Automation and Top Income Inequality

Ömer Faruk Koru*

University of Pennsylvania

June 8, 2020

Abstract

Top income inequality has been increasing in the US. Hence, the Pareto parameter associated with the top income distribution is decreasing. In this paper, we provide a theory that links automation technology to the Pareto parameter of the top income distribution. We construct a model in which the span of control is defined by the measure of labor used in production. We model this as a convex cost of labor. This convex cost generates a decreasing returns to scale production function. An improvement in automation enables entrepreneurs to substitute labor with capital and decreases the severity of diseconomies of scale. This leads to higher returns to entrepreneurial skills, a decrease in the Pareto parameter, and an increase in top income inequality. We rationalize the convex cost of labor using a theory of efficiency wages. Using cross-industry and cross-country data, we show that there is a significant correlation between automation and top income inequality.

JEL classification: E23, J23, J3, O33.

Keywords: automation, top income inequality, entrepreneurship, efficiency wage, superstars, Pareto distribution, span of control, labor share.

---

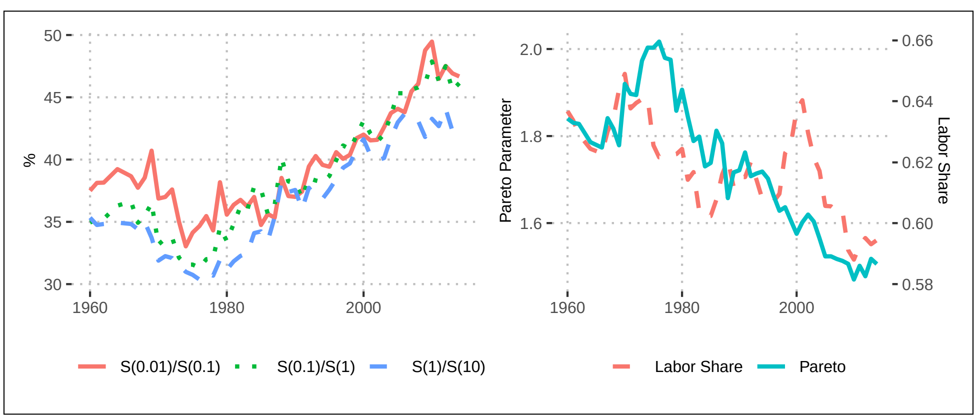
*Email: okoru@sas.upenn.edu. The author thanks Guido Menzio, Kenneth Burdett, Dirk Krueger, Harold Cole, Pascual Restrepo, Joachim Hubmer, Görkem Bostancı, Tomas Larroucau, Sergio Villalvazo Martin, Mustafa Doğan, Gabrielle Vasey and participants of the UPenn Macro Lucnh Seminar, Midwest Macro Meeting Spring 2019, NASMES 2019, GCER 2019, GW Student Research Conference 2019.

1 Introduction

Income inequality among the top earners, measured as the ratio of the top 0.1% income share to the top 1% income share, has been increasing in the US for the last forty years. In other words, the income gap between the super-rich and rich individuals is increasing. During the same time period, there was a significant improvement in technology, specifically in automation. We define automation in a broad sense, including information technology. In this paper, we provide a theory that links the automation level to top income inequality.

The left panel of the figure 1 plots the relative income share at the top income distribution. The solid line shows the ratio of the income share of the top 0.01% to 0.1%. The dotted line shows the ratio of the income share of the top 0.1% to 1%. The dashed line shows the ratio of the income share of the top 1% to top 10%. The figure shows that top income inequality has been increasing since the 1980s. It is true whether we define the top income as the top 10%, the top 1%, or the top 0.1%. It can be seen from the figure that top income inequality has increased by almost half.

Figure 1: Top Income Inequality and the Labor Share



Note: *Left panel: the solid line shows the ratio of income share of top* 0.01% *to* 0.1%*. The dotted line shows the ratio of income share of top* 0.1% *to* 1%*. The dashed line shows the ratio of income share of t op* 1% *to top* 10%*. Right panel: The solid line shows the Pareto parameter implied by the relative income share of top* 0.1% *to top* 1%*. The dashed line is the labor share of income (right axis).*
Source: Relative income shares: World Inequality Database; Labor share: Penn World Table 9.1.

It is well-known that top income distribution of income is well approximated by a Pareto distribution.[1] An implication of the Pareto distribution is that the relative income share is a function of the Pareto parameter. The increase in top income inequality implies that the Pareto parameter of top income distribution is decreasing. Our model links the Pareto parameter to automation technology. In the model, we use a production function similar to Acemoglu & Restrepo (2018b). An implication of this production function is that the labor share of income is a function of the automation level, which is also true in our model. In other words, the labor share of income provides us with a measure of automation in the economy. Therefore, we plot the labor share and the Pareto parameter on the right panel of figure 1. It illustrates that there is a high correlation between automation and top income inequality.

There are well-established theories explaining why the right tail of the income distribution is approximated by a Pareto distribution. We build our theory on the diseconomies of scale argument in Lucas (1978) and Rosen (1981). In these models, top income inequality depends on the severity of diseconomies of scale. As the severity of diseconomies of scale decreases, top skilled entrepreneurs scale up their production and increase their market share. Hence inequality at the top percentile rises. However, in these models, the decreasing returns to scale parameter is exogenously given. In this paper, we endogenize this parameter and show how it changes with automation.

The idea that technological innovation allows firms to scale up their production is not new; it goes back to Rosen (1981). However, the main models of technological progress have no implication for the change in the scalability. Technological improvement is usually modeled as either increase in the productivity of some factor (for example Acemoglu (2002)), change in capital share (for example Acemoglu & Restrepo (2018b)) or decrease in the price of capital (for example Autor & Dorn (2013)). None of these affect the decreasing returns to scale of the production function, hence the change in the Pareto parameter.

In our model, the reason for diseconomies of scale is information friction. In the model, an entrepreneur does not know whether his employee is working or shirking. In order to provide his

[1] The CDF of a Pareto distribution with scale parameter c and shape parameter (Pareto parameter) λ is given by $F(x) = 1 - (c/x)^{\lambda}$.

employees with an incentive to exert effort, an entrepreneur needs to spend additional resources, such as investing in monitoring technology or paying efficiency wages to workers. If the cost of additional resources is convexly increasing with the number of workers, then the profit function exhibits decreasing returns to scale. First, we provide the result for any convex cost function in labor, and then we show that efficiency wage theory provides a microfoundation for this convex cost.

We define the level of automation as the share of tasks that can be produced by capital, as in Acemoglu & Restrepo (2018b). An entrepreneur needs to complete a set of tasks (such as designing, engineering, accounting, etc.) in order to produce the final good. While some of these tasks can be automated (i.e., it can be produced by capital), some of them can only be produced by labor. As automation technology improves, dependency on labor decreases and convexity of production cost decreases. Therefore, the scalability problem is associated with the level of automation technology. Our main result links the Pareto parameter of the right tail of the income distribution to the skill distribution in the population, automation technology, and the severity of the convexity of the monitoring cost. We show that as automation technology improves, inequality at the top rises.

To see the mechanism, consider the extreme case: none of the tasks can be automated. In such a case, the cost of production is the price of labor plus the monitoring cost, which is convex. Therefore, the top skilled entrepreneur could only serve a portion of the market, which enables lower-skilled entrepreneurs to enter the market. Now, consider the other extreme in which any task can be automated. In such a case, the only cost of production is the price of capital, which is linear. Therefore, an entrepreneur has no problem scaling up his output, hence the top skilled individual captures the entire market. While in the first scenario there is lower inequality thanks to the existence of other entrepreneurs, in the second scenario there is perfect inequality since the top talented entrepreneur owns the entire market. When the economy converges from one extreme to the other, thanks to automation, the income inequality at the top increases.

The model implies that the automation level can be measured by the capital share of income and the Pareto parameter of the top income distribution is proportional to the labor share of income. We test our model's prediction using two different data. First, we look at the cross-industry cross-time variation of the labor share and the Pareto parameter in the US. Second, we look at the cross-country, cross-industry variation in OECD countries. Our regression results support the model's prediction.

Our model is static and automation technology is exogenously given. We show the impact of an exogenous change in automation technology. Although the model is static, the main result applies to a dynamic model in which the only dynamic choice is capital accumulation and the entrepreneur's problem is static since the result does not depend on the capital stock or prices.

Related Literature: This paper is related to several strands of literature. First, we contribute to the literature on the impact of automation on the labor market outcome (Acemoglu & Restrepo, in press; Autor & Dorn, 2013; Goos et al., 2014; Hémous & Olsen, 2018; Moll et al., 2019). This literature mainly focuses on wage inequality between high and low skill workers. In this paper, we look at the inequality among high skilled individuals. The closest article to the current study is Moll et al. (2019). Here, they study the impact of automation on income and wealth distribution. In their model, automation gives rise to higher returns to wealth which eventually leads to thicker income and wealth distribution. Our mechanism is different from this argument. Here, we focus on the increase in entrepreneurial income which is an important part of the increase in top income inequality (Guvenen & Kaplan, 2017; Smith et al., 2019). In our model, automation impacts top income inequality through the increase in return to entrepreneurial skills. Smith et al. (2019) show that entrepreneurs are an important fraction of top income earners and their skill is an integral part of their firm's performance. In this regard, we believe that the change in the return to entrepreneurial skills is important to understand the dynamics of top income inequality.

Second, there is a growing literature about the determinants of top income inequality and the change in top income inequality. Gabaix & Landier (2008) and Tervio (2008) explore the connection between the change in the firm size distribution and an increase in CEO compensation using assignment models. However, the Pareto parameter in those models is constant, whereas we are interested in the change in the Pareto parameter. Several other articles study the impact of the decrease in the top marginal tax rate on the share of top income percentile (Piketty et al., 2014; Kim, 2015; Aoki & Nirei, 2017). Aghion et al. (2018) and Jones & Kim (2018) show that innovation and creative destruction are important factors for top income inequality. Geerolf (2017) shows that change in information technology can increase top income inequality in the knowledge-based hierarchies model of Garicano (2000) and Garicano & Rossi-Hansberg (2006).

This paper focuses on the change in the return to entrepreneurial skill as the main driver of top income inequality. We contribute to the literature on innovation and top income inequality by providing a mechanism that can explain increase in innovation. An alternative interpretation of entrepreneurial skill in our model is the quality of an idea. We do not consider innovation in our model, though in a richer model where entrepreneurs exert effort to increase their productivity or the quality of their idea, we believe that it can be shown that an increase in the return to entrepreneurial skill leads to higher the return to innovation, hence more innovation.

The third strand of literature considers the impact of change in factor's share of income on inequality. Piketty (2014) argues that capital income is more concentrated than the labor, hence an increase in capital income share leads to higher inequality. Bengtsson & Waldenström (2018) shows that there is a positive relationship between capital share in national income and income share of the top 1%. In our model, the increase in capital income share leads to an increase in top income inequality. However, the reason for this is not that top income owners are the owners of capital, but because automation increases the return to entrepreneurial skill. Indeed, since the 1960s, the share of business income inside the top 0.1% almost doubled (Piketty & Saez, 2003).

In another related paper, Dogan & Yildirim (2017) study the impact of automation on compensation schemes of workers. In their model, replacing labor with capital leads to a reduction in peer monitoring, hence firms change the compensation scheme in order to incentivize workers to exert effort. In this paper, we also consider the monitoring problem of workers. However, our main focus is on top income inequality.

The structure of the paper is as follows: Section 2 presents the reduced form model and the main results. Section 3 tests the model prediction. Section 4 provides a microfoundation for the model in section 2. And section 5 concludes.

2 The Model

We consider a static model economy. In order to understand the impact of improvement in automation technology, we characterize top income distribution and we consider the comparative statistic with respect to the automation level.

There is a unit mass of individuals, each endowed with two types of skill: labor and entrepreneurial. The labor skill is the same for all individuals, whereas the entrepreneurial skill, denoted by z, is distributed with some cumulative distribution function G with support $[z_{min}, z_{max}] \subset \mathbb{R}_+$. There is a fixed amount of capital stock in the economy, owned by individuals.[2]

Each individual can either become a worker or an entrepreneur. If an individual becomes a worker, he supplies labor inelastically and earns wage w. If he becomes an entrepreneur, he rents capital and hires labor in order to produce output and enjoy a profit, $\pi(z)$, which is determined in equilibrium. Individuals choose their occupations to maximize their income.

2.1 The Entrepreneur's Problem

Each entrepreneur has access to production technology. We use a task-based framework similar to Zeira (1998) and Acemoglu & Restrepo (2018b). In order to produce a unique final good, an entrepreneur needs to complete a measure one of tasks, $i \in [0, 1]$. There is no market for tasks, hence the entrepreneur needs to complete all of the tasks inside the firm.[3]

[2]Since the model is static, the distribution of capital is not important.

[3]Assume transportation cost is high enough so that no one wants to trade tasks.

Tasks are complements and they are aggregated into output by a unit elastic aggregator (i.e., Cobb-Douglas):

$$lnY = \int_0^1 lny(i)di, \tag{1}$$

where Y is the total output, and $y(i)$ is the level of task i used in the production.

Given a task i, capital and labor are perfect substitutes. However, there is a technological constraint on the usage of capital. Some of the tasks are not technologically automated, meaning that they cannot be produced by capital. There is an automation technology frontier I such that task $i \leq I$ can be produced either by capital or by labor, while task $i > I$ can be produced only by labor. Formally, the production function for task i is:

$$y(i) = \begin{cases} k_i + \gamma_i \ell_i & \text{if } i \leq I, \\ \gamma_i \ell_i & \text{if } i > I, \end{cases} \tag{2}$$

where k_i and ℓ_i denote capital and labor, γ_i is the productivity of labor in task i. We assume that γ_i is increasing in i.[4] In other words, i denotes the complexity of the task, and labor has a comparative advantage relative to capital in high-index tasks.

Since capital and labor are perfect substitutes, only one of them is going to be used to produce a task. In a sense, automation is labor replacing. Once a task is automated, capital might replace labor for that task. Because γ_i is increasing, it is optimal to automate (i.e. produce by using capital) the low-index tasks first. In other words, if it is optimal to automate task i, then it is optimal to automate task $j < i$. Let $I^\star \leq I$ be the automation decision of the entrepreneur,

[4] For simplicity, we assume that capital has the same productivity for each task, which is normalized to 1. However, as long as the ratio of labor productivity to capital productivity is increasing the following analysis holds.

so that any $i < I^\star$ is automated.[5] Then, by combining (1) and (2), the output is:

$$lnY = \int_0^{I^\star} lnk_s ds + \int_{I^\star}^{1} ln\left(\gamma_i \ell_i\right) di. \tag{3}$$

Apart from the technological constraint, there is another difference between labor and capital: the entrepreneur has limited ability to manage the labor. As the employment size increases, the entrepreneur losses control over the labor. The usage of capital does not affect the span of control of the entrepreneur, only the measure of labor affects it. We represent the loss of control as a cost paid by the entrepreneur. In order to sustain control, he needs to spend additional resources. Let $v(\int_{I^\star}^{1} \ell_i di)$ denote this cost and assume that it is strictly increasing and convex: $v' > 0$, $v'' > 0$. Moreover, we assume that $v(0) = 0$ and $v'(0) = 0$. We discuss the interpretation of this additional cost in the next sub-section.

The entrepreneur's objective is to maximize profit. He decides which tasks are to be automated, $I^\star$, how much capital to hire for each automated task, k_s for $s < I^\star$, and how much labor to hire for tasks that are not automated, ℓ_i for $i \geq I^\star$. Formally, the entrepreneur's problem is:

$$\pi(z) = \max_{\substack{I^\star,\{\ell_i\}_{i\in[I^\star,1]},\\ \{k_s\}_{s\in[0,I^\star)}}} zY - w\int_{I^\star}^{1} \ell_i di - v\left(\int_{I^\star}^{1} \ell_i di\right) - R\int_0^{I^\star} k_s ds \tag{4}$$

$$s.t. \quad 0 \leq I^\star \leq I,$$

$$\ell_i \geq 0, \;\; k_s \geq 0.$$

where z is an entrepreneurial skill, w is the wage rate and R is the rental rate of capital.

[5] We assume that the least productive tasks can be automated. However, Autor & Dorn (2013) argue that it is the middle-skilled jobs that are more prone to automation. For the more general cases, suppose M is the set of tasks that can be automated. For the main result of this paper, the only important parameter is the measure of tasks that cannot be automated, $1 - |M|$. For ease of interpretation and mathematical computation, we assume the set of automated tasks is connected, $M = [I_{min}, I)$. For simplicity, let $I_{min} = 0$, since it is always optimal to start automating the least productive task.

Our main mechanism works through the convex cost of labor, v. This additional convex cost makes the profit function decreasing returns to scale. Because the production function, zY, is constant returns to scale and v is convex, $zY - v$ is decreasing returns to scale. If every task is automated, $I^\star = 1$, then the production function is constant returns to scale. If there is no automation technology, $I^\star = 0$, then the model is similar to the span-of-control model of Lucas (1978). Hence, the level of automation determines the severity of diseconomies of scale.

2.1.1 Interpretation of v

The main mechanism of this paper depends on the convex of v. Therefore, we discuss what v represents. The convex cost v represents, in a reduced form, the loss of control over labor. The idea of v is that labor has an additional cost and this cost is convexly increasing with employment size. There are several interpretations of this additional cost of labor.

One interpretation of v is the monitoring cost. Since workers can shirk, the entrepreneur needs to spend additional resources to prevent workers from shirking as in the efficiency wage theory (Shapiro & Stiglitz, 1984; Calvo, 1985). If the probability of monitoring is inversely related to the labor force, that leads to a convex cost of the labor force. Since capital cannot shirk, the size of capital does not affect the monitoring cost. Therefore v only depends on the labor force.

There are other possible interpretations of v. For example, v can be thought of as convex hiring and firing cost (Hopenhayn, 1992). Since there is no friction in the capital market, the cost of capital is just the price. Another interpretation might be the problem-solving cost of the entrepreneur (Garicano, 2000). If automation means that the task has a well-defined objective now and capital can solve the problem by itself, then only labor encounters problems that he cannot solve by himself. He asks about these problems to the entrepreneur, which costs the entrepreneur time and effort to solve those problems.

The direct evidence of the convex cost of labor is firm-size-wage-premium. Large firms pay higher wages than smaller firms, even after controlling for worker heterogeneity (Oi & Idson, 1999). If the wage rate depends on the employment size, then firms face with convex cost in labor.

Lemma 1. *If wage rate $w(L)$ is strictly increasing in L and positive, then $w(L)L$ is strictly convex.*

Here, we interpret v as the monitoring cost. In section 4, we provide a micro foundation for v using the efficiency wage theory, which leads to firm size wage premium. However, since the efficiency wage leads to wage distribution for workers and we are only interested in top income inequality, for the tractability of the model we start with the reduced form model.[6]

2.2 The Equilibrium

Now, we are in a position to define an equilibrium.

Definition 1. *For a given automation technology I, skill distribution G with support $[z_{min}, z_{max}]$ and capital stock $\bar{K}$ an equilibrium consists of prices $\{R, w\}$, the set of entrepreneurs $E \subset [z_{min}, z_{max}]$, labor and capital demand $\{\ell_i^\star(z)\}_{i\in[I^\star,1]}, \{k_s^\star(z)\}_{s\in[0,I^\star)}$, automation technology $I^\star(z)$ for $z \in E$ such that:*

- *$\pi(z) \geq w$ for all $z \in E$;*
- *$\{\ell_i^\star(z)\}_{i\in[I^\star,1]}, \{k_s^\star(z)\}_{s\in[0,I^\star)}, I^\star(z)$ solves the entrepreneur's problem (4);*
- *Labor market clears: $\int\limits_E \int\limits_{I^\star(z)}^{1} \ell_i^\star(z) di dG(z) = 1 - |E|$;*
- *Capital market clears: $\int\limits_E \int\limits_0^{I^\star(z)} k_s^\star(z) ds dG(z) = \bar{K}$.*

Proposition 1. *For a given automation technology $0 < I < 1$, capital stock $\bar{K}$, and skill distribution G with support $[z_{min}, z_{max}] \subset \mathbb{R}_+$, there exists a unique equilibrium.*

[6] What happens to v depends on the interpretation: it can be a part of the compensation scheme for labor, or it can be an effort cost incurred by the entrepreneur. In the reduced-form model, we assume v is incurred by the entrepreneur, in the model with efficiency wage v/L is paid to labor as compensation to not shirk; however, our result does not depend on what happens to v.

2.3 Characterization of the Equilibrium

We left the details of the characterization of the equilibrium to the appendix. Here, we point out the main features.

2.3.1 Optimal Occupational Choice

It is easy to see that profit $\pi(z)$ is increasing in z, hence there is a cutoff $z^\star$ such that any individual with $z > z^\star$ becomes entrepreneur and others become a worker.

2.3.2 Optimal Allocation of Capital and Labor of an Entrepreneur

An entrepreneur uses the same measure of labor in non-automated tasks and the same measure of capital in automated tasks. To see this, consider the first-order conditions of (4) with respect to ℓ_i and k_s:

$$[\ell_i] : \frac{zY}{\ell_i} = w + v'\left(\int_{I^\star}^{1} \ell_i di\right) \implies \ell_i = \ell_j = \ell \ \forall i, j \geq I^\star. \tag{5a}$$

$$[k_s] : \frac{zY}{k_s} = R \implies k_s = k_t = k \ \forall s, t < I^\star. \tag{5b}$$

The first condition equates the marginal product of labor in task i to the marginal cost of labor. Since marginal cost is the same for each task that is not automated, marginal products must be equalized across tasks. Hence, this condition implies that the measure of labor used in each task that is not automated is the same. Similarly, the second condition implies that the capital used for each task that is automated is the same. For automated tasks, this is easy to see. Since there is no productivity difference between the tasks, an entrepreneur should be indifferent to allocating resources to each task, therefore he distributes the capital across tasks uniformly. This is also true for labor because of unit elasticity of substitution between tasks. Unit elasticity leads to the productivity of labor to be in multiplicative form. Once the automation level is fixed, labor productivity behaves as if it is total factor productivity. Formally, effective TFP becomes $zC(I^\star)$, where $C(I^\star) = exp\left(\int_{I^\star}^{1} ln\gamma_i di\right)$. Hence the productivity level of a task affects each task in the same way, and optimal labor is the same across non-automated tasks.

Optimal solution to the entrepreneur's problem induces the output to Cobb-Douglas looking function: $zC(I^\star)k^{I^\star}\ell^{1-I^\star}$.

2.3.3 Optimal Automation Level of an Entrepreneur

Now we characterize the optimal automation level of an entrepreneur. Taking the first order condition of 4 with respect to $I^\star$ and imposing optimality condition for labor and capital leads to the following equation:

$$\int_{I^\star}^{1} ln\gamma_i di - (1 - I^\star) ln\gamma_{I^\star} + ln(z) = ln(R). \tag{6}$$

The solution to this equation is the unconstrained optimal automation level, $I^\star$. The entrepreneur chooses $I^\star$ if it is less than automation constraint I, otherwise, he chooses I. The solution only depends on the rental rate of capital R and the skill of entrepreneur, z. Wage rate does not affect the automation level, but it affects the labor and capital decision, that is why it impacts the level of production indirectly. Let $\tilde{I}(z)$ be the solution to (6).

One thing to notice is that for low productive tasks, using labor might never be optimal. Consider a low productive task i such that $w/\gamma_i > R$. Even without any labor, the effective cost of labor is higher than the capital. Therefore, the entrepreneur does not have any incentive not to automate this task. So, all tasks $i < \underline{I} := max\{0, \gamma^{-1}(w/R)\}$ are automated in the equilibrium, where $\gamma^{-1}(x)$ is the task that has the labor productivity x.

Proposition 2. *Optimal choice of automation level, $I^\star(z)$ is increasing in z and given by:*

$$I^\star(z) = \begin{cases} I & \text{if } z \geq \tilde{z}, \\ \tilde{I}(z) & \text{if } \underline{z} < z < \tilde{z}, \\ \underline{I} & \text{if } z \leq \underline{z}. \end{cases} \tag{7}$$

where $\tilde{I}(\tilde{z}) = I$ and $\tilde{I}(\underline{z}) = \underline{I}$.

Proposition 2 tells that more skilled entrepreneurs automate more tasks. The reason is as follows: as discussed above, labor productivity appears like a TFP in the optimal production, $C(I^\star)$. Hence, there is a tradeoff for automation. On the one hand, automation enables entrepreneurs to use cheaper factor. On the other hand, it decreases the productivity gain from the labor, $C(I)$. Low productive entrepreneurs automate less to benefit from total productivity. As z increases, the benefit of labor productivity decreases, hence the entrepreneurs prefer the cost-effective inputs. Therefore, $I^\star$ is increasing in z. The reason is similar to the argument of Zeira (1998). In a similar model, he studies technological adaptation across countries. He shows that low productive countries have lower wages and hence lower technological adoption. In our model, the wage rate is the same for all firms, the only difference is productivity.

2.4 Top Income Distribution

Now, we can characterize the top income distribution. Individuals with skill level below $z^\star$ become a worker and earn wage w and individuals with skill level above $z^\star$ become entrepreneurs and earn profit $\pi(z)$. Because $\pi(z) \geq w$ for entrepreneurs, the top income percentile consists of entrepreneurs. As a result, we only need to characterize the profit function for top income distribution.

Recall from the first order conditions that $zY^\star = k^\star R$ and $zY^\star = \ell^\star(w + v'(L^\star))$. If we multiply the first one with $I^\star$ and the second one with $(1 - I^\star)$ and sum them, we get:

$$zY^\star = RI^\star k^\star + L^\star(w + v'(L^\star)).$$

Hence, the profit function is given by:

$$\begin{aligned}\pi(z) =& zY^\star - RI^\star k^\star - wL^\star - v(L^\star) \\ =& v'(L^\star)L^\star - v(L^\star).\end{aligned}$$

In order to get a closed form solution, we need more structure. Assume that $v(L) = L^\alpha$, where $\alpha > 1$. Then:

$$\pi(z) = (\alpha - 1)L^{\star\alpha}.$$

Imposing functional form of v into entrepreneur's problem gives us:

$$L(z) = \left[\left(\frac{zC(I^\star)}{R^{I^\star}}\right)^{\frac{1}{1-I^\star}} - w\right]^{\frac{1}{\alpha-1}} \left(\frac{1}{\alpha}\right)^{\frac{1}{\alpha-1}}.$$

Because we are interested in top income, consider high z. Suppose that $\bar{z} > R$, then clearly automation technology binds for top skilled entrepreneurs. To see this, consider $z > R$. If he automates all tasks, then he has a linear production function and makes infinite profit. Hence, I binds for high enough z.

By plugging labor demand and $I^\star = I$ into the profit function, we get:

$$\pi(z) = (\alpha - 1)\left[\left(\frac{zC(I)}{R^I}\right)^{\frac{1}{1-I}} - w\right]^{\frac{\alpha}{\alpha-1}} \left(\frac{1}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}. \tag{8}$$

The profit function is convex in z. $I \in (0, 1)$ implies that $1/(1 - I) > 1$. Similarly, $\alpha > 1$

implies $\alpha/(\alpha-1) > 1$. Convexity implies that there is a superstar effect (Rosen, 1981): the profit is increasing in z disproportionately. A high productive entrepreneur's earning is much higher than a low productive one.

Observe that the convexity of profit function is increasing with automation technology I. The reason for this is that automation constraint in the entrepreneur's problem binds stronger for high skilled than low skilled. Hence, once this constraint is relaxed, the return is higher for a high skilled entrepreneur. For simplicity, consider two entrepreneurs, one with high z so that automation technology binds and one with low z that does not automation all automatable tasks. An increase in I does not affect the choice of low z, whereas now a high z can enlarge its production and increase its profit. Therefore, the value of relaxing the automation constraint is increasing at z. This implies that an improvement in automation technology increases the convexity of the profit function.

The convexity of profit function is also increasing with a reduction in the monitoring cost, i.e. a decrease in α. The monitoring cost is the main reason behind the decreasing returns to scale. As the monitoring problem is relaxed, entrepreneurs can enlarge their span of control. Since the enlargement is bigger for high productive entrepreneurs, this leads to an increase in the convexity of the profit function.

To characterize the distribution of profits, we need to know how productivity z is distributed. It is well-known that top income distribution of income is well approximated by a Pareto distribution. Moreover, the power of a Pareto distribution is also a Pareto distribution. Therefore, if z is Pareto distributed, then the convexity of profit function implies that the distribution of profit has a Pareto tail.

Proposition 3. *Suppose the distribution of entrepreneurial productivity, z, is Pareto with shape parameter λ, monitoring cost function is $v(L) = L^\alpha$, and $\lambda(1-I)(\alpha-1) > 1$.*[7] *Then, the distribution of profits has a Pareto tail with shape parameter $\lambda(1-I)\frac{\alpha-1}{\alpha}$.*[8]

[7]Proposition 1 can be extended to any unbounded distributions as long as the labor demand remains finite. For the Pareto distribution, we need $\lambda(1-I)(\alpha-1) > 1$ to have an equilibrium.

[8]We say that the tail distribution of F is distributed by G if $F(x)/G(x) \to 1$ as $x \to \infty$. Observe that including

Pareto parameter gives us a measure for inequality (Gabaix, 2016). Lower tail parameter means higher inequality. In this model, the Pareto parameter of the profits has three components: entrepreneurial skill distribution (λ), the convexity of the labor cost function (α), and the automation technology (I). Since $(\alpha-1)/\alpha < 1$, both automation and labor cost make income distribution thicker. As I increases or α decreases, the Pareto parameter goes down, hence inequality at the top increases.

Here, decreasing returns comes from automation technology and labor cost. As we discussed before, the convexity of profit function increases with improvement in I and decreases with α. This leads top skilled entrepreneurs to capture the higher share of total profits. As the severity of diseconomies of scale decreases, top income inequality increases.

Under the assumption of the constant returns to scale profit function, only the highest skilled entrepreneur produces and others would work for her since there is no limit to scaling the production function. If everything can be automated, i.e. $I = 1$, or there is no convex cost of labor, i.e. $\alpha = 1$, then profit function is constant returns to scale. In such a situation there is no limit for entrepreneurs to scale up their production, hence only the most productive person becomes an entrepreneur. Hence as I goes up, inequality also increases because the limit on scaling up the production lessened.

2.5 A Measure for Automation: Labor Share

In our model, $1 - I$ corresponds to the labor share of income. To see this, consider the entrepreneur's first order condition with respect to capital. It implies that the capital share of production within a firm is I, $I^\star(z)zY = RI^\star(z)k^\star(z)$. If there is no entrepreneur that automation level does not bind, then in the aggregate, capital share of income is I:

$$\int I^\star(z)zY(z)dG(z) = \int RI^\star(z)k^\star(z)dG(z) \implies I = \frac{R\bar{K}}{\int zY(z)dG(z)}.$$

capital income, RK, does not impact the tail of income distribution of entrepreneurs.

Therefore, the remaining part, $1-I$ accrues to labor and entrepreneur.

Proposition 4. *If automation technology binds for every entrepreneur, then the labor share of income (wage share + entrepreneurial share) is* $1-I$.

2.6 Back of the Envelope Calculation

Now, we do back of the envelope calculation for the impact of the improvement in automation technology on the change in the Pareto parameter of the top income distribution. Recall that our theory states that $\beta = \lambda(1-I)\frac{\alpha-1}{\alpha}$. We know that the labor share in the US decreased from around 64% to 59% from the 1970s to 2010s. Assuming the Pareto parameter for skill distribution and the convexity of monitoring cost function has not changed, i.e. all the decrease comes from the change in I, this implies that:

$$\frac{\hat{\beta}_{2010}}{\hat{\beta}_{1970}} = \frac{1-I_{2010}}{1-I_{1970}} \approx 0.92.$$

In other words, the model predicts 8% decrease in the Pareto parameter. In the WID data, the estimated Pareto parameter decrease from 2 to 1.5 in the same period. This corresponds to approximately 25% decrease. In other words, our model can explain a third of the decrease in the Pareto parameter.

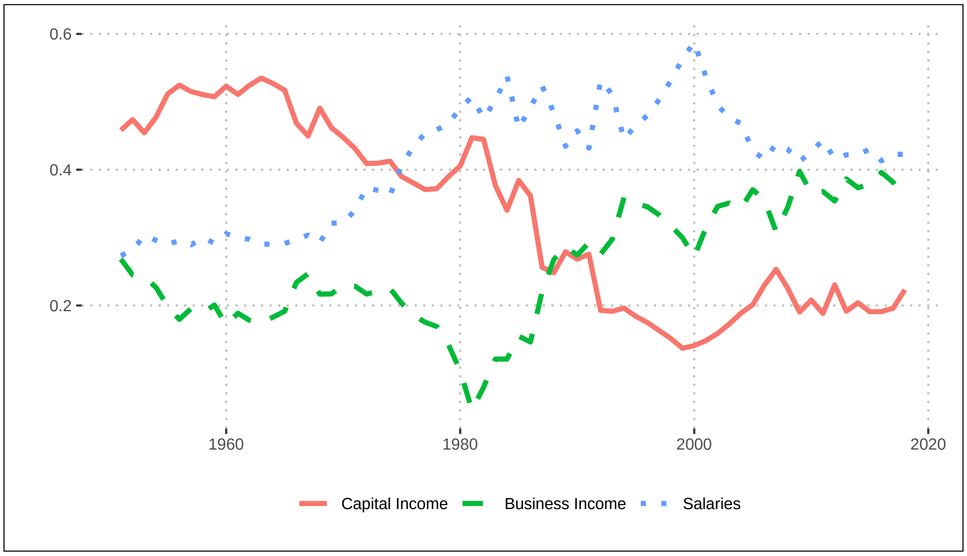
3 Test for Empirical Predictions

3.1 Evidence Justifying Assumptions of the Model

In this subsection, we discuss the importance of the change in the return to entrepreneurial skill for the dynamics of top income inequality.

Figure 2 shows the income composition of the top 0.1% (excluding capital gains) across

Figure 2: Income Composition of Top 0.1%



Note: *Capital gains are excluded. Source: Piketty & Saez (2003)*

the last 50 years. The share of business income (dashed line in the graph) almost doubled since the 1960s. Together with wages and salaries, they account for 80% of the income of top income earners (Piketty & Saez, 2003; Atkinson & Lakner, 2017). Moreover, the major component of the increase in the top income share can be accounted for by the increase in business income (Guvenen & Kaplan, 2017; Smith et al., 2019; Bakija et al., 2012). 60% of the growth of income share of top earners can be accounted for by managers, executives, entrepreneurs, supervisors, and financial professionals (Bakija et al., 2012). Therefore, the change in the return to entrepreneurial skills is the main driver of the top income inequality.

3.2 Testing Model Predictions

In this subsection, we test the main theory of the paper. The implication of our main result is that

$$log\beta = log\lambda + log(1 - I) + log(\alpha - 1) - log(\alpha).$$

This implies that there is a one-to-one relationship between the Pareto parameter and the labor share, $1-I$. The model predicts that a percentage increase in $1-I$ leads to a percentage increase of β.

We test this prediction with two different cases. First, we consider the industry-level panel data for the US. Second, we consider the country-level panel data. To test our theory, we regress estimate the following equation:

$$\Delta log\beta_{it} = \gamma\Delta log(\text{labor share}_t) + \tau + \Delta\epsilon_{it},$$

where i is industry or country, t is time, τ is time trend and Δ is the first difference operator. Under the assumption that skill distribution and monitoring cost remains constant across time, our theory predicts that γ is equal to one.

3.2.1 Measure for Labor Share

Labor share of income is defined as the total compensation of workers divided by the total income. For the US industrial level data, we use the compensation of employees as a share of the value-added GDP for each industry, using BEA's industry-level GDP data (U.S. Bureau of Economic Analysis, 2018). For international level analysis, we use the labor share estimates of Penn World Table 9.1 (Feenstra et al., 2015). BEA's data starts in 1987, therefore we consider the years between 1987 and 2005. For international comparison, we consider 1961-2005.

Observe that labor share in the model consists of two parts: wage and entrepreneurial income. Though, in reality, accounting for self-employment income is not straightforward since it is hard to distinguish what fraction of income is returned to entrepreneurial skill and what fraction is returned to own capital. How to incorporate self-employment income into factor share calculations is an important discussion (Gollin, 2002). PWT divides self-employment income between labor and capital using the split ratio in the corporate sector (Feenstra et al., 2015).

3.2.2 Measure for Top Income Inequality

The main source of top income shares is the World Inequality Database (WID)[9]. WID relies on the tax data and available for a wide range of countries. In the international level evidence, we use data from WID. Specifically, we use pre-tax income (equally split between spouses) shares for top 0.1% and top 1%.

We get relative income share from WID and then estimate the Pareto parameter using the following relation:

$$\hat{RIS} = 10^{\frac{1-\hat{\beta}}{\hat{\beta}}} \times 100.$$

Unfortunately, the tax data usually does not have information on the industry, making it hard to get a good estimate of top income share by industry. Therefore, we use the CPS ASEC microdata extracted from IPUMS (Flood et al., 2018). As we discussed before, business income is an important component of income for top income earners and our model is also about business income. For this reason, we considered the distribution of income of self-employment workers. Assuming that the right tail of income distribution follows a Pareto distribution, we estimated the Pareto parameter using the maximum likelihood estimator.

A major drawback of public use micro data is that the income is top coded in the data. Since we are interested in the right tail of the income distribution, a significant fraction of the observations is top-coded. In order to estimate Pareto parameter with a top-coded data, we follow the strategy of Clemens et al. (2017). Let x_i be the income of person i and let $\bar{x}$ be the top code. Observed income in the data is then:

$$\tilde{x}_i = \begin{cases} x_i & \text{if } x_i \leq \bar{x}, \\ \bar{x} & \text{if } x_i > \bar{x}. \end{cases}$$

[9] We retrieved data from https://wid.world/.

Assume that income distribution after q^{th} percentile is distributed by a Pareto with shape parameter β. Maximum likelihood estimator for scale parameter is q^{th} percentile of the data. Let's denote it by x_q. Then maximum likelihood estimator for the Pareto parameter is

$$\hat{\beta} = argmax \ \ \Pi \left[\left(\frac{\beta x_q^{\beta}}{x_i^{\beta+1}} \right) \right]^{D_i} \left(\frac{x_q}{\bar{x}} \right)^{\beta(1-D_i)},$$

where D_i indicated whether $x_i \leq \bar{x}$ or not. Solution to this problem is

$$\frac{1}{\hat{\beta}} = \frac{1}{N_{unc}} \left[\sum ln \left(\frac{x_i}{x_q} \right) + N_{cen} ln \left(\frac{\bar{x}}{x_q} \right) \right],$$

where N_{unc} is the number of uncensored observations and N_{cen} is the number of censored observations.

One problem with this strategy is that there is not always enough observation to consistently estimate the Pareto parameter. We only estimated the Pareto parameter if there are more than 15 observations and we fit the Pareto distribution to distribution after top $0.85th$ percentile. In total, we have 648 estimated parameter for 18 industries between the years 1970-2005.[10]

We take the 5-year averages to decrease the short-run fluctuations and to reduce the noise of the data. And because BEA starts in 1987, we end up having 4 observations for each industry.

3.2.3 Result

Table 1 shows the regression results. As can be seen, all of the coefficients are positive and significant. This means that there is a positive correlation between the labor share and the Pareto parameter, both at the industry level and at the country level. Moreover, when we control for time trends, coefficients are very close to 1, as predicted by the model. A percentage increase in the labor share increases the Pareto parameter by one percent. In this regard, the data supports the main prediction of the model.

[10] We exclude utilities, public administration and other services.

Table 1: Results

	log(Pareto Parameter)			
	US - Industry Level		Cross-country	
log(Labor Share)	1.002	1.02	1.33	0.99
	(0.57)	(0.6)	(0.37)	(0.44)
Time Trend	-	0.01	-	-0.03
	-	(0.03)	-	(0.01)
Nobs	46	46	198	198

Note: All standard errors clustered at industry or country level. Rows are independent variables and columns are dependent variables. The Pareto parameter for industries are estimated by fitting Pareto parameter to top business income distribution using CPS ASEC data. The Pareto parameter for countries are taken from World Inequality Database. The labor share for industries are taken from BEA. The labor share for countries are taken from Penn World Table version 9.1.

3.3 Beyond Top Income Inequality

In this section, we discuss our model's implications other than top income inequality.

CEO compensation: Even though our model is about entrepreneurs, it is possible to consider entrepreneurs as CEOs also. In an extended model, if we assume that there is a competitive market for CEOs and firms are competing to hire CEOs, then Bertrand competition among firms leads the CEO to capture all the surplus. In this regard, the model predicts that an improvement in automation technology leads to an increase in CEO compensation. Indeed, there is also a significant increase in CEO compensation in the US, especially between the mid-1970s and the 2000s(Frydman & Jenter, 2010). Median CEO compensation was $1.2 million in the 1970s and it increased to $9.2 million in the 2000s.

Market concentration: Our model predicts that market concentration, either measured as top firms' share in sales or in employment, increases with automation. Even though we did not prove formally, it is straightforward to see that top firms' share in sales and employment size is increasing with the automation level. Autor et al. (2020) show that "superstar" firms are capturing a larger share of the market and this phenomenon is more pronounced in the industries where labor share is falling faster. Though, they interpret the increase in market power as an important driver of the decrease in the labor share. In our model, the causality is reversed. Here, automation leads to a decline in the labor share and an increase in market concentration. There is also a significant correlation between information technology (IT) intensity and market concentration (Brynjolfsson et al., 2008; Bessen, 2017). We believe that IT is an important part of automation technology hence high IT intensity is an implication of more automation. In this regard, these observations are consistent with our model's prediction on market concentration.

Decreasing Entrepreneurship Rate: One important margin in the model is the occupational choice. An increase in automation technology has two counteracting impacts on the wage rate. Similar to Acemoglu & Restrepo (2018a), there are replacement and productivity effects. First, as capital replaces labor, demand for labor decreases, and this dampens the wage. Second, automation enables firms to allocate workers to more productive tasks, therefore it increases productivity and wages. Due to occupational choice, there is an additional effect in this model. Automation increases the return to entrepreneurial skill, hence it incentivizes workers to become an entrepreneur. Depending on the change in the wage rate and return to entrepreneurial skill, the marginal individual might change his occupation. In the early stage of automation, productivity effect dominates and hence the marginal entrepreneur becomes a worker. In the later stage of automation, the replacement effect dominates, and this reverses the decision of the marginal individual. Hence, it is expected to see a decreasing business dynamism in the early stage of automation. This is in line with the decrease in the start-up rate in the US (Decker et al., 2014; Pugsley & ahin, 2019; Salgado, 2019).

4 Endogenizing v: Efficiency Wage

In the previous section, we introduced a convex cost function for labor, v, as the source of the diseconomies of scale, but did not provide why there is this additional cost and why it is convex. In this section, we provide a micro foundation for v using the efficiency wage theory similar to

Shapiro & Stiglitz (1984).

Suppose that time is continuous and there is a measure one of the risk-neutral individuals who discounts future with rate $r > 0$. Each individual has two types of skills: labor and entrepreneurial skill. Labor skill is the same for all individuals, whereas, entrepreneurial skill, denoted by z, is distributed with some cumulative distribution function G. There is a fixed amount of capital, $\bar{K}$. To avoid the capital accumulation decision of individuals, we assume that capital is owned by outsiders.

Each individual can either become a worker or an entrepreneur. An entrepreneur rents capital hires labor to produce output and enjoys a profit, $\pi(z)$. A worker can be in one of two states at any point in time: employed or unemployed.

4.1 The Problem of the Worker

An employed worker earns a flow wage w until he is separated from the job. The separation can happen in two ways: exogenous separation that happens with Poisson rate δ or getting caught while shirking. A worker is monitored with a Poisson rate of q. Hence, a worker who shirks leaves the job with Poisson rate $\delta + q$, and a worker who exerts effort leaves the job with Poisson rate δ.

Let U denotes the value of being unemployed, $V_e(w, q)$ denotes the value of exerting effort in a job that pays wage w and monitoring probability is q, and $V_s(w, q)$ denotes the value function for shirking on the job that pays wage w and monitoring probability is q. Then, V_e and V_s satisfy the following equations:

$$rV_s(w, q) = w + (\delta + q)\left[U - V_s(w, q)\right], \tag{9a}$$

$$rV_e(w, q) = w - c + \delta\left[U - V_e(w, q)\right], \tag{9b}$$

where c is the cost of exerting effort.

An employed worker exerts effort if and only if $V_e(w,q) \geq V_s(w,q)$. This implies that the worker exerts effort if the wage rate satisfies:

$$w \geq rU + c + \frac{(r+\delta)c}{q}. \tag{NSC}$$

This is the so-called *no-shirking condition*. This condition says that wage rate should compensate for disutility of working, $rU + c$. Moreover, there is a premium to induce the worker to work, $(r+\delta)c/q$.

An unemployed worker enjoys a flow unemployment benefit b and finds a job with Poisson rate μ. Let $x = (w,q)$ be the characteristics of the jobs. Then, U satisfies the following equation:

$$rU = b + \mu \int_{\mathcal{X}} \left[\max\{V(x) - U, 0\}\right] dF(x), \tag{10}$$

where $\mathcal{X}$ is the set of active firms characteristics, F is the distribution of job openings and $V(w,q) = \max\{V_e(w,q), V_s(w,q)\}$.

4.2 The Problem of the Entrepreneur

Now consider an entrepreneur. The entrepreneur has a limited ability to monitor his employees. Assume that as the measure of employees increases, the probability of being monitored for a single worker decreases, i.e. $q(L)$ is a decreasing function of L. [11] The entrepreneur wants his employees to exert effort, otherwise, they produce nothing. Therefore, he needs to take into account the moral hazard problem. An entrepreneur with L labor needs to pay his workers $w(L)$ such that $V_e(w,q(L)) \geq V_s(w,q(L))$ so that workers exert effort. In other words, at the optimum w should

[11] In this model monitoring is only done by the entrepreneur. The monitoring cost is only important if the entrepreneur cannot identify the shirking worker from the non-shirking one. Alternatively, it might be possible to use a contract that depends on the performance of peers. This way, the entrepreneur can incentivize her employees to monitor each other (Che & Yoo, 2001). Replacing labor with capital would decrease the peer monitoring, which might lead to a change in the compensation scheme to induce workers to exert effort (Dogan & Yildirim, 2017). However, in this paper, we only consider monitoring entrepreneurs.

satisfy (NSC) with equality. Hence, define the optimal wage policy as:

$$w(L) = rU + c + \frac{(r+\delta)c}{q(L)}.$$

In this setting, the wage premium to induce a worker to exert effort is a function of firm size: larger firms need to pay a higher wage.

The problem of an entrepreneur in this setting is:

$$\pi(z) = \max_{\substack{I^\star, \{\ell_i\}_{i\in[I^\star,1]}, \\ \{k_s\}_{s\in[0,I^\star)}}} zY - w\left(\int_{I^\star}^{1} \ell_i di\right) \int_{I^\star}^{1} \ell_i di - R \int_{0}^{I^\star} k_i di \tag{11}$$

$$s.t. \;\; 0 \leq I^\star \leq I,$$

$$\ell_i \geq 0, \;\; k \geq 0.$$

An individual would become an entrepreneur instead of a worker if $\pi(z) \geq rU$. Since $\pi(z)$ is increasing in z, there exist a marginal entrepreneur $z^\star$ such that any individual with $z' > z^\star$ becomes an entrepreneur.

4.3 Equilibrium

Definition 2. *For a given automation technology I, skill distribution G with support $[z_{min}, \infty)$ and capital stock $\bar{K}$, the steady state equilibrium of the economy consists of prices $\{R, w(.)\}$, value functions $\{U, V_e(z), V_s(z)$, the marginal entrepreneur $z^\star$, labor and capital demand $\{\ell^\star(z), k^\star(z)\}$ for $z \geq z^\star$, automation technology $I^\star(z)$ for $z \geq z^\star$, matching process μ, vacancy distribution F, and unemployment u such that:*

- *Value functions satisfy* (10), (9a),(9b);
- $\pi(z^\star) = rU$;

- *$\ell(z)$, $k(z)$ and $I^\star(z)$ solve the entrepreneur's problem* (11)*;*

- *inflow to unemployment should be equal to outflow from unemployment:* $\delta(G(z^\star) - u) = \mu u$*;*

- *labor market clears:* $\int\limits_{z^\star}^{\infty}(1 - I^\star(z))\ell^\star(z)dG(z) = G(z^\star) - u$*;*

- *capital market clears:* $\int\limits_{z^\star}^{\infty} I^\star(z)k^\star(z)dG(z) = \bar{K}$*;*

- *wage function:* $w(L) = rU + c + \frac{(\delta+r)c}{q(L)}$*;*

- *vacancy posting:* $F(w) = \frac{\int\limits_{\underline{w}}^{w} g(z(w'))L(z(w'))}{\int\limits_{\underline{w}}^{\infty} g(z(w'))L(z(w'))}dw'$ *for* $w \geq \underline{w} = w(L(z^\star))$*,*

where $z(w) = L^{-1}\left(q^{-1}\left(\frac{w-c-rU}{(r+\delta)c}\right)\right)$*, i.e. entrepreneurial skill level that offers wage rate* w*,* L^{-1} *is the inverse function of labor demand and* q^{-1} *is the inverse of monitoring probability.*

4.4 Characterization of the Income Distribution

We can separate the wage function in two components, fixed and variable part. Define $w_0 = rU + c$ and $w_v(L) = \frac{(r+\delta)c}{q(L)}$ so that $w(L) = w_0 + w_v(L)$. The labor cost in this setting can be mapped to the labor cost in previous setting by defining $v(L) = w_v(L)L$. Hence, if $L/q(L)$ is convex, the entrepreneur's problem would be the same in both settings.

Even though nothing has changed on the entrepreneur's side, the labor supply side has changed. First, to discipline the workers, there must be unemployment. Without unemployment, workers can immediately find a new job after being fired, then there is no cost of being fired. Therefore, unemployment is needed to discourage workers from shirking. Second, there is a wage dispersion. In the previous section, we assumed v as a waste; in contrast, here we assume that it is paid to the workers as compensation. Since monitoring in larger firms is harder, an entrepreneur with a larger labor force needs to pay more to provide workers an incentive to exert effort. Therefore, the firm size distribution would lead to a wage dispersion.

To characterize distributions, we need more structure. Recall that $q(L)$ is the probability that a worker is being monitored. One intuitive way to define q is to think that the entrepreneur randomly selects workers and monitors them. Let M denote the measure of workers that an entrepreneur can monitor in a given time. Then, the Poisson rate that a worker in a firm with L employees is monitored is $q(L) = M/L$. This leads to $v(L) = ML^2$, in other words letting $\alpha = 2$ in the previous setting would give the same entrepreneur's problem. Therefore, similar results follow in this setting:

Corollary 1. *If z follows a Pareto distribution with parameter λ, then for large enough π, profit distribution can be approximated by Pareto distribution with parameter $\lambda(1-I)/2$.*

As we discussed, there is going to be a wage dispersion even for workers. Since there is one-to-one relation with wage level and firm size, wage distribution mimics the firm size distribution:

Corollary 2. *If z follows a Pareto distribution with parameter λ, then for large enough w, wage distribution can be approximated by Pareto distribution with parameter $\lambda(1-I)$.*

Observe that only the curvature of $q(L)$ is important for profit distribution. In this formulation, the efficiency of monitoring, M, does not have an impact on the tail parameter. If an entrepreneur can monitor a higher measure in a given time, this would not change the convexity of the profit function. Therefore, if we think M as the monitoring technology or communication technology as in Garicano (2000), then it has no impact on top income inequality. On the other hand, automation technology I still has the same impact on the right tail of the income distribution. Furthermore, now it not only impacts profit distribution but also leads to thicker wage distribution.

5 Conclusion

After the 1980s, income distribution in the US has become more skewed. While rich people have been getting richer, the super-rich has become even more so. In this paper, we argue that improvements in automation technology contributed to the widening gap between the top earners. Our theory states that if the cost of labor is convex, then entrepreneurs have decreasing returns to scale production function. As automation technology improves, dependence on labor deteriorates and the importance of convex cost decreases. This lessens the severity of diseconomies of scale and increases the return to entrepreneurial skill. Therefore, income inequality among the top earners' increases. Using industry-level data for the US and cross country data, we provide evidence that an improvement in automation technology leads to a lower Pareto parameter.

According to our model, the Pareto parameter of top income distribution is a function of three parameters: automation level, skill distribution, and convexity of labor cost. In this paper, we discussed the impact of the change in automation level, because we know for a fact that automation has increased. However, we believe that the other two parameters are also important and need attention.

We provide one explanation for convex cost of labor: efficiency wage. However, any theory that leads to firm size premiums should deliver similar results. It is shown that firm size wage premium is decreasing, the gap between large firms and small firms is decreasing (Bloom et al., 2018; Cobb & Lin, 2017). This might be a piece of evidence that monitoring cost is decreasing, hence it also contributing to top income inequality.

Even though we see this paper as a model of automation, the model is open to other ways of interpretation. With slight modification (by specifying the supply and price of the outsourcing), it can be seen as a model of outsourcing. We believe that outsourcing is an important subject as automation and it is crucial to distinguish them. Unfortunately, this paper remains short in that aspect.

References

Acemoglu, D. (2002). Directed technical change. *The Review of Economic Studies*, *69*(4), 781–809.

Acemoglu, D., & Restrepo, P. (2018a). *Artificial intelligence, automation and work* (Working Paper No. 24196). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w24196

Acemoglu, D., & Restrepo, P. (2018b). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, *108*(6), 1488-1542.

Acemoglu, D., & Restrepo, P. (in press). Robots and jobs: Evidence from us labor markets. *Journal of Political Economy*, *0*(ja), null. doi: 10.1086/705716

Aghion, P., Akcigit, U., Bergeaud, A., Blundell, R., & Hémous, D. (2018). Innovation and top income inequality. *The Review of Economic Studies*, *86*(1), 1–45.

Aoki, S., & Nirei, M. (2017). Zipf's law, pareto's law, and the evolution of top incomes in the united states. *American Economic Journal: Macroeconomics*, *9*(3), 36–71.

Atkinson, A., & Lakner, C. (2017). *Capital and labor: The factor income composition of top incomes in the united states, 1962–2006.* The World Bank.

Autor, D., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review*, *103*(5), 1553-97.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Van Reenen, J. (2020, 02). The Fall of the Labor Share and the Rise of Superstar Firms*. *The Quarterly Journal of Economics*, *135*(2), 645-709.

Bakija, J., Cole, A., & Heim, B. T. (2012). *Jobs and income growth of top earners and the causes of changing income inequality: Evidence from us tax return data.* Retrieved from https://web.williams.edu/Economics/wp/BakijaColeHeimJobsIncomeGrowthTopEarners.pdf

Bengtsson, E., & Waldenström, D. (2018). Capital shares and income inequality: Evidence from the long run. *The Journal of Economic History*, *78*(3), 712–743.

Bessen, J. E. (2017). *Information technology and industry concentration* (Working Paper No. 17-41). Boston Univ. School of Law, Law and Economics. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3044730

Bloom, N., Guvenen, F., Smith, B. S., Song, J., & von Wachter, T. (2018, May). The disappearing large-firm wage premium. *AEA Papers and Proceedings*, *108*, 317-22.

Brynjolfsson, E., McAfee, A., Sorell, M., & Zhu, F. (2008). *Scale without mass: business process replication and industry dynamics* (Working Paper No. 07-016). Harvard Business School Technology & Operations Mgt. Unit. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=980568

Calvo, G. A. (1985). The inefficiency of unemployment: The supervision perspective. *The Quarterly Journal of Economics*, *100*(2), 373–387.

Che, Y.-K., & Yoo, S.-W. (2001). Optimal incentives for teams. *American Economic Review*, *91*(3), 525–541.

Clemens, J., Gottlieb, J. D., Hémous, D., & Olsen, M. (2017). *The spillover effects of top income inequality.* mimeo, University of Zurich.

Cobb, J. A., & Lin, K.-H. (2017). Growing apart: The changing firm-size wage premium and its inequality consequences. *Organization Science*, *28*(3), 429–446.

Decker, R., Haltiwanger, J., Jarmin, R., & Miranda, J. (2014). The role of entrepreneurship in us job creation and economic dynamism. *Journal of Economic Perspectives*, *28*(3), 3–24.

Dogan, M., & Yildirim, P. (2017). *Managing automation in teams.* Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3035849

Feenstra, R. C., Inklaar, R., & Timmer, M. P. (2015, October). The next generation of the penn world table. *American Economic Review*, *105*(10), 3150-82. Retrieved from http://www.aeaweb.org/articles?id=10.1257/aer.20130954 doi: 10.1257/aer.20130954

Flood, S., King, M., Rodgers, R., Ruggles, & Warren, J. R. (2018). *Integrated public use microdata series, current population survey: Version 6.0.* doi: 10.18128/D030.V6.0

Frydman, C., & Jenter, D. (2010). Ceo compensation. *Annual Review of Financial Economics*, *2*, 75–102.

Gabaix, X. (2016). Power laws in economics: An introduction. *Journal of Economic Perspectives*, *30*(1), 185–206.

Gabaix, X., & Landier, A. (2008). Why has ceo pay increased so much? *The Quarterly Journal of Economics*, *123*(1), 49–100.

Garicano, L. (2000). Hierarchies and the organization of knowledge in production. *Journal of political economy*, *108*(5), 874–904.

Garicano, L., & Rossi-Hansberg, E. (2006). Organization and inequality in a knowledge economy. *The Quarterly Journal of Economics*, *121*(4), 1383–1435.

Geerolf, F. (2017). *A theory of pareto distributions* [Working Paper].

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

Goos, M., Manning, A., & Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review*, *104*(8), 2509–26.

Guvenen, F., & Kaplan, G. (2017, April). *Top income inequality in the 21st century: Some cautionary notes* (Working Paper No. 23321). National Bureau of Economic Research. Retrieved from `http://www.nber.org/papers/w23321` doi: 10.3386/w23321

Hémous, D., & Olsen, M. (2018). *The rise of the machines: Automation, horizontal innovation and income inequality.* Retrieved from `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2328774`

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society*, 1127–1150.

Jones, C. I., & Kim, J. (2018). A schumpeterian model of top income inequality. *Journal of Political Economy*, *126*(5), 1785–1826.

Kim, J. (2015). *The effect of the top marginal tax rate on top income inequality* [Working Paper]. mimeo, KAIST. Retrieved from `http://jiheekim.net/files/Kim2015Nov-TopTaxInequality.pdf`

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

Moll, B., Rachel, L., & Restrepo, P. (2019). *Uneven growth: automations impact on income and wealth inequality.* mimeo, Princeton University.

Oi, W. Y., & Idson, T. L. (1999). Chapter 33 firm size and wages. In *Handbook of labor economics* (Vol. 3, p. 2165 - 2214). Elsevier. Retrieved from `http://www.sciencedirect.com/science/article/pii/S1573446399300195` doi: https://doi.org/10.1016/S1573-4463(99)30019-5

Piketty, T. (2014). *Capital in the 21st century.* Harvard University Press Cambridge, MA.

Piketty, T., & Saez, E. (2003). Income inequality in the united states, 1913–1998. *The Quarterly journal of economics*, *118*(1), 1–41.

Piketty, T., Saez, E., & Stantcheva, S. (2014). Optimal taxation of top labor incomes: A tale of three elasticities. *American economic journal: economic policy*, *6*(1), 230–71.

Pugsley, B. W., & ahin, A. (2019). Grown-up business cycles. *The Review of Financial Studies*, *32*(3), 1102–1147.

Rosen, S. (1981). The economics of superstars. *The American Economic Review*, *71*(5), 845–858.

Salgado, S. (2019). *Technical change and entrepreneurship.* mimeo, Wharton School, University of Pennsylvania.

Shapiro, C., & Stiglitz, J. E. (1984). Equilibrium unemployment as a worker discipline device. *The American Economic Review*, *74*(3), 433–444.

Smith, M., Yagan, D., Zidar, O., & Zwick, E. (2019, 07). Capitalists in the Twenty-First Century. *The Quarterly Journal of Economics*, *134*(4), 1675-1745.

Tervio, M. (2008). The difference that ceos make: An assignment model approach. *American Economic Review*, *98*(3), 642–68.

U.S. Bureau of Economic Analysis. (2018). *Components of value added by industry as a percentage of value added.* Retrieved from `https://apps.bea.gov/iTable/iTable.cfm?reqid=51&step=51&isuri=1&table_list=7` (accessed January 15, 2020)

Zeira, J. (1998). Workers, machines, and economic growth. *The Quarterly Journal of Economics*, *113*(4), 1091–1117.

A Online Appendix

Lemma 1. *If wage rate $w(L)$ is strictly increasing in L and positive, then $w(L)L$ is strictly convex.*

Proof. *Let $L_1 > L_2$ and define $L_\lambda = \lambda L_1 + (1-\lambda)L_2$ for some $\lambda \in (0,1)$. Suppose $w(L)L$ is not is not convex, then there exist $\lambda \in (0,1)$ such that*

$$w(L_\lambda)L_\lambda \geq \lambda w(L_1)L_1 + (1-\lambda)w(L_2)L_2$$
$$(1-\lambda)[w(L_\lambda) - w(L_2)]L_2 \geq \lambda[w(L_1) - w(L_\lambda)]L_1.$$

Since $w(L)$ is strictly increasing and positive, $w(L_\lambda) > w(L_1)$ and $w(L_\lambda) < w(L_2)$. This implies that left hand size is negative and right hand size is positive. Hence, it leads to a contradiction. This prooves that $w(L)L$ is a convex function. ■

A.1 Proof of Proposition 1

In order to prove the existence and the uniqueness of equilibrium, first we consider a constraint problem, in which cutoff for being entrepreneur is fixed. In such setting, we show that there exist a cutoff skill level $\bar{z}$ such that for $z > \bar{z}$ there exist no positive prices that clears labor and capital market at the same time. This allows us to bind the set of skill level. Then, we show that there exist $z^\star$ such that market clearing wage rate and profit for cutoff entrepreneur is same, hence $z^\star$ together with associated wage and rental rate constitute the equilibrium.

For a given prices $\{w, R\}$, labor demand for entrepreneur z is the solution to $v'(L(z)) = \left(\frac{zC(I^\star(z))}{R^{I^\star(z)}}\right)^{1/(1-I^\star(z))} - w$, where $C(I) = exp\left(\int\limits_I^1 ln\gamma_i\right)$. Since v is twice continuously differentiable and strictly convex, inverse of v' exists. Define $\phi := v'^{-1}$. For ease of notation, we drop argument for $I^\star(z)$ and write it as $I^\star$. Define $M(R,z) = \left(\frac{zC(I^\star)}{R^{I^\star}}\right)^{\frac{1}{1-I^\star}}$. Since labor demand is decreasing with R, and increasing with z, $M_R < 0$, $M_z > 0$, where M_i is partial derivative with respect to i.

Observe that boundary condition for labor demand is not satisfied, i.e. for a given positive R as the wage rate converges to 0 labor demand does not diverge. Therefore, decrease in the wage rate might not be sufficient to clear the market. We are going consider this boundary case in order to find when markets are not going to clear.

The labor market and capital market clearing conditions when the cutoff skill is z' are:

$$\int_{\tilde{z}}^{z_{max}} \phi\left[M(R,z) - w\right] dG(z) = G(z^\star), \tag{12a}$$

$$\int_{\tilde{z}}^{z_{max}} \frac{I^\star}{1-I^\star} \phi\left[M(R,z) - w\right] \frac{M(R,z)}{R} dG(z) = \bar{K}, \tag{12b}$$

where $\tilde{z} = max\{z', w^{1-I^\star}R^{I^\star}/C(I^\star)\}$ is the least productive active entrepreneur, given prices $\{R, w\}$, i.e. $M(R,\tilde{z}) = w$ if $\tilde{z} \neq z'$. Anyone above $\tilde{z}$ hires positive mass of labor, and anyone below $\tilde{z}$ does not hire.

Define $R_\ell(w, z')$ as the labor market clearing rental rate when the wage is w and individuals with $z > z'$ are entrepreneur. Define similar object $R_k(w, z')$ for the capital market. Observe that both $R_\ell(w, z)$ and $R_k(w, z)$ is decreasing in z, since increase in z decreases the total demand, but does not decrease the supply, hence R must decrease for a fixed wage. The intersection of these two curves is the rental rate that clears both markets for a given w and z'.

Observe that boundary condition for labor demand is not satisfied, i.e. for a given positive R as the wage rate converges to 0 labor demand does not diverge. Therefore, decrease in the wage rate might not be sufficient to clear the market. We are going consider this boundary case in order to find when markets are not going to clear.

Now we show that there exist unique $\bar{z}$ such that $R_\ell(0,\bar{z}) = R_k(0,\bar{z})$. To do this, first we show that $R_k(0, z)$ is bounded above, whereas $R_\ell(0, z)$ is not. Second, we show that for high enough z, $R_k(0, z) > R_\ell(0, z)$, hence, by the intermediate value theorem, they must intersect. Lastly, we show that at the point where they intersect, derivative of R_ℓ is higher than R_k, i.e.

around $\bar{z}$, $R_\ell(0,\bar{z}) - R_k(0,\bar{z})$ is decreasing, so that they can only intersect once. Notice that $\tilde{z} = z'$ when the wage rate is zero.

Lemma 2. *As $z \to z_{min}$, $R_\ell(0,z) \to \infty$, and $R_k(0,z) \to t < \infty$.*

Proof. *Let $z \to z_{min}$. Suppose the contrary, $R_\ell(0,z) \to p < \infty$, and $p > 0$. Take small $\epsilon > 0$, by continuity of $R_\ell(0,z)$, there exist $\delta > 0$, such that $R_\ell(0,z') \in (p-\delta, p+\delta)$ for any $z' \in (z_{min}, z_{min}+\epsilon)$. Define $k := \phi\left[M(p+\delta, z_{min}\right] > 0$. Let z' be such that $(1-G(z'))k > G(z')$, and $z' \in (z_{min}, z_{min}+\epsilon)$. Because $k > 0 = G(z_{min})$, such z' exists. Since $R_\ell(z') < p+\delta$ and labor demand is decreasing with R, labor demand for each z is higher than k. Hence, for small enough z':*

$$\int_{z'}^{z_{max}} \phi\left[M(R_\ell, z)\right] dG(z) > \int_{z'}^{\infty} k dG(z) = (1-G(z'))k > G(z'),$$

which contradicts that R_ℓ clears the market. Therefore, with a finite R_ℓ, the labor market cannot be cleared. Hence $R_\ell(0,z) \to \infty$ as $z \to z_{min}$.

Now consider the capital market condition (12b) when $z^\star = z_{min}$. As R_k converges to zero, demand goes to infinity, and as R_k diverges, demand converges to 0. Hence, there exist $R_k(0, z_{min} < \infty$ that clears the capital market. By continuity, $R_k(0,z) \to R(0,z_{min})$ as $z \to z_{min}$.

■

Since $R_\ell(0,z)$ diverges and $R_k(0,z)$ converges to some positive number as $z \to z_{min}$, this implies that for low enough z, $R_\ell(0,z) > R_k(0,z)$. We now show that inequality must be flipped for high enough z.

Lemma 3. *For high enough z', $R_k(0,z') > R_\ell(0,z')$.*

Proof. *Observe that as $z' \to z_{max}$, $R_\ell(0,z')$ and $R_k(0,z')$ converge to 0. To see this, for a positive R, both total labor demand and total capital demand converges to 0, in contrast capital supply is fixed and labor supply converges to 1. Hence, R_ℓ and R_k converge to 0 in order to clear the market. As R_k converges to 0, $I^\star$ converges to I, every entrepreneur automates all possible* *tasks. Then, capital demand is:*

$$\frac{I}{1-I} R_k^{-\frac{1}{1-I}} \int_{z'}^{z_{max}} \phi \left[M(R_k, z) - w \right] (zC(I))^{\frac{1}{1-I}} \, dG(z) = \bar{K}.$$

Since $R_k^{-\frac{1}{1-I}}$ *diverges, it must be the case that integral converges to 0 in order to have finite demand. Observe that* $\phi(M(R,z))(zC(I))^{\frac{1}{1-I}} > \phi(M(R,z))(z_{min}C(I))^{\frac{1}{1-I}} > 0$ *for* $z > z_{min}$*. Therefore,*

$$\int_{z'}^{z_{max}} \phi \left[M(R_k, z) \right] (zC(I))^{\frac{1}{1-I}} \, dG(z) > \int_{z'}^{z_{max}} \phi \left[M(R_k, z) \right] (z_{min}C(I))^{\frac{1}{1-I}} \, dG(z) \to 0.$$

However, labor demand must be equal to labor supply $G(z')$*, close to* 1 *for large* z'*. Hence, for large enough* z'*, it must be the case that:*

$$1 \approx \int_{z'}^{z_{max}} \phi \left[M(R_\ell, z) \right] dG(z) > \int_{z'}^{z_{max}} \phi \left[M(R_k, z) \right] dG(z) \approx 0. \tag{13}$$

Since M *is decreasing in* R*, it must be the case that* $R_\ell(0, z') < R_k(0, z')$ *for large* z'*.* ■

Lemma 4. *Let* $R_\ell(0, \bar{z}) = R_k(0, \bar{z})$*. Then* $|R'_\ell(0, \bar{z})| > |R'_k(0, \bar{z})|$*. In other words,* $R_\ell(0, \bar{z}) - R_k(0, \bar{z})$ *is decreasing around* $\bar{z}$*.*

Proof. *Let* $R_\ell(0, \bar{z}) = R_k(0, \bar{z}) = \tilde{R}$*. Using implicit function theorem, taking derivative of labor market condition* (12a) *with respect to* $\bar{z}$ *gives us:*

$$-\phi \left[M(\tilde{R}, \bar{z}) \right] g(\bar{z}) + \int_{\bar{z}}^{x_{max}} \phi' \left[M(\tilde{R}, z) \right] M_R(\tilde{R}, z) R'_\ell(\bar{z}) dG(z) = g(\bar{z}). \tag{14}$$

Similarly, derivative of capital market condition with respect to $\bar{z}$ *is:*

$$0 = -\frac{I^\star(\bar{z})}{1-I^\star(\bar{z})}\phi\left[M(\tilde{R},\bar{z})\right]\left(\frac{M(\tilde{R},\bar{z})}{\tilde{R}}\right)g(\bar{z}) + \int_{\bar{z}}^{z_{max}}\left[\frac{I^\star}{1-I^\star}\phi'\left[M(\tilde{R},z)\right]M_R(\tilde{R},z)R_k'(\bar{z})\frac{M(\tilde{R},z)}{R}\right.$$

$$+\frac{I^\star}{1-I^\star}\phi\left[M(\tilde{R},z)\right]\frac{M_R(\tilde{R},z)R_k'(\bar{z})\tilde{R}-M(\tilde{R},z)R_k'(\bar{z})}{\tilde{R}^2}$$

$$\left.+\frac{I_R^\star R_k'(\bar{z})}{(1-I^\star)^2}\phi\left[M(\tilde{R},z)\right]\frac{M(\tilde{R},z)}{\tilde{R}}\right]dG(z).$$

Recall that $I^\star(z)$ *is decreasing in* R*, hence* $I_R^\star \leq 0$*. Since* $R_k' < 0$*, the last term of the integrand is positive for all* z*. Similarly, the second term of the integrand is also positive, since* $M_R < 0$*. Therefore:*

$$\frac{I^\star(\bar{z})}{1-I^\star(\bar{z})}\phi\left[M(\tilde{R},\bar{z})\right]\left(\frac{M(\tilde{R},\bar{z})}{\tilde{R}}\right)g(\bar{z})$$

$$> \int_{\bar{z}}^{z_{max}}\left[\frac{I^\star}{1-I^\star}\phi'\left[M(\tilde{R},z)\right]M_R(\tilde{R},z)R_k'(\bar{z})\frac{M(\tilde{R},z)}{R}\right]dG(z).$$

Moreover, $I^\star/(1-I^\star)$ *and* $M(R,z)$ *are increasing in* z*. We can simplify the above expression by replacing them with* $I^\star(\bar{z})/(1-I^\star(\bar{z}))$ *and* $M(R,\bar{z})$ *:*

$$\phi\left[M(\tilde{R},\bar{z})\right]g(\bar{z}) > \int_{\bar{z}}^{z_{max}}\left[\phi'\left[M(\tilde{R},z)\right]M_R(\tilde{R},z)R_k'(\bar{z})\right]dG(z).$$

Using derivative of labor market condition, equation (14)*:*

$$\int_{\bar{z}}^{\infty}\phi'\left[M(\tilde{R},z)\right]M_R(\tilde{R},z)R_\ell'(\bar{z})dG(z) > \int_{\bar{z}}^{\infty}\left[\phi'\left[M(\tilde{R},z)\right]M_R(\tilde{R},z)R_k'(\bar{z})\right]dG(z),$$

$$\int_{\bar{z}}^{\infty}\phi'\left[M(\tilde{R},z)\right]M_R(\tilde{R},z)(R_\ell'(\bar{z})-R_k'(\bar{z}))dG(z) > 0.$$

Which implies that $R_\ell'(\bar{z}) - R_k'(\bar{z}) < 0$*, since* $\phi' > 0$ *and* $M_R < 0$*. In other words, if* $\tilde{R}$ *clears both labor and capital market, when we increase* z*, rental rate that clears the labor market* *decreases much faster than capital market.* ■

Lemma 5. *There exists a unique $\bar{z}$, such that $R_\ell(0,\bar{z}) = R_k(0,\bar{z})$*

Proof. *Lemma 2 shows that $R_\ell(0,z) > R_k(0,z)$ for low z, and lemma 3 shows that $R_\ell(0,z) < R_k(0,z)$ high z, therefore, they must intersect. Lemma 4 shows that they can at most intersect once, since at the point they intersect, $R_\ell(0,\bar{z}) - R_k(0,\bar{z})$ is decreasing. If they intersect once more, then it must be the case that the difference is increasing in the second intersection. Hence $\bar{z}$ exists and it is unique.* ■

Now, we prove that for $z^\star < \bar{z}$, there exist positive prices that clears the market.

Lemma 6. *For $z^\star < \bar{z}$, there exist $w > 0$ and $R > 0$ that labor market condition, equation* (12a)*, and capital market condition, equation* (12b)*, hold.*

Proof. *First, notice that $R_\ell(w,z^\star)$ and $R_k(w,z^\star)$ are decreasing in w. To see this, assume w increases but R does not decrease, $\tilde{z}$ weakly increases by definition and demand strictly decreases for each entrepreneur. But then market conditions cannot be satisfied. Therefore, $R_i(w,z^\star)$ must be strictly decreasing in w. Since $z^\star < \bar{z}$, $R_\ell(0,z^\star) > R_k(0,z^\star)$. We need to show that for large enough w, $R_k(w,z^\star) > R_\ell(w,z^\star)$.*

Observe that as w diverges, R converges to 0, otherwise market clearing condition cannot be satisfied. As R converges to 0, automation technology binds for everyone: $I^\star \to I$. Using the similar argument with proof of lemma 3:

$$G(z^\star) = \int_{\tilde{z}_\ell}^{z_{max}} \phi\left[M(R_\ell, z) - w\right] dG(z) > \int_{\tilde{z}_k}^{z_{max}} \phi\left[M(R_k, z) - w\right] dG(z) \to 0,$$

where $\tilde{z}_i = max\{z^\star, w^{1-I}R_i^I/C\}$. If $R_\ell \geq R_k$, $M(R_\ell,z^\star) \leq M(R_k,z^\star)$ and $\tilde{z}_\ell \geq \tilde{z}_k$. Therefore, above inequality cannot hold. Hence $R_k(w,z^\star) > R_\ell(w,z^\star)$ for large enough w. Using the intermediate value theorem, continuity of R_ℓ and R_k implies that there exist $w^\star > 0$ such that $R_\ell(w^\star,z^\star) = R_k(w^\star,z^\star)$.

Using similar argument with the proof of lemma 4, one can easily show that $R_\ell(w,z^\star) -$

$R_k(w, z^\star)$ is decreasing around $w^\star$, hence R_ℓ and R_k can at most intersect once. Therefore, prices are unique for a given $z^\star$. ■

Lemma 7. *For $z^\star > \bar{z}$, there does not exist positive prices that clears the market.*

Proof. *As we discussed in the previous lemma, R_k and R_ℓ can only intersect if $R_\ell > R_k$ for low wage rates, since R_ℓ cross R_k from above. However, $R_\ell(0, z^\star) < R_k(0, z^\star)$ since , $z^\star > \bar{z}$. Therefore, they cannot intersect when $w > 0$.* ■

Up to now, we know that for any $z \in (z_{min}, \bar{z}]$, there exist unique prices $w(z), R(z)$ that clears the market. To find the equilibrium, we need $z^\star$ to be indifferent between occupations, i.e. $\pi(z^\star|R(w^\star), w(w^\star)) = w(z^\star)$ where $\pi(z|R(w), w(w))$ is the profit of entrepreneur with skill z when prices are $\{R, w\}$. Clearly, if there are inactive entrepreneurs, then it cannot be equilibrium. Recall that an entrepreneur with skill level z' is inactive if $M(R(z), z') < 0 = w(z)$. Let's define $A := \{z|M(R(z), z) > w(z), z \leq \bar{z}\}$, so that $z \in A$ implies every entrepreneur is active when z is cutoff entrepreneur.

Lemma 8. *There exists $z_0 \in (z_{min}, \bar{z})$, such that for $z < z_0$, there exist inactive entrepreneurs, and for $z > \bar{z}$, every entrepreneur is active.*

Proof. *Let $z' \in A$. If $M(R(z'), z') - w(z')$ is increasing around z', then $z'' > z'$ implies $z'' \in A$. To show $M(R(z'), z') - w(z')$ is increasing take derivative with respect to z':*

$$dM/dz = M_R R_z + M_z - w_z.$$

Since M_z is positive, it is sufficient to show that $M_R R_z - w_z$ is positive. Now consider labor market clearing condition. By definition, $\tilde{z}' = z'$. Then:

$$\int_{z'}^{z_{max}} \phi \left[M(R(z'), z) - w(z') \right] dG(z) = G(z').$$

Derivative with respect to z' leads to:

$$\int_{z'}^{z_{max}} \phi'\left[M(R(z'),z)-w(z')\right](M_R(R(z'),z)R_z - w_z)\,dG(z) = g(z')+\phi\left[M(R(z'),z')-w(z')\right] > 0.$$

Claim 1. $\frac{\partial M(R,z)}{\partial R \partial z} < 0$.

Proof. *Suppose $I^\star(z) < I$, then $M(R,z) = R\gamma_{I^\star}$. Then:*

$$\frac{\partial M(R,z)}{\partial R \partial z} = \gamma'_{I^\star} I^\star_z + R\gamma''_{I^\star} I^\star_R I^\star_z + R\gamma_{I^\star} I^\star_{Rz}.$$

Recall that optimal $I^\star$ solves the following equality:

$$\int_{I^\star}^{1} ln\gamma_i - (1-I^\star) ln\gamma_{I^\star} = ln(R/z).$$

Using implicit function theorem twice, one for derivative of R, and second for derivative of z, we could get:

$$-\gamma_{I^\star} I^\star_z = -I^\star_z R\gamma'_{I^\star} I^\star_R + (1-I^\star)R\gamma''_{I^\star} I^\star_z I^\star_R + (1-I^\star)R\gamma'_{I^\star} I^\star_{Rz}.$$

By rearranging, one can get:

$$\frac{\partial M(R,z)}{\partial R \partial z} = I^\star_z R\gamma'_{I^\star} I^\star_R - I\gamma'_{I^\star} I^\star_z < 0$$

since $I^\star_z > 0$, $I^\star_R < 0$ and $\gamma_{I^\star} > 0$.

Now suppose technology binds, hence $M(R,z) = (ZC/R^I)^{1/(1-I)}$. Since I does not change with small changes in R and z, it is straight forward to show that $M_{Rz} < 0$. ■

By claim 1, M_R is decreasing in z, hence $M_R(R(z'),z') > M_R(R(z'),z)$ for $z > z'$. ϕ is *strictly increasing, hence derivative is positive. Thus*

$$(M_R(R(z'), z')R_z - w_z) \int_{z'}^{z_{max}} \phi' \Big[M(R(z'), z) - w(z') \Big] dG(z) > 0.$$

Therefore, it must be the case that $M_R(R(z'), z')R_z - w_z$ *is positive, which implies that* $M(R(z'), z') - w(z')$ *is increasing. Define* $z_0 := inf A$. $M(R(z), z) - w(z)$ *is increasing implies that* A *is connected, for* $z \in A \iff \bar{z} \geq z > z_0$*, if such* z_0 *exists.*

Next, we show that z_0 *exists and is in* $(z_{min}, \bar{z})$*.B y definition,* $M(R(\bar{z}), \bar{z}) > 0 = w(\bar{z})$*, hence* $\bar{z} \in A$*. Continuity of* R, w, M *implies that* $z_0 < \bar{z}$*.*

To show $z_0 > z_{min}$*, suppose the contrary. Notice that as* $z \to z_{min}$*, labor supply shrinks, so demand converges to 0. Hence, it cannot be possible that both* $R(z)$ *and* $w(z)$ *converges to a finite number, which leads to positive labor demand. Since* $R_k(0, z) < \infty$ *by lemma 2, it must be the case that* $w(z) \to \infty$ *and* $R(z) \to 0$ *as* $z \to z_{min}$*. Since, by assumption, every entrepreneur is active, then* $z^\star > w^{1-I}R^I/C$*, hence* $w^{1-I}R^I/C$ *is finite . However, this implies that the labor demand,* $\phi\left[\frac{(zC)^{1/(1-I)} - (w^{1-I}R^I)^{1/(1-I)}}{R^{I/(1-I)}}\right]$*, diverges. Therefore,* $w^{1-I}R^I$ *must diverge, which implies for small* z*, there are inactive entrepreneurs. Hence* $z_0 > z_{min}$*.* ■

Proposition 1. *For a given automation technology* $0 < I < 1$*, capital stock* $\bar{K}$*, and skill distribution* G *with support* $[z_{min}, z_{max}] \subset \mathbb{R}_+$*, there exists a unique equilibrium.*

Proof. *By lemma 6 and 7, and due to the fact that every entrepreneur is active in the equilibrium, we know that* $z^\star \in (z_0, \bar{z})$*. Define profit of cutoff entrepreneur as* $\tilde{\pi}(z) = \pi(z|R(z), w(z))$*.* $z^\star, R(z^\star), w(z^\star)$ *is equilibrium if* $\tilde{\pi}(z^\star) = w(z^\star)$*. Optimality conditions imply* $\tilde{\pi}(z) = v'\left[M(R(z), z) - w(z)\right]\phi\left[M(R(z), z) - w(z)\right] - v(\phi\left[M(R(z), z) - w(z)\right])$*. Derivative with respect to* z *gives us :*

$$\left[M(R(z), z) - w(z)\right]\phi'\left[M(R(z), z) - w(z)\right]\left[M_R R_z + M_z - w_z\right] > 0$$

where first term is positive since $z \in A$*, second term is positive because* ϕ *is increasing and last* *term is positive by lemma 8. Therefore, $\tilde{\pi}$ is strictly increasing in $(z_0, \bar{z})$, with $\tilde{\pi}(z_0) = 0$ and $\tilde{\pi}(\bar{z}) > 0$.*

On the other hand, $w(\bar{z}) = 0$ by definition, and $w(z_0) > 0$ by lemma 6. By the intermediate value theorem and continuity of $\tilde{\pi}(z)$ and $w(z)$, they must intersect.

To show that it is unique, we want to show that $w(z)$ is decreasing in z. Fix z' and $w' = w(z')$. Take the derivative of the market clearing conditions with respect to z fixing w constant, around $R_\ell(w', z')$ and $R_k(w', z')$. Using similar idea to lemma 4, one can get:

$$\int_z^{\prime} \phi' M_R(R'_\ell(w', z') - R_k(w', z')) > 0.$$

Since M_R is negative, it must be the case that $(R'_\ell - R'_k) < 0$. By definition of derivative, this implies that:

$$\frac{R_\ell(w', z' + \epsilon) - R_k(w', z + \epsilon)}{\epsilon} < 0$$

for small $\epsilon > 0$. But then, R_ℓ and R_k cannot intersect at $w >\geq w'$, since $R_\ell(w, z) - R_k(w, z)$ must be decreasing in w around market clearing wage rate. This implies that $w(z)$ is strictly decreasing.

This concludes that $w(z)$ and $\tilde{\pi}$ intersects only once, hence the equilibrium is unique. ■

A.2 Proof of Proposition 3

Proposition 3. *Suppose the distribution of entrepreneurial productivity, z, is Pareto with shape parameter λ, monitoring cost function is $v(L) = L^\alpha$, and $\lambda(1 - I)(\alpha - 1) > 1$.*[12] *Then, the distribution of profits has a Pareto tail with shape parameter $\lambda(1 - I)\frac{\alpha-1}{\alpha}$.*[13]

[12]Proposition 1 can be extended to any unbounded distributions as long as the labor demand remains finite. For the Pareto distribution, we need $\lambda(1 - I)(\alpha - 1) > 1$ to have an equilibrium.

[13]We say that the tail distribution of F is distributed by G if $F(x)/G(x) \to 1$ as $x \to \infty$. Observe that including capital income, RK, does not impact the tail of income distribution of entrepreneurs.

Proof. *The distribution of profit* $\Pi = \pi(z)$ *is given by:*

$$P(\Pi > \pi) = D\left[\alpha\left(\frac{\pi}{\alpha-1}\right)^{\frac{\alpha-1}{\alpha}} + w\right]^{\lambda(1-I)} \frac{R^{\lambda I}}{C^{\lambda}}. \tag{15}$$

By dividing to $\tilde{D}\pi^{\lambda(1-I)\frac{\alpha-1}{\alpha}}$*, where* $\tilde{D} = DR^{\lambda I}/C(I)^{\lambda}\left[\alpha/(\alpha-1)^{(\alpha-1)/\alpha}\right]^{\lambda(1-I)}$*, we can get:*

$$\frac{P(\Pi > \pi)}{\tilde{D}\pi^{\lambda(1-I)\frac{\alpha-1}{\alpha}}} = \left[1 + \frac{w(\alpha-1)^{\frac{\alpha-1}{\alpha}}}{\alpha\pi^{\frac{\alpha-1}{\alpha}}}\right]^{\lambda(1-I)} \to 1 \quad as\ \pi \to \infty.$$

Hence $\Pi \sim Pareto(\lambda(1-I)\frac{\alpha-1}{\alpha})$*.*

■