Repercussions of Endogenous Fast Rising Top Inequality

Seungjin Han*
McMaster University
May 10, 2018

Abstract

This paper develops a fully-solvable equilibrium matching model of incomplete information with early skill acquisition to provide general theoretical insights into fast rising top income inequality observed in the United States. Fast rising top income inequality is endogenously accommodated: In response to a change in each factor contributing to rising inequality, the equilibrium percentage changes in skill investment, income, and firm earnings are monotonically increasing in individual type, switching from negative to positive at respective cutoff types. Rising income inequality is shown to have serious repercussions on welfare and efficiency. A change in each factor contributing to rising inequality makes individuals of type below a corresponding cutoff type worse off but individuals of type above better off. However, only changes in firm-related factors necessarily improve efficiency.

JEL (C78, D31, D82, J30)

*Acknowledgement to be added. Address: Kenneth Taylor Hall, Room 421, Department of Economics, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada, L8S 4M4. Phone: 905-525-9140 ext. 23818. Email: hansj@mcmaster.ca. Fax: N/A.
1 Introduction

In 1965, CEO of major U.S. companies earned 20 times more than a typical worker (Mishel and Davis (2014)). This ratio has grown dramatically over time. From its 2009 low, the worker-to-worker compensation ratio in 2013 had reached the value of 303.4-to-1, a rise of 107.6 since 2009.

The rapid rise in the top income inequality in the United States for the past forty years or so has coincided with the fattening of the right tail of the income distribution (Piketty and Saez (2003), Atkinson, Piketty and Saez (2011)). In particular, if the income distribution at high levels is well approximated by a Pareto distribution, the top income inequality has a fractal pattern, which means that the fraction of income earned by the top 0.1% attributed to the top 0.01% is equal to the fraction of income earned by the top 1% attributed to the top 0.1%, which is equal to the fraction of income earned by the top 10% attributed to the top 1%, and so on. According to Jones and Kim (2017), the U.S. top income inequality is indeed well characterized by the fractal pattern and the fractal share of income is around 25% prior to 1980 but it is closer to 40% in 2015.

However, one might ask why rising top income inequality even matters. Inequality is only a relative concept, so rising top inequality may be even Pareto improving: Everyone is better off but top income earners are a lot better off. If this is the case, the society as a whole would be less resistant to rising top inequality. Therefore, individual welfare analysis in relation to rising top income inequality is quite important. How about the relation between rising top income inequality and efficiency? Can rising top income inequality also tell us the change in the income inequality based on the whole income distribution? Can we address those questions in a model that endogenously accommodates fast rising top income inequality observed in the United States?

For these studies, we incorporate pre-match investment competition (Burdett and Coles (2001), Chade and Lindenlaub (2017), Cole, Mailath and Postlewaite (2001), de Meza and Lockwood (2010), Felli and Roberts (2016), Hopkins (2012), Hoppe, Moldovanu and Sela (2009), Iyigun and Walsh (2007), Nöldeke and Samuelson (2015)) to worker-firm matching with earlier skill acquisition and market imperfections. Our model is sufficiently tractable for us to systematically examine the impacts of a change in each factor contributing to rising top income inequality on individual welfare, efficiency, the Lorenz curve, and the rate of changes in income and other variables across different individuals.

In our model, workers and firms all differ in terms of type and size respectively. A worker’s type, which reflects her inherent ability, is not observable but her talent consists of both type and skill. Workers acquire skill prior to matching with firms and subsequently stable matching occurs in the market.

Our equilibrium matching model is fully solvable given model primitives with
a general class of power functions for a firm’s revenue, a worker’s investment cost, and her talent composition. This is particularly useful in that many variables in economics and finance vary as a power of other variables. For example, a power function has been used to study the distribution of city sizes (Gabaix (1999)), firm sizes (Luttmer (2007)), and the emergence of a Cobb-Douglas function as the production function (Jones (2005)).

Let us explain our key findings. We identify that an advance in production technology, a reduction in investment costs, and an increase in the heterogeneity of firm size relative to worker type\(^1\), induced by changes in the values of power parameters in the model, all increase income inequality.\(^2\) Importantly, the equilibrium elasticities of skill investment, income, and firm revenue derived in our paper show that, in response to a change in each factor contributing to rising inequality, the percentage changes in skill investment, income, and firm revenue are all monotonically increasing in worker type, switching from negative to positive at respective cutoff types. This is robust to the observability of worker type, although cutoff types under observable type are different.

This result suggests that if we treat the distribution of incomes after a change in a factor contributing to rising inequality as the new stationary distribution, the income at the top end of the distribution must grow at a faster rate in the transition. Therefore, fast rising top inequality observed in the U.S. data is endogenously accommodated in our model. This contrasts with Gabaix, Lasry, Lions, and Moll (2016) who notice that the speed of convergence to a new stationary distribution after a shock in the standard random growth model is too slow to match the rapid rise in top income inequality observed in the U.S. data. As a solution to the slow transition, they exogenously introduce heterogeneous mean growth rates, allowing some individuals to grow rapidly.

We then show that fast rising top inequality has serious non-trivial repercussions on welfare and efficiency. Specifically, a change in each factor contributing to rising inequality decreases the equilibrium utility of workers of type below a respective cutoff type but increases the equilibrium utility of workers of type above. This result is also robust to the possibility of the observability of worker type. Therefore, our welfare analysis shows that low income earners actually hurt, whereas high income earners benefit from a change in each factor contributing to rising inequality.

Because of the unobservability of the worker’s type, we can also come up with

---

\(^1\)Firm size can be measured by a firm’s market value, market capitalization, total assets, etc.

\(^2\)The recent literature in the random growth model suggests that entrepreneurs’ innovation and creative destruction (Jones and Kim (2017)), human capital accumulation (Kim (2015)), and knowledge diffusion and accumulation (Luttmer (2015)) are determinants of the (fractal) top income inequality in the U.S., which can be broadly incorporated into the set of factors contributing to rising income inequality identified in our paper.
the measure of efficiency that shows how close our equilibrium is to the equilibrium without market imperfection. Clearly, all workers invest inefficiently more in skill than they would without market imperfections because their investment may signal their inherent type in the market on top of enhancing their talent directly. We show that only changes in firm related factors contributing to rising income inequality (i.e., an increase in the relative heterogeneity of firm size or the homogeneity of the firm’s revenue function with respect to firm size) necessarily improve efficiency.\(^3\)

Firm size in our model reflects the firm’s productivity-related characteristic and the distribution of firm size is treated exogenously. Later in the paper, we generalize our equilibrium matching model for two-sided investment where firms and workers all make investment in their observable characteristics prior to matching and both types of firms and workers are unobservable. The results derived in one-sided investment are qualitatively preserved in two-sided investment but there are important differences.

In two-sided investment, a change in a factor contributing to rising inequality affects the worker’s investment through two different channels. It directly affects her investment given the firm’s investment but it also changes the firm’s investment, which affects the worker’s investment as well. The latter is the secondary impact that is absent in one-sided investment. A change in any factor, except for the relative heterogeneity of firm type, contributing to inequality creates a positive secondary impact on the worker’s skill investment, and subsequently makes equilibrium percentage changes in skill investment, firm revenue and worker pay all increase in worker type more in two-sided investment than they do in one-sided investment. This in turn makes income inequality increase more in two-sided investment.

However, an increase in the relative heterogeneity of firm type has a negative secondary impact. If the firm type distribution is more spread to the right, it is a bad thing for a firm because it now can hire only a manager of a lower type. Other things equal, this dampens the firm’s incentive to invest, which in turn makes the secondary impact on the worker’s investment negative. Subsequently, an increase in the relative heterogeneity of firm type makes equilibrium percentage changes in skill investment, firm revenue, worker pay all increase in worker type in two-sided investment but not as much as in one-sided investment, and income inequality increases in two-sided investment but not as much as well.

Among the literature on pre-match investment, Hoppe, Moldovanu and Sela (2009) and Hopkins (2012) are notably based on the same type of market imperfections as ones in our paper. In Hoppe, Moldovanu and Sela (2009), market

\(^3\)If we apply our results to Landier and Gabaix (2008), we can conclude that the rapid rise in CEO pay due to an increase in the relative heterogeneity of firm size over the past forty year brings efficiency improvement but serious asymmetric welfare effects in the CEO market. See Section 6.4 for details.
imperfections are imposed on market participants on both sides of the market. They engage in pre-match, non-productive investment competition for pure signaling in a non-transferrable utility framework where how to split match surplus is exogenously fixed.\textsuperscript{4} However, a non-transferrable utility framework seems reasonable for European markets with various institutional or government regulations on pay. Our model is rather based on a transferrable utility framework because it captures the nature of worker-firm matching in the United States better in that worker pay in the U.S. market is quite competitively and endogenously determined by freely transferring part of the firm’s revenue to its worker through bargaining.

Our paper is more closely related to Hopkins (2012), who studies only one-sided investment by workers with unobservable type. In his analysis on a transferrable utility model, a worker’s investment and firm characteristic are separable in match surplus, which makes it possible apply the wage determination in the case of complete information to the case of incomplete information.

Our paper however assume that firm size and worker talent (skill and type) are complementary in the firm’s revenue so that larger firms can pay out larger compensations to their workers, as the complementarity seems a driving force of increasingly large worker pay by a larger firm. The main technical innovation of our paper is that we do not apply the wage determination in the case of complete information but we derive the equilibrium worker pay conditional only on worker skill by directly integrating the two components of the marginal return on skill accumulation in the case of incomplete information: the direct productivity effect and the information effect. This makes our equilibrium matching model fully solvable with a general class of power functions and leads to rich applications.

Chade and Lindenlaub (2017) introduce risks in returns on pre-match investments in an environment where investment determines a distribution from which skill is drawn. Their interests differ from ours in that their applications focus on how earnings uncertainty or skill-biased technological changes induce changes in pre-college investment or the gender gap in educational attainment.

Lindenlaub (2017) studies multidimensional matching with no pre-matching investment to show that cognitive skill or task biased technological changes can account for a significant part of the increase in the wage dispersion in the U.S. However, just like the literature on the skill-biased technological changes in the one-dimensional setting, her approach is silent on the fractal property of the top income inequality - the key feature of the income inequality in the U.S., whereas the income inequality derived in our paper is fractal.

\textsuperscript{4}Their innovation is rather the development of a technique for equilibrium derivation with a finite number of agents. Also, see Cole, Mailath and Postlewaite (1995) and Rege (2008) for pre-match investment with a non-transferrable utility framework.
2 Model

We frame an equilibrium matching model with early skill acquisition. There is a continuum of workers. They differ in their inherent ability, called type, denoted by $\theta \in \Theta$, where $\Theta \subset \mathbb{R}^+$ is a connected interval and its infimum is zero. $G(\theta)$ denotes the measure of workers whose types are less than or equal to $\theta$ with positive finite mean and positive density $g(\theta)$ for all $\theta \in \Theta$.

A worker can acquire her skill $y$. A worker’s overall talent $t(y, \theta)$ is composed of both her type $\theta$ and skill level $y$. A worker’s utility by acquiring skill $y$ is her pay $p$ net of the utility cost of acquiring skill $c(y, \theta)$:

$$U(y, p, \theta) = p - c(y, \theta).$$ (1)

Firms differ in terms of size. Firm $x$ has a size $x \in X$, where $X \subset \mathbb{R}^+$ is a connected interval and its infimum is zero. $H(x)$ denotes the measure of firms whose sizes are less than or equal to $x$ with positive finite mean and positive density $h(x)$ for any $x \in X$. If firm $x$ hires a worker with talent $t$, she can generate revenue $f(x, t)$. Therefore, when a worker with talent $t$ works for firm $x$, the firm’s profit becomes

$$\Lambda(x, t, p) = f(x, t) - p.$$

The reservation profit for firms are normalized to zero. We assume a limited liability property for both firms and workers. If we assume worker type and firm size are bounded above, then we denote the supremums of $\Theta$ and $X$ by $\bar{\theta}$ and $\bar{x}$ respectively.

2.1 Power functions

Many variables in economics and finance vary as a power of another. For example, a power function has been used to study the distribution of city sizes (Gabaix (1999)), firm sizes (Luttmer (2007)), and the endogenous emergence of a Cobb-Douglas function as the production function (Jones (2005)).

Jones (2005) views a production function as a choice from a menu of ideas for production technology at each possible input bundle. He shows that if the distribution of ideas in the menu follows a Pareto or Poisson distribution, the production function that results from a choice of ideas is Cobb-Douglas. Following the approach in Jones (2005), we assume that the menus of ideas for revenue technology are the same across worker-firm matches and hence the revenue function in every potential match follows a Cobb-Douglas function:

$$f(x, t) = \alpha t^e x^d$$ (2)
with $\alpha,d > 0$ and $0 < e < 1$. The talent function also takes the form of a Cobb-Douglas function:

$$t(y,\theta) = y^r \theta^s$$  \hspace{1cm} (3)

with $s > 0$ and $0 \leq r < 1$. If $r = 0$, then the investment in $y$ is for pure signaling without enhancing the worker’s talent. If $r > 0$, investment in $y$ not only signals the worker’s type but also her talent as well. Given $f$ and $t$, we can redefine the revenue function in terms of firm size, a worker’s skill and type:

$$\pi(x,y,\theta) = \alpha(y^r \theta^s)^e x^d = \alpha y^b \theta^c x^d,$$  \hspace{1cm} (4)

where $b = re$ and $c = se$. Note that if $r = 0$, then $b = 0$. In this case, skill accumulation is only for pure signaling and is not productive. If $b > 0$, skill accumulation not only signals the worker’s type but also enhances her talent directly so that she creates more revenue.

Note that $\pi_{xy}(x,y,\theta) > 0$ and $\pi_{x\theta}(x,y,\theta) > 0$. Hopkins (2012) assumes $\pi_{xy}(x,y,\theta) = 0$ in a transferable utility framework, but it does not seem reasonable in the literature on worker pay (Gabaix and Landier (2008), Tervio (2008)) in which firm size and talent is complementary in the firm’s revenue (i.e., $f_{xt}(x,t) > 0$). Such complementary is the deriving force for competitive market pay and it ensures that firm size and worker skill are complementary in the firm’s revenue ($\pi_{xy}(x,y,\theta) > 0$) because talent is increasing in worker skill.

The worker’s investment cost function follows a power function:

$$c(y,\theta) = \beta y^m \theta^n$$  \hspace{1cm} (5)

with $\beta, n > 0$ and $m > 1$. We assume that $m-b > 0$. Note that $\lim_{\theta \to 0} c_y(y,\theta) = \infty$ for all $y > 0$.

Note that the talent function $t(y,\theta) = y^r \theta^s$ in (3) implies that skill and type are complementary in talent formation (i.e., $t_{y\theta}(y,\theta) > 0$). Generally, this complementarity is not needed to induce the assortative worker-firm matching in equilibrium in terms of worker type and firm size as Appendix shows. For the assortative matching, we only need the single crossing property of the investment cost function (i.e., $c_{y\theta}(y,\theta) < 0$) and the complementarity between firm size and worker talent in the firm’s revenue (i.e., $f_{xt}(x,t) > 0$).

---

5Jones (2005) sets up the local production function as $\tilde{F}(a_iK,b_iL)$, where $a_i$ and $b_i$ are the choice variables from a menu. Therefore, the global production function we observed is $F(K,L) = \max_{i \in I} \tilde{F}(a_iK,b_iL)$, where $I$ is the menu. $F$ exhibits the constant returns to scale when $\tilde{F}$ exhibits the constant returns to scale such as Cobb-Douglas with the homogeneity of degree one or the Leontief production function. We do not impose the constant returns to scale for general comparative statics.

6We believe that it can be shown that the investment cost function follows a power function if it is a choice of ideas from a menu for investment costs.
Further, restrictions on the ranges of $e$, $r$, and $m$ imply that the revenue and talent functions are concave in $t$ and $y$ respectively but the utility cost function of investment is convex in $y$. The convexity assumption of the cost function is reasonable. The concavity of the revenue and talent function are not needed if worker type is unbounded so that the set of equilibrium skill levels acquired by all workers can span the entire $\mathbb{R}_+$. As shown in Appendix, the concavity assumption kicks in only when we show that it is not profitable for a worker to deviate to acquire skill beyond the highest level observed in equilibrium of the type set $\Theta$ is bounded.

3 Separating Stable Matching Equilibrium

We formulate one-to-one matching between firms and workers through transferring part of revenue from firms to their workers. Let $y(\theta)$ be the worker skill function that characterizes the level of skill that each worker acquires prior to matching as a function of her type $\theta$. The image set of function $y$, denoted by $Y$, includes all levels of skill acquired by workers. Therefore, skill levels in $Y$ are observable in the market. Given the set of all skill levels $Y$ chosen by workers, let

$$\bar{y} := \max Y.$$ 

In practice, workers have established themselves as talented workers in the market with plenty of observable skills and in particular, a more talented worker tends to have a higher level of skill. This makes the market believe that a worker’s type is increasing in skill. Therefore, we make the following monotone assumption on every firm’s belief function $\mu(y)$ of a worker’s type conditional on her skill level.$^7$

**Assumption 1.** $\mu(y)$ is increasing if $y \leq \bar{y}$ and non-decreasing if $y > \bar{y}$.

Because worker type is not observable, there may be multiple equilibria such as separating, pooling or partially pooling equilibrium. Given Assumption 1, the equilibrium turns out to be unique and it is separating in the sense that a worker of a higher type acquires a higher level of skill and the market’s perception of the worker’s talent conditional on her skill is correct (See Proposition 1).

First consider worker-firm matching. Given $\{y, \mu\}$, stable worker-firm matching is characterized by a market matching function $x(y)$ and a market pay function $p(y)$. $x(y)$ specifies the size of the firm for whom a worker works, as an injective function of her skill level $y$, for all $y \in Y$ and $p(y)$ be a pay function that specifies a pay for a worker as a function of her skill level $y$, for all $y \geq 0$.

$^7$Where $\mu(y)$ denotes a worker’s type that the firm believes conditional on her skill $y$. 

7
Since a worker with skill $y$ can get paid $p(y)$ in the market, firm $x$’s problem is to find the skill level of a worker who it wants to hire and her pay as follows.

$$\max_{(y,p)} \left[ f(x, t(y, \mu(y)) - p \right] \text{ subject to } p \geq p(y). \tag{6}$$

Let $\tilde{y}(x)$ be the skill level of the worker who works for firm $x$ in equilibrium, that is, $x(\tilde{y}(x)) = x$. In stable matching, the solution for firm $x$’s problem (6) is $\tilde{y}(x)$ and $p(\tilde{y}(x))$.

Consider the matching decision of a worker with skill $y$. If a worker wants to work for firm $x$, she ensures that the firm’s profit should be at least as high as $f(x, t(\tilde{y}(x), \mu(\tilde{y}(x)))) - p(\tilde{y}(x))$. Therefore, the matching problem for a worker with skill $y$ is to find the size of the firm she wants to work for and pay she is willing to accept as follows.

$$\max_{(x,p)} p \text{ subject to } f(x, t(y, \mu(y)) - p \geq f(x, t(\tilde{y}(x), \mu(\tilde{y}(x)))) - p(\tilde{y}(x)). \tag{7}$$

In stable matching, the solution for problem (7) that a worker with skill $y$ faces in the matching market is $x(y)$ and $p(y)$.

$\{x, p\}$ derived by solving problems (6) and (7) ensures that there are no pairs of a firm and a worker who, by matching and transferring part of the perceived revenue to the worker, can make themselves strictly better off. It leads to stable matching and a worker $y$ works for firm $x(y)$ at pay $p(y)$.

It is worthwhile to mention that Hopkins (2012) formulates the worker’s wage as $w(x, y, \theta)$ with the assumption of the separability between $x$ and $y$ in the match surplus, which allows him to apply wage determination in the complete information case to that with incomplete information. From $w(x, y, \theta)$, one can derive the worker’s wage only conditional on $y$. Our approach that formulates worker pay $p(y)$ as only conditional on $y$ works with or without the separability in the match surplus by directly integrating equilibrium marginal returns on investment in $y$ (See details in Section 3.2).

Now, consider the worker’s decision on skill accumulation beyond the level that is required in her current position. Given $p$, a worker of type $\theta$ solves the following effort exerting problem in the current position in order to acquire skill:

$$\max_y [p(y) - c(y, \theta)]. \tag{8}$$

Problem (8) captures the investment component of acquiring skill $y$ in the current worker position.

We define a separating stable matching equilibrium (SSME) based on the notion of perfect Bayesian equilibrium as follows.
Definition 1 \( \{y, \mu, x, p\} \) constitutes an SSME if

1. \( y(\theta) \neq y(\theta') \) if \( \theta \neq \theta' \),
2. for all \( \theta \), \( \mu(y(\theta)) = \theta \)
3. for all \( \theta \), \( y(\theta) \) solves the worker's problem (8),
4. Given \( \{y, \mu\}, \{x, p\} \) is the pair of the market matching function and the market pay function if \( \{\tilde{y}(x), p(\tilde{y}(x))\} \) is a solution to problem (6) for firm \( x \) and \( \{x(y), p(y)\} \) is a solution to problem (7) for a worker with \( y \).

Part 1 of Definition 1 implies that every worker’s skill fully reveals her type. Part 2 shows that the firm’s belief on the worker type is based on Bayes’ rule. Part 3 defines the worker’s equilibrium effort exerting decision to acquire skill and part 4 defines the market matching function and the market pay function that lead to stable job matching in the worker market.

3.1 Equilibrium characterization

We characterize an SSME. Proposition 1 fully characterizes the worker’s skill accumulation prior to matching and the subsequent stable matching in the market.

Proposition 1 Given monotone belief \( \mu \) stated in Assumption 1, any equilibrium is an SSME \( \{y, \mu, x, p\} \) that satisfies Conditions 1 to 3 below. Condition 4 specifies \( p \) as non-decreasing functions even beyond \( \bar{y} \).

1. \( y \) is a strictly increasing function; \( y(0) = 0 \) and, for all \( \theta > 0 \), the necessary and sufficient condition for a worker of type \( \theta \) to acquire \( y(\theta) \) is
   \[
   p'(y) - c_y(y, \theta) = 0 \text{ at } y = y(\theta) \tag{9}
   \]
2. for all \( \theta \), \( \mu(y(\theta)) = \theta \) and \( \mu(y) = \bar{\theta} \) if \( y > \bar{y} \)
3. Given \( \{y, \mu\}, \{x, p\} \) is the pair of the market matching function and the market pay function if and only if it satisfies (a) and (b) below
   
   (a) for all \( \theta \), \( 1 - H(x(y(\theta))) = 1 - G(\theta) \),
   (b) for all \( y \in Y \)
   \[
   p'(y) = f_t(x(y), t(y, \mu(y)))t_y(y, \mu(y)) + f_t(x(y), t(y, \mu(y)))t_\theta(y, \mu(y))\mu'(y) \tag{10}
   \]
   with \( p(0) = 0 \)
4. for all \( y > \bar{y} \), \( p(y) \) satisfies

\[
f(\bar{x}, t(y, \bar{\theta})) - p(y) = f(\bar{x}, t(\bar{y}, \bar{\theta})) - p(\bar{y})
\]  

(11)

**Proof.** See Appendix. ■

Consider Condition 1. Given the monotone belief \( \mu \) stated in Assumption 1, the worker’s skill accumulation is increasing in her type because of the single crossing property of the cost function. Further, \( y(0) = 0 \) because we assume that \( \lim_{\theta \to 0} c_y(y, \theta) = \infty \) for all \( y > 0 \). Therefore, the SSME is the only type of equilibria that occurs. (9) is self-explanatory for equilibrium skill accumulation given the single crossing property of the cost function.

Condition 2 implies that the market correctly infers the type of a worker conditional on her skill level in equilibrium. Given Assumption 1, the only admissible belief \( \mu(y) \) for all \( y > \bar{y} \) is equal to \( \bar{\theta} \), because \( \bar{y} \) is the skill level acquired by a worker of the highest type, \( \bar{\theta} \).

Condition 3 characterizes stable matching given \( \{\bar{y}, \mu\} \). Condition 3.(a) shows that given the single crossing property of the cost function and the complementarity between firm size and talent in the firm’s revenue function, the market matching function \( x \) satisfies positive assortative matching.

Conditions 3.(b) and 4 characterizes the market pay function. A marginal increase in a worker’s skill \( y \) has two effects. The first effect is the marginal increase in a firm’s revenue due to the increase in her talent directly through the increase in her skill. This is the direct productivity effect of skill accumulation and the first term on the right-hand side of (10) captures it. The second effect is due to the change in a firm’s perception \( \mu(y) \) on her type conditional on \( y \). As a firm hires a worker with a slightly higher \( y \), it believes that the worker’s type is slightly higher. Because a worker’s talent is also increasing in her type, there is a marginal increase in a firm’s revenue due to an increase in the worker’s talent through a marginal increase in her type. This is the informational effect of skill accumulation and the second term on the right-hand side of (10) captures it. If the worker type is observable, the informational effect is absent and the marginal increase in the pay includes the direct productivity effect only. Note that the earning created in the smallest firm is equal to \( \lim_{x \to 0} \alpha t^e x^d = 0 \) given the Cobb-Douglas revenue function. This implies that \( \lim_{y \to 0} p(y) = 0 \), (i.e., \( p(0) = 0 \)) given the positive assortative matching and the limited liability.

There may be some leeway to specify the market pay function beyond the highest observed skill level. Condition 4 specifies the market pay function for a worker with skill beyond the highest level observed in the equilibrium. The right hand side of (11) is the equilibrium profit for the largest firm. Because, for all \( y \geq \bar{y} \), \( \mu(y) = \bar{\theta} \), (11) shows that \( p(y) \) for all \( y > \bar{y} \) tracks down the largest firm’s iso-profit curve associated with hiring a worker of the highest type in the
space of \( y \) and \( p \). Conditions 3.(b) and 4 together make the market pay function non-decreasing for all \( y \), even beyond \( \bar{y} \).

Note that we assume in Section 2 that worker type and firm size are bounded below by \( \bar{\theta} \) and \( \bar{x} \). This is partly because we like to theoretically examine (i) how the market belief for a worker’s type should be formed conditional on her skill level that is never observed in equilibrium and (ii) how the market pay should be determined for a worker with such a skill level. These are characterized by \( \mu(y) = \bar{\theta} \) if \( y > \bar{y} \) in Condition 2 and Condition 4. Neither worker type nor firm size need to be bounded. Proposition 1 without \( \mu(y) = \bar{\theta} \) if \( y > \bar{y} \) in Condition 2 and Condition 4 can characterize the SSME when worker type and firm size are not bounded.

3.2 Equilibrium derivation

Let us show how to derive \( \{y, \mu, x, p\} \) given the characterization of the SSME in Proposition 1.

**Worker skill function:** We start with deriving the *worker skill function*. Define

\[
\tilde{x}(\theta) := x(y(\theta))
\]  

(12)

as the size of the firm who hires a worker of type \( \theta \) in equilibrium. This is derived by solving \( 1 - H(\tilde{x}(\theta)) = 1 - G(\theta) \) for \( \tilde{x}(\theta) \) at all \( \theta \) by Condition 3.(a) in Proposition 1. Since every worker fully reveals her type by acquiring skill in equilibrium (i.e., \( \mu(y(\theta)) = \theta \)), we have \( \mu'(y(\theta)) = 1/y'(\theta) \). Given \( \tilde{x}(\theta) \) and \( \mu'(y(\theta)) = 1/y'(\theta) \), combining (9) and (10) yields

\[
f_t(\tilde{x}(\theta), t(y, \theta)) \left( t_y(y, \theta) + \frac{t_\theta(y, \theta)}{y'} \right) - c_y(y, \theta) = 0.
\]  

(13)

This yields a first-order differential equation \( y' = \phi(y, \theta) \), where

\[
\phi(y, \theta) := \frac{-f_t(\tilde{x}(\theta), t(y, \theta))t_\theta(y, \theta)}{f_t(\tilde{x}(\theta), t(y, \theta))t_y(y, \theta) - c_y(y, \theta)}
\]  

(14)

The initial condition is \( y(0) = 0 \) according to Condition 1. If \( \phi(\cdot, \cdot) \) is continuous in \( \theta \) and Lipshitz continuous in \( y \), then the first-order differential equation \( y' = \phi(y, \theta) \) has the unique solution for \( y \) given an initial condition, according to the Picard-Lindelof Theorem (See Teschl (2012)).

**Firm’s belief function:** The worker skill function \( y \) is a strictly increasing function in the range of \( Y \) according to Condition 1. We can derive the *firm’s belief function* \( \mu(y) \) by deriving the inverse function of the worker skill function in the
range of $Y$, i.e., $\mu(y(\theta)) = \theta$. If $y > \bar{y}$, we set up $\mu(y) = \bar{\theta}$ as stated in Condition 2.

**Market matching function**: Given the firm’s belief function $\mu(y)$, we can derive the *market matching function* $x(y)$ according to Condition 3.(a), $1 - H(x(y)) = 1 - G(\mu(y))$.

**Market pay function**: Finally, consider the *market pay function* $p(y)$. For $y \in Y$, $p(y)$ can be derived by integrating the right-hand-side of (10) with the initial condition $p(0) = 0$ at the bottom match. For $y \geq \bar{y}$, $p(y)$ keeps track of the iso-profit curve of the largest firm according to (11). Therefore, $p(y)$ can be specified as follows:

$$p(y) = \begin{cases} \int_0^y f(x(s), t(s, \mu(s))) (t_0(s, \mu(s)) + t_0(s, \mu(s))\mu'(s)) ds & \text{if } y \leq \bar{y}, \\ f(\bar{x}, t(y, \mu(y))) - f(\bar{x}, t(\bar{y}, \mu(\bar{y}))) + p(\bar{y}) & \text{if } y > \bar{y}. \end{cases} \quad (15)$$

As explained above, the derivation of the SSME is clear. However, in order to derive the SSME, we need to know the functional form of the matching function $\tilde{x}(\theta)$ specified in (12), which is not possible to know unless we know the distributions of firm size and worker type, $H$ and $G$. We assume that $\tilde{x}(\theta)$ follows a power function:

$$\tilde{x}(\theta) = k\theta^q, \quad (16)$$

where $k > 0$ and $q > 0$. $k$ is the “shift” parameter and $q$ is the “relative spacing” parameter. Given $k$, the relative spacing parameter $q$ shows the relative heterogeneity of firm size to worker type. This functional form can be derived under several reasonable distributions for firm size and worker type such as log-normal distributions, exponential distributions, Weibull distributions, Fréchet distributions, and Singh-Maddala distributions (Singh and Maddala (1976)). For example, assume that the distributions of firm size and worker type follow a class of Weibull distributions (Singh and Maddala (1976)). Then, we have $1 - G(\theta) = \exp[-(\theta/\lambda_1)^{z_1}]$ and $1 - H(x) = \exp[-(x/\lambda_2)^{z_2}]$.

In this case, $q = z_1/z_2$ and $k = \lambda_2/\lambda_1^{z_1/z_2}$.

If worker type is bounded above, it would be hard to get a closed-form solution of the market matching function. Section 7.2 examines the case where firm size follows a Pareto distribution with the shape parameter according to Zipf’s Law (Axtell (2001), Luttmer (2007)), but worker type follows either a truncated Pareto or the Gabaix-Landier specification (2008) based on the extreme value theory. Our numerical exercise shows that the empirically estimated matching function $\tilde{x}(\theta) = k\theta^q$, based on the method of the nonlinear least square, explains the assortatively matched pairs of firms and workers, randomly drawn from the specified distributions, extremely well (Think about whether we need to drop the Gabaix-Landier specification (2008)).
We derive the equilibrium outcome by trial and error. It is most important to derive the worker skill function $y$. Then, other equilibrium functions are subsequently derived. We first guess the worker skill function $y$ as a power function:

$$y(\theta) = A\theta^z$$

(17)

and then determine $A$ and $z$ that satisfy (13) (equivalently (14)).

**Proposition 2** The SSME is characterized as follows.

1. $y(\theta)$ follows the form of (17) with

$$A = \left[ \frac{\alpha k^d \left( \frac{bdq + bn + cm}{dq + c + n} \right)}{m\beta} \right]^\frac{1}{m-\beta},$$

$$z = \frac{dq + c + n}{m - b}.$$

2. $\mu(y)$ is the inverse of $y(\theta)$ for all $y$. Given $\tilde{x}$ and $\mu$, $x(y)$ is derived from (12) and it is

$$x(y) = k \left[ \frac{m\beta}{\alpha k^d} \left( \frac{dq + c + n}{bdq + bn + cm} \right) \right]^{\frac{q}{dq + c + n}} y^{\frac{q(m - b)}{y^{dq + c + n}}}$$

(18)

3. for all $y$

$$p(y) = \int_0^y \pi_y(x(z), z, \mu(z))dz + \int_0^y \pi_\theta(x(z), z, \mu(z))\mu'(z)dz$$

(19)

$$= \frac{b(dq + c + n)}{mdq + cm + bn} \pi(x(y), y, \mu(y)) + \frac{c(m - b)}{mdq + cm + bn} \pi(x(y), y, \mu(y))$$

$$= \frac{bdq + cm + bn}{mdq + cm + bn} \pi(x(y), y, \mu(y)),$$

where

$$\pi(y, \mu(y), x(y)) = \alpha k^d \left[ \frac{m\beta}{\alpha k^d} \left( \frac{dq + c + n}{bdq + bn + cm} \right) \right]^{\frac{dq + c}{dq + c + n}} y^{\frac{mdq + cm + nb}{dq + c + n}}.$$

(20)

**Proof.** We can determine $A$ and $z$ that satisfy (13) (equivalently (14)). Because $\lim_{\theta \to 0} c(y, \theta) = \infty$, we must have $y(\theta) = 0$, which is satisfied by (17). Then the market’s belief $\mu(y)$ on worker type is the inverse of $y(\theta)$. $x(y)$ is the size of the firm that hires a worker with skill level $y$. Given $\tilde{x}$ and $\mu$, we can derive $x(y)$ by

$$x(y) = \tilde{x}(\mu(y)),$$

13
which comes from (12). Following (15) yields our solution $\mathbf{p}(y)$ in (19), where $\pi(y, \mu(y), x(y))$ is the revenue created in a match with a worker with skill level $y$.

Before starting the analysis of inequality, note that worker pay is a fixed proportion of the revenue created in her match, as shown in item 3 in Proposition 2, and this proportion is

$$B := \frac{bdq + cm + bn}{mdq + cm + bn} \in (0, 1).$$

This ratio is less than one because $m > b$. One can think of this as the labour income share determined from model primitives.

4 Inequality

4.1 Fractal top income inequality

We first consider top income inequality. Since Pareto (1896), it is well known that an income distribution at high levels is well approximated by a Pareto distribution. Saez (2001) shows that the wage and salary income from the U.S. income tax records in the early 1990s is well approximated by a Pareto distribution. With a Pareto distribution, the probability that income $W$ is higher than $w$ is given by

$$\Pr[W > w] = \left(\frac{w}{w_0}\right)^{-\xi}$$

provided that income is at least above some high level $w_0$.

The share of income going to the top $a$ percentile, denoted by $\tilde{S}(a)$, is $\left(\frac{a}{100}\right)^{1-1/\xi}$ given a Pareto distribution. The larger the value of $\xi$ is, the lower the top income inequality is. Therefore, one can define the top income inequality as follows

$$\psi := \frac{1}{\xi}.$$  \hspace{1cm} (23)

A larger $\psi$ implies higher top income inequality. We can then rewrite $\tilde{S}(a)$ as

$$\tilde{S}(a) = \left(\frac{100}{a}\right)^{\psi - 1}.$$

The top income inequality measure $\psi$ has the fractal nature. Let $S(a) = \tilde{S}(a)/\tilde{S}(10a)$ denote the fraction of income earned by the top $10a$ percent of individuals that actually goes to the top $a$ percent and this is equal to

$$S(a) = 10^{\psi - 1}.$$
(a) does not depend on $a$ and hence $S(0.01) = S(0.1) = S(1)$ and so on. Therefore, the fraction of income earned by the top 0.1% attributed to the top 0.01% is equal to the fraction of income earned by the top 1% attributed to the top 0.1%, which is equal to the fraction of income earned by the top 10% attributed to the top 1%, and so on. These fractions are all the same. Jones and Kim (2017) show that these fractal predictions hold remarkably well in the U.S. (See Figure 5 in their paper): Prior to 1980, the fractal shares are around 0.25, meaning one quarter of the top 10\(\%\)’s income goes to the top 1\(\%\). By the end of the sample in 2015, this fractal share is closer to 0.4.

If underlying worker types at high levels are well approximated by a Pareto distribution with the inequality parameter $\psi_\theta$, our solution for the worker skill function $y(\theta) = A\theta^z$ implies that worker skills at high levels are well approximated by a Pareto distribution with the inequality parameter,

$$\psi_y := z\psi_\theta = \frac{dq + c + n}{m - b} \psi_\theta.$$ 

Equations (19) and (20) show that worker pay $p(y)$ takes the form of $Ry^\nu$ with $\nu := \frac{mdq + cm + nb}{dq + c + n}$. If worker skills at high levels are well approximated by a Pareto distribution with the inequality parameter $\psi_y$, worker pay at high levels is also approximated by a Pareto distribution with inequality parameter,

$$\psi_p := \nu\psi_y = T\psi_\theta,$$ 

where

$$T := \nu z = \frac{dqm + nb + cm}{m - b} > 0.$$ 

This shows that inequality parameter $\psi_p$ is increasing in $T$. Therefore, $T$ is the measure of the fractal income inequality given the underlying type inequality parameter $\psi_\theta$.

Given $\psi_\theta$, income inequality parameter $\psi_p$ is increasing in $T$, which depends systematically on all power parameters $((b, c, d), (m, n), q)$ in the model. $b$, $c$, and $d$ determine production technology. $m$ and $n$ determine a worker’s skill investment cost, and $q$ determines the relative heterogeneity of firm size to worker type in the matching function. We can summarize the effect of changes in power parameters on the top income inequality as follows.

**Proposition 3** The comparative statics on $T$ with respect to power parameters are

$$\frac{\partial T}{\partial b} > 0, \frac{\partial T}{\partial c} > 0, \frac{\partial T}{\partial d} > 0, \frac{\partial T}{\partial m} > 0, \frac{\partial T}{\partial q} < 0.$$

Proposition 3 shows that an advance in production technology ($b \uparrow$, $c \uparrow$, $d \uparrow$), a reduction of investment costs ($n \uparrow$, $m \downarrow$), and an increase in the relative heterogeneity of firm size ($q \uparrow$), induced by changes in the values of power parameters, all increase the fractal top income inequality.
4.2 Lorenz income inequality

High incomes are well approximated by a Pareto distribution, but it is well known that small to mid range incomes are well approximated by a log-normal distribution (Cowell (1995), Lubrano (2016)). The fractal property of inequality does not hold for this class of distributions and hence it is hard to expect that the fractal property of inequality would hold for the entire income distribution in the economy.

We will show that the fractal top income inequality measure $T$ is indeed equal to the Lorenz inequality measure with or without a Pareto assumption on the whole income distribution. The Lorenz inequality measures the inequality at every income percentile. It is the most comprehensive inequality measure because it is based on the Lorenz curve that provides the whole distribution of income shares. Our result implies that we can infer changes in the income inequality at every income percentile from changes in the fractal top income inequality measure $T$ even though the Lorenz income inequality may not have the fractal property.

In general, we can consider the Lorenz curve, $L(u)$, for income inequality analysis over the entire income distribution. It measures the proportion of the aggregate income attributed to the bottom $100u\%$ for every $u \in [0,1]$. Given $y$ and $p$ derived in Proposition 2, one can express worker pay in terms of worker type,

$$\tilde{p}(\theta) := p(y(\theta)) = R\theta^T$$

with a constant $R > 0$ and $T$ defined in (25). Let

$$\eta := \theta^T.$$

Let $F$ and $f$ be the CDF and PDF of this new random variable $\eta$. Let $\eta_u$ be the value of $\eta$ that satisfies $u = F(\eta_u)$ for $u \in [0,1]$. The quantile function is given by $Q(u) = \inf \{ \eta \in \mathbb{R} : F(\eta) \geq u \}$. Then, the Lorenz curve for the income distribution can be derived as follows:

$$L(u) = \frac{\int_0^{\eta_u} R\eta f(\eta)d\eta}{\int_0^\infty R\eta f(\eta)d\eta} = \frac{\int_0^{\eta_u} \eta f(\eta)d\eta}{\int_0^\infty \eta f(\eta)d\eta} = \frac{\int_0^u Q(s)ds}{\int_0^1 Q(s)ds}.$$  

(27)

If the Lorenz curves of two income distributions do not intersect, they can be ordered without ambiguity (Sarabia (2008)). Consider two income distributions $W$ and $\hat{W}$. $W$ Lorenz dominates $\hat{W}$, $W \leq_L \hat{W}$, denoted by if

$$L_W(u) \geq L_{\hat{W}}(u)$$

for all $u$.

If $W$ Lorenz dominates $\hat{W}$ (equivalently $\hat{W}$ is Lorenz dominated by $W$), then $W$ exhibits less inequality than $\hat{W}$ in the Lorenz sense. When we do comparative statistics on inequality in terms of the Lorenz ordering, we utilize the theorem established by Fellman (1976):
Theorem 1 (Fellman (1976)) Given a non-negative random variable $X$ with finite mean, let $\omega(X)$ be continuous and increasing with finite mean. If $\frac{\omega(X)}{X}$ is increasing, $X \leq_L \omega(X)$

This theorem shows that if $\frac{\omega(X)}{X}$ is increasing, then $\omega(X)$ exhibits more inequality than $X$. We can apply this theorem into our inequality analysis.

As one can see in (27), the Lorenz curve for the income distribution is essentially based on the distribution of $\eta = \theta^T$, where $T$ is defined in (25) and it is the fractal top inequality measure. Suppose that changes in exogenous parameters lead to an increase in $T$. Then, applying Theorem 1, we can show that the income inequality worsens in the Lorenz sense.

Proposition 4 If $\hat{T} > T$, $\hat{\eta} = \theta^{\hat{T}}$ is Lorenz dominated by $\eta = \theta^T$.

Proof. Note that $\hat{\eta} = \omega(\eta) = \eta^{\hat{T}}$. Because $\frac{\hat{T}}{T} > 1$, $\frac{\omega(\eta)}{\eta} = \eta^{\hat{T}-1}$ is monotone increasing in $\eta$. Therefore, we have $\eta \leq_L \omega(\eta)$ by Theorem 1. \[\Box\]

The measure $T$, defined in (25), characterizes the fractal top income inequality when high incomes are well approximated by a Pareto distribution. Proposition 4 shows that the income inequality at every income percentile is indeed generally increasing in $T$ in the Lorenz sense, whether or not the whole income distribution follows a Pareto distribution. Therefore, we can say that $T$ is the Lorenz income inequality measure that characterizes overall income inequality based on the whole distribution of income shares even though the fractal property may not hold for mid- or low-range of income. In other words, looking at the changes in the (fractal) top income inequality measure $T$, we can learn whether or not the overall income inequality increases or decreases and the comparative statics on the top income inequality go through for the inequality over the whole income distribution.

The annual report of *Income and Poverty in the United States: 2016* published by U.S. Census Bureau provides the household income shares of quintiles and top 5 percent. Further the Gini Index is also provided and it was 0.403 in 1980 but it went up to 0.481 in 2016. Given income shares reported, the cumulative income shares are shown in the table below.

<table>
<thead>
<tr>
<th></th>
<th>1st q.</th>
<th>≤ 2nd q.</th>
<th>≤ 3rd q.</th>
<th>≤ 4th q.</th>
<th>≤ top 5 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>4.2</td>
<td>14.4</td>
<td>31.2</td>
<td>55.9</td>
<td>83.5</td>
</tr>
<tr>
<td>2016</td>
<td>3.1</td>
<td>11.4</td>
<td>25.6</td>
<td>48.7</td>
<td>55.9</td>
</tr>
</tbody>
</table>

Table 1. Cumulative income shares (q. stands for quintile)\(^8\)

\(^8\)For example, the income of the bottom 80% households is 55.9% of the total household income in 1980.
Although we need to construct the exact Lorenz curve at every percentile, five observations of income shares in a given year tell us quite a lot. Because the Lorenz curve is always non-decreasing, the cumulative income shares shown on the table imply that the 2015 Lorenz curve lies below the 1980 Lorenz curve before top 5 percent. It is very unlikely for the Lorenz curves to cross each other after top 5% given the rapid rise in the fractal top income inequality given the same period of time (Jones and Kim (2017)). Therefore, the U.S. data suggest that the rise in the fractal top income inequality is associated with a downward shift in the Lorenz curve. This is exactly the prediction of Proposition 4.

5 Disproportional Changes and Welfare

To explain the fast rise in top income inequality observed in the U.S., Gabaix, Lasry, Lions and Moll (2016) exogenously introduce heterogeneous mean growth rates, making it possible for some workers to grow rapidly. While our model is not fully dynamic, it is important to see whether or not our model can incorporate, for example, the idea that the increasing rate of income for high types are higher than that for low types in response to a change in a factor contributing to income inequality. Since a change in such a factor would eventually move the economy to another stationary distribution, the higher increasing rate of income for high types can be interpreted as a faster average growth rate of income for high types in the transition periods.

On the other hand, income inequality is a relative concept, so it is possible that inequality rises even when every individual’s income rises. Rising income inequality can be even Pareto improving so that everyone is better off but high income earners may be far better off. If this is the case, then rising income inequality may not be so bad and everyone could be willing to accept it. Therefore, it is also important to analyze individual welfare changes associated with rising income inequality.

To see those two points in our model, we consider worker skill, firm revenue,

---

9They introduce the type dependent earnings growth rate and disproportional impacts of shocks on high incomes.
and worker pay and her utility in terms of worker type in the SSME: For all $\theta \in \Theta$,

$$y(\theta) = \left[ \frac{\alpha k^d (bdq + bn + cm)}{m\beta \left( \frac{dq + c + n}{dq + c + n} \right)} \right]^{\frac{1}{m-b}} \frac{dq + c + n}{\theta^{m-b}}$$

$$\tilde{\pi}(\theta) = \alpha k^d \left[ \frac{\alpha k^d (bdq + bn + cm)}{m\beta \left( \frac{dq + c + n}{dq + c + n} \right)} \right]^{\frac{b}{m-b}} \frac{dmdq + c + m}{\theta^{m-b}}$$

$$\tilde{p}(\theta) = \frac{bdq + cm + bn}{mdq + cm + bn} \alpha k^d \left[ \frac{\alpha k^d (bdq + bn + cm)}{m\beta \left( \frac{dq + c + n}{dq + c + n} \right)} \right]^{\frac{b}{m-b}} \frac{dmdq + c + m}{\theta^{m-b}}$$

$$U(\theta) = \alpha k^d \left[ \frac{\alpha k^d (bdq + bn + cm)}{m\beta \left( \frac{dq + c + n}{dq + c + n} \right)} \right]^{\frac{b}{m-b}} \frac{dmdq + c + m}{\theta^{m-b}}$$

where $\tilde{\pi}(\theta) = \pi(y(\theta), \mu(y(\theta)), x(y(\theta)))$, $\tilde{p}(\theta) = p(y(\theta))$ and $U(\theta) = \tilde{p}(\theta) - c(y(\theta), \theta)$. $\tilde{\pi}(\theta)$ is the revenue of the firm that the worker of type $\theta$ works for. Note that $\tilde{p}(\theta)$ is indeed the same as $\frac{bdq + cm + bn}{mdq + cm + bn} \tilde{\pi}(\theta)$, where $\frac{bdq + cm + bn}{mdq + cm + bn}$ is the proportion of the revenue that goes to the worker and it is defined as the worker's bargaining power, $B$ in (21).

All four functions take the form of a power function

$$y = \varphi(\kappa) \theta^{\rho(\kappa)}, \quad (28)$$

where $\varphi(\kappa) > 0$ is the “shift” coefficient, and $\rho(\kappa) > 0$ is the “spacing” exponent with a parameter vector $\kappa = [b, c, d, q, m, n]$. The partial derivative of $y$ with respect to $i = b, c, d, q, m, n$ is

$$\frac{\partial y}{\partial i} = \theta^{\rho(\kappa)} \left[ \frac{\varphi_i(\kappa) + \varphi(\kappa) \rho_i(\kappa) \ln \theta}{\partial \theta} \right], \quad (29)$$

where $\varphi_i(\kappa)$ and $\rho_i(\kappa)$ are the partial derivatives of $\varphi(\kappa)$ and $\rho(\kappa)$ with respect to $i$ respectively. This means that if $\rho_i(\kappa) \geq 0$, then

$$\frac{\partial y}{\partial i} \geq 0 \text{ if } \theta > \theta_i^* = \exp \left( -\frac{\varphi_i(\kappa)}{\varphi(\kappa) \rho_i(\kappa)} \right), \quad (30)$$

$$\frac{dy}{dt} \leq 0 \text{ if } \theta < \theta_i^* = \exp \left( -\frac{\varphi_i(\kappa)}{\varphi(\kappa) \rho_i(\kappa)} \right). \quad (31)$$

Therefore, if $\rho_i(\kappa) > 0$ ($\rho_i(\kappa) < 0$), then an increase in $i$ increases (decreases) the value of $y$ for $\theta$ that is above the cutoff value $\theta_i^*$ but decreases (increases) the value of $y$ for $\theta$ that is below the cutoff value $\theta_i^*$.

Figure 1 shows the impact of a change in $\kappa$ on a power function $y = \varphi(\kappa) \theta^{\rho(\kappa)}$. For simplicity, assume that $\kappa$ is a scalar with $\rho(\kappa) > 0$ and $\varphi(\kappa) > 0$. Whether a
power function is concave or convex, an increase in $\kappa$ ($\kappa$ to $\kappa'$) changes the graph of the power function from the red solid curve to the blue dotted curve so that such a change results in a decrease in the value of $y$ for $\theta < \theta^*_i$, but an increase in the value of $y$ for $\theta > \theta^*_i$.

If $\rho'(\kappa) < 0$, then the impact of an increase in $\kappa$ is the opposite: An increase in $\kappa$ decreases the value of $y$ for $\theta$ that is above the cutoff value $\theta^*_i$ but increases the value of $y$ for $\theta$ that is below.

More interestingly, we can derive the elasticity of $y$ with respect to $i = b, c, d, q, m, n$ at $\theta$ and it is equal to

$$
\epsilon_{yi}(\theta) = \frac{\partial y_i}{\partial \theta} = \frac{i\varphi_i(\kappa)}{\varphi(\kappa)} + i\rho_i(\kappa) \ln \theta
$$

(32)

This measures the percentage change in $y$ in response to a 1% change in $i = b, c, d, q, m, n$ and we have that if $\rho_i(\kappa) \gtrless 0$, then

$$
\epsilon_{yi}(\theta) \gtrless 0 \text{ if } \theta > \theta^*_i,
$$

(33)

$$
\epsilon_{yi}(\theta) \lesssim 0 \text{ if } \theta < \theta^*_i,
$$

(34)

where $\theta^*$ is the same cutoff value of type in (30) and (31). (33) and (34) express (30) and (31) in terms of elasticity: If $\rho_i(\kappa) > 0$ ($\rho_i(\kappa) < 0$), then $\epsilon_{yi}(\theta)$ is positive (negative) for $\theta > \theta^*_i$ but negative (positive) for $\theta < \theta^*_i$. Most importantly, if $\rho_i(\kappa) > 0$ ($\rho_i(\kappa) < 0$), $\epsilon_{yi}(\theta)$ is monotonically increasing (decreasing) in $\theta$. This feature is important in that it can incorporate the faster rise in top income inequality (Gabaix, Lasry, Lions and Moll (2016)) if we treat the income distribution after a change in $i$ as the new stationary distribution that the economy reaches. This point will be clearer in the next subsection.
5.1 Disproportional changes

We first look at the effect of a change in values of power parameters on the worker’s skill investment and the revenue of the firm that the worker works for. With a slight abuse of notation, let $\epsilon_{yi}(\theta) = \frac{\partial y(\theta)}{\partial i} \frac{i}{y(\theta)}$ denote the elasticity of skill investment for a worker of type $\theta$ with respect to each power parameter $i = b, c, d, q, m, n$. Let $\epsilon_{\pi i}(\theta) = \frac{\partial \tilde{\pi}(\theta)}{\partial i} \frac{i}{\tilde{\pi}(\theta)}$ be the elasticity of the revenue of the firm that a worker of type $\theta$ works for, with respect to each power parameter $i = b, c, d, q, m, n$.

**Proposition 5** For the worker skill function $y(\theta)$, there exists a cutoff type $\theta_i^c$ for each power parameter $i = b, c, d, q, m, n$ such that

1. for $i = b, c, d, q, n$, $\epsilon_{yi}(\theta)$ is negative (positive) if $\theta < \theta_i^c$ ($\theta > \theta_i^c$),
2. for $i = m$, $\epsilon_{yi}(\theta)$ is positive (negative) if $\theta < \theta_i^c$ ($\theta > \theta_i^c$), and
3. $\epsilon_{yi}(\theta)$ is increasing in $\theta$ for $i = b, c, d, q, n$ but decreasing in $\theta$ for $i = m$

For the firm’s revenue function $\tilde{\pi}(\theta)$, there exists a cutoff type $\theta_i^\Delta$ for each power parameter $i = b, c, d, q, m, n$ such that

4. for $i = b, c, d, q, n$, $\epsilon_{\pi i}(\theta)$ is negative (positive) if $\theta < \theta_i^\Delta$ ($\theta > \theta_i^\Delta$),
5. for $i = m$, $\epsilon_{\pi i}(\theta)$ positive (negative) if $\theta < \theta_i^\Delta$ ($\theta > \theta_i^\Delta$), and
6. $\epsilon_{\pi i}(\theta)$ is increasing in $\theta$ for $i = b, c, d, q, n$ but decreasing in $\theta$ for $i = m

**Proof.** $y(\theta)$ follows the form of the power function with

$$\varphi(\kappa) = \left[\frac{\alpha k^d}{m \beta \left(\frac{dq + bn + cm}{dq + c + n}\right)}\right]^{\frac{1}{m-b}} \text{ and } \rho(\kappa) = \frac{dq + c + n}{m - b}.$$ 

The partial derivatives of $\rho(\kappa)$ with respect to $b, c, d, q,$ and $n$ are all positive, whereas the partial derivative with respect to $m$ is negative. Therefore, applying (33) and (34) yield items 1 and 2 above.10 Item 3 follows from (32). Items 4, 5, 6, 7, 8, and 9 are proved similarly.

Proposition 3 in Section 4 shows that a technology advance ($b \uparrow, c \uparrow, d \uparrow$), a reduction of investment costs ($n \uparrow, m \downarrow$), and an increase in the relative heterogeneity of firm size ($q \uparrow$), induced by changes in values of power parameters, all increase income inequality. Combining with Proposition 3, items 1 and 2 in Proposition 5 show that a change in each factor contributing to rising inequality ($b \uparrow,$

---

10All the cutoff type values can be derived in closed forms. For example, we have that $\theta_q^c = \exp \left(\frac{1}{dq + c + n}\right).$
$c \uparrow, d \uparrow, n \uparrow, m \downarrow, q \uparrow$) discourages skill investment by workers of type below the corresponding cutoff type but encourages that by workers of type above. Similarly, items 4 and 5 in Proposition 5 show that a change in each factor $i$ contributing to rising inequality makes the revenue of the firm decrease when its size is smaller than the corresponding cutoff size $\bar{X}(\theta^*_{i})$ but makes the revenue of the firm increase when its size is greater.$^{11}$ Items 3 and 6 in Proposition 5 show that in response to a change in each factor contributing to rising inequality, the percentage changes in skill investment and revenue monotonically increase in worker type, switching from positive to negative at the respective cutoff types.

Let $\epsilon_{pi}(\theta) = \frac{\partial \hat{p}(\theta)}{\partial \theta} \frac{i}{\hat{p}(\theta)}$ be the elasticity of the pay of a worker of type $\theta$ with respect to each power parameter $i = b, c, d, q, m, n$. 

**Proposition 6** For the worker pay function $\hat{p}(\theta)$, there exists a cutoff type $\theta^*_i$ for each power parameter $i = b, c, d, q, m, n$ such that

1. for $i = b, c, d, q, n$, $\epsilon_{pi}(\theta)$ is negative (positive) if $\theta < \theta^*_i$ ($\theta > \theta^*_i$),
2. for $i = m$, $\epsilon_{pi}(\theta)$ is positive (negative) if $\theta < \theta^*_i$ ($\theta > \theta^*_i$), and
3. $\epsilon_{pi}(\theta)$ is increasing in $\theta$ for $i = b, c, d, q, n$ but decreasing in $\theta$ for $i = m$

**Proof.** The proof is similar to the proof of Proposition 5. ■

From Proposition 5 and 6, we can see that a change in each factor contributing to rising inequality makes workers of low types reduce skill investment, accompanied by a reduction in their firm’s revenue and their pay but workers of high types increase skill investment, accompanied by an increase in their firm’s revenue and their pay.$^{12}$ Importantly, items 3 and 6 in Proposition 5 and item 3 in Proposition 6 show that in response to a change in each factor contributing to rising inequality, the percentage changes in skill investment, revenue, and pay are all monotonically increasing in worker type, starting from a negative value, switching from negative to positive at the respective cutoff types.

Most recently, Gabaix, Lasry, Lions and Moll (2016) show that the growth rate of top incomes is too slow in the standard random growth model, compared to the fast rise for the past thirty years or so in the U.S. This is because it assumes that Gibrat’s law holds for income dynamics everywhere in the state space, implying that the growth rate of income is independent of the income level. Instead

---

11Our theory is based on no government intervention. Therefore, whether or not the profits of smaller firms decrease with respect to such a change may depend on tax practices such as corporate tax credits for small businesses.

12Whether or not the pays of managers of low type decrease in practice may depend not only on pay floor (e.g., minimum wage) but also tax credit on small businesses because the pay a proportion of the firm’s profit.
of (exogenously) introducing heterogeneous mean growth rates, our model can endogenously accommodate fast rising top income inequality as shown in Proposition 6, which implies that the percentage change in income is increasing order in worker type. Therefore, if we treat the distribution of incomes after a change in a factor contributing to rising inequality as a new stationary distribution, the income at the top end of the distribution must grow at a higher average rate in the transition periods.

5.2 Welfare analysis

A change in each factor contributing to rising inequality brings changes in both skill investment and pay. Therefore, it is not sufficient to focus on the changes in pay for the welfare effects. Because the equilibrium utility function also follows the form of power function in (28), we can conduct the welfare analysis based on (30) and (31).

Our next Proposition shows that a change in each factor contributing to rising inequality brings serious asymmetric welfare effects.

Proposition 7 For the equilibrium utility function $U(\theta)$, there exists a cutoff type $\hat{\theta}_i$ for each power parameter $i = b, c, d, q, m, n$ such that

1. for $i = b, c, d, q, n$, $\frac{\partial U(\theta)}{\partial i}$ is negative (positive) if $\theta < \hat{\theta}_i$ ($\theta > \hat{\theta}_i$) and
2. for $i = m$, $\frac{\partial U(\theta)}{\partial i}$ is positive (negative) if $\theta < \hat{\theta}_i$ ($\theta > \hat{\theta}_i$).

Proof. We can use (30) and (31) for the proof. ■

Proposition 7 shows that a change in each factor contributing to rising inequality makes workers of type below the corresponding cutoff value worse off but workers of type above better off. This is because the reduction in investment costs for low-type workers associated with less skill investment is not high enough to compensate the reduction in their worker pay. On the other hand, the increase in worker pay for high-type workers is more than the increase in the investment costs associate with more skill investment. It suggests that rising income inequality is not Pareto improving. Its welfare implication is quite severe in that low income earners actually hurt, whereas high income earners benefit. It is also worthwhile to mention that our welfare analysis is done in the model that endogenously accommodates fast rising top inequality observed in the U.S. data.
6 Efficiency

6.1 Benchmark: observable type

Now we formulate the measure of efficiency. For that, we first consider the benchmark case where a worker’s type is observable. It is straightforward to show that in this case, the stable matching equilibrium is still assortative in terms of firm size and worker type. Let \( p_e(y) \) be the equilibrium worker pay. In the case of observable type, the equilibrium worker pay includes only the direct productivity effect but not the informational effect. Therefore, the equilibrium marginal worker pay becomes

\[
p_e'(y) = f_t(x_e(y), t(y, \mu(y)))t_y(y, \mu(y))
\]

where \( x_e(y) \) is the assortative market matching function.

Given this marginal return on investment, the optimal investment \( y_e(\theta) \) by a worker of type \( \theta \) satisfies

\[
p_e'(y_e(\theta)) - c_y(y_e(\theta), \theta) = \pi_y(x_e(y_e(\theta)), \mu(y_e(\theta)), y_e(\theta)) - c_y(y_e(\theta), \theta) = 0.
\]

Because \( x_e(y_e(\theta)) = \tilde{x}(\theta) \) defined in (16), solving the equation above yields

\[
y_e(\theta) = \left(\frac{\alpha bk^d}{m\beta}\right)^{\frac{1}{m-\beta}} \theta^{\frac{\alpha d+1}{m-\beta}}.
\] (35)

Then, we can derive the market matching function \( x_e(y) \) as

\[
x_e(y) = k \left(\frac{m\beta}{\alpha bk^d}\right)^{\frac{q}{dp+e+n}} \frac{q(m-b)}{y^{dp+e+n}}.
\] (36)

The equilibrium worker pay is then

\[
p_e(y) = \int_0^y \pi_y(x_e(z), \mu(z), z)dz = \frac{bdq + bc + bn}{mdq + mc + bn} \pi(x_e(y), \mu(y), y),
\] (37)

where

\[
\pi(x_e(y), \mu(y), y) = \alpha k^d \left(\frac{m\beta}{\alpha bk^d}\right)^{\frac{dq+e}{dp+e+n}} y^{\frac{dnq+cm+bn}{dp+e+n}}.
\] (38)

Note that the equilibrium worker pay is a fixed proportion of the earning \( \pi(x_e(y), \mu(y), y) \), which is

\[
B_e := \frac{bdq + cb + bn}{mdq + cm + bn} \in (0, 1)
\]
This can be seen as the worker’s bargaining power with observable type. It is easy to check $B_e < B$ given $m > b$ so that the labor income share in the case of observable type is lower that that in the case of unobservable type.

The (fractal or Lorenz) inequality measure can be derived similar to the inequality measure $T$ defined previously in the case of unobservable type. Note that the coefficients of $y_e$ and $p_e$ are different from those of $y$ and $p$ respectively, but their powers are identical to those of $y$ and $p$ respectively. Because the inequality measure only depends on powers, the inequality measure in the case of observable type, denoted by $T_e$, is the same as the inequality measure in the case of unobservable type:

$$T_e := \frac{dqm + nb + cm}{m - b} = T.$$  

Therefore, the observability of worker type does not affect inequality, whereas it does affect the labor income share.

### 6.2 Measure of efficiency

To formulate the measure of efficiency, first consider the composition of worker pay in (19). Worker pay consists of two parts. The first term in the first line of (19) is the part of worker pay attributed to the direct productivity of skill, whereas the second term is the part of worker pay attributed to the informational effect of skill. The presence of the informational effect drives over-investment by workers. The equilibrium market matching is assortative in terms of firm size and worker type regardless of the observability of worker type. Therefore, the measure of efficiency is measured by how close the worker’s investment in the case of unobservable type is to that in the case of observable type, which is the efficient investment.

For efficiency analysis, Hopkins (2012) examines whether the equilibrium skill accumulation increases or decreases when the distribution of agents on either side changes (See Proposition 8 p.312 in Hopkins (2012)). However, such analysis is incomplete. Both the equilibrium skill accumulation and the efficient skill accumulation change as an exogenous change occurs. Therefore, a correct measure of efficiency is the ratio of the efficient skill accumulation to the equilibrium skill accumulation. Given our solutions for $y_e(\theta)$ and $y(\theta)$, this ratio does not depend on $\theta$ and it is

$$E := \frac{y_e(\theta)}{y(\theta)} = R^\frac{1}{m-\sigma},$$  (39)

where $R$ is defined as the proportion of worker pay attributed to the direct productivity effect, which is the ratio of the first term on the right-hand side of (19)
to the sum of the two terms on the right-hand side:

\[ R := \frac{\text{pay due to direct productivity effect}}{p(y)} = \frac{b(dq + c + n)}{mdq + cm + bn} \pi(x(y), y, \mu(y)) + \frac{bdq + bc + bn}{bdq + cm + bn} \pi(x(y), y, \mu(y)) = \frac{bdq + bc + bn}{bdq + cm + bn} \] (40)

Because \( m > b \), \( R \) is less than one and \( 1 - R \) is the proportion of worker pay attributed to the informational effect. As \( R \) increases, the proportion of worker pay attributed to the direct productivity effect increases, whereas that attributed to the informational effect decreases. (39) shows that efficiency is positively related to the proportion of worker pay attributed to the direct productivity effect.

### 6.3 Inequality versus efficiency

Since the measure of efficiency \( E \) is fully characterized by power parameters, we can conduct comparative statics as follows

**Proposition 8** The comparative statics on \( E \) with respect to power parameters are as follows

\[
\frac{\partial E}{\partial c} < 0, \quad \frac{\partial E}{\partial d} > 0, \quad \frac{\partial E}{\partial q} > 0, \quad \frac{\partial E}{\partial n} > 0
\]

but the signs of \( \frac{\partial E}{\partial b} \) and \( \frac{\partial E}{\partial m} \) are inconclusive.

First of all, equality and efficiency do not always go in the opposite directions. Propositions 3 and 8 show that an advance in production technology associated with unobservable worker type (\( c \uparrow \)) decreases efficiency but increases inequality as well. A reduction of investment costs does not necessarily induce an increase in efficiency. It does only when it comes from an increase the value of \( n \), in which case both efficiency and inequality increase.

If a reduction of investment cost comes from a decrease in the value of \( m \) (i.e., the homogeneity of the cost function with respect to investment), a change in efficiency is unclear. To see this, recall that the measure of efficiency is

\[ E = R^{\frac{1}{m-b}} = \left[ \frac{bdq + bc + bn}{bdq + cm + bn} \right]^{\frac{1}{m-b}} \]

As \( m \) decreases, \( R \), the ratio of worker pay attributed to the direct productivity effect increase, which increases \( E \) given the value of power \( \frac{1}{m-b} \). However, a decrease
in \( m \) increases the value of power \( \frac{1}{m-b} \), which decreases \( E \) because \( R \) is less than one. In fact, the partial derivative of \( E \) with respect to \( m \) is

\[
\frac{\partial E}{\partial m} = \frac{E}{m-b} \left[ -\frac{c}{bdq + cm + bn} - \frac{1}{m-b} \ln R \right]
\]

and it shows the two opposite effects of a change in \( m \): the firm term in the bracket is negative but the second term is positive because \( 0 < R < 1 \) and hence \( \ln R < 0 \). Therefore, the effects of changes in investment costs are not clear unless we can clearly say changes in investment costs only comes from changes in \( n \).

For the same reason for an ambiguous sign of \( \frac{\partial E}{\partial m} \), the effect of an increase in \( b \) (i.e., an increase in the homogeneity of the revenue function with respect to observable skill) on efficiency is unclear. Therefore, the effects of changes in the homogeneity of the firm’s revenue with respect to talent are not clear unless changes in the homogeneity comes only from inherent type-related talent (\( c \)) but not skill-related talent (\( b \)).

**Remark 1** Propositions 3 and 8 show that only changes in firm-related factors contributing rising income inequality (i.e., an increase in the homogeneity of the firm’s revenue with respect to firm size (\( d \uparrow \)) and an increase in the relative heterogeneity of firm size (\( q \uparrow \)) necessarily increase efficiency.

Remark 1 has important implications on studies on worker pay (e.g., Gabaix and Landier (2008)).

### 6.4 Gabaix and Landier (2008)

Gabaix and Landier (2008) show that there has been a significant increase in firm size with the right tail getting increasingly fatter over the last thirty years. In our formation, it is reflected by an increase in \( q \), i.e., an increase in the relative heterogeneity of firm size. Given their parsimonious set up of the revenue function, \( f(x,t) = \alpha t x^d \) with no market imperfections and no investment, they focus on the CEO pay and find that the rapid increase in CEO pay in the United States between 1980 and 2003 can be fully attributed to the six fold increase in the firm size. However, their result does not address inequality nor efficiency because the focus is the relation between the increase in the firm size and the rapid increase in CEO pay.

Our analysis shows that the increase in the relative heterogeneity of firm size observed in the last thirty years or so not only has increased the top income inequality but it has shifted the whole Lorenz curve downward even when the income distribution is not a Pareto distribution. Further, Propositions 3 and 8 show that such an increase in the relative heterogeneity of firm size increases
both efficiency and inequality (not only the top fractal income inequality but the overall inequality based on the Lorenz ordering). In fact, we can have the following convergence theorem for the SSME.

**Theorem 2** As \( q \to \infty \) (or \( d \to \infty \)),

1. the SSME under incomplete information converges to the stable matching equilibrium under complete information with the efficiency measure \( E \) converging to one,

2. the inequality measure \( T \) in the SSME under incomplete information increases indefinitely, and

3. worker’s equilibrium bargaining power \( B \) converges to \( \frac{b}{m} \).

**Proof.** Consider the two equilibria, one under incomplete information and the other under complete information. Then, we can see that

\[
\lim_{q \to \infty} \frac{y_e(\theta)}{y(\theta)} = \lim_{q \to \infty} E = 1,
\]

\[
\lim_{q \to \infty} \frac{x_e(y)}{x(y)} = 1,
\]

\[
\lim_{q \to \infty} \frac{p_e(y)}{p(y)} = \lim_{q \to \infty} R = 1.
\]

Because \( \mu_e(y) \) and \( \mu(y) \) are the inverses of \( y_e(\theta) \) and \( y(\theta) \) respectively, it is also easy to check that

\[
\lim_{q \to \infty} \frac{\mu_e(y)}{\mu(y)} = 1.
\]

The limit results above show the convergence of the SSME under incomplete information to the stable matching equilibrium under complete information as \( q \to \infty \). Furthermore, we have that \( \lim_{q \to \infty} T = \infty \) and \( \lim_{q \to \infty} B = \frac{b}{m} \). The limit results are the exactly same as \( d \to \infty \).

Theorem 2 show that the homogeneity \( (d) \) of the revenue function with respect to firm size and the relative heterogeneity \( (q) \) of firm size to worker type both have the same qualitative effects on efficiency, inequality, and worker’s bargaining power. It is also worthwhile mentioning the worker’s bargaining power. Even when the worker market has the same measures of firms and workers, an increase in \( q \) means that the distribution of firm size is more skewed towards larger sizes and it gives a more bargaining power to firms and less to workers. So does an increase in \( d \).
Another interesting aspect of worker’s bargaining power, \[ B = \frac{bdq + cm + bn}{mdq + cm + bn} \] is that for all \( q \) or for all \( d \), it is bounded above by \( b/m \) given the values of the other power parameters in the case of productive investment \((b > 0)\). However, it uniformly converges to zero as \( q \to \infty \) or \( d \to \infty \) when a worker’s investment is for pure signaling \((b = 0)\) \((B = \frac{c}{dq+c} \) in this case).

7 Robustness

7.1 Complete information

Our results in Section 5 depart from Gibrat’s law in that in response to a change in each factor contributing to rising income inequality, the equilibrium percentage changes in skill investment, revenue, and pay are all monotonically increasing in worker type, switching from negative to positive at respective cutoff types. A change in each factor contributing to rising income inequality also makes workers of type above a corresponding cutoff value better off but workers of type below worse off.

While our analysis is conducted in the case of unobservable worker type, we can conduct the same analysis in the case of observable type where the worker skill, firm revenue, worker pay and her utility in terms of worker type in equilibrium are as follows:

\[
\begin{align*}
y_e(\theta) &= \left(\frac{\alpha bk^d}{m^\beta}\right)^{\frac{1}{m-\delta}} \theta^\frac{qd+e+n}{m-b}, \\
\tilde{\pi}_e(\theta) &= \alpha k^d \left[\frac{\alpha k^d}{m^\beta}\right]^\frac{b}{m-\delta} \theta^\frac{mdq+cm+nb}{m-b}, \\
\tilde{p}_e(\theta) &= \frac{bdq + bc + bn}{mdq + mc + bn} \alpha k^d \left(\frac{\alpha bk^d}{m^\beta}\right)^\frac{b}{m-\delta} \theta^\frac{dq + e + n}{mdq + mc + bn} - \frac{1}{m} \theta^\frac{dmq + cm + bn}{m-b}, \\
U_e(\theta) &= \alpha k^d b \left(\frac{\alpha bk^d}{m^\beta}\right)^\frac{b}{m-\delta} \left(\frac{dq + e + n}{mdq + mc + bn} - \frac{1}{m}\right) \theta^\frac{dmq + cm + bn}{m-b},
\end{align*}
\]

where \( \tilde{\pi}_e(\theta) = \pi(y_e(\theta), \mu(y_e(\theta)), x_e(y_e(\theta))) \), \( \tilde{p}_e(\theta) = p_e(y_e(\theta)) \) and \( U_e(\theta) = p_e(\theta) - c(y_e(\theta), \theta) \). According to (30), (31), (33), and (34), the sign of the partial derivative of the exponent of \( \theta \) with respect to \( i = b, c, d, q, m, n \) determines whether the functional value increases or decreases as variable \( i \) increases, and equivalently whether the elasticity of the functional value with respect to \( i \) is positive or negative. Because the exponents of \( \theta \) in the worker skill, pay and utility functions are exactly the same as the counterparts in the case of unobservable worker type, Propositions 5, 6, and 7 hold in the case of observable worker type with different cutoff types.
Furthermore, the exponents of $\theta$ in both $\tilde{p}_e(\theta)$ and $\tilde{p}(\theta)$ are in fact the measure of fractal top income inequality $T$. Applying the result in Section 4.2, we can conclude that the measure of the fractal top income inequality is also the measure of the Lorenz income inequality in the case of observable worker type. Therefore, the results in Section 5 are robust to the possibility of the observability of worker type.

7.2 Bounded worker type

The analytical results for the SSME are based on the assumption that the matching pattern in terms of firm size and worker type follows a power function $\tilde{x}(\theta) = k\theta^q$. As mentioned in Section 3.2, this functional form can be derived under several reasonable distributions for firm size and worker type such as log-normal distributions, exponential distributions, Weibull distributions, Fréchet distributions, and Singh-Maddala distributions.

The random variables characterized by those distributions are not bounded above. This seems fine for firm size since it is well characterized by an unbounded distribution, in particular an (unbounded) Pareto distribution, following Zipf’s Law (Axtell (2001), Luttmer (2007)). A Pareto distribution is characterized by shape and scale parameters, $\lambda$ and $z_m$ given the cumulative distribution

$$\Pr(Z > z) = \left(\frac{z_m}{z}\right)^\lambda.$$  \hspace{1cm} (41)

Zipf’s Law implies that firm size follows a Pareto distribution with $\lambda = 1$.

However, one may wonder whether or not unobservable worker type can be well characterized by an unbounded distribution. If we think of our model as an approximation of large but finite number of firms and workers, an unbounded distribution for worker type may well approximate the finite economy, based on the finite number of observations that are randomly drawn from the unbounded distribution.

If the distribution for worker type is bounded above, whereas the distribution for firm size is not, it is hard to derive a closed-form solution of the matching function $\tilde{x}(\theta) = k\theta^q$. In order to see the validity of our assumption on the matching function, $\tilde{x}(\theta) = k\theta^q$, in this case, we do a numerical exercise based on firms and workers that are randomly drawn from distributions of firm size and worker type we choose.

Based on Zipf’s law, we assume that firm size follows a Pareto distribution with shape parameter $\lambda = 1$. We fix the scale parameter equal to 0.1. If worker type follows the Pareto distribution with the same shape and scale values, we can derive the closed-form solution of the matching function $\tilde{x}(\theta) = k\theta^q$ as $\tilde{x}(\theta) = \theta$, i.e., $k = 1$ and $q = 1$ so that it is a 45-degree line.
Figure 2
Suppose that worker type follows a (truncated) Pareto distribution given the values of shape and scale parameters equal to 1 and 0.1 respectively. Scatter plots in Figure 2 show the assortative matching pattern based on 999 observations that are randomly generated by the Pareto distribution for firm size and the truncated Pareto distribution for worker type: Figure 2.(a) with the upper bound 10 for worker type and Figure 2.(b) with upper bound 100 for worker type. Green curves represent the non-parametrically fitted matching functions, whereas orange curves represent the matching function, $\tilde{x}(\theta) = k\theta^q$, estimated by the method of nonlinear least square (NLS). Straight black dotted lines in both figures are the matching function, $\tilde{x}(\theta) = \theta$ when both firm size and worker type are unbounded.

First of all, the figures show that the NLS estimation of $\tilde{x}(\theta) = k\theta^q$ explains the actual assortatively matched pairs of firms and workers very well as does the non-parametrically fitted matching function. When the upper bound of worker type is 10, the NLS provides $k = 0.002$ and $q = 4.971$. When the upper bound of worker type is 100, it provides $k = 0.481$ and $q = 1.347$. Therefore, as the upper bound of worker type increases, the value of $q$ decreases and gets closer to 1, which is the value of $q$ when both worker type and firm size are unbounded.

An alternative exercise is possible. Based on the extreme value theory, Gabaix and Landier (2008) suggest that worker talent, especially at the top end, may follow the bounded distribution

$$
\text{Pr}(T > t) = B (T_{\text{max}} - t)^{1/\beta}
$$

where $T_{\text{max}}$ is the upper bound of talent so that $t \in [0, T_{\text{max}}]$. Given the competitive assignment model without skill accumulation and market imperfections, they argue that $\beta = 2/3$, based on the pay to firm-size elasticity, taking the distribution of talent as given.

However, talent is a function of both observable skill and unobservable type in our model and hence it is the type distribution that we need to fix, but not the talent distribution. We assume that worker type also follows the class of distributions satisfying (42). The power parameter, $\beta$ for the distribution of worker type should be smaller than 2/3 because (i) talent is not only determined by worker type but also enhanced by skill accumulation and (ii) skill accumulation has a higher impact on talent under unobservable type than it has under observable type.

Figure 3 provides scatter plots of assortatively matched pairs of firms and workers, non-parametrically fitted curve, and the NLS estimation of the matching function $\tilde{x}(\theta) = k\theta^q$ for each of four values of $\beta = 1/12, 1/9, 1/6, 1/3$ with

$^{13}$Given $T_{\text{max}}$, we set $B$ equal to $\frac{1}{T_{\text{max}}^{1/\beta}}$ based on $\text{Pr}(T > 0) = 1$.

$^{14}$This second reason stems from the fact that managers accumulate more skill under unobservable type.
Figure 3
$T_{\text{max}} = 10$, while maintaining the Pareto distribution for firm size with the shape and scale parameters being 1 and 0.1 respectively.\textsuperscript{15} The NLS estimates of $k$ and $q$ are as follows.

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$1/12$</th>
<th>$1/9$</th>
<th>$1/6$</th>
<th>$1/3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>$2.13 \times 10^{-2}$</td>
<td>$1.56 \times 10^{-4}$</td>
<td>$9.63 \times 10^{-7}$</td>
<td>$8.59 \times 10^{-16}$</td>
</tr>
<tr>
<td>$q$</td>
<td>$7.25$</td>
<td>$7.93$</td>
<td>$9.57$</td>
<td>$10.78$</td>
</tr>
</tbody>
</table>

Table 2 NLS estimates of the matching function

As one can see, the matching functions $k\theta^q$ based on the NLS estimation explain the scatter plots of assortatively matched pairs of firms and worker very well, as non-parametrically fitted matching functions do in all values of $\beta$. It is noticeable that $k$ is decreasing in $\beta$ but $q$ is increasing in $\beta$. In particular, the pattern of $q$ is interesting. A higher value of $\beta$ implies that a fewer workers are on the high end of the type distribution but a lot more workers are on the low end of type distributions. Therefore, an increase in $\beta$ translates into an increase of $q$ in the matching function.

8 Two-sided investment

Firm size can be interpreted as something that reflects the firm’s productivity-related characteristic. We assumed that the distribution of firm size is given but it is very well possible that firms also invest in their productivity-related characteristic. Then, firms and workers both invest their observable productivity-related characteristics, $x$ and $y$, respectively prior to matching. We can extend our analysis to the case with two-sided investment as shown in Appendix B. Let me explain key results with two-sided investment.

8.1 Equilibrium derivation

The firm’s revenue function is

$$\pi(x, y, \theta, \delta) = \alpha x^a y^b \theta^c \delta^d,$$

with $a, b \geq 0$ and $\alpha, c, d > 0$. $\theta > 0$ denotes worker type as before, whereas $\delta > 0$ denotes firm type. The investment cost functions for the worker of type $\theta$ and the

\textsuperscript{15}999 observations for both firm size and manager type were randomly generated from the distributions for each case.
firm of type $\delta$ are respectively given by
\[
c(y, \theta) = \beta^y \frac{m}{\theta^n},
\]
\[
e(x, \delta) = \tau^x \frac{u}{\delta^v},
\]
with $\beta, \tau > 0, m, u > 1$ and $n, v > 0$. Further we assume that $(u - a)(m - b) > ab$ with $u > a$ and $m > b$.

Because types are both not observable, the market forms beliefs on types conditional on observable $x$ and $y$. As before, let $\mu$ denote the market’s belief on the worker’s type conditional on her skill $y$, whereas we use $\rho$ for the market’s belief on the firm’s type conditional on its observable characteristic $x$. Suppose that the market’s beliefs $\mu$ and $\rho$ are increasing. Then, a separating equilibrium is a unique equilibrium. For a separating equilibrium, we can redefine the firm’s revenue function in terms of observable $x$ and $y$.

\[
S(x, y) := \pi(x, y, \mu(y), \rho(x)).
\] (43)

Because of the form of the firm’s revenue function $\pi$ and the monotone beliefs $\mu$ and $\rho$, $S(x, y)$ is supermodular, i.e., $S_{xy}(x, y) > 0$ for all $x, y$.

Given the market’s beliefs $\mu$ and $\rho$ on types, stable matching is characterized by $\{x, p\}$, where $x$ is the matching function in terms of observable characteristics so that $x(y)$ is the characteristic $x$ of the firm that a worker with $y$ works for and $p(y)$ is the pay for a worker with $y$. Importantly, a firm is the residual claimant in that it claims the remaining revenue after paying its worker. Therefore, for a firm with $x$, the stable matching partner $y$ can be found by solving

\[
\max_{y \geq 0} S(x, y) - p(y).
\] (44)

For a worker with $y$, the stable matching partner $x$ can be found by solving

\[
\max_{x \geq 0} S(x, y) - [S(x, \tilde{y}(x)) - p(\tilde{y}(x))],
\] (45)

where $\tilde{y}$ is the inverse of $x$. For a worker with $y$, $x(y)$ is the solution for her matching problem (45). For a firm with $x$, $\tilde{y}(x)$ is the solution for its matching problem (44).

Given the beliefs $\mu$ and $\rho$, the matching function $x$ is increasing and the worker pay function $p$ satisfies

\[
p'(y) = S_y(x(y), y)
\]
\[
= \pi_y(x(y), y, \mu(y), \rho(x(y))) + \pi_{\theta}(x(y), y, \mu(y), \rho(x(y))) \mu'(y)
\]
and $p(0) = 0$. The pay determination in (46) is comparable to one in (10).
Given stable matching \( \{x, p\} \) that would occur in the market, we can study pre-match investment problems. Firms and workers believe that workers will be compensated by the worker pay function \( p \) satisfying (46), whereas firms are residual claimants in their respective matches. This implies that the worker of type \( \theta \) makes her investment \( y(\theta) \) that solves

\[
\max_{y \geq 0} p(y) - c(y, \theta),
\]  

(47)

whereas the firm of type \( \delta \), as a residual claimant, makes its investment \( g(\delta) \) that solves

\[
\max_{x \geq 0} \pi(x, \tilde{y}(x), \mu(\tilde{y}(x)), \delta) - p(\tilde{y}(x)) - e(x, \delta),
\]  

(48)

where \( \tilde{y}(x) \) is the inverse of the market matching function \( x(y) \). Then, the worker’s equilibrium investment satisfies that for all \( \theta \),

\[
p'(y) - c_y(y, \theta) = \pi_x(x, \tilde{y}(x), \mu(\tilde{y}(x)), \delta) + \pi_\theta(x, \tilde{y}(x), y, \mu(y), \rho(x(y)))\mu'(y) - c_y(y, \theta) = 0
\]  

at \( y = y(\theta) \). This is the first-order condition for the worker’s equilibrium investment and the worker inefficiently overinvests because of the informational effect of her investment, \( \pi_\theta(x(y), y, \mu(y), \rho(x(y)))\mu'(y) \).

On the other hand, the first-order condition for the equilibrium investment made by the firm of type \( \delta \), denoted by \( g(\delta) \), satisfies that for all \( \delta \),

\[
\pi_x(x, \tilde{y}(x), \mu(\tilde{y}(x)), \delta) - e_x(x, \delta) + [\pi_y(x, \tilde{y}(x), \mu(\tilde{y}(x)), \delta) + \pi_\theta(x, \tilde{y}(x), \mu(\tilde{y}(x)), \delta)\mu'(\tilde{y}(x)) - p'(\tilde{y}(x))]\tilde{y}'(x) = 0
\]  

at \( x = g(\delta) \). Because of (46), this first-order condition is reduced to

\[
\pi_x(x, \tilde{y}(x), \mu(\tilde{y}(x)), \delta) - e_x(x, \delta) = 0 \text{ at } x = g(\delta).
\]  

(51)

(51) shows that the marginal return on the firm’s investment includes only the direct productivity effect of investment and hence the firm’s equilibrium investment is constrained efficient conditional on its worker’s investment and her type. This is because any potential adverse affect of the asymmetric information on the firm’s type is internalized by its investment decision as the residual claimant.

We continue to assume that the matching function \( \delta(\theta) \) in terms of types is a power function,

\[
\tilde{\delta}(\theta) = k\theta^\eta.
\]  

(52)

Given this power matching function, we can fully solve the SSME when both firms and workers make pre-match investment (See Appendix B for details). The
solution is provided in Proposition 10 in Appendix B and the investment functions take the form of $g(\delta) = D\delta^r$ and $y(\theta) = A\theta^2$ with

$$r = \frac{dm + vm - bv}{um - ub - am} + \frac{bn + cm}{q(um - ub - am)}$$

and $z = \frac{(ar + d)q + c + n}{m - b}$

and coefficients $D$ and $A$ are functions of parameters in the model.\(^\text{16}\)

The market matching function becomes $x(y) = Dk r A^{-\frac{m}{m-b}} y^{\frac{m}{m-b}}$ and the worker pay function becomes

$$p(y) = \frac{(ar + d) bq + bn + cm}{(ar + d)mq + bn + cm} S(g(y), y)$$

$$= \frac{(ar + d) bq + bn + cm}{(ar + d)mq + bn + cm} \alpha k^{ar + d} D^a y^{\frac{(ar + d)qm + nb + cm}{(ar + d)q + c + n}}.$$

The equilibrium worker pay conditional on the worker’s type is

$$\tilde{p}(\theta) = p(y(\theta)) = K \theta^{\frac{(ar + d)qm + nb + cm}{m - b}},$$

where coefficient $K$ is a function of parameters in the model. As shown earlier in (26) in Section 4.2, the inequality measure is the exponent of $\theta$ in $\tilde{p}(\theta)$, which is

$$T = \frac{(ar + d)qm + nb + cm}{m - b}.$$  \hspace{1cm} (53)

### 8.2 Two-sided investment versus one-sided investment

$T$ specified in (53) depends on all nine power parameters $(a, b, c, d, m, n, u, v, q)$ in the model and also on $r$, which is the exponent of $\delta$ in the firm’s investment function $g(\delta)$. Because $r$ is also a function of the power parameters in the model, comparative statics for inequality can be started by looking at the following partial total derivatives

$$\frac{\partial T}{\partial i} = T_r \times r_i + T_i$$

for $i = a, b, c, d, m, n, u, v, q.$  \hspace{1cm} (54)

(54) shows the two channels of changes in inequality in two-sided investment as the value of power parameter $i$ changes. A change in $i$ directly changes $T$ through the second term in (54) but it also changes the firm’s investment, which results in a change in $r$. This changes will lead to a change in $T$ through the first term in (54), which is absent in one-sided investment by workers.

\(^{16}\)To see the effect of the firm’s investment on the manager’s investment later in this section, it is better not to expand the term $r$ in $z$, where $r$ is the exponent in the firm’s investment function.
For comparative analysis, we can first derive the partial derivatives of \( r \):

\[
r_a > 0, r_b > 0, r_c > 0, r_d > 0, r_v > 0, r_u < 0, r_n > 0, r_m < 0, r_q < 0.
\]

(55)

All partial derivatives are intuitive. One interesting partial derivative is \( r_q < 0 \). As \( q \) increases, the distribution of firm type is relatively more spread to the right. For a worker, it is a good thing since she can match with a firm of a higher type. For a firm, it is a bad thing because a firm of type \( \delta \) now matches with a worker of a type lower than the type of the worker it would match with prior to an increase in \( q \). This dampens every firm’s incentive to invest, which is reflected in \( r_q < 0 \).

Partial derivatives of \( T \) are

\[
T_a > 0, T_b > 0, T_c > 0, T_d > 0, T_v > 0, T_u < 0, T_n > 0, T_m < 0, T_q > 0, T_r > 0.
\]

(56)

**Proposition 9** The comparative statics on inequality are as follows.

\[
\frac{\partial T}{\partial q} = T_r r_i + T_i
\]

(57)

\[\begin{array}{cccccccc}
  a & b & c & d & v & u & n & m & q \\
  + & + & + & + & + & - & + & - & +
\end{array}\]

Proof. (55) and (56) determine the signs of the partial total derivatives \( \frac{\partial T}{\partial q} \) for \( i = a, b, c, d, v, u, n, m \), whereas they do not for \( i = q \). This is because \( T_r r_q < 0 \) and \( T_q > 0 \). For the change in \( T \) with respect to \( q \), we directly take the derivative of

\[
T = \frac{(ar + d)qm + nb + cm}{m - b}
\]

(58)

\[
= \frac{\left( a \left( \frac{dm + cm - bw}{um - ub - am} + \frac{bn + cm}{qm - ub - am} \right) + d \right) qm + nb + cm}{m - b}
\]

\[
= \frac{\left( a \left( \frac{q(dn + cm - bw)}{um - ub - am} + \frac{bn + cm}{um - ub - am} \right) + dq \right) m + nb + cm}{m - b}
\]

It is easy to see that \( T \) is increasing in \( q \) from (58). This implies that \( T_q \) dominates \( T_r r_q \), which yields \( \frac{\partial T}{\partial q} > 0 \).

According to Proposition 9, an advance in production technology (\( a \uparrow, b \uparrow, c \uparrow, d \uparrow \)), a reduction of investment costs (\( v \uparrow, u \downarrow, n \uparrow, m \downarrow \)) on either side of the market, and an increase in the relative heterogeneity of firm type (\( q \uparrow \)), induced by changes in the values of power parameters, all increase income inequality.

Therefore, Proposition 9 extends the corresponding results in one-sided investment derived in Proposition 3. In fact, all the other results derived in one-sided investment are qualitatively preserved in two-sided investment. In particular, we
extend Propositions 5, 6, and 7 with different cutoff type values, and also Proposition 8 (See Propositions 11, 12, 13, and 16 in Appendix B).

Therefore, even in two-sided investment, in response to a change in each factor contributing to rising inequality, equilibrium percentage changes in worker skill investment, firm revenue, and worker pay monotonically increasing in worker type, switching from negative to positive at corresponding cutoff type values. A change in each factor contributing to rising inequality makes individuals of type below a corresponding cutoff type worse off but individuals of type above better off. However, only changes in firm-type related factors \((d \uparrow, q \uparrow)\) necessarily improve efficiency. However there are important differences as well. We explain them now.

In one-sided investment by workers, the firm’s investment does not change so that \(g(\delta) = D\delta^r\) is fixed. This implies that in response to a change in factor \(i\) contributing to rising income inequality, the partial derivative of \(r\) with respect to \(i\) is zero, \(r_i = 0\). Therefore, \(\frac{T_i}{T_i} = T_r r_i + T_i\) is the change in income inequality in two-sided investment. The difference in the change in income inequality due to two-sided investment is

\[
\frac{\delta T_i}{\delta r} - T_i = T_r r_i
\]

\(T_r r_i\) is positive for all \(i \neq q\), whereas it is negative for \(i = q\). Therefore, for all \(i \neq q\), a change in factor \(i\) contributing to rising income inequality increases income inequality more in two-sided investment than it does in one-sided investment. However, an increase in \(q\) increases income inequality less in two-sided investment than it does in one-sided investment. As mentioned above, an increase in \(q\) dampens the firm’s incentive to invest because it now matches with a worker of lower type. This is translated into \(r_q < 0\). If other things equal, this indeed also dampens the worker’s incentive to invest. Therefore, as \(q\) increase, the worker’s investment increases but not as much as it would in the one-sided investment. As shown later, this is reflected into the firm revenue, worker pay and income inequality: Because of \(T_r r_q < 0\), we have that \(\frac{\delta T_q}{\delta q} < T_q\), where \(T_q\) is a change in \(T\) in one-sided investment where investment by firms is exogenously given.

Proposition 14 in Appendix B provides the analysis on how equilibrium changes in worker skill investment, firm revenue, and worker pay differ in two-sided investment. Recall that given a general form of a power function, \(y = \varphi(\kappa)\theta^\rho(\kappa)\), the elasticity of \(y\) with respect to \(i\) is

\[
\epsilon_{yi}(\theta) = \frac{\partial y}{\partial i} \frac{i}{y} = \frac{i\varphi_i(\kappa)}{\varphi(\kappa)} + i\rho_i(\kappa) \ln \theta
\]

and the partial derivative of \(\epsilon_{yi}(\theta)\) with respect to \(\theta\) is

\[
\frac{\partial \epsilon_{yi}(\theta)}{\partial \theta} = \frac{i\rho_i(\kappa)}{\theta}.
\]

39
For the worker skill function $y(\theta) = A \theta^z$, we have that $\rho(\kappa) = z = \frac{(ar+d)q+c+n}{m-b}$.

Because $r$ is a function of power parameters, the partial total derivative of $z$ with respect to $i$ is $\frac{\partial z}{\partial r_i} = z_i r_i + z_i$. Let $\epsilon_{yi}^2(\theta)$ and $\epsilon_{yi}^1(\theta)$ denote the elasticities of skill investment with respect to $i$ in two-sided investment and one-sided investment respectively. Because $r_i = 0$ in one-sided investment, we have that

$$\frac{\partial \epsilon_{yi}^2(\theta)}{\partial \theta} - \frac{\partial \epsilon_{yi}^1(\theta)}{\partial \theta} = \frac{iz_r r_i}{\theta}$$

Because $z_r > 0$, the partial derivatives of $r$ given in (55) induce that for $i = a, b, c, d, n, v$,

$$\frac{\partial \epsilon_{yi}^2(\theta)}{\partial \theta} - \frac{\partial \epsilon_{yi}^1(\theta)}{\partial \theta} > 0 \quad (59)$$

but for $i = m, u, q$,

$$\frac{\partial \epsilon_{yi}^2(\theta)}{\partial \theta} - \frac{\partial \epsilon_{yi}^1(\theta)}{\partial \theta} < 0 \quad (60)$$

(59) and (60) show that in response to a change in any factor other than $q$ contributing to rising inequality, equilibrium skill increases in worker type at a higher rate in two-sided investment than it does in one-sided investment. However, in response to an increase in $q$, equilibrium skill increases in worker type at a lower rate in two-sided investment.

For the elasticity of worker pay with respect to $i$, we have that

$$\frac{\partial \epsilon_{pi}^2(\theta)}{\partial \theta} - \frac{\partial \epsilon_{pi}^1(\theta)}{\partial \theta} = \frac{i T_r r_i}{\theta}$$

and the same results are derived as ones derived in (59) and (60). Because worker pay is a fixed proportion of the firm’s revenue, the differences in partial derivatives of the elasticities of firm revenue with respect to $\theta$, $\frac{\partial \epsilon_{pi}^2(\theta)}{\partial \theta} - \frac{\partial \epsilon_{pi}^1(\theta)}{\partial \theta}$, also have the same results as well.

In summary, in response to a change in any factor other than $q$ contributing to rising inequality, equilibrium skill investment, firm revenue, and worker pay all increase in worker type at a higher rate in two-sided investment than it does in one-sided investment. However, in response to an increase in $q$, they increase in worker type at a lower rate in two-sided investment.

For efficiency analysis, our intuition suggests that we only need to look at the ratio of efficient skill investment to equilibrium skill investment because the firm’s investment is constrained efficient and the equilibrium matching pattern is efficient (assortative). Our intuition is confirmed in Proposition 16 in Appendix B in that efficiency analysis is exactly the same whether it is based on the ratio of efficient skill investment to equilibrium skill investment or the ratio of the firm’s efficient investment to its equilibrium investment: Only changes in firm-type related factors $(d \uparrow, q \uparrow)$ necessarily improve efficiency.
It is also interesting to see how efficiency changes differ in two-sided investment in a response to an increase in \( d \) or \( q \). The ratio of equilibrium skill investment to equilibrium skill investment in two-sided investment is \( E^y_2 = R^\sigma \) with

\[
R = \frac{(ar + d) bq + cb + nb}{(ar + d) bq + cm + nb} \quad \text{and} \quad \sigma = \frac{u - a}{mu - ma - bu},
\]

where subscript 2 in \( E^y_2 \) stands for two-sided investment.\(^{17}\) The elasticity of \( E^y_2 \) with respect to \( i = d \) or \( q \) is

\[
\frac{\partial E^y_2}{\partial i} i E^y_2 = \frac{\sigma}{\sigma R} \left[ R, r_i + R \right]
\]

In one-sided investment, \( r_i = 0 \) and hence the elasticity of \( E^y_1 \) with respect to \( i = d, q \) is

\[
\frac{\partial E^y_1}{\partial i} i E^y_1 = \frac{\sigma}{R} R_i
\]

Therefore, the difference is

\[
\frac{\partial E^y_2}{\partial i} i E^y_2 - \frac{\partial E^y_1}{\partial i} i E^y_1 = \frac{\sigma}{R} R_i r_i
\]

Because \( R_d = R_q, R_r > 0, r_d > 0, \) and \( r_q < 0 \), we can show that

\[
\frac{\partial E^y_2}{\partial q} q E^y_2 < \frac{\partial E^y_1}{\partial d} d E^y_1 \quad \frac{\partial E^y_2}{\partial q} q E^y_2 < \frac{\partial E^y_1}{\partial d} d E^y_1.
\]

Therefore, an increase in \( q \) increases efficiency at a lower rate in two-sided investment than it does in one-sided investment, whereas an increase in \( d \) increases efficiency at a higher rate in two-sided investment. Further, the impact of an increase in \( q \) on efficiency is lower than the impact of an increase in \( d \) in two-sided investment, whereas they are the same in one-sided investment.

\(^{17}\)According to Proposition 15 in Appendix B, the efficient skill investment takes the form of

\[
y_e(\theta) = A_e \theta^\varepsilon
\]

with

\[
A_e = \left( \frac{a \alpha}{w^\tau} \right) \frac{\alpha^{a+\frac{a}{(2+2a)}}}{\frac{a}{m-b-a-m} \left[ \frac{\alpha k^{ar+d} ((ar + d) bq + cm + nb)}{m^\beta ((ar + d) q + c + n)} \right]^{\frac{u-d}{m-b-a-m}}}
\]

On the other hand, according to Proposition 10, the equilibrium skill investment is \( y(\theta) = A \theta^\varepsilon \) with

\[
A = \left( \frac{a \alpha}{w^\tau} \right) \frac{\alpha^{a+\frac{a}{(2+2a)}}}{\frac{a}{m-b-a-m} \left[ \frac{\alpha k^{ar+d} ((ar + d) bq + cm + nb)}{m^\beta ((ar + d) q + c + n)} \right]^{\frac{u-d}{m-b-a-m}}}
\]

Therefore, the ratio of \( y_e(\theta) \) to \( y(\theta) \) at all \( \theta \) is

\[
E^y_2 = \frac{y_e(\theta)}{y(\theta)} = \left[ \frac{(ar + d) bq + cb + nb}{(ar + d) bq + cm + nb} \right]^{\frac{u-d}{m-b-a-m}}.
\]
9 Conclusion

While there is a growing number of the literature in the random growth model (Gabaix, Lasry, Lions, and Moll (2016), Jones and Kim (2017), Kim (2015), Luttmer (2015)) that study determinants and dynamics of top income inequality observed in the United States, less known are how all factors contributing to top income inequality systematically affects individual welfare, efficiency and the Lorenz curve and whether it is possible to address these questions in a model that can endogenously accommodate the fast growth rate of income observed in the U.S. data.

The equilibrium matching model with pre-match investment and market imperfections developed in this paper is fully solvable and extremely tractable so that it provides clear repercussions of rising top income inequality in the SSME that endogenously accommodates fast rising top income inequality observed in the United States, departing from Gibrat’s law. In response to a change in each factor contributing to rising inequality, the percentage changes in investment skill, revenue, and pay are all monotonically increasing in worker type, switching from negative to positive at respective cutoff types. Rising inequality has serious non-trivial repercussions on welfare and efficiency: A change in each factor contributing to rising inequality makes individuals of type below a certain value worse off but individuals of type above better off. However, only changes in firm-related factors necessarily improve efficiency.

Naturally, one can ask why the United States experiences faster rising top inequality compared to other developed countries such as France since early 1980s. One possible explanation is regulations. If regulations on wage or other forms of income are so severe that firm revenue is not freely transferable to worker through competitive bargaining, they effectively make utility non-transferable. This may make competition to match with a better partner less fierce and have implications on inequality. Taxes on wage and other forms of income may also affect income inequality as well. In order to see the comprehensive impacts of regulations or tax policies, it is important how those policies also affect individual’s investment decisions. Our analysis is based on comparative statics, so it does not directly keep track of the transition or dynamics of income inequality. It would be interesting to see the dynamics of income inequality based on our fully solvable matching model with skill acquisition. We leave these research questions for future research.

References


APPENDIX

Seungjin Han
May 10, 2018

A Proof of Proposition 1

We first present the proof of Proposition 1. Throughout the appendix, the equation numbers less than (61) are ones defined in the paper. Proposition 1 can be established based on general functional forms of the firm’s revenue function $f(x,t)$, talent function $t(y,\theta)$, and utility cost function of investment $c(y,\theta)$ with the following assumptions.

Function $t(y,\theta)$ is concave in $y$ and non-decreasing given any $\theta > 0$, and twice continuously differentiable. It is increasing in $\theta$ given any $y > 0$ with $t(y,0) = 0$. If $t(y,\theta) = t(y',\theta)$ for all $(y,y',\theta)$, then $y$ is for pure signaling of type without enhancing talent. For now, we will assume that $t$ is increasing in $y$ given any $\theta > 0$ so that investment in $y$ not only increases the worker’s talent but also signals her type.

The cost function $c(y,\theta)$ is twice continuously differentiable, convex in $y$ given any $\theta > 0$, and increasing in $y$ given any $\theta > 0$, but decreasing in $\theta$ given any $y > 0$. We also assume that $\lim_{\theta \to 0} c_{y}(y,\theta) = \infty$ for all $y > 0$.

The revenue function $f(x,t)$ is increasing in $x$ given any $t > 0$ and in $t$ given any $x > 0$. We assume that $f(x,t) \geq 0$ for all $(x,t)$, it is concave in $t$ and twice continuously differentiable.

We impose the submodular property on the cost function $c$ (i.e., the single crossing property of $c$) and the supermodular assumption on the revenue function $f$ (i.e., complementarity between firm size and talent in $f$):

Assumption A1. $c_{y\theta}(y,\theta) < 0$ for all $(y,\theta)$.

Assumption A2. $f_{xt}(x,t) > 0$ for all $(x,t)$.

It is worth pointing out that it seems reasonable to assume the convexity of $c(y,\theta)$ in $y$. However, the concavity of $f(x,t)$ in $t$ and the concavity of $t(y,\theta)$ in $y$ are not needed if worker type is unbounded so that the set of equilibrium skill levels acquired by all workers can span the entire $\mathbb{R}_+$. As shown later, the concavity assumption kicks in only when we show that it is not profitable for a worker to deviate to acquire skill beyond the highest level observed in equilibrium of the type set $\Theta$ is bounded.
A.1 Proof of Conditions 1, 2, and 4

We first prove the property of the worker skill function \( y(\theta) \) stated in Condition 1 of Proposition 1. It is clear that \( y(0) = 0 \) because we assume that \( \lim_{\theta \to 0} c_y(y, \theta) = \infty \) for all \( y > 0 \).

Consider two workers, one of type \( \theta \) and the other of type \( \theta' \) with \( \theta > \theta' \). Since each worker can earn the other worker’s pay by acquiring the same level of skill, we must have

\[
\begin{align*}
\mathcal{P}(y(\theta)) - c_y(y(\theta), \theta) &\geq \mathcal{P}(y(\theta')) - c_y(y(\theta'), \theta') \quad \text{in equilibrium.}
\end{align*}
\]

Assumption 1 shows that the market belief \( \mu \) uniquely specifies the worker type for any skill level observed in the market since \( \mu \) is an increasing function. This implies that \( y(\theta) \neq y(\theta') \). Because of Assumption 1, (61) and \( y(\theta) \neq y(\theta') \) imply that \( y(\theta) > y(\theta') \) if \( \theta > \theta' \). Therefore, \( y \) is increasing.

Consider the skill accumulation decision by a worker of type \( \theta \). Since the market pay function is given by \( \mathcal{P} \), it must satisfy the first-order condition in (9) in the range of \( Y \):

\[
\mathcal{P}'(y(\theta)) - c_y(y(\theta), \theta) = 0.
\]

Since the skill accumulation decision by a worker of type \( \theta' \) also satisfies the first-order condition, we have

\[
\mathcal{P}'(y(\theta')) - c_y(y(\theta'), \theta') = 0. \tag{62}
\]

Suppose that \( \theta' < \theta \). By Assumption A1, (62) implies

\[
\mathcal{P}'(y(\theta')) - c_y(y(\theta'), \theta) > 0. \tag{63}
\]

Suppose that \( \theta' > \theta \). By Assumption A1, (62) implies

\[
\mathcal{P}'(y(\theta')) - c_y(y(\theta'), \theta) < 0. \tag{64}
\]

Equations (63) and (64) show that no worker has an incentive to change her skill level to those that other workers acquire if and only if (9) is satisfied for all \( \theta \).

What if a worker deviates to acquire \( y > \bar{y} = \max Y' \)? Let \( \bar{\theta} \) be the supremum of \( \Theta \). The market pay function defined for \( y > \bar{y} \) should provide no incentive for such a deviation. We consider the market pay function that satisfies Condition 4 for \( y > \bar{y} \). Note that the market pay function defined in Condition 4 incorporates
the firm’s belief \( \mu(y) = \tilde{\theta} \) on the worker’s type, conditional on \( y > \bar{y} \) and it is based on the iso-profit curve for the firm with the biggest size associated with hiring a worker of the highest type, when her skill level is \( y > \bar{y} \). First, consider a deviation by a worker of type \( \tilde{\theta} \). Applying (11) yields the right-hand limit as follows:

\[
p'_{\tilde{\theta}}(\bar{y}) = f_t(x, t(\bar{y}, \tilde{\theta}))t_t(\bar{y}, \tilde{\theta}) - c_{\tilde{\theta}}(\bar{y}, \tilde{\theta}).
\]

This measures the rate of change in the worker’s utility when she marginally increases her skill level above \( \bar{y} \). Combining (9) and (10), we can induce

\[
f_t(x, t(\bar{y}, \tilde{\theta}))t_t(\bar{y}, \tilde{\theta}) + f_t(x, t(\bar{y}, \tilde{\theta}))t_\theta(\bar{y}, \tilde{\theta})\mu'(\bar{y}) - c_{\tilde{\theta}}(\bar{y}, \tilde{\theta}) = 0.
\]

Equation (65) implies that

\[
p'_{\tilde{\theta}}(\bar{y}) - c_{\tilde{\theta}}(\bar{y}, \tilde{\theta}) = f_t(x, t(\bar{y}, \tilde{\theta}))t_t(\bar{y}, \tilde{\theta}) - c_{\tilde{\theta}}(\bar{y}, \tilde{\theta}) < 0. \tag{66}
\]

Therefore, the utility for a worker of type \( \tilde{\theta} \) indeed decreases as she marginally increase her skill level above \( \bar{y} \). Because \( f \) is concave in \( t \), \( t \) is concave in \( y \), and \( c \) is convex in \( y \), (66) implies that for all \( y > \bar{y} \), we have

\[
p'(y) - c_{y}(y, \tilde{\theta}) = f_t(\bar{x}, t(y, \tilde{\theta}))t_t(y, \tilde{\theta}) - c_{\tilde{\theta}}(y, \tilde{\theta}) < 0.
\]

Therefore, the utility for a worker of type \( \tilde{\theta} \) monotonically decreases as she increases her skill level beyond \( \bar{y} \). Therefore, we can conclude that, for all \( y > \bar{y} \)

\[
p(y) - c(y, \tilde{\theta}) < p(\bar{y}) - c(\bar{y}, \tilde{\theta}). \tag{67}
\]

Due to Assumption 1, (67) induces that for all \( \theta < \tilde{\theta} \) and all \( y > \bar{y} \)

\[
p(y) - c(y, \theta) < p(\bar{y}) - c(\bar{y}, \theta). \tag{68}
\]

Because no worker wants to change her skill level to any other worker’s skill level, we have that for all \( \theta < \tilde{\theta} \),

\[
p(\bar{y}) - c(\bar{y}, \theta) < p(y(\theta)) - c(y(\theta), \theta). \tag{69}
\]

Equations (68) and (69) yield that for all \( \theta < \tilde{\theta} \) and all \( y > \bar{y} \)

\[
p(y) - c(y, \theta) < p(y(\theta)) - c(y(\theta), \theta).
\]

Thus, by acquiring \( y > \bar{y} \) the utility levels for all other workers are also lower than their equilibrium payoffs. Therefore, no worker has incentives to acquire \( y > \bar{y} \).

Condition 2 is about the firm’s belief function \( \mu(y) \). Condition 1 shows that \( y \) is increasing. Given increasing \( y \), a worker’s skill fully reveals her type and hence we have that, for all \( \theta \), \( \mu(y(\theta)) = \theta \). Given Condition 1, the firm’s belief function is increasing up until \( \bar{y} \) and \( \mu(\bar{y}) \) = \( \tilde{\theta} \). Because \( \mu(\bar{y}) = \tilde{\theta} \), \( \mu(y) \) for all \( y > \bar{y} \) cannot be lower than \( \tilde{\theta} \) given Assumption 1. Since \( \tilde{\theta} \) is the highest type level, it implies that for all \( y > \bar{y} \), \( \mu(y) = \tilde{\theta} \).
A.2 Proof of Condition 3

Consider the property of the market matching function $x(y)$ in Condition 3. For condition 3.(a), consider two workers, one with skill level $y$ and the other with $y'$. Suppose that the worker with $y$ works for firm $x$ and the worker with $y'$ works for firm $x'$ in equilibrium such that $x > x'$. If firm $x'$ were to hire the executive with $y$, the maximum pay that it is willing to offer is $f(x', t(y, \mu(y))) = [f(x', t(y', \mu(y'))) - p(y')]$ because it will not hire the worker with $y$ if the profit associated with hiring her is less than the profit associated with hiring the worker with $y'$ at the market pay $p(y')$. Therefore, if the worker with $y$ wants to work for firm $x$ at the market pay $p(y)$ instead of working for firm $x'$, then the following condition is satisfied in equilibrium:

$$p(y) \geq f(x', t(y, \mu(y))) - [f(x', t(y', \mu(y'))) - p(y')].$$  \hspace{1cm} (70)

Similarly, if the worker with $y'$ wants to work for firm $x'$ at the market pay $p(y')$ instead of working for firm $x$, the following condition is satisfied in equilibrium as well:

$$p(y') \geq f(x, t(y', \mu(y'))) - f(x, t(y, \mu(y))) + p(y).$$  \hspace{1cm} (71)

Summing up (70) and (71) yields

$$f(x', t(y', \mu(y'))) + f(x, t(y, \mu(y))) \geq f(x', t(y, \mu(y))) + f(x, t(y', \mu(y'))).$$  \hspace{1cm} (72)

Since the firm’s belief function $\mu(y)$ is an increasing function, $t(y, \mu(y)) \geq t(y', \mu(y'))$ if $y \geq y'$. Because $x > x'$ and $f$ satisfies Assumption A2, (72) implies that $y > y'$. Therefore, the worker with a higher skill level matches with a firm with a bigger size.

Consider two firms, one hiring a worker with $y$ and the other hiring a worker with $y'$ such that $y > y'$, that is, firms $x(y)$ and $x(y')$. Since each firm can hire a worker by paying a pay that is at least as high as the market pay, we must have

$$f(x(y), t(y, \mu(y))) - p(y) \geq f(x(y), t(y', \mu(y'))) - p(y'),$$

$$f(x(y'), t(y', \mu(y'))) - p(y') \geq f(x(y'), t(y, \mu(y))) - p(y).$$  \hspace{1cm} (73) \hspace{1cm} (74)

in equilibrium. Summing up (73) and (74) yields

$$f(x(y), t(y, \mu(y))) + f(x(y'), t(y', \mu(y'))) \geq f(x(y), t(y', \mu(y'))) + f(x(y), t(y, \mu(y)))$$  \hspace{1cm} (75)

Since $y > y'$, we have $\mu(y) > \mu(y')$ and hence $t(y, \mu(y)) > t(y', \mu(y'))$. Because $f$ satisfies Assumption A2, (75) implies that $x(y) > x(y')$. Therefore, a firm with a bigger size matches with a worker with a higher skill level. From both sides
of the market, we can then conclude that the market matching function must be positively assortative, which is stated in Condition 3.(a)

For Condition 3.(b), consider firm \( x(y) \) who hires a worker with \( y = y(\theta) \) for each \( y \in Y \). It is necessary that the first-order condition (10) holds for the firm’s hiring problem (6):

\[
f_x(x(y), t(y, \mu(y))) (t_y(y, \mu(y)) + t_\theta(y, \mu(y))\mu'(y)) - p'(y) = 0.
\]

Now consider a worker with some skill level \( y \) in \( Y \). The first-order condition for the solution to the worker’s matching problem (7) is

\[
f_x(x, t(y, \mu(y))) - f_x(x, t(\tilde{y}(x), \mu(\tilde{y}(x)))) [t_y(\tilde{y}(x), \mu(\tilde{y}(x))) + t_\theta(\tilde{y}(x), \mu(\tilde{y}(x)))\mu'(\tilde{y}(x))] \tilde{y}'(x)
- f_x(x, t(\tilde{y}(x), \mu(\tilde{y}(x)))) + p'(\tilde{y}(x))\tilde{y}'(x) = 0. \tag{76}
\]

at \( \tilde{y}(x) = y \). Because

\[
f_x(x, t(y, \mu(y))) = f_x(x, t(\tilde{y}(x), \mu(\tilde{y}(x))))
\]

at \( \tilde{y}(x) = y \) and \( \tilde{y} \) is an increasing function, (76) is equivalent to (10) at \( \tilde{y}(x) = y \). This shows that it is also necessary that the first-order condition (10) holds for the worker’s matching problem (7). Finally, we generally have \( 0 \leq p(0) \leq f(0, 0) \) in an equilibrium because \( f(0, 0) \) is the revenue created at the bottom match between the smallest firm and the least talented executive with \( t(0, 0) = 0 \). Given our specific functional form for \( f \), \( f(0, 0) = 0 \), which implies that \( p(0) = 0 \). Therefore, Condition 3 (i.e., 3.(a) and 3.(b)) is the necessary condition for \( \{x, p\} \) to be the pair of the market matching function and the market pay function.

Let us prove that Condition 3 is the sufficient condition as well. Suppose that the worker with \( y = \tilde{y}(x) \) wants to works for firm \( x' \). Then the first-order condition for the worker with \( y \) at \( x' \) becomes

\[
f_x(x', t(y, \mu(y))) - f_x(x', t(\tilde{y}(x'), \mu(\tilde{y}(x')))) [t_y(\tilde{y}(x'), \mu(\tilde{y}(x'))) + t_\theta(\tilde{y}(x'), \mu(\tilde{y}(x')))\mu'(\tilde{y}(x'))] \tilde{y}'(x')
- f_x(x', t(\tilde{y}(x'), \mu(\tilde{y}(x'))) + p'(\tilde{y}(x'))\tilde{y}'(x'). \tag{77}
\]

Since (10) must be satisfied for the worker with \( y' = \tilde{y}(x') \), (77) is equivalent to

\[
f_x(x', t(y, \mu(y))) - f_x(x', t(\tilde{y}(x'), \mu(\tilde{y}(x'))) = 0. \tag{78}
\]

If \( x' > x \), \( \tilde{y}(x') > y = \tilde{y}(x) \) and \( \mu(\tilde{y}(x')) > \mu(y) \) because \( \tilde{y} \) is an increasing function given Condition 3.(a) and \( \mu \) is a increasing function. Therefore, if \( x' > x \), then \( t(\tilde{y}(x'), \mu(\tilde{y}(x'))) > t(y, \mu(y)) \). This implies that (78) is negative by Assumption
1. If \( x' < x \), \( \tilde{y}(x') < y = \tilde{y}(x) \) and \( \mu(\tilde{y}(x')) < \mu(y) \). Therefore, if \( x' < x \), then \( t(\tilde{y}(x'), \mu(\tilde{y}(x')) < t(y, \mu(y)) \) (78) is positive by Assumption 1. Therefore, (10) in Condition 3.(b), together with Condition 3.(a) (i.e., positively assortative matching), is the necessary and sufficient condition for a worker with \( y \) to work for firm \( x(y) \) at pay \( p(y) \).

Finally, consider the firm’s decision. Since (10) must hold for any firm \( x(y') \) who hires a worker with \( y' = y(\theta') \).

\[
f_t(x(y'), t(y', \mu(y'))) [t_y(y', \mu(y')) + t_\theta(y', \mu(y'))] - p'(y') = 0. \tag{79}
\]

Suppose that firm \( x(y) \) considers hiring a worker with \( y' = y(\theta') \). Then, the first-order condition for firm \( x(y) \) at \( y' \) is

\[
f_t(x(y), t(y', \mu(y'))) [t_y(y', \mu(y')) + t_\theta(y', \mu(y'))] = p'(y'). \tag{80}
\]

If \( y' \gtrless y \), \( x(y') \gtrless x(y) \) because of Condition 3.(a). Then, the supermodular assumption on \( f \) implies that if \( y' \gtrless y \), then

\[
f_t(x(y'), t(y', \mu(y'))) \gtrless f_t(x(y), t(y', \mu(y'))). \tag{81}
\]

Because \( \mu'(y') > 0 \) and \( t_y \) and \( t_\theta \) are also positive, (81) and (79) imply that if \( y' > y \), then (80) is negative. On the other hand, if \( y' < y \), then (80) is positive. It means that firm \( x(y) \) does not want to hire a worker other than a worker with \( y \). Therefore, (10) in Condition 3.(b), together with Condition 3.(a), is also the necessary and sufficient condition for firm \( x(y) \) to hire a worker with \( y \) at pay \( p(y) \).

B Two-sided Investment

We can show that both equilibrium investment functions \( g(\delta) \) and \( y(\theta) \) are monotonically increasing, given the firm’s revenue function, and investment cost functions. Because they are increasing and the redefined revenue function \( S(x, y) \) shows complementarity between characteristics \( x \) and \( y \), stable matching pattern is assortative, i.e., \( x(y) \) is increasing as well.

For equilibrium derivation, we define the type of a firm as a function of the type of the worker with whom it matches:

\[
\hat{\delta}(\theta) := \rho(x(y(\theta))) \tag{82}
\]

for all \( \theta \). Since every worker fully reveals her type by her skill investment in an SSME (i.e., \( \mu(y(\theta)) = \theta \)), we have \( \mu'(y(\theta)) = 1/y'(\theta) \).

Given \( \hat{\delta}(\theta) \) and \( \mu'(y(\theta)) = 1/y'(\theta) \), (49) becomes

\[
\pi_y(g(\hat{\delta}(\theta)), y, \theta, \hat{\delta}(\theta)) + \pi_\theta(g(\hat{\delta}(\theta)), y, \theta, \hat{\delta}(\theta))/y' - c_y(y, \theta) = 0. \tag{83}
\]
This yields a first-order differential equation \( y' = \phi(y, \theta) \), where

\[
\phi(y, \theta) := \frac{-\pi_\theta(g(\hat{\delta}(\theta)), y, \theta, \hat{\delta}(\theta))}{\pi_g(g(\hat{\delta}(\theta)), y, \theta, \hat{\delta}(\theta))} - c_y(y, \theta)
\]

with the initial condition \( y(0) = 0 \).

Note that (84) is derived given the firm’s investment function \( g(\delta) \), which is also endogenously determined. Let \( \tilde{\theta} \) denote the inverse of \( \hat{\delta} \). Then, (51) is rewritten as

\[
\pi_x(x, y(\tilde{\theta}(\delta)), \tilde{\theta}(\delta), \delta) - e_x(x, \delta) = 0
\]

Therefore, we need to simultaneously solve (84) and (85) together for \( y(\theta) \) and \( g(\delta) \). Once we derive both investment functions, we can derive belief functions \( \mu \) and \( \rho \) by inverting them. We can then derive the worker pay function \( p(\theta) \) by integrating the right-hand-side of (46):

\[
p(y) = \int_0^y \{ \pi_y(x(z), z, \mu(z), \rho(x(z))) + \pi_\theta(x(z), z, \mu(z), \rho(x(z)))\mu'(z) \} dz.
\]

with \( p(0) = 0 \).

**B.1 Equilibrium derivation**

It is most important to derive the investment functions \( g(\delta) \) and \( y(\theta) \) because belief, matching and worker pay functions can be derived from the investment functions. We first guess the forms of the investment functions as

\[
g(\delta) = D\delta^r \text{ for } \delta > 0 \text{ and } y(\theta) = A\theta^z \text{ for } \theta > 0
\]

Note that \( \lim_{\theta \to 0} c_y(y, \theta) = \infty \) and \( \lim_{\delta \to 0} e_x(x, \delta) = \infty \) so that we must have \( \lim_{\theta \to 0} y(\theta) = \lim_{\delta \to 0} g(\delta) = 0 \). The functional forms in (87) satisfy these properties. Now we present the values of \( D, r, A, \) and \( z \) by using equilibrium conditions.

**Proposition 10** The SSME is unique and characterized as follows.

1. \( g \) and \( y \) take the forms of (87) and \( \rho \) and \( \mu \) are their inverses respectively
with
\[
\begin{align*}
  r &= \frac{dm + bm - bv}{um - ub - am} + \frac{bn + cm}{q(um - ub - am)}, \\
  z &= \frac{(ar + d)q + c + n}{m - b}, \\
  D &= \left(\frac{a\alpha}{u^\tau} \right) \frac{m-b}{um - ub - am} \left[ \alpha k^{ar+d} \left( ((ar + d)bq + cm + nb) \right) \right] \frac{m-b}{m^2}, \\
  A &= \left[ \frac{\alpha k^{ar+d} D^a ((ar + d)bq + cm + nb)}{m^2} \right] \frac{1}{m-b} \frac{m-b}{m^2} \\
  &= \left(\frac{a\alpha}{u^\tau} \right) \frac{m-b}{um - ub - am} \left[ \alpha k^{ar+d} \left( ((ar + d)bq + cm + nb) \right) \right] \frac{m-b}{m^2},
\end{align*}
\]

2. for all \( y \geq 0 \), \( x(y) = DK^r A^{-\frac{m}{\tau^2}} y^{\frac{m}{\tau^2}} \) and
\[
\begin{align*}
  p(y) &= \frac{(ar + d)bq + bn + cm}{(ar + d)mq + bn + cm} S(x(y), y) \\
  &= \frac{(ar + d)bq + bn + cm}{(ar + d)mq + bn + cm} \alpha k^{ar+d} D^a y^{\frac{(ar + d)qm + nb + cm}{m^2}}.
\end{align*}
\]

**Proof.** Given \( g(\delta) = D\delta^r \), we derive \( A \) and \( z \) that satisfy (84) (equivalently (83)). Then, \( A \) and \( z \) are expressed in terms of parameters, including \( D \) and \( r \). We apply these solutions for \( A \) and \( z \) to \( y(\theta) \) and then use (85) to derive the solutions for \( D \) and \( r \), which are expressed in terms of parameters in the model primitives. Given (52) and solutions for \( y(\theta) \) and \( g(\delta) \), we can derive \( x(y) \). Finally, \( p(y) \) is derived by (86) with \( p(0) = 0 \). Note that \( r \) is positive by the assumption of \((u - a)(m - b) > ab\) with \( u > a \) and \( m > b \). This in turn makes \( z \) positive.

The measure of the worker’s bargaining power is
\[
B := \frac{(ar + d)bq + bn + cm}{(ar + d)mq + bn + cm},
\]
which is the proportion of the match surplus that goes to the worker in the form of pay. Given \( y(\theta) \), one can express the worker pay in terms of the worker’s type,
\[
\tilde{p}(\theta) = p(y(\theta)) = K\theta^{\frac{(ar + d)qm + nb + cm}{m^2}},
\]
with a constant \( K > 0 \). Following the earlier approach in Section 4, the measure of income inequality becomes
\[
T = \frac{(ar + d)qm + nb + cm}{m - b}.
\]
Note that $T$ measures the fractal top income inequality when worker types at high levels are well approximated by a Pareto distribution and it also represents the Lorenz income inequality even without a Pareto assumption on the whole distribution.

The matching pattern in the SSME is efficient in that matching is positively assortative in terms of types and also characteristics $x$ and $y$. However, investment levels are not within each match. Because part of the worker pay includes the informational effect (i.e., the second term on the right hand side of the second line in (86)), the worker’s investment level is inefficiently high, whereas the firm’s investment level is constrained efficient.

### B.2 Disproportional changes and welfare

Details on the analysis of income inequality is provided in Section 8.2. Now, let us examine the impacts of a change in each factor contributing to rising inequality on equilibrium percentage changes in worker skill investment, firm revenue, worker pay, and utility levels of workers. We first express worker skill investment, firm revenue, and worker pay and her utility in terms of worker type in the SSME: For all $\theta \in \Theta$,

\[
y(\theta) = A\theta^{\frac{(ar+d)q+c+n}{m-b}}, \tag{88}
\]

\[
\tilde{\pi}(\theta) = \alpha k^{\frac{(ar+d)q+c+n}{m-b}} A^{\frac{(ar+d)q+c+n}{m-b}} \theta^{\frac{(ar+d)q+c+n}{m-b}}, \tag{89}
\]

\[
\bar{p}(\theta) = \frac{(ar+d)q+c+n}{m-b} \theta^{\frac{(ar+d)q+c+n}{m-b}}, \tag{90}
\]

\[
U(\theta) = \left( (ar+d)q+c+n \theta^{\frac{(ar+d)q+c+n}{m-b}} - \beta A^m \right)^\theta^{\frac{(ar+d)q+c+n}{m-b}}. \tag{91}
\]

**Proposition 11** For the worker skill function $y(\theta)$, there exists a cutoff type $\bar{\theta}_{i}^\circ$ for each power parameter $i = b, c, d, q, n, v$ such that

1. for $i = b, c, d, q, n, v$, $\epsilon_{yi}(\theta)$ is negative (positive) if $\theta < \bar{\theta}_{i}^\circ$ ($\theta > \bar{\theta}_{i}^\circ$),

2. for $i = m, u$ $\epsilon_{yi}(\theta)$ is positive (negative) if $\theta < \bar{\theta}_{i}^\circ$ ($\theta > \bar{\theta}_{i}^\circ$), and

3. $\epsilon_{yi}(\theta)$ is increasing in $\theta$ for $i = b, c, d, q, n, v$ but decreasing in $\theta$ for $i = m, u$.

For the firm’s revenue function $\tilde{\pi}(\theta)$, there exists a cutoff type $\bar{\theta}_{i}^\triangledown$ for each power parameter $i = b, c, d, q, n, v$ such that

4. for $i = b, c, d, q, n, v$, $\epsilon_{\pi i}(\theta)$ is negative (positive) if $\theta < \bar{\theta}_{i}^\triangledown$ ($\theta > \bar{\theta}_{i}^\triangledown$),

5. for $i = m, u$, $\epsilon_{\pi i}(\theta)$ positive (negative) if $\theta < \bar{\theta}_{i}^\triangledown$ ($\theta > \bar{\theta}_{i}^\triangledown$), and
6. $\epsilon_{pi}(\theta)$ is increasing in $\theta$ for $i = b, c, d, q, n, v$ but decreasing in $\theta$ for $i = m, u$

**Proof.** The exponent of $\theta$ in the worker skill function $y(\theta)$ is $z = \frac{(ar + d)q + c + n}{m - b}$. Because $r$ is a function of power parameters in the model, we consider the partial total derivative of $z$ with respect to $i = a, b, c, d, m, n, u, v, q$:

$$\frac{\partial z}{\partial i} = z_r r_i + z_i.$$

The signs of partial derivatives of $r$ are provided in (55), whereas the signs of partial derivatives of $z$ are as follows:

$$z_r > 0, z_a > 0, z_b > 0, z_c > 0, z_d > 0, z_v = 0, z_a = 0, z_n > 0, z_m < 0, z_q > 0. \quad (92)$$

Combining (55) and (92) determines the signs of the partial total derivatives of $z$ for all $i = a, b, c, d, m, n, u, v$ except for $i = q$. The sign of $\frac{\partial z}{\partial q}$ is not clear because $z_r r_q < 0$ and $z_q > 0$. To determine the sign, we can express $z$ as follows.

$$z = \frac{(ar + d)q + c + n}{m - b} = \frac{(a \left( \frac{dm + vm - bw}{um - ub - am} + \frac{bn + cm}{q(um - ub - am)} \right) + d)q + c + n}{m - b} = \frac{aq(dm + vm - bw) + a(bn + cm)}{am - ub - am} + dq + c + n$$

It is easy to see that the derivative of $z$ with respect to $q$ is positive, which means that $z_q$ dominates $z_r r_q$. Therefore, we have the following signs of the partial total derivatives:

<table>
<thead>
<tr>
<th>$\frac{\partial z}{\partial q}$</th>
<th>$a$</th>
<th>$b$</th>
<th>$c$</th>
<th>$d$</th>
<th>$v$</th>
<th>$u$</th>
<th>$n$</th>
<th>$m$</th>
<th>$q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_r r_i + z_i$</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

(93)

Given (93), we can show that items 1 and 2 hold, applying (30) and (31). Given (93), we can also show that item 3 holds, applying (33) and (34). The proofs of items 4, 5 and 6 are analogous to those of items 1, 2, and 3 given (57).

We also have the qualitatively same results for worker pay.

**Proposition 12** For the worker pay function $\tilde{p}(\theta)$, there exists a cutoff type $\tilde{\theta}_i^*$ for each power parameter $i = b, c, d, q, m, n$ such that

1. for $i = b, c, d, q, n, v$, $\epsilon_{pi}(\theta)$ is negative (positive) if $\theta < \tilde{\theta}_i^*$ ($\theta > \tilde{\theta}_i^*$),

2. for $i = m, u$, $\epsilon_{pi}(\theta)$ is positive (negative) if $\theta < \tilde{\theta}_i^*$ ($\theta > \tilde{\theta}_i^*$), and
3. $\epsilon_{pi}(\theta)$ is increasing in $\theta$ for $i = b, c, d, q, n, v$ but decreasing in $\theta$ for $i = m, u$

**Proof.** The proof is similar to the proof of Proposition 11 given (57).

Proposition 9 shows that an advance in production technology ($a \uparrow, b \uparrow, c \uparrow, d \uparrow$), a reduction of investment costs ($v \uparrow, u \downarrow, n \uparrow, m \downarrow$) on either side of the market, and an increase in the relative heterogeneity of firm type ($q \uparrow$), induced by changes in the values of power parameters, all increase the income inequality, $T$. Given this result, Propositions 11 and 12 show that in a response to a change in each factor contributing to rising inequality, equilibrium percentage changes in worker skill investment, firm revenue, and worker pay are monotonically increasing in worker type, switching from negative to positive at the corresponding cutoff values of worker type.

Finally, we can analyze individual welfare changes.

**Proposition 13** For the equilibrium utility function $U(\theta)$, there exists a cutoff type $\tilde{\theta}_i$ for each power parameter $i = b, c, d, q, m, n$ such that

1. for $i = b, c, d, q, n, v$, $\frac{\partial U(\theta)}{\partial i}$ is negative (positive) if $\theta < \tilde{\theta}_i$ ($\theta > \tilde{\theta}_i$) and
2. for $i = m, u$, $\frac{\partial U(\theta)}{\partial i}$ is positive (negative) if $\theta < \tilde{\theta}_i$ ($\theta > \tilde{\theta}_i$).

**Proof.** We can use (30) and (31) for the proof given (57).

Proposition 13 shows that a change in each factor contributing to rising inequality makes workers of type below the corresponding cutoff value worse off but workers of type above better off. The impacts of a change in each factor contributing to rising inequality on equilibrium changes in worker skill investment, firm revenue, worker pay, and her utility in two sided investment are qualitatively the same as those in one-sided investment in that Propositions 11 and 12 extends Propositions 5 and 6 with different cutoff type values. Welfare analysis in two-sided investment is also qualitatively the same as that in one-sided investment since Proposition 13 extends Proposition 7 with different cutoff type values.

However, there are important differences in two-sided investment in terms of equilibrium percentage changes in worker skill investment, firm revenue, and worker pay. To see this, recall (32). It shows that given a power function, $y = \varphi(\kappa)\theta^{\rho(\kappa)}$, the elasticity of $y$ with respect to $\theta$ is

$$\epsilon_{yi}(\theta) = \frac{\partial y}{\partial i} \frac{i}{y} = \frac{i \varphi(\kappa)}{\varphi(\kappa)} + i \rho(\kappa) \ln \theta$$

and the partial derivative of $\epsilon_{yi}(\theta)$ with respect to $\theta$ is

$$\frac{\partial \epsilon_{yi}(\theta)}{\partial \theta} = \frac{i \rho(\kappa)}{\theta}$$

(94)
Consider the elasticities of worker skill $y$ with respect to $i = a, b, c, d, v, u, n, m, q$. $\epsilon_y^2(\theta)$ denotes the elasticity in two-sided investment and $\epsilon_y^1(\theta)$ in one-sided investment. For the elasticities of firm revenue and worker pay, $\epsilon_{\pi i}(\theta)$ and $\epsilon_{p i}(\theta)$, we also use the superscripts of 1 and 2 to differentiate one-sided investment and two-sided investment. First, we can establish the following results on elasticities.

**Proposition 14** Elasticiies of worker skill, firm revenue, and worker pay satisfy the following properties,

1. for $i = a, b, c, d, n, v, \theta$,
   \[
   \frac{\partial \epsilon_y^2(\theta)}{\partial \theta} > \frac{\partial \epsilon_y^1(\theta)}{\partial \theta} \quad \text{for all } L = y, \pi, p.
   \]

2. and, for $i = m, u, q, \theta$,
   \[
   \frac{\partial \epsilon_y^2(\theta)}{\partial \theta} < \frac{\partial \epsilon_y^1(\theta)}{\partial \theta} \quad \text{for all } L = y, \pi, p.
   \]

**Proof.** For the elasticities of worker skill, first note that $y(\theta) = A\theta^z$ with $z = \frac{ar + d}{m - b} - c + n$. According to (94), the partial derivative of $\epsilon_y^2(\theta)$ with respect to $\theta$ is
   \[
   \frac{\partial \epsilon_y^2(\theta)}{\partial \theta} = \frac{i z_i}{\theta} = \frac{i (z r_i + z_i)}{\theta} \quad \text{for all } i = a, b, c, d, v, n, m, u, q.
   \]

If firms do not invest, then $r_i = 0$ for all $i$. Therefore, the partial derivative of $\epsilon_y^1(\theta)$ with respect to $\theta$ is
   \[
   \frac{\partial \epsilon_y^1(\theta)}{\partial \theta} = \frac{i z_i}{\theta} \quad \text{for all } i = a, b, c, d, n, v, m, u, q.
   \]

and hence
   \[
   \frac{\partial \epsilon_y^2(\theta)}{\partial \theta} - \frac{\partial \epsilon_y^1(\theta)}{\partial \theta} = \frac{i z r_i}{\theta} \quad \text{for all } i = a, b, c, d, n, v, m, u, q. \tag{95}
   \]

Combining (55) and $z r > 0$ in (92) yield that
   \[
   z r_i > 0 \quad \text{for } i = a, b, c, d, v, n, \tag{96}
   \]
   \[
   z r_i < 0 \quad \text{for } i = u, m, q. \tag{97}
   \]

(95), (96), and (97) yield the results for changes in elasticities of worker skill in items 1 and 2.

Now consider the elasticities of firm revenue and worker pay. According to (89) and (90), firm revenue and worker pay functions, $\tilde{\pi}(\theta)$ and $\tilde{p}(\theta)$, are both power
functions with $T = \frac{(ar+d)qm+nb+cm}{m-b}$ as the exponent of $\theta$. Following (94), we can show that for all $i = a, b, c, d, n, v, m, u, q,$

$$\frac{\partial \epsilon_2^{\epsilon_1}(\theta)}{\partial \theta} - \frac{\partial \epsilon_1^{\epsilon_1}(\theta)}{\partial \theta} = \frac{\partial \epsilon_2^{\epsilon_1}(\theta)}{\partial \theta} - \frac{\partial \epsilon_1^{\epsilon_1}(\theta)}{\partial \theta} = \frac{iT_r r_i}{\theta}. \tag{98}$$

Combining (55) and $T_r > 0$ in (56) yield that

$$T_r r_i > 0 \text{ for } i = a, b, c, d, v, n, \tag{99}$$

$$T_r r_i < 0 \text{ for } i = u, m, q. \tag{100}$$

(98), (99), and (100) yield the results for changes in elasticities of firm revenue and worker pay in items 1 and 2. ■

The worker’s worker skill investment depends on the firm’s investment as well in that the exponent $z$ of the skill investment function $y(\theta) = A\theta^z$ also depends on the exponent $r$ of the firm’s investment function. Similarly, worker pay and firm revenue functions $\tilde{P}(\theta)$ and $\tilde{\pi}(\theta)$ also depend on the firm’s investment in that the exponents $T$ of those functions depends on the exponent $r$ of the firm’s investment function. Therefore, in response to a change in power parameter $i$ given in the model ($a, b, c, d, n, m, u, v, q$), changes in $z$ and $T$ are expressed as

$$\frac{8z}{8i} = z_r r_i + z_i \text{ and } \frac{8T}{8i} = T_r r_i + T_i$$

In both partial total derivatives $\frac{8z}{8i}$ and $\frac{8T}{8i}$ above, the firm terms $z_r r_i$ and $T_r r_i$ represent the additional changes in $z$ and $T$ caused by the changes in the firm’s investment due to a change in power parameter $i$. These terms are absent in one-sided investment and it is these terms that make elasticities of worker skill investment, firm revenue and worker pay in two-sided investment differ from corresponding elasticities in one-sided investment.

Note that increases in $a, b, c, d, v, n,$ and $q$ increase income inequality, whereas decreases in $u$ and $m$ increase it. Therefore, Proposition 14 implies that in a response to a change in each factor contributing to rising inequality except for the relative heterogeneity of firm type ($q$), equilibrium percentage changes in worker skill investment, firm revenue, and worker pay all increases in worker type at a higher rate in two-sided investment than they do in one-sided investment.

However, in a response to an increase in the relative heterogeneity of firm type ($q \uparrow$), equilibrium percentage changes in worker skill investment, firm revenue, and worker pay increases in worker type at a lower rate in two-sided investment than they do in one-sided investment. An increase in $q$ is a good thing for a worker because she can now match with a firm of a higher type, whereas it is a bad thing for a firm because it now matches with a worker of a lower type. This dampens
the firm’s incentive to invest, which is captured in \( r_q < 0 \). If other things equal, it has an adverse effect on the worker’s investment, which is captured in \( z_r q < 0 \). Therefore, the equilibrium percentage changes in worker skill is still increasing in worker type but not as much as it would in one-sided investment. This in turn makes the equilibrium percentage changes in firm revenue and worker pay still increase in worker type but not as much as they would in one-sided investment.

### B.3 Efficiency

For efficiency analysis, first consider the equilibrium when types on both sides of the market is observable. The equilibrium matching pattern continues to be assortative so we continue to use \( \delta(\theta) = k\theta^q \) defined in (52) and its inverse \( \tilde{\theta}(\delta) \). Within a match, the efficient equilibrium investment \( g_e(\delta) \) and \( y_e(\theta) \) are determined by

\[
\begin{align*}
\pi_x(g_e(\delta), y_e(\tilde{\theta}(\delta)), \delta) - e_x(x_e(\delta), \delta) &= 0 \quad (101) \\
\pi_y(g_e(\delta(\theta)), y_e(\theta), \theta, \tilde{\theta}(\delta)) - \epsilon(\theta) &= 0 \quad (102)
\end{align*}
\]

Now we can simultaneously solve the system of equations (101) and (102).

**Proposition 15** The efficient investment functions \( g_e \) and \( y_e \) take the forms of \( g_e(\delta) = D_e \delta^r \) for \( \delta \geq 0 \) and \( y_e(\theta) = A_e \theta^z \) for \( \theta \geq 0 \) with

\[
\begin{align*}
r_e &= r \text{ and } z_e = z \\
D_e &= \left( \frac{a\alpha}{u\tau} \right)^{\frac{m-b}{m-b}} k^{- \frac{(m-b)(m-b)}{q(m-u-b-a)}} \left[ \frac{\alpha b k^{a+d}}{m\beta} \right] ^{\frac{b}{m-b}} \\
A_e &= \left( \frac{a\alpha}{u\tau} \right)^{\frac{a}{m-u-b-a}} k^{- \frac{a(b+c)}{q(m-u-b-a)}} \left[ \frac{\alpha b k^{a+d}}{m\beta} \right] ^{\frac{a}{m-u-b-a}}
\end{align*}
\]

**Proof.** Given the forms of \( g_e(\delta) = D_e \delta^r \) and \( y_e(\theta) = A_e \theta^z \), (101) and (102) imply that

\[
\begin{align*}
A_e &= \left[ \frac{\alpha b k^{a+d}}{m\beta} D_e \right] ^{\frac{1}{m-b}} \text{ and } z_e = \frac{(ar_e + d)q + c + n}{m - b} \\
D_e &= \left( \frac{a\alpha}{u\tau} A_e \right) ^{\frac{b}{u-a}} k^{- \frac{a(b+c)}{q(u-a)}} \text{ and } r_e = \frac{zeb + c + (d+v)q}{(u-a)q}
\end{align*}
\]

We can simultaneously solve \( z_e \) and \( r_e \) and the solutions for \( z_e \) and \( r_e \) are in fact \( r_e = r \) and \( z_e = z \). By plugging \( A_e \) and \( z_e \) into the equation for \( D_e \) and then solving it for \( D_e \) gives us the solution for \( D_e \), which is expressed in terms of parameters in the model. Finally plugging \( D_e \) into the equation for \( A_e \) gives us the expression of \( A_e \) in terms of parameters in the model. \( \blacksquare \)
Efficiency can be measured by the ratio of efficient investment under observable types to equilibrium investment under unobservable types.

\[
E^y := \frac{y_e(\theta)}{y(\theta)} = R^{\frac{u-a}{m-a-bu}} \quad \text{and} \quad E^x = \frac{g_e(\delta)}{g(\delta)} = R^{\frac{b}{m-a-bu}} \quad (103)
\]

with \( R \) being defined as

\[
R := \frac{(ar + d) bq + cb + nb}{(ar + d) bq + cm + nb}.
\]

Note that the exponents of \( R \) in both \( E^y \) and \( E^x \) are positive and \( R < 1 \) because \( m > b \). Therefore, \( E^y \) and \( E^x \) are both less than one, which means that both equilibrium investments \( g(\delta) \) and \( y(\theta) \) are strictly greater than \( g_e(\delta) \) and \( y_e(\theta) \) respectively.

Now let us consider efficiency. For the worker’s investment, consider the efficiency measure

\[
E^y = R^\sigma,
\]

where

\[
\sigma = \frac{u-a}{m-a-bu}.
\]

First of all, \( R \) is a function of power parameters and \( \sigma \) is also a function of power parameters. Therefore, the partial derivative of \( E^y \) with respect to \( i = a, b, c, d, m, n, u, v, q \) is given by

\[
\frac{\partial E^y}{\partial i} = E^y \left[ \frac{\sigma \frac{\delta R}{\delta i}}{R} \frac{\delta R}{\delta i} + \frac{\partial \sigma}{\partial i} \ln R \right] \quad (104)
\]

where \( \frac{\delta R}{\delta i} = R_r r_i + R_q \) and \( \ln R < 0 \) because \( 0 < R < 1 \). The efficiency measure of the firm’s investment is \( E^x = R^\varsigma \), where \( \varsigma = \frac{b}{m-a-bu} \). Then, the partial derivative of \( E^x \) with respect to \( i = a, b, c, d, m, n, u, v, q \) is

\[
\frac{\partial E^x}{\partial i} = E^x \left[ \varsigma \frac{\delta R}{\delta i} \frac{\delta R}{\delta i} + \frac{\partial \varsigma}{\partial i} \ln R \right] \quad (105)
\]

**Proposition 16** The comparative statics on \( E^y \) and \( E^x \) with respect to power parameters are as follows

<table>
<thead>
<tr>
<th>( \frac{\partial E^y}{\partial i} )</th>
<th>( \frac{\partial E^x}{\partial i} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>inc</td>
<td>inc</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>inc</td>
</tr>
<tr>
<td>+</td>
<td>inc</td>
</tr>
<tr>
<td>inc</td>
<td>+</td>
</tr>
</tbody>
</table>

where “inc” stands for inconclusive.

**Proof.** Because \( R \) is a function of power parameters in the model and also \( r \). Because \( r \) is also a function of power parameters, the partial total derivative of \( R \) with respect to \( i \) is given by

\[
\frac{\delta R}{\delta i} = R_r r_i + R_i \quad (107)
\]
We have the following signs of the partial derivatives of $R$:

$$R_a > 0, R_b > 0, R_c < 0, R_d > 0, R_v = 0, R_u = 0, R_n > 0, R_m < 0, R_q > 0, R_r > 0$$  \(108\)

Combining (55), (107), and (108) yields the definite signs of the partial total derivatives of $R$ with respect to $i = a, b, c, d, v, u, n, m$. However, $\frac{\partial R}{\partial i} = R_r r_q + R_q$ is not certain because $R_r > 0, r_q < 0$ and $R_q > 0$. To determine the sign, $R$ can be expanded as follows:

$$R = \frac{(ar + d) bq + cb + nb}{(ar + d) bq + cm + nb} + \frac{abq (dm + vm - bv)}{um - ub - am} + \frac{ab (bn + cm)}{(um - ub - am)} + dbq + cb + nb + \frac{abq (dm + vm - bv)}{am - ub - am} + \frac{ab (bn + cm)}{(am - ub - am)} + dbq + cm + nb$$

The derivative of $R$ with respect to $q$ is then

$$\left(\frac{abq (dm + vm - bv)}{um - ub - am} + db\right) c (m - b) (\frac{abq (dm + vm - bv)}{um - ub - am} + \frac{ab (bn + cm)}{(um - ub - am)} + dbq + cm + nb)^2 > 0$$

Therefore, we can finally complete the signs of the partial total derivatives as follows:

$$\begin{array}{cccccccc}
a & b & c & d & v & u & n & m & q \\
+ & + & - & + & + & - & - & + & +
\end{array}$$ \(109\)

It is easy to see the sign of the partial derivatives of $\sigma$ and $\varsigma$ as follows:

$$\begin{array}{cccccccc}
a & b & c & d & v & u & n & m & q \\
+ & + & 0 & 0 & 0 & - & 0 & - & 0
\end{array} \quad (110)$$

Applying (109) and (110) to (104) and (105) yields (106).  

First of all, (106) shows that the qualitative impacts of changes in power parameters on the efficiency of the firm’s equilibrium investment coincide with those on the efficiency of the worker’s equilibrium investment. Because the equilibrium matching pattern is efficient (i.e. assortative), it implies that we can simply use the efficiency of the worker’s investment to measure the efficiency of the SSME outcome. This is intuitive because firms are residual claimants of the remaining revenue after paying their workers and hence their investments are constrained efficient. Therefore, we only need to look at the efficiency of equilibrium investments made by workers. This implies that the measure of efficiency we employed in the
paper based on the worker’s investment is robust to the possibility of two-sided investment.

Second, the impact of a change in some power parameters on efficiency is not conclusive. For example, an increase in the homogeneity \((a)\) of the firm’s revenue function with respect to the firm’s characteristic acquired by its costly investment has two opposite effects on efficiency. On one hand, it increases \(R\) (i.e. \(\frac{\partial R}{\partial a} > 0\)), which increases the efficiency measure \(E^y = R^\sigma\) given \(\sigma\). This is the impact captured in the first term in the bracket on the right hand side of (104). On the other hand, it also increases \(\sigma\), which reduces \(E^y\) given \(R\) because \(R\) is less than one. This is the impact captured in the second term in the bracket on the right hand side of (104). The sign of \(\frac{\partial E^y}{\partial a}\) is determined by the relative magnitudes of those two effects.

Third, efficiency analysis in two-sided investment produces the qualitative same result as that in one-sided investment in that only changes in firm-related factors \((d \uparrow \text{ and } q \uparrow)\) necessarily increases efficiency, according to Proposition 16. The impact of changes in power parameters related to talent \((b \text{ and } c)\) on efficiency is not clear unless we can clearly say that the change comes from \(c\). Similarly, the impact of changes in power parameters related to investment costs \((v, u, n, \text{ and } m)\) is not clear unless we know that changes comes from either \(v\) or \(n\).