Automation and Top Income Inequality

Ömer Faruk Koru*

Pennsylvania State University

September 2023

Abstract

For almost 40 years, top income inequality has increased sharply in the US. At the same time, there have been major improvements in automation technology. It is well-known that the distribution of top income is well approximated by a Pareto distribution. In this paper, we provide a theory that links automation technology to the Pareto tail of the income distribution. We construct a model in which managing labor is more difficult than managing capital. We model this as a convex cost of labor, resulting in 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 on entrepreneurial skills, a decrease in the Pareto parameter of income distribution, and an increase in top income inequality. We microfound the convex cost of labor using a theory of efficiency wages. Using cross-industry and cross-country data, we provide evidence 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.

*Visiting Assistant Professor. Email: oqk5084@psu.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 Lunch Seminar, Midwest Macro Meeting Spring 2019, NASMES 2019, GCER 2019, GW Student Research Conference 2019. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

1 Introduction

Over the past four decades, income inequality among the highest earners in the United States has witnessed a significant rise. This inequality is measured by the ratio of the income share belonging to the top 0.1% to that of the top 1%, highlighting the widening income gap between the super-rich and the rich. Notably, this period coincided with substantial technological advancements, particularly in the field of automation. In this paper, we present a theory that establishes a connection between the level of automation and top income inequality.

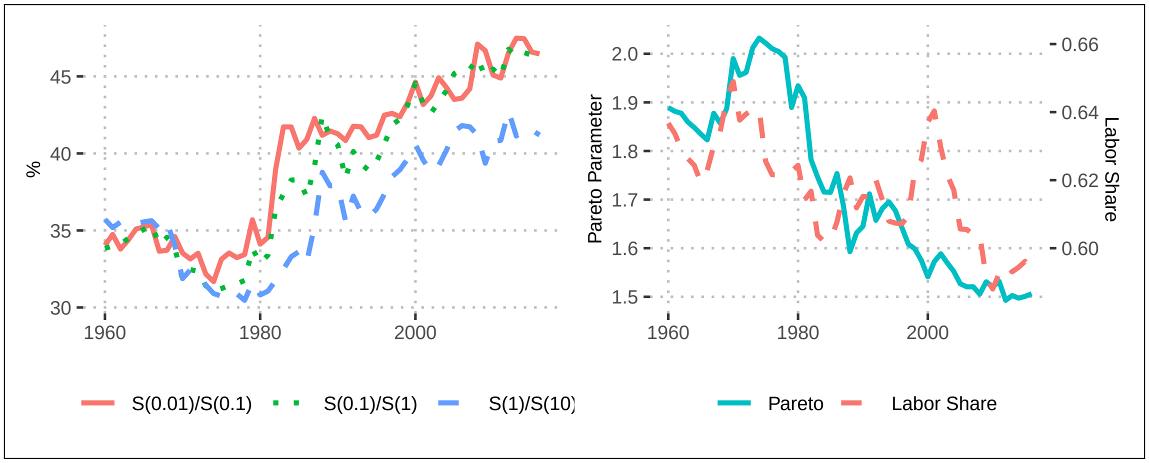
The left panel of Figure 1 illustrates the relative income shares within the top income distribution. The solid line depicts the ratio of the income share of the top 0.01% to that of the top 0.1%, while the dotted line represents the ratio of the income share of the top 0.1% to the top 1%. Furthermore, the dashed line showcases the ratio of the income share of the top 1% to the top 10%. The figure reveals a clear upward trend in top income inequality since the 1980s. This trend holds true regardless of whether we define the top income as the top 10%, top 1%, or top 0.1%. As demonstrated in Figure 1, top income inequality has increased by nearly half over the course of the past four decades.

It is well-known that the top income distribution of income is well approximated by a Pareto distribution.[1] An important implication of the Pareto distribution is that the relative income share is determined by the Pareto (shape) parameter. The observed increase in top income inequality suggests that the Pareto parameter associated with the top income distribution has been decreasing over time.

Our model establishes a link between the Pareto parameter and automation technology. Specifically, we adopt a task-based production function similar to the one employed by Acemoglu & Restrepo (2018b). This production function implies that the labor share of income is a function of the level of automation, which holds true in our model as well. Acemoglu & Restrepo (2021)

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

Figure 1: Top Income Inequality and Labor Share



Note: *Left panel: The solid line shows the ratio of the income share of the top* 0.01% *to the top* 0.1%. *The dotted line shows the ratio of the income share of the top* 0.1% *to the top* 1%. *The dashed line shows the ratio of the income share of the top* 1% *to the 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.*

and Moll et al. (2021) use this result to estimate the impact of automation on economic inequality using labor share as a proxy. Consequently, we depict the relationship between the labor share and the Pareto parameter in the right panel of Figure 1. The figure reveals a strong correlation between automation and top income inequality.

In our interpretation, automation encompasses a broad range of technologies that replace labor. This definition encompasses computers, software (or information technology, IT, in general), as well as industrial robots. As depicted in the figure, top income inequality began to increase in the late 1970s, coinciding with the dawn of the IT revolution. This observation supports the notion that the rise in top income inequality can be attributed to the advent of labor-replacing technologies.

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 span of control argument proposed by 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 contrast, in this paper, we endogenize this parameter and show how it changes with automation. We incorporate the span of control problem of entrepreneurs to the task-based framework of automation decision, as modeled by Acemoglu & Restrepo (2018b).

The model is based on the assumption that labor is the main reason for diseconomies of scale. We assume that managing labor is more difficult than managing capital, and replacing labor with capital enables entrepreneurs to scale up their production. Improvement in automation technology helps entrepreneurs to replace labor with capital and, hence, reduces the severity of diseconomies of scale. One reason for the cost of managing labor might be moral hazard. If entrepreneurs do not know whether their employees are working or shirking, then entrepreneurs need to spend additional resources to provide an incentive for their workers to exert effort, such as investing in monitoring technology or paying efficiency wages. If the cost of additional resources is convexly increasing with the number of workers, then the profit function exhibits decreasing returns to scale. In the main text, we provide the result for any convex cost function in labor. In the appendix, we show that an efficiency wage theory provides a microfoundation for this convex cost.

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 implications for a change in scalability. Technological improvement is usually modeled as either an increase in the productivity of some factor (for example, Acemoglu (2002)), a change in capital share (for example Acemoglu & Restrepo (2018b)), or a decrease in the price of capital (for example, Autor & Dorn (2013)). Because none of these affect the decreasing returns to scale of the production function, they do not impact the tail of income distribution.

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.) to produce the final good. While some of these tasks can be automated (i.e., they can be produced by capital), some can only be produced by labor. As automation technology improves, dependency on labor decreases, and the 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 understand the mechanism, consider the extreme case when none of the tasks can be automated. Then, the cost of production is the price of labor plus the monitoring cost, which is convex. Therefore, the top-skilled entrepreneur can only serve a portion of the market, which enables lower-skilled entrepreneurs to enter the market. Now, consider the other extreme case when any task can be automated. Now, the only cost of production is the price of capital, which is linear. Therefore, an entrepreneur has no problem scaling up his output, and, hence, the top-skilled individual captures the entire market. While in the first scenario, there is lower inequality due to the presence of other entrepreneurs, in the second scenario there is perfect inequality because the top talented entrepreneur controls the entire market. When the economy shifts from one extreme to the other, thanks to automation, income inequality at the top increases.

The main result of our model is consistent with the data. One important feature of the task-based framework is that it endogenizes factor shares of income. Acemoglu & Restrepo (2018b) and Martinez (2021) show that the labor share of income is a function of the automation level. This is also true in our model. Our model predicts that the Pareto parameter is proportional to the share of non-automated tasks, which is equal to the labor share of income. We test this relationship using two different datasets. First, we examine the cross-industry cross-time variation of the labor share and the Pareto parameter in the US. Second, we study the cross-country, cross-industry variation in OECD countries. Our regression results confirm the model's prediction.

Even though our focus is an entrepreneurial income, an alternative interpretation of the profit in the model is CEO compensation. In other words, if entrepreneurs in the model are CEOs of large corporations, our model predicts that the Pareto parameter of CEO compensations should be lower in industries with lower capital share. We test this result using CEO compensation data from the Compustat Execucomp dataset and find that there is a positive relationship between the Pareto tail and the labor share. Furthermore, using firm-CEO level data, we show that there is a positive relationship between the capital intensity of the firm and CEO compensation.

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 extends to a dynamic model without any friction in which the only dynamic choice is capital accumulation. In such a model, the entrepreneur's problem is static. Furthermore, because the main result of the static model does not depend on the prices, it is true in this dynamic version of the model.

Related Literature: This paper is related to several strands in the literature. First, we contribute to the literature on the impact of automation on labor market outcomes (Acemoglu & Restrepo, 2020; Autor & Dorn, 2013; Goos et al., 2014; Hémous & Olsen, 2018; Moll et al., 2021). This literature mainly focuses on wage inequality between high and low-skill workers. Martinez (2021) also considers the impact of automation on the span of control; however, the main focus of his paper is the decline in the labor share, which does not drive the right tail of the income distribution. In this paper, in contrast, we focus on the impact of automation on the Pareto parameter of the top income distribution.

The study most closely related to ours is that undertaken by Moll et al. (2021), who examine the impact of automation on income and wealth distribution. In their model, automation gives rise to higher returns to wealth, and, hence, it increases the incentive to accumulate wealth. Because of the birth and death process, some individuals are lucky enough to live long enough to accumulate wealth exponentially and end up in the top percentile of income and wealth distribution. As in our paper, the thickness of the income distribution is a function of the automation level. Yet we identify a different mechanism. 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 crucial fraction of top income earners and their skill is an integral part of their firm's performance. Thus, we believe that to understand the dynamics of top income inequality it is important to first understand the change in the return to entrepreneurial skills.

Second, a growing literature examines the determinants of top income inequality and the change in top income inequality. Gabaix & Landier (2008) and Tervio (2008) use assignment models to explore the connection between changes in firm size distribution and increases in CEO compensation. 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 decline in the top marginal tax rate on the share of the 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 the knowledge-based hierarchies model of Garicano (2000) and Garicano & Rossi-Hansberg (2006) generates a Pareto tail at the top of income distribution and that the Pareto parameter is inversely related to the number of layers. However, Rajan & Wulf (2006) provide evidence that the number of layers in corporations in the US has decreased over time.

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

In a 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 to incentivize workers to exert effort. In this paper, we, too, consider the monitoring problem of workers. However, our main focus is on top income inequality.

The structure of the remaining paper is as follows: Section 2 presents the reduced form model and the main results. Section 3 provides motivating facts, and section 4 concludes.

2 The Model

We consider a static model economy. To understand the impact of improvement in automation technology, we characterize top income distribution and 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.

Each individual can either become a worker or an entrepreneur. If an individual becomes a worker, she supplies labor inelastically and earns wage w. If she becomes an entrepreneur, she 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 that of Zeira (1998) and Acemoglu & Restrepo (2018b). To produce a unique final good, an entrepreneur needs to complete a measure one of the tasks, $i \in [0, 1]$. There is no market for tasks, and so the entrepreneur needs to complete all of the tasks inside the firm.[2]

Tasks are complements and they are aggregated into output by a unit elastic aggregator (i.e., Cobb-Douglas). The log of the production function is given by:

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

where $y(i)$ is the level of task i used in the production. For a given entrepreneurial skill z, the total output is zY.

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, and γ_i is the productivity of labor in task i. We assume that γ_i is increasing in i.[3] In other words, i denotes the complexity of the task, and labor has a comparative advantage relative to capital in high-index tasks.

Because 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 technology. 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

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

[3] For the sake of 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.

entrepreneur, so that any $i < I^\star$ is automated.[4] Then, by combining (1) and (2), the log of the production function 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 labor. As the employment size increases, the entrepreneur losses control over 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, who, to sustain control, 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. she decides which tasks are to be automated, $I^\star$; how much capital to rent 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:

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

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

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

and the production is subject to (3), where w is the wage rate and R is the rental rate of capital.

[4] We assume that the least productive tasks can be automated. However, Autor & Dorn (2013) argue that middle-skilled jobs 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 $I_{min} = 0$ because 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 leads to a decreasing returns to scale profit function. 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 the diseconomies of scale.

2.1.1 Interpretation of v

The main mechanism of this paper relies on the concept of convex cost represented by v, which captures the diminishing control over labor as employment size increases. The interpretation of v varies and offers insights into different aspects of labor costs.

One interpretation of v relates to monitoring costs. In order to mitigate issues such as shirking, entrepreneurs incur additional expenses to monitor their workers, as described in the efficiency wage theory (Shapiro & Stiglitz, 1984; Calvo, 1985). If the probability of effective monitoring decreases as the labor force expands, it leads to a convex cost associated with labor. Notably, capital does not face shirking concerns, hence the monitoring cost is unaffected by the size of capital. As a result, v depends solely on the labor force.

Another interpretation of v is related to convex hiring and firing costs (Hopenhayn, 1992). Given the absence of frictions in the capital market, the cost of capital is determined solely by its price. In contrast, labor encounters convex costs due to the complexities involved in the hiring and firing processes. Additionally, v can be understood as the problem-solving cost borne by the entrepreneur (Garicano, 2000). With the introduction of automation, tasks become more well-defined, and capital gains the ability to independently resolve these problems. However, labor still encounters challenges that workers are unable to solve independently. To address these issues, workers seek the advice and assistance of the entrepreneur, resulting in a cost in terms of the entrepreneur's time and effort.

It is worth noting that the typical employment size of top entrepreneurs is approximately 100 Smith et al. (2019), which may imply that issues related to span-of-control do not pose a significant problem for them. However, we posit that the lack of substantial scaling in their businesses, despite their success, is indicative of underlying frictions. Given their success and wealth, financial constraints are unlikely to be the primary barrier. Instead, we argue that the absence of scaling can be attributed to challenges related to span-of-control. Consequently, a reduction in this problem through automation emerges as a significant factor influencing top income inequality.

One piece of evidence supporting the idea of convex labor costs is the significant interest entrepreneurs show in seeking advice on employee relations, particularly among large firms. According to the "Business Advice and Planning" module of the Annual Survey of Entrepreneurs conducted by the Census in 2016, more than a quarter of entrepreneurs who sought guidance sought advice specifically on employee relations (U.S. Census Bureau, 2016). In contrast, only 20% sought advice on business finances, and merely 9% sought guidance on investment and access to capital. Notably, it is primarily entrepreneurs running large firms who seek advice on employee relations. On average, the employment size of these entrepreneurs' firms is 65% larger than those seeking any type of advice, while their revenue figures are twice as high. On the other hand, there is no difference between the average firm size of entrepreneurs who seek advice on business finances and those seeking any type of advice. These findings emphasize the significance of employee relations for owners of large firms.

Another evidence of the convex cost of labor is the 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 a convex cost in labor[5].

[5]Labor market power can be seen as an alternative explanation for the rising wage schedules observed. We acknowledge that this simple model may not be sufficient for analyzing the impact of labor market power on inequality. To comprehensively address labor market power, additional structural considerations need to be incorporated. For instance, the fate of workers who become unemployed due to wage cuts should be specified. This assumption would outline the strategic interaction between firms, resulting in a more complex model. The current model can be interpreted as a model with a high degree of labor market power. Under the assumptions that: i) the labor market is segmented by regions and industries, ii) workers cannot move across regions and industries, iii) each region-industry combination has only one entrepreneur and iv) occupational choice is absent, the model's

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

In this paper, we interpret v as the monitoring cost. In appendix B, we provide a micro foundation for v using the efficiency wage theory, which leads to the firm size wage premium. Because the use of the efficiency wage as a micro foundation provides no additional insight, to make the model more tractable, we focus on the reduced form model[6].

The largest firms in the United States are not typically owned by entrepreneurs, as stated by (Smith et al., 2019). According to their research, businesses owned by the top 0.1% of individuals are generally regional enterprises with approximately \$20 million in sales, while C corporations with profits exceeding \$250 million primarily dominate the market. To account for this observation, one can incorporate another sector into the model. This additional sector would consist of large public firms operating under a constant returns to scale production function. Importantly, the primary findings of this paper remain valid even when considering this extended model.

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 $E \subset [z_{min}, z_{max}]$, automation technology $I^\star(z)$, and labor and capital demand $\{\ell_i^\star(z)\}_{i\in[I^\star,1]}, \{k_s^\star(z)\}_{s\in[0,I^\star)}$ 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);*

results are valid. An increase in labor market power can be interpreted as an increase in the convexity of v.

[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 for not shirking. However, our result does not depend on what happens to v.

- *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.*

All proofs are documented in the appendix.

2.3 Characterization of the Equilibrium

The details of the characterization of the equilibrium are provided in the appendix. Here we point out the main features.

2.3.1 Optimal Occupational Choice

As usual in this type of model, there is a cutoff productivity level that determines the optimal occupation. It is easy to see that profit $\pi(z)$ is increasing in z, and, hence, there is a cutoff $z^\star$ such that any individual with $z > z^\star$ becomes an entrepreneur, and others become workers.

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. Because there is no productivity difference between the tasks, an entrepreneur should be indifferent to allocating resources to each task; therefore, she distributes the capital across tasks uniformly. This is also true for labor because of the unit elasticity of substitution between tasks. Unit elasticity leads to the productivity of labor being 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 $C(I^\star) = exp\left(\int\limits_{I^\star}^{1} ln\gamma_i di\right)$. Hence, the productivity level of a task affects each task in the same way, and optimal labor is the same across non-automated tasks.

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

2.3.3 Optimal Automation Level of 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 the optimality condition for labor and capital, leads to the following equation:

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

The solution to this equation is the unconstrained optimal automation level, $I^\star$. Entrepreneurs choose $I^\star$ if it is less than the automation constraint I; otherwise, they choose I. Let $\tilde{I}(z)$ be the solution to (6).

Note 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. *The optimal choice of automation level, $I^\star(z)$ is increasing in z and given by:*

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

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

Proposition 2 indicates that highly skilled entrepreneurs automate more tasks than lowly skilled entrepreneurs. This is so because, as noted 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 the cheaper factor. On the other hand, it decreases the productivity gain from the labor, $C(I)$. Low-productive entrepreneurs automate less in order to benefit from total productivity. As z increases, the benefit of labor productivity decreases; hence, the entrepreneurs prefer cost-effective inputs. Therefore, $I^\star$ is increasing in z. Our explanation resembles one posited by Zeira (1998), who, by studying technological adoption across countries, demonstrates that countries of low productivity have lower wages, and, hence, lower technological adoption than those of high productivity countries. In our model, the wage rate is the same for all firms, and firms differ only in their productivity.

Observe that in partial equilibrium (when R is fixed), the convexity of v does not impact the optimal choice of I because it does not appear in equation (6). This is thanks to two assumptions: constant returns to scale production function and competitive capital market. To see this, first, observe that if optimal $I^\star$ is interior, then the capital intensity is a function of automation choice: $k/\ell = \gamma_I$. The interior solution implies that the entrepreneur is indifferent between using capital or labor for the marginal task. Because effective costs are the same for both technology (an implication of combining conditions (5a) and (5b)), the level of production must be the same, which leads to $k/\ell = \gamma_I$. Thanks to the constant returns to scale assumption implies that the marginal product of capital is a function of capital intensity, productivity, and automation choice. Having established that capital intensity depends on automation level, the marginal product of capital is a function of productivity and automation choice. At the optimum, marginal product of capital must be equal to rental rate. The competitive market assumption for capital market leads rental rate to be independent of the entrepreneur's choice. Hence, automation choice is function of productivity and rental rate.

$$\frac{k}{\ell} = \gamma_{I^\star} \;\; \& \;\; MPK(k/\ell, z, I^\star) = R \implies MPK(z, I^\star) = R \implies I^\star(z, R).$$

Therefore, if R is fixed (for example, in an open economy) the underlying reason for convex labor cost v is not important for the choice of automation. However, it impacts employment level, hence profit. Therefore, in general equilibrium of a closed economy, the shape of v alters the capital and labor demand, changing the rental rate and automation decision of entrepreneurs.

2.4 Top Income Distribution

Now, we can characterize the top income distribution. Individuals whose skill level is below $z^\star$ become workers and earn wages w while individuals whose skill level is 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 need only 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:

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

To develop 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 the functional form of v into the 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, let's focus on high z. Suppose that $z_{max} > R$, in this case, automation technology clearly binds for top skilled entrepreneurs. To see this, consider $z > R$. If an entrepreneur automates all tasks, then she has a linear production function and makes an 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 highly productive entrepreneur's earnings are much higher than those of an entrepreneur of low productivity.

Observe that the convexity of the profit function increases with automation technology I. This is so because the 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 highly skilled entrepreneur. For simplicity, consider two entrepreneurs: one has a high z so that automation technology binds, and the other has a low z that does not automate all automatable tasks. An increase in I does not affect the choice of the entrepreneur who has low z, whereas now the entrepreneur who has high z can enlarge its production and increase its profit. Therefore, the value of relaxing the automation constraint increases with z. This implies that an improvement in automation technology increases the convexity of the profit function.

The convexity of the profit function also increases with a reduction in the monitoring cost-i.e., a decrease in α. The monitoring cost is the principal cause of the decreasing returns to scale. As the monitoring problem is relaxed, entrepreneurs can enlarge their span of control. Since the enlargement is bigger for highly 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 the 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 the 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, the monitoring cost function is $v(L) = L^\alpha$, and $\lambda(1 - I)(\alpha - 1) > 1$.*[7] *Then, the*

[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.

distribution of profits has a Pareto tail with shape parameter $\lambda(1 - I)\frac{\alpha-1}{\alpha}$.[8]

The Pareto parameter gives us a measure for inequality (Gabaix, 2016). The 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 the labor cost make income distribution thicker. As I increases or α decreases, the Pareto parameter goes down, and, consequently, inequality at the top increases.

Under these conditions, decreasing returns come from automation technology and the cost of labor. As discussed above, the convexity of the profit function increases with improvement in I and decreases with α. This leads to top-skilled entrepreneurs capturing a higher share of total profits. As the severity of the diseconomies of scale decreases, top income inequality increases.

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

This result does not depend on occupational choice. Because the most productive entrepreneurs always become entrepreneurs, independent of automation technology, occupational choice does not impact the mechanism here. The same result could be achieved using a fixed type of individuals, with no heterogeneity among workers. An important assumption for this result is that an individual can only have a claim on a single firm, and so the top income distribution mimics the top profit distribution.

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

The primary aim of this paper is not to formulate a theory specifically for a Pareto tail but rather to develop a theory that explains the potential changes occurring within the tail. In pursuit of this objective, we assume that productivity follows a Pareto distribution. One theory that generates a Pareto distribution for productivity is the random growth model, as outlined in the work by Gabaix (2009). In the random growth model, an entrepreneur's productivity evolves following Brownian motion, with a constant growth rate accompanied by stochastic diffusion. Under steady-state conditions, the productivity distribution exhibits a Pareto tail. The shape parameter of this tail is inversely related to the growth rate and positively associated with the exit rate. Our model can be seen as the steady state of a random growth model. However, Gabaix et al. (2016) has shown that the basic random growth model does not adequately account for the rapid decline in the Pareto parameter of income distribution. Consequently, they introduce two modifications that enhance the model's performance. One of these modifications relates to the change in scale dependence or the convexity of the return to skill. Our theory can be viewed as a variation of this scale dependence, allowing for the generation of dynamics in top income inequality that align with data.

2.5 A Measure for Automation: Capital Share

We build our model using the task-based framework developed by Acemoglu & Restrepo (2018b). An implication of this type of production function is that I corresponds to the capital share of income. This is also true in our model. To see this, consider the entrepreneur's first-order condition with respect to capital. This condition implies that the capital share of production within a firm is I, $I^\star(z)zY = RI^\star(z)k^\star(z)$. If automation level binds for every entrepreneur, then in the aggregate, the capital share of income is I:

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

Therefore, the remaining part, $1 - I$, accrues to labor and the entrepreneur. If automation technology binds for every entrepreneur, then the automation level would be equal to capital share. Even though the capital share does not change for low productive entrepreneurs, because the market share of top productive entrepreneurs increases, overall capital share increases. This implies that the measure of automation is highly correlated with the labor share.

The model suggests that an increase in automation technology results in a corresponding increase in capital share. However, ongoing discussions in the academic literature focus on the changes in factor distribution. For instance, Barkai (2020) argues that the decline in labor share has not been compensated by capital share. Nonetheless, these findings are sensitive to underlying assumptions and time frames. Karabarbounis & Neiman (2019), on the other hand, propose that measurement errors in capital might be a more plausible explanation than changes in profit share. As a result, the debate on who benefits from the disappearing labor share remains unresolved.

Furthermore, other factors could contribute to the decline in labor share (see Grossman & Oberfield (2022)). Even though automation is the only reason for the change in capital share in the model, when we test the model prediction in the next section, we control for other possible channels, namely markups and the price of capital. Our approach aligns with the existing literature that employs a task-based production function. Acemoglu & Restrepo (2021) estimate the impact of automation on wage inequality using a task-based framework and also consider labor share of income as an automation measure while controlling for other variables. Similarly, Moll et al. (2021) adopt a similar strategy to analyze the impact of automation on wealth inequality. We follow this literature and use capital share as a proxy for automation and control for other confounding factors.

2.6 Back of Envelope Calculation

Next, we do a 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 approximately 64% to 59% between the 1970s and the 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 an 8% decrease in the Pareto parameter. In the WID data, the estimated Pareto parameter decreased from 2 to 1.5 during the same period. This corresponds to an approximately 25% decrease. In other words, our model can explain a third of the decrease in the Pareto parameter.

Clearly, the automation level constraint might not bind for all entrepreneurs. However, as automation increases, the market share of highly productive (hence, low labor share) entrepreneurs increases. Thanks to the reallocation channel, we expect labor share to decline at a higher rate than the change in automation level. Furthermore, there are other factors that lead to a decline in labor share. In this regard, this analysis provides an upper bound.

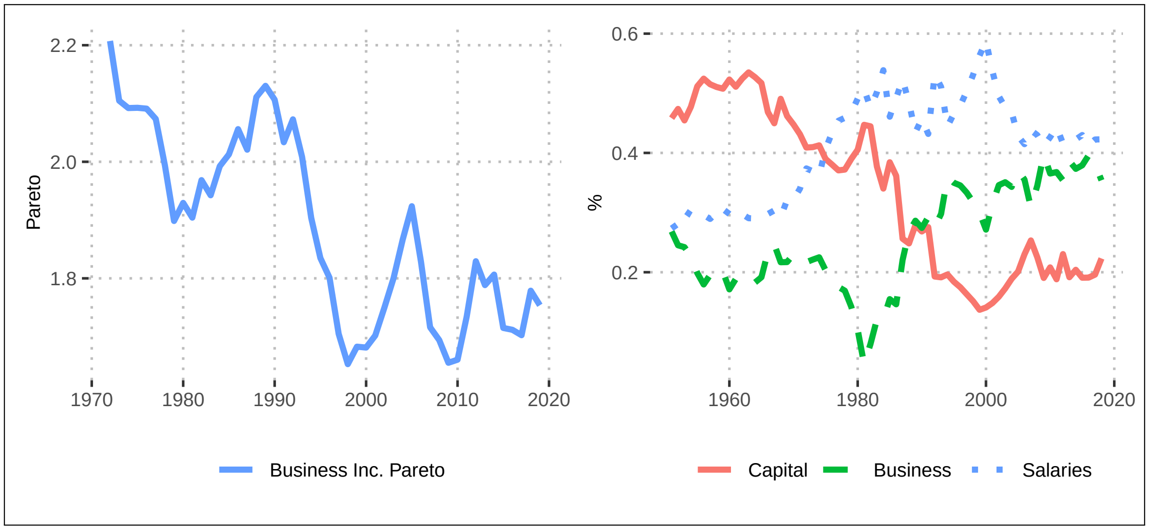
3 Motivating Facts

3.1 Importance of Business Income

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

The right panel of Figure 2 shows the income composition of the top 0.1% (excluding capital gains) across the last 50 years. Since the 1960s, the share of business income (dashed line in the graph) has almost doubled. 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). Sixty percent

Figure 2: Pareto Parameter for Business Income and Income Composition of Top 0.1%



Left panel: 5-year moving averages of the fitted Pareto parameter for business income for top entrepreneurs. Right panel: income composition of top 0.1%. Capital gains are excluded from income.
Source: Author's calculation using IPUMS CPS ASEC and Piketty & Saez (2003).

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). In other words, since the 1960s, the change in the return to entrepreneurial skills has been the main driver of the top income inequality.

The right panel, which shows the importance of business income, does not imply that inequality in business income increased. To show that inequality increased, the left panel of Figure 2 plots the fitted Pareto parameter for business income for top entrepreneurs. It is calculated using IPUMS CPS ASEC (Flood et al., 2018). Details of the estimation are discussed in Section 3.3. We plot the five-year moving averages. The figure shows that since the 1970s, there has been a 10% decrease in the Pareto parameter for business income. This implies an increase in business income inequality among top entrepreneurs. The Pareto parameter shows a similar pattern in the business income share: both inequality and share increased significantly in the mid-1980s, and after 2000 they stabilized.

3.2 Automation and Firm Size Distribution

One of the main assertions of this paper is that automation enables entrepreneurs to scale up their production. This implies that the average firm size increases with the automation level. Equation (8) shows that profit is a power function of employment. Hence, as the convexity of the profit function increases, the convexity of the optimal labor choice as a function of productivity increases. This implies that the employment share of highly productive entrepreneurs is increasing, and, thus, that the average firm size in the entrepreneurial sector is increasing.

To determine whether higher automation leads to larger firms in the data, we regress the change in firm size to a change in automation. Because this result is for the entrepreneurial sector, we only consider the employment distribution across private firms. We obtain the data from the Amadeus database of Bureau van Dijk/Moody's Analytics, which provides information about private firms in European countries. For each industry-country pair, we calculate two measures: average firm size and the share of the top 1% of firms in employment. Because we do not have a direct measure of automation, we use information technology intensity, defined by total IT capital over total capital. Eden & Gaggl (2018) show that an increase in IT intensity is associated with reallocation of income from routine labor to non-rountine labor. This is inline with an increase in automation technology, hence, following Eden & Gaggl (2018), we use IT intensity as a measure for automation. We construct this measure using the data from EU KLEMS. We consider the changes between 2006 and 2016 because the number of observations in the Amadeus database is significantly low for previous years.

Table 1 presents the results. All of the measures of change in firm size distribution are positively correlated with IT intensity. This implies that industries that observed a higher rate of IT growth also observed a higher rate of firm size growth. A percentage increase in the growth of IT intensity leads to an increase in the growth of the average firm size in the employment of about 0.6%. Also, the growth rate of the employment share of the top 1% of firms increases by 0.3%. Furthermore, as Bessen (2017) and Brynjolfsson et al. (2008) show, inter terms of sales, higher IT intensity leads to higher market concentration. Stiebale et al. (2020) estimate that the

Table 1

	Dependent variable:	
	Δlog(Ave. FS)	Δlog(Top Emp Share)
	(1)	(2)
Δlog(IT Intensity)	0.581***	0.276*
	(0.147)	(0.141)
Nobs	182	182
Note:		*p<0.1; **p<0.05; ***p<0.01

impact of robots on productivity and sales is greater in larger firms than in smaller firms. Hence, the data are in line with the model's prediction that automation enables entrepreneurs to grow their businesses.

3.3 Labor Share and the Pareto Parameter

In this subsection, we consider the relationship between labor share and the Pareto parameter in the data. As we discussed above, capital share of income provides a measure for automation level in them model. Following Acemoglu & Restrepo (2021) and Moll et al. (2021), we use labor share for the measure of $1 - I$.

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 in 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:

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

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

3.3.1 Measure for Labor Share

The labor share of income is defined as the total compensation of workers divided by the total income. For the US industry-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 our international level analysis, we use the labor share estimates of Penn World Table 10.0 (Feenstra et al., 2015). Because BEA's data start in 1987, we consider the years between 1987 and 2016. For international comparison, we consider 1961-2015.

Observe that labor share in the model consists of two parts: wage and entrepreneurial income. Accounting for self-employment income is not straightforward because it is difficult 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 methodological issue (Gollin, 2002). Using the split ratio in the corporate sector, PWT divides self-employment income into labor and capital (Feenstra et al., 2015).

3.3.2 Measure for Top Income Inequality

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

We obtain the 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 about the industry; thus, it is difficult to obtain a good estimate of top income share by industry. Therefore, we use the CPS ASEC microdata extracted from IPUMS (Flood et al., 2018). As discussed above, business income is an important component of income for top income earners and our model is also about business income. For this reason, we consider the distribution of income of self-employment workers. Assuming that the right tail of income distribution follows a Pareto distribution, we estimate the Pareto parameter using the maximum likelihood estimator.

A major drawback of public-use microdata is that the income is top coded in the data. Because we are interested in the right tail of the income distribution, a significant fraction of the observations is top-coded. To estimate the Pareto parameter with 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. The maximum likelihood estimator for the scale parameter is q^{th} percentile of the data. We denote it by x_q. Then the 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 indicates whether or not $x_i \leq \bar{x}$ or not. The solution to this problem is

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

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

Because a Pareto distribution can only approximate the right tail of the income distribution, we consider the top 10%. To consistently estimate, we fit the Pareto tail if there are more than 30 observations in an industry-year pair. If there is an insufficient number of observations, we fit the Pareto distribution to the top 15% and decrease the required number of observations to 20. In total, we have 396 estimated parameters for non-agricultural industries[10] between 1987 and 2016.

We take the 3-year averages in order to decrease the short-run fluctuations and reduce the noise of the data. Because the labor share series of BEA starts in 1987, we obtain 9 observations for each industry.

3.3.3 Result

Table 2 shows the regression results. The first three columns record the results using US industry-level data, and the last column records the result using cross-country data. All columns control for time-fixed effects. Columns 1 and 4 show that there is a positive correlation between the labor share and the Pareto parameter, both at the industry level and at the country level. However, the

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

Table 2: Results

	Dependent variable:			
	log(Pareto Parameter)			
	US Industries			International
	(1)	(2)	(3)	(4)
log(Labor Share)	1.312*	1.314*	1.324*	0.723**
	(0.736)	(0.742)	(0.777)	(0.317)
log(Markup)		0.022	0.057	
		(0.292)	(0.291)	
log(Rel. Price of Equipment)			0.178	
			(0.250)	
Time FE	Yes	Yes	Yes	Yes
Observations	124	124	124	288
R^2	0.172	0.172	0.174	0.138

Note:*All standard errors are clustered at the industry or country level. Rows are independent variables and columns are dependent variables. The Pareto parameter for industries is estimated by fitting the Pareto parameter to the top business income distribution using CPS ASEC data. The Pareto parameter for countries is taken from the World Inequality Database. The labor share for industries is taken from BEA. The labor share for countries is taken from Penn World Table version 10.0.*
*p< 0.1; **p< 0.05; ***p< 0.01

estimation is not exactly 1, as predicted by the model.

Automation is not the only cause that might lead to a decrease in labor share. Other reasons for the decrease in the labor share might be the rise in markups (De Loecker et al., 2020), or a decrease in the relative price of equipment Karabarbounis & Neiman (2013). In the third column, we control for the change in markups and attribute all other changes in labor share to automation. To estimate the average markup for each industry and year we employ the estimation method of De Loecker et al. (2020). Including the change in markup does not impact the coefficient of labor share. In the fourth column, we also include the change in the relative price of equipment. We constructed an investment price deflator using BEA's NIPA table and divide it with the personal consumption expenditure deflator to get the relative price. The result does change, the coefficient of labor share remains the same and other parameters are insignificant. In other words, changes in the markup or relative price of capital do not explain the decrease in the Pareto tail. All changes in the Pareto tail can be attributed to changes in labor share that are not related to markup or capital price.

3.4 CEO Compensation

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

In this subsection, we consider the impact of automation on the distribution of CEO compensation. We use Compustat Execucomp for CEO compensation. The data is available since 1992, but the first year has a small sample size, so we dropped the first year. Since the data is not top-coded, it is straightforward to calculate the Pareto parameter for each industry and year. Because these are public firms, we believe that the top of the CEO distribution is populated by the CEOs in this dataset. We repeat the same exercise with the previous subsection by only changing the dependent variable.

Panel A in table 3 shows that in the industries where labor share decreased, the Pareto parameter of CEO compensation decreased. The positive result remains there even after we control for the change in industry-level markups and the relative price of capital goods. In other words, this table provides evidence that the distribution of CEO compensation is also impacted by a

Table 3: CEO Compensation

	Dependent variable:		
	log(Pareto Parameter)		
	(1)	(2)	(3)
Panel A: Pareto			
log(Labor Share)	1.025**	1.311***	1.312***
	(0.406)	(0.509)	(0.506)
log(Markup)		0.863	0.864
		(0.639)	(0.638)
log(Rel. Price of Capital)			0.027
			(0.242)
Panel B: P90/P50			
log(Labor Share)	0.252	0.356	0.362
	(0.344)	(0.354)	(0.351)
log(Markup)		0.313	0.319
		(0.218)	(0.218)
log(Rel. Price of Capital)			0.184*
			(0.107)
Panel C: P99P90			
log(Labor Share)	−1.165***	−1.057**	−1.065**
	(0.302)	(0.427)	(0.431)
log(Markup)		0.323	0.317
		(0.545)	(0.546)
log(Rel. Price of Capital)			−0.210
			(0.255)
Observations	325	325	325
R^2	0.079	0.087	0.087
Note:			*p<0.1; **p<0.05; ***p<0.01

change in labor share that is not related to markup or a change in capital price.

Panel B and panel C show the impact on different percentiles. In panel B, we consider the inequality between the median and the 90^{th} percentile. It turns out that changes in labor share have no impact on any specification. However, there is a negative impact on the difference between 90^{th} and 99^{th} percentile. As predicted by the model, the gap between super-rich CEO and rich CEO is increasing as labor share goes down, and this impact is affected by controlling markup and relative price of capital. This implies that the main impact of automation is for the top of the CEO compensation distribution.

Table 4: Capital Intensity vs CEO Compensation

	Dependent variable:
	log(CEO Compensation)
log(Capital Intensity)	0.047**
	(0.019)
Firm x CEO FE	Yes
Time FE	Yes
Note:	*p<0.1; **p<0.05; ***p<0.01

Another implication of the model is that CEO compensation increases with the capital intensity of the firm. Without the convex labor cost, the homothetic production function leads to the same capital per labor across firms. Because of the convex cost, firm size impacts the relative marginal cost of inputs, hence the capital per labor ratio. Highly productive firms use more capital-intensive technology, and hence, there is a positive relationship between capital per labor and CEO compensation.

To check this relationship in the data, we use the Net Property, Plant, and Equipment from the Compustat data set. We use the investment price deflator to convert it to real value. The price deflator is calculated using BEA's NIPA table. Similarly, we deflated the CEO compensation using the personal consumption expenditure deflator, which we accessed from FRED. We estimated the impact of capital intensity on CEO compensation using firm-CEO pair fixed effects. Table 4 shows that there is indeed a positive relationship between these two variables. As firms increase their capital intensity, they offer higher compensation for their CEOs.

Overall, the pattern in CEO compensation is in line with the prediction of the model. In the industries where labor share is lower, the tail of CEO compensation is thicker and the gap between 99^{th} and 90^{th} percentile is increasing faster than the gap between 90^{th} percentile and the median.

3.5 Monitoring Cost or Automation?

The main result of this paper does not allow us to distinguish the impact of change in the convexity of labor cost function (α) from changes in automation (I). In this section, we present evidence supporting automation as the primary channel. One implication of the model is that a change in α directly affects the capital intensity of firms. However, our analysis of Compustat data shows that the elasticity of capital intensity with respect to firm size has been increasing since the 1980s, which contradicts the model's implication.

To see how α impacts the elasticity of capital intensity with respect to firm size, let's consider the optimal capital intensity of a firm (combining (5a) and (5b) and $L = (1 - I)\ell$, $K = Ik$):

$$log\left(\frac{K}{L}\right) = log\left(\frac{1-I}{I}\frac{w + v'(L)}{R}\right).$$

This implies that for large firms:

$$log\left(\frac{K}{L}\right) \approx constant + (\alpha - 1)log(L).$$

This result indicates that the elasticity of capital intensity with respect to labor is a function of α. The model predicts that as α decreases, this elasticity decreases as well. Intuitively, the optimality condition implies that the marginal rate of technological substitution (MRTS) should be equal to relative marginal costs. Since the production function is homothetic, MRTS is a function of capital intensity. Hence, capital intensity across firms depends on the relative marginal cost. The primary source of difference in marginal costs among firms is α. As α decreases, marginal costs become similar across firms, resulting in similar capital intensity.

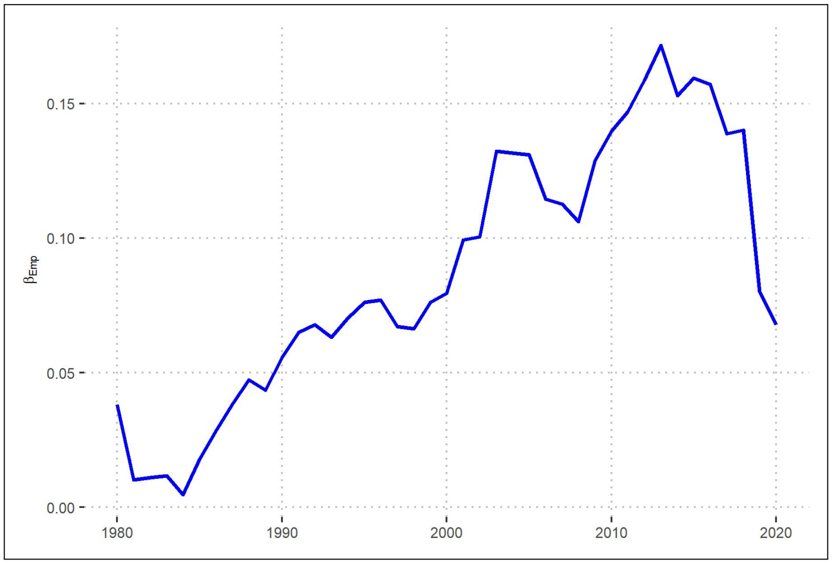
To assess how α is changing, we examine the relationship between capital intensity and employment level. For each year using Compustat data, we estimate the following equation:

$$log(\text{Capital Intensity}_{it}) = \beta_t + \beta_{Emp,t} log(\text{Employment}) + \epsilon_{it}.$$

Figure 3 plots the estimated coefficient. It shows that the elasticity of capital intensity with respect to employment level has been increasing since the 1980s and only started to decline in the late 2010s. Thus, our model predicts that the main channel is not the decrease in α, but the increase in the automation level I. In other words, the difference in capital intensity is increasing across firms. In the model, this implies that wage premium is becoming more important. This might mean that monitoring cost is actually increasing. One explanation for this might be that workers have more ways to shirk. For example: playing games on their phone, messaging with friends, browsing the internet, etc.

It's important to note that this analysis does not entirely rule out the possibility of a decline in α. As discussed earlier, the convex cost of labor might be influenced by labor market power, and our current analysis cannot account for changes in local labor market power. The Compustat dataset only includes data on large firms, and thus, it might not capture firms with local labor market power. Despite this limitation, the evidence still supports the importance of automation, rather than α, in explaining the increase in top income inequality.

Figure 3: Elasticity of capital intensity



3.6 Beyond Top Income Inequality

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

Market concentration: Our model predicts that market concentration, either measured as the top firms' share in sales or as employment, increases with automation. Although we did not prove it formally, it is clear that the share of the top firm in sales and employment size increases as the automation level increases. Autor et al. (2020) show, first, that "superstar" firms are capturing a larger share of the market and, second, this phenomenon is more pronounced in industries in which labor share is falling faster than in other industries. Autor et al. (2020) 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. Moreover, there is also a significant correlation between information technology (IT) intensity and market concentration (Brynjolfsson et al., 2008; Bessen, 2017). We believe 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 of market concentration.

Decreasing Entrepreneurship Rate: One crucial margin in the model is an occupational choice. There are two channels through which automation impacts the entrepreneurship rate. First, it increases the return to entrepreneurial skill, and, hence, leads to a higher entrepreneurship rate. Second, it changes the outside option by altering the equilibrium wage rate. Automation might increase or decrease the wage rate. Similar to Acemoglu & Restrepo (2018a), there are replacement and productivity impacts. First, as capital replaces labor, labor demand decreases, dampening the wage. Second, automation enables firms to allocate workers to more productive tasks, and, therefore, it increases productivity and wages. Due to occupational choice, there is an additional effect in this model. If the increase in wage rate is more significant than the return to entrepreneurship, then the share of individuals who own a business goes down. Otherwise, the entrepreneurship rate declines. Which one dominates depends on the parameters of the model. Numerical exercises (available upon request) illustrate a U-shape relationship between automation level and entrepreneurship rate. In the early stage of automation, the productivity effect dominates, and hence, the marginal entrepreneur becomes a worker. In the later stage of automation, the replacement effect dominates, reversing the marginal individual's decision. Hence, we expect to see a decreasing business dynamism in the early stage of automation. This aligns with the decrease in the start-up rate in the US (Decker et al., 2014; Pugsley & Sahin, 2019; Salgado, 2019).

Wealth Inequality: Wealth inequality in the US also increased during the same period (Saez & Zucman, 2016). For two reasons, entrepreneurial income is one of the important drivers of wealth concentration (Quadrini, 2000; Cagetti & De Nardi, 2006). First, it is risky: each period the entrepreneurs' businesses might fail and lose their business income. Second, because of financial frictions (for example they need to provide collateral), the return to capital is higher for entrepreneurs. This provides an incentive for entrepreneurs to save, and consequently, it generates wealth concentration more than income concentration. Our model can explain the rise in wealth concentration. Because an improvement in automation technology increases income concentration, it leads entrepreneurs to face larger business income risks. When entrepreneurs' businesses fail, they become regular worker ad loses a significant share of their income. Through the first channel, this causes higher wealth concentration. Furthermore, as automation increases, entrepreneurs demand more capital and financial constraints become more severe. Koru (2020) shows that a dynamic version of this model with financial frictions can explain one-fourth of the rise in the share of the top 1% in wealth.

4 Conclusion

Since the 1980s, income distribution in the US has become more skewed. While rich people have been getting richer, the super-rich have become especially rich. In this paper, we argue that improvements in automation technology contributed to the widening gap between top earners. Our theory states that if the cost of labor is convex, then entrepreneurs have a decreasing returns to scale production function. As automation technology improves, dependence on labor deteriorates and the importance of the 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 the convexity of labor cost. We know that automation has increased since the 1970s, and, therefore, we focus on the impact of automation in this paper. However, we believe that the other two parameters are also important and deserve attention.

We provide one explanation for the convex cost of labor: efficiency wage. However, any theory that leads to firm size premiums should deliver similar results. We show that when the firm size wage premium decreases, the gap between large firms and small firms decreases (Bloom et al., 2018; Cobb & Lin, 2017). This might constitute evidence that the monitoring cost decreasing, thus also contributing to top income inequality. However, as we show above, the change in the elasticity of capital intensity to employment size indicates that convexity of labor cost is not decreasing.

Another interpretation of convex cost is labor market power. Even though the current model is too simple to analyze any market power, it provides a first step in that direction. Yeh et al. (2022) show that labor market power in the manufacturing sector was declining between the 1970s and early 2000s, and started to increase sharply afterward. This can be interpreted as a decline in the convexity of the labor cost function. Our model predicts that that leads to a decline in the Pareto tail until the 2000s and a sharp increase afterward. The model fails to explain why top income inequality is not decreasing recently. We think that impact of labor market power and top income inequality is an important question.

Although we see this paper as a model of automation, the model is open to other interpretations. With slight modification (by specifying the supply and price of the outsourcing), it can be seen as a model of outsourcing. We believe that as a research subject, outsourcing is as important as automation, and it is crucial to distinguish them. The outsourcing ramifications of his paper remain to be examined.

Statement: During the preparation of this work the author used ChatGPT in order to improve readability and language. After using this tool/service, the author reviewed and edited the content as needed, and he takes full responsibility for the content of the publication.

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. (2020). Robots and jobs: Evidence from us labor markets. *Journal of Political Economy*, *128*(6), 2188-2244.

Acemoglu, D., & Restrepo, P. (2021, June). *Tasks, automation, and the rise in us wage inequality* (Working Paper No. 28920). National Bureau of Economic Research. Retrieved from `http://www.nber.org/papers/w28920` doi: 10.3386/w28920

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

Barkai, S. (2020). Declining labor and capital shares. *The Journal of Finance*, *75*(5), 2421-2463. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12909 doi: https://doi.org/10.1111/jofi.12909

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

Cagetti, M., & De Nardi, M. (2006). Entrepreneurship, frictions, and wealth. *Journal of political Economy*, *114*(5), 835–870.

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.

De Loecker, J., Eeckhout, J., & Unger, G. (2020, 01). The Rise of Market Power and the Macroeconomic Implications*. *The Quarterly Journal of Economics*, *135*(2), 561-644. Retrieved from `https://doi.org/10.1093/qje/qjz041` doi: 10.1093/qje/qjz041

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

Eden, M., & Gaggl, P. (2018). On the welfare implications of automation. *Review of Economic Dynamics*, *29*, 15–43.

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. R., & 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. (2009). Power laws in economics and finance. *Annual Review of Economics*, *1*(1), 255-294. Retrieved from `https://doi.org/10.1146/annurev.economics.050708.142940` doi: 10.1146/annurev.economics.050708.142940

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.

Gabaix, X., Lasry, J.-M., Lions, P.-L., & Moll, B. (2016). The dynamics of inequality. *Econometrica*, *84*(6), 2071–2111.

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.

Grossman, G. M., & Oberfield, E. (2022). The elusive explanation for the declining labor share. *Annual Review of Economics*, *14*, 93–124.

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.

Karabarbounis, L., & Neiman, B. (2013, 10). The Global Decline of the Labor Share*. *The Quarterly Journal of Economics*, *129*(1), 61-103. Retrieved from https://doi.org/10.1093/qje/qjt032 doi: 10.1093/qje/qjt032

Karabarbounis, L., & Neiman, B. (2019). Accounting for factorless income. *NBER Macroeconomics Annual*, *33*(1), 167–228.

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`

Koru, O. F. (2020). Automation and top wealth inequality.

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

Martinez, J. (2021). Putty-clay automation. *DP16022*.

Moll, B., Rachel, L., & Restrepo, P. (2021). *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., & Sahin, A. (2019). Grown-up business cycles. *The Review of Financial Studies*, *32*(3), 1102–1147.

Quadrini, V. (2000). Entrepreneurship, saving, and social mobility. *Review of economic dynamics*, *3*(1), 1–40.

Rajan, R. G., & Wulf, J. (2006). The flattening firm: Evidence from panel data on the changing nature of corporate hierarchies. *The Review of Economics and Statistics*, *88*(4), 759–773.

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

Saez, E., & Zucman, G. (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *The Quarterly Journal of Economics*, *131*(2), 519–578.

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.

Stiebale, J., Suedekum, J., & Woessner, N. (2020). *Robots and the rise of European superstar firms* (Discussion Paper No. DP15080). CEPR. Retrieved from `https://cepr.org/active/publications/discussion_papers/dp.php?dpno=15080`

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)

U.S. Census Bureau. (2016). *Annual survey of entrepreneurs.* Business Advice and Planning module. Retrieved from `https://www.census.gov/programs-surveys/ase/data/tables.html`

Yeh, C., Macaluso, C., & Hershbein, B. (2022, July). Monopsony in the us labor market. *American Economic Review*, *112*(7), 2099-2138. Retrieved from `https://www.aeaweb.org/articles?id=10.1257/aer.20200025` doi: 10.1257/aer.20200025

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 proves that $w(L)L$ is a convex function. ■

A.1 Proof of Proposition 1

The proposition assumes the distribution of z is bounded above. Here, we prove for unbounded distribution, under the assumption that total labor demand is finite. For the Pareto distribution, parameters must satisfy $\lambda(1-I)(\alpha-1) > 1$, otherwise labor demand is infinite since $\mathbb{E}(L) = \infty$. Clearly, for bounded z, labor demand is finite for positive prices.

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.

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

$$\int\limits_{\tilde{z}}^{\infty} \phi\left[M(R,z) - w\right] dG(z) = G(z^\star), \tag{9a}$$

$$\int\limits_{\tilde{z}}^{\infty} \frac{I^\star}{1-I^\star}\phi\left[M(R,z) - w\right] \frac{M(R,z)}{R} dG(z) = \bar{K}, \tag{9b}$$

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':*

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

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 (9b) *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 $z' \to \infty$, $R_\ell(0,z')$ 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:*

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

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,*

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

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:*

$$1 \approx \int_{z'}^{\infty} \phi\left[M(R_\ell, z)\right] dG(z) > \int_{z'}^{\infty} \phi\left[M(R_k, z)\right] dG(z) \approx 0. \tag{10}$$

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* (9a) *with respect to* $\bar{z}$ *gives us:*

$$-\phi\left[M(\tilde{R}, \bar{z})\right] g(\bar{z}) + \int_{\bar{z}}^{\infty} \phi'\left[M(\tilde{R}, z)\right] M_R(\tilde{R}, z) R'_\ell(\bar{z}) dG(z) = g(\bar{z}). \tag{11}$$

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

$$
\begin{aligned}
0 = &-\frac{I^\star(\bar z)}{1-I^\star(\bar z)}\phi\left[M(\tilde R,\bar z)\right]\left(\frac{M(\tilde R,\bar z)}{\tilde R}\right)g(\bar z) + \int_{\bar z}^{\infty}\left[\frac{I^\star}{1-I^\star}\phi'\left[M(\tilde R,z)\right]M_R(\tilde R,z)R_k'(\bar z)\frac{M(\tilde R,z)}{R}\right.\\
&+\frac{I^\star}{1-I^\star}\phi\left[M(\tilde R,z)\right]\frac{M_R(\tilde R,z)R_k'(\bar z)\tilde R - M(\tilde R,z)R_k'(\bar z)}{\tilde R^2}\\
&\left.+\frac{I_R^\star R_k'(\bar z)}{(1-I^\star)^2}\phi\left[M(\tilde R,z)\right]\frac{M(\tilde R,z)}{\tilde R}\right]dG(z).
\end{aligned}
$$

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:

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

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 $M(R,\bar z)$:

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

Using derivative of labor market condition, equation (11):

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

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* (9a)*, and capital market condition, equation* (9b)*, 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:

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

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:

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

Derivative with respect to z' leads to:

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

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

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

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 con-

nected, 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 :*

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

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, the monitoring cost function is $v(L) = L^\alpha$, and $\lambda(1 - I)(\alpha - 1) > 1$.*[11] *Then, the distribution of profits has a Pareto tail with shape parameter $\lambda(1 - I)\frac{\alpha-1}{\alpha}$.*[12]

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

$$P(\Pi > \pi) = D \left[\alpha \left(\frac{\pi}{\alpha - 1} \right)^{\frac{\alpha-1}{\alpha}} + w \right]^{\lambda(1-I)} \frac{R^{\lambda I}}{C^\lambda}. \tag{12}$$

[11] 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.

[12] 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.

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

■

B 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 the 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.

Individuals 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.

B.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:

$$rV_s(w,q) = w + (\delta + q)\left[U - V_s(w,q)\right], \tag{13a}$$

$$rV_e(w,q) = w - c + \delta\left[U - V_e(w,q)\right], \tag{13b}$$

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 the wage rate should compensate for the 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:

$$rU = b + \mu \int_{\mathcal{X}} \left[\max\{V(x) - U, 0\}\right] dF(x), \tag{14}$$

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

B.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. [13] 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 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:

$$\pi(z) = \max_{\substack{I^\star,\{\ell_i\}_{i\in[I^\star,1]},\\ \{k_s\}_{s\in[0,I^\star)}}} zY - w\left(\int_{I^\star}^{1} \ell_i di\right)\int_{I^\star}^{1} \ell_i di - R\int_{0}^{I^\star} k_i di \tag{15}$$

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

$$\ell_i \geq 0,\ \ k \geq 0,$$

[13] 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 his employees to monitor each other (Che & Yoo, 2001). Replacing labor with capital would decrease 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.

and the output is given by (3).

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.

B.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* (14), (13a),(13b)*;*
- $\pi(z^\star) = rU$*;*
- $\ell(z)$, $k(z)$ *and* $I^\star(z)$ *solve the entrepreneur's problem* (15)*;*
- *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:* $F(w) = \frac{\int_{\underline{w}}^{w} g(z(w'))L(z(w'))}{\int_{\underline{w}}^{\infty} g(z(w'))L(z(w'))}dw'$ *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.*

B.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 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 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 wage dispersion even for workers. Since there is a one-to-one relation between 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.