Occupational choice and matching in the labor market

Eric Mak
Shanghai University of Finance and Economics

Aloysius Siow
University of Toronto

June 2016 (preliminary)

Three invariant features of earnings distributions

Despite differences in the distributions of employers by technologies and workers by skills across regions, industries and time,

- **Old:** There are many occupations. Occupational earnings distributions are single peaked and right skewed. Labor economists everywhere run log earnings regressions.
- **Recent:** There are firm/establishment fixed effects in log earnings regressions (E.g. Groshen, AKM). How do we interpret these effects?

  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)

- **New:** Recent changes of earnings inequality in many countries, either increasing or decreasing, are primarily due to changes 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).

Earnings decompositions

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 $\overline{w}_t^j$ respectively.

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

$$J_t : \text{ number of firms in year } t$$

$$P_t^j : \ j's \text{ share of employment in year } t$$

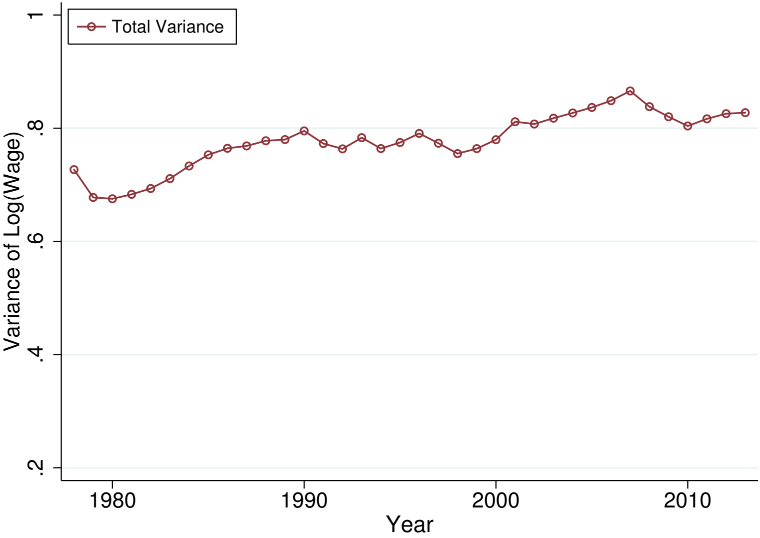
Let W^i_{pt} be the mean of w^{ij}_t of all workers in the $p'th$ percentile in the earnings distribution in year t. Let $\overline{W}^j_{pt}$ be the mean of $\overline{w}^j_t$ for each worker in the $p'th$ percentile. Then:

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

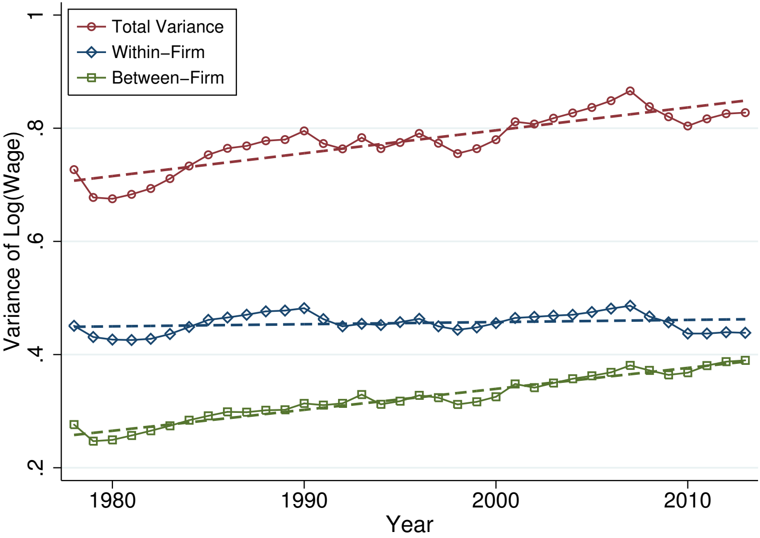
The change in earnings inequality by percentile from year t to year t' is:

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

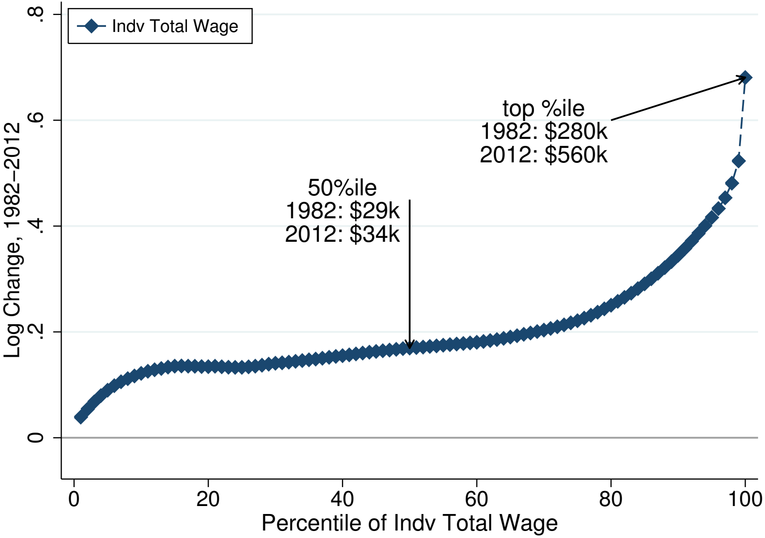
Total Wage Inequality



Total, Between- and Within-Firm Inequality

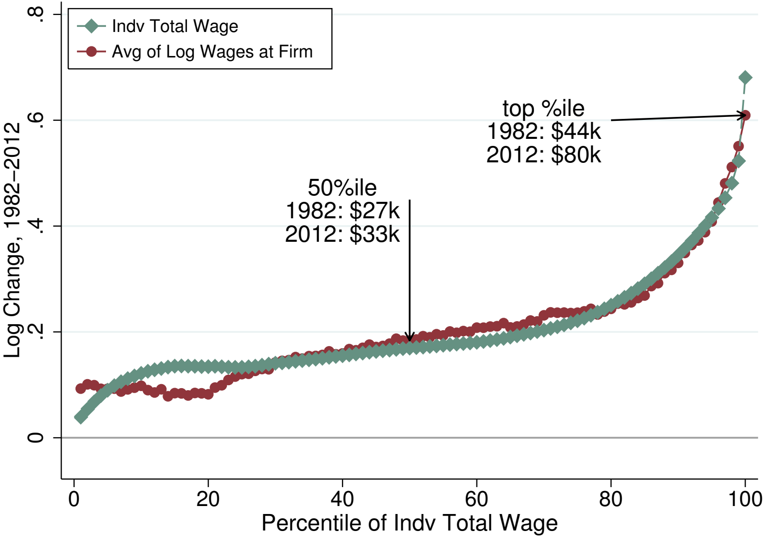


Wage Inequality: By Percentile



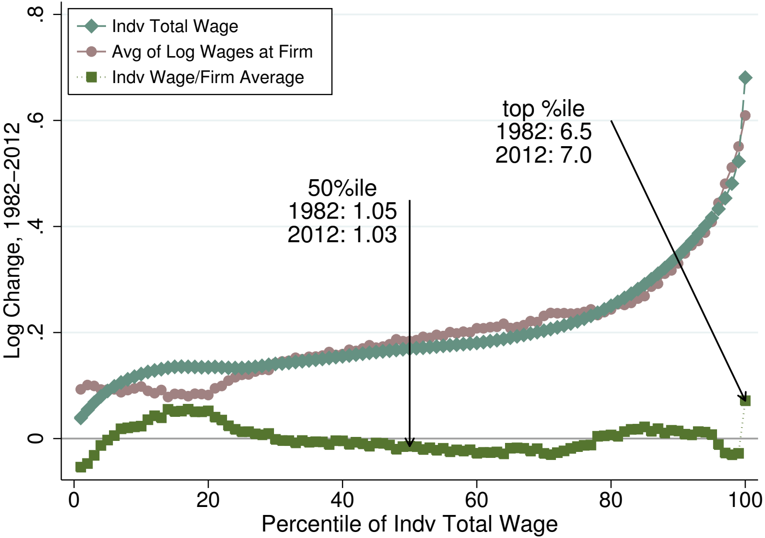
Note: Sample contains workers in firms with 20+ full-time equivalent employees.

Wage Inequality: Between Firms



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

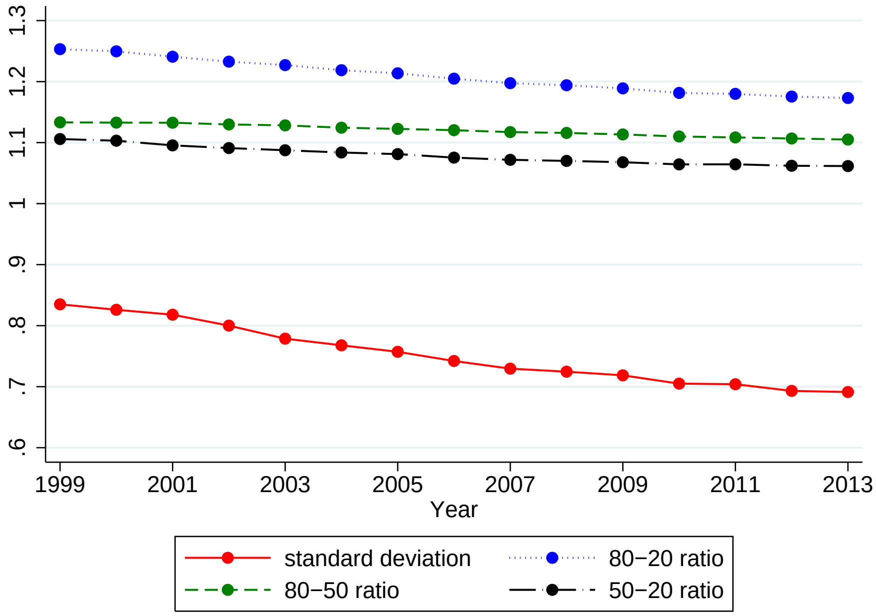
Wage Inequality: Within Firms



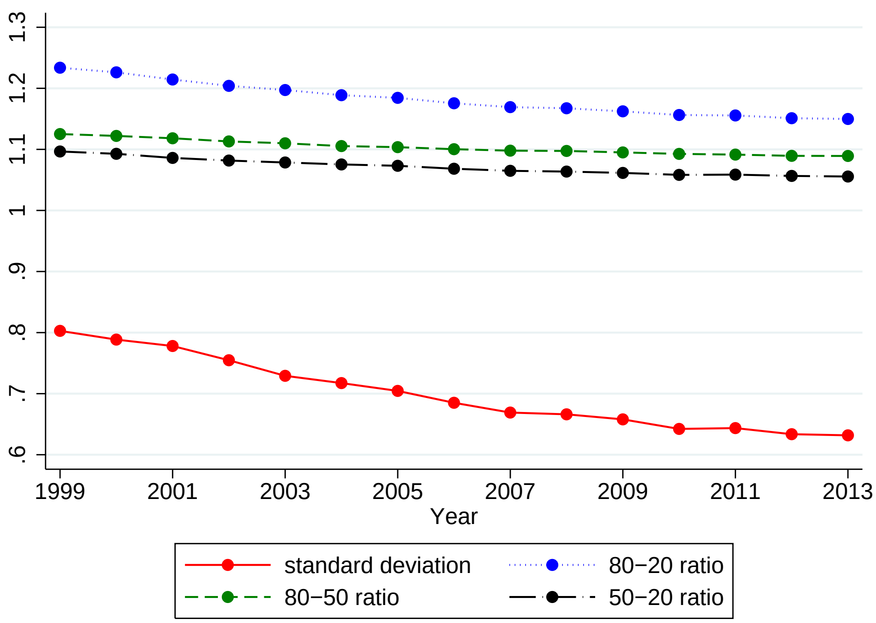
Note: Sample contains workers in firms with 20+ full-time equivalent employees.

Figure 1: Earnings Inequality, 1999 - 2013.

PANEL A: All sectors.



PANEL B: Excluding government, education and health.

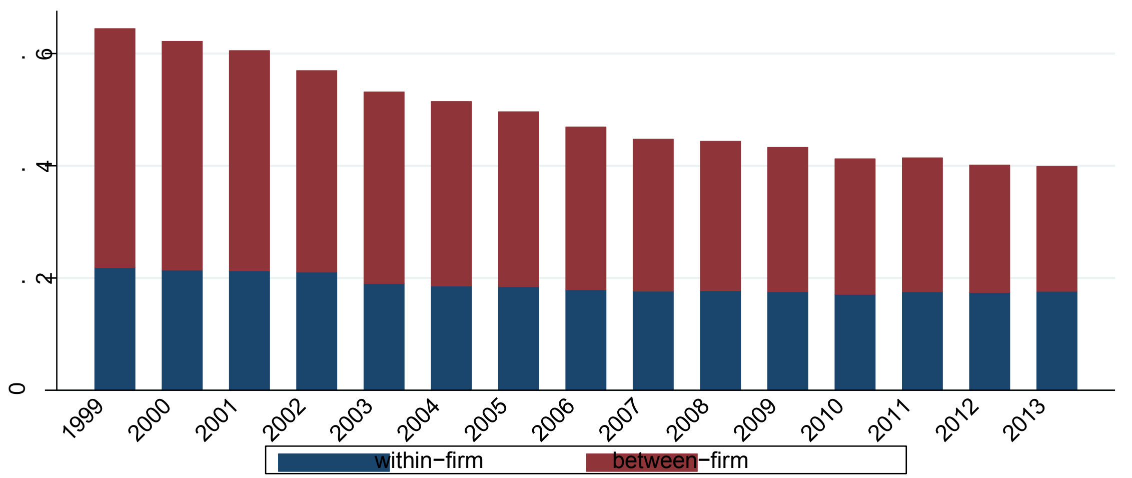


Notes: This graph shows the evolution of earnings inequality over the 1999-2013 period measured by the standard deviation and the 80-20, 80-50 and 50-20 percentile ratios of log monthly earnings.

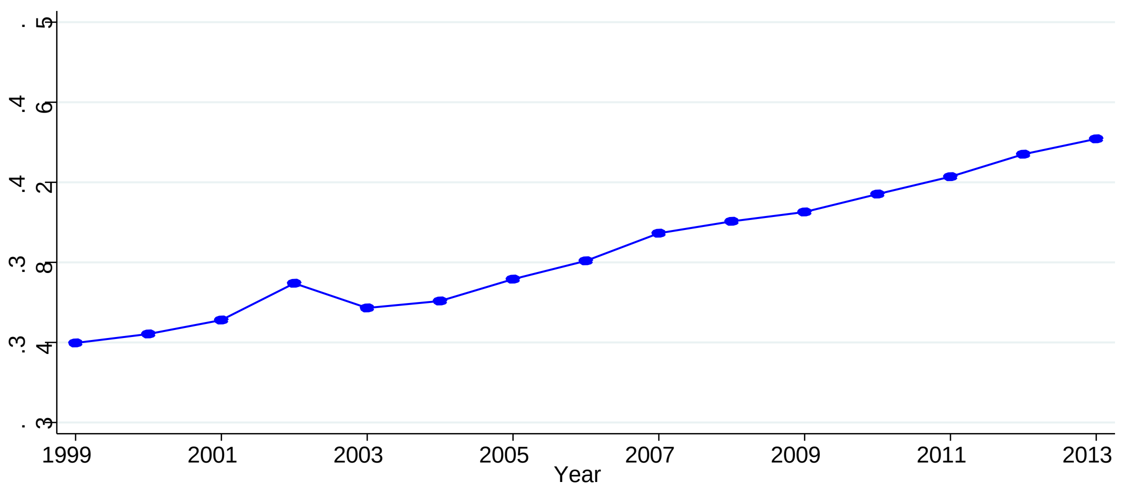
BRAZIL (Benguria 2015)

Figure 6: Between-Firm and Within-Firm Earnings Inequality, 1999 - 2013.

PANEL A: Between-Firm and Within-Firm Inequality.

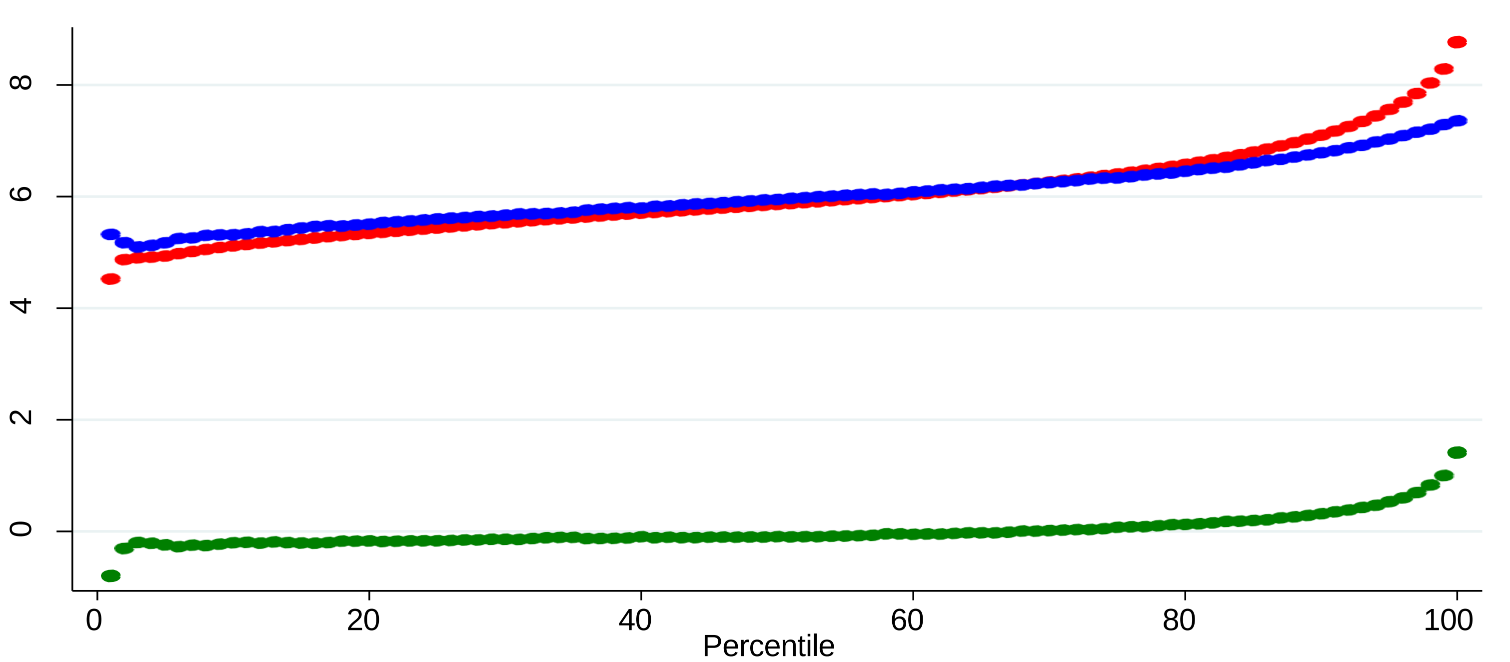


PANEL B: Share of Within-Firm Inequality.



Notes: This graph shows the evolution of within-firm (blue) and between-firm (red) earnings inequality over the 1999-2013 period measured according to equation 1. The sums of both bars corresponds to overall earnings inequality.

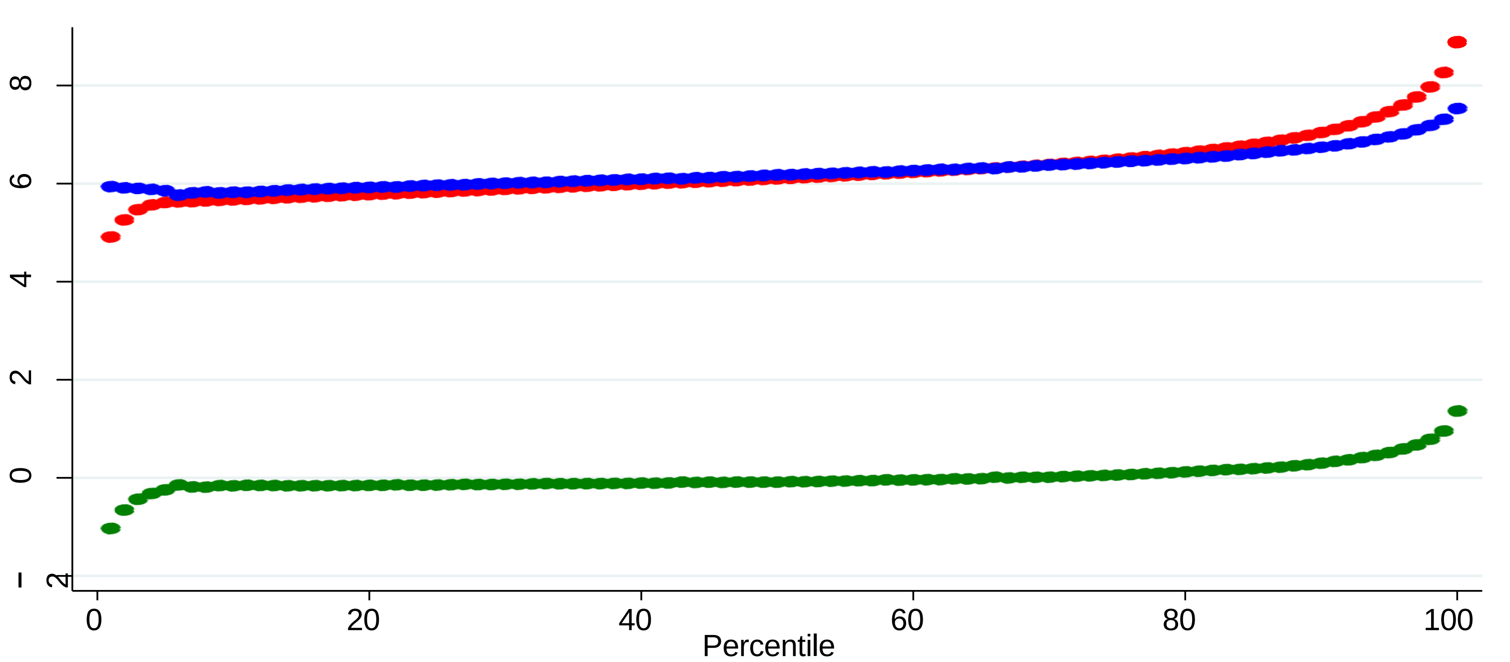
Figure 11: Analysis by Percentiles, 1999 Cross-Section.



RED: individuals, BLUE: firms, GREEN: individual/firm

Notes: This graph shows the earnings for different percentiles of the earnings distribution. For the red line, earnings are based on the average earnings of individuals in each percentile. For the blue line, earnings are based on the average earnings of the firms that employ individuals in each percentile. For the green line, earnings are based on the average of the difference in earnings between workers and their employers.

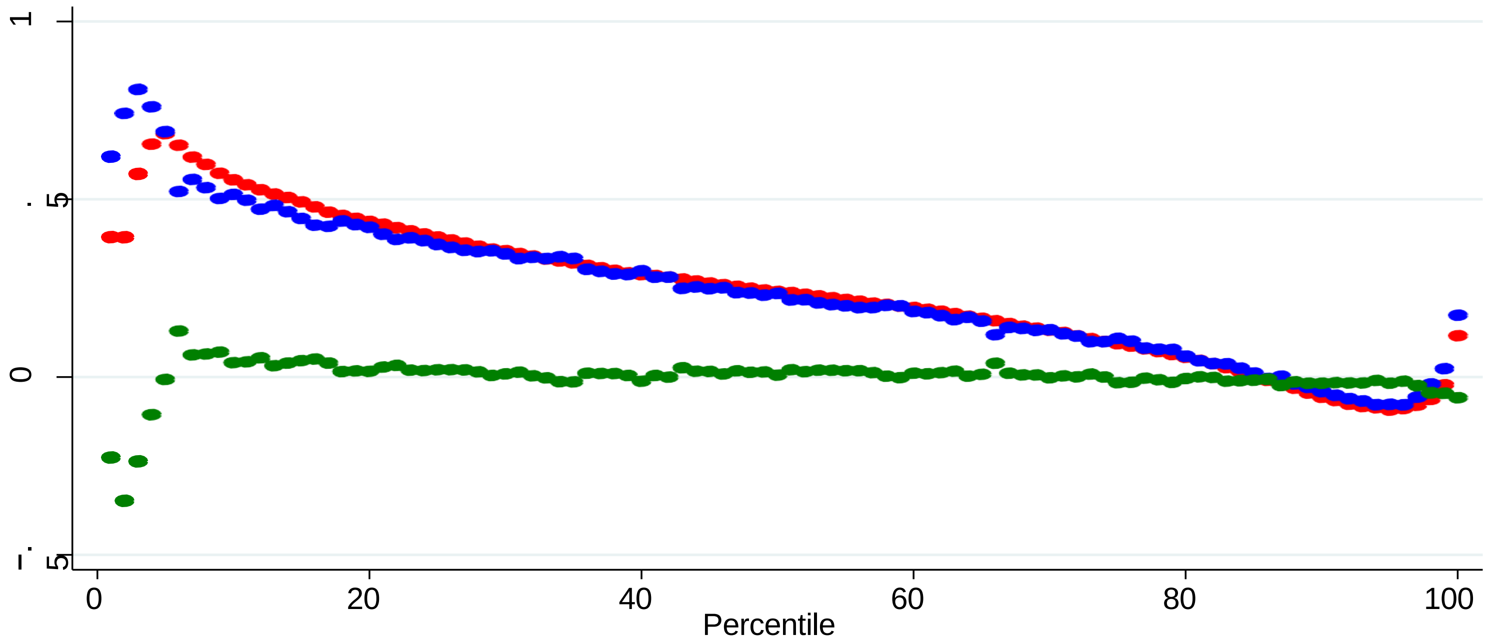
FIGURE 11: ANALYSIS BY PERCENTILES, 2013 CROSS-SECTION.



RED: INDIVIDUALS, BLUE: FIRMS, GREEN: INDIVIDUAL/FIRM

NOTES: This graph shows the earnings for different percentiles of the earnings distribution. For the red line, earnings are based on the average earnings of individuals in each percentile. For the blue line, earnings are based on the average earnings of the firms that employ individuals in each percentile. For the green line, earnings are based on the average of the difference in earnings between workers and their employers.

Figure 11: Analysis by Percentiles, 1999 - 2013.



RED: individuals, BLUE: firms, GREEN: individual/firm

Notes: This graph shows the 1999-2013 growth in earnings for different percentiles of the earnings distribution. For the red line, growth in earnings is based on the average earnings of individuals in each percentile. For the blue line, growth in earnings is based on the average earnings of the firms that employ individuals in each percentile. For the green line, growth in earnings is based on the average of the difference in earnings between workers and their employers. A positively sloped curve reflects a growth in inequality - individuals at the top of the distribution earn more (red line), work in firms paying more (blue line) or earn more in comparison to their employers' average wage.

Ubiquitous demand/supply concerns and arbitrage

- The exact distributions of firm (demand) and worker (supply) heterogeneity must be second order in explaining those earnings regularities across occupations, industries, time and space. **The levels of demand and supply must be second order.**
- The log earnings function is a pricing function. Invariant pricing features depend on ubiquitous demand/supply concerns and arbitrage arguments. E.g. Mincer schooling model and the experience earnings profile are based on supply side concerns.

This paper:

- Ubiquitous demand concerns:
  1. There are gains to specialization which implies multiple occupations and team revenue is convex in occupational skills.
  2. Team revenue supermodular in occupational skills.
- Arbitrage: (1) occupational choice, and (2) matching.

What we do

- Static frictionless labor market.
- No firm heterogeneity.
- Workers choose occupations by comparative advantage (**Roy**).
- Across occupations, workers choose team mates (matching: **Becker**).
- As proof of concept, our quantitative model, fitted to the Brazilian earnings distribution has the discussed invariant earnings regularities. And it suggests that the decline of earnings inequality in Brazil is due to her increased educational attainment.

Model's inputs:

- Bivariate distribution of workers' skills: compact domain, continuous density (other shape restrictions?).
- Occupational skill aggregation functions: occupational skill indexes must not be monotone transforms of each other.
- Technology for team revenue: supermodular and convex in occupational skills.

Cognitive and non-cognitive skills in the labor market

1. There are 2 roles per team, k for key role and s for support role. Team member in role, n, has cognitive skill c_n and non-cognitive skill r_n.
2. The revenue of a team is:

$$R(c_k, r_k, c_s, r_s) = c_k^{\alpha_k} r_k^{\beta_k} c_s^{\alpha_s} r_s^{\beta_s}$$

3. In the key role model, the cognitive ability of the person in the support role does not affect profit. The key role person is the owner of the team who solves

$$\pi(r_k c_k) = \max_{r_s} (r_k c_k)^{\alpha} r_s - w(r_s) \tag{1}$$

$$\pi(\kappa) = \max_{r} \kappa^{\alpha} r - w(r)$$

4. Three important assumptions:
   - The two occupational skill indices, κ and r, are not monotone transforms of each other.
   - Revenue is supermodular in the occupational skill indices.
   - Revenue is convex in occupational skills.

Deriving the matching function

$$\pi(\kappa) = \max_r \kappa^\alpha r - w(r)$$
$$\text{FOC: } \kappa^\alpha = w'(r^*) \Rightarrow r = \mu(\kappa)$$

- $r = \mu(\kappa)$ is the matching function.
- Due to supermodularity of revenue function, we get PAM by occupational skills (Becker). $\mu' > 0 \Rightarrow w(r)$ is convex in r.
- A convex revenue function in occupational skills implies occupational earnings are convex in $w(r)$ and $\pi(\kappa)$.
- Increasing returns to occupational skills imply occupational specialization. I.e. we should not observe workers in simultaneous multiple occupations.

Occupational choice: Derive separating function

- A worker with skills (κ, r) will choose the occupation which solves:

$$max[\pi(\kappa) - \eta, w(r)]$$

  where η is the cost of tuition to become a key role worker.

- $r = \phi(\kappa)$ defines the separating function (Roy), where workers with characteristics $(\kappa, \phi(\kappa))$ are indifferent between the two occupations:

$$\pi(\kappa) - \eta = w(\phi(\kappa)) \tag{2}$$

- Analytically, we integrate occupational choice (Roy) with matching (Becker).

Equilibrium

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

1. All key role workers choose support role workers to maximize their net earnings, i.e. solve equation (1).
2. All workers choose occupations which maximize their net earnings, i.e. solve equation (2).
3. The labor market clears: Every worker of type (κ, r) can find the job which maximizes their net earnings. And every key role worker of type κ can hire a support worker of type $\mu(\kappa)$ at wage $w(\mu(\kappa))$.

Due to PAM, labor market clearing can be written as:

$$H(\kappa; \phi(.)) = G(\mu(\kappa); \phi(.)) \; \forall \; \kappa \tag{3}$$

$$\int_{\underline{\kappa}}^{\kappa} \int_{\underline{r}}^{\min(\phi(u), \overline{r})} f(u, v) dv du = \int_{\underline{r}}^{\mu(\kappa)} \int_{\phi^{-1}(v)}^{\overline{\kappa}} f(u, v) du dv \; \forall \; \kappa$$

Equation (3) says that for every κ, the mass of key role workers up to skill κ must be equal to the mass of support role workers up to skill $\mu(\kappa)$.

Equilibrium

Theorem An equilibrium, consisting of four unique functions, an earnings function for support workers, $w(r)$, an earnings function for key role workers, $\pi(\kappa)$, a separating function, $\phi(\kappa)$, and a matching function, $\mu(\kappa)$, exists. Furthermore,

1. $w(r)$ is increasing and convex in r.
2. $\pi(\kappa)$ is increasing and convex in κ.
3. $\phi(\kappa)$ is weakly increasing in κ.
4. $\mu(\kappa)$ is increasing in κ.
5. $\phi(\kappa)$ and $\mu(\kappa)$ solves equations (2) and (3).

The Social Planner's Linear Programming Problem

- Let the revenue function of a type $t_1 \equiv (\kappa_1, r_1)$ person in the key role and a type $t_2 \equiv (\kappa_2, r_2)$ person in the support role be:

$$R(t_1, t_2) \equiv R((\kappa_1, r_1), (\kappa_2, r_2)) = \kappa_1^{\alpha} r_2, \quad \forall t_1, t_2 \in \mathcal{T}$$

- A coupling $m(t_1, t_2)$ is the mass of (t_1, t_2) matches. Let the mass of individuals for each type $t = (k, r) \in \mathcal{T}$ be specified by $f(t)$. Then feasible couplings must satisfy:

$$\int_{\mathcal{T}} m(t; \tau) d\tau + \int_{\mathcal{T}} m(\tau; t) d\tau \equiv f(k, r), \forall (k, r) \in \mathcal{T} \tag{4}$$

- The social planner considers the following problem:

$$\max_{m} \int_{\mathcal{T}} \int_{\mathcal{T}} R(t_1, t_2) m(t_1, t_2) dt_1 dt_2 \tag{5}$$

subject to feasibility couplings (4).

Calibration parameters

- The skills distributions, r and κ, are independent.
- κ, the key role skill distribution is the schooling distribution by years in Brazil (Benguria).
- r, the support role distribution is a symmetric turncated normal distribution (at 3sd) with the same support as the schooling distribution.
- The revenue function is

$$
\begin{aligned}
R(\kappa, r) &= A\kappa^{\alpha} r \\
A &= 0.158 \\
\alpha &= 2.8
\end{aligned}
$$

- Both parameters are chosen to match the intercept and slope of the aggregate earnings distribution by percentile in 1999 in Benguria.

Matching Brazil

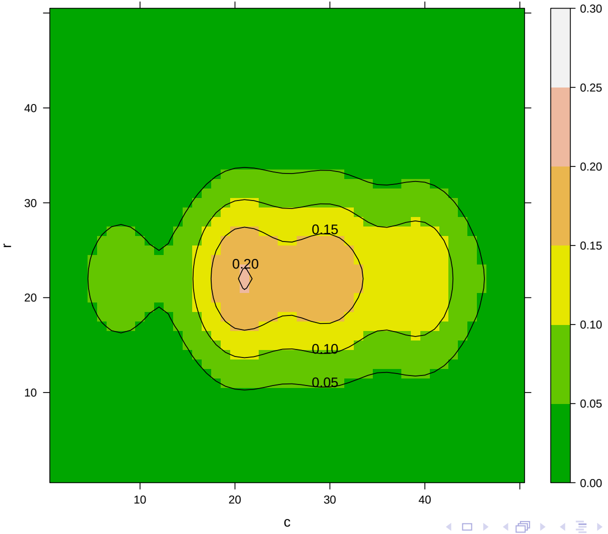
- α is matched to 1999 Overall Percentile's slope.
- A is matched to 1999 Overall Intercept.

	Benguria (1999)	Simulated (1999)
Overall		
Intercept	4.598	4.598
Slope	0.029	0.029
Gini	0.427	0.427
Between Firm		
Intercept	5.087	4.783
Slope	0.018	0.025
	Benguria (2013)	Simulated (2013)
Overall		
Intercept	5.146	5.59
Slope	0.022	0.019
Gini	0.295	0.286
Between Firm		
Intercept	5.01	5.632
Slope	0.006	0.017

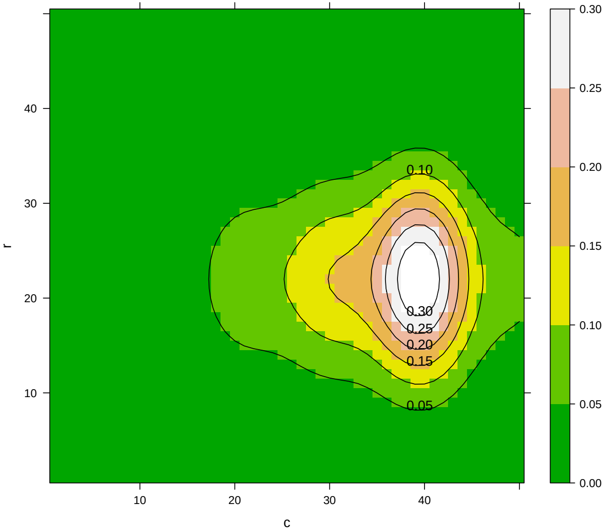
Type Distribution

1999

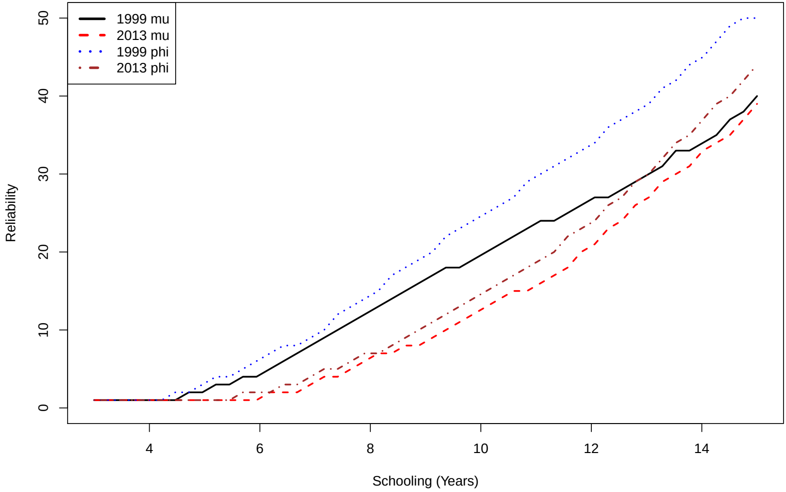
50 $\times$ 50 square grid for (c, r).



2013

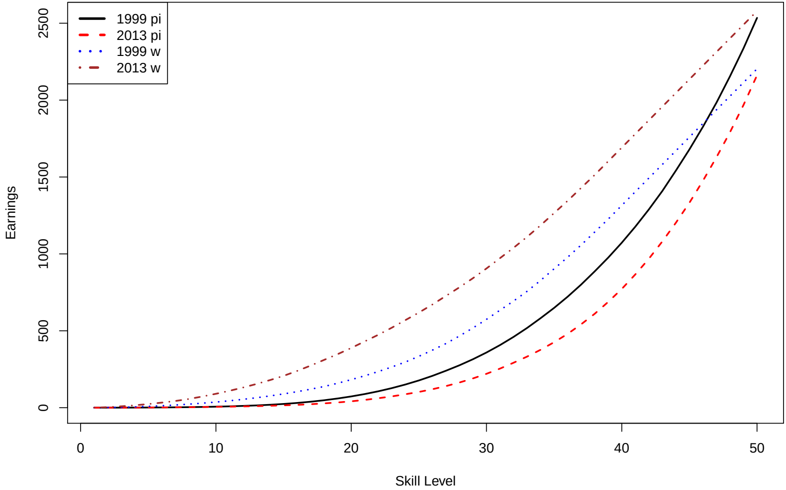


Plots of μ and ϕ



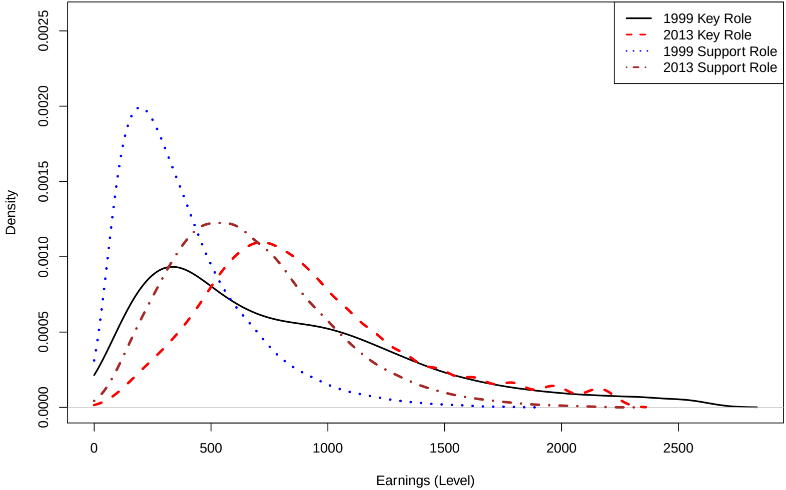
2000 IPUMS

Median Earnings in IPUMS data (Year 2000) is 300 BRL, or 5.7 in natural logs. Or about 600 USD.

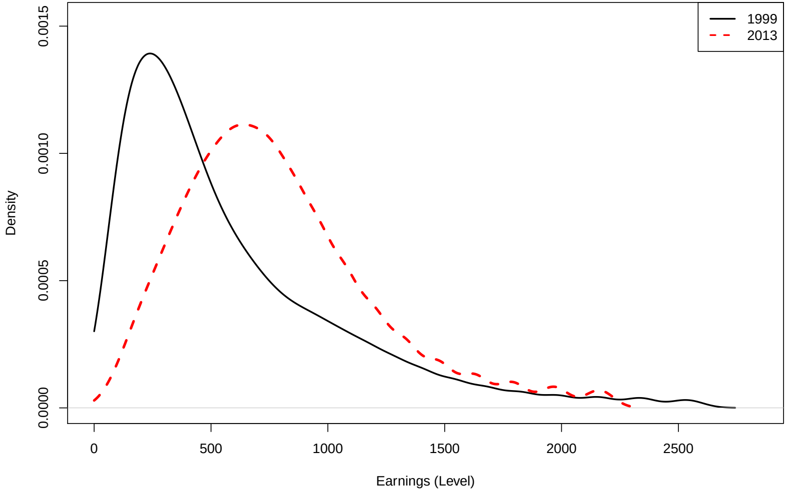


Level Earnings

By Occupation



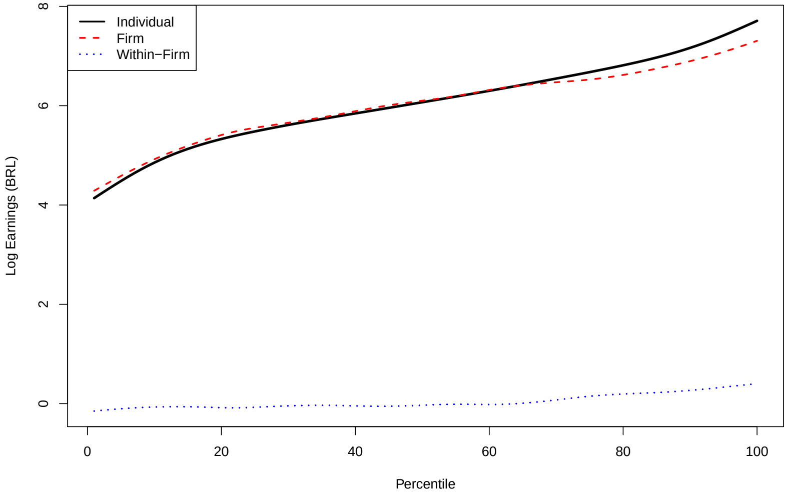
Aggregate



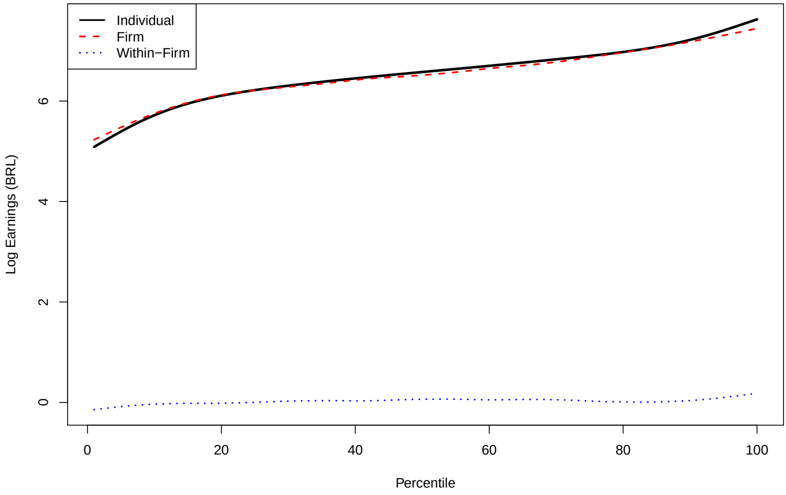
1999
2013
Density
0.0000
0.0005
0.0010
0.0015
0
500
1000
1500
2000
2500
Earnings (Level)

Percentile Plot

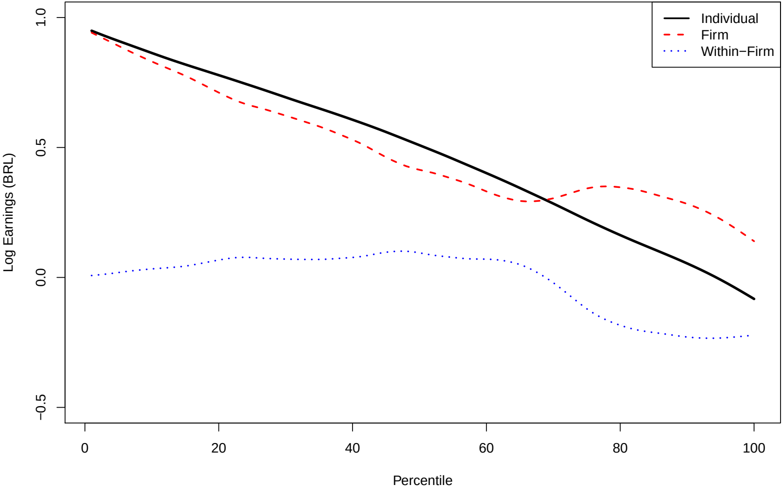
1999



2013


Individual
Firm
Within−Firm
Log Earnings (BRL)
0
2
4
6
0
20
40
60
80
100
Percentile

Change



Individual
Firm
Within−Firm
Log Earnings (BRL)
−0.5
0.0
0.5
1.0
0
20
40
60
80
100
Percentile

Earnings Regression

By Occupation

	1999 Key Role	2013 Key Role	1999 Support Role	2013 Support Role
	(1)	(2)	(3)	(4)
Schooling	0.362***	0.343***	0.105***	0.099***
	(0.002)	(0.003)	(0.006)	(0.004)
Constant	2.570***	2.500***	5.040***	5.470***
	(0.023)	(0.033)	(0.037)	(0.040)
Observations	2,075	2,196	2,016	1,943
R^2	0.932	0.878	0.134	0.217

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

In the simulation we smoothed the raw schooling distribution to get the distribution of c. Here in the regressions we use raw schooling (with support $\{1, 3, 6, 7, 9, 10, 12, 13, 15\}$).

Aggregate

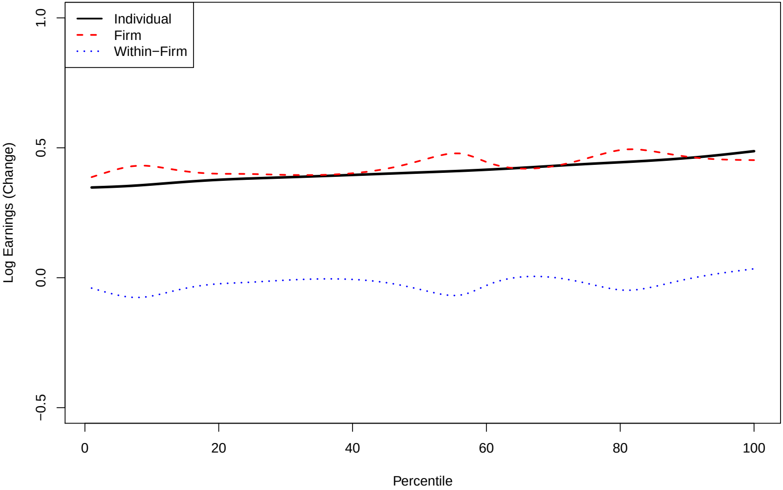
	1999				2013			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Schooling	0.187***	0.225***	0.108***	0.009***	0.138***	0.168***	0.051***	−0.001***
	(0.003)	(0.004)	(0.001)	(0.000)	(0.002)	(0.003)	(0.001)	(0.000)
Occupation		0.328***		−0.669***		0.223***		−0.320***
		(0.024)		(0.003)		(0.016)		(0.001)
Constant	4.480***	4.000***	3.020***	3.600***	5.050***	4.630***	4.280***	4.640***
	(0.023)	(0.041)	(0.197)	(0.049)	(0.027)	(0.040)	(0.105)	(0.027)
Firm FE	N	N	Y	Y	N	N	Y	Y
Observations	4,091	4,091	4,091	4,091	4,139	4,139	4,139	4,139
R^2	0.548	0.568	0.940	0.996	0.432	0.458	0.959	0.997

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Occupation dummy: (1 if key role, 0 if support role).

Skill Biased Technical Change

Now we increase α from 2.8 to 2.976 to match 2013's aggregate median earnings. Here we hold skill distribution at 1999 level.



What mitigates changes in within firm inequality?

- In our model, as the distribution of key role skills shift to the right, the distribution of support role skills remain unchanged.
- So how can within firm inequality remain roughly unchanged?
- **Occupation choice is key:** When the key role skill distribution shifts to the right, holding $w(r)$ constant, more workers want to be key role workers. So we have a shortage of support workers which means that the earnings of support workers must increase to keep enough support workers around.
- Remaining low skill key role workers can switch to be support workers.
- There are now more high skill key role workers. Their demand for high skill support workers will increase the skill gradient for support workers because the supply of skilled support workers have not changed much.
- The above three forces will tend to mitigate aggregate and within firm inequality in the face of changes in the distribution of workers' skills.

Literature review

- We build on Smith, Ricardo, Roy and Becker. Formally, we integrate Roy and Becker.
- We are not the first to investigate occupational choice and matching in teams.
- Fully multidimensional occupational choice and matching is hard to characterize. E.g. McCann and Trokhimtchouk; Lindenlaub.
- Kremer and Maskin, and others have one dimensional occupational choice and matching. Because both occupational choice and matching rely on absolute advantage, behavioral results are sensitive to fine demand/supply considerations. Lucas; Garicano and Rossi-Hansberg provide compelling behavioral rationalizations.
- Roy reduces a two dimensional occupational choice problem into a one dimensional occupational skill index for matching following Becker.
- So we can use comparative advantage to do occupational choice and absolute advantage to do matching.

Conclusions

1. The paper argues that occupational choice and matching, convexity and supermodularity of occupational skills in the team revenue function are first order features of labor markets.
2. These features can generate occupational earnings distributions which are single peaked and right skewed, firm fixed effects in log earnings regressions, and mitigate within firm earnings inequality as aggregate inequality changes.
3. Recent increased educational attainment is likely a significant factor in reducing earnings inequality in Brazil.
4. Our model and empirical exercise suggest that data on earnings distributions alone, both within and across firms, are unlikely to point identify structural parameters of a labor market.

Open questions

- What distribution of skills will generate single peak right skewed earnings distributions when wages are convex in occupational skills?
- We did not show that firm heterogeneity is unimportant for explaining the earnings distribution.
- So need to add heterogenous firm productivity and firm size. McCann, Shi, Siow and Wolthoff provides a lead.
- AKM models generate significant firm effects. How do we decompose those firm effects into firm heterogeneity versus worker heterogeneity with matching and occupational choice?
- How do we think about the consequences of increasing the minimum wage versus strengthening unionization?
- What about lifecycle concerns including job mobility?
- Can we extend to multi-industry and multi-occupation? Consequences for trade?

Empirical papers

Benguria, Felipe. "Inequality Between and Within Firms: Evidence from Brazil." Available at SSRN 2694693 (2015).

Pijoan-Mas, Josep, and Virginia Sánchez-Marcos. "Spain is different: Falling trends of inequality." *Review of Economic Dynamics* 13.1 (2010)

Song, Jae, et al. Firming up inequality. No. w21199. NBER, 2015.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. "Firms and Labor Market Inequality: Evidence and Some Theory." SSRN, 2016.

Khanna, Gaurav. "Large scale education reform in general equilibrium: Regression discontinuity evidence in India". University of Michigan manuscript.

Occupational choice and matching

Eeckhout, Jan, and Philipp Kircher. "Assortative matching with large firms: Span of control over more versus better workers." (2012).

Garicano, Luis, and Esteban Rossi-Hansberg. "Inequality and the Organization of Knowledge." American Economic Review 94.2 (2004)

Grossman, Gene M., Elhanan Helpman, and Philipp Kircher. Matching, Sorting, and the Distributional Effects of International Trade. 2014.

Gola, Pawel. "Supply and demand in a two sector matching model". Cambridge University manuscript. 2016.

Geerolf, Francois. "A Static Theory of Pareto Distributions." UCLA manuscript 2014.

Kremer, Michael, and Eric Maskin. "Wage inequality and segregation by skill." No. w5718. National Bureau of Economic Research, 1996.

McCann, Robert J., Shi, Siow and Wolthoff. "Becker meets ricardo: Multisector matching with communication and cognitive skills." Journal of Law, Economics, and Organization (2015):

McCann, Robert J., and Maxim Trokhimtchouk. "Optimal partition of a large labor force into working pairs." Economic theory 42.2 (2010)