Occupational Choice and Matching in the Labor Market*

Eric Mak¹ and Aloysius Siow²

¹Shanghai University of Finance and Economics

²University of Toronto

November 3, 2016

1 Introduction

Despite differences in the distributions of employers by technologies and worker by skills across industries, regions and time, earnings distributions have some invariant characteristics:

1. In every market economy, there are many occupations. Most occupational earnings distributions are single peaked and right skewed. As a consequence, most empirical economists use log earnings regressions to investigate earnings distributions.
2. 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. \(1999\)](#)).

*We thank Felipe Benguria for providing unpublished figures on the Brazilian data. We also thank Robert McCann, Ronald Wolthoff, Marcin Peski, seminar and conference participants for useful comments. We gratefully acknowledge SSHRC for financial support.

3. 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. \(2010\)](#); [Skans et al. \(2009\)](#) (Sweden).

Characteristic 1 is well known. Figure 1, which is obtained from the US 2015 Current Population Survey, 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. All the earnings distributions have earnings which are weakly convex by percentile.¹ Since the distributions of the demand and supply of skills to an occupation will likely affect the sex composition, size and average earnings, the demand and supply distributions cannot be a first order determinant of convexity. Rather, there must be a common mechanism across occupations which generates convexity in the occupational earnings distributions. A single peaked right skewed distribution such as the log normal earnings distribution will approximate the earnings distributions on Figure 1.

¹The least convex distributions are dominated by non-competitive (government/unionized) employment.

Figure 1: Current Population Survey (United States, 2015)

Characteristic 2 was discovered shortly after economists started estimating 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 (Groschen, 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 and Airbnb in the last 18

months. (NYT Aug 18, 2015)

Characteristic 3 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.² What about changes in across and within firm earnings inequality? In an important recent paper by [Song et al. \(2015\)](#), with Social Security Administration earnings data for more than 100 million workers per year, showed that for 99.8% of the working population, there was no change in within firm earnings inequality from 1982 to 2012. [Benguria \(2015\)](#) showed that for male workers in the formal sector in Brazil, aggregate earnings inequality have fallen significantly from 1999 to 2013. To a first order, there was also no change in within firm earnings inequality. We will describe the findings from both papers in more detail below.³ Since aggregate changes in earnings inequality across countries recently occurred in both directions, how can there be so little change to within firms inequality?

In order to discuss earnings inequality within and across firms, 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 a higher skill worker and not more workers of the same skill. We assume all firms produce the same quantity of output. As such, our model is not a complete model of the labor market.

In our model, there is a distribution of workers, each characterized by his or her cognitive and non-cognitive skills. There are two occupations: key role and support role. The occupational skill for each worker is an exogenous aggregation of the worker's cognitive and non-cognitive skills. The aggregation functions for the two roles are not monotone transforms of each

²Declining earnings inequalities were and are common in Latin America ([Lustig et al. \(2013\)](#)). For Spain, see [Pijoan-Mas and Sánchez-Marcos \(2010\)](#).

³In the face of large increases in aggregate earnings inequality, [Faggio et al. \(2010\)](#) also showed that there was little change in within firm inequality in the UK from 1984 to 1999. For Sweden where the wage setting institutions are significantly different than the US, [Skans et al. \(2009\)](#) concludes that “the trend in between-plant variance makes up the entire increase in wage dispersion over the period (i.e. 1985-2000)”.

other. Although each individual has two skills, each of their occupational skill is unidimensional. We start the analysis of equilibrium behavior with the bivariate distribution of occupational skills because the aggregation of individual skills into occupational skills is exogenous.

Workers work in firms/teams of two, one in a key role and one in a support role. Since we fix every team size at two, we interpret differences in revenue produced across teams as due to quality, and not quantity, differences in output produced across teams. There is free entry of teams which implies that the revenue of a team is divided between its workers. Each worker chooses the occupation which maximizes his or her earnings. Given their occupational choices, they also choose which team to work for.

This model integrates the Roy model of occupational choice with Becker's model of frictionless positive assortative matching (PAM). In a Roy model with two occupations, workers, each with two occupational skills, choose which occupation to enter based on maximizing their earnings. In Becker's model of PAM, market participants are exogenously divided into two sides of the market. Within each side of the market, agents are ranked by their occupational skill. An agent on each side matches with an agent on the other side to produce output which is divided between them. When team revenue is supermodular in occupational skills and the market for matches is frictionless, PAM by abilities between agents on opposite sides of the market obtains. Assuming everyone wants to be matched and cannot choose occupations, agents on the long side of the market will be unmatched in equilibrium. When we integrate Roy with Becker, workers who are potentially unmatched in one occupation will switch occupations. So the equilibrium populations of workers in both occupations are the same. Due to PAM in teams, this model rationalizes establishment/firm earnings effects.

We will also assume that each team's revenue function is convex in occupational skills. Convexity of the team's revenue function will imply that earnings for each occupation is convex in occupational skills. As Adam Smith pointed out long ago, workers are specialized by occupations. From a labor market point of view, what is the nature of wages which workers

observe such that they will willingly specialize? If occupational wages are convex in occupational skills, workers will want to specialize in skills investments. Microfoundations for such specialization using indivisibilities and increasing returns are provided by [Rosen \(1983\)](#). Our reduced form revenue function assumes convexity in occupational skills. Since differences in team revenues are due to differences in the quality of team output, we are assuming that team quality (measured in money) is convex occupational skills. E.g. we are assuming that a better cake is not produced by adding more bakers but by hiring a better baker.

Occupational earnings distributions which are single peaked and right skewed will also depend on the distribution of occupational skills. Our simulation section using Brazilian data shows that the occupational earnings distributions are single peaked and right skewed.

SBTC is the preferred explanation of the recent increase in earnings inequality in the US. What happens when there is skill biased technical change in the model? When key role workers exogenously become more productive, more workers will want to become key role workers leading to a scarcity of support workers. Wages for support workers have to increase when workers from both occupations are necessary for production. PAM in the labor market also means that at the old wage gradient for support workers, enhanced key role workers want to match with better qualified support workers than before. This increased competition for better support team mates will lead to an increase in earnings inequality across support workers. The two effects will mitigate an increase in within teams earnings inequality.

Finally, we use our framework to study the recent evolution of the distributions of earnings in Brazil. The average years of educational attainment doubled in Brazil from 1999 to 2013. Using educational attainment as a proxy for cognitive skill, we first estimate 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. We can rationalize a significant share of the recent decline in Brazilian earnings inequality

from her increase in educational attainment. As a test of external validity, we show that our model can reproduce the schooling earnings relationship at both the individual and firm level in Brazil.

This paper builds on previous work on frictionless occupation choice and matching in the labor market. We discuss our debt to the literature in the next section.

2 Literature Review

2.1 Empirical Studies

This section reviews the above two papers which pertain to our work.⁴ 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 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”.

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

⁴For the US, [Barth et al. \(2014\)](#) has similar results by establishments.

earnings and a within firm variance of log earnings:

$$var(w_t^{ij}) = var(\bar{w}_t^j) + \sum_{j=1}^{J_t} P_t^j \times var(w_t^{ij} | i \in j)$$

J_t : number of firms in year t

P_t^j : j 's share of employment in year t

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.

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 2: U.S. Trends of Total, Between and Within Inequality

Figure 3: Brazilian Trends of Total Inequality

Figure 4: Brazilian Between versus Within Firm Inequality

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.

For a finer decomposition within a year, the authors use another decomposition of earnings inequality for year t . Let W_{pt}^i be the mean of w_t^{ij} of all workers in the p' th percentile in the earnings distribution in year t . Let $\bar{W}_{pt}^j$ be the mean of $\bar{w}_t^j$ for each worker in the p' th percentile. Then:

$$W_{pt}^i = \bar{W}_{pt}^j + (W_{pt}^i - \bar{W}_{pt}^j)$$

W_{pt}^i is decomposed into a part which is the mean of the firms in which these workers work in, $\bar{W}_{pt}^j$, and a residual, $(W_{pt}^i - \bar{W}_{pt}^j)$, 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:

$$W_{pt'}^i - W_{pt}^i = \bar{W}_{pt'}^j - \bar{W}_{pt}^j + (W_{pt'}^i - \bar{W}_{pt'}^j) - (W_{pt}^i - \bar{W}_{pt}^j)$$

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.⁵ 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

⁵Their paper showed that there was a large change in within firm inequality at the 99.8 percentile.

Note: Sample contains workers in firms with 20+ full-time equivalent employees.

Figure 5: Percentile Decomposition of US Wage Inequality

Figure 6: Percentile Decomposition of Brazil Wage Inequality

1999 to 2013 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.

We have only discussed the parts of the above papers which are relevant for our purpose. There is a larger literature which uses the AKM methodology to quantify different mechanisms for explaining changes in earnings inequality. There are two reasons why it is difficult to use our model to discuss AKM based decompositions. First, the AKM methodology requires a panel data of workers and firms, and inter firm mobility by workers. Our model is static and has no inter-firm mobility. Extending our model to include dynamics and interfirm mobility is left for future research. Second, AKM decomposes the residual of individual earnings into a firm, individual, match and other effects. We focus on own and co-worker effects.

2.2 Theoretical Studies

Our model builds on classics in labor economics. Adam Smith already emphasized specialization and the division of labor which are fundamental to what is discussed here. Ricardo first recognized that individuals may have more than one dimension of skills and argued that occupational choice should be based on comparative advantage. Roy (1951) sparked the analytic literature on occupational choice. French and Taber (2011) has a recent survey on Roy models.

Building on Katz and Murphy, Acemoglu and Autor provides a survey of the standard model of SBTC and changes in earnings inequality. These

models assume that firm output satisfies constant returns to scale in occupational skills. One skill worker is perfectly substitutable to two lesser skill workers in the same occupation. So labor market equilibrium is determined by an aggregate production function and the aggregate supply of skills to each occupation. By construction, the standard model is silent on earnings inequality between and within firms, and output quality differences across firms. We differ from them by ignoring variation in output across firms and focus only on quality differences.

Our revenue function assumes convexity in occupational skills. Micro-foundations for such specialization using indivisibilities and increasing returns are provided by Rosen (1983). Also see Yang and Borland (1991). The standard model assumes that firm revenue is proportional to occupational skills. This is an important analytic difference between the two models. In order to motivate our different returns to scale assumption, we posit that different firms produce different quality goods whereas the standard model posits that different firms produce different quantity of output. Integrating the two models is left for future research. McCann, et. al. (2015) provides a start.

Becker (1973, 1974) started the modern literature on frictionless matching by studying marriage matching. In his model,

1. Absolute Advantage: men and women are separately ordered by a one-dimensional index of ability.
2. The marital output production function is supermodular (complementary) in spousal abilities.

Under these two assumptions, there is positive assortative matching (PAM) in spousal abilities in equilibrium. Eeckhout and Kircher (2012) provides a state of the art summary and application to labor markets. PAM results in team fixed effects in earnings which are invariant to differences in technology and the distributions of workers' abilities. Most matching models of the labor market study one dimensional matching. Lindenlaub (2015) is an exception. The optimal transport literature in mathematics

also studies matching models. Galichon (2015) provides an accessible survey of this literature for economists.

Except for McCann and Trokhimtchouk (2010) and McCann, et. al. (2015), most previous work on occupational choice and matching build on Kramer and Maskin (1993) in which workers differ by one dimension of heterogeneity. In one dimensional models, there is no comparative advantage in occupational choice. So both occupational choice and matching are based on absolute advantage which makes general characterization difficult to come by (E.g. Tommaso (2016)). Garicano and Rossi-Hansberg (2004) and Lucas (19xx) provide elegant behaviorally motivated special cases. Erlinger, et. al. (2015) also studies a one dimensional occupational choice and matching model with investment and two labor market sectors. Geeroff showed that a Garicano model with occupational choice and matching generates an earnings distribution with a single peak and right skewness. His techniques may also be useful in our context.

Our model has two dimensions of skills. For each worker, the two skills aggregate into two unidimensional occupational skills. There is PAM by occupational skills. With two skills, occupational choice is based on comparative advantage and PAM by occupational skills is due to absolute advantage.

McCann and Trokhimtchouk (2010) has a general multidimensional skills model of occupational choice and matching. Our model is a special case of their framework and we owe our existence and uniqueness results to them. We differ from them by reducing our two dimensional skills problem into one dimensional occupational skill indices problem with which we provide sharper characterizations of equilibrium. For e.g., we work with separating and matching functions which do not always obtain in their setup.

McCann, et. al. (2015) study a model of schooling investment, occupational choice and matching where workers differ by cognitive and communication skills. That model is richer in terms of behavior because there is also investment, multi-sector considerations and heterogeneity in firm size. This paper builds on that work. Our environment here is simpler

which leads to more transparent analytic results. McCann, et. al. provides a way to model variations in the quantities and qualities of firm output. Building on McCann, et. al., Melynk and Turner (2016) estimates a model of occupational choice, time use, matching in both labor and marriage market and where individuals differ by cognitive and communication skills.

Gola (2016) studies a two dimensional occupational choice problem similar to ours. He differs by studying the matching of firms to workers in each separate occupation. He derives comparative statics results for changes in the distribution of skills and/or the revenue functions which may be useful here.

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 $[\underline{c}, \bar{c}] \times [\underline{r}, \bar{r}]$. 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 worker. Consider a team with a key role worker with characteristics (c_1, r_1) and a support 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 assumptions 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 in teams in a frictionless labor market of the type that we study.

Assumption 2 (Increasing Returns). R is strictly convex in k_1, s_2 .

Recall that our assumption is that the revenue of a team increases as the quality of output increases. I.e. higher skill workers generate higher quality output which results in increased revenue. So the convexity assumption is that quality is a convex function of occupational skills. This embodies two assumptions. First, we are assuming that higher quality is not due to adding more workers to the production process. Rather, the team is hiring higher quality workers. As noted earlier, we hire better bakers if we want to bake a better cake. The second assumption is that 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. As noted earlier, Rosen, Yang and Borland, and others have provided microfoundations for this convexity assumption. We will show below that the convexity of the production function in occupational skills will result in occupational wages which are convex in occupational skills.

Both the PAM result and wages being convex in occupational skills will obtain without strong restrictions on the base skills distribution.

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 $\Omega \equiv [\underline{k}, \bar{k}] \times [\underline{s}, \bar{s}]$. 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. Workers choose the occupation which will maximize their net earnings. So a worker of type (k, s) will earn:

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

If $\pi(k) > w(s)$, the worker will be a key role worker. If $\pi(k) < w(s)$, the worker will be a support worker. If $\pi(k) = w(s)$, the worker will be indifferent between the two roles. Therefore, the type space Ω 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 π, w are continuous, strictly increasing functions in their respective arguments (which we shall justify later), Ω_{ks} is an upward sloping line in Ω 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 $\phi : [\underline{k}, \bar{k}] \rightarrow [\underline{s}, \bar{s}]$ such that

$$\phi(k) = \min\{w^{-1}(\pi(k)), \bar{s}\}$$

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

$$w(\phi(k)) = \pi(k) \quad (2)$$

While if $\phi(k) = \bar{s}$, then a worker with skill k will always prefer the key role regardless of his s . In the discussion below and in the simulation, $\phi(k) < \bar{s}$ 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:

$$H(k) = \int_{\underline{k}}^k \int_{\underline{s}}^{\phi(u)} f(u, v) dv du \quad (3)$$

The cumulative distribution of support role workers from ability $\underline{s}$ to s is

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

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:

$$\max_{\tilde{k}, \tilde{s}} R(\tilde{k} : \tilde{s}) - \pi(\tilde{k}) - w(\tilde{s}) \quad (5)$$

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

$$R_k(k; s) = \pi'(k) \quad (6)$$

$$R_s(k; s) = w'(s) \quad (7)$$

We can invert either (6) or (7) to get the matching function, $s = \mu(k)$, where a key role worker of skill k will employ a worker of type s . The matching function is a central concept in Becker's matching model and here as well.

Since we have assumed that π, w is convex, we have Becker's famous PAM result:

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 such that $(k_A; s_A), (k_B; s_B)$ both satisfy the first-order conditions. Due to free entry of the entrepreneur,

$$R(k_A; s_A) + R(k_B; s_B) - \pi(k_A) - w(s_A) - \pi(k_B) - w(s_B) = 0 \quad (8)$$

Suppose that an rearrangement $(k_A; s_B), (k_B; s_A)$ also satisfies the first-

order condition. Then

$$R(k_A; s_B) + R(k_B; s_A) - \pi(k_A) - w(s_B) - \pi(k_B) - w(s_A) = 0 \quad (9)$$

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 . $\square$

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 will now provide a link between the separating function, ϕ of occupational choice and the matching function μ . Differentiating (2) with respect to k ,

$$w'(\phi(k))\phi'(k) = \pi'(k) \quad (10)$$

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

$$\phi'(k) = \frac{R_k(k; \mu(k))}{R_s(\mu^{-1}(\phi(k)); \phi(k))} \quad (11)$$

Equation (11) provides a restriction between the separating function and matching function that depends on the technology. $\phi'(k)$ is the slope of the separating line, representing how 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.

Notably, the two marginal products are different in the equilibrium because $\mu \neq \phi$.

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, π , a separating function, ϕ , and a matching function, μ , 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:*

$$H(k) = G(\mu(k)), \forall k \quad (12)$$

Equation (12) says that for every k , the mass of key role workers up to skill κ must be equal to the mass of support role workers up to skill $\mu(\kappa)$.

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, $\pi(\kappa)$, a separating function, $\phi(\kappa)$, and a matching function, $\mu(\kappa)$, exists.*

3.2 Characterizations

We offer the characterization of the competitive equilibrium as follows.

Proposition 3 (Identical Outside Options for the Worst Match). *$\phi(0) = \mu(0) = 0$, such that the worst type $(0; 0)$ self-matches and splits the output evenly, so that $\pi(0) = w(0) = 0$, and the earnings inequality within this team is zero.*

Proof. If $\phi(0) = 0$, the type $t_1 \equiv (0; 0)$ self-matches. Then $\mu(0) = 0, \pi(0) = w(0) = R(0, 0)/2 = 0$, and our proposition holds.

Now suppose not, and consider the case where $\phi(0) = s^* > 0$. Then $(0,0) \in \Omega_k$, such that this type works exclusively as key role workers.

Given that ϕ is strictly increasing, the only type in $\Omega_{ks} \cup \Omega_s$ in which $s = s^*$ is $(0, s^*)$. Hence $(0,0)$ will be matched to $(0, s^*)$ under PAM. So $\mu(0) = s^*$.

The type $(0, s^*)$ is indifferent to being a key role worker and a support role worker. Hence $\pi(0) = w(s^*)$. Whereas the type $(0,0)$ strictly prefers the key role. Hence $\pi(0) > w(0)$.

Since w is strictly increasing, $w(0) < w(s^*)$, resulting in a contradiction. $\square$

Convexity of the revenue function in occupational skills implies that occupational earnings are convex in occupational skills:

Proposition 4 (Convex Earning Schudules). *$w(s)$ 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, $w''(s) > 0$. $\square$

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 $\phi(k) > \mu(k)$. $\square$

$\phi(k) > \mu(k)$ implies that the worker type $(k, \mu(k)) \in \Omega_k$, so that this type of worker 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.

4 Social Planner's Problem

Readers who are primarily interested in the calibration and simulation results with respect to the Brazilian data may skip ahead to the next section.

Our model of frictionless occupational choice and matching is equivalent to a social planner choosing 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 and T'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 a competitive model. This is how we produce the calibration and simulation results of the next section.

Let $m : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$ be a density function such that $m(k_1; s_2)$ states the mass of team $(k_1; s_2)$, and let $\sigma_k, s_r : \mathcal{T} \rightarrow \mathbb{R}_+$ be density functions such that $\sigma_k(t), s_r(t)$ record the mass of agents of type t working in key role and support role respectively.

We have two accounting constraints:

$$\text{(Accounting Constraint for Key Role): } \int_{\tilde{s}} m(k; \tilde{s}) d\tilde{s} = \int_{\tilde{s} \in \mathbb{R}_+} \sigma_k(k, \tilde{s}) d\tilde{s}, \forall k \in \mathcal{K}$$

$$\text{(Accounting Constraint for Support Role): } \int_{\tilde{k}} m(\tilde{k}; s) d\tilde{k} = \int_{\tilde{k} \in \mathbb{R}_+} \sigma_k(\tilde{k}, s) d\tilde{k}, \forall s \in \mathcal{S}$$

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:

$$\sigma_k(t) + \sigma_r(t) = f(t) \quad \forall t \in \mathcal{T}$$

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 Ω . The social planner's problem is:

$$S = \max_{\{\hat{m}, \hat{\sigma}_k, \hat{\sigma}_w\} \in \Omega} \int_{\mathcal{T}} R(k : s) \hat{m}(k : s) dk ds \quad (13)$$

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

$$\begin{aligned} \mathcal{L} = & \int_{\mathcal{T}} R(k; s) \hat{m}(k : s) dk ds \\ & + \int_{\mathcal{K}} \left\{ \pi(k) \int_{\mathcal{S}} [\sigma_k(k, s) - \hat{m}(k; s)] ds \right\} dk \\ & + \int_{\mathcal{S}} \left\{ w(s) \int_{\mathcal{K}} [\sigma_s(k, s) - \hat{m}(k; s)] dk \right\} ds \\ & + \int_{\mathcal{T}} \psi(k, s) [f(k, s) - \hat{\sigma}_k(k, s) - \hat{\sigma}_s(k; s)] dk ds \\ & + \int_{\mathcal{T}} \lambda_m(k, s) \hat{m}(k; s) dk ds \\ & + \int_{\mathcal{T}} \lambda_{\sigma_k}(k, s) \hat{\sigma}_k(k, s) dk ds + \int_{\mathcal{T}} \lambda_{\sigma_s}(k, s) \hat{\sigma}_s(k, s) dk ds \end{aligned} \quad (14)$$

where, abusing notation, $\{\pi(k)\}_{k \in \mathbb{R}_+}, \{w(s)\}_{s \in \mathbb{R}_+}$ are the collection of Lagrangian multipliers for the accounting constraints; $\{\psi(k, s)\}_{(k, s) \in \mathbb{R}_+^2}$ is the collection of Lagrangian multipliers for the resource constraints, and

$$\{\lambda_m(k, s)\}_{(k, s) \in \mathbb{R}_+^2}, \{\lambda_{\sigma_k}(k, s)\}_{(k, s) \in \mathbb{R}_+^2}, \{\lambda_{\sigma_s}(k, s)\}_{(k, s) \in \mathbb{R}_+^2}$$

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

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

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

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:

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

$$w(s) - \psi(k, s) + \lambda_{\sigma_s}(k, s) = 0 \quad (17)$$

For a type $(k, s) \in \mathcal{T}$ 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:

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

which is the occupational choice equation in the competitive equilibrium.

Note that $\psi(k, s) = \partial \mathcal{L}(\cdot) / \partial f(k, s)$. Therefore, $\psi(k, s)$ is the social cost of employing an individual of the type $(k, s) \in \mathcal{T}$ at the margin. Also, $\psi(k, s) = \max\{\pi(k), w(s)\}$. 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.

5 Calibrations and Simulations

5.1 Main Specification

This section presents a calibration of the model using 1999 Brazilian data. We also simulate the model to see how it matches the 2013 data.

The base skills distributions, c and r , are independent. c is the schooling by years in Brazil, taken from Benguria. Non-cognitive skill r follows a symmetric truncated normal distribution (at 3 standard deviations) with the same support as the schooling distribution, representing a generic uni-modal distribution.

See Figure 7 for a plot of the density functions of base skills. The curve with dot markers shows the density of cognitive skill c in the year 1999, corresponding to the years of schooling. The curve with triangle markers

shows the density of cognitive skill in the year 2013. The two curves reveals that the schooling distribution shifts to the right from 1999 to 2013 in Brazil significantly; the average schooling almost doubled. The red curve in Figure 7 shows the density of non-cognitive skill r , which is assumed to be invariant across the two years 1999 and 2013.

Figure 7: Base Skill Distributions

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

$$k_1 = c_1^{\beta_k} r_1^{1-\beta_k} \quad (19)$$

$$s_2 = c_2^{\beta_s} r_2^{1-\beta_s} \quad (20)$$

The revenue function is

$$R(k_1, s_2) = A k_1^{\alpha_k} s_2^{\alpha_s} \quad (21)$$

So the five parameters of the model consist of $\{A, \alpha_k, \alpha_s, \beta_k, \beta_s\}$. A is a scaling parameter; α_k, α_s control the (marginal) productivity/revenue of k

and s ; β_k, β_s control how cognitive skill c and non-cognitive skill r aggregate into role-specific skills k and s .

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

The simulation starts with a 50×50 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 50×50 square grid. This invariance with respect to a change in β_k, β_s is convenient for our simulations and is without loss of generality.⁶

The roles k, s can be relabelled, such that if (α_k, β_k) and (α_s, β_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.

The parameters of the model are calibrated to 1999 individual and between-firm inequality. Specifically, the sum of squares deviation, evaluated at each percentile between the simulated curves and the actual corresponding curves from Benguria, is minimized by a Nelder-Mead optimizer.⁷

The main specification simulates the model using the following calibrated parameters: $A = 2.22$, $\alpha_k = 1.7438$, $\alpha_s = 0.998$, $\beta_k = 0.827$, $\beta_s = 0.4126$. The first panel of Figure 8 shows how well our calibrated model fits

⁶This parameterization derives from a general Cobb-Douglas production function of (c_1, r_1, c_2, r_2) without loss of generality. To see this, consider a production function

$$R(c_1, r_1; c_2, r_2) = A c_1^{\gamma_{11}} r_1^{\gamma_{12}} c_2^{\gamma_{21}} r_2^{\gamma_{22}} \quad (22)$$

where $\gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22} > 0$. This production function can be transformed as

$$R(c_1, r_1; c_2, r_2) = A (c_1^{\frac{\gamma_{11}}{\gamma_{11}+\gamma_{12}}} r_1^{\frac{\gamma_{12}}{\gamma_{11}+\gamma_{12}}})^{\gamma_{11}+\gamma_{12}} (c_2^{\frac{\gamma_{21}}{\gamma_{21}+\gamma_{22}}} r_2^{\frac{\gamma_{22}}{\gamma_{21}+\gamma_{22}}})^{\gamma_{21}+\gamma_{22}}$$

Letting $\beta_k = \gamma_{11}/(\gamma_{11} + \gamma_{12})$, $\alpha_k = \gamma_{11} + \gamma_{12}$, $\beta_s = \gamma_{21}/(\gamma_{21} + \gamma_{22})$, $\alpha_s = \gamma_{21} + \gamma_{22}$ yields our parameterization.

⁷In the actual implementation, we transform the parameters as follows: A kept in levels; $\alpha_k = \exp(a_k)$, $\alpha_s = \exp(a_s)$ where $a_k, a_s \in \mathbb{R}$; $\beta_k = 0.5 + 0.5\Phi(b_k)$ where $b_k \in \mathbb{R}$, $\beta_s = 0.5 + 0.5\Phi(b_s)$ where $b_s \in \mathbb{R}$. This transformation yields an objective function unconstrained in the parameters, while imposing $\beta_k > 0.5 > \beta_s$.

the 1999 data. The deviations are mostly at the top and bottom percentiles. Quantitatively, the R-squared fits for individuals and firms are 0.7856 and 0.8751.

For 2013, we hold the calibrated parameters constant, only allowing the schooling distribution to shift to its 2013 values. This exercise tests whether this distributional shift alone can produce the observed changes in inequalities. The second panel of Figure 8 shows how well our calibrated model predicts the 2013 data. Qualitatively, the simulated data replicates the slopes of individual and firm quantiles well, despite there is a misfit in levels. As is apparent from the figures, the R-squared fits for individuals and firms are 0.457 and 0.2291 which are worse than for 1999. This is not surprising because the parameters of the model were chosen to match the 1999 figure.

The graphs of simulated matching function μ and separating function ϕ are plotted in Figure 9 for 1999 and 2013. For both years, μ is strictly increasing in k , implying positive assortative matching. ϕ is strictly increasing in k as well. They are in line with our theoretical results. Across the two periods, μ shifted to the right less than ϕ . For any k , the average before-after difference in $\mu(k)$ is about 1.36 units of s (relative to a grid of 50 units). Whereas the average before-after difference in $\phi(k)$ is 2.51, which is larger.

Figure 8: Goodness of Fit Plots

Figure 9: Plots of Matching and Separating Functions in (k, s) space

Next we examine the simulated earnings functions π and w . Figure 10 shows their corresponding plots. In the first panel of Figure 10, the horizontal axis is skill level (k for key role, s for support role). Both $\pi(k)$ and $w(s)$ are strictly increasing and convex with respect to their arguments. In the second panel of Figure 10, we plot π, w by rank instead. Due to positive assortative matching, a key role worker and support role worker would match if and only if their respective ranks are equal. The second panel shows that a key role worker at any rank is earning more than his support role partner. Across the two periods, the difference in earnings diminishes.

Figure 10: Earnings Schedule

The earnings functions, being convex in their respective skills, do not necessarily lead to convex earning distributions which also depend on the underlying distribution of skills. Hence we plot the earnings distributions in Figure 11. The upper panel of Figure 11 shows that for both key role and support role, the earnings distributions are skewed to the right. Aggregating across the two roles, the earnings distributions for 1999 and 2013 are both shown in the lower panel of Figure 11. The earnings distribution in 2013 is less skewed than that of 1999, corresponding to a decreased individual inequality.

Figure 11: Earnings Distributions

Given our goodness of fit results, it is not surprising that we can also replicate the changes in earnings inequality across individuals and firms between 1999 and 2013. The simulated changes in inequality is shown in Figure 12 which largely resembles the actual data (Figure 6).

Figure 12: Percentile Plot of Change (2013-1999)

As discussed by Benguria and reviewed here, the Brazilian labor market changed significantly from 1999 to 2013. In particular, there was a marked decline in individual earnings inequality.

We have shown that a five parameter model of the Brazilian labor market can fit individual and across firm earnings inequality in 1999. The only observed heterogeneity of this model is the educational distribution. Based only on a shift in the educational distribution, we can, to a first order, replicate the changes in individual and across firms earnings inequality from 1999 to 2013. In spite of large changes in the distribution of skills and distribution of earnings, our model can also generate the lack of change in within firm earnings inequality observed between the two periods.

6 Educational decompositions

6.1 Firm Decomposition of Education

Since the relationship between education and earnings is central to our explanation of the Brazilian experience, we provide some evidence on the actual relationship as well as our modeled relationship.

Let $c_t(i)$ be the years of education (cognitive skill) of worker i in year t . c_{pt} is the mean education of all workers in the percentile p in the earnings distribution in year t . Let worker i in wage percentile p work in firm $j(i)$. c_t^j is the mean education of workers in a particular firm j . Then we have the following decomposition of education for any given percentile p :

$$c_{pt} = c_{pt}^j + (c_{pt} - c_{pt}^j) \quad (23)$$

c_{pt}^j in (23) associates for each individual i his firm j and thus its firm education average c_t^j , averaged over all individuals in the percentile p .

Figures (13) and (14) show the plots of c_{pt} , c_{pt}^j and $(c_{pt} - c_{pt}^j)$ for Brazil in 1990 and 2013 respectively. c_{pt} is our non-parametric version of the Mincer schooling earnings relationship. In both years, c_{pt} is quantitatively and qualitatively close to c_{pt}^j . So the line $(c_{pt} - c_{pt}^j)$, to a first order, is zero. Figure (15) shows these changes between 1990 and 2013. Due to the increase in schooling attainment, $c_{p2013} - c_{p1999}$ and $c_{p2013}^j - c_{p1999}^j$ are downward sloping. The time difference in $(c_{pt} - c_{pt}^j)$ is essentially zero for

FIGURE 2: DECOMPOSITION OF MEAN YEARS OF EDUCATION OF WORKERS IN PTH PERCENTILE INTO MEAN YEARS OF EDUCATION FOR FIRMS OF WORKERS IN PTH PERCENTILE AND WITHIN-PERCENTILE DEVIATION; 1999

Figure 13: Education Decomposition 1999

FIGURE 4: DECOMPOSITION OF MEAN YEARS OF EDUCATION OF WORKERS IN PTH PERCENTILE INTO MEAN YEARS OF EDUCATION FOR FIRMS OF WORKERS IN PTH PERCENTILE AND WITHIN-PERCENTILE DEVIATION; 2013

Figure 14: Education Decomposition 2013

FIGURE 6: DECOMPOSITION OF MEAN YEARS OF EDUCATION OF WORKERS IN PTH PERCENTILE INTO MEAN YEARS OF EDUCATION FOR FIRMS OF WORKERS IN PTH PERCENTILE AND WITHIN-PERCENTILE DEVIATION; DIFFERENCE 1999-2013

Figure 15: Time change in Education Decomposition

most wage percentiles.

Figure (16) shows the same figures for our simulated model. Focusing on the last graph, which plots the time changes in between 1990 and 2013, $c_{p2013} - c_{p1999}$ and $c_{p2013}^j - c_{p1999}^j$ are downward sloping. Due to numerical problems, the differences in $(c_{pt} - c_{pt}^j)$ is noisy but centered around zero.

We did not use any actual educational earnings relationship/target to estimate the parameters of our model. Thus our simulated educational earnings relationships provides some external validity for our explanation of the recent decline in earnings inequality in Brazil.

Figure 16: Firm Decomposition of Education

References

- ABOWD, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): “High wage workers and high wage firms,” *Econometrica*, 67, 251–333.
- BARTH, E., A. BRYSON, J. C. DAVIS, AND R. FREEMAN (2014): “It’s where you work: Increases in earnings dispersion across establishments and individuals in the US,” Tech. rep., National Bureau of Economic Research.
- BENGURIA, F. (2015): “Inequality Between and Within Firms: Evidence from Brazil,” *Available at SSRN 2694693*.
- CARD, D., J. HEINING, AND P. KLINE (2013): “Workplace Heterogeneity and the Rise of West German Wage Inequality,” *The Quarterly Journal of Economics*, 128, 967–1015.
- FAGGIO, G., K. G. SALVANES, AND J. VAN REENEN (2010): “The evolution of inequality in productivity and wages: panel data evidence,” *Industrial and Corporate Change*, 19, 1919–1951.
- GROSHEN, E. L. (1991): “Sources of intra-industry wage dispersion: How much do employers matter?” *The Quarterly Journal of Economics*, 869–884.
- LUSTIG, N., L. F. LOPEZ-CALVA, AND E. ORTIZ-JUAREZ (2013): “Declining inequality in Latin America in the 2000s: the cases of Argentina, Brazil, and Mexico,” *World Development*, 44, 129–141.
- MCCANN, R. J., X. SHI, A. SIOW, AND R. WOLTHOFF (2015): “Becker meets ricardo: Multisector matching with communication and cognitive skills,” *Journal of Law, Economics, and Organization*, ewv002.
- PIJOAN-MAS, J. AND V. SÁNCHEZ-MARCOS (2010): “Spain is different: Falling trends of inequality,” *Review of Economic Dynamics*, 13, 154–178.
- ROSEN, S. (1983): “Specialization and human capital,” *Journal of Labor Economics*, 43–49.

- SKANS, O. N., P.-A. EDIN, AND B. HOLMLUND (2009): “Wage dispersion between and within plants: Sweden 1985-2000,” in *The structure of wages: An international comparison*, University of Chicago Press, 217–260.
- SONG, J., D. J. PRICE, F. GUVENEN, N. BLOOM, AND T. VON WACHTER (2015): “Firming up Inequality,” Tech. rep., National Bureau of Economic Research.