} \left(\frac{1}{W_s}\right)^{\frac{(1-\theta)(1-\alpha)\gamma(1-\tau)}{1-\gamma(1-\tau)}} \end{aligned} \tag{43}$$

- Unskilled labor

$$\begin{aligned} &u(z, W_u, W_s, R) = \\ &\Phi z^{\frac{(1-\gamma)(1-\tau)}{1-\gamma(1-\tau)}} \left(\frac{\alpha}{\theta(1-\alpha)}\right)^{\frac{\alpha\gamma(1-\tau)}{1-\gamma(1-\tau)}} \left(\frac{1-\theta}{\theta}\right)^{\frac{(1-\theta)(1-\alpha)\gamma(1-\tau)}{1-\gamma(1-\tau)}} \left(\frac{1}{R}\right)^{\frac{\alpha\gamma(1-\tau)}{1-\gamma(1-\tau)}} \\ &\left(\frac{1}{W_u}\right)^{\frac{1-\gamma(1-\tau)+(1-\alpha)\gamma\theta(1-\tau)}{1-\gamma(1-\tau)}} \left(\frac{1}{W_s}\right)^{\frac{(1-\theta)(1-\alpha)\gamma(1-\tau)}{1-\gamma(1-\tau)}} \end{aligned} \tag{44}$$

- Skilled labor

$$
\begin{aligned}
&s(z, W_u, W_s, R) = \\
&\Phi z^{\frac{(1-\gamma)(1-\tau)}{1-\gamma(1-\tau)}} \left(\frac{\alpha}{\theta(1-\alpha)}\right)^{\frac{\alpha\gamma(1-\tau)}{1-\gamma(1-\tau)}} \left(\frac{1-\theta}{\theta}\right)^{\frac{1-\gamma(\alpha+\theta-\alpha\theta)(1-\tau)}{1-\gamma(1-\tau)}} \left(\frac{1}{R}\right)^{\frac{\alpha\gamma(1-\tau)}{1-\gamma(1-\tau)}} \\
&\left(\frac{1}{W_u}\right)^{\frac{\theta\gamma(1-\alpha)(1-\tau)}{1-\gamma(1-\tau)}} \left(\frac{1}{W_s}\right)^{\frac{1-\gamma(\alpha+\theta-\alpha\theta)(1-\tau)}{1-\gamma(1-\tau)}}
\end{aligned} \tag{45}
$$

- Profit

$$
\pi(z, W_u, W_s, R) = \Omega \left(\frac{1}{R}\right)^{\frac{\alpha\gamma(1-\tau)}{1-\gamma(1-\tau)}} \left(\frac{1}{W_u}\right)^{\frac{\theta\gamma(1-\alpha)(1-\tau)}{1-\gamma(1-\tau)}} \left(\frac{1}{W_s}\right)^{\frac{(1-\alpha)(1-\theta)\gamma(1-\tau)}{1-\gamma(1-\tau)}} \tag{46}
$$

where

$$
\Phi = (A\theta(1-\alpha)\gamma(1-\tau))^{\frac{1}{1-\gamma(1-\tau)}}
$$

and

$$
\Omega = (1-\gamma(1-\tau))Az(A\theta(1-\alpha)\gamma(1-\tau))^{\frac{\gamma(1-\tau)}{1-\gamma(1-\tau)}} \left(\frac{\alpha}{\theta(1-\alpha)}\right)^{\frac{\alpha\gamma(1-\tau)}{1-\gamma(1-\tau)}} \left(\frac{1-\theta}{\theta}\right)^{\frac{\gamma(1-\alpha)(1-\theta)(1-\tau)(1-\alpha\gamma(1-\theta)(1-\tau))}{1-\gamma(1-\tau)}}
$$

.

Appendix 4

In this appendix, I explain in more details the data source and methods used to identify moments for calibration.

American Community Survey

The American Community Survey (ACS) is an ongoing survey that provides vital information on a yearly basis about our nation and its people. Information from the survey generates data that help determine how more than 675 billion dollars in federal and state funds are distributed each year. Through the ACS, we know more about jobs and occupations, educational attainment, veterans, whether people own or rent their home, and other topics. Public officials, planners, and entrepreneurs use this information to assess the past and plan the future. When you respond to the ACS, you are doing your part to help your community plan hospitals and schools, support school lunch programs, improve emergency services, build bridges, and inform businesses looking to add jobs and expand to new markets, and more.

One important fact to remember about the ACS is that the forms are not mailed to specific people, but rather to specific addresses. The sample is designed to ensure good geographic coverage and does not target individuals. By focusing on quality geographic coverage, the ACS can produce a good picture of the community's people and housing by surveying a representative sample of the population.The Census Bureau selects a random sample of addresses to be included in the ACS. Each address has about a 1-in-480 chance of being selected in a month, and no address should be selected more than once every 5 years. The Census Bureau mails questionnaires to approximately 295,000 addresses a month across the United States. This is a small number of households considering there are more than 180 million addresses in the United States and an address that receives ACS instructions will not likely find a neighbor or friend who has also received them.

The Census Bureau informs people living at an address that they have been selected to participate in the ACS. Shortly thereafter (for most U.S. addresses), instructions for completing the survey online are mailed. In Puerto Rico and some hard to reach areas in the U.S., only a paper questionnaire is mailed. Households are asked to complete the survey online or to mail the completed paper questionnaire back to the Census Bureau's National Processing Center in Jeffersonville, Indiana. If the Census Bureau does not receive a completed survey within a few weeks, it will mail an additional paper survey questionnaire.

Data Cleaning and Variable Selection

The process of cleaning the data and variables that are used in the regressions are outlined in this section.

-**Year**: The sample years are 2014-16. The reason for the small selection of years is that I do want to pick up as little as possible any employment effect resulting from the Great Recession. Hours of work, occupational standings, wages and worker composition may differ very much form year to year close to the recession and although I control the year effects with year-dummies, anything related to the business cycle is not relevant for my purposes here. I do a robustness check and run the regression just for the year 2015 and results are basically the same.

-**Hours**: The workers report their hours worked in the reference week in the survey. Part-time workers are those who work less that 35 hours per week and full-time workers are those who work 35 hours and more. I exclude all the part-time workers.

-**Age**: I focus on prime age, namely from 40 to 54. The reason is that my model does not have any implications for the changes in earnings, hours and occupation over the life-cycle and most of these transitions have died out for the prime age workers.

-**Class of Worker**: The workers are categorized by the ACS based on the nature of their job az "works for wages" and "self employed". This category based on a survey question that asks about the type of job with the following categories: a private for-profit company or business, or of an individual, for wages, salary, or commissions,a private non-for-profit, tax-exempt, or charitable organization, a local government employee (city, county, etc.), a state government employee, a Federal government employee, self-employed in own not incorporated business, professional practice, or farm, self-employed in own incorporated business, professional practice, or farm, working without pay in family business or farm. The workers who are not in the last two categories (self-employed in own incorporated business and working without pay in family) are listed as "works for wages". The self-employed workers who work in their own incorporated business are categorized as "self employed".

-**Labor Earnings**: The labor earnings in the ACS consists of wages, salary, commission, tips, pay-in-kind, or piece rates for a private, for-profit employer or a private not-for-profit, tax-exempt or charitable organization. These include private and salary workers and government employees. Self-employed people whose business was incorporated are included with private wage and salary workers because they are paid employees of their own companies. For the "self-employed" workers, roughly two-third of their profit can be attributed to wage earnings based on a simple aggregate neoclassical production function. This adjustment excludes the earning from capital. None of the earnings for workers include asset earnings of any kind such as housing, bonds and securities. All the earnings data were deflated by Consumer Price Index (CPI) from Bureau of Labor Statistics (BLS).[24]

I will also follow a procedure to eliminate badly incomplete or highly implausible observations from the sample following Heathcote *et al.* (2010) and Heathcote *et al.* (2014). Based on the hours worked and wage earnings, I can impute the hourly wage for the sample. I drop observations with positive income bu zero annual hours. Also, those with hourly wage income less than half of the

[24]https://www.bls.gov/cpi/data.htm.

federal minimum wage.[25] Finally I drop observations with more that $7 \times 12 = 84$ hours of work per week.

-Occupation: I categorize my sample into workers and managers as follows. First, all the self employed people are considered managers. Also, the people who work in broad categories of "Management Occupations", which corresponds to the 2010 Census Occupation codes of 0010-0430, are considered managers as well. The rest are categorized as workers.

-Education: The people in the sample are categorized based on educational level into two groups of skilled and unskilled. The unskilled people are those who finished at most the high school (grade 12). Any education more than high school is considered skilled. These include vocational and trainings, college dropout, college graduates and graduate level degrees among other forms of postsecondary education.

Regression Framework

I use a simple regression framework to estimate the relative earnings of skilled manager, unskilled managers, skilled workers and unskilled workers. Since the ACS data is not a panel, I cannot use the panel specifications. However, I can pool the ACS data for a couple of years, as repeated cross sections, and use year dummies to control for year fixed effects.

$$\begin{aligned} \log(\text{labor earnings}) = \alpha + \beta X_{it} + \xi_t + \delta_1 \mathbb{1}\{\text{skilled manager}\} + \delta_2 \mathbb{1}\{\text{unskilled manager}\} \\ + \delta_3 \mathbb{1}\{\text{skilled worker}\} + \epsilon_{it}. \end{aligned} \tag{47}$$

The vector X_{it} contains individual-specific controls such as log of hours, age, age-squared, gender, race, marital status, veteran status and detailed occupational codes. The inclusion of a variable for years of education is redundant since the skilled vs unskilled classification controls partially for years of education. The variables ξ_t are year fixed effects. I run the above regression under two specifications: One that includes self employed people in the sample and one that excludes them and only focus on managers with the 2010 Census codes classification. My preferred specification is the one that includes self-employed people. As a robustness check, I run the regression for the 2015 cross section and the whole sample from 2014 to 2016. The results stays basically the same.

2015 Cross Section

The result of the above regression without the year dummy for the cross-section of 2015 is summarized in table (1).

[25]The federal hourly minimum wage is 7.5 \$ in 2014 to 2016.

Table 1: Relative Earnings of managers and skilled workers to unskilled workers (2015 cross section)

	skilled manager	unskilled manager	skilled worker
with self-employed	1.91	1.07	1.61
without self-employed	2.22	1.40	1.55

The skilled workers and all managers have a premium over the unskilled workers in terms of earnings, no matter how I define earnings and occupation. Based on these premia, the other moments in the model are summarized in table (2). Given that the inclusion of self-employed people reduces the premia of all managers, it is reasonable that the skill premium and managerial premium increase in the absence of self-employed people as is the case in the first two columns of table (2). The share of managers is higher with the inclusion of self-employed people which is expected. The share of skilled managers among the pool of managers increases with the exclusion of self-employed people, suggesting that the majority of self-e,ployed people are unskilled. Finally, the share of unskilled labor, which is the sum of unskilled worker and managers, changes very little.

Table 2: Moments for calibration under different specifications (2015 cross section)

	skill premium	managerial premium	manager share	skilled manager share	unskilled labor
with self employed	1.70	1.14	0.17	0.77	0.40
without self employed	1.57	1.42	0.13	0.81	0.40

The above regression correctly gives the required moments. The only concern is that given the nature of a cross section, I cannot separate the effect of year fixed effect from the effect of age. Specifically, since I am only looking at prime age individuals, the effect of cohort is eliminated. What remains is the effect of the age of individuals, which I control for, and the effect of years. In order to see whether there is something special about the year 2015, I have to pool the ACS data from several years and run the regression(47) with year fixed effects and access the sensitivity of the results to the particular year of 2015. I present these results in the following section.

2014-2016 Pooled Cross Section

The following tables illustrate the results of running the regression (47) for the pooled cross section of ACS from 2014 to 2016 with year fixed effects. Comparing table (3) with table (1) reveals that the cross section and pooled data regressions result in fairly similar calculations. In other words, there is not much change in the cross section during 2014-2016. The same comparison is also true for tables (4) and (2).

Table 3: Relative Earnings of managers and skilled workers to unskilled workers (pooled data)

	skilled manager	unskilled manager	skilled worker
with self-employed	1.92	1.07	1.62
without self-employed	2.33	1.44	1.62

Table 4: Moments for calibration under different specifications (pooled data)

	skill premium	managerial premium	manager share	skilled manager share	unskilled labor
with self employed	1.71	1.14	0.17	0.77	0.40
without self employed	1.62	1.44	0.12	0.81	0.40

Appendix 5

Table 5: Parameter values

Parameter	Notation	Value
Discount factor	β	0.966
Population growth rate	g_L	0.009
Span-of-control parameter	γ	0.770
Capital Share	α	0.43
Depreciation Rate	δ	0.074
Unskilled labor share	θ	0.287
Skill efficiency parameter	B	0.998
Skill curvature parameter	ϕ	0.180
Schooling talent distribution	λ_a	9.724
Managerial talent distribution	λ_z	10.022
Correlation between talents	ρ	0.021

Note: The numbers are the values of the calibrated parameters for the benchmark economy. Discount factor and the depreciation rate are reported at the annual rate.

Table 6: Empirical targets: model and data (U.S.)

Statistic	Model	Data
Capital-to-output ratio	0.80	0.80
Investment rate	0.27	0.27
Fraction of unskilled labor	0.40	0.40
Expenditure per tertiary student (% GDP per worker)	0.25	0.25
Managerial rate	0.17	0.17
Share of skilled managers	0.77	0.77
Skill premium	1.71	1.71
Managerial premium	1.14	1.14
Income of skilled manager to unskilled worker	2.20	2.10
Income of unskilled manager to skilled worker	0.67	0.69

Note: The numbers are the values for the U.S. statistics used as moment targets as the benchmark economy. Capital-to-output ratio and the investment rate are reported at the annual rate.

Table 7: Results of a 20% reduction in exogenous productivity on steady state equilibrium in the model

		Productivity	
		$A = 1$	$A = 0.8$
Panel A	***Model versus Facts in the Data***		
Fact 1	Skilled labor Share	100	66
Fact 1	Share of skilled managers (among managers)	100	85
Fact 2	Managerial premium	100	108
Fact 3	Managerial rate	100	105
Fact 4	Mean Size	100	90
Fact 5	Skill premium	100	132
Panel B	***Output and Shares***		
	Output per Capita	100	88
	Unskilled Labor Share	100	106
	Unskilled Manager Share	100	122
	Skilled Labor Share	100	92
Panel C	***Quality and Size***		
	Investment in skills	100	89
	Average Quality of All Managers	100	72
	Average Quality of Skilled Managers	100	88
	Average Quality of Unskilled Managers	100	58
	Relative Average Quality of Skilled to Unskilled Managers	100	156
	Average Plant Size of Skilled Managers	100	99
	Average Plant Size of Unskilled Managers	100	47
	Relative Average Plant Size of Skilled to Unskilled Managers	100	214

Note: All the shares are reported as the relevant shares in the workforce, i.e. managers plus workers. I also indicate the shares that are calculated only among managers.

Table 8: Effects of size-dependent distortions on steady state equilibrium in the model

		Size-dependent Distortions	
Panel A	***Model versus Facts in the Data***	$\tau = 0$	$\tau = 0.05$
Fact 1	Skilled labor share	100	60
Fact 1	Share of Skilled Managers (among managers)	100	20
Fact 2	Managerial premium	100	149
Fact 3	Managerial rate	100	142
Fact 4	Mean Size	100	16
Fact 5	Skill premium	100	177
Panel B	***Output and Shares***		
	Output per Capita	100	33
	Unskilled Manager Share	100	300
	Skilled Labor Share	100	15
Panel C	***Quality and Size***		
	Investment in skills	100	56
	Average Quality of All Managers	100	20
	Average Quality of Skilled Managers	100	34.5
	Average Quality of Unskilled Managers	100	14.5
	Relative Average Quality of Skilled to Unskilled Managers	100	276
	Average Plant Size of Skilled Managers	100	22
	Average Plant Size of Unskilled Managers	100	9.6
	Relative Average Plant Size of Skilled to Unskilled Managers	100	230

Note: All the shares are reported as the relevant shares in the workforce, i.e. managers plus workers. I also indicate the shares that are calculated only among managers.

Table 9: Model implications for cross-country comparison

	Top 10% / Bottom 10%	Top 20% / Bottom 20%	Top 35% / Bottom 35%
-Output			
Data & Model	10.50	6.65	4.99
-Share of Unskilled			
Data & Model	0.80	0.84	0.88
-Managerial rate			
Data	4.21	3.24	2.31
Model	4.03	3.11	2.09
-Skilled managers			
Data	5.05	2.98	2.09
Model	4.89	2.71	1.94
-Managerial Premium			
Data	0.48	0.53	0.61
Model	0.55	0.68	0.81
-Skill premium			
Data	0.66	0.75	0.81
Model	0.63	0.71	0.79
Average managerial quality	3.05	2.44	2.04
Investment in skills	1.86	1.61	1.42
Capital per worker	8.55	6.47	3.16

Note: Two parameters, productivity (A) and distortions (τ), are calibrated to match two moments in each country: GDP per worker and the share of unskilled. I compare each statistic in the data for three sections of the distribution of GDP per worker: top 10% to bottom 10%, top 20% to bottom 20% and top 35% to bottom 35%. I report three more statistics, average managerial quality, investment in skills and capital per worker. I do not have data on these statistics and I report what the model generates.

Table 10: The importance of investment in skills and complementarities between productivity and size-dependent distortions

		$A = 1$ $\tau = 0$	$A = 0.9$ $\tau = 0$	$A = 1$ $\tau = 0.01$	$A = 0.8$ $\tau = 0.01$
$\phi = 0$	Y/L	32	31	28	25
	Skill premium	56	61	64	71
	Managerial Premium	68	72	76	83
$\phi = 0.05$	Y/L	69	64	59	44
	Skill premium	73	62	54	41
	Managerial Premium	76	81	85	91
$\phi = 0.1$	Y/L	75	68	62	53
	Skill premium	86	79	71	57
	Managerial Premium	85	89	92	98
$\phi = 0.17$	Y/L	100	94	89	71
	Skill premium	100	118	135	169
	Managerial Premium	100	108	116	129
$\phi = 0.25$	Y/L	119	102	87	61
	Skill premium	129	144	168	184
	Managerial Premium	112	121	132	144

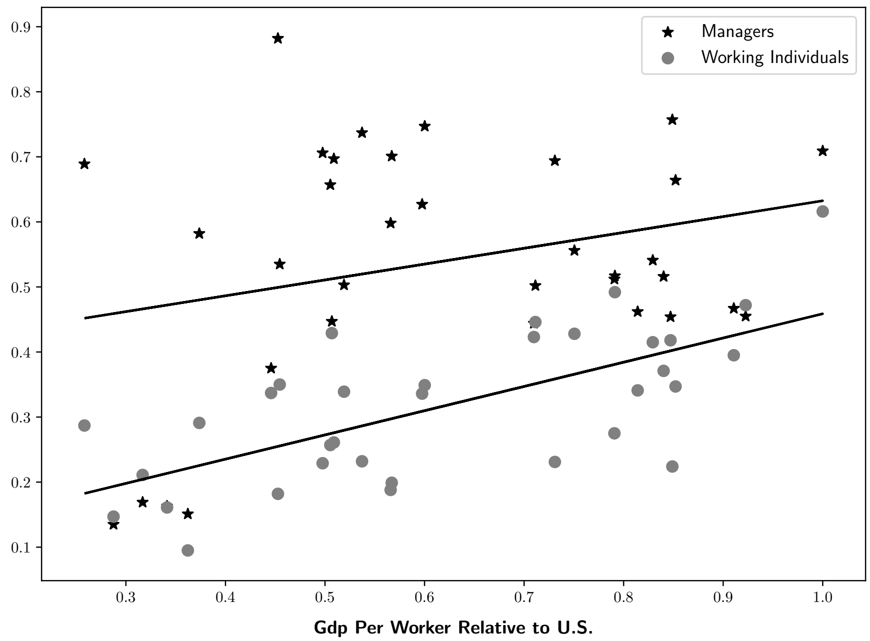
Table 11: Data sources

Country	Source	Sample size	Year
Austria	(EU-SILC)	3177	2016
Belgium	(EU-SILC)	2806	2016
Bulgaria	(EU-SILC)	4015	2016
Switzerland	(EU-SILC)	4313	2016
Cyprus	(EU-SILC)	1919	2016
Czech Republic	(EU-SILC)	2815	2016
Germany	(EU-SILC)	6962	2016
Denmark	(EU-SILC)	1904	2016
Estonia	(EU-SILC)	3512	2016
Greece	(EU-SILC)	7806	2016
Spain	(EU-SILC)	8044	2016
Finland	(EU-SILC)	3038	2016
France	(EU-SILC)	6104	2016
Croatia	(EU-SILC)	3338	2016
Hungary	(EU-SILC)	3991	2016
Ireland	(EU-SILC)	2439	2016
Italy	(EU-SILC)	10927	2016
Lithuania	(EU-SILC)	2735	2016
Latvia	(EU-SILC)	2986	2016
Malta	(EU-SILC)	2032	2016
Netherlands	(EU-SILC)	3856	2016
Norway	(EU-SILC)	2140	2016
Poland	(EU-SILC)	5640	2016
Portugal	(EU-SILC)	5756	2016
Romania	(EU-SILC)	3960	2016
Serbia	(EU-SILC)	2429	2016
Sweden	(EU-SILC)	1668	2016
Slovenia	(EU-SILC)	2240	2016
Slovakia	(EU-SILC)	3780	2016
United Kingdom	(EU-SILC)	4656	2016
United States	IPUMS-USA	445485	2016
Canada	IPUMS-International	31413	2010
South Africa	IPUMS-International	76191	2007
Brazil	IPUMS-International	1281770	2010
Puerto Rico	IPUMS-International	5046	2010
Panama	IPUMS-International	47815	2010
Mexico	IPUMS-International	1030948	2010

Note: The data for Eurpopean countries are from European Union Statistics on Income and Living Conditions (EU-SILC). The IPUMS-International uses the national statistics agencies as follows: Canada: Statistics Canada, South Africa: Statistics South Africa, Brazil: Institute of Geography and Statistics, Puerto Rico: U.S. Bureau of the Census, Panama: Census and Statistics Directorate, Mexico: National Institute of Statistics, Geography, and Informatics.

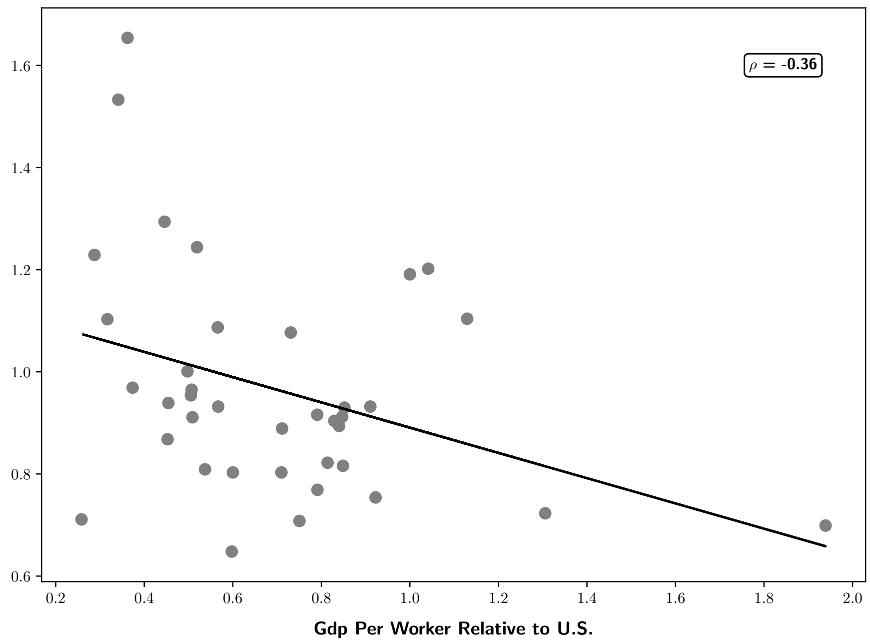
Appendix 6

Figure 2: **Educational Attainment of Managers and Working Individuals**



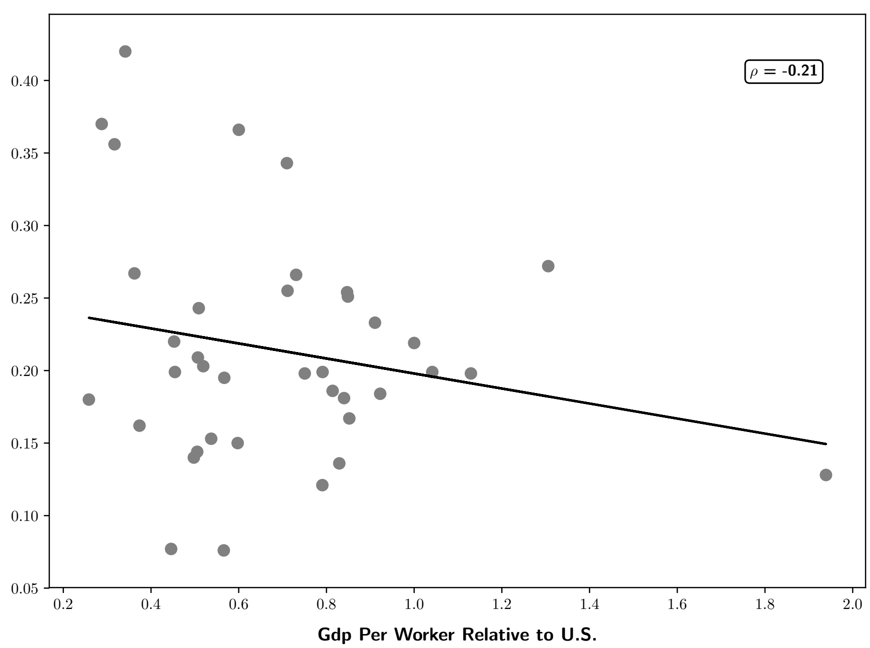
Notes: The y-axis is the percentage of managers and working individuals (age 25-64) who has at least a high school degree. The stars represent managers and the dots represent population. The correlation coefficient is 0.59 for managers and 0.67 for working individuals.
Source: Author's calculation based on individual-level survey data.

Figure 3: **Managerial Premium**



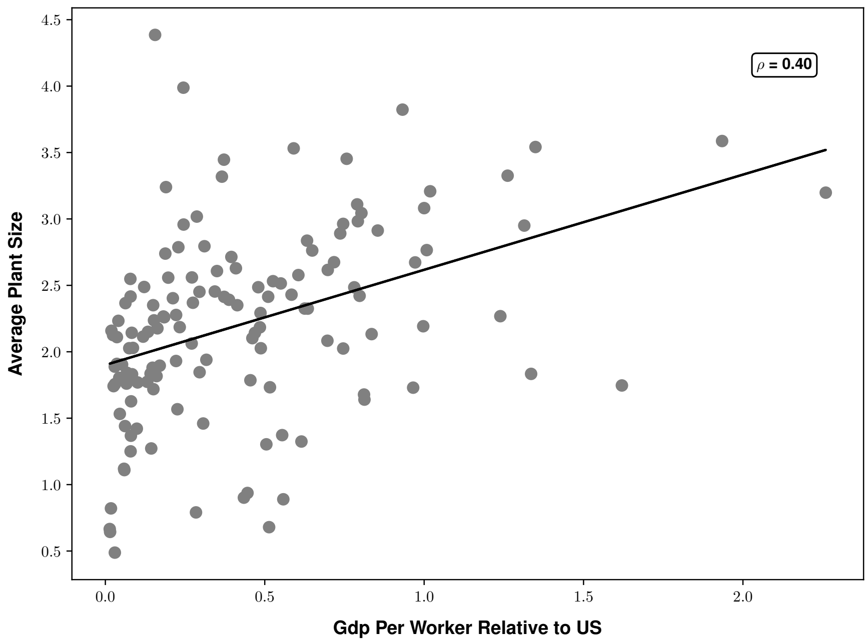
Notes: The y-axis is the relative income of managers to non-managers.
Source: Author's calculation based on individual-level survey data.

Figure 4: **Managerial Rate**



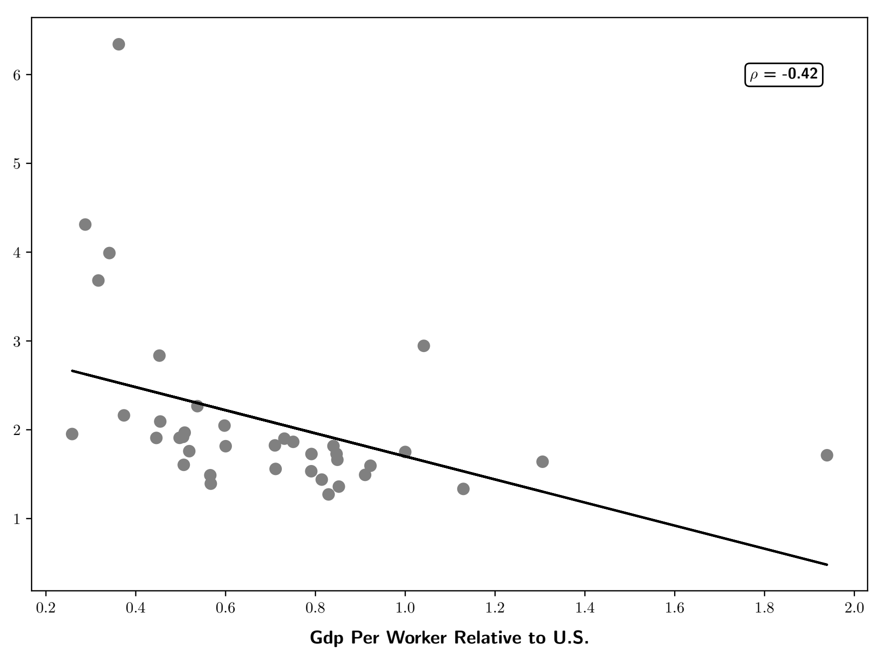
Notes: The y-axis is the share of managers in working individuals (age 25-64).
Source: Author's calculation based on individual-level survey data.

Figure 5: **Mean Establishment Size**



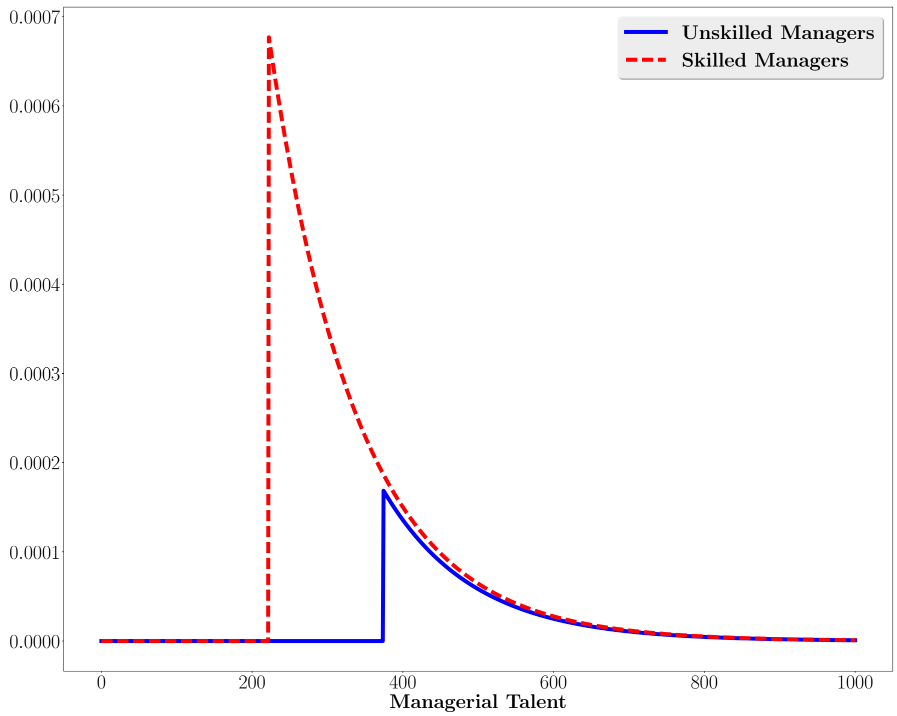
Notes: The y-axis is the mean establishment size of the manufacturing plants in logs.
Source: Bento and Restuccia (2017).

Figure 6: **Skill Premium**



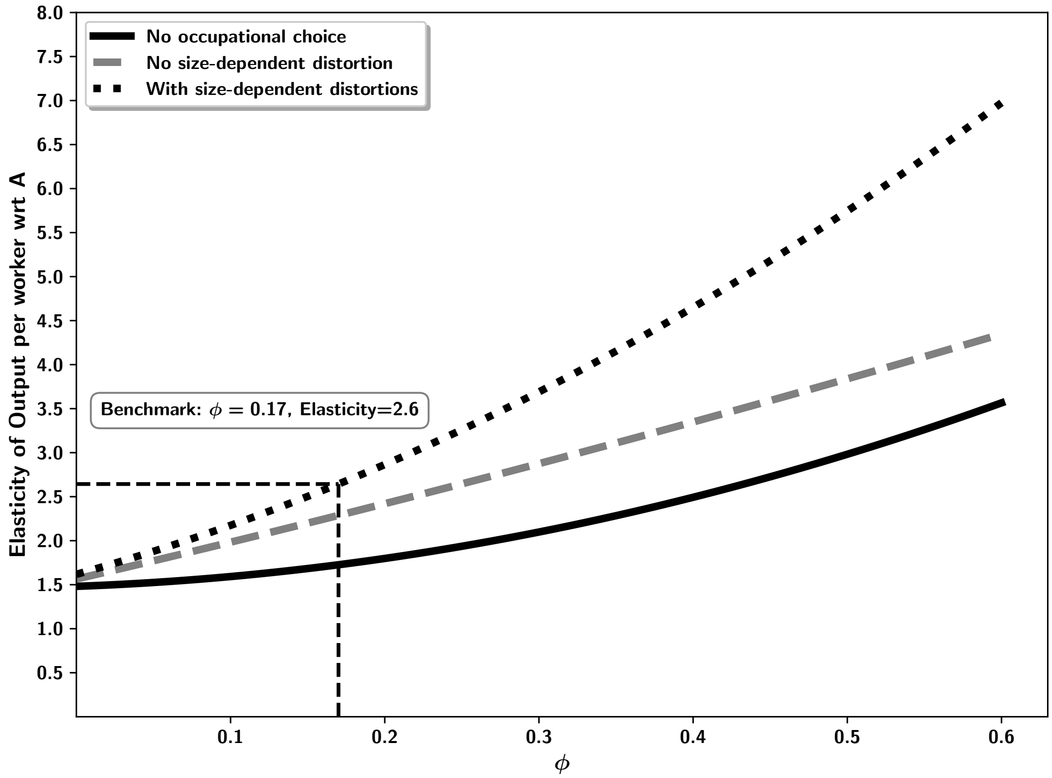
Notes: The y-axis is the relative income of skilled to unskilled working individuals.
Source: Author's calculation based on individual-level survey data.

Figure 7: **Steady State Equilibrium Distribution of Managers in the Benchmark Economy**



Notes: The average managerial talent of unskilled managers is 46% lower than the average mangerial talent of skilled managers.

Figure 8: **Amplification from skill investment and occupational choice**



Notes: The x-axis shows different values for the returns to investment in skills. The y-axis is elasticity of output with respect to productivity term (A). The distortion parameter τ in the case with size-dependent distortions is set to 0.015.