Occupational Choice, Matching and Earnings Inequality*

Eric Mak\textsuperscript{1} and Aloysius Siow\textsuperscript{2}

\textsuperscript{1}Shanghai University of Finance and Economics
\textsuperscript{2}University of Toronto

March 23, 2018

Abstract

Integrating Roy with Becker, this paper studies occupational choice and matching in the labor market. Our model generates occupation earnings distributions which are right skewed, have firm fixed effects, and large changes in aggregate earnings inequality without significant changes in within firm inequality. The calibrated model fits the earnings distribution both across and within firms in Brazil in 1999. It shows that the recent decrease in aggregate Brazilian earnings inequality is largely due to the increase in her educational attainment over the same years. A simulation of skilled biased technical change in the model increased earnings inequality without increasing within firm inequality, consistent with recent US experience.

1 Introduction

In recent decades, earnings inequality has increased in Canada, UK and the US; fallen in Japan, Spain and Latin America; and fluctuated in Germany, South Korea and Mexico\textsuperscript{1}. The variety of experiences suggest that different aggregate shocks are operative.
in different countries. In spite of rapid changes in labor markets, earnings distributions have some invariant characteristics:

1. Occupational earnings densities are continuous, single peaked and right skewed.

2. A major determinant of individual earnings variation in an economy, both across and within firms, are due to the different occupations which workers hold.

3. Firm/establishment fixed effects continue to explain a significant fraction of the variance of log earnings after controlling for individual characteristics, industry and occupation fixed effects (E.g. Groshen (1991), Abowd et al. (1999)).

4. Recent changes of earnings inequality in many countries, either increasing or decreasing, are primarily due to changes in earnings inequality across and not within firms. E.g. Song et al. (2015) (United States); Benguria (2015) (Brazil); Faggio et al. (2010) (UK); Skans et al. (2009) (Sweden); Lee et al. (in process).

Characteristic 1 is well known. Figure 1, which is obtained from the US 2000 Census, plots the earnings distributions for different occupations selected by three different criteria: Occupations by sex ratios (men to women), size (measured by number of workers) and average earnings ranked at the 80th, 50th and 20th percentiles. Except for the lowest percentiles, the earnings distributions have earnings which are weakly convex by percentile. Since the distributions of the demand and supply of skills to an occupation will likely affect the sex composition, size and average earnings, and the share of workers in an occupation in an economy fluctuates over time, occupational demand and supply conditions at a point in time cannot be a first order determinant of convexity.

Concerning characteristic 2, Groshen (1991) estimates that 11-35% of the variation in worker’s earnings are due to their occupations. A significant fraction of the gender wage gap is due to different occupational choices (E.g. Goldin (2014)). Occupational differences within firms account for most of the differences in earnings inequality within firms. Handwerker (in process) estimates that more than $\frac{2}{3}$ of the average variance of within establishment wages in the US are due to workers occupying 98 occupations (3 digits SOCS codes). This estimate is a significant lower bound on within establishment variances of ln(wage) cover 86,935,498 individuals at the establishment level by 6-digit occupation by wage interval level, weighted by the number of employees in each wage interval and by benchmarked establishment weights roughly corresponding to the inverse of sampling probabilities.

\[2\]
establishment wage variation which can be explained by occupational wage differences because 3 digit occupation codes are broad and skill differences within an occupation are not distinguished. E.g., all doctors and nurses are grouped in occupation 291. All teaching staff in universities are in occupational code 251.

Characteristics 3 was discovered shortly after economists started estimating log earnings regressions. After controlling for individual characteristics, industry and occupational effects, firm/establishment fixed effects explain a significant fraction of the residual variance of cross section log earnings Groshen (1991). Following Abowd et al. (1999), economists extended the analysis with panel data to estimate log earnings regressions with both worker and firm/establishment fixed effects. The explanatory power of the firm/establishment effects remain large. Some researchers (e.g. Card et al. (2013)), but not all, show that the correlation between individuals’ and firms’ fixed effects is quantitatively large. I.e. controlling for observables, including occupation, workers with high average earnings work primarily in firms with high average earnings. This potentially high correlation imply that there is positive assortative matching of co-workers by ability. The popular press also noticed this correlation:

*The recruiting is not confined to the best engineers; sometimes it spills over to nontechnical employees too. Two of the chefs who prepared meals for Googlers, Alvin San and Rafael Monfort, have been hired away by Uber*

Data were collected between Fall 1999 and November 2015 and are pooled together. Occupations used in dividing variance into within and between occupation are at the 3 digit level.
Characteristic 4 is a recent discovery. In recent decades, labor earnings inequality within many countries have changed significantly. For many countries, including the US, earnings inequality have risen. For other countries, as will be shown below for Brazil, it has fallen.\footnote{Declining earnings inequalities were and are common in Latin America (Lustig et al. (2013)). For Spain, see Pijoan-Mas and Sánchez-Marcos (2010).} What about changes in across and within firm earnings inequality? To a first order, these papers show that recent aggregate changes in earnings inequality were primarily due to changes in earnings inequality across firms and non-within firms.

Countries differ in their labor market institutions such as degree of unionization, size of the informal sector, centralized versus decentralized wage bargaining, minimum wages, etc.. Different countries also experience different recent changes in earnings inequality due to different changes in country specific institutions, impact of foreign trade, skill distributions, skill biased technical change, and other aggregate shocks. Thus the invariances we discussed above must be due to first order factors which apply to every economy.

More than 200 years ago, Adam Smith proposed an answer. He observed specialization and the division of labor within firms in all but the smallest firms. He also proposed that compensating wage differentials sorted workers into different occupations. He did not study how workers are sorted across firms.

In every economy, each worker chooses an occupation and a firm to match with. The objective of this paper is to study how these choices affect the distribution of earnings in the economy. We integrate two classic models, the Roy model of occupational choice (Roy, 1951) and the Becker model of matching (Becker, 1973). The Roy model of occupational choice based on comparative advantage is the standard model of occupation choice (E.g. French and Taber (2011)). It has been extended in several directions including being used to study labor market clearing without matching (E.g. Heckman and Seldacek (1985); Lee and Wolpin (2006)). Becker’s model of matching is widely used to study matching between workers and firms (E.g., Chade et al. (2017),Eeckhout (2017)). In an influential study, Kremer (1993) applied it to occupational matching in teams. Our integration of these two models is new.

Due to supermodularity of a firm’s revenue function in occupational skills, Becker’s matching model can generate convex earnings where good workers match and rein-
force each other in good firms (characteristic 2) (Kremer, 1993); this also explains firm effects in log earnings regressions (characteristic 3). But there is no guarantee that the employees in a firm, working in different occupations, have an invariant relative difference in earnings (characteristic 4). When an exogenous shock affects one occupation, the earnings of one occupation may fall relative to that of the other occupation, thereby changing relative wages within firms. With occupational choice, there will be an influx of workers from the lowered earnings occupation to the high earnings occupation; this arbitrage will drive the occupational earnings distributions closer together. This equilibration argument is of course familiar from Roy. Positive assortative matching by occupational skills within a firm then buys the further implication that wages across occupations within firms remain stable.

We are not the first to study occupational choice and matching. In the classic Kremer Maskin model (KM hereafter) (Kremer and Maskin, 1996), individuals differ by one dimension of skill and they work as either managers or workers. With one dimension of heterogeneity, KM is a model of absolute advantage in occupational choice where, as the skill level of an individual increases, the individual will switch from a low skill occupation to a high skill occupation. The individual who is indifferent between the two occupations is the highest paid low skill worker and the lowest paid high skill worker. What this means is that KM often predicts low skill occupational earnings densities which truncates on the right instead of having long right tails as well as unconnected occupational earnings densities (E.g. Porzio (2015)). These discontinuous features may be reasonable descriptions of hierarchical labor markets where the earnings of workers in low skill occupations in a hierarchy have a small support. Lucas Jr (1978); Rosen (1981); Garicano and Rossi-Hansberg (2004) provide such models. Porzio (2015) has a survey. These discontinuities are not a robust feature of occupational earnings as shown in Figure 1. There is also overwhelming evidence that individuals have multi-dimensional skills, e.g. cognitive, strength, dexterity and social skills, which are valued in the labor market. Equilibrium multifactor occupational choice and matching models are difficult to characterize (E.g. McCann and Trokhimtchouk (2010); Lindenlaub (2016)). Our paper is a compromise which allows for multidimensional occupational choice and retains one dimensional matching of workers within a firm by occupational skills.

In order to discuss earnings inequality within and across firms, workers in dif-

\footnote{Heckman and his collaborators led this literature among economists (E.g. Borghans et al. (2008)). Deming (2017) is a recent application.}
ferent firms have to differ in essential ways. In this paper, different firms produce
different qualities of output. A firm can only produce higher quality output by hiring
higher skill workers but not more workers of the same skill\(^5\). I.e. we are assuming that
a bakery produces high quality cakes by hiring a better baker and not more bakers.
For analytic convenience, we ignore quantity concerns\(^6\).

This paper is not a general model of the labor market. We ignore the roles of labor
market frictions, firm rents, heterogeneity in firm size and dynamic considerations,
considerations which are important features of the labor market (E.g., Chade et al.
(2017); Eeckhout (2017)). Rather, we focus on the effects of occupational choice and
matching on earnings inequality.

Our paper complements the existing literature on skill biased technical change
(SBTC) and earnings inequality.\(^7\) These models assume that firm output satisfies con-
stant returns to scale in occupational skills. Workers’ skills are measured in efficiency
units and workers in the same occupation are perfect substitutes. Labor market equi-
librium is determined by an aggregate production function and the aggregate supply of
skills to each occupation. By construction, the standard model, which was constructed
before characteristic 4 was discovered, is silent on earnings inequality between and
within firms, and differences in product quality across firms.

In order to investigate the properties of our model quantitatively, we use our frame-
work to study the recent decline in earnings inequality in Brazil. The average years
of educational attainment doubled in Brazil from 1999 to 2013. We ask whether this
change in schooling can explain the observed decline in earnings inequality and also
exhibit the four invariant features of the earnings distributions\(^8\). Using educational
attainment as a proxy for cognitive skill, we first calibrate the parameters of our model
with the distributions of individual earnings and average earnings by firm in 1999.
Then we simulate the earnings distributions predicted by our estimated model with
the educational distribution in 2013. Our simulation replicates the Brazilian data
where for both years, the occupational earnings distributions are single peaked and
right skewed. Earnings inequality in both distributions are almost entirely due to
the between-firm component. We also rationalize a significant decline in Brazilian

\(^5\)Related studies include Verhoogen (2008); Helpman et al. (2017); Orr (2018)
\(^6\)Eeckhout (2017) provides a general framework of one factor matching models with variable firm
size.
\(^7\)E.g. Katz and Murphy (1992). Acemoglu and Autor (2011) provides a survey of the standard model
of skill biased technical change and changes in earnings inequality.
\(^8\)We build on Lam and Levison (1991) and Galiani et al. (2017).
earnings inequality between 1999 and 2013, as well as little change in within firm inequality between the two periods\textsuperscript{9}. For comparison, we will also calibrate the Kremer Maskin model to the same data.

What happens when there is skill biased technical change (SBTC)? Using our quantitative model, we simulate the effect of SBTC. Compared with the 2013 benchmark, occupational choice and matching significantly mitigates the earnings inequality due to SBTC, particularly among the lesser skilled workers. Moreover, occupational choice reduces earnings inequality within firms.

In addition to the above cited papers, this paper also builds on worker firm matching models (E.g., Chade et al. (2017); Eeckhout (2017)). Most of the worker firm matching literature focus on other issues than what we discuss here, such as earnings dynamics, inter firm mobility, and firm size. McCann et al. (2015) study a model of schooling investment, occupational choice and matching where workers differ by cognitive and communication skills. Our environment here is simpler which allows us to focus on different questions. Gola (2016) studies a two dimensional occupational choice problem similar to ours. He differs by studying the matching of firms to workers in occupational specific industries.

We present the empirical facts concerning earnings inequality in Section 2. Section 3 presents our model and Section 4 provides some characterizations. Section 5 shows the equivalence between our competitive model and a utilitarian social planner’s linear programming problem. Section 6 provides our quantitative rationalization of the Brazilian experience. Our exploration of SBTC is in Section 7. Section 8 concludes.

2 Two Empirical Studies

This section reviews Song et al. (2015) and Benguria (2015) which pertain to our work.\textsuperscript{10} The two papers use the same empirical strategy. Song et al. (2015) analyzed data from the US Social Security Administration master file from 1982 to 2012. This data consists of the W2 forms filed annually by every employer for each employee to the US tax authority, the IRS. Their sample has between 66 million to 153 million workers per year, and between 0.8 million to 1 million firms per year. Individuals in firms

\textsuperscript{9}Helpman et al. (2017) estimated a model of international trade and Brazilian earnings inequality. Engbom and Moser (2017) study the impact of an increase in the minimum wage on earnings inequality. Neither paper discusses changes in within firm inequality.

\textsuperscript{10}For the US, Barth et al. (2014) has similar results by establishments.
with less than 10 workers were excluded. The earnings information per worker includes wages and salaries, bonuses, exercised stock options, the dollar value of vested restricted stock units and other sources of income. Wage earnings are top coded at the 99.999th percentile.

Benguria (2015) uses the Relação Anual de Informações (RIAS) data from 1999-2013. This data is filed annually by all registered employers on their employees. He analyzes a 10% random sample of the data set. The sample has over 5 million workers in 2013. The reported average monthly earnings “are gross and include not only regular salary but also bonuses and other forms of compensation”. All earnings are deflated using the Brazilian consumer price index.

For any year $t$, let the log earnings of worker $i$ and the mean of log earnings in firm $j$ be $w_{ij}^t$ and $\bar{w}_j^t$ respectively. For year $t$, we can decompose the variance of individual log earnings into a between firms variance of log earnings and a within firm variance of log earnings:

$$
\text{var}(w_{ij}^t) = \text{var}(\bar{w}_j^t) + \sum_{j=1}^{J_t} P_j^t \times \text{var}(w_{ij}^t | i \in j)
$$

$J_t$: number of firms in year $t$

$P_j^t$: $j$'s share of employment in year $t$

![Figure 2: U.S. Trends of Total, Between and Within Inequality](image)
All US figures are from Song et al. (2015). Figure 2 shows the evolution of the variance decomposition for the US from 1980 to 2012. The top line is the evolution of the total variance over time. It has a significant upward trend which shows the well known increase in aggregate inequality in earnings in the US in recent decades. The middle line is the evolution of the variance of within firm earnings. Finally, the bottom line is the evolution of the between firm variance of earnings. Note that the variance of within firm earnings is larger than that of between firms earnings. So there are significant differences in earnings within firm. On the other hand, the slope of within firm earnings over time is significantly flatter than the slope for overall inequality. Rather, the slope of the between firm variance over time has the same slope as the slope for overall inequality.

![Figure 3: Brazilian Trends of Total Inequality](image)

All Brazil figures are from Benguria (2015). Figure 3 shows the evolution of the total variance of earnings for Brazil from 1999 to 2013. The slope of this (top) line is completely different from what happened in the US. The variance of aggregate earnings fell by 32% (21 log points) over the period 1999 to 2013.

Figure 4 shows the decomposition of the decline of the aggregate variance into between and across firms variation. Their panel A shows that most of the decline in aggregate variance is reflected in the decline in across firms variance. There is at best a modest decline in within firm variance. Unlike the US, there is more between firm inequality than within firm inequality.
Figure 4: Brazilian Between versus Within Firm Inequality

For a finer decomposition within a year, the authors use another decomposition of earnings inequality for year $t$. Let $W^i_{pt}$ be the mean of $w_{ij}^t$ of all workers in the $p'th$ percentile in the earnings distribution in year $t$. Let $W^j_{pt}$ be the mean of $w^j_t$ for each worker in the $p'th$ percentile. Then:

$$W^i_{pt} = W^j_{pt} + (W^i_{pt} - W^j_{pt})$$

$W^i_{pt}$ is decomposed into a part which is the mean of the firms in which these workers work in, $W^j_{pt}$, and a residual, $(W^i_{pt} - W^j_{pt})$, which is how these workers’ mean log earnings deviate from their firms’ mean. The change in earnings inequality by percentile from year $t$ to year $t'$ is:

$$W^i_{pt'} - W^i_{pt} = W^j_{pt'} - W^j_{pt} + (W^i_{pt'} - W^j_{pt'}) - (W^i_{pt} - W^j_{pt})$$

Figure 5 shows the changes in earnings inequality by percentile from 1982 to 2012 in the US. The diamond line represents the well known increase in overall inequality. The circle line, which is essentially on top of the diamond line, is the change in firm inequality by percentile. Since the difference between the two lines is the residual, the bottom line is the change in within firm inequality by percentile. What is remarkable is that there is, to a first order, no change in within firm inequality by percentile.11 The change in overall inequality has received significant attention from the policy makers and researchers. It represents significant changes to how the US labor market evolved.

Figure 6 shows the changes in earnings inequality by percentile from 1999 to 2013.

---

11Their paper showed that there was a large change in within firm inequality at the 99.8 percentile.
in Brazil. Earnings inequality fell significantly from 1999 to 2013. That is, earnings for the top percentiles did not increase as much as for the lower percentiles. On the other hand and similar to the US experience, there was essentially no change in within firm inequality except at the lowest percentiles.

Taken together, the above figures and the studies by Faggio et al. (2010) for the UK and Skans et al. (2009) for Sweden, showed that total inequality within a country can change significantly in a few decades and in different ways. In spite of sometimes large overall changes, there was essentially no change in within firm inequality by percentile. These effects are not well captured by existing models of the labor market.
3 The Model

3.1 The Setup

Consider a labor market with a unit mass of workers. Each worker has two base skills \((c, r)\), his or her cognitive skill and non-cognitive skill respectively. \(c\) and \(r\) are distributed according to the continuous bivariate density \(b(c, r)\), such that \(\bar{b} > b(c, r) > 0\) with positive domain \([c, \bar{c}] \times [r, \bar{r}]\). There is no atom in the density.

Production takes place in a team of two workers. One worker is a key role worker and the other is a support role worker. Consider a team with a key role worker with characteristics \((c_1, r_1)\) and a support role worker with characteristics \((c_2, r_2)\). The revenue they produce is:

\[
\bar{R}(c_1, r_1; c_2, r_2) = R(k_1; s_2)
\]

The cognitive skill of the key role worker, \(c_1\), and her non-cognitive skill, \(r_1\), interacts to form an index of key role skill, \(k_1 = g_k(c_1, r_1)\). Analogously, the skill index of the support role worker is \(s_2 = g_s(c_2, r_2)\). Consider a worker with base skills \((c, r)\). We assume that \(g_k(c, r)\) is not a monotone transform of \(g_s(c, r)\) so that the two occupations rank at least some of the same workers differently.

We impose the following assumption on the technology \(R\):

**Assumption 1 (Supermodularity).** \(R\) is strictly supermodular in \(k_1, s_2\).

As is well known since Becker and we will also show below, the supermodularity assumption will result in PAM by occupational skills.

The density of \((k, s)\) is a continuous function \(f\), derived from \(b\) after a transform of variables, is also strictly bounded above and its domain remains a rectangle \(\Omega = [k, \bar{k}] \times [s, \bar{s}]\). We assume that \(R(k_1; s_2) > 0\) for all \((k, s) \in \Omega\).

Let \(\pi(k)\) be the earnings of a key role worker with skill \(k\). Let \(w(s)\) be the earnings of a support worker with support role skill \(s\). For the moment, assume that the earnings functions for both types of workers are increasing and convex in their occupational specific skills.

The following discusses occupational choice and matching given \((\pi, w)\). We first show that the optimal occupational choice is characterized by complete segregation, with the \((k, s)\) space partitioned into two equal halves by a separating line. Then we show that matching follows PAM. Therefore, the competitive equilibrium is characterized by a separating line and a matching line.
3.2 Occupational Choice

Workers choose the occupation which will maximize their net earnings. So a worker of type \((k,s)\) will earn:

\[ y(k,s) = \max[\pi(k), w(s)] \] (1)

If \(\pi(k) > w(s)\), the worker will be a key role worker. If \(\pi(k) < w(s)\), the worker will be a support role worker. If \(\pi(k) = w(s)\), the worker will be indifferent between the two roles. Therefore, the type space \(\Omega\) is partitioned into three sets \(\Omega_k \equiv \{(k,s) \in \Omega | \pi(k) > w(s)\}\), \(\Omega_s \equiv \{(k,s) \in \Omega | \pi(k) < w(s)\}\), and \(\Omega_{ks} \equiv \{(k,s) \in \Omega | \pi(k) = w(s)\}\). Presuming that \(\pi, w\) are continuous, strictly increasing functions in their respective arguments (which we shall justify later), \(\Omega_{ks}\) is an upward sloping line in \(\Omega\) since for a higher \(k\) worker, he would be indifferent between the two roles only if his \(s\) is also higher. Accordingly, we define the separating function \(\phi : [\underline{k}, \overline{k}] \rightarrow [\underline{s}, \overline{s}]\) such that

\[ \phi(k) = \min\{w^{-1}(\pi(k)), \overline{s}\} \]

As such, if \(\phi(k) < \underline{s}\), then workers with characteristics \((k, \phi(k))\) are indifferent between the two occupations:

\[ w(\phi(k)) = \pi(k) \] (2)

While if \(\phi(k) = \overline{s}\), then a worker with skill \(k\) will always prefer the key role regardless of his \(s\). In the discussion below and in the simulation, \(\phi(k) < \underline{s}\) for all \(k\). So we focus on (2).

The separating function is a central concept of the Roy model of occupational choice. We summarize the above discussion in Proposition 1.

**Proposition 1 (Separating Function).** Consider a worker with characteristics \((k,s)\). If \(s > \phi(k)\), the worker will choose to be a support role worker. If \(s < \phi(k)\), the worker will choose to be a key role worker. If \(s = \phi(k)\), the worker will be indifferent between the two occupations. \(\phi(k)\) is non-decreasing in \(k\).

Given \(\phi(k)\), the cumulative distribution of key role workers from ability \(\underline{k}\) to \(k\) is:

\[ H(k) = \int_{\underline{k}}^{k} \int_{\underline{s}}^{\phi(u)} f(u,v) dv du \] (3)
The cumulative distribution of support role workers from ability $s$ to $s$ is

$$G(s) = \int_{\bar{s}}^{s} \int_{\bar{k}}^{k} \phi^{-1}(u) f(u,v) du dv$$  \hspace{1cm} (4)

### 3.3 Matching

Since we are in a competitive environment and there is no cost of entry of firms/teams, a firm would hire a key role worker and a support role worker according to:

$$\max_{(\hat{k}, \hat{s}) \in \Omega} R(\hat{k}; \hat{s}) - \pi(\hat{k}) - w(\hat{s})$$  \hspace{1cm} (5)

The optimal choice of $(k; s)$ will satisfy:

$$R_k(k, s) = \pi'(k)$$  \hspace{1cm} (6)

$$R_s(k, s) = w'(s)$$  \hspace{1cm} (7)

Since the technology is strictly increasing in $k$ and $s$, $\pi, w$ are strictly increasing functions as we presumed.

We can invert either (6) or (7) to get the matching function $\mu : [k, \bar{k}] \rightarrow [s, \bar{s}]$ such that $s = \mu(k)$, where a key role worker of skill $k$ will match with a support worker of type $s$. The matching function is a central concept in Becker’s matching model and here as well.

**Proposition 2** (Matching Function). $\mu' > 0$: There is PAM between key role workers and support workers by occupational skills.

**Proof.** The proof is by contradiction. Consider two key role workers, $k_A$ and $k_B$, $k_A > k_B$ and two support workers $s_A$ and $s_B$, $s_A > s_B$ such that $(k_A; s_A), (k_B; s_B)$ both satisfy the first-order conditions. Then we have PAM. Due to free entry of the entrepreneur,

$$R(k_A; s_A) + R(k_B; s_B) - \pi(k_A) - \pi(k_B) - w(s_A) - w(s_B) = 0$$  \hspace{1cm} (8)

Suppose that a non-PAM rearrangement $(k_A; s_B), (k_B; s_A)$ also satisfies the first-order condition. Then

$$R(k_A; s_B) + R(k_B; s_A) - \pi(k_A) - \pi(k_B) - w(s_B) - w(s_A) = 0$$  \hspace{1cm} (9)
Together, they imply that 
\[ R(k_A; s_A) + R(k_B; s_A) - [R(k_A; s_B) + R(k_B; s_A)] = 0, \]
which violates supermodularity of \( R \).

Unlike one factor models of matching and occupational choice, there is no conflict between PAM and occupational choice. The reason for our lack of conflict is because in our two factor model of skills, occupational choice is due to comparative advantage and PAM is due to absolute advantage.

We are now ready to define an equilibrium for this labor market.

**Definition 1.** An equilibrium consists of an earnings function for support workers, \( w(r) \), an earnings function for key role workers, \( \pi \), a separating function, \( \phi \), and a matching function, \( \mu \), such that:

1. All workers choose occupations which maximize their net earnings, i.e. solve equation (1).
2. A free-entry entrepreneur chooses key role workers and support role workers to maximize its net earnings (which is zero), i.e. solve equation (5).
3. The labor market clears. I.e. every worker of type \((k, s)\) can find the job which maximizes his or her net earnings. Due to PAM, the labor market clearing condition can be written as:

\[
H(k) = G(\mu(k)), \forall k
\]  

Equation (10) says that for every \( k \), the mass of key role workers up to skill \( k \) must be equal to the mass of support role workers up to skill \( \mu(k) \).

### 4 Characterizations

First, we appeal to McCann et al. (2015) for the existence of competitive equilibrium.

**Theorem 1 (Existence).** An equilibrium, consisting of four unique functions, an earnings function for support workers, \( w(s) \), an earnings function for key role workers, \( \pi(k) \), a separating function, \( \phi(k) \), and a matching function, \( \mu(k) \), exists.

Second,
**Proposition 3** (Convex Earning Schedules). If \( R \) is strictly increasing and convex in \( k_1, s_2 \), then \( w(s) \) and \( \pi(k) \) are convex in \( s \) and \( k \) respectively.

**Proof.** From the optimal choices of key role workers, taking the second derivatives yields:

\[
\pi''(k) = R_{kk}(k; \mu(k)) + R_{ks}(k; \mu(k)) \mu'(k)
\]

Convexity of the revenue function in occupational skills imply \( R_{kk} > 0 \) and supermodularity implies \( R_{ks} > 0 \). PAM implies \( \mu' > 0 \). So \( \pi''(k) > 0 \). By symmetry, \( w''(s) > 0 \).

While sufficient to obtain convexity of the occupational earnings functions, \( R_{kk} > 0 \) is not necessary. E.g. \( R_{kk} = 0 \) will work as well. Our empirical results obtain convex earnings function without imposing convexity of the revenue function. The convexity assumption is that quality is a convex function of occupational skills. This embodies two assumptions:

First, higher quality output/revenue is due to higher quality workers.

Second, convexity of skills is necessary if we want workers to specialize in their occupational skills investments. If the returns to skills are not convex, workers will diversify in the skills investment which is not what we see. Rosen (1983), Yang and Borland (1991), and others have provided microfoundations for this convexity assumption.

We then discuss links between the separating and matching functions.

**Proposition 4** (Identical Outside Options for the Worst Match). \( \phi(k) = \mu(k) \), such that the worst match is a self-match. Since self-match splits the output evenly, the earnings inequality within this team is zero.

**Proof.** If \( \phi(k) = s \), the type \( (k, s) \) is the worst type among key role workers and support role workers in the equilibrium. Hence under PAM this type self-matches, equally splitting the output. Hence our proposition holds.

Now suppose not, where \( \phi(k) = s^* > s \). Then \( (k, s) \in \Omega_k \), such that this type works exclusively in the key role. Given that \( \phi \) is strictly increasing, the lower support of the support role workers has support role skill \( s^* \). Hence under PAM, \( \mu(k) = s^* = \phi(k) \).

**Proposition 5** (Within-Firm Inequality). \( \pi(k) - w(\mu(k)) > 0 \) if and only if \( \phi(k) > \mu(k) \).
Proof. $\pi(k) = w(\phi(k))$. As $w$ is strictly increasing, $w(\phi(k)) - w(\mu(k)) > 0$ if and only if $\phi(k) > \mu(k)$.\qed

$\phi(k) > \mu(k)$ implies that the worker type $(k, \mu(k))$ works exclusively in the key role. This means that although key role workers $k$ will match with support role workers $\mu(k)$, the type $(k, \mu(k))$ will not self-match, i.e. there is specialization within the firm. This can only be the case when $\pi(k) > w(\mu(k))$, such that working in the key role is strictly better off for this type.

Proposition 5 implies that within-firm inequality depends on the wedge between the separating line and the matching line. Given that $\phi(k) = \mu(k)$, we are interested in how $\phi$ and $\mu$ differ up along the ranks. We will now provide a link between the separating function, $\phi$ of occupational choice and the matching function $\mu$. Differentiating the indifference condition (2) with respect to $k$,

$$w'(\phi(k))\phi'(k) = \pi'(k) \quad (11)$$

Substituting (6), (7) and $s = \mu(k)$ yields:

$$\phi'(k) = \frac{R_k(k; \mu(k))}{R_s(\mu^{-1}(\phi(k)); \phi(k))} \quad (12)$$

Equation (12) is an optimality condition. It provides a restriction between the separating function and matching function that depends on the technology $R$. $\phi'(k)$ is the slope of the separating line, representing marginally how workers separate into key role and support role. The fraction on the right hand side is a ratio of marginal products: $R_k(k; \mu(k))$ is the marginal product of the indifferent type $(k, \phi(k))$ when he works as a key role worker; $R_s(\mu^{-1}(\phi(k)); \phi(k))$ is the marginal product of $(k, \phi(k))$ when he works as a support worker. The larger this fraction, optimally locally more types should be assigned to the key role.

### 4.1 A Closed-Form Example

To get further insights, next we consider a special case of our model to study its analytical properties. We assume the skill distribution of $(k, s)$ is bivariate standard normal; the revenue function is:

$$R(k; s) = \exp(k)^{\alpha_k} \exp(s)^{\alpha_s} = \exp(\alpha_k k + \alpha_s s)$$
where \( \alpha_k, \alpha_s > 0 \).\(^{12}\)

There is no loss in generality in assuming that \( k, s \) being independent, since otherwise one can project \( s \) onto \( k \) and redefine \( s \) as the residual \( s - E[s | k] \). For convenience, we standardize the unit of measurement for the skills such that \( \alpha_k + \alpha_s = 1 \). This technology is supermodular in \((k, s)\) since \( R_{ks} = \alpha_k \alpha_s \exp(\alpha_k k + \alpha_s s) > 0 \).

4.1.1 Case 1: Random Occupational Choice and Matching

If occupational choice is random, for any type \((k, s)\), half of its population mass works in the key role, while the other half works in the support role. Without selection, the occupational choice distribution of key role workers and support role workers are univariate standard normals, and that they are independent.

Suppose that the matching is also random, such that key role and support role workers are randomly paired up regardless of their types. Consequently,

\[
\alpha_k k + \alpha_s s \sim N(0, \alpha_k^2 + \alpha_s^2)
\]

so that

\[
R(k; s) \sim LN(0, \alpha_k^2 + \alpha_s^2)
\]

That is, the revenue is log-normal distributed with variance \( \alpha_k^2 + \alpha_s^2 \).

4.1.2 Case 2: Random Occupational Choice, PAM

Case 2 maintains the assumption of random occupational choice, but here we introduce frictionless matching between key role workers and support role workers. Because of supermodularity of the technology, positive assortative matching (PAM) results.

Since the occupational skill distributions of \( k, s \) are identical, PAM implies that the matching function is \( \mu(k) = k \), a 45 degree straight line through the origin of the \((k, s)\) space.

As a result,

\[
\alpha_k k + \alpha_s \mu(k) = (\alpha_k + \alpha_s)k \sim N(0, 1)
\]

and hence

\[
R(k, s) \sim LN(0, 1)
\]

So that the revenue is log-normally distributed. Relative to the Case 1, the resulting

\(^{12}\)The type space here is unbounded unlike what was assumed, but one can impose finite but arbitrarily large bound to make this assumption true.
revenue distribution is more skewed than the case without PAM since \( \alpha_k^2 + \alpha_s^2 < 1 \). This result agrees with Kremer (1993).

The revenue between any two teams is split in a fixed ratio of \( \alpha_k : \alpha_s \) since:

\[
\pi(k) = \int_{-\infty}^k R_k(u; \mu(u)) du = \int_{-\infty}^k \alpha_k \exp((\alpha_k + \alpha_s)u) du = \frac{\alpha_k}{\alpha_k + \alpha_s} \exp(k) = \alpha_k \exp(k)
\]

and

\[
w(s) = \int_{-\infty}^s R_s(\mu^{-1}(v); v) dv = \int_{-\infty}^s \alpha_s \exp((\alpha_k + \alpha_s)\nu) dv = \frac{\alpha_s}{\alpha_k + \alpha_s} \exp(s) = \alpha_s \exp(s)
\]

and that

\[
\frac{\pi(k)}{w(\mu(k))} = \frac{\alpha_k \exp(k)}{\alpha_s \exp(k)} = \frac{\alpha_k}{\alpha_s}
\]

### 4.2 Case 3: Optimal Occupational Choice, PAM

Occupational choice introduces censoring, so that the occupational skill distribution differs from the original (marginal) skill distribution. Here we restrict \( \alpha_k = \alpha_s = 1/2 \), so that the technology is symmetric.

**Proposition 6.** Assume that \( \alpha_k = \alpha_s \) and there is optimal occupational choice and PAM. At the optimum, the matching-separating function is the 45 degree line \( \mu(k) = \alpha_k = k \), and that occupational skill distributions are skewed normal (skewed to the right). There is no within-firm inequality.

**Proof.** Consider a candidate separating function which is a straight line passing through the origin \( \phi(k) = k \).

The occupational skill distribution is \( p(k|k \geq s) \) Similarly for the support role, the occupational skill distribution is: \( p(s|k < s) \). Let \( z = s - k \sim N(0, 2) \). The occupational skill distributions are determined by Bayes’ Rule:

\[
p(k|z < 0) = \frac{Pr(z < 0|k)\phi(k)}{Pr(z > 0)}
\]

\[
p(s|z > 0) = \frac{Pr(z > 0|s)\phi(s)}{Pr(z < 0)}
\]
Since the total population must be divided equally into the two roles, we have
\( Pr(z > 0) = Pr(z < 0) = 0.5 \). The next step is to determine \( Pr(z < 0 | k) \) and \( Pr(z > 0 | s) \). Note that \( z|k, z|s \) are normally distributed. Also, \( E[z|k] = \frac{\text{cov}(z,k)}{\text{var}(k)} k = -\phi_0 k, E[z|s] = \frac{\text{cov}(z,s)}{\text{var}(s)} s = s \), and that \( \text{var}(z|k) = \text{var}(s) = 1, \text{var}(z|s) = \text{var}(-k) = 1 \). As a result:

\[
Pr(z < 0|k) = \Phi\left(\frac{0 - E[z|k]}{sd(z|k)}\right) = \Phi(k)
\]

\[
Pr(z > 0|s) = 1 - \Phi\left(0 - \frac{E[z|s]}{sd(z|s)}\right) = \Phi(s)
\]

where \( \Phi(.) \) is the cumulative density function of the standard normal distribution. Therefore,

\[
p(k|z < 0) = 2\Phi(k)\phi(k)
\]

\[
p(s|z > 0) = 2\Phi(s)\phi(s)
\]

so that both occupational skill distributions are skew normally distributed. This then implies that the matching function is \( \mu(k) = k \), and that \( \frac{\pi(k)}{\omega(\mu(k))} = 1 \).

Because of the skewness of the occupational skill distributions, the resulting occupational earnings distributions is more skewed than a lognormal. This result is the source of convexity generated by our model.

The last step is to check the optimality condition (12). Given the functional form assumptions and the candidate separating function \( \phi(k) = k \), we require

\[
\frac{0.5 \exp(0.5k + 0.5\mu(k))}{0.5 \exp(0.5\mu^{-1}(\phi(k)) + 0.5\phi(k))} = \phi(k) = k
\]

to hold. Given that \( \mu(k) = \phi(k) = k \), this condition is true.

\[ \square \]

4.3 Kremer Maskin

Kremer Maskin is a one factor model of both occupational choice and matching. Our model is a two factor model of occupational choice and a one factor model in occupational matching. Starting from two base skills, \( c \) and \( r \), we use skill aggregation functions, \( g_k(c,r) \) and \( g_s(c,r) \) to generate one factor occupational skills for each occupation. We assume that \( g_k(c,r) \) is not a monotone transform of \( g_s(c,r) \) so that the two occupations rank at least some of the same workers differently.
We obtain our version of the Kremer Maskin model by assuming \( g_k(c, r) = g_s(c, r) = g(c, r) \) instead. What this means is that each worker of type \( c \) and \( r \) has one skill \( g(c, r) \) which the worker can apply to either occupation. Put another way, if worker A has higher skill than B in the key role, A will also have higher skill than B in the support role and vice versa. Although we can compute the Kremer Maskin equilibrium, from a behavioral perspective, Kremer Maskin is degenerate because the Roy model is specifically a two factor model of occupational choice whereas Kremer Maskin is a one factor model of occupational choice.

We will not discuss the analytic properties of our Kremer Maskin model because it has been discussed in the literature already. Rather, we will calibrate our Kremer Maskin model to our Brazilian data and contrast its quantitative properties to our two factor model.

5 Social Planner’s Problem and its Dual

Our model of frictionless occupational choice and matching is equivalent to a social planner determining occupational choices and matching for the population to maximize total revenue of the economy. In fact, the social planner’s problem is a linear programming problem. McCann et al. (2015)’s proof of existence and uniqueness uses this equivalence. This section develops the Social Planner’s problem in detail because a linear program is a much easier problem to numerically solve than looking for a fixed point of our competitive model. This is how we estimate the model and also produce simulation results.

Let \( m : \Omega \rightarrow \mathbb{R}_+ \) be a density function such that \( m(k_1; s_2) \) states the mass of team \((k_1; s_2)\), and let \( \sigma_k, \sigma_r : \Omega \rightarrow \mathbb{R}_+ \) be density functions such that \( \sigma_k(t), \sigma_r(t) \) record the mass of agents of type \( t = (k, s) \in \Omega \) working in key role and support role respectively.

We have two accounting constraints:

(Accounting Constraint for Key Role):
\[
\int_{\bar{s} \in [s, \bar{s}]} m(k; \bar{s})d\bar{s} = \int_{\bar{s} \in [s, \bar{s}]} \sigma_k(k, \bar{s})d\bar{s}, \forall k \in [\bar{k}, \bar{k}]
\]

(Accounting Constraint for Support Role):
\[
\int_{\bar{k} \in [k, \bar{k}]} m(\bar{k}; s)d\bar{k} = \int_{\bar{k} \in [k, \bar{k}]} \sigma_r(\bar{k}, s)d\bar{k}, \forall s \in [s, \bar{s}]
\]

The first accounting constraint states that the total mass of teams that involves key role workers \( k \) must be equal to the total mass of individuals whose key role skill is \( k \), and that they select to work in the key role. The second accounting constraint is
similarly defined.

We also have a resource constraint:

$$\sigma_k(t) + \sigma_r(t) = f(t) \forall t \in \Omega$$

Given \{R, f\}, the Social Planner allocates agents in teams to maximize social output, defined as the integral of team outputs. Let the space of \{m, \sigma_k, \sigma_w\} under the resource and accounting constraints defined above as \Delta. The social planner’s problem is:

$$S = \max_{\hat{m}, \hat{s}_k, \hat{s}_w} \int_{\mathcal{F}} R(k; s) \hat{m}(k; s) dk ds$$ (13)

The Lagrangian for the Social Planner’s Problem is as follows:

$$\mathcal{L} = \int_{\Omega} R(k; s) \hat{m}(k; s) dk ds + \int_{[k, \bar{k}]} \left\{ \pi(k) \int_{\mathcal{K}} [\sigma_k(k, s) - \hat{m}(k; s)] ds \right\} dk + \int_{[s, \bar{s}]} \left\{ w(s) \int_{\mathcal{K}} [\sigma_s(k, s) - \hat{m}(k; s)] dk \right\} ds + \int_{\Omega} \psi(k, s) \left[ f(k, s) - \hat{\sigma}_k(k, s) - \hat{\sigma}_s(k, s) \right] dk ds + \int_{\Omega} \lambda_m(k, s) \hat{m}(k; s) dk ds + \int_{\Omega} \lambda_{\sigma_k}(k, s) \hat{\sigma}_k(k, s) dk ds + \int_{\Omega} \lambda_{\sigma_s}(k, s) \hat{\sigma}_s(k, s) dk ds$$ (14)

where, abusing notation, \{\pi(k)\}_{k \in [k, \bar{k}]}, \{w(s)\}_{s \in [s, \bar{s}]} are the collection of Lagrangian multipliers for the accounting constraints; \{\psi(k, s)\}_{(k, s) \in \Omega} is the collection of Lagrangian multipliers for the resource constraints, and

$$\{\lambda_m(k, s)\}_{(k, s) \in \Omega}, \{\lambda_{\sigma_k}(k, s)\}_{(k, s) \in \Omega}, \{\lambda_{\sigma_s}(k, s)\}_{(k, s) \in \Omega}$$

are the collection of Lagrangian multipliers for the non-negativity constraints.

The first-order condition with respect to \(\hat{m}(k; s)\) is:

$$R(k; s) - [\pi(k) - w(s)] + \lambda_m(k, s) = 0$$ (15)

If \(m(k; s) > 0\), then \(\lambda_m(k; s) = 0\) due to complementary slackness. This first-order condition is the same as that in the competitive equilibrium.
The first-order conditions with respect to $\sigma_k(k,s), \sigma_s(k,s)$ are:

$$
\pi(k) - \psi(k,s) + \lambda_{\sigma_k}(k,s) = 0 \quad (16)
$$
$$
\omega(s) - \psi(k,s) + \lambda_{\sigma_s}(k,s) = 0 \quad (17)
$$

For a type $(k,s) \in \Omega$ who works in both roles such that $\sigma_k(k,s) > 0, \sigma_s(k,s) > 0$, then $\lambda_{\sigma_k}(k,s) = \lambda_{\sigma_s}(k,s) = 0$. Hence:

$$
\pi(k) = \omega(s) = \psi(k,s) \quad (18)
$$

which is the occupational choice equation in the competitive equilibrium.

Note that $\psi(k,s) = \frac{\partial \mathcal{L}(\cdot)}{\partial f(k,s)}$. Therefore, $\psi(k,s)$ is the social cost of employing an individual of the type $(k,s) \in \Omega$ at the margin. Also, $\psi(k,s) = \max\{\pi(k), \omega(s)\}$. Hence, the social marginal cost, due to occupational choice, is the maximum of the cost of hiring the individual as a key role worker and that as a support role worker.

Some further insights can be obtained by inspecting the dual of the Social Planner’s Problem, which is:

$$
\min_{\pi,\omega,\lambda} \int \lambda(k,s)dF(k,s)
$$

such that

$$
\pi(k) + \omega(s) \geq R(k,s) \quad (19)
$$
$$
\pi(k) \leq \lambda(k,s) \quad (20)
$$
$$
\omega(s) \leq \lambda(k,s) \quad (21)
$$

for all $(k,s)$.

The constraints also imply that

$$
\lambda(k,s) \geq \max\{\pi(k), \omega(s)\}
$$

which in the optimum must be binding because the social planner is minimizing with respect to only $\lambda$ and that $\pi,\omega$ do not directly enter the dual objective function. We
can eliminate $\lambda$ to yield the following equivalent problem:

$$\min_{\pi, w} \int \max\{\pi(k), w(s)\} dF(k, s)$$

such that (19) holds.

The Leontief form of the problem punishes the Social Planner whenever $\pi(k) \neq w(s)$ for any $(k, s)$, weighted by the type distribution $F(k, s)$. Therefore, there is an incentive for the Social Planner to overlap the two occupational earnings distributions together as long as (19) is not being violated.

For reference, the dual objective of Kremer Maskin is:

$$\min_{\pi, w} \int \min\{\pi(q), w(q)\} dF(q)$$

(22)

such that $\pi(q) + w(q') \geq R(q, q')$ for any $q, q'$ where $q$ is a univariate skill. There is no punishment in setting $\pi(q) \neq w(q')$ for any $q \neq q'$. This explains why the occupational earnings distributions tend to be non-overlapping. The Kremer Maskin simulation below illustrates the same point.

## 6 Calibrations and Simulations

The objectives of this section are to:

1. Calibrate the model to 1999 Brazilian data and then simulate the model to show how it matches the 2013 data.

2. Calibrate and simulate a version of Kremer and Maskin. We compare this one factor model to our two factors model.

3. Simulate our model with and without occupational choice as it responds to SBTC.

A summary of the data is given below. All earnings are deflated by the Brazilian consumer price index.

### 6.1 Summary Statistics

Table 1 below provide some summary statistics:
<table>
<thead>
<tr>
<th>Year</th>
<th>Variable</th>
<th>Percentiles (0, 25, 50, 75, 100%)</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>Individual earnings</td>
<td>4.6854, 5.4283, 5.784, 6.4146, 8.5836</td>
<td>log BRL</td>
</tr>
<tr>
<td>1999</td>
<td>Average firm earnings</td>
<td>5.1131, 5.6575, 5.9567, 6.3694, 7.365</td>
<td>log BRL</td>
</tr>
<tr>
<td>1999</td>
<td>Education</td>
<td>1, 6, 9, 12, 15</td>
<td>years</td>
</tr>
<tr>
<td>2013</td>
<td>Individual earnings</td>
<td>4.9457, 5.7813, 6.0755, 6.5491, 8.9069</td>
<td>log BRL</td>
</tr>
<tr>
<td>2013</td>
<td>Average firm earnings</td>
<td>5.7981, 5.9634, 6.1926, 6.4579, 7.4683</td>
<td>log BRL</td>
</tr>
<tr>
<td>2013</td>
<td>Education</td>
<td>1, 9, 12, 12, 15</td>
<td>years</td>
</tr>
</tbody>
</table>

Table 1: Summary Statistics

The main feature of interest in the data is that earnings inequality had decreased between 1999 to 2013. Compared with 1999, the education distribution on 2013 shifted significantly to the right. The median years of schooling went from 9 to 12 years. This shift was due to the expansion of high school education in Brazil. One should expect that earnings inequality in Brazil would have fallen over that period because the supply of cognitive skills in the labor force increased significantly. This results in a more compressed cognitive skill distribution, which implies a more compressed earnings distribution as well.

### 6.2 Main Specification

This subsection section calibrates the model using 1999 Brazilian data.

The base skills distributions, \( c \) and \( r \), are independent. \( c \) is the schooling by years in Brazil, taken from Benguria. In our quantitative exercises, \( r \) is the error term in our model because we do not have any other unobservable. We assume that the non-cognitive skill \( r \) follows a normal distribution which has the same mean and variance as the cognitive (schooling) distribution in the same year.\(^{13}\) How to scale a non-cognitive skill distribution is an open question in psychology and economics. Normality of the non-cognitive distribution is a common assumption.\(^{14}\) We chose to match the first two moments of the cognitive and non-cognitive skills distributions so that any differences in calibrated parameter values between the two skills are due to the data

---

\(^{13}\)In our first draft, we held the non-cognitive skill distribution fixed to the base year. The fit of the model is similar. Thus the choice of a fixed versus variable non-cognitive skill distribution over our short time period is not quantitatively important.

weighting the two skills differently in explaining observed behavior.

See Figure 7 for a plot of the density functions of base skills. The red curve with dot markers shows the density of cognitive skill \( c \) in the year 1999, corresponding to the years of schooling.\(^{15}\) The green curve with triangle markers shows the density of cognitive skill in the year 2013. Since years of schooling are discrete, we smooth the distributions to get continuous cognitive skill distributions. The two curves reveals that the schooling distribution shifts to the right from 1999 to 2013 in Brazil significantly; median years of education went from 9 to 12 years. The blue and purple curves in Figure 7 show the densities of non-cognitive skill \( r \) in 1999 and 2013 respectively.

The two occupations (key role and support role) have the following aggregators:

\[
\begin{align*}
    k_1 &= c_1 \beta_k r_1^{1-\beta_k} \\
    s_2 &= c_2 \beta_s r_2^{1-\beta_s}
\end{align*}
\]

The simulation starts with a \( 50 \times 50 \) square grid for \((c, r)\). Because the aggregation is constant returns to scale with equal support of \( c \) and \( r \), the grid for \((k, s)\) is also a \( 50 \times 50 \) square grid.

\(^{15}\)We interpret cognitive skill as the schooling plus an unobserved component. So we interpolate a discrete schooling distribution by first simulate 10000 draws. For each draw, we add normally distributed errors, then we estimate its kernel density. The variance of the error term is not critical to our results.
The production function in \((k, s)\) is
\[
R(k_1, s_2) = Ak_1^{\alpha_k}s_2^{\alpha_s}
\] 
(25)

So the five parameters of the model consist of \(\{A, \alpha_k, \alpha_s, \beta_k, \beta_s\}\). \(A\) is a scaling parameter; \(\alpha_k, \alpha_s\) control the (marginal) productivity of \(k\) and \(s\); \(\beta_k, \beta_s\) control how cognitive skill \(c\) and non-cognitive skill \(r\) aggregate into role-specific skills \(k\) and \(s\). The key role and support role can be relabelled, such that if \((\alpha_k, \beta_k)\) and \((\alpha_s, \beta_s)\) are swapped, an equivalent model would result. To resolve this labelling issue, we impose \(\beta_k > 0.5 > \beta_s\), such that the key role demands more cognitive skill than non-cognitive skill, and the support role demands more non-cognitive skill than cognitive skill. Hence \(\theta \in \Theta \equiv (\mathbb{R}_+^3 \times [0, 1]^2, \beta_k > 0.5 > \beta_s)\).

The Cobb-Douglas form of the revenue function assumes supermodularity in \((k_1, s_2)\) as the cross-derivative \(R_{ks} = A\alpha_k\alpha_s k_1^{\alpha_k-1} s_2^{\alpha_s-1} > 0\). Convexity of the revenue function is not assumed.

There are three 1999 percentile lines, \(q\), to be matched as calibration targets:

- \(q = ind\): the average individual wage by percentile in 1999;
- \(q = firm\): the firm average wage by individual percentile 1999;
- \(q = edu\): average education in years by individual percentile 1999.

These five parameters are calibrated to the three percentile lines. The calibration is done by minimizing the weighted sum of squares deviations (SSR) between the actual percentile line and the simulated counterparts given \(\theta\), evaluated at 1-100th percentiles. The SSR is defined as follows:

\[
SSR = \sum_{q \in \mathcal{Q}} w_q \sum_{p=1}^{100} (y_{q,sim,p} - y_{q,actual,p})^2
\]

In this expression, \(y_{q,sim,p}\) denotes the \(p\)-th percentile of a simulated percentile line \(q \in \mathcal{Q}\) given \(\theta\), while \(y_{q,actual,p}\) denotes the actual counterpart;

\[
w_q = \frac{1/var(y_{q,actual})}{\sum_{q' \in \mathcal{Q}} (1/var(y_{q,actual}))}
\]

is a inverse variance weight for line \(q \in \mathcal{Q}\). In order to provide a sense of goodness of
fit, we define the explained proportion for line \( q \in \mathcal{Q} \) as:

\[
R_q = 1 - \frac{\sum_{p=1}^{100} (y_{q,\text{sim},p} - y_{q,\text{actual},p})^2}{\sum_{p=1}^{100} (y_{q,\text{actual},p} - \bar{y}_{q,\text{actual}})^2}
\]

As discussed above, the grid for \((k, s)\) is a 50 × 50 square grid that is invariant to \( \theta \). The simulation maps \((k, s)\) to this 50 × 50 grid, so that the minimizing \( \theta \) corresponds to the transformed skill space. The simulation is done by solving the primal (social planner) problem, with the earnings of each simulated individual evaluated using the Lagrangian multipliers of the primal solution.

### 6.3 Calibration results

The calibrated parameters are:

<table>
<thead>
<tr>
<th>( A )</th>
<th>( \alpha_k )</th>
<th>( \alpha_s )</th>
<th>( \beta_k )</th>
<th>( \beta_s )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.104</td>
<td>1.4158</td>
<td>0.597</td>
<td>0.5257</td>
<td>0.0283</td>
</tr>
</tbody>
</table>

The point estimate for \( \alpha_s < 1 \) which means that revenue is not convex in support skills. As discussed earlier, convexity of the revenue function in occupational is not necessary for convexity of occupational earnings. As will be shown below, the density of support skill earnings is single peaked and right skewed.

The overall goodness of fit for our calibration is in column (1) of Table 2 below:

<table>
<thead>
<tr>
<th>( SQR )</th>
<th>( R_{\text{ind}} )</th>
<th>( R_{\text{firm}} )</th>
<th>( R_{\text{edu}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.7036</td>
<td>0.7306</td>
<td>0.8442</td>
<td>0.8968</td>
</tr>
<tr>
<td>7.6539</td>
<td>0.7846</td>
<td>0.8898</td>
<td>0.9088</td>
</tr>
<tr>
<td>12.848</td>
<td>0.7360</td>
<td>0.8002</td>
<td>0.7643</td>
</tr>
</tbody>
</table>

Table 2: Goodness of Fit

Comparing the fit of the three lines, we are least able to fit the distribution of individual earnings. The best fitting line is the distribution of education by earnings. To provide a finer view of fit, the percentile graphs (individual, firm, education), with the simulated plotted along with the actual are given below.
Columns (1) to (3) in Table 3 provide statistics on how we fit earnings inequality in Brazil in 1999.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SD (ln inc)</td>
<td>0.53</td>
<td>0.49</td>
<td>0.08</td>
<td>0.37</td>
<td>0.30</td>
<td>0.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gini</td>
<td>0.31</td>
<td>0.24</td>
<td>0.27</td>
<td>0.22</td>
<td>0.17</td>
<td>0.26</td>
<td>0.27</td>
<td>0.15</td>
</tr>
<tr>
<td>90-10 Ratio</td>
<td>4.24</td>
<td>3.36</td>
<td>0.23</td>
<td>2.57</td>
<td>2.29</td>
<td>0.11</td>
<td>3.86</td>
<td>0.09</td>
</tr>
<tr>
<td>50-10 Ratio</td>
<td>1.84</td>
<td>2.10</td>
<td>-0.13</td>
<td>1.46</td>
<td>1.67</td>
<td>-0.13</td>
<td>2.18</td>
<td>-0.17</td>
</tr>
<tr>
<td>90-50 Ratio</td>
<td>2.30</td>
<td>1.60</td>
<td>0.36</td>
<td>1.75</td>
<td>1.37</td>
<td>0.25</td>
<td>1.77</td>
<td>0.26</td>
</tr>
<tr>
<td>SD (ln inc) within firm</td>
<td>0.30</td>
<td>0.05</td>
<td>1.71</td>
<td>0.29</td>
<td>0.09</td>
<td>1.17</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Fitting Earnings Inequality

Unsurprisingly, the 1999 model underestimates earnings inequality as represented by the Gini coefficient, the 90-10 and 90-50 earnings ratios. It fits the 50-10 earnings ratio relatively well. We also underestimate the standard deviation of within firm earnings. This discrepancy is not unexpected because the number of workers within firms are so different between the data and the model.

Graphically, Figure 8 shows that the misfit of the earnings distribution is after the 75th percentile where the model underestimates the long tail of the actual distribution. We suspect that the main reason for this misfitting is because the observed years of schooling distribution is too crude an approximation of the cognitive skill distribution. The firm line fits better because the average earnings in firms by percentile does not have as long a tail. Finally, education fits the best.

6.4 Occupational choice and matching

Based on our estimates of the parameters of the aggregation functions, Figure 9 presents the joint distribution of \((k, s)\). There is a positive correlation between key role and support role skills.

Figure 10 shows the separating and matching functions, \(\phi(k)\) and \(\mu(k)\) respectively. In general, the separating line is above the matching line. This means that an individual will occupation skills on the matching line strictly prefers to be a key role worker. Also the key role worker will earn more than his or her support role employee.
In our calibration, we restricted the aggregation functions such that the marginal product of cognitive skills is higher for the key role relative to the support role, $\beta_k > \beta_s$. We did not restrict the marginal product of the key role to be higher than that of the support role in the revenue function. Our estimates show $\alpha_k \gg \alpha_s$. However since there is occupational choice, occupational earnings cannot be too different.

Panel a of Figure 11 shows that both occupational earnings densities are single peaked and right skewed.
Figure 9: Joint Distribution of $(k, s)$

Figure 10: Separating and Matching Functions

Figure 11: Occupational Earnings Distributions

(a) Becker-Roy

(b) Kremer-Maskin
6.5 2013

Here, we shift the education and non-cognitive skill distributions to that of 2013, holding all other parameters at the 1999 level. The change in the \((k,s)\) distribution is shown in Figure 9; the separating and matching lines for 2013 are also shown in Figure 10. The overall goodness of fit for our calibration is in Table 2 column 2. Surprisingly, the simulated fit of our 2013 model is better than the 1999 calibrated model. Comparing the fit of the three lines, we are least able to fit the distribution of individual earnings. The best fitting line is the distribution of education by earnings. To provide a finer view of fit, the percentile graphs (individual, firm, education), with the simulated plotted along with the actual are given below.

Columns (4) to (6) in Table 3 provide statistics on how we fit earnings inequality in Brazil in 2013. Again, the model underestimates earnings inequality as represented by the Gini coefficient. But there is a significant improvement in fit in the 90-10 and 90-50 ratios which is due to the fact that actual earnings inequality in 2013 have fallen.
We next show figures for the changes in the earnings distribution by percentile and changes in average firm earnings by percentile (Figure 13). These are declining in both the simulated and actual data.

Figure 14 shows the changes in within firm earnings inequality over time. Largely, there is little change in both actual and simulated within firm earnings inequality over time. There is a significant decline in the change actual within firm earnings inequality after the 75% of earnings. Our simulated model replicates this decline qualitatively but the quantitative magnitude is not as large. Our model provides a potential explanation of this decline. The education distribution shifted to the right
significantly in 2013. This shift change the distribution of occupational skills with much more high skilled key role workers as shown in Figure 3. What this means is that in 2013, a high skilled key role worker will have many more competitors than in 1999. But high skilled support workers did not increase as much in 2013. So the demand by high skilled key role workers for high skilled support partners will drive up the wages of high skilled support workers, significantly lowering within firm inequality in high wage firms in 2013. This relative decrease in within firm earnings inequality will be mitigated by workers with skills on the 1999 separating line who in 2013 will strictly prefer to be support workers.

In summary, our model is able to quantitatively generate the four characteristics of the earnings distributions which we discussed in the introduction:

1. Occupational earnings distributions are single peaked and right skewed.

2. Individual earnings variation, both across and within firms, are due to the different occupations which workers hold.

3. Firm/establishment fixed effects explain a significant fraction of the variance of log earnings.

4. Changes of individual earnings inequality which are primarily due to changes in earnings inequality across and not within firms.

Furthermore, to a first order, our calibrated model fits the three 1999 earnings percentile lines. We also replicate changes in the earnings distribution both within and across firms between 1999 and 2013 using our calibrated model and the observed shift in the educational distribution. It also qualitatively replicated the drop in changes in within firm inequality in high wage firms in Brazil.

There are other features of the earnings distribution which we do not explain such as earnings dynamics within and across firms, the relationship between firm size and earnings, the effects of minimum wages and other labor market distortions. And as with Helpman et al. (2017) and Engbom and Moser (2017) which use the data to study earnings inequality in Brazil, we ignore the informal sector. Thus our model do not provide a full picture of Brazilian earnings inequality.

Rather our goal is to show that occupational choice and matching are quantitatively important determinants of the recent evolution Brazilian earnings inequality.
6.6 Calibrating Kremer Maskin

This subsection calibrates the KM model to 1999 data. The KM model has one less parameter than our model. The calibrated parameters are:

\[
A, \quad \alpha_k, \quad \alpha_s, \quad \beta_k = \beta_s
\]

\[
1.6469, \quad 1.1071, \quad 0.8701, \quad 0.2789
\]

Given the estimates of our larger model, there is nothing surprising about the KM parameter values. We expect the KM model to fit worse and it does as shown in column (3) of Table 2. The sum of squares residual is 32\% higher than in our model. Comparing \(R_{ind}\) in Table 2, our model and the KM model fit comparably well on individual earnings inequality. On the other hand, \(R_{firm}\) and \(R_{edu}\) of Table 2 show that the KM model provide a significantly worse fit for the firm earnings and education earnings relationships in 1999.

In general, KM model can also generate characteristics (2) and (3). There is a difference in its ability to match the changes in within firm inequality. Figure 15 shows that KM is able to replicate the overall lack of change in within firm earnings inequality between the two periods. Unlike our Figure 14, the KM model is unable to replicate the drop in change in within firm earnings inequality among high wage firms. Our explanation is as follows. A change in the educational distribution changes the one factor skill distribution in 2013 which is available to both occupations. Due to
PAM, high skilled key role workers will still match with relatively high skilled support workers. If occupational wages within firm differ too much, workers will switch occupations, independent of the one factor skill distribution. This arbitrage restrict changes in within firm inequality across the two periods. Our two factor Roy model of occupational choice weakens this arbitrage.

There is another difference between KM and our model. Recall that in the KM model, there is one distribution of skill and individuals with different skills will choose different occupations. Figure 16 shows how different individuals with increasing skills choose between key roles and support roles in 1999. The main feature of the occupational choices by skill is that there are many non-overlapping occupational segments: The lowest skill individuals choose the key roles. As skills increase, individuals switch from a key role to a support role. With further increase in skills, individuals switch back to a key role and so on. Neither set of workers, key role or support role, form a connected set. What this means is that both 1999 density plots of occupational earnings have disconnected multiple peaks, a counterfactual prediction of occupational earnings densities in general. See Panel b of Figure 11.

As discussed in earlier papers and here, these multiple peaks occupational earnings densities and lack of convexity of occupational earnings by earnings percentiles are generic features of the KM model and due to its one factor skill structure.

This non-overlapping feature of KM model critically depends on $\beta_k = \beta_s$. In Fig-
ure 17, we reduce $\beta_s$ slightly by 0.025 while holding $\beta_k$ unchanged. The resulting occupational earnings distributions overlaps again.

Figure 17: Occupational Earnings Distributions ($\beta_k = 0.2789, \beta_s = 0.2539$)

The KM model is a classic. It was the first model to recognize the importance of occupational choice and matching in the labor market. Outside modelling hierarchies\(^\text{16}\), the assumption of one dimension of heterogeneity in workers’ skills is too restrictive to make it a reasonable quantitative model of the earnings distribution with multiple occupations. Our contribution is to build on KM and relax its one factor skill structure.

7 Skill biased technical change

The predominant explanation for the recent increase in the US is SBTC. See the survey by Acemoglu and Autor (2011). SBTC is assumed to increase the marginal productivity of college educated workers relative to non-college educated workers. This divergence in productivity translates to divergence in earnings. In the standard model of SBTC, college enrollment is assumed to adjust slowly if at all to the change in earnings inequality over time.

Our model can be used to model the short and long run effect of SBTC. In the short run, there is no occupational choice as SBTC change occurs. Key role workers

\(^{16}\)KM is reasonable if occupations are different levels of hierarchical firms and workers in higher levels of a hierarchy have more of the one dimensional skill (e.g. Caliendo et al. (2015)).
and support workers cannot change occupations. They can change firms in response to SBTC. Since SBTC increases the marginal product of key role workers, at the old wage gradient for support workers, key role workers will want to hire higher skill support workers bidding up the earnings of support workers. This effect will mitigate the increase in earnings inequality in the short run. In the long-run, individuals can switch occupations. Thus we expect the increase in earnings inequality due to SBTC will be even more muted.

To study how SBTC affects earnings inequality in our framework, we use the 2013 simulated model as a benchmark. To model SBTC, we increase $\alpha_k$, which raises the marginal rate of technical substitution $R_k/R_s = (\alpha_k/\alpha_s)(s/k)$, so that the key role becomes relatively more productive than before. We increase the value of $\alpha_k$ from 1.41584 to 1.5.

In the short-run, key role and support workers cannot change occupations in response to SBTC. The short run separating function is the same as before SBTC. Earnings will change in response to SBTC to maintain PAM in the new equilibrium. So short run equilibrium matching in teams is the same as before. Since $R_k(k, \mu(k)) = \pi'(k)$ and $\pi(0) = 0$ after SBTC, the key role earnings schedule rotates upwards. All key role workers earn more than before. Support role workers, whose earnings schedule is governed by $R_s(\mu^{-1}(s), s) = w'(s)$, will also increase. Column (3) of Table 4 below shows the log change in earnings inequality from short run SBTC relative to the 2013 benchmark model. All measures of earnings inequality have increased. Also within firm inequality as measured by the standard deviation of within firm earnings increased by 0.90 log points, i.e. more than doubled compared with the benchmark.

In the long run, there are both occupational choice and matching. The first order effect of allowing for occupational choice is that allocative efficiency in the economy increases. This shows up clearly in Figure 18 where the long run earnings line is mostly above the short run line by wage percentile.

At the highest percentiles, the short run earnings line is higher. This is explained as follows. In the long run, some previously high skill support workers who also have high key role skills will switch to the key role occupation. This increase in high skill key role workers will reduce their earnings compared to the short run. To maintain labor market equilibrium, previously low skill key role workers who also have low support role skills switch to support role occupations. For these previously key role workers to switch in spite of their increased productivity after SBTC, it must be the
case that their support role wages increase significantly. Thus we expect low skilled workers, in both key and support roles, to earn higher wages in the long run compared with the short run.

Unsurprising, comparing with column (1) with that of column (4) show that every measure of earnings inequality have increased relative to the benchmark model. Row (2) of columns (2) and (4) show that the long run standard deviation of earnings fell relative to the short run. The Gini coefficients remain essentially unchanged. Since the Gini is affected by both level and variance changes in the earnings distribution, the Gini and standard deviation do not have to move in the same direction. Row 5 of columns (2) and (4) show that the ratio of earnings between 50th percentile to the 10th percentile fell in the long run relative to the short run. Row 6 of columns (2) and (4) show that the ratio of earnings between 90th percentile to the 50th percentile increased in the long run relative to the short run. So SBTC increased long run earnings inequality between the high skilled workers, key role or support role, and the median skilled worker, and it decreased earnings inequality between the median and low skilled workers compared with the short run.

Row 7 of columns (2) and (4) showed a significant reduction in within firm earnings inequality in the long run. Because workers match by occupational skill rank, a decrease in within firm inequality also means a decrease in occupational wage differences by skill rank. So occupational choice reduced earnings inequality between occupations but does not reduce earnings inequality everywhere when SBTC occurs. Figure 19 shows that there is less changes in within firm inequality by earnings per-

<table>
<thead>
<tr>
<th>Moment</th>
<th>(1) 2013 SBTC Model</th>
<th>(2) Long Run ln(2)_T</th>
<th>(3) ln(2)_T</th>
<th>(4) Short Run ln(2)_T</th>
<th>(5) ln(4)_T</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD (ln inc)</td>
<td>0.30</td>
<td>0.59</td>
<td>0.68</td>
<td>0.51</td>
<td>0.53</td>
</tr>
<tr>
<td>Gini</td>
<td>0.17</td>
<td>0.23</td>
<td>0.30</td>
<td>0.24</td>
<td>0.30</td>
</tr>
<tr>
<td>90-10 Ratio</td>
<td>2.29</td>
<td>3.26</td>
<td>0.35</td>
<td>3.53</td>
<td>0.43</td>
</tr>
<tr>
<td>50-10 Ratio</td>
<td>1.67</td>
<td>2.17</td>
<td>0.26</td>
<td>2.10</td>
<td>0.23</td>
</tr>
<tr>
<td>90-50 Ratio</td>
<td>1.37</td>
<td>1.5</td>
<td>0.10</td>
<td>1.68</td>
<td>0.20</td>
</tr>
<tr>
<td>SD (ln inc) within firm</td>
<td>0.09</td>
<td>0.20</td>
<td>0.80</td>
<td>0.06</td>
<td>-0.40</td>
</tr>
</tbody>
</table>

Table 4: SBTC and Earnings Inequality
Figure 18: SBTC Earnings by Percentile

centile due to SBTC for the long run relative to the short run.
8 Conclusion

This paper integrates Roy’s model of occupational choice with Becker’s model of matching in the labor market. We used it to discuss invariant characteristics of the earnings distribution. The model is parameterized to quantitatively fit the aggregate and between firm earnings inequality in Brazil in 1999. Our simulation of the model for 2013 shows that the large increase in educational attainment between 1999 and 2013 was a first order factor in reducing aggregate Brazilian earnings inequality over that period.

SBTC in the model can also qualitatively rationalize the changes in the US earnings distribution discussed by Song et al. (2015).

Our model is highly stylized, ignoring some important features of labor markets. First, we do not consider variation in firm size. This does not allow us to discuss variation in the quantity of output across firms, an important concern of the standard model of SBTC. Second, we take the underlying skill (education) distribution as exogenous without considering why the Brazilian schooling has shifted. Third, we have a static model and we ignore search frictions in both occupational choice and matching. We leave these important concerns for further research.

There are national level linked employer and employee datasets which include occupational codes for workers. An important extension of this model is to estimate our model using observed occupations.


References


42


