

w o r k i n g
p a p e r

17 06R

**Organizations, Skills, and
Wage Inequality**

Roberto Pinheiro and Murat Tasci



FEDERAL RESERVE BANK OF CLEVELAND

ISSN: 2573-7953

Working papers of the Federal Reserve Bank of Cleveland are preliminary materials circulated to stimulate discussion and critical comment on research in progress. They may not have been subject to the formal editorial review accorded official Federal Reserve Bank of Cleveland publications. The views stated herein are those of the authors and are not necessarily those of the Federal Reserve Bank of Cleveland or the Board of Governors of the Federal Reserve System.

Working papers are available on the Cleveland Fed's website:

<https://clevelandfed.org/wp>

Firms, Skills, and Wage Inequality

Roberto Pinheiro and Murat Tasci

We present a model with search frictions and heterogeneous agents that allows us to decompose the overall increase in US wage inequality in the last 30 years into its within- and between-firm and skill components. We calibrate the model to evaluate how much of the overall rise in wage inequality and its components is explained by different channels. Output distribution per firm-skill pair more than accounts for the observed increase over this period. Parametric identification implies that the worker-specific component is responsible for 85 percent of this, compared to 15 percent that is attributable to firm-level productivity shifts.

JEL Codes: D02, D21, J2, J3.

Keywords: Multi-agent firms, skill distributions, wage inequality.

Suggested citation: Pinheiro, Roberto, and Murat Tasci, 2019. “Firms, Skills, and Wage Inequality,” Federal Reserve Bank of Cleveland, Working Paper no. 17-06R. <https://doi.org/10.26509/frbc-wp-201706r>.

Roberto Pinheiro is at the Federal Reserve Bank of Cleveland (Roberto.Pinheiro@clev.frb.org). Murat Tasci is at the Federal Reserve Bank of Cleveland (Murat.Tasci@clev.frb.org). The authors would like to thank Jim Albrecht, Sanjay Chugh, Ben Craig, Bruce Fallick, Andreas Hornstein, Pawel Krolikowski, Etienne Lalé, Rasmus Lentz, Richard Rogerson, and Susan Vroman for their comments, as well as the participants at the ROA Seminar at Maastricht University, Cleveland State University, Kent State University, Cleveland Fed Microeconomics Workshop, 2017 Applied Micro Fed System Conference, 2017 MEA-Sole, 2017 IIOC, 2017 SOLE Conference, 2017 European Econometric Society Meetings, and 2018 Joint Central Bankers’ Conference. They would like to thank Regis Barnichon for providing them with the Composite Help-Wanted Index data.

*First version May 2017 under the title “Organizations, Skills, and Wage Inequality.”

The increase in wage inequality over the last three decades is a well-documented pattern in the US data. While most of the literature has focused on the widening gap between college-educated and high-school-educated workers, the data also show a significant increase in wage inequality among observationally identical workers, in particular for highly skilled workers (Lemieux (2006a)). More recently, empirical work has documented the key role played by firms characteristics and quality in the rise in wage inequality (Song et al. (2019) and Barth et al. (2016)). While the empirical literature has shown progress documenting the role of firm and worker fixed effects, currently there is no theoretical framework that allows us to coherently address these different patterns presented in the data.

In this paper, we build a novel and tractable model that allows us to decompose the increase in US wage inequality over the past 30 years into its four main components and identify the contribution of each to the increase. The first component is the inequality between groups with different skill - the between-group component of wage-inequality discussed in papers addressing the college wage premium.¹ Second, we look into the wage dispersion among equally skilled workers - usually called within-group or residual wage inequality. Third, we look at the wage inequality among workers employed at the same firm - usually called within-firm wage inequality. Finally, we focus on the wage inequality driven by different employers - between-firm wage inequality.

We calibrate the models parameters to match the US economy for the years 1985 and 2015 and then use the calibrated model to evaluate how changes in the economy's fundamentals affected each one of these components. We focus on three different mechanisms: 1) changes in the output of education-firm pairs; 2) changes in market composition in terms of the educational attainment distribution; and 3) changes in labor market frictions. Moreover, we use our model's rank-preserving property - i.e., that a firm's rank in the firm productivity distribution is the same as its rank in the wage-posting distribution - and some additional parametric assumptions to recover firms' size and productivity distributions from the wage data. As a result, our methodology allows us to disentangle the impact of changes in output in terms of its key components: labor and firm productivities.

Our counterfactual exercises show that the key driver of the increase in overall wage dispersion between 1985 and 2015 is the change in the distribution of output per education-firm pairs between these two years. This measure shows the output flow generated by the match between a worker with a given educational level and a firm with a given productivity level. We show that the increase in overall, within-, and between-group wage inequality

¹Different skill levels in our model map into different educational attainment levels in the data as well as in our empirical exercises.

would have been significantly higher than that observed in 2015 if the only change in the model’s parameters was due to changes in the underlying distributions of output by education-firm pairs. In fact, changes in market composition - in particular the overall increase in average educational attainment in the labor force - partially mitigated the effect of the changes in the education-firm output distributions. Finally, changes in labor market frictions by themselves are able to explain between 40 percent to 60 percent of the increase in overall, within-, and between-group wage inequality.

Moreover, we disentangle the impact of changes in output in terms of its key components: labor and firm productivities. We show that changes in labor productivity explain the bulk of the impact of the changes in output distribution on overall, within-, and between- group wage inequality. Hence, changes in the firm-productivity distribution play a minor role in the increase in wage inequality. Furthermore, our decomposition of overall wage inequality in terms of within- and between-firm components shows that the increase in the between-firm component is responsible for the majority of the increase in overall wage inequality (67 percent). Consequently, our counterfactual exercises jointly with the within- and between-firm variance decomposition show that the increase in overall wage inequality through the between-firm component is explained by a combination of the sorting of high-skill workers in high-productivity firms and a large increase in high-skill workers’ labor productivity over the period.

We also believe the model is an important contribution by itself. We build a tractable model with frictional labor markets in which firms with different productivities compete for workers with different skills who are allowed to search on the job. Both worker skill and the firm productivity affect a match’s output. The model’s steady-state equilibrium presents patterns in line with those found by the empirical literature. We show that, in equilibrium, given some parametric assumptions that are in line with what we find empirically, high-productivity firms are larger, employing more workers and paying higher wages at all skill levels. Moreover, high-productivity firms hire proportionately more high-skill workers than their low-productivity competitors. Furthermore, high-productivity firms weakly hire a wider range of skills than their low-productivity peers. On the workers’ side, high-skill workers face a better wage offer distribution and earn, on average, higher wages than their low-skill peers. We show that these equilibrium features are in line with the observed patterns in Quarterly Workforce Indicators (QWI). Another important feature of our model is its rank-preserving property, i.e., the fact that we have a one-to-one relationship between a firm’s rank in the firm-productivity distribution and its rank in the wage-posting distribution. This modeling feature imposes enough structure to allow us to

estimate the output distribution relying solely on wage data, without the need for employer-employee matched data. As a result, the model can be quite useful in situations in which employer-employee matched data are not available, for example, when estimating models for developing economies.

Our empirical results also highlight that the increase in within-group dispersion that occurred between 1985 and 2015 has been concentrated among high-education groups (college grads and post-graduates). Our counterfactual exercises show that, because of the growth in high-skill labor productivity, matching with a high-productivity firm became increasingly more important for high-skill workers. Consequently, the increase in within-group wage inequality concentrated at the top of the educational attainment distribution can be explained by some high-skill workers being lucky to land a job at a high-productivity firm, while others are unlucky and work for low-productivity firms, generating low output due to skill-firm complementarities. As pointed out by Uren and Virag (2011), the fact that luck is a leading cause of within-group inequality may be a feature that helps the model fit the empirical evidence. The empirical evidence shows that the increase in within-group dispersion has often been transitory (see Gottschalk and Moffitt (1994) and Kambourov and Manovskii (2009), among others). This transient pattern is at odds with an explanation based on unobserved skills, since unobserved skills are usually either constant or move slowly over time.²

Our paper contributes to the literature on skill heterogeneity and wage inequality in several ways. We present and calibrate an extended on-the-job search model that allows us to evaluate the importance of labor frictions, labor and firm productivity, and market composition in explaining the increase in overall wage dispersion. Moreover, we are also able to evaluate the impact of these variables on the within- and between-group and within- and between-firm components of wage dispersion. To our knowledge, this is the first time that this quantitative exercise has been presented, in particular using a model that allows us to pin down both decompositions. In this sense, our model adds to the discussion on within- and between-group inequality presented by Lemieux (2006a,b, 2008) and Acemoglu and Autor (2011), as well as the models presented by Uren and Virag (2011) and Albrecht and Vroman (2002). In the same vein, we are able to address the factors that contributed to the increase in between-firm wage inequality, highlighted by Barth et al. (2016) and Song et al. (2019). Finally, we are also able to discuss the different patterns of job mobility be-

²The literature also considers that changes in the price of unobserved skills due to skill-biased technology changes are usually persistent, reinforcing the idea that increases in within-group wage inequality should be long-lasting.

tween skill levels, as discussed by Dolado, Jansen, and Jimeno (2009), while also addressing and extending the discussion on the relationship between firm size, productivity, and skill distribution presented by Eeckhout and Pinheiro (2014).

Section 1 presents the model and equilibrium properties. Section II discusses the data as well as our calibration for the labor market frictions and educational- attainment distribution, along with the non-parametric and parametric calibration of the output per skill-firm pairs. Section 3 presents the benchmark results for our non-parametric and parametric approaches. Section 4 presents counterfactual exercises for our non-parametric and parametric calibrations. Section 5 concludes the paper. All proofs, as well as additional evidence and explanations, are presented in the online appendix.

1 Model

Consider an economy with a measure 1 of firms. We assume that firms differ in their productivity level x . In particular, we assume that a firm's productivity follows a continuous distribution $\Gamma(x)$ with support $[x, \bar{x}]$ and no mass points. Firms are risk neutral, infinitely lived, and discount the future at rate $r > 0$.

There is a measure M of workers in the economy. Workers are heterogeneous in their skill levels. Worker skill is observable. There are $\mathcal{I}$ skills in the economy, ranked from the lowest to the highest ($1 < 2 < \dots < \mathcal{I} - 1 < \mathcal{I}$). The measure of workers of skill i in the economy is m_i . As a result, we have $\sum_{i=1}^{\mathcal{I}} m_i = M$. All workers are risk neutral, discount the future at rate $r > 0$ and exit the market at rate $d > 0$, regardless of their skill level. Exiting workers are replaced by unemployed workers with the same skill level. The cost of exiting the market is normalized to zero.

Labor markets are frictional and search is random. Labor markets are segmented across skills.³ In particular, firms can costlessly post vacancies across different markets. As in Burdett and Mortensen (1998), we assume that a vacancy is attached to a posted wage. We assume that firms can simultaneously post different wages in different markets. As a result, we assume a firm actively posts in any given skill's labor market whenever the output flow produced by the match is above the worker's reservation wage. Workers can costlessly search for a job offer regardless of their employment status. Moreover, search efficiency is not affected by the worker's employment status. Consequently, the flow of matches in labor market i is given by $\mathcal{M}_i(m_i, V_i)$, where V_i is the measure of firms actively posting vacancies

³In a previous version, we consider the case of non-segmented markets, which generates qualitatively identical results.

in skill i 's labor market. Defining $\underline{x}(i)$ as the firm with the lowest productivity among the ones that hire skill level i in equilibrium, we have that $V_i = [1 - \Gamma(\underline{x}(i))]$. Consequently, the arrival rate of job offers to a skill i worker is given by $\lambda(i) = \frac{\mathcal{M}_i(m_i, V_i)}{m_i}$. Job offers consist of a wage rate w . Offers must be accepted or rejected on the spot, with no recall.

Once a match between a firm with productivity x and a worker of skill level i is formed, they produce an output flow of $p(x, i)$. As a result, the output flow created by a match is not affected by either the size or the composition of the firm's labor force.⁴ We assume that $\frac{\partial p(x, i)}{\partial x} > 0$ and $\frac{\partial p(x, i)}{\partial i} \geq 0$. Additionally, unless otherwise stated, we also assume that $p(x, i)$ is supermodular. Finally, a match with a worker of skill i is exogenously destroyed at the rate $\delta_i \in (0, 1)$.

While searching for a job, unemployed workers engage in home production. We assume that home production by a worker of skill level i produces an output flow rate $b(i)$ with $b'(i) \geq 0$. We also assume that $p(\underline{x}, 1) \geq b(1)$, which implies that all firms are active at least at the lowest rung of the skill ladder.

1.1 Worker's Problem

Given the environment described above, the expected discounted lifetime income when a skill- i worker is unemployed, $U(i)$, can be expressed as the solution of the following equation:

$$(r + d)U(i) = b(i) + \lambda(i) \int \max\{J(w', i) - U(i), 0\} dF_i(w') \quad (1)$$

where $b(i)$ is the flow value of home production, d is the rate at which workers exit the market, $\lambda(i)$ is the arrival rate of a job offer, and $F(\cdot)$ is the distribution of posted wages faced by a skill- i worker.

Once a skill- i worker is employed at a firm paying a wage rate w , the value of holding a job at this company is

$$(r + d)J(w, i) = w + \lambda(i) \int_w [J(w', i) - J(w, i)] dF_i(w') + \delta_i [U(i) - J(w, i)] \quad (2)$$

where δ_i is the rate at which matches with skill- i workers are exogenously destroyed. Notice

⁴As pointed out by Hagedorn, Law, and Manovskii (2017): "This model of the firm, as simplistic as it is, represents the current state of the art in this literature" As Lentz and Mortensen (2010) (pp. 593-594) put it: "All the analyses that we know of assume that output of any given job-worker match is independent of the firm's other matches. Furthermore, firm output is the sum of all the match outputs. Hence, the identification challenge reduces to that of identifying worker and firm contributions over matches and a common match production function. Of course, as the research frontier moves to improve our understanding of multiworker firms, it is likely and appropriately an assumption that will be challenged."

that we already incorporated in the value function the fact that a worker will accept any offer as long as it is above her current wage. This is the case because of the commitment that firms make to a flat wage rate and no counteroffers to outside opportunities an employee may receive.

As $J(w, i)$ is strictly increasing in w whereas $U(i)$ is independent of it, a reservation wage for a skill- i worker, $R(i)$, exists and is defined by $J(R(i), i) = U(i)$. Then, from equations (1) and (2), we obtain that $R(i) = b(i)$, which is driven by the fact that on-the-job search is as effective as out-of-the-job search.

Finally, let's consider the measures of employed and unemployed workers in a steady state equilibrium. Notice that the flows of skill- i workers in and out of employment in the steady state is given by

$$d(m_i - u_i) + \delta_i(m_i - u_i) = \lambda(i) [1 - F_i(R(i))] u_i \quad (3)$$

where the left-hand side (henceforth LHS) of the above equality represents the inflow of skill i workers into the unemployment pool, while the right-hand side (henceforth RHS) represents the outflow.

In order to simplify notation, let's define $\kappa_i = \frac{\lambda(i)}{d + \delta_i}$. Following Jolivet (2009), we call κ_i the *search friction index* for skill i , corresponding to the ratio between the arrival of positive and adverse shocks faced by a worker. Similarly, define $\bar{F}(x) = 1 - F(x)$. As a result, we can rewrite equation (3) as:

$$u_i = \frac{m_i}{1 + \kappa_i \bar{F}_i(R(i))} \quad (4)$$

1.2 Firm's Problem

In this subsection, we take the behavior of workers as given and derive the firms' optimal response. Firms post wages that maximize their profits, taking as given the distribution of wages posted by their competitors ($F_i(w)$, $i \in \{1, 2, \dots, \mathcal{I}\}$) and the distribution of wages that employed workers are currently earning at other firms, given by $G_i(w)$, $i \in \{1, 2, \dots, \mathcal{I}\}$. We assume here that all distributions are stationary and well-behaved.

When a firm is choosing its optimal wage level at each skill level, it has to take into consideration the number of workers it can attract at any given wage and whether it is optimal to attract a given skill level in the first place. For this reason, before we analyze the firm's wage decision, let's derive the distribution of earned wages $G_i(w)$, $\forall i \in \{1, 2, \dots, \mathcal{I}\}$ and the firm's labor force for each skill level i . In order to do that, let's start with the firm's

decision to hire a given skill level or not. In order to hire a skill- i worker, the firm must pay a wage that is at least as high as the reservation wage $R(i)$. Consequently, a firm with productivity level x will only hire skill i if and only if $p(x, i) \geq R(i)$. Then, substituting $R(i) = b(i)$, we have that $p(x, i) \geq b(i)$. Since $p(x, i)$ is strictly increasing on x , $\underline{x}(i)$, the lowest productivity-level firm that hires skill level i , is determined by $p(\underline{x}(i), i) = b(i)$. Given $\frac{\partial p(x, i)}{\partial x} > 0$, for any $x > \underline{x}(i)$, the demand for skill i is strictly positive. The following lemma shows that $\underline{x}(i)$ must be nondecreasing in i .

Lemma 1 $\underline{x}(i)$ is nondecreasing in i .

In principle, since we do not impose a free-entry condition, we may have $\underline{x} \geq \underline{x}(i)$. In this case, all firm types will hire skill i . In order to take this possibility into account, we define $\underline{x}^*(i) \equiv \max\{\underline{x}, \underline{x}(i)\}$.

We now consider the distribution of employed skill- i workers who earn a wage less than w , $G_i(w)$. This distribution evolves over time according to:

$$\begin{aligned} \frac{dG_i(w, t)(m_i - u_i(t))}{dt} &= \lambda(i)[F_i(w, t) - F_i(R(i), t)]u_i(t) \\ &- \{d + \delta_i + \lambda(i)\bar{F}_i(w, t)\} G_i(w, t)(m_i - u_i(t)) \end{aligned} \quad (5)$$

The first term on the RHS of (5) describes the inflow at time t of skill- i unemployed workers into firms offering a wage no greater than w to skill- i workers at time t , whereas the second term represents the outflow into death, unemployment, and higher-paying jobs, respectively. Since in the steady state $\frac{dG_i(w, t)}{dt} = 0$, this distribution can be rewritten as

$$G_i(w) = \frac{[F_i(w) - F_i(R(i))]}{[1 + \kappa_i \bar{F}_i(w)] \bar{F}_i(R(i))} \quad (6)$$

Finally, the steady state measure of skill- i workers earning in the interval $[w, w - \varepsilon]$ is represented by $[G_i(w) - G_i(w - \varepsilon)](m_i - u_i)$, while the measure of firms offering wages in the same interval is $\bar{\Gamma}(\underline{x}^*(i)) [F_i(w) - F_i(w - \varepsilon)]$. Notice that $\bar{\Gamma}(\underline{x}^*(i))$ represents the measure of firms actively hiring skill- i workers. Thus, the measure of skill- i workers earning a wage w and employed at a firm actively hiring skill i can be expressed as

$$l(w; i) = \frac{(m_i - u_i)}{\bar{\Gamma}(\underline{x}^*(i))} \lim_{\varepsilon \rightarrow 0} \frac{G_i(w|i) - G_i(w - \varepsilon|i)}{F_i(w|i) - F_i(w - \varepsilon|i)}$$

Solving it and substituting u_i , we obtain the following:

$$l(w; i) = \frac{\kappa_i m_i}{\bar{\Gamma}(\underline{x}^*(i)) \{1 + \kappa_i \bar{F}_i(w)\}^2} \quad (7)$$

1.3 The Firm's Problem and Labor Market Equilibrium

Firms face the problem of picking the wage that maximizes their steady state profits. If a firm pays a higher wage than its peers, *ceteris paribus*, workers will more highly value being employed in this firm relative to other firms. As a result, workers employed in other firms are more likely to move into the firm in question whenever they receive this firm's wage offer. Similarly, when the firm's own workers receive alternative offers themselves, they are more likely to reject those offers and stay with the firm. Therefore, with higher wages facilitating recruitment and retention, the firm employs more workers in the steady state. While this force pushes overall profits upward, it comes at the cost of earning lower profits per worker, which pushes overall profits downward. At the optimal wage, the firm balances these two counteracting forces, maximizing total steady state profits.

As we mention before, firms have different levels of productivity x . The distribution of productivity levels in the economy is given by a continuous $\Gamma(\cdot)$ with support $[\underline{x}, \bar{x}]$. Firms can offer different wages for different skill levels. Therefore, the profit function for a firm that has productivity level x and posts wages $\{w(i)\}_{i=1}^T$ is given by

$$\pi(x) = \sum_{i \in A(x)} [p(x, i) - w(i)] l(w(i); i) \quad (8)$$

where $A(x)$ denotes the set of skills in which a productivity- x firm is actively searching, i.e., posts a wage above the reservation wage of that skill level. It is easy to see that a firm with productivity x would post a wage above the reservation wage $R(j)$ of a given skill j if $p(x, j) \geq R(j)$. Therefore, we can define

$$A(x) = \{i \in \{1, \dots, N\} \mid p(x, i) \geq R(i)\} \quad (9)$$

Therefore, the firm's problem is to pin down the wage schedule $\{w(i)\}_{i \in A(x)}$ in order to maximize $\pi(x)$. Because of the linearity of the profit function in skills, firms can pin down the wage posted for each skill level separately, i.e., firms can solve

$$\pi(x; i) = \max_w (p(x, i) - w) l(w; i), \quad \forall i \in A(x) \quad (10)$$

Then, the optimal wage posted by a firm of productivity x for a skill level $i \in A(x)$ is given by $w = K(x; i)$. From the first-order condition, i.e., $\frac{\partial \pi(x, w(i); i)}{\partial w(i)} = 0$, substituting equation (7) and manipulating it, we obtain

$$2\kappa_i f_i(w)(p(x, i) - w) - [1 + \kappa_i \bar{F}_i(w)] = 0 \quad (11)$$

which provides us with an implicit equation for w as a function of $p(x, i)$ and $F(\cdot | i)$. At this stage, we proceed as if the second-order conditions are satisfied. In corollary A.1 – presented in the appendix – we verify that this condition is met by the equilibrium solution.

Let's now consider the optimal wage function $w(x, i) = K(x, i)$. The flow profit $\pi(x; i)$ that a skill level i generates for a firm of productivity level x offering wage $K(x; i)$ is $(p(x, i) - K(x; i))l(K(x; i); i)$. Then, a differentiation with respect to x , using the envelope theorem, gives us

$$\frac{\partial \pi(x; i)}{\partial x} = \frac{\partial p(x, i)}{\partial x} l(K(x; i); i) \quad (12)$$

The boundary condition associated with the differential equation presented in (12) depends on $\underline{x}^*(i)$. We solve the differential equation, while substituting (7) and taking into account that $\bar{F}_i(K(x'; i)) = \frac{\bar{\Gamma}(x')}{\bar{\Gamma}(\underline{x}^*(i))}$ (see Bontemps, Robin, and Van den Berg (2000), Proposition 3). Moreover, in the appendix we show in detail, the arguments that allow us to take into account the boundary condition $[p(\underline{x}^*(i), i) - K(\underline{x}^*(i), i)]l(w(\underline{x}^*(i), i))$. As a result, we obtain

$$\pi(x; i) = \int_{\underline{x}(i)}^x \frac{\partial p(x', i)}{\partial x'} \frac{\kappa_i m_i}{\bar{\Gamma}(\underline{x}^*(i)) \left[1 + \kappa_i \frac{\bar{\Gamma}(x')}{\bar{\Gamma}(\underline{x}^*(i))}\right]^2} dx' \quad (13)$$

Finally, once $\pi(x; i) = (p(x, i) - K(x; i))l(K(x; i); i) \Rightarrow K(x; i) = p(x, i) - \frac{\pi(x; i)}{l(K(x; i); i)}$. Then, substituting $\pi(x; i)$, we have

$$K(x; i) = p(x, i) - \int_{\underline{x}(i)}^x \frac{\partial p(x', i)}{\partial x'} \left(\frac{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x')}{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x')} \right)^2 dx' \quad (14)$$

Notice that $K(\underline{x}^*(i); i) = b(i) = R(i)$, i.e., the minimum wage optimally posted is equal to the skill's reservation wage. We are now ready to present an important initial result.

Lemma 2 (rank-preserving) *Consider two firms with productivity levels x and $\tilde{x}$ with $x > \tilde{x}$ that hire workers of skill level i . The optimal offered wages are strictly increasing in firm productivity.*

Lemma 2 delivers us a rank-preserving property, once it shows that there is a one-to-one relationship between a firm’s productivity x and its posted wages. As a result, given that $\gamma(x)$ is a continuous distribution with no mass point, the “rank” that a x -productivity firm has in the productivity distribution (its quantile position in the distribution) is the same as the one it has in the wage-posting distribution, once adjusted by the measure of firms that decide to hire a particular skill i . In particular, we corroborate the previously mentioned result in Bontemps, Robin, and Van den Berg (2000) that:

$$F_i(K(x; i)) = \frac{\Gamma(x) - \Gamma(\underline{x}(i))}{\bar{\Gamma}(\underline{x}(i))} \quad (15)$$

This property will be particularly important during our parametric calibration exercise, presented in Section 2.3.2. Moreover, notice that $F_i(K(\underline{x}(i); i)) = F_i(R(i)) = 0$, as expected.

In order to pin down the labor distribution within and across firms, define $l_i(x) \equiv l_i(K(x, i))$. Then, we get the additional results:

Lemma 3 *For any two firms with productivity levels x and x' with $x > x'$, we must have that $l_i(x) \geq l_i(x')$, $\forall i \in \mathcal{I}$. Consequently, high-productivity firms hire more at all skill levels.*

Then, define the size of a firm with productivity x by $S(x) = \sum_{i=1}^{\mathbf{I}(x)} l_i(x)$, where $\mathbf{I}(x)$ is the highest skill level hired by a firm with productivity level x . Then, a simple corollary of lemmas 1 and 3 follows:

Corollary 1 *Firm size is increasing in x .*

Now, let’s define the skill distribution at a x -type firm by $\Phi_x(i) = \frac{\sum_{i'=1}^i l_{i'}(x)}{S(x)}$. The following proposition allows us to associate a firm-skill distribution with its productivity level.

Proposition 1 *Assume that, for any two skill levels i and j with $i > j$, we have $\kappa_i \geq \kappa_j$. In this case, high-productivity firms hire proportionately more high-skill workers.*

First, based on the proof for Proposition 1, we can see that the assumption on κ ’s is a sufficient condition. Moreover, while κ ’s are partially endogenous through λ ’s – which are pinned down in equilibrium – important components of κ (m , d , and δ) are exogenously determined.⁵ In fact, Cairo and Cajner (2017) show that the differences in κ across educa-

⁵The model presented in this paper is an example of what Lopes de Melo (2018) calls a “piece-rate” model. As Lopes de Melo (2018) shows in Sections 3.1 and 3.2, models don’t present sorting in equilibrium, once reservation strategies and meeting rates are independent of firm type. Consequently, the differences in κ ’s are likely a necessary condition for Proposition 1. Nevertheless, in Lopes de Melo (2018)’s calibration, κ is also increasing with education.

tional groups is mostly due to differences in δ . Our empirical results presented in Section 2.4 corroborate their results.

Theorem 1 summarizes all the results presented up to now.

Theorem 1 *In equilibrium, compared to their low-productivity counterparts, high-productivity firms:*

1. *Hire a wider range of skills;*
2. *Pay higher wages at all skill levels;*
3. *Hire more at all skill levels;*
4. *Are bigger;*
5. *Hire proportionately more high-skill workers if the κ 's are increasing in skill.*

For completeness, in the appendix we present our equilibrium concept. This is a standard concept, following the same general properties presented in Bontemps, Robin, and Van den Berg (2000), apart from a few simplifications due to our parameter restrictions.

1.3.1 Wage distributions by skill type

We present how the equilibrium distributions of posted and earned wages vary with skills, highlighting how job market opportunities and outcomes are impacted by workers' skills.

We start showing how firms' wage-posting strategies vary with workers' skills. In particular, from the optimal wageposting strategy presented in equation (14), we obtain the following result.

Lemma 4 (wages increasing in skill) *Assume that, for any two skill levels i and j with $i > j$, we have $\kappa_i \geq \kappa_j$. Then, for any firm with productivity level x that hires both skill levels, we must have $K(x, i) > K(x, j)$.*

Therefore, firms with higher productivity offer higher wages at all skill levels, and workers with higher skill levels receive higher wage offers at all different productivity levels. Notice that this does not mean that there is no overlap between the distributions of offered wages for different skills, once the wage offered by firms with different productivity levels to workers with different skill levels may be the same. In fact, as a corollary of Lemma 4, we know that if two firms offer the same wage to workers of different skill levels, the firm that offers this particular wage to the higher-skill worker must have lower productivity.

Corollary 2 *If a given wage w is offered for both skills i and j , $i > j$, the firm offering wage w for skill level i has lower productivity than the firm offering the wage for skill j .*

Similarly, gathering our previous results and our assumptions on $b(\cdot)$ and $p(\cdot, \cdot)$ and taking into account the traditional Burdett and Mortensen (1998) arguments (so distributions are continuous with connected supports), we also have the following corollary:

Corollary 3 *If $i > j$, we have that the support of offered wages is given by $[b(i), K(\bar{x}, i)]$ and $[b(j), K(\bar{x}, j)]$, where $b(i) \geq b(j)$ and $K(\bar{x}, i) > K(\bar{x}, j)$.*

Then, considering the distribution of posted wages by firms and how a firm's productivity affects its posted wages, we can show the following result:

Proposition 2 *If $i > j$, $F_i(w)$ dominates stochastically in first-order $F_j(w)$.*

As a straightforward consequence of first-order stochastic dominance (henceforth, F.O.S.D.), we have

Corollary 4 *The average offered wage increases with skill.*

Once we have obtained these results for the distribution of posted wages, we are now able to present some key characteristics of the distributions of earned wages by skill level. From equation (6) and Lemma 2's implication that $F_i(R(i)) = 0$, we have that the cumulative distribution and density function of earned wages are given by

$$G_i(w) = \frac{F_i(w)}{1 + \kappa_i \bar{F}_i(w)} \quad \text{and} \quad g_i(w) = \begin{cases} \frac{(1+\kappa_i)f_i(w)}{[1+\kappa_i\bar{F}_i(w)]^2} & \text{if } w \in [\underline{w}_i, \bar{w}_i] \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

respectively. We are now able to show the following proposition:

Proposition 3 *Assume that, for any two skill levels i and j with $i > j$, we have $\kappa_i \geq \kappa_j$. In this case, the distribution of earned wages for skill i F.O.S.D. the distribution for skill j .*

Corollary 5 *Assume that, for any two skill levels i and j with $i > j$, we have $\kappa_i \geq \kappa_j$. Then, average earned wages increase with skill level.*

In appendix Section A.2, we derive economy-wide wage distributions. Moreover, in the appendix Section A.3 we show how the model delivers wage variance decompositions both within- and between-skills as well as within- and between-firms. These decompositions are then calculated in Sections 3 and 4 using the calibrated parameters and estimated wage, firm productivity, and output distributions.

Moreover, in appendix Section A.4 we present empirical evidence based on the Quarterly Workforce Indicators (QWI). The section presents summary statistics and empirical patterns by firm size that are in line with the results we have obtained from our model. While the evidence is mostly suggestive, it offers an indication in favor of the model. Our empirical approach relies on the identification of firm-level properties using our model and only data from the worker-level observations conditional on skill. Appendix Section A.4 highlights that the model’s implications are nevertheless consistent with the firm-level observations, albeit at an aggregate level.

2 Calibration and Data

2.1 Data

We calibrate our steady state model for the US economy at two different points in time (1985 and 2015). The unit of time considered is one month. To calibrate our parameters, we use wage and educational data from the Current Population Survey’s Merged Outgoing Rotation Groups (CPS MORG) for 1985 and 2015. For both years, we use the National Bureau of Economic Research (NBER) CPS labor extracts.⁶

We use data from the CPS MORGs of 1985 and 2015 in order to obtain stock consistent data on flows, following the methodology presented by Frazis et al. (2005). In the estimation of our matching function, we particularly use the Unemployment-to-Employment (UE) flows, while in the calibration of the job destruction rates, we use the Employment-to-Unemployment (EU) flows. Finally, we use Barnichon (2010)’s composite Help-Wanted Index (HWI), which combines information from the Conference Board’s “print” HWI, the Conference Board’s “online” HWI, as well as JOLTS.⁷

2.2 Skill Distribution

We assume that worker heterogeneity in terms of skill is observable and well measured by educational attainment. This assumption allows us to calibrate our model in terms of skill endowments directly from the data. The caveat is that our paper is unable to discuss issues driven by unobserved heterogeneity among workers beyond the patterns generated by labor

⁶Data reference is U.S. Census Bureau.

⁷According to a note from Barnichon’s website: “Because of divergence in the online HWI series and the JOLTS (see Cajner, , and Ratner (2016)), I now rely solely on the JOLTS V (number of non-farm ads) past 2001Q1 (...) The composite HWI index over 1951Q1-2000Q4 was converted in units of job openings rate, i.e. (number of V)/(size of the Labor Force).”

frictions.

We consider five educational-attainment groups: Less than High School, High School Graduates, Some College, College Graduates, and Post-Graduates. Since the 1985 CPS MORG does not have a variable that pins down the highest degree attained, we follow Jaeger (1997)’s method and attribute a given educational-attainment category based on the highest grade of school attended.⁸ The distributions of workers across the education groups are then used as our calibration targets (normalized to add up to 1) and are presented at the end of this section.

2.3 Productivity Distribution

Calibrating firm-level heterogeneity in terms of productivity x is a challenging endeavor. Due to data restrictions, estimating a firm-level productivity distribution from establishment-level data is beyond the scope of this paper. To solve this issue, we present the results for two approaches: non-parametric and parametric, each one with its benefits and caveats. The benefit of the non-parametric approach is that it allows us to estimate output distributions without imposing restrictions on $p(x, i)$ ’s functional form. On the other hand, the non-parametric approach neither allows us to recover the underlying firms’ productivity distribution ($\Gamma(x)$) nor allows us to evaluate how labor productivity varies across skills and over time. The benefit of the parametric approach is to allow us to recover the underlying firm productivity distribution ($\Gamma(x)$) and labor productivities up to a normalization. The drawback of the parametric approach is that we must impose a relatively simple functional form on $p(x, i)$.

2.3.1 Non-parametric Estimation of $p(x, i)$

We follow Bontemps, Robin, and Van den Berg (2000) and Launov (2006) to estimate $p(x, i)$. In particular, we nonparametrically (kernel estimation) estimate the distributions

⁸We consider as part of the educational-attainment category “less than high school” workers whose highest grade of school completed was either 11 years or less, as well as workers who attended grade 12 but did not complete it. Similarly, the educational-attainment category “some college” would comprise workers whose highest completed grades were 13, 14, and 15 or workers who attended but did not complete grade 16. Educational-attainment category “college graduates” would comprise workers who completed grade 16 or at least attended grade 17. Finally, educational-attainment category “post-graduates” would comprise workers who attended grades 18 and up. In 1992 the BLS switched from a years of schooling measure to a credential-oriented measure. As a result, for the 2015 CPS MORG, we are able to use the credential oriented measure recorded in variable *grade92* in order to pin down our educational-attainment categories. In particular “Less than High School” is pinned down by $grade92 \leq 38$, “high school graduate” by $grade92 = 39$, “some college” by $grade92 \in \{40, 41, 42\}$, “college graduate” by $grade92 = 43$, and “post-graduate” by $grade92 \geq 44$.

of earned wages ($G_i(w)$). Owing to top-coding, we use Launov (2006)’s procedure to obtain an estimate of the upper tail using a Pareto distribution. After performing adjustments in order to obtain distributions that are closest to the estimated ones but that still satisfy the model’s conditions, we use our estimates of the earned wages distribution and the calibrated parameters of the model in order to pin down output ($p(x, i)$) and output density functions ($v_i(p(x, i))$). We assume that all firms hire all skill levels, i.e., $\underline{x}^*(i) = \underline{x}$, $\forall i$. Appendix Section A.6 describes our methodology in detail.

The estimated cumulative distributions for output ($\Upsilon_i(p)$) are presented in Figure 1. Notice that the 2015 estimated output distributions dominate their 1985 counterparts for high educational-attainment groups. This is more pronounced for groups 4 (college graduates) and 5 (post graduates), but is still present in minor degrees for groups 2 (high school graduates) and 3 (some college). On the contrary, for the lowest levels of educational attainment, we see some additional mass concentrated at the lowest output levels.

Moreover, Figure 2 clearly shows the effect of F.O.S.D. in panel (a), with the larger increases in average output for college graduates and post-graduates. Differently, there is a decline in average output for employed workers with less than a high school education. Panel (b) shows that this increase in average output was accompanied by a large increase in within-group output inequality for high-education workers. In fact, between 1985 and 2015, college graduates and post-graduates have seen an increase in the standard deviation of output flow of 32.18 and 30.21 percent, respectively. Differently, workers with a high school diploma or some college have seen an increase in the standard deviation of the output flow in the same period of only 9.53 and 11.66 percent, respectively. In sharp contrast, workers with less than a high school education have actually seen a drop in output dispersion of -20.69 percent.

2.3.2 Parametric Estimation of $p(x, i)$

In this case, we assume a parametric production function $p(x, i)$ that has a functional form $p(x, i) = A(i)x$, where $A(i)$ is the labor productivity of the skill i worker and x is the firm’s productivity.

As in Section 1, we assume that x follows a distribution $\Gamma(x)$ continuous with no mass points. Notice that there is a unique underlying productivity distribution $\Gamma(\cdot)$, even though our non-parametric distribution delivered us five output distributions $\Upsilon_i(p(x, i))$, $i \in \{1, 2, 3, 4, 5\}$. Given that $A(i)$ is the same for all workers with the same education and the functional form of $p(x, i)$, we have that each $\Upsilon_i(p(x, i))$ is a simple transformation of the underlying $\Gamma(x)$. Moreover, this is a monotonic transformation, since both x and $A(i)$ are

positive terms. Consequently, the transformation between $\Upsilon_i(p(x, i))$, $i \in \{1, 2, 3, 4, 5\}$, and $\Gamma(x)$ is rank-preserving.

While both x and $A(i)$ are unobserved, the rank-preserving property of our non-parametric estimates for output distributions allows us to parametrically estimate $\Gamma(x)$ and $A(i)$, $i \in \{1, \dots, 5\}$. To illustrate the importance of the rank-preserving property, consider the following example. Assume that all firms hire skills i and j ($\underline{x}^*(i) = \underline{x}^*(j) = \underline{x}$). Consider two output levels $\tilde{p}_i$ and $\hat{p}_j$, observed for skills i and j , respectively. In this case, if $\Upsilon_i(\tilde{p}_i) = \Upsilon_j(\hat{p}_j)$, we have that the underlying firm productivity x is the same for $\tilde{p}_i$ and $\hat{p}_j$. In other words, if the output levels $\tilde{p}_i$ and $\hat{p}_j$ are at the same quantile of output distributions for skills i and j , they must share the same firm productivity level x , conditional on $\underline{x}^*(i) = \underline{x}^*(j)$.

We use this rank-preserving property in our parametric estimation of x and $A(i)$ s. In particular, we assume $\underline{x}(i) = \underline{x}$, $\forall i \in \{1, \dots, 5\}$. Then, we take our nonparametric estimates of the output levels $\hat{p}_i$, $i \in \{1, \dots, 5\}$ and output distributions $\hat{\Upsilon}_i(p)$, $i \in \{1, \dots, 5\}$ at 500 different quantile points $q \in [0, 1]$. For any quantile q , we have that if $\Upsilon_i(p_i) = \Upsilon_j(p_j) = q$, we must have $\frac{p_i}{p_j} = \frac{p(x, i)}{p(x, j)} = \frac{A(i)}{A(j)}$. Consequently, using our kernel density estimations for $\hat{p}_i$ and $\hat{\Upsilon}_i(\cdot)$ and using the less than high school category as our baseline (defined as $j = 1$), we have that, for each quantile q , we must have

$$\frac{\hat{p}_i}{\hat{p}_1} = \frac{A(i)}{A(1)} + \varepsilon_i \quad (17)$$

where $\varepsilon_i = \frac{\hat{p}_i}{\hat{p}_1} - \frac{p(x, i)}{p(x, 1)}$. As a result, running linear regressions represented by equation (17) across different quantiles for each skill level $i \in \{2, \dots, 5\}$ allows us to pin down estimates for $\frac{A(i)}{A(1)}$. We estimate (17) by median regression. The results are presented in Table 1.⁹

All education groups show an increase in labor productivity relative to high school dropouts over the period. Moreover, labor productivity growth was concentrated in the highest-education groups. In particular, while the (relative) labor productivity growth rate in the period for high school graduates and workers with some college education were 8.7 percent, 8.3 percent, respectively, the labor productivity growth rate for workers with college degrees and post-graduate education were 14.9 percent and 20.2 percent.

In order to pin down estimates for firms' productivity levels, define $\underline{p}_i = p(\underline{x}^*(i), i)$. Given the functional form for $p(x, i)$ and the fact that labor productivity is constant across

⁹Median regressions are more robust to outliers than least squares regression. Moreover, median regressions are semiparametric, avoiding assumptions about the parametric distribution of the error process. For these reasons, we present the results for the median regression in Table 1. However, results for the O.L.S. estimates are qualitatively the same.

Table 1: Est. Labor Productivity – 1985 vs. 2015
(95% confidence intervals in parenthesis)

1985				
$\frac{A_2}{A_1}$	$\frac{A_3}{A_1}$	$\frac{A_4}{A_1}$	$\frac{A_5}{A_1}$	$A_1 \underline{x}$
1.012 (1.009, 1.015)	1.034 (1.030, 1.038)	1.175 (1.169, 1.180)	1.301 (1.281, 1.321)	1277.47 (1269.63, 1285.30)
2015				
$\frac{A_2}{A_1}$	$\frac{A_3}{A_1}$	$\frac{A_4}{A_1}$	$\frac{A_5}{A_1}$	$A_1 \underline{x}$
1.100 (1.087, 1.113)	1.120 (1.110, 1.130)	1.350 (1.337, 1.363)	1.564 (1.548, 1.581)	1255.23 (1248.60, 1261.86)

workers with the same education, we have that $\frac{\hat{p}_i}{\underline{\hat{p}}_i} = \frac{p(x,i)}{p(\underline{x}^*(i),i)} = \frac{x}{\underline{x}^*(i)}$, $\forall i \in \{1, \dots, 5\}$. Consequently, using our kernel density estimations for $\hat{p}_i$ and keeping our maintained assumption that $\underline{x}(i) = \underline{x}$, $\forall i \in \{1, \dots, 5\}$, we have that, for each quantile q , we must have

$$\frac{\hat{p}_i}{\underline{\hat{p}}_i} = \frac{x}{\underline{x}} + v_i \quad (18)$$

where $v_i = \frac{\hat{p}_i}{\underline{\hat{p}}_i} - \frac{p(x,i)}{p(\underline{x},i)}$. Then, at a given quantile q , we are able to run a linear regression represented by equation (18) across different educational groups in order to estimate $\frac{\hat{x}}{\underline{x}}$. Repeating the procedure at several different quantile values allows us to estimate both the support for $\frac{x}{\underline{x}}$ as well as the distribution $\hat{\Gamma}(\cdot)$. As before, this methodology allows us to pin down the firm productivity distribution relative to the minimum productivity firm operating in the economy.

Our estimates for $\frac{\hat{x}}{\underline{x}}$ for 1985 and 2015 are presented in Figure 3. Differently from what we observed in the case of output distributions, we see a change in the productivity distribution between 1985 and 2015 that is closer to a mean-preserving spread.¹⁰

There are two remaining parameters to estimate: $A(1)$ and $\underline{x}$. Although we are unable

¹⁰This result is in line with the findings in the previous literature using data on public firms for the US and the UK (see İmrohoroglu and Tüzel (2014) and Faggio, Salvanes, and Van Reenen (2010), respectively). However, differently from what was previously found in the literature, the increase in the standard deviation that we found was somewhat small. In particular, while we observe a minor decline in average productivity between 1985 and 2015 – from 1.492 to 1.483, a decline of 0.59 percent – we see a modest increase in dispersion – the standard deviation increases by 3.60 percent in the period. Nevertheless, the reader should be aware of a couple of caveats to the comparison with the previous literature. First, the results we present in Figure 3 are all relative to the productivity of the least productive active firm in the economy. So changes in this productivity lower bound over time may significantly alter the results. Second, the results from the previous literature were based on a biased sub-sample of firms – public firms – which tend to be significantly larger and possibly more productive than their private counterparts.

to separately estimate these parameters, we can easily pin down their product. There are many ways to estimate $\widehat{A(1)\underline{x}}$. In particular, we take into account that our parametric estimates are given by $\tilde{p}_i = \frac{A(i)}{A(1)} \frac{\underline{x}}{\underline{x}}$. Then, we run a regression across all skill groups but less than high school (again properly lining up observations across quantiles), i.e., $\hat{p}_i = \beta \tilde{p}_i + u_i$, where $\hat{p}_i$ is our nonparametric estimate. Notice that $\hat{\beta} = \widehat{A(1)\underline{x}}$, gives us the joint estimate. We present the estimated values for 1985 and 2015 in table 1. As we can see, our estimates for 1985 and 2015 are similar, with the point estimate declining by 1.74 percent between the two years.¹¹ However, since changes in $\widehat{A(1)\underline{x}}$ represent the net effect of changes in $A(1)$ and $\underline{x}$, we can't say whether the changes in the underlying parameters were small or just mostly offset each other.

Keep in mind that, in order to do counterfactuals with respect to either changes in labor productivity across educational groups ($A(i)$) or changes in the firms' productivity distribution ($\Gamma(x)$), we must take a stand in how $\widehat{A(1)\underline{x}}$ is split between $\widehat{A(1)}$ and $\hat{x}$. In Sections 3.2 and 4.2, we focus on the case in which we normalize $A(1) = 1$ and attribute all changes in $\widehat{A(1)\underline{x}}$ to changes in the underlying firm productivity.

Given all the presented estimates in this section, we are able to create parametric counterparts for $p(x, i)$, for $i \in \{1, \dots, 5\}$. In the appendix, we present figures highlighting how much the parametric estimates deviate from the nonparametric ones. While the deviations tend to be overall small, they are larger for 1985 estimates. Moreover, compared to nonparametric estimates, the parametric estimates tend to underestimate low-skill workers' output and overestimate high-skill workers' output.

Finally, even though our assumption on $p(x, i)$'s functional form is restrictive, it is more general than initially perceived. For example, it encompasses Cobb-Douglas output functions such as $p(x, i) = A(i)^\beta x^\alpha$ by simply redefining $\tilde{A}(i) = A(i)^\beta$ and $\tilde{x} = x^\alpha$.¹²

2.4 Labor Frictions

As mentioned in Section 1, we pin down the job arrival rate for a skill i worker through the estimation of a matching function. In particular, we assume $\lambda(i) = \frac{\mathcal{M}_i(m_i, V_i)}{m_i}$. In order to estimate the matching function $\mathcal{M}_i(m_i, V_i)$, we follow Jolivet (2009) and Petrongolo and Pissarides (2001). In particular, we assume that $\mathcal{M}_i(m_i, V_i)$ has the following Cobb-Douglas

¹¹Although it is worthwhile to point out that the confidence interval is somewhat wider in 2015.

¹²In a previous version, we assumed a Cobb-Douglas functional form for $p(x, i)$ and used the estimates of the productivity distribution for public companies from İmrohoroglu and Tüzel (2014). This approach had some caveats in addition to those for the current parametric calibration, relying on the estimates from a selected sample of firms in order to obtain the productivity distribution and imposing additional *ad hoc* conditions in order to pin down the exponents α and β . Nevertheless, the results were qualitatively similar.

functional form:

$$\mathcal{M}_i(m_i, V_i) = \mathcal{A}_i(E_i + U_i)^{\beta_i} V_i^{1-\beta_i} \quad (19)$$

which implies that employed workers (E_i) and unemployed workers (U_i) are not only as effective while searching for a job, but also create congestion externalities for each other (as pointed out by Petrongolo and Pissarides (2001) in their Section 5). Moreover, we are clearly assuming constant returns to scale in the matching function.

Given this assumption, we can define the number of new matches for workers coming out of unemployment (M_{U_i}) as $M_{U_i} = \frac{U_i}{U_i + E_i} \mathcal{M}_i(U_i + E_i, V_i)$. Replacing $\mathcal{M}_i(U_i + E_i, V_i)$'s functional form, rearranging it, and including a time trend as suggested in Jolivet (2009), we obtain the following specification¹³:

$$\ln \frac{M_{U_i}}{U_i} = \ln \mathcal{A}_i + (1 - \beta_i) \ln \left(\frac{V}{U_i + E_i} \right) + \phi t + \varepsilon_t \quad (20)$$

where t is a time trend. Notice that we are assuming that $V_i = V$, $\forall i \in \{1, \dots, 5\}$, which is in line with the maintained assumption that $\underline{x}(i) = \underline{x}$, $\forall i \in \{1, \dots, 5\}$ and that all firms exercise the same search effort across all submarkets. These assumptions are necessary, since we don't have information on vacancy creation across educational group submarkets. In this sense, we use the same information on vacancy creation based on the Composite Help-Wanted Index provided by Barnichon (2010) for all submarkets. Results for the matching function estimates, as well as the adjustments in order to pin down the job finding rates in each submarket and ultimately λ_i (following Tasci (2012) and Shimer (2012), respectively), are detailed in appendix Section A.7.

We calibrate the death rate d in order to match the average overall death probability rate at the median age of 37 throughout the entire period (1985-2015). Moreover, for each time year (1985 and 2015) we calibrate the job destruction rate δ_i such that $\delta_i + d$ match the average EU transition rate in the data for each educational group.

Results for the adjusted job finding rates (λ_i), job destruction rates (δ_i), and death rates (d) are presented in Table 2. Similarly to Cairo and Cajner (2017), we observe that $\lambda(i)$'s are relatively similar, while job destruction rates (δ_i) are quite different across groups. As a result, *search friction indexes* (κ_i) are strictly increasing in skill. Moreover, notice that adjusted job finding rates declined significantly between 1985 and 2015, while job

¹³Following Jolivet (2009), we consider the following econometric model:

$$\mathcal{M}_i(m_i, V_i) = \mathcal{A}_i(E_i + U_i)^{\beta_i} V_i^{1-\beta_i} e^{\phi t + \varepsilon_i}$$

Table 2: Parameters

<i>A: 1985</i>					
	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$
m_i	0.1843	0.4039	0.1998	0.1234	0.0886
$b(i)$	0.2220	0.3398	0.4359	0.5201	0.5965
$\lambda(i)$	0.5168	0.4955	0.5747	0.5574	0.5359
δ_i	0.0584	0.0287	0.0233	0.0125	0.0093
d	.000115	.000115	.000115	.000115	.000115
κ_i	8.83	17.20	24.53	44.33	56.85
<i>B: 2015</i>					
	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$
m_i	0.0936	0.2665	0.2917	0.2258	0.1223
$b(i)$	0.1513	0.2267	0.2871	0.3396	0.3868
$\lambda(i)$	0.4561	0.4102	0.4381	0.3993	0.4040
δ_i	0.0528	0.0279	0.0220	0.0117	0.0093
d	.000115	.000115	.000115	.000115	.000115
κ_i	8.63	14.62	19.79	33.90	42.80

destruction rates weakly declined. As a result *search friction indexes* declined, in particular for high-skill workers.

Finally, the parameters for $b(i)$ are pinned down such that the lowest productivity firm makes zero profits by hiring any of the skills. We present the calibrated parameters in Table 2.

3 Benchmark Estimates: 1985 and 2015

3.1 Non-parametric Approach

We now present our results for the non-parametric calibrations for 1985 and 2015. In terms of changes in the wages per worker as a function of educational attainment, results are as expected. The model shows that average real wages have increased significantly for college graduates and post-graduates – by 12.10 and 14.13 percent, respectively. Differently, the model shows that real wages have barely changed for workers with a high school education and some college in the same period, with changes of -0.49 percent and 0.01 percent, respectively. In a remarkable contrast, the model shows a decline of 13 percent in real average wages for workers with less than a high school education. In terms of the within-group

Table 3: Variance Decomposition - Nonparametric
(Calculations using values in \$1000 2015 USD)

<i>Skills - Magnitudes</i>			
	Overall	Within-group	Between-group
1985 - Model	5.79 (100%)	4.68 (81%)	1.10 (19%)
1985 - Data	5.84 (100%)	4.81 (82%)	1.03 (18%)
2015 - Model	10.60 (100%)	8.01 (76%)	2.59 (24%)
2015 - Data	11.37 (100%)	9.08 (80%)	2.30 (20%)
% Change 1985-2015 : Model	78%	72%	101%
% Change 1985-2015 : Data	95%	89%	124%

standard deviation, the model shows an increase in the standard deviation of earned real wages of 16.25 and 18.77 percent between 1985 and 2015 for college graduates and post-graduates, respectively. In contrast, high school graduates and workers with some college have seen an increase in standard deviation of real wages by only 6.56 and 12.27 percentage points, respectively. Moreover, high school dropouts have actually seen a drop in real wage dispersion by -16.87 percent. In order to facilitate comparing benchmark results with the results from the counterfactual exercises, Figure 4 shows model-generated average wages and standard deviations in panels (a) and (b), respectively.

In terms of how much of the overall variation observed in the data our non-parametric calibration captures, Table A.4 in the appendix compares the average wage and within-group standard deviation for the model and data in 1985 and 2015. As presented in appendix's Table A.4, the model does a decent job of capturing the wage variation in both time periods and across the educational groups – particularly for the low educational attainment groups.¹⁴ While this good match should be expected, since the nonparametric calibration uses the entire wage distribution in order to obtain estimates for $p(x, i)$ and $\gamma(x)$, the benchmark model with the nonparametric calibration seems a good starting point to our counterfactual exercises.

Table 3 presents the overall wage variance in the data and in the model for both years, along with the breakdown in terms of its within- and between-group components. The per-

¹⁴Moreover, results do not seem to depend fundamentally on the imputation method used to obtain the censored upper tail. In an online appendix, we follow Dustmann, Ludsteck, and Schönberg (2009) and use different methods to impute censored wages. All methods deliver results that are qualitatively the same.

centages presented between parentheses underneath each number correspond to the fraction of the overall variance that is explained by each component. Table 3 shows a large increase in wage dispersion between 1985 and 2015. Moreover, the overall breakdown between within- and between-group components stayed reasonably stable over time, with both components growing at a pace similar to that of the overall variance. However, this breakdown is heavily influenced by the tail estimates of censored observations, so caution is necessary when interpreting the results.

3.2 Parametric Approach

We now present our results for the parametric calibrations for 1985 and 2015. In terms of changes in the wages per worker as a function of educational attainment, the model’s results are not as close as the nonparametric calibration to what was observed in the data. However, while the model is quite parsimonious in terms of the number of parameters, it is still able to capture around 40 percent to 60 percent of the within-skill wage dispersion observed in the data, measured by the standard deviation of earned wages. Appendix table A.5 presents further details on the comparison results.

In terms of the variance decomposition within and between educational groups, we see in Table 4 that the parametric model does a much better job detecting the between-group variation than the within-group variation observed in the data. This result is not surprising, given that we proxy skill heterogeneity by educational attainment, which is a coarse proxy, leaving a significant level of within-group heterogeneity unexplained. Given this caveat, we can see that a combination of firm heterogeneity and labor frictions is able to explain about 20 to 25 percent of the within-group heterogeneity observed in the data.

As previously discussed in Section 2.3.2, the parametric calibration allows us to identify the firms based on the model’s rank-preserving properties. As a result, we can present a model-based variance decomposition in terms of its within- and between-firm components, using equations 31 and 32 presented in the appendix. Results are presented in table 5. In the presented case, we normalize $A(1) = 1$ in both years. Our results indicate that the bulk of the total variance observed for real earned wages comes from the between-firm component (83 percent in 1985 and 77 percent in 2015). Moreover, while the within-firm component has grown faster than the between-firm component in the period 1985-2015, the increase in the between-firm component is responsible for most of the increase in overall variance (67 percent). This result is in line with what has been found in the empirical literature (see Song et al. (2019) and Barth et al. (2016)).

Moreover, we are able to pin down the firm size distribution. We are particularly

Table 4: Variance Decomposition - Parametric
(Calculations using values in \$1000 2015 USD)

<i>Skills - Magnitudes</i>			
	Overall	Within-group	Between-group
1985 - Model	1.87 (100%)	1.18 (63%)	0.69 (37%)
1985 - Data	5.84 (100%)	4.81 (82%)	1.03 (18%)
2015 - Model	2.86 (100%)	1.77 (62%)	1.09 (38%)
2015 - Data	11.37 (100%)	9.08 (80%)	2.30 (20%)
% Change 1985-2015 : Model	53%	50%	58%
% Change 1985-2015 : Data	95%	89%	124%

Table 5: Variance Decomposition - Firms
(Calculations using values in \$1000 2015 USD)

	Overall	Within-firm	Between-firm
1985 - Model	1.87 (100%)	0.33 (17%)	1.55 (83%)
2015 - Model	2.86 (100%)	0.65 (23%)	2.21 (77%)
% Change 1985-2015 : Model	53%	101%	43%

interested in seeing how the distribution has changed over time. Figure A.9 in the appendix shows that large firms have become significantly larger, while small and mid-size firms have declined in size. Even more important, the large firms that have seen positive growth in their labor force are concentrated in the top decile of the size distribution. This result is in line with the empirical pattern observed in appendix Figure A.4 based on the BDS data.

4 Counterfactual Experiments

4.1 Non-parametric Approach

In this section, we present our counterfactual exercises. Our model has three sets of parameters. First are the parameters for labor market frictions, defined by the matching function parameters $\mathcal{A}_j$ and β_i , the job destruction rate δ_i , and death rate d , for $i \in \{1, 2, \dots, 5\}$.

Table 6: Std. Deviation of Earned Wages
(Calculations using values in \$1000 2015 USD)

	>HS	HS	SC	Col	Col+
1985 Benchmark	1.20	1.62	1.81	2.83	4.29
2015 Benchmark	1.00	1.58	2.03	3.45	5.09
Difference	-0.20	-0.05	0.23	0.63	0.80
<i>Counterfactuals:</i>					
<i>Production</i>	91%	-315%	258%	218%	286%
<i>Labor Frictions</i>	-15%	-59%	319%	155.%	68%
<i>Market Composition</i>	-17%	128%	-118%	-82%	-59%

The second set includes parameters for market composition, defined by the educational-attainment distributions, characterized by $m_i, i \in \{1, 2, \dots, 5\}$. Finally, parameters for the output per skill-firm $p(x, i)$ distributions, which, in the case of the nonparametric approach, are obtained through the nonparametric estimation using data on earned wages. In our counterfactual exercises, we calculate how much of the increase in wage inequality – both overall as well as within- and between-skill groups – can be explained by changes in each one of the set of parameters separately. In order to do that, we start from the 1985 benchmark calibration inputs and consider that only one of the three sets of parameters (labor market frictions, market composition, and output distributions) is changed to its 2015 counterpart. We then compare the obtained wage distribution against the 1985 and 2015 benchmark results.

Table 6 and Figure 5 summarize the results. Table 6 shows the model-based standard deviation of earned wages for each educational group in 1985 and in 2015 and their difference. The table also shows how much of this difference can be explained by each counterfactual channel. Notice that the production counterfactual explains more than 100 percent of the difference for all but workers with less than a high school education. This result shows that changes in output not only explain the bulk of the result, but also that, if labor frictions and market composition had been kept at their 1985 levels, within-group inequality should be even higher for most education groups. Panel (b) in Figure 5 gives a visual perspective for the result. Similarly, Figure 5’s panel (a) shows that between-group wage inequality should be even higher without the changes in labor market friction and market composition.

In terms of the counterfactuals for labor market frictions and market composition, Table 6 and Figure 5 paint a more nuanced picture, as the effect varies significantly across groups. First, workers without a high school diploma seem to have benefited from both changes.

In fact, the labor friction indexes (κ_i) for workers with less than a high school education are significantly higher for both counterfactuals (labor frictions and market composition) than the 2015 counterpart.¹⁵ As a result, both average wages and the standard deviation of real wages are higher for workers without a high school diploma in these counterfactual exercises, compared to the 2015 benchmark. Differently, for all other education groups but high school graduates, the market composition counterfactual shows lower average wages and standard deviations than both of the 1985 and 2015 counterparts. The leading reason for this decline – in particular for workers with at least some college education – is the increase in the measure of workers in each educational group, reducing market tightness. Differently, in the case of the labor friction counterfactual, high educational-attainment groups see higher average wages and standard deviation of average wages than in the 1985 benchmark. In this case, the increase in efficiency, through higher $\mathcal{A}_i$ and lower β_i , implies higher labor friction indexes (κ_i) than their 2015 benchmark counterparts.

Table 7 shows not only the difference in overall variance across the different time periods and counterfactuals, but also how the overall wage inequality is decomposed in terms of within- and between-group components. As in Table 6, the percentages in the counterfactual rows show how much of the difference between the 1985 and 2015 values is observed in each counterfactual scenario. As one can see, just the changes in output distributions more than explain the overall increase in wage dispersion, generating an overshooting. In the opposite direction, the changes in market composition by themselves would in fact imply a decline in wage dispersion in the same period. Finally, the labor frictions counterfactual shows that the changes in labor market frictions by themselves are able to explain about half of the increase in wage inequality.

4.2 Parametric Approach

In this section we use the parametric approach to disentangle the impact of changes in output production in terms of its two components: labor productivity ($A(i)$) and firm productivity (x). As we have shown in Section 4.1, changes in the output flow by themselves were able to induce a large increase in wage dispersion, both within and between education groups.

¹⁵Notice that the drivers of labor friction improvements are different in each counterfactual. The market composition counterfactual has a higher κ_i for less than high school workers than the 2015 benchmark due to a better market tightness and positive time trend. The labor friction counterfactual has a higher κ_i for less than high school workers than the 2015 benchmark due to a smaller impact of the time trend, even though it faces a worse market tightness.

Table 7: Variance Decomposition - Skills
(Calculations using values in \$1000 2015 USD)

	$Var_G(w)$	<i>Within-group</i>	<i>Between-group</i>
1985	5.79	4.68	1.10
2015	10.60	8.01	2.59
Difference	4.81	3.33	1.48
Counterfactuals:			
<i>Production</i>	138%	124%	169%
<i>Labor Frictions</i>	53%	59%	40%
<i>Market Composition</i>	-6%	-4%	-10%

In fact, Tables 6 and 7 and Figure 5 show that the observed changes in market composition attenuated the effect of changes in output flow. In other words, without other changes in the labor market, inequality driven by changes in output would have been larger than the observed pattern. Consequently, further decomposing the effect of output flow changes may allow us to understand a key driver of wage inequality.

Table 8 shows the change in the standard deviation of earned wages between 1985 and 2015, based on our parametric calibration. We consider three different counterfactual exercises. *Production* considers that both $A(i)$ and x change to their 2015 levels, following the output counterfactual in the non-parametric case.¹⁶ *Labor Productivity* considers that only $A(i)$'s are at their 2015 levels, with all other parameters at the 1985 levels. *Firm Productivity* considers that only $\Gamma(x)$ is at its 2015 levels, with all other parameters at 1985 levels. Similarly to what we showed in Table 6, we observe that the production counterfactual generates a wide increase in standard deviation of earned wages across all educational groups. Different from the nonparametric case, we can decompose the effect of changes in output flow in terms its components. According to Table 8 and Figure 6's panel (b), the vast majority of the effect of output changes on the standard deviation of earned wages is through changes in labor productivity. The only educational group in which we see firm productivity having a larger impact is among workers without a high school diploma. However, this result is artificial owing to our normalization of $A(1) = 1$ in both periods.

Similarly, Tables 9 and 10 show that the key driving force behind the increase in wage inequality through changes in output flow are the changes in labor productivity. In fact, labor productivity is the key driver not only for the increase in the overall wage, but also for its within and between educational groups and firm components. Changes in firm

¹⁶Notice that results here are different from those of the nonparametric approach, since we are still assuming a parametric functional form for $p(x, i)$.

Table 8: Std. Deviation of Earned Wages
(Calculations using values in \$1000 2015 USD)

	>HS	HS	SC	Col	Col+
1985 Benchmark	0.71	0.94	1.07	1.42	1.65
2015 Benchmark	0.68	0.95	1.11	1.62	2.02
Difference	-0.04	0.02	0.03	0.21	0.38
Counterfactuals:					
<i>Production</i>	77%	504%	411%	168%	143%
<i>Labor Productivity</i>	0%	472%	286%	102%	88%
<i>Firm Productivity</i>	77%	30%	115%	57%	45%

productivity explain only a minor share of the difference between 1985 and 2015 benchmark values. In terms of the differences in average wages across educational groups, Figure 6's panel (a) shows that changes in labor productivity across educational groups induced a large increase in average wages for high-skill workers that may have been somewhat attenuated by changes in firm productivity.

Table 9: Variance Decomposition - Skills
(Calculations using values in \$1000 2015 USD)

	$Var_G(w)$	Within-group	Between-group
1985	1.87	1.18	0.69
2015	2.86	1.77	1.09
Difference	0.99	0.59	0.40
Counterfactuals:			
<i>Production</i>	106%	77%	148%
<i>Labor Productivity</i>	85%	51%	136%
<i>Firm Productivity</i>	15%	19%	9%

In summary, the results from our parametric counterfactual exercises are in line with Bowlus and Robinson (2012), i.e., it is the fact that high-skill workers became more productive (have more human capital than they used to) that drives the increase in wage inequality. Moreover, our results also imply that while we do observe that between-firm inequality explains the bulk of the overall increase in wage inequality, changes in firm productivity by themselves can only explain a limited fraction of the increase in wage inequality and dispersion over time.

Table 10: Variance Decomposition - Firms
(Calculations using values in \$1000 2015 USD)

	$Var_G(w)$	Within-firm	Between-firm
1985	1.87	0.33	1.55
2015	2.86	0.65	2.21
Difference	0.99	0.33	0.66
Counterfactuals:			
<i>Production</i>	106%	124%	98%
<i>Labor Productivity</i>	85%	108%	74%
<i>Firm Productivity</i>	15%	10%	18%

5 Conclusion

In this paper, we present a model that allows us to decompose the overall wage dispersion into dispersion between and within skill groups, as well as within and between firms, while delivering most of the properties discussed in the empirical literature on organizations as equilibrium properties. We calibrate the model both parametrically and non-parametrically using wage data from the CPS MORG for the years 1985 and 2015. Our results show that changes in the output flow produced by worker-firm pairs induced both an increase in the average wage and wage dispersion for highly educated workers, while depressing average wages and reducing dispersion among low-education workers. Moreover, as we disentangle the changes in output flows in terms of its components – labor and firm productivities – labor productivity changes explain the vast majority of the effect of output changes on wage dispersion.

In contrast, changes in market composition – in particular the increase in the fraction of the labor force with high educational attainment – attenuated the impact of changes in labor productivity. Finally, changes in labor frictions by themselves were able to explain about half of the change in overall wage inequality, as well as about half of the within- and between-group components.

We also present some methodological contributions. Our model depicts an equilibrium with a rank-preserving property, in which there is a one-to-one relationship between a firm’s rank in the productivity distribution and its rank in the wage-posting distribution. Consequently, we are able to recover information about the firm from the wage distribution, subject to parametric assumptions on the production function. We calculated firm size distributions for both 1985 and 2015 and showed that firms in the top decile of the size dis-

tribution grew significantly during this period, while most small and mid-size firms shrunk – a result in line with what we observed in the BDS. Finally, this new method allowed us to decompose the overall variance of earned wages in their within- and between-firm components. We see that most of the increase in the overall variance is due to an increase in between-firm inequality. This result is in line with what was found by Barth et al. (2016) and Song et al. (2019). However, our calibration shows that while we do observe that between-firm inequality explains the bulk of the overall increase in wage inequality, changes in firm productivity by themselves can only explain a limited fraction of the increase in wage inequality and dispersion.

References

- Abowd, John A., Francis Kramarz, and David Margolis (1999). “High wage workers and high wage firms.” *Econometrica*, 67(2), pp. 251–333. doi:10.1111/1468-0262.00020.
- Acemoglu, Daron and David Autor (2011). “Skills, tasks and technologies: Implications for employment and earnings.” In *Handbook of Labor Economics*, volume 4, pp. 1043–1171. Elsevier. doi:10.1016/s0169-7218(11)02410-5.
- Albrecht, James and Susan Vroman (2002). “A matching model with endogenous skill requirements.” *International Economic Review*, 43(1), pp. 283–305. doi:10.1111/1468-2354.t01-1-00012.
- Barnichon, Regis (2010). “Building a composite help-wanted index.” *Economics Letters*, 109(3), pp. 175–178. doi:10.1016/j.econlet.2010.08.029.
- Barth, Erling, Alex Bryson, James C Davis, and Richard Freeman (2016). “It’s where you work: Increases in the dispersion of earnings across establishments and individuals in the united states.” *Journal of Labor Economics*, 34(S2), pp. S67–S97. doi:10.1086/684045.
- Bontemps, Christian, Jean-Marc Robin, and Gerard J Van den Berg (2000). “Equilibrium search with continuous productivity dispersion: Theory and nonparametric estimation.” *International Economic Review*, 41(2), pp. 305–358. doi:10.1111/1468-2354.00066.
- Bowlus, Audra J and Chris Robinson (2012). “Human capital prices, productivity, and growth.” *American Economic Review*, 102(7), pp. 3483–3515. doi:10.1257/aer.102.7.3483.
- Burdett, Kenneth and Dale T Mortensen (1998). “Wage differentials, employer size, and unemployment.” *International Economic Review*, 39(2), pp. 257–273. doi:10.2307/2527292.

- Cairo, Isabel and Tomaz Cajner (2017). “Human capital and unemployment dynamics: Why more educated workers enjoy greater employment stability.” *Economic Journal*, 128(609), pp. 652–682. doi:10.1111/ecoj.12441.
- Cajner, Tomaz, , and David Ratner (2016). “A cautionary note on the help wanted online data.” *FEDS Notes*, 2016(1795). doi:10.17016/2380-7172.1795.
- Dolado, Juan J, Marcel Jansen, and Juan F Jimeno (2009). “On-the-job search in a matching model with heterogeneous jobs and workers.” *Economic Journal*, 119(534), pp. 200–228. doi:10.1111/j.1468-0297.2008.02210.x.
- Dustmann, Christian, Johannes Ludsteck, and Uta Schönberg (2009). “Revisiting the german wage structure.” *Quarterly Journal of Economics*, 124(2), pp. 843–881. doi:10.1162/qjec.2009.124.2.843.
- Eeckhout, Jan and Roberto Pinheiro (2014). “Diverse organizations and the competition for talent.” *International Economic Review*, 55(3), pp. 625–664. doi:10.1111/iere.12065.
- Faggio, Giulia, Kjell G Salvanes, and John Van Reenen (2010). “The evolution of inequality in productivity and wages: Panel data evidence.” *Industrial and Corporate Change*, 19(6), pp. 1919–1951. doi:10.1093/icc/dtq058.
- Foster, Lucia, John Haltiwanger, and Chad Syverson (2008). “Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?” *American Economic Review*, 98(1), pp. 394–425. doi:10.1257/aer.98.1.394.
- Frazis, Harley J, Edwin L Robison, Thomas D Evans, and Martha A Duff (2005). “Estimating gross flows consistent with stocks in the cps.” *Monthly Labor Review*, 128, p. 3. URL <https://www.bls.gov/opub/mlr/2005/09/art1full.pdf>.
- Gottschalk, Peter and Robert A Moffitt (1994). “Welfare dependence: Concepts, measures, and trends.” *American Economic Review*, 84(2), pp. 38–42. URL <http://www.jstor.org/stable/2117798>.
- Hagedorn, Marcus, Tzuo Hann Law, and Iourii Manovskii (2017). “Identifying equilibrium models of labor market sorting.” *Econometrica*, 85(1), pp. 29–65. doi:10.3982/ecta11301.
- Handwerker, Elizabeth Weber and James R Spletzer (2016). “The role of establishments and the concentration of occupations in wage inequality.” In Solomon Polachek and Konstantinos Tatsiramos, editors, *Inequality: Causes and Consequences*, volume 43 of

- Research in Labor Economics*, pp. 167–193. Emerald Group Publishing Limited. doi:10.1108/s0147-912120160000043013.
- İmrohoroglu, Ayşe and Şelale Tüzel (2014). “Firm-level productivity, risk, and return.” *Management Science*, 60(8), pp. 2073–2090. doi:10.1287/mnsc.2013.1852.
- Jaeger, David A (1997). “Reconciling the old and new census bureau education questions: Recommendations for researchers.” *Journal of Business & Economic Statistics*, 15(3), pp. 300–309. doi:10.2307/1392334.
- Jolivet, Grégory (2009). “A longitudinal analysis of search frictions and matching in the us labor market.” *Labour Economics*, 16(2), pp. 121–134. doi:10.1016/j.labeco.2008.09.002.
- Kambourov, Gueorgui and Iourii Manovskii (2009). “Occupational mobility and wage inequality.” *Review of Economic Studies*, 76(2), pp. 731–759. doi:10.1111/j.1467-937x.2009.00535.x.
- Launov, Andrey (2006). *Structural Estimation of Search Equilibrium Models with Wage Posting*. phdthesis, Ludwig Maximilians University. URL https://edoc.ub.uni-muenchen.de/5110/1/Launov_Andrey.pdf.
- Lazear, Edward P. and Kathryn L. Shaw (2008). “Wage structure, raises and mobility: International comparison of the structure of wages within and between firms.” In *The Structure of Wages*, pp. 1–58. The University of Chicago Press. doi:10.7208/chicago/9780226470511.003.0001.
- Lemieux, Thomas (2006a). “Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill?” *American Economic Review*, 96(3), pp. 461–498. doi:10.1257/aer.96.3.461.
- Lemieux, Thomas (2006b). “Postsecondary education and increasing wage inequality.” *American Economic Review*, 96(2), pp. 195–199. doi:10.1257/000282806777211667.
- Lemieux, Thomas (2008). “The changing nature of wage inequality.” *Journal of Population Economics*, 21(1), pp. 21–48. doi:10.1007/s00148-007-0169-0.
- Lentz, Rasmus and Dale T Mortensen (2010). “Labor market models of worker and firm heterogeneity.” *Annual Review of Economics*, 2(1), pp. 577–602. doi:10.1146/annurev.economics.102308.124511.

- Lopes de Melo, Rafael (2018). “Firm wage differentials and labor market sorting: Reconciling theory and evidence.” *Journal of Political Economy*, 126(1), pp. 313–346. doi:10.1086/695505.
- Petrongolo, Barbara and Christopher A Pissarides (2001). “Looking into the black box: A survey of the matching function.” *Journal of Economic Literature*, 39(2), pp. 390–431. doi:10.1257/jel.39.2.390.
- Shimer, Robert (2012). “Reassessing the ins and outs of unemployment.” *Review of Economic Dynamics*, 15(2), pp. 127–148. doi:10.1016/j.red.2012.02.001.
- Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter (2019). “Firming Up Inequality.” *Quarterly Journal of Economics*, 134(1), pp. 1–50. doi:10.1093/qje/qjy025. URL <https://dx.doi.org/10.1093/qje/qjy025>.
- Tasci, Murat (2012). “The ins and outs of unemployment in the long run: Unemployment flows and the natural rate.” Working Paper 12-24, Federal Reserve Bank of Cleveland. doi:10.26509/frbc-wp-201224.
- Uren, Lawrence and Gabor Virag (2011). “Skill requirements, search frictions, and wage inequality.” *International Economic Review*, 52(2), pp. 379–406. doi:10.1111/j.1468-2354.2011.00632.x.
- U.S. Census Bureau (). “Current Population Survey, Merged Outgoing Rotation Group (CPS MORG).” <https://www.nber.org/morg/annual/>.
- U.S. Census Bureau (2019). “Quarterly Workforce Indicators (1998-2016) [Computer file].” <https://ledextract.ces.census.gov>. Longitudinal-Employer Household Dynamics Program [distributor].

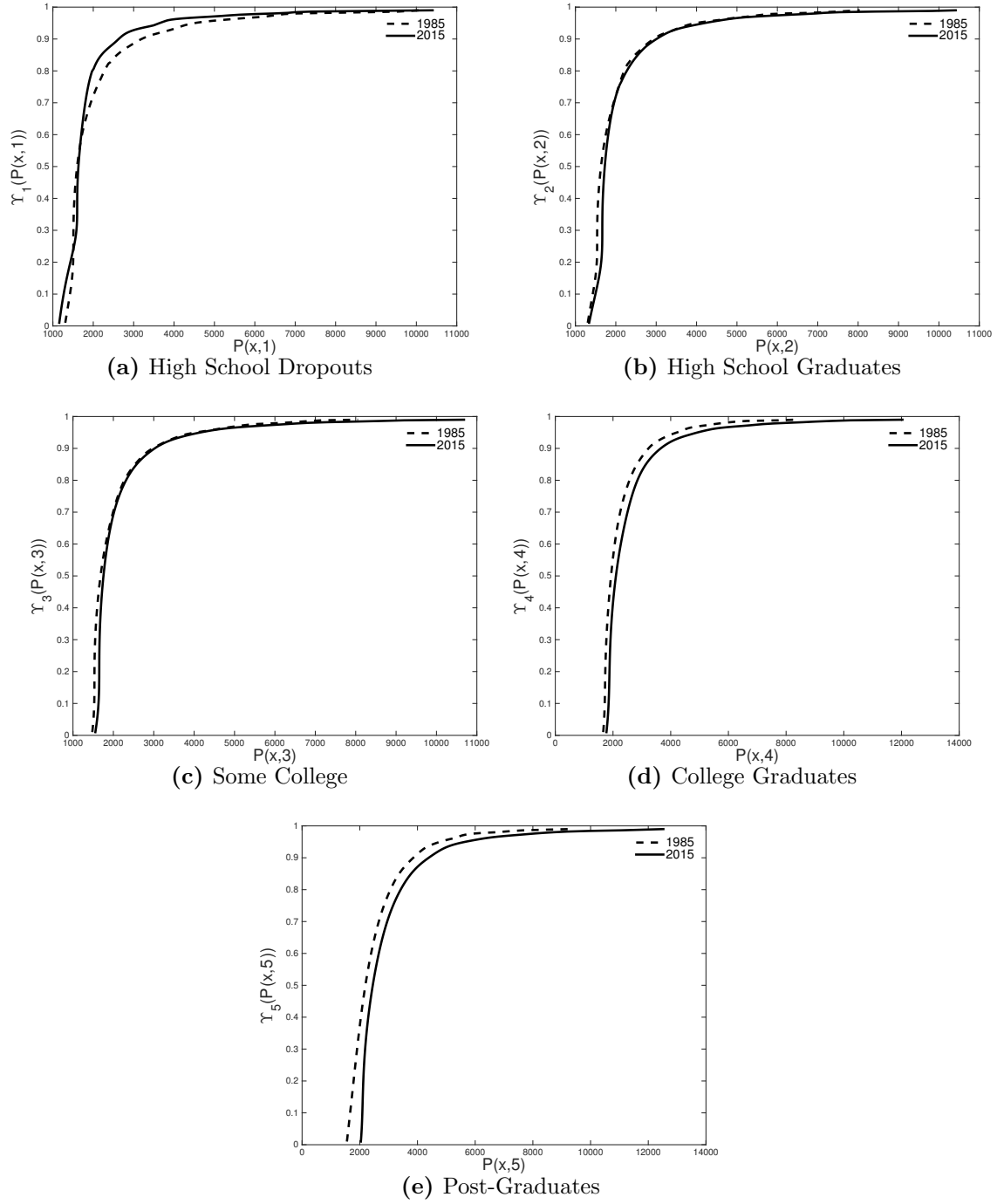


Figure 1: Comparison between Cumulative Distributions of Outputs for Different Skill Levels – 1985 vs. 2015

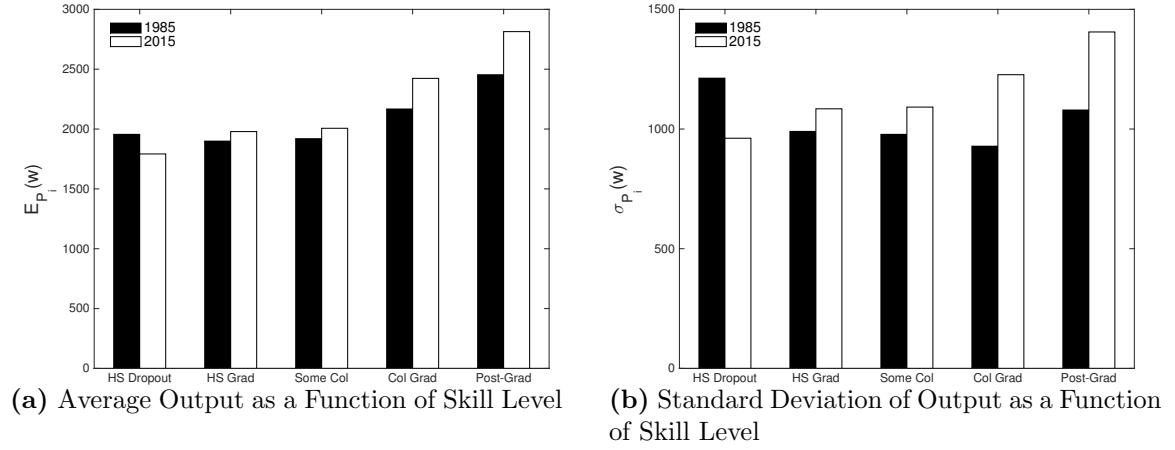


Figure 2: Average and Standard Deviation of Output – 1985 vs. 2015

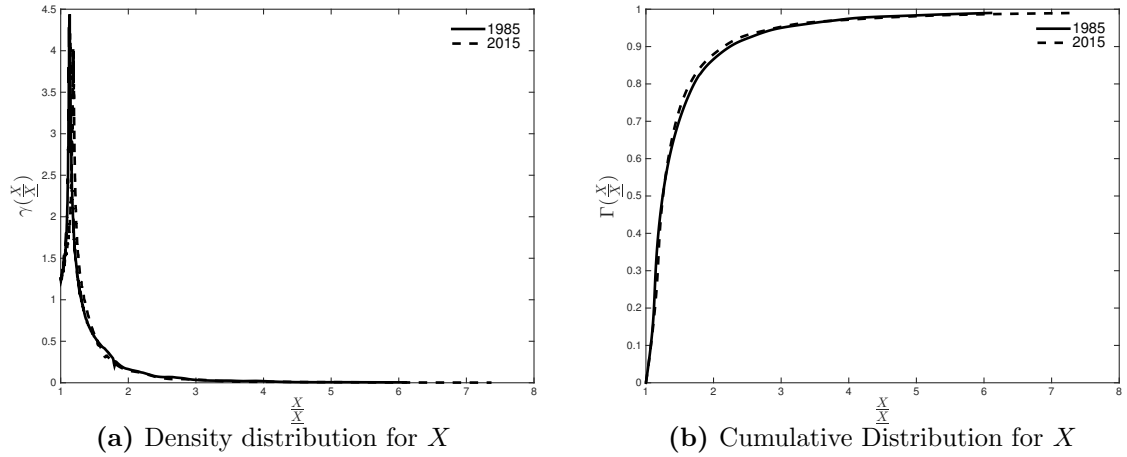


Figure 3: Estimated productivity Distributions – 1985 vs. 2015

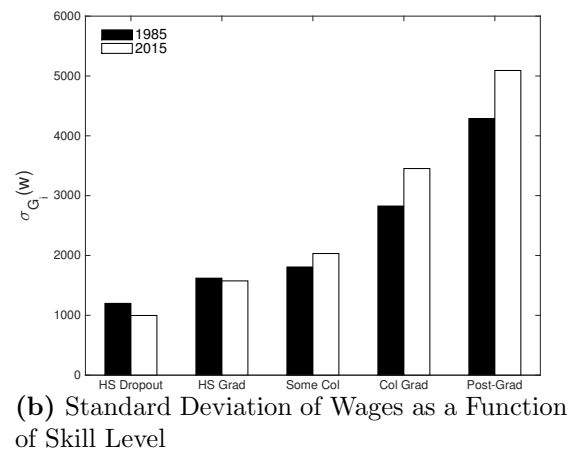
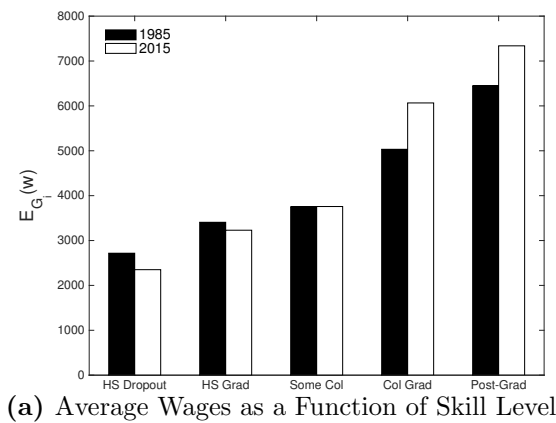
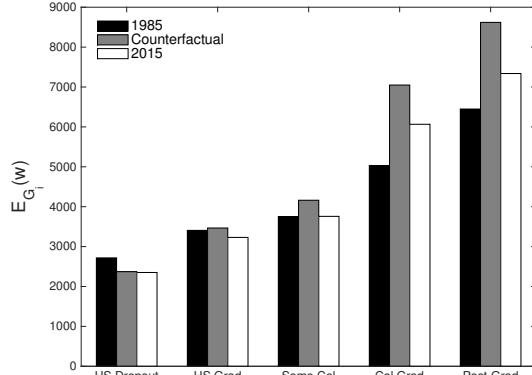
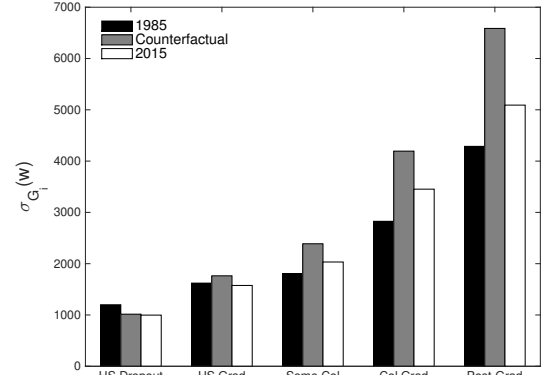


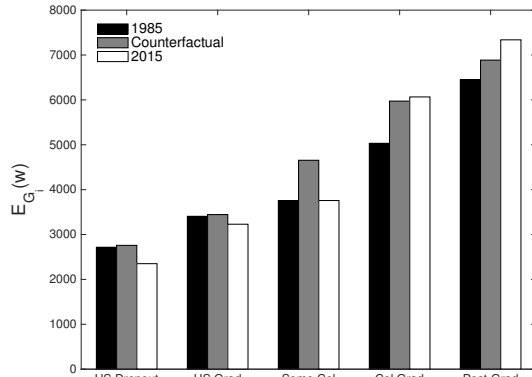
Figure 4: Average and Standard Deviation of Wages – 1985 vs. 2015



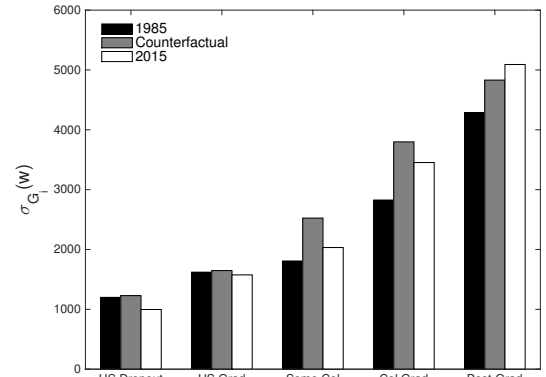
(a) Average Wage - Output Counterfactual



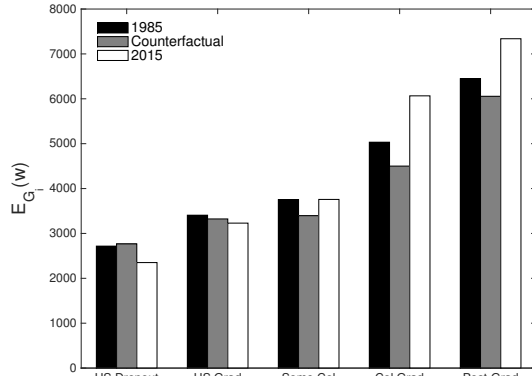
(b) Standard Deviation of Wages - Output Counterfactual



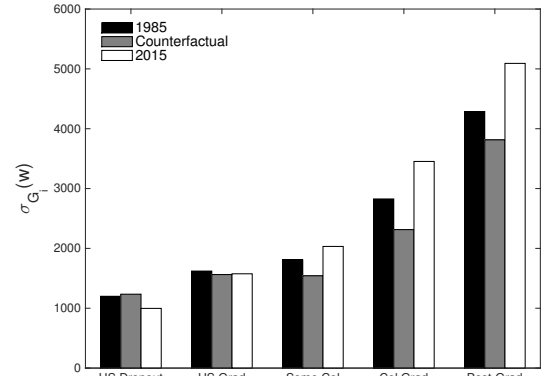
(c) Average Wage - Labor Frictions Counterfactual



(d) Standard Deviation of Wages - Labor Frictions Counterfactual

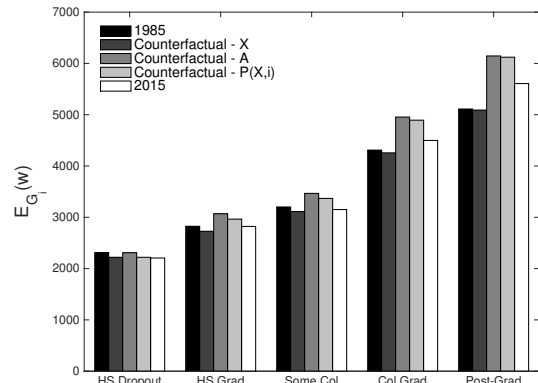


(e) Average Wage - Market Composition Counterfactual

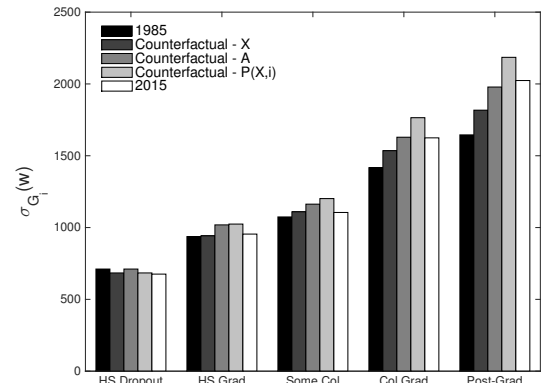


(f) Standard Deviation of Wages - Market Composition Counterfactual

Figure 5: Average and Standard Deviation of Wages - Counterfactuals



(a) Average Wages as a Function of Skill Level



(b) Standard Deviation of Wages as a Function of Skill Level

Figure 6: Average and Standard Deviation of Wages - Parametric Counterfactuals

A Appendix - For online Publication

A.1 Market Equilibrium Definition

Definition 1 Consider an economy with heterogeneous workers and firms, in which the firm-productivity distribution $\Gamma(x)$ is continuous with no mass points. A market equilibrium is a set of $\mathcal{I}$ triples – one for each skill i submarket – $(R(i), \underline{x}^*(i))$, such that

1. Only firms with productivity greater than $\underline{x}^*(i)$ are active, and the distribution of firm productivity in the market is given by

$$\frac{\Gamma(x) - \Gamma(\underline{x}^*(i))}{\bar{\Gamma}(\underline{x}^*(i))}$$

2. The distribution of wage offers in submarket i is given by $F_i(\cdot)$. In particular, the profit-maximizing wage of a productivity- x firm, given other firms' and workers' strategies, is given by:

$$K(x; i) = \arg \max_w \{ \pi(x; i) | b(i) \leq w \leq p(x, i) \}$$

where $\pi(x; i)$ is defined in equation (10).

3. $R(i)$ is workers' best response to firms' aggregate behavior. Given that search efficiency is unaffected by employment status, $R(i) = b(i)$.

A.2 Economy-wide Wage Distributions

In this section, we present the economy-wide distributions of offered and earned wages. Our goal is to show how the wage distributions per skill type interact in order to build their economy-wide counterpart and, consequently, how changes in the composition of the labor force may affect the economy-wide wage distribution. We start with the distribution of posted wages followed by the distribution of earned wages.

AGGREGATED DENSITY OF POSTED WAGES

First, let's be very precise about the densities for each skill level i :

$$f_i(\tilde{w}) = \begin{cases} \frac{\gamma(K^{-1}(\tilde{w}, i))}{\bar{\Gamma}(K^{-1}(\underline{w}_i, i)) \frac{\partial K(K^{-1}(\tilde{w}, i), i)}{\partial K^{-1}(\tilde{w}, i)}} & \text{if } \tilde{w} \in [\underline{w}_i, \bar{w}_i] \\ 0 & \text{otherwise} \end{cases} \quad (21)$$

Then, notice that because not all firms offer jobs at all skill levels, we need to weight the

wage distributions per skill by the measure of firms that actively post wages at that particular skill level. Once weights are properly included, the aggregated cumulative distribution of offered wages in the economy is given by

$$F(\tilde{w}) = \frac{\sum_{i=1}^{\mathcal{I}} \bar{\Gamma}(K^{-1}(\underline{w}_i, i)) F_i(\tilde{w})}{\sum_{i=1}^{\mathcal{I}} \bar{\Gamma}(K^{-1}(\underline{w}_i, i))} \quad (22)$$

Consequently, the density function associated with this cumulative distribution is given by

$$f(\tilde{w}) = \frac{\sum_{i=1}^{\mathcal{I}} \bar{\Gamma}(K^{-1}(\underline{w}_i, i)) f_i(\tilde{w})}{\sum_{i=1}^{\mathcal{I}} \bar{\Gamma}(K^{-1}(\underline{w}_i, i))} \quad (23)$$

Finally, the average posted wage for the overall economy is given by

$$E_F[w] = \int_{\underline{w}}^{\bar{w}} \tilde{w} f(\tilde{w}) d\tilde{w} = \frac{\sum_{i=1}^{\mathcal{I}} \bar{\Gamma}(K^{-1}(\underline{w}_i, i)) E_{F_i}[w]}{\sum_{i=1}^{\mathcal{I}} \bar{\Gamma}(K^{-1}(\underline{w}_i, i))} \quad (24)$$

AGGREGATED DENSITY OF EARNED WAGES

In this case, we focus on the wage actually earned by workers. Consequently, instead of tracking the measure of firms posting a job at a given skill level, we need to track the measure of employed workers at each skill level as a proportion of all employed workers. Based on the cumulative distribution of earned wages per skill presented in equation (6), we have that the aggregated cumulative distribution and density function of earned wages are given by

$$G(\tilde{w}) = \frac{\sum_{i=1}^{\mathcal{I}} G_i(\tilde{w})(m_i - u_i)}{\sum_{i=1}^{\mathcal{I}} (m_i - u_i)} \quad \text{and} \quad g(\tilde{w}) = \frac{\sum_{i=1}^{\mathcal{I}} g_i(\tilde{w})(m_i - u_i)}{\sum_{i=1}^{\mathcal{I}} (m_i - u_i)} \quad (25)$$

respectively. Consequently, the average aggregate wage in this economy is

$$E_G[w] = \int_{\underline{w}}^{\bar{w}} w g(w) dw = \frac{\sum_{i=1}^{\mathcal{I}} (m_i - u_i) E_{G_i}[w]}{\sum_{i=1}^{\mathcal{I}} (m_i - u_i)} \quad (26)$$

where $E_{G_i}[w]$ can be derived after manipulating equations (15) and (16) as well as implementing a change of variables taking into account that $w = K(x, i)$:

$$E_{G_i}[w] = \int_{\underline{x}(i)}^{\bar{x}} K(x', i) \frac{(1 + \kappa_i) \bar{\Gamma}(\underline{x}^*(i)) \gamma(x')}{[\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x')]^2} dx' \quad (27)$$

Finally, let's consider the economy-wide variance of earned wages:

$$Var_G(w) = \int_{\underline{w}}^{\overline{w}} (w - E_G(w))^2 g(w) dw \quad (28)$$

Again, after some algebra and a change of variables considering $w = K(x, i)$, along with defining $M - U = \sum_{i=1}^{\mathcal{I}} (m_i - u_i)$, after substituting equation (7) we have:

$$Var_G(w) = \frac{1}{M - U} \int_{\underline{x}}^{\overline{x}} \sum_{i=1}^{\mathcal{I}} (K(x, i) - E_G(w))^2 l_x(i) \gamma(x) dx \quad (29)$$

A.3 Wage Variance Decompositions

As we described in the introduction, one of the most important questions that our model can address pertains to the source of wage inequality across workers in the overall economy. In particular, the decomposition of the source of wage dispersion in terms of the within- versus between-firm components is one that has been the focus of most of the recent empirical literature. In particular, according to Lazear and Shaw (2008), the total variance in wages, σ^2 , is given by

$$\sigma^2 = \sum_{j=1}^J p_j \sigma_j^2 + \sum_{j=1}^J p_j (\bar{w}_j - \bar{\bar{w}})^2 \quad (30)$$

The first term on the RHS of equation (30) is the within-firm component of the variance. Notice that p_j is the share of workers in the economy employed in firm j , while σ_j^2 is the variance of wages in firm j . The second term on the RHS of equation (30) represents the between-firm component of the wage variance. In this expression, $\bar{w}_j$ is the mean wage in firm j , and $\bar{\bar{w}}$ is the mean wage in the economy.

In this section, we apply their decomposition to our model. Let's start with the within-firm component. Assume that we partition the interval $[\underline{x}, \overline{x}]$ in N intervals of length Δ . Then, we have $x_{i+1} = x_i + \Delta$. Moreover, the measure of type x_{i+1} firms is given by $\Gamma(x_{i+1}) - \Gamma(x_i)$, while the share of employed workers in the economy at type x_{i+1} is $\frac{S(x_{i+1})}{M-U}$. Then, we have that $p_j \sigma_j^2$ can be rewritten as $Var_{\phi_{x_{i+1}}}(w) \frac{S(x_{i+1})}{M-U} \Gamma(x_{i+1}) - \Gamma(x_i)$. Multiplying and dividing by Δ and adding across x_S , we have

$$\sum_{i=1}^N Var_{\phi_{x_{i+1}}}(w) \frac{S(x_{i+1})}{M-U} \frac{[\Gamma(x_{i+1}) - \Gamma(x_i)]}{\Delta} \Delta$$

Taking $\Delta \rightarrow 0$, we have

$$\text{Within-firm component} = \int_{\underline{x}}^{\bar{x}} \frac{S(x)}{M-U} \text{Var}_{\phi_x}(w) \gamma(x) dx \quad (31)$$

Following the same procedure for the between-firm component, we obtain

$$\text{Between-firm component} = \int_{\underline{x}}^{\bar{x}} \frac{S(x)}{M-U} (E_{\phi_x}(w) - E_G(w))^2 \gamma(x) dx \quad (32)$$

Therefore, the entire decomposition can be rewritten as

$$\text{Var}_G(w) = \int_{\underline{x}}^{\bar{x}} \frac{S(x)}{M-U} \text{Var}_{\phi_x}(w) \gamma(x) dx + \int_{\underline{x}}^{\bar{x}} \frac{S(x)}{M-U} (E_{\phi_x}(w) - E_G(w))^2 \gamma(x) dx \quad (33)$$

where $E_G(w)$ is presented in equation (26), while $\text{Var}_{\phi_x}(w)$ and $E_{\phi_x}(w)$ are defined by

$$E_{\phi_x}[w] = \frac{\sum_{i=1}^{\mathcal{I}} K(x, i) l_x(i)}{S(x)} \quad (34)$$

and

$$\text{Var}_{\phi_x}[w] = \frac{\sum_{i=1}^{\mathcal{I}} (K(x, i) - E_{\phi_x}[w])^2 l_x(i)}{S(x)} \quad (35)$$

Moreover, we can adapt Lazear and Shaw (2008)'s decomposition in order to decompose the overall wage variance in terms of the within and between educational groups. Using the same logic, we obtain the following decomposition:

$$\sigma^2 = \sum_{i=1}^I m_i^e (\sigma_i^e)^2 + \sum_{i=1}^I m_i^e (\bar{w}_i^e - \bar{\bar{w}})^2 \quad (36)$$

Similar to the decomposition in terms of within and between firms, the first term in the RHS of equation (36) represents the within-skill component of the overall wage inequality and the second term in the RHS is the between-skill component. m_i^e is the fraction of the employed labor force with skill level i , σ_i^e is the standard deviation of wages for workers of skill i , $\bar{w}_i^e$ is the average wage for employed workers with skill level i , and $\bar{\bar{w}}$ is the mean wage in the economy.

Mapping this decomposition to our model, we have $\bar{w}_i^e = E_{G_i}[w]$, where $E_{G_i}[w]$ is given

by equation (27) and $(\sigma_i^e)^2$ is given by

$$(\sigma_i^e)^2 \equiv Var_{G_i}(w) = (1 + \kappa_i) \bar{\Gamma}(\underline{x}^*(i)) \int_{\underline{x}^*(i)}^{\bar{x}} \left[\frac{K(x, i) - E_{G_i}[w]}{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x)} \right]^2 \gamma(x) dx \quad (37)$$

Finally, the fraction of employed workers who have skill level i in our model is given by

$$m_i^e = \frac{\frac{\kappa_i m_i}{1 + \kappa_i}}{\sum_{j=1}^{\mathcal{I}} \frac{\kappa_j m_j}{1 + \kappa_j}} \quad (38)$$

Consequently, the decomposition of overall wage variance in terms of within- and between-skill components is

$$\begin{aligned} \text{Within-skill component} &= \sum_{i=1}^{\mathcal{I}} m_i^e Var_{G_i}(w) & \text{Between-skill component} &= \sum_{i=1}^{\mathcal{I}} m_i^e (E_{G_i}[w] - E_G[w])^2 \end{aligned} \quad (39)$$

where m_i^e , $Var_{G_i}(w)$, $E_{G_i}[w]$, and $E_G[w]$ are given by equations (38), (37), (27), and (26), respectively.

A.4 Empirical Evidence: Quarterly Workforce Indicators (QWI)

We present summary statistics and empirical patterns by firm-size that are in line with the results we have obtained from our model. While this evidence is mostly suggestive, it offers an initial indication in favor of our model, before we present our calibration exercise. Our empirical approach relies on the identification of firm-level properties using our model and only data from the worker-level observations conditional on skill. This section highlights that implications from our model are nevertheless consistent with the firm-level observations, albeit at an aggregate level.

The database for our analysis in this section comes from the Quarterly Workforce Indicators (henceforth QWI).¹⁷ The QWI provides local labor statistics by industry, worker demographics, employer age, and size. The source data for the QWI are the Longitudinal Employer-Household Dynamics (LEHD), which is a linked employer-employee longitudinal database covering over 95 percent of the US private sector jobs. While the QWI reports education information only for workers who are age 25 and up, this still suits our analysis well, once we focus on workers who completed their education. Consequently, our analysis considers just the universe of employed workers age 25 and up. Moreover, while the QWI provides data at the establishment level as well as at the firm level, we focus on the

¹⁷Data reference is U.S. Census Bureau (2019).

firm-level information, since it is more closely related to the model’s analysis. Firm size in the QWI is defined at the national level and it refers to the national employment size of the firm on March 12th of the previous year for existing firms. For new firms, firm size is measured as the current year’s March employment (or the employment in the first month of positive employment if born after March). Firm size is reported only for private-sector firms. Finally, our measures of employment and earnings hinge on the concept of stable (or full-quarter) employment, which focuses on individuals that receive earnings from the same employer for at least three consecutive quarters. Consequently, stable employment suggests that an employee has an ongoing relationship with the employer throughout the quarter.

Recall from our Theorem 1, that there is a tight link between a firm’s productivity and its size. While the aggregate data from the QWI are not suitable to show this point, there is a large literature that documents the joint distribution of firm productivity and its size, both in the US (Foster, Haltiwanger, and Syverson (2008)) and elsewhere (Abowd, Kramarz, and Margolis (1999)). Thus, we take this tight link between a firm’s productivity and its size as given and focus on the aggregate evidence related to the other relevant properties implied by Theorem 1.

Results displayed in Table A.1 present the average number of workers per educational group per firm size. These averages are calculated by dividing the number of workers per education group employed at firms within the size class. The QWI data allow for four distinct education levels for workers: less than high school (LHS), high school (HS), some college education (SC) and college educated and above (BA+). Average sizes are normalized by the number of firms in each firm size class, giving us the size of a typical firm in each size class. Overall, results presented in Table A.1 are in line with the equilibrium characteristics presented in Theorem 1. In particular, larger firms hire more at all educational groups.¹⁸ Results for other years (1993–2014) are qualitatively the same as the ones for 2015 so we omitted them here, but they are available upon request.

Similarly, Table A.2 displays the average monthly earnings of workers at different firm size categories by their education level. Consistent with our Theorem 1, average earnings increase by education level for each firm-size group. Moreover, a typical worker in each education level earns more at larger firms. These aggregate patterns are all consistent with the model implications.

¹⁸The model presented in this paper is an example of what Lopes de Melo (2018) calls a “piece-rate” model. As Lopes de Melo (2018) shows in Sections 3.1 and 3.2, models don’t present sorting in equilibrium, once reservation strategies and meeting rates are independent of firm type. Consequently, the differences in κ ’s are likely a necessary condition for Proposition 1. Nevertheless, in Lopes de Melo (2018)’s calibration κ is also increasing with education.

Table A.1: Avg. No. of workers across firm sizes - 2015

	1 to 19	20 to 49	50 to 249	250 to 499	500+
LHS	0.58	3.47	11.66	39.61	303.99
HS	1.09	6.43	21.48	73.46	604.70
SC	1.19	6.84	23.41	81.55	702.71
BA+	0.98	5.28	18.30	65.14	620.39

Table A.2: Avg. Monthly Earnings - 2015

	1 to 19	20 to 49	50 to 249	250 to 499	500+
LHS	2534	2961	3183	3289	3518
HS	2905	3437	3738	3881	4196
SC	3331	4000	4375	4583	5091
BA+	5067	6371	7141	7730	8925

Another important result from Theorem 1 states that larger firms hire disproportionately more workers with higher educational attainment. This implies a concentration of higher-education workers in large firms. Hence, we also want to highlight this concentration pattern in the aggregate QWI data. In fact, in order to take into account the economy-wide distribution of workers across different educational groups, we present two alternative measures of concentration across education groups. The first one is a measure based on the location quotient measures (LQ), which is calculated as follows:

$$LQ_i = \frac{\left(\begin{array}{c} \text{fraction of firm-size group} \\ \text{employees in education group } i \end{array} \right)}{\left(\begin{array}{c} \text{fraction of employed workers} \\ \text{in the economy in education group } i \end{array} \right)} \quad (40)$$

Consequently, if $LQ_i > 1$ we have that the firm's labor force is more concentrated in education group i than in the economy as a whole. We present results for the QWI database in Figure A.2. Panels A.2(a) and A.2(b) present the results focusing on firm sizes and education groups, respectively. As we can see, compared to the economy-wide distribution, low-education groups are overrepresented in small firms, while high-education workers are overrepresented in large firms. As with the evidence presented in Figure A.1, the patterns observed in Figure A.2 also corroborate our findings from Theorem 1.

Finally, we create Herfindahl-Hirschman Indexes (henceforth HHIs) to measure the dis-

tribution of employment by education group within firms.¹⁹ In particular, we compute the following indexes:

$$HHI_{firm\ j} = \sum_i \left(\frac{\text{share of education group } i \text{ in}}{\text{firm-size group } j \text{ total employment}} \right)^2 \quad (41)$$

$$HHI_{educ\ i} = \sum_l \left(\frac{\text{share of firm-size group } i \text{ in}}{\text{education group } i \text{ total employment}} \right)^2 \quad (42)$$

Figure A.3 presents our results for $HHI_{firm\ j}$ and $HHI_{educ\ i}$ across several years in the database. As expected, higher values of HHI indicate higher degrees of concentration. As we can see from panel A.3(a), large firms (more than 500 workers) tend to be more concentrated than smaller ones, even though the degree of concentration has decreased over time. Similarly, panel A.3(b) shows that the distribution of high-education workers across firm size groups is more concentrated. Moreover, the concentration of workers across firm-size groups has actually increased over time. A possible explanation for the different inter-temporal patterns presented in panels A.3(a) and A.3(b) is shown in Figure A.4, in which we use data from the Business Dynamics Statistics (BDS) in order to calculate the share of the employed labor force across the different firm-size groups over time. As we can see, large firms (500+) have absorbed an increasing share of the employed labor force, possibly explaining the increasing concentration at the $HHI_{educ\ i}$ measure.

A brief analysis of the aggregate patterns in the QWI data confirms that our model is consistent with firm-level observations. In the rest of the paper, we will solely rely on worker level outcomes (employment, earnings) to calibrate our model and conduct our counterfactual experiments to uncover the main channels responsible for rising wage inequality in the US.

A.5 Proofs

Proof of Lemma 1

Proof. Toward a contradiction, assume that $j > i$ but $\underline{x}(i) > \underline{x}(j)$. From the increasing differences property of supermodular functions, we have:

$$p(\underline{x}(i), j) - p(\underline{x}(i), i) > p(\underline{x}(j), j) - p(\underline{x}(j), i) \quad (43)$$

¹⁹We follow procedures presented in Handwerker and Spletzer (2016).

Rearranging it, we have:

$$p(\underline{x}(i), j) - p(\underline{x}(j), j) > p(\underline{x}(i), i) - p(\underline{x}(j), i) \quad (44)$$

Since $\frac{\partial p(\underline{x}, i)}{\partial x} > 0$, we have that the RHS of equation (44) is larger than zero. As a result:

$$p(\underline{x}(i), j) > p(\underline{x}(i), i) \Rightarrow p(\underline{x}(i), j) > b(j) \quad (45)$$

which is a contradiction. Consequently, if $j > i$, we must have $\underline{x}(j) \geq \underline{x}(i)$, concluding the proof. ■

Calculations for equation (13)

Now, focusing on the boundary condition, notice that if $\underline{x}^*(i) = \underline{x}(i)$, we have that $\pi(K(\underline{x}(i); i); i) = 0$ and the boundary condition is also zero. Similarly, if $\underline{x}^*(i) = \underline{x}$, we have that:

$$l(K(\underline{x}; i); i) = \frac{\kappa_i m_i}{[1 + \kappa_i]^2} \quad (46)$$

Consequently, notice that

$$\int_{\underline{x}(i)}^{\underline{x}} \frac{\partial p(x'; i)}{\partial x'} \frac{\kappa_i m_i}{[1 + \kappa_i]^2} dx' = [p(\underline{x}, i) - b(i)] \frac{\kappa_i m_i}{[1 + \kappa_i]^2} \quad (47)$$

Following an argument similar to the one presented in Burdett and Mortensen (1998), Section 4 (which we will verify in our presentation later), we have that $K(\underline{x}; i) = b(i)$. As a result, regardless of whether $\underline{x}^*(i) = \underline{x}(i)$ or $\underline{x}^*(i) = \underline{x}$, we have:

$$\pi(x; i) = \int_{\underline{x}(i)}^x \frac{\partial p(x'; i)}{\partial x'} \frac{\kappa_i m_i}{\bar{\Gamma}(\underline{x}^*(i)) \left[1 + \kappa_i \frac{\bar{\Gamma}(x')}{\bar{\Gamma}(\underline{x}^*(i))}\right]^2} dx' \quad (48)$$

which is equation (13) presented in the main text. □

Proof of Lemma 2

Proof.

Since $x > \tilde{x} \Rightarrow \bar{\Gamma}(x) < \bar{\Gamma}(\tilde{x})$. Then, from equation (14), we have:

$$K(x; i) - K(\tilde{x}; i) = \left\{ \begin{aligned} & p(x, i) - p(\tilde{x}, i) \\ & - \int_{\underline{x}(i)}^{\tilde{x}} \frac{\partial p(x', i)}{\partial x'} \left[\frac{\left(\frac{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x)}{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x')} \right)^2}{\left(\frac{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(\tilde{x})}{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x')} \right)^2} - 1 \right] dx' \\ & - \int_{\tilde{x}}^x \frac{\partial p(x', i)}{\partial x'} \left(\frac{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x)}{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x')} \right)^2 dx' \end{aligned} \right\} \quad (49)$$

Since $\bar{\Gamma}(x) < \bar{\Gamma}(\tilde{x})$, we have that:

$$\left[\left(\frac{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x)}{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x')} \right)^2 - \left(\frac{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(\tilde{x})}{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x')} \right)^2 \right] < 0 \quad (50)$$

Similarly, since $x' \leq x$, we have that:

$$\left(\frac{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x)}{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x')} \right)^2 \leq 1 \quad (51)$$

with equality only if $x' = x$, once $\gamma(\cdot)$ is continuous with no mass points.

Consequently, we have that:

$$K(x; i) - K(\tilde{x}; i) > \int_{\underline{x}(i)}^{\tilde{x}} \frac{\partial p(x', i)}{\partial x'} \left[\frac{\left(\frac{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(\tilde{x})}{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x')} \right)^2}{\left(\frac{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x)}{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x')} \right)^2} - 1 \right] dx' > 0 \quad (52)$$

Establishing that $K(x; i) > K(\tilde{x}; i)$. ■

Corollary A.1 *The second-order condition of the firm's profit maximization problem is satisfied.*

Proof. From F.O.C. in equation (11), we obtain the following second-order condition (henceforth S.O.C.)

$$2\kappa_i f'_i(w)(p(x, i) - w) - \kappa_i f_i(w) \quad (53)$$

Substituting equation (11) and rearranging, we have that the S.O.C. is satisfied if:

$$\kappa_i f_i(w)^2 - f'_i(w) [1 + \kappa_i \bar{F}_i(w)] > 0 \quad (54)$$

Without loss of generality, assume a skill level $i \in \{1, \dots, \mathcal{I}\}$. Consider two levels of firm productivity x_1 and $x_2 \in [\underline{x}^*(i), \bar{x}]$, with $x_2 > x_1$. From F.O.C. in equation (11), after a

couple of manipulations, we have:

$$p(x, i) = w + \frac{[1 + \kappa_i \bar{F}_i(w)]}{2\kappa_i f_i(w)} \quad (55)$$

Subtracting the F.O.C. for x_1 from the F.O.C. for x_2 , we have

$$p(x_2, i) - p(x_1, i) = w_2 - w_1 + \frac{[1 + \kappa_i \bar{F}_i(w_2)]}{2\kappa_i f_i(w_2)} - \frac{[1 + \kappa_i \bar{F}_i(w_1)]}{2\kappa_i f_i(w_1)} \quad (56)$$

Notice that the LHS of equation (56) is positive, since $\frac{\partial p(x, i)}{\partial x} > 0$. As a result the RHS of equation (56) must also be positive. This result is true for any x_1 and $x_2 \in [\underline{x}^*(i), \bar{x}]$, with $x_2 > x_1$. Moreover, from Lemma 2 we have that $w_2 = K(x_2; i) > K(x_1; i) = w_1$. Consequently, we have that the RHS. of equation (56) must be strictly increasing in w . Therefore

$$\frac{\partial \left\{ w + \frac{[1 + \kappa_i \bar{F}_i(w)]}{2\kappa_i f_i(w)} \right\}}{\partial w} = \frac{\kappa_i f_i(w)^2 - f_i'(w) [1 + \kappa_i \bar{F}_i(w)]}{2\kappa_i f_i(w)^2} > 0 \quad (57)$$

Consequently, we must have that:

$$\kappa_i f_i(w)^2 - f_i'(w) [1 + \kappa_i \bar{F}_i(w)] > 0 \quad (58)$$

Implying that the S.O.C. is satisfied. ■

Proof of Lemma 3

Proof. Consider two productivity levels x and x' with $x > x'$ and a skill level i . If $x < \underline{x}(i)$, we have that $l_i(x) = l_i(x') = 0$. Differently, if $x' < \underline{x}(i) < x$, we have that $0 = l_i(x') < l_i(x)$. Finally, consider the case in which $x' > \underline{x}(i)$. From equation (7) and the definition that $l_i(x) = l(w(x, i), i)$, we have:

$$\frac{l_i(x)}{l_i(x')} = \left(\frac{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x')}{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x)} \right)^2 > 1 \quad (59)$$

Since $\bar{\Gamma}(x') > \bar{\Gamma}(x)$. Consequently, for any $x, x' \in [\underline{x}, \bar{x}]$ and $x > x'$, we have that $l_i(x) \geq l_i(x')$. ■

Proof of Proposition 1

Proof. From Lemma 1, we know the support of skills for a type x firm is $\{1, \dots, \mathbf{I}(x)\}$ while the support of skills hired by firm x' is $\{1, \dots, \mathbf{I}(x')\}$ with $\mathbf{I}(x') \leq \mathbf{I}(x)$. Consequently, if $\mathbf{I}(x') \leq i \leq \mathbf{I}(x)$ we have that $\Phi_{x'}(i) = 1$ and $\Phi_x(i) \leq 1$ and consequently $\Phi_x(i) \leq \Phi_{x'}(i)$.

Similarly, if $i > \max\{\mathbf{I}(x), \mathbf{I}(x')\}$, we must have $\Phi_{x'}(i) = \Phi_x(i) = 1$. Let's now consider the case in which $i < \mathbf{I}(x')$. In this case, both firms hire this particular skill. In particular, consider the skill level $\tilde{I} < \mathbf{I}(x')$. In this case, we have:

$$\Phi_x(\tilde{I}) - \Phi_{x'}(\tilde{I}) = \frac{\sum_{i=1}^{\tilde{I}} l_i(x)}{\sum_{j=1}^{\mathbf{I}(x)} l_j(x)} - \frac{\sum_{i=1}^{\tilde{I}} l_i(x')}{\sum_{j=1}^{\mathbf{I}(x')} l_j(x')} \quad (60)$$

Rearranging it, we have:

$$\Phi_x(\tilde{I}) - \Phi_{x'}(\tilde{I}) = \frac{1}{S(x)S(x')} \left\{ \sum_{i=1}^{\tilde{I}} l_i(x) \sum_{j=1}^{\mathbf{I}(x')} l_j(x') - \sum_{i=1}^{\tilde{I}} l_i(x') \sum_{j=1}^{\mathbf{I}(x)} l_j(x) \right\} \quad (61)$$

Simplifying it, we have:

$$\Phi_x(\tilde{I}) - \Phi_{x'}(\tilde{I}) = \frac{1}{S(x)S(x')} \left\{ \begin{aligned} &\sum_{i=1}^{\tilde{I}} \sum_{j=\tilde{I}+1}^{\mathbf{I}(x')} l_i(x) \left[l_j(x') - \frac{l_i(x')}{l_i(x)} l_j(x) \right] \\ &- \sum_{i=1}^{\tilde{I}} \sum_{j=\mathbf{I}(x)+1}^{\mathbf{I}(x')} l_i(x') l_j(x) \end{aligned} \right\} \quad (62)$$

As a result, a sufficient condition for $\Phi_x(\tilde{I}) - \Phi_{x'}(\tilde{I}) \leq 0$ is:

$$l_j(x') - \frac{l_i(x')}{l_i(x)} l_j(x) \leq 0, \quad \forall i \in \{1, \dots, \tilde{I}\} \text{ and } \forall j \in \{\tilde{I} + 1, \mathbf{I}(x')\} \quad (63)$$

Rearranging equation (63), we have:

$$\frac{l_j(x)}{l_j(x')} \geq \frac{l_i(x)}{l_i(x')}, \quad \forall i \in \{1, \dots, \tilde{I}\} \text{ and } \forall j \in \{\tilde{I} + 1, \mathbf{I}(x')\} \quad (64)$$

Replacing equation (59) into equation (64), we have:

$$\left(\frac{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x')}{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x)} \right)^2 \geq \left(\frac{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x')}{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x)} \right)^2 \quad (65)$$

According to Lemma 3, we have $l_j(x) > l_j(x')$ and $l_i(x) > l_i(x')$. So the above inequality can be simplified as

$$\left(\frac{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x')}{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x)} \right) \geq \left(\frac{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x')}{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x)} \right) \quad (66)$$

Rearranging it and simplifying it, we have:

$$\left[\bar{\Gamma}(x') - \bar{\Gamma}(x)\right] \left[\kappa_j \bar{\Gamma}(\underline{x}^*(i)) - \kappa_i \bar{\Gamma}(\underline{x}^*(j))\right] \geq 0 \quad (67)$$

Since $x > x'$, we have that $\left[\bar{\Gamma}(x') - \bar{\Gamma}(x)\right] \geq 0$. Similarly, since $j > i$, from Lemma 1 we have that $\underline{x}^*(j) \geq \underline{x}^*(i)$. Consequently $\left[\kappa_j \bar{\Gamma}(\underline{x}^*(i)) - \kappa_i \bar{\Gamma}(\underline{x}^*(j))\right] \geq 0$. Finally, given $\kappa_j \geq \kappa_i$, we have that the inequality presented in equation (63) is satisfied and $\Phi_x(\tilde{I}) - \Phi_{x'}(\tilde{I}) \leq 0$.

In summary, $\Phi_x(i) - \Phi_{x'}(i) \leq 0, \forall i \in \{1, \dots, \mathcal{I}\}$. Consequently, $\Phi_x(i)$ dominates stochastically in first order $\Phi_{x'}(i)$. ■

Proof of Lemma 4

Proof.

From the integral on the RHS of equation (14), notice that, given that $x' \leq x$, we have that $\left(\frac{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x)}{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x')}\right) \leq 1$, with equality only at the limit $x' = x$. Moreover, consider $j < i$. From Lemma 1, we have $\underline{x}(j) \leq \underline{x}(i) \Rightarrow \bar{\Gamma}(\underline{x}(j)) \geq \bar{\Gamma}(\underline{x}(i))$. Then, we can easily show that, given $\kappa_i \geq \kappa_j$, we have:

$$\left(\frac{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x)}{\bar{\Gamma}(\underline{x}^*(i)) + \kappa_i \bar{\Gamma}(x')}\right) \leq \left(\frac{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x)}{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x')}\right) \quad (68)$$

As a result, from equation (14), we have:

$$K(x; i) \geq p(x, i) - \int_{\underline{x}(i)}^x \frac{\partial p(x', i)}{\partial x'} \left(\frac{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x)}{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x')}\right)^2 dx' \quad (69)$$

Consequently:

$$K(x; i) - K(x; j) \geq \left[\begin{aligned} & p(x, i) - p(x, j) \\ & - \int_{\underline{x}(i)}^x \left(\frac{\partial p(x', i)}{\partial x'} - \frac{\partial p(x', j)}{\partial x'}\right) \left(\frac{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x)}{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x')}\right)^2 dx' \\ & + \int_{\underline{x}(j)}^{\underline{x}(i)} \frac{\partial p(x', j)}{\partial x'} \left(\frac{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x)}{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x')}\right)^2 dx' \end{aligned} \right] \quad (70)$$

Since $x' \leq x$, we have that $\frac{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x)}{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x')} \leq 1$ with equality only at $x' = x$. Consequently, we have that:

$$\int_{\underline{x}(i)}^x \left(\frac{\partial p(x', i)}{\partial x'} - \frac{\partial p(x', j)}{\partial x'} \right) \left(\frac{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x)}{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x')} \right)^2 dx' < \int_{\underline{x}(i)}^x \left(\frac{\partial p(x', i)}{\partial x'} - \frac{\partial p(x', j)}{\partial x'} \right) dx' = (p(x, i) - p(x, j)) - (p(\underline{x}(i), i) - p(\underline{x}(i), j)) \quad (71)$$

Substituting back into (70), we have:

$$K(x; i) - K(x; j) > \left\{ \begin{array}{l} p(\underline{x}(i), i) - p(\underline{x}(i), j) + \\ + \int_{\underline{x}(j)}^{\underline{x}(i)} \frac{\partial p(x', j)}{\partial x'} \left(\frac{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x)}{\bar{\Gamma}(\underline{x}^*(j)) + \kappa_j \bar{\Gamma}(x')} \right)^2 dx' \end{array} \right\} \quad (72)$$

Since $i > j \Rightarrow p(\underline{x}(i), i) - p(\underline{x}(i), j) > 0$, we have that $K(x; i) > K(x; j)$. ■

Proof of Proposition 2

Proof. From Lemmas 2 and 4, we know that for all x , $K(x, i) > K(x, j)$, and for all skill i , we have $K(x, i) > K(x', i)$ if $x > x'$. Moreover, from Corollary 3, we have that the support of wages offered to i and j -type workers is $[b(i), K(\bar{x}, i)]$ and $[b(j), K(\bar{x}, j)]$, where $b(i) \geq b(j)$ and $K(\bar{x}, i) > K(\bar{x}, j)$. Moreover, the distributions are continuous with no mass points and with a connected support, as shown by Burdett and Mortensen (1998). Now let's consider $F_i(w)$ and $F_j(w)$. If $w \in (b(j), b(i))$, we have that $F_j(w) > 0$ and $F_i(w) = 0$, so $F_i(w) \leq F_j(w)$. Similarly, if $w \in (K(\bar{x}, j), K(\bar{x}, i))$, we have that $F_j(w) = 1$ and $F_i(w) < 1$, so $F_i(w) \leq F_j(w)$. Finally, if $w \in [b(i), K(\bar{x}, j)]$, so the wage is offered to both skill- i and skill- j workers. From equation (15), we have that

$$F_i(w) = \frac{\Gamma(K^{-1}(w, i)) - \Gamma(K^{-1}(\underline{w}_i, i))}{\bar{\Gamma}(K^{-1}(\underline{w}_i, i))} \quad (73)$$

From Corollary 2, we have $K^{-1}(w, i) < K^{-1}(w, j)$, since the firm offering a given wage w for a higher-skill worker must have a lower productivity than the firm offering the same wage for a lower-skill worker. Consequently, $\Gamma(K^{-1}(w, i)) < \Gamma(K^{-1}(w, j))$. On the other side, Lemma 1 shows that $\underline{x}(i)$ is weakly increasing in i , and we have that $\underline{x}(i) \equiv K^{-1}(\underline{w}_i, i) \geq K^{-1}(\underline{w}_j, j) \equiv \underline{x}(j)$. Consequently, $\Gamma(K^{-1}(\underline{w}_i, i)) \geq \Gamma(K^{-1}(\underline{w}_j, j))$. Then, we have that

$$F_i(w) - F_j(w) = \frac{\Gamma(K^{-1}(w, i)) - \Gamma(K^{-1}(\underline{w}_i, i))}{\bar{\Gamma}(K^{-1}(\underline{w}_i, i))} - \frac{\Gamma(K^{-1}(w, j)) - \Gamma(K^{-1}(\underline{w}_j, j))}{\bar{\Gamma}(K^{-1}(\underline{w}_j, j))} \quad (74)$$

Manipulating it, we have

$$F_i(w) - F_j(w) = \frac{\begin{Bmatrix} \Gamma(K^{-1}(\underline{w}_j, j))\bar{\Gamma}(K^{-1}(w, i)) \\ -\Gamma(K^{-1}(\underline{w}_i, i))\bar{\Gamma}(K^{-1}(w, j)) \\ +\Gamma(K^{-1}(w, i)) - \Gamma(K^{-1}(w, j)) \end{Bmatrix}}{\bar{\Gamma}(K^{-1}(\underline{w}_j, j))\bar{\Gamma}(K^{-1}(\underline{w}_i, i))} \quad (75)$$

Adding and subtracting one to the numerator and manipulating, we obtain

$$F_i(w) - F_j(w) = \frac{\begin{Bmatrix} \bar{\Gamma}(K^{-1}(\underline{w}_i, i))\bar{\Gamma}(K^{-1}(w, j)) \\ -\bar{\Gamma}(K^{-1}(\underline{w}_j, j))\bar{\Gamma}(K^{-1}(w, i)) \end{Bmatrix}}{\bar{\Gamma}(K^{-1}(\underline{w}_j, j))\bar{\Gamma}(K^{-1}(\underline{w}_i, i))} < 0 \quad (76)$$

since $\bar{\Gamma}(K^{-1}(\underline{w}_i, i)) < \bar{\Gamma}(K^{-1}(\underline{w}_j, j))$ and $\bar{\Gamma}(K^{-1}(w, j)) < \bar{\Gamma}(K^{-1}(w, i))$. Consequently, $F_i(w) \leq F_j(w)$, $\forall w$. Therefore, $F_i(w)$ first-order stochastically dominates $F_j(w)$. ■

Proof of Corollary 4

Proof.

$$\begin{aligned} E_{F_i}[w] &= \int_{\underline{x}(i)}^{\bar{x}} K(x', i) \frac{\gamma(x')}{1-\Gamma(\underline{x}(i))} dx' \geq \int_{\underline{x}(i)}^{\bar{x}} K(x', j) \frac{\gamma(x')}{1-\Gamma(\underline{x}(i))} dx' \geq \\ &\geq \int_{\underline{x}(j)}^{\bar{x}} K(x', j) \frac{\gamma(x')}{1-\Gamma(\underline{x}(j))} dx' = E_{F_j}[w] \end{aligned} \quad (77)$$

where the first inequality comes from 2, while the second inequality comes from $F_i(w)$ F.O.S.D. $F_j(w)$. Simplifying the above expression, we have $E_{F_i}[w] \geq E_{F_j}[w]$. ■

Proof of Proposition 3

Proof. From Lemmas 2 and 4, we know that for all x , $K(x, i) > K(x, j)$, and for all skill i we have $K(x, i) > K(x', i)$ if $x > x'$. Moreover, from Corollary 3, we have that the support of wages offered to i and j -type workers is $[b(i), K(\bar{x}, i)]$ and $[b(j), K(\bar{x}, j)]$, where $b(i) \geq b(j)$ and $K(\bar{x}, i) > K(\bar{x}, j)$. Moreover, the distributions are continuous with no mass points and with a connected support, as shown by Burdett and Mortensen (1998). Now let's consider $G_i(w)$ and $G_j(w)$. If $w \in (b(j), b(i))$, we have that $G_j(w) > 0$ and $G_i(w) = 0$, so $G_i(w) \leq G_j(w)$. Similarly, if $w \in (K(\bar{x}, j), K(\bar{x}, i))$ we have that $G_j(w) = 1$ and $G_i(w) < 1$, so $G_i(w) \leq G_j(w)$. Finally, if $w \in [b(i), K(\bar{x}, j)]$, so the wage is earned by both skill- i and skill- j workers. In this case, we have

$$\frac{G_i(w)}{G_j(w)} = \frac{F_i(w)}{F_j(w)} \times \frac{1 + \kappa_j \bar{F}(w)}{1 + \kappa_i \bar{F}_i(w)} \quad (78)$$

From 2, we have $F_i(w) \leq F_j(w) \Rightarrow \bar{F}_i(w) \geq \bar{F}_j(w)$. Given the assumption that $\kappa_i \geq \kappa_j$

both terms in the RHS of equation (78) are smaller than one. Hence $G_i(w) \leq G_j(w)$ and $G_i(w)$ F.O.S.D. $G_j(w)$. ■

A.6 Methodology of Non-parametric Estimation of $p(x, i)$

We enact the following steps:

1. We nonparametrically (kernel estimation) estimate $G_i(w)$. We use Launov (2006)'s procedure to obtain an estimate of the upper tail using a Pareto distribution.
2. We adjust the estimates in order to obtain distributions that are closest to the estimated ones but still satisfy the model's conditions, i.e., the distribution must be single-peaked and

$$3\kappa_i g_i(w)^2 - g'_i(w) [1 + \kappa_i G_i(w)] > 0 \quad (79)$$

3. We assume all firms hire all skill levels: $\underline{x}^*(i) = \underline{x}$, $\forall i$.

Based on our estimates for the earned wage distributions and the parameters $\lambda(i)$, $\delta(i)$, and d , estimated from the job flows and death rates, we pin down output and output density functions ($v_i(p(x, i))$) as

$$p(x, i) = w + \frac{1 + \kappa_i G_i(w)}{2\kappa_i g_i(w)} \quad (80)$$

and

$$v_i(p(x, i)) = \frac{2\kappa_i(1 + \kappa_i)g_i(w)^3}{\left\{ \begin{array}{l} 3\kappa_i \{1 + \kappa_i G_i(w)\}^2 g_i(w)^2 \\ -g'_i(w)\{1 + \kappa_i G_i(w)\}^3 \end{array} \right\}} \quad (81)$$

Note that the recovered output distributions – $\Upsilon_i(\cdot)$, $i \in \{1, \dots, 5\}$ – are consistent with the labor market turnover rates and the average unemployment rates by different skill groups. Between conditions 2 and 3 described above, 2 ends up being the most material. It effectively guarantees that wages are non-decreasing with firm output for a given skill type. This does not stand out as a very restrictive assumption a priori, but we find this to be violated in the data, especially at the low-skill levels and at the low end of the wage distribution among low-skill worker types. The single-peak feature guarantees that for wages higher than the mod-wage, this condition is satisfied just by virtue of $g'_i(w) < 0$. We ensure this by effectively flattening the local peaks in the relevant domain for each skill type, keeping the aggregate mass constant. This step is rather straightforward and does not require a calibrated wage density that is far off from the empirical one. The problem is a bit complicated on the left side of the distribution for wages lower than the mod-wage.

Intuitively, our restriction in 2 implies that the calibrated density cannot be increasing “very fast” in that region. It turns out that this seems to be a feature of the data anyway for high-skill types. In the end, we can easily calibrate output distributions so that we can get the wage densities exactly for workers with a college degree or more, both for 1985 and 2015. For the low-skill workers, though, our best fit seems to fail to account for the rapid increase in the density early on.

We present the empirical estimates of $g_i(w)$ and the closest distributions we can get following our methodology that satisfy conditions 2 and 3 in Figures A.5 and A.6 for 1985 and 2015, respectively. As we can see, the adjusted distributions, serving as the inputs in our non-parametric calibration of the output distributions, are quite close to the empirical counterparts. Our largest “mismatch” is for high school dropouts in 1985, and that yields a fitted wage density that has 5.6 percent of its mass away from the empirical density. For the rest of the skill types, it rapidly declines to 3 percent for high school graduates and 1 percent for workers with some college. By these measures, our fit is much better for 2015. We obtain a calibrated density that only relocates 4.5 percent of the mass for the lowest skill type and 1.5 percent for the high school grads. The mismatch for workers with some college education is less than 0.5 percent. For both years, our non-parametric match to the empirical density for college graduates and workers with post-graduate degrees requires a distortion that is less than 0.03 percent of the respective mass. In summary, our inputs to the calibration of the output distributions present only small deviations from the empirical counterparts.

A.7 Matching Function Estimates: Adjustments for Job Finding and Arrival Rates

The matching function estimates are presented in Table A.3. The estimates for 1985 are based on monthly data for the sample period 02/1976 – 12/1990, while the 2015 estimates are based on monthly data for the sample period 01/2000 – 12/2015. The estimates, presented in Table A.3, show that matching efficiency has increased across all submarkets, although at different rates. Differently, the elasticity of the matching function with respect to labor supply have gone down in all submarkets, showing that the congestion caused by workers on each other within each submarket has gone up across all markets. Furthermore, we observe a positive trend in the matching function in the first period, while the second period has a negative and significant downward trend. While coefficients are small, these trends significantly affect the job finding rates.

In order to calculate the job-finding rates in each submarket, we follow Tasci (2012) and

Table A.3: Matching Function Estimates

Matching Efficiency					
	$\log(\mathcal{A}_1)$	$\log(\mathcal{A}_2)$	$\log(\mathcal{A}_3)$	$\log(\mathcal{A}_4)$	$\log(\mathcal{A}_5)$
1985	2.503*** (0.883)	6.887*** (0.501)	3.953*** (0.542)	2.833*** (0.640)	1.837** (0.809)
2015	5.764*** (0.439)	8.116*** (0.546)	9.230*** (0.534)	6.163*** (0.701)	5.230*** (0.711)
Matching Elasticity					
	$1-\beta_1$	$1-\beta_2$	$1-\beta_3$	$1-\beta_4$	$1-\beta_5$
1985	0.267*** (0.058)	0.521*** (0.031)	0.346*** (0.036)	0.279*** (0.044)	0.222*** (0.056)
2015	0.464*** (0.028)	0.577*** (0.033)	0.641*** (0.033)	0.462*** (0.044)	0.423*** (0.047)
Time Trend ϕ					
	1985	0.001*** (0.000)		2015	-0.002*** (0.000)

introduce the adjustment presented in equation (82):

$$\text{jfr}_i = \frac{\mathcal{M}_i(m_i, V_i)}{m_i} + \frac{\text{flow}U_i N_i}{U_i} \times \frac{\text{flow}N_i E_i}{\text{flow}N_i E_i + \text{flow}N_i U_i} \quad (82)$$

The adjustment in the RHS of equation (82) captures the created matches that involve a spell in not in the labor force (N) that may be missed otherwise in our estimation.

Finally, in order to pin down λ_i , we follow Shimer (2012) and the adjustment presented in equation (83):

$$\lambda_i = -\frac{\log(1 - \text{jfr}_i)}{1 - \delta_i} \quad (83)$$

However, our results are qualitatively the same if we ignore these adjustments in our counterfactual exercises.

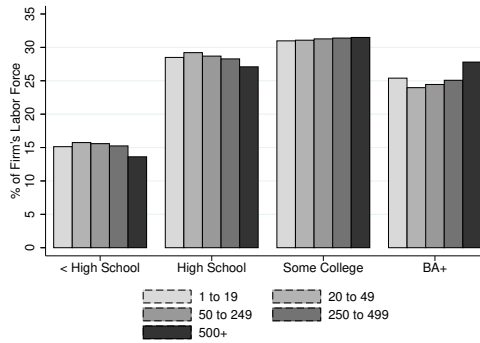
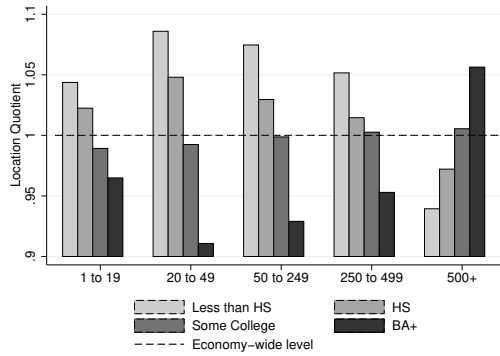
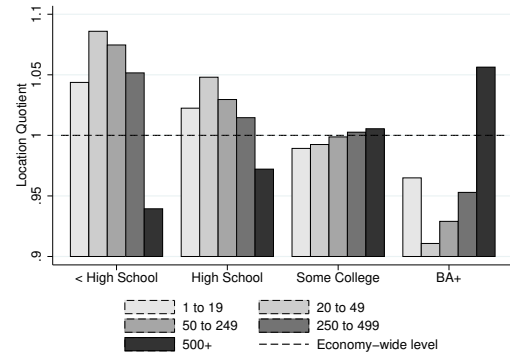


Figure A.1: Percentage of firm's labor force in each educational group

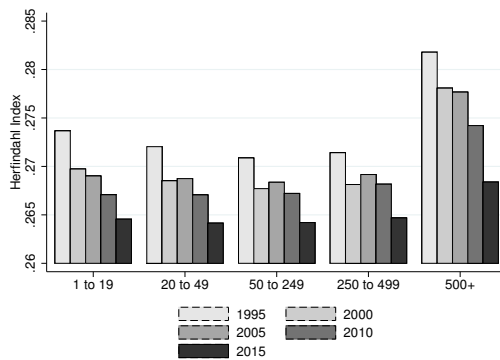


(a) Location Quotient - Firms

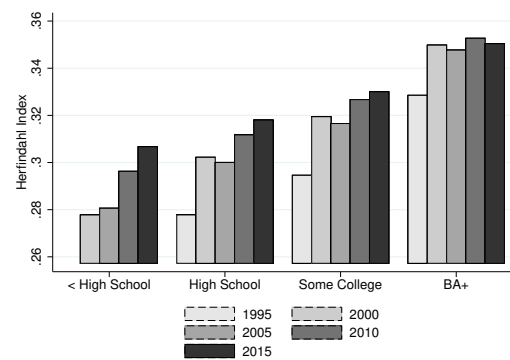


(b) Location Quotient - Workers

Figure A.2: Location quotient of educational-attainment groups across firm-size groups



(a) HHI - Firms



(b) HHI - Workers

Figure A.3: Herfindahl-Hirschman Indexes of Concentration: Education and Firm-Size Groups

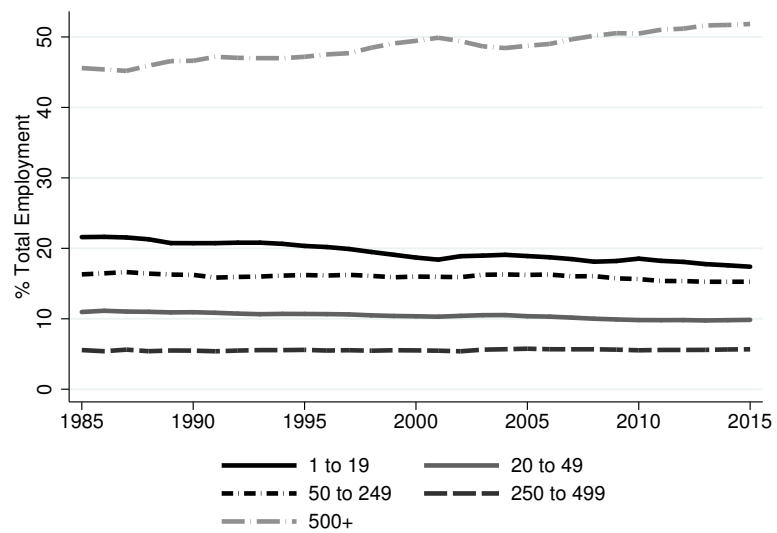


Figure A.4: Evolution of the Employed Labor Force across Firm Sizes

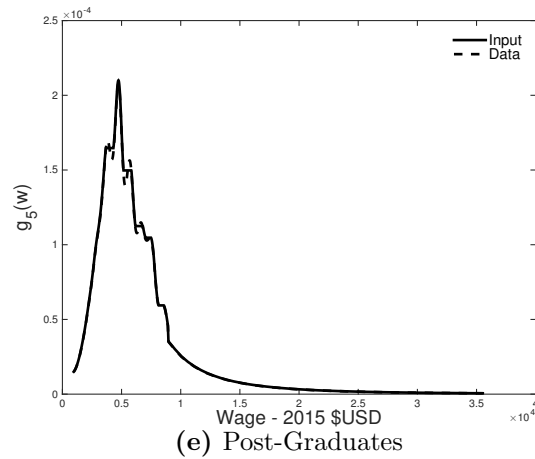
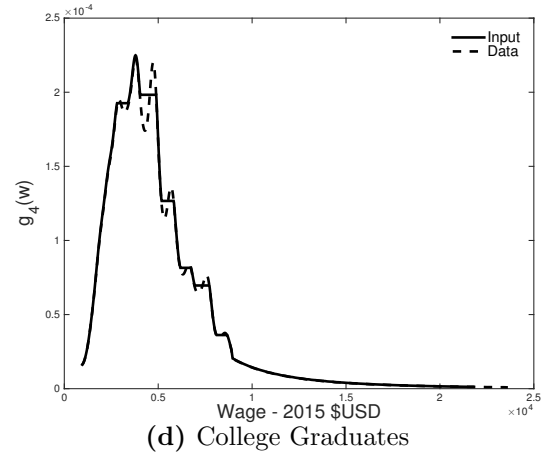
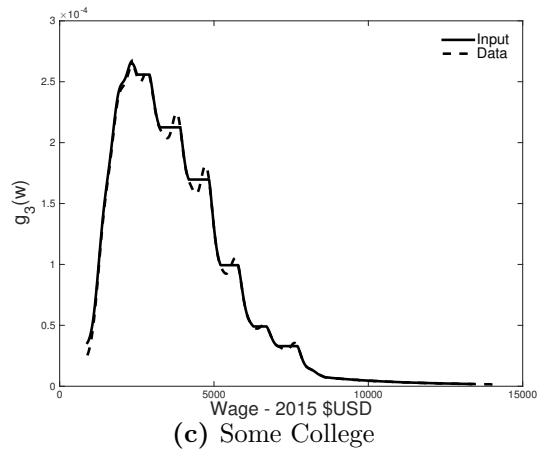
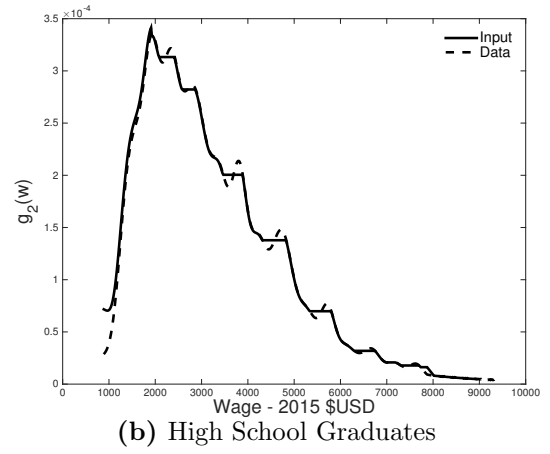
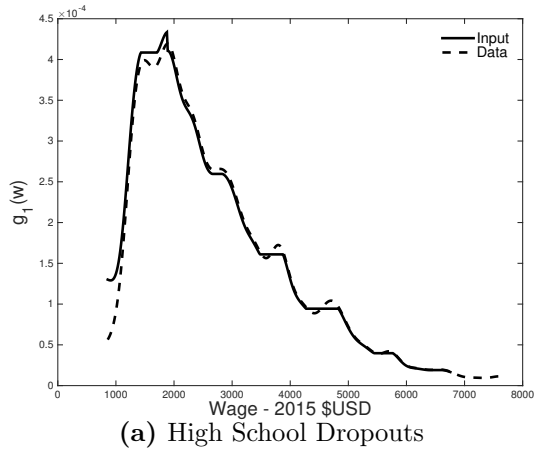
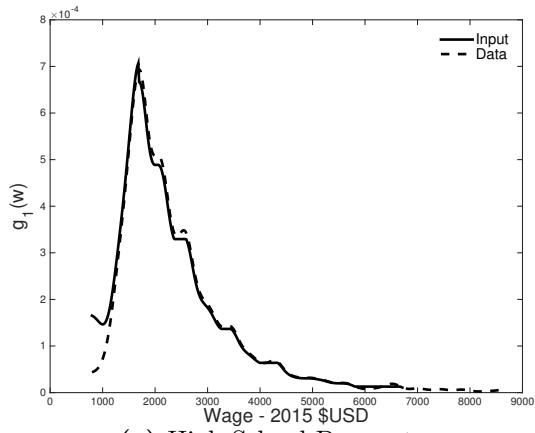
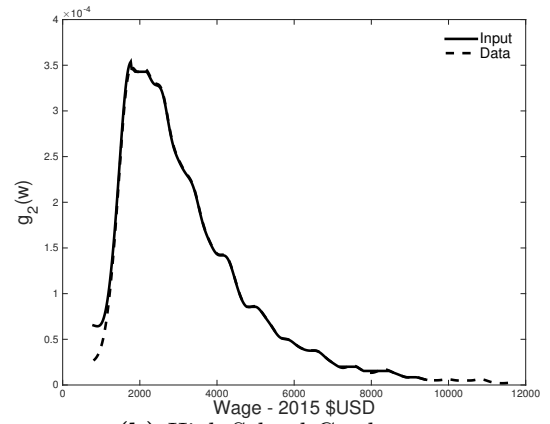


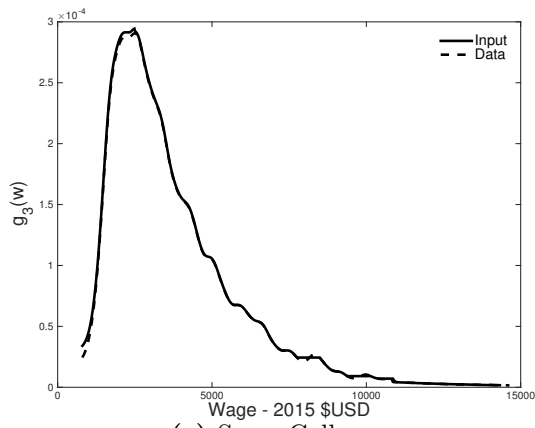
Figure A.5: Comparison between Empirical Wage Distributions and Adjusted Distributions that Fulfill Model Requirements – 1985



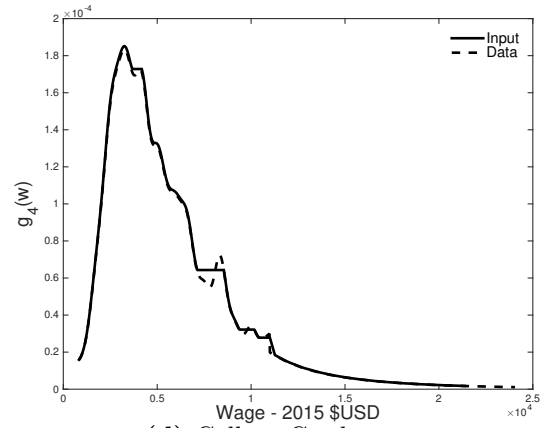
(a) High School Dropouts



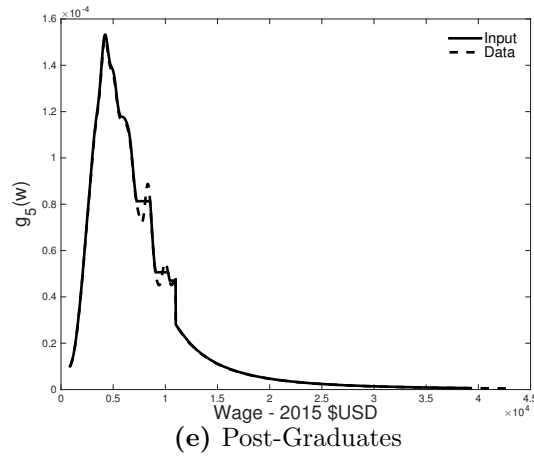
(b) High School Graduates



(c) Some College

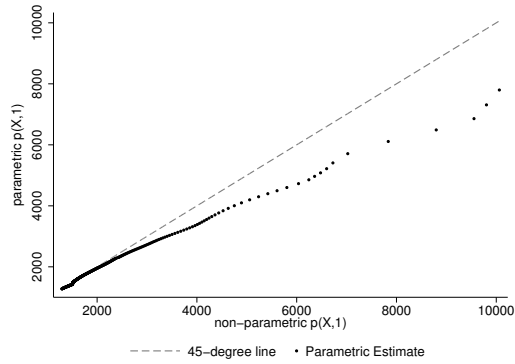


(d) College Graduates

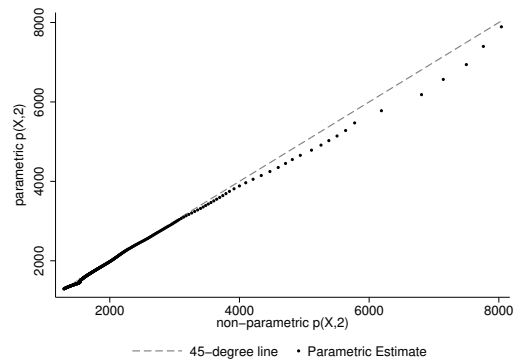


(e) Post-Graduates

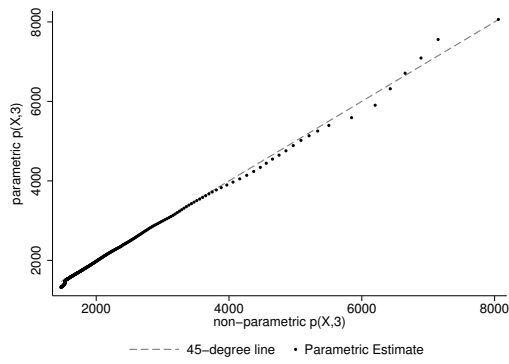
Figure A.6: Comparison between Empirical Wage Distributions and Adjusted Distributions that Fulfill Model Requirements – 2015



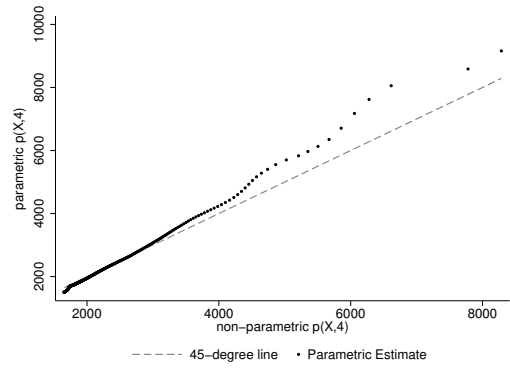
(a) High School Dropouts



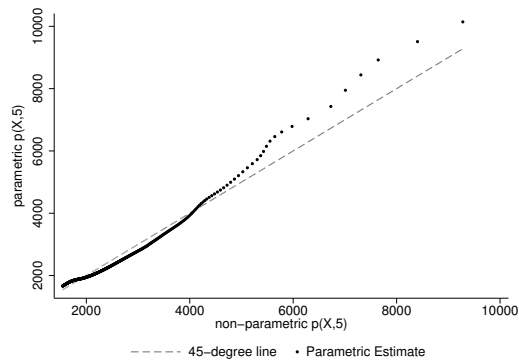
(b) High School Graduates



(c) Some College

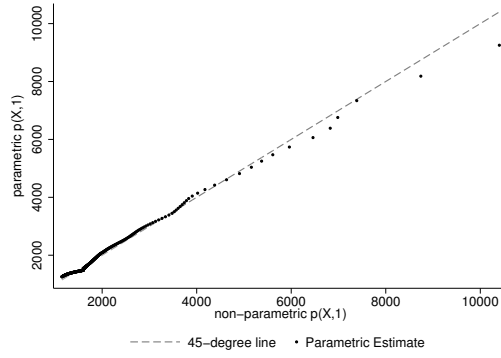


(d) College Graduates

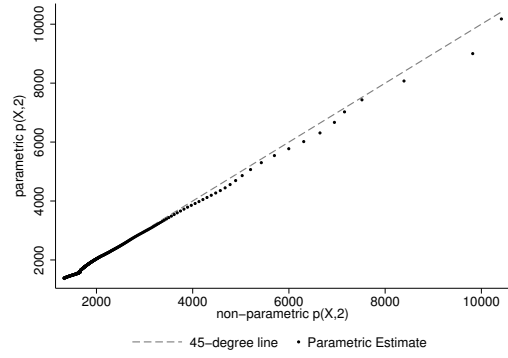


(e) Post-Graduates

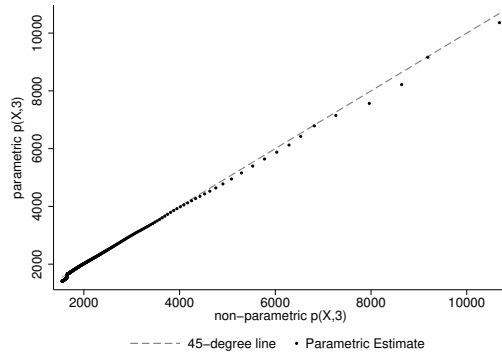
Figure A.7: Comparison between Parametric and Non-parametric Estimates of Output – 1985



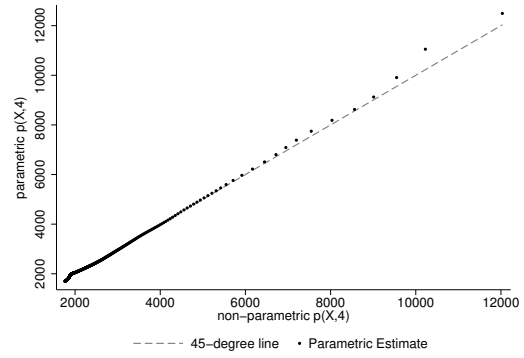
(a) High School Dropouts



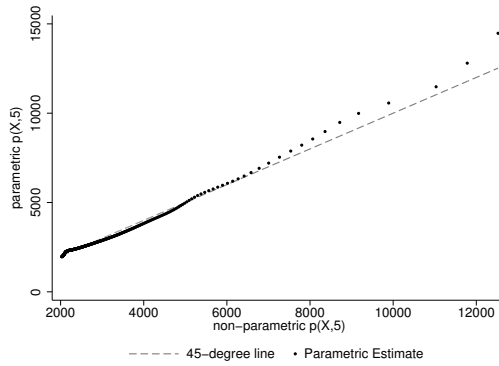
(b) High School Graduates



(c) Some College



(d) College Graduates



(e) Post-Graduates

Figure A.8: Comparison between Parametric and Non-parametric Estimates of Output – 2015

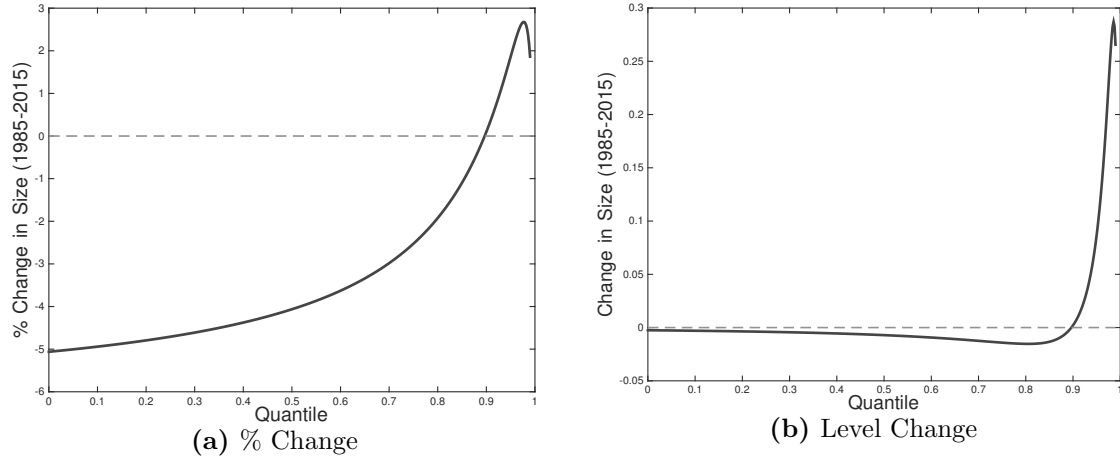


Figure A.9: Change in Size Distribution Across Quantiles – 1985 vs. 2015

Table A.4: Comparison Model vs. Data – Non-parametric Approach
(Calculations using values in \$1000 2015 USD)

<i>A. Average Earned Wage</i>						
	All	>HS	HS	SC	Col	Col+
1985 - Model	3.84	2.72	3.41	3.75	5.03	6.45
1985 - Data	3.79	2.76	3.22	3.71	5.00	6.43
2015 - Model	4.49	2.35	3.23	3.76	6.07	7.34
2015 - Data	4.48	2.43	3.30	3.69	5.60	7.38
<i>B. Std. Deviation of Earned Wages</i>						
	All	>HS	HS	SC	Col	Col+
1985 - Model	2.41	1.20	1.62	1.81	2.83	4.29
1985 - Data	2.42	1.31	1.54	1.89	2.97	4.34
2015 - Model	3.26	1.00	1.58	2.03	3.45	5.09
2015 - Data	3.37	1.15	1.75	2.10	3.49	5.36

Table A.5: Comparison Model vs. Data – Parametric Approach
(Calculations using values in \$1000 2015 USD)

<i>A. Average Earned Wage</i>						
	All	>HS	HS	SC	Col	Col+
1985 - Model	3.21	2.31	2.82	3.20	4.31	5.11
1985 - Data	3.79	2.76	3.22	3.71	5.00	6.43
2015 - Model	3.61	2.21	2.82	3.15	4.50	5.61
2015 - Data	4.48	2.43	3.30	3.69	5.60	7.38
<i>B. Std. Deviation of Earned Wages</i>						
	All	>HS	HS	SC	Col	Col+
1985 - Model	1.37	0.71	0.94	1.07	1.42	1.65
1985 - Data	2.42	1.31	1.54	1.89	2.97	4.34
2015 - Model	1.69	0.68	0.95	1.11	1.62	2.02
2015 - Data	3.37	1.15	1.75	2.10	3.49	5.36