working paper

17 06

Organizations, Skills, and Wage Inequality

Roberto Pinheiro and Murat Tasci

FEDERAL RESERVE BANK OF CLEVELAND

ISSN: 2573-7953

Working Paper 17-06 May 2017

Organizations, Skills, and Wage Inequality

Roberto Pinheiro and Murat Tasci

We extend an on-the-job search framework in order to allow firms to hire workers with different skills and skills to interact with firms' total factor productivity (TFP). Our model implies that more productive firms are larger, pay higher wages, and hire more workers at all skill levels and proportionately more at higher skill types, matching key stylized facts. We calibrate the model using five educational attainment levels as proxies for skills and estimate nonparametrically firm-skill output from the wage distributions for different educational levels. We consider two periods in time (1985 and 2009) and three counterfactual economies in which we evaluate how the wage distribution would have evolved if we kept one of the following key characteristics at its 1985's levels: firm-skill output distribution, labor market frictions, and skill distribution. Our results indicate that 79.2 percent of the overall wage dispersion and 72.6 percent of the within-group component can be attributed to a shift in the firm-skill output distribution. Once we assume a parametric calibration of the output per skill-TFP pair, we are able to show that most of the effect of changes in the output distribution is due to an increase in the average labor productivity of college graduates and post-graduates.

JEL Codes: D02, D21, J2, J3.

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

Suggested citation: Pinheiro, Roberto, and Murat Tasci, 2017. "Organizations, Skills, and Wage Inequality," Federal Reserve Bank of Cleveland, Working Paper no. 17-06. https://doi.org/10.26509/frbc-wp-201706.

Roberto Pinheiro is at the Federal Reserve Bank of Cleveland (Roberto.Pinheiro@clev.frb.org). Murat Tasci is at at the Federal Reserve Bank of Cleveland (Murat.Tasci@clev.frb.org). The authors thank Jim Albrecht, Bruce Fallick, and Susan Vroman for their comments, as well as the participants of the ROA Seminar at Maastricht University, Cleveland State University, Kent State University, Cleveland Fed Microeconomics Workshop, 2017 MEA-Sole, and 2017 IIOC.

1 Introduction

The role played by observable workers' characteristics, such as education, in accounting for the increase in wage inequality is of primary concern for economists and policymakers alike. Moreover, the increasing wage inequality among otherwise observationally identical workers, in particular for highly skilled workers, complicates even more the understanding of the driving forces behind wage inequality. In fact, not only workers with different educational attainment levels have different wage profiles and labor market transition rates, but wage outcomes also differ even among observationally equivalent workers, depending on what type of firms they work for. In this paper, we revisit this issue of wage inequality by taking seriously these empirical facts about labor market experiences of different skill/education groups of workers. Our primary contribution is to present and calibrate a structural model with labor market frictions that is consistent with these empirical regularities in order to evaluate the role played by different explanations proposed in the literature for rising wage inequality during the previous three decades.

Many of these facts have been already widely documented in the literature. Workers with different skills have different career paths. In particular, more-educated workers experience lower unemployment rates and lower employment volatility. They also face a higher chance of being offered a better job, being more likely to move on the job. For example, Fallick and Fleischman (2001) find that, while the total separation rates fall with workers' educational attainment, in relative terms job-to-job transitions account for a much larger share of total separations for college-educated workers (50 percent) than for high school dropouts (30 percent). Similar evidence has been found by Nagypal (2008), who reports that around 55% of total separations of workers with a college degree are due to job-to-job transitions.

Moreover, highly educated workers are more likely to be employed at large firms that pay higher wages in general. Studies that look at the labor skill distribution among firms show that larger firms hire over a wider range of skills and as a result have more levels in their hierarchy. Even more, large firms employ proportionally more high-skill workers, i.e., the skill distribution at large firms dominates the skill distribution at smaller firms in first order. Consequently, large firms pay on average higher wages. In fact, large firms pay higher wages than smaller firms for workers with the same observable characteristics. Empirically, these patterns were found by Caliendo, Monte, and Rossi-Hansberg (2015) using data from French manufacturing firms; by Tag (2013) using Swedish data; and Colombo and Delmastro (1999) using Italian data.

However, the data show that education by itself is not a guarantee of a successful career. Although average wages are increasing with educational attainment, within-group inequality is also increasing in the level of education. As presented by Budría and Telhado-Pereira (2011), education decisions embody wage risks once recent empirical research has shown that returns to schooling are subject to an important degree of variation across individuals.[1] More importantly, these risks are concentrated at the higher levels of education. For example, Lauer (2004) uses data for France and Germany to examine differences across educational qualifications regarding the amount of (unexplained) within-group dispersion. For both countries, she finds that wage dispersion is the lowest among workers with a vocational qualification and highest among workers with a tertiary education. Similar results were found by Budría and Telhado-Pereira (2011) for most European countries, among others.[2] Moreover, changes in residual or within-group inequality in the last 30 years have been concentrated at the top-end of the skill distribution. For instance, Lemieux (2006) shows that within-group inequality grew substantially among college-educated workers but changed little for most other groups. Similarly, Autor et. al. (2005) find that "top-end" residual inequality – e.g., the difference between 90^{th} and 50^{th} percentile of the distribution of residuals – increased substantially, while residual inequality at the low-end – the 50-10 gap – actually decreased.

In this paper, we construct a model that allows us to study the market characteristics that generate the patterns for workers' career paths and organizations' sizes and shapes. In our model, not only do we allow for firm and worker heterogeneity, but we also allow for firms to hire workers at different skill levels, generating within-firm skill distributions. In this sense, we are able not only to discuss within-group and between-group wage inequality, but also to discuss the contribution of within- and between-firm wage dispersion on the overall increase in wage inequality (as discussed by Song et. al. (2016) and Barth et. al. (2016)). Once we calibrate our parameters for two distinct years of the US economy – 1985 and 2009 – we are able to disentangle how much changes in the output per skill-firm pairs, educational attainment distribution, and labor market frictions have contributed to the overall increase in wage inequality. Moreover, we can also evaluate how much those changes may have impacted the within-and between-skill group wage inequality.

More specifically, we build a model with frictional labor markets in which firms with different levels of total factor productivity (TFP) compete for differently skilled workers that are allowed to search while on the job. In our framework, a worker's skill not only affects her productivity, but also how easily she adapts to the changing nature of tasks. New tasks arrive at a Poisson rate $\delta > 0$, and a worker is able to solve the new task depending on his or her skill level. We show that in this environment, the patterns found by the empirical literature that connect firm size, within-firm hierarchy, and wage distributions arise endogenously in equilibrium. In particular, we show that, in equilibrium, firms with higher TFP are larger, employ more workers at all skill levels, and pay higher wages at all skill levels. Furthermore, if firms have skill distributions with a shared support, the skill distribution of a firm with

[1] Maier et. al. (2004) report that between 20 percent and 30 percent of German male workers earn a negative return from schooling while more than 25 percent earn a return exceeding 15 percent.

[2] Similar results were found by Buchinsky (1994, 1998) for the United States; Machado and Mata (1997) and Hartog, Pereira, and Vieira (2001) for Portugal; Abadie (1997) for Spain; and Gosling et. al. (2000) for the United Kingdom.

higher TFP stochastically dominates in first order the skill distribution of a lower TFP firm.

In terms of wage distributions, we show that average wages are higher for more skilled workers, independent of a firm's TFP levels. However, the support of the distributions of wages received by different skill levels may overlap once workers with different skill levels may receive the same pay at firms with different TFP levels. In addition, our model allows us to calculate the decomposition of within- and between-firm components of wage dispersion, following previous work by Lazear and Shaw (2007). We also adapt their methodology and decompose the overall wage variance in within- and between-educational group components.

We present two calibrations for our model. First, we follow Bontemps et al. (2000) and Launov (2005) and nonparametrically estimate the firm-skill output function. This methodology has the advantage of allowing us to pin down the output without imposing an ad hoc functional form for the production function or demanding matched employer-employee data. As a drawback, we are only able to evaluate the aggregate impact of changes in the overall output per firm-skill on the wage distribution. Consequently, this calibration is unable to pin down the mechanism through which changes in output impact wage dispersion. In order to disaggregate the impact of changes in output into its components – i.e., changes in the TFP distribution and changes in labor productivity – we present a parametric calibration, as well. In this calibration, we use the TFP distribution estimates from previous studies – in particular, Imrohoroglu and Tuzel (2014) – and a Cobb-Douglas production function. We calibrate the production function parameters to match the BLS's estimates for labor share in 1985 and 2009. Finally, we calibrate labor productivity for each skill in order to match the average wage per skill in each period. The parametric calibration allows us not only to disentangle the impact of changes of the TFP distribution and labor skill productivity on average wages and wage dispersion, but it also allows us to discuss the changes in within- and between-firm components of the overall wage inequality. The caveat of the parametric exercise is that not only our estimates for TFP distribution are based on a sample of public companies – consequently a subsample of the overall universe of U.S. firms that is biased toward large firms – but we also have to impose a particular production function. As we show in section 3.2, the model fit based on the parametric calibration is significantly worse than the one obtained through the nonparametric specification.

Our nonparametric calibration and counterfactual exercises show that the bulk of the increase in overall wage dispersion (79.2 percent) between 1985 and 2009 can be explained by differences in output per skill-TFP pairs across the two periods. Changes in the underlying output distributions for different skill types account for 72.6 percent of the within group variation and more than account for the between group variation. Moreover, we see that the output distributions for all groups except high school dropouts in 2009 dominate stochastically in first order their 1985 counterparts. This result indicates that skill-biased technological change (SBTC) played a key role for the increase in between-group wage inequality.

We also highlight an increase in dispersion (measured by the standard deviation of earned wages) between 1985 and 2009 that has been concentrated among high-educational-attainment groups (college grads and post-graduates). This increase in dispersion, jointly with the importance of changes in output distribution in explaining the increase in wage dispersion, indicates that firms with different levels of TFP had quite distinct levels of success leveraging the potential of highly-skilled workers. Consequently, the increase in intra-group wage inequality concentrated at the top of the educational attainment distribution can be explained by some highly skilled workers being lucky to land a job at a firm able to extract a large output from them, while others are unlucky and work for firms in which SBTC was not as intense. 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 many times the increase in within-group dispersion has been transitory (see Gottschalk and Moffitt (1994) and Kambourov and Manovskii (2009), among others), at odds with an explanation based on unobserved skills, since unobserved skills are usually constant or moving slowly over time.[3]

Our results also show that changes in labor market frictions partially undo the effects of the technological changes. Based on our estimates from the CPS MORG, we observe a reduction in labor market frictions at the low educational attainment groups (high school dropouts and high school graduates) and an increase in frictions for higher educational groups. This feature manifests itself as higher transition rates into employment and lower separation rates for the low skill groups. As our counterfactual results show, average wages and wage dispersion would be even lower for high school dropouts and higher for college graduates and post graduates if labor market frictions were kept at 1985 levels.

Finally, our parametric calibration and counterfactual exercises allow us to explore by which particular mechanism SBTC affects the increase in overall wage dispersion. Our results indicate that the increase in the labor productivity of highly skilled workers explains the bulk of the increase in wage dispersion and the increase in average wages earned by highly skilled workers during the period. In fact, changes in the TFP distribution, while contributing to the overall wage dispersion, had a negative effect on average wages earned by workers with a high school degree or more education. Moreover, once we look at the average compensation and wage dispersion across firms, we also see that changes in labor productivity are responsible not only for high-TFP firms paying significantly higher average wages over time, but also for an increase in within-firm wage dispersion at high-TFP firms.

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

[3]The literature also considers that changes in the price of unobserved skills because of SBTC are usually persistent, reinforcing the idea that increases in intragroup wage inequality should be long-lasting.

of labor frictions, SBTC, and educational attainment distributions to the increase in overall wage dispersion. Moreover, we are able to decompose the effect in terms of within- and between-group and within- and between-firm components. 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 decomposition. In this sense, our model adds to the discussion on within- and between-group inequality presented by Lemieux (2006a, 2006b, 2008a, 2008b) and Autor and Acemoglu (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. (2016). We are also able to discuss the different patterns of job mobility between 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). Finally, our model connects the search literature with the models of knowledge hierarchy (Garicano (2000), and Garicano and Rossi-Hansberg (2006), among others) and organization adaptability (Dessein and Santos, 2006).

The paper is divided to 6 sections. Section 2 presents the model and our analytical results characterizing the equilibrium. Section 3 discusses our calibration for the labor market frictions and educational attainment distribution, along with the nonparametric and parametric calibration of the output per skill-TFP pairs. Section 4 presents the benchmark results for our nonparametric approach and our counterfactual exercises that allow us to evaluate how much of the overall increase in within and between-group wage inequality between 1985 and 2009 can be explained by changes in the distribution of educational attainment, labor market frictions, and output distributions by TFP-skill pairs across the two periods. Section 5 presents the results for the benchmark and counterfactual exercises for our parametric calibration. We focus on the contributions of changes in labor productivity and TFP distributions between 1985 and 2009 in order to disentangle the overall effect on wage inequality of changes in the output distribution by TFP-skill pairs that we obtained from our nonparametric calibration. Section 6 concludes the paper.

2 Model

Consider an economy with a measure 1 of firms. Firms have different levels of productivity. We assume a continuous distribution of productivity types $\Gamma(x)$ with support $[\underline{x}, \overline{x}]$. Firms are risk-neutral, infinitely lived, and discount future at rate $r > 0$. There is a measure M of workers. Workers are heterogeneous in their skills. There are $\mathcal{I}$ skills, and the measure of workers with skill i is given by m_i with $\sum_{i=1}^{\mathcal{I}} m_i = M$. We assume that the skill distribution in the economy is given. All workers are risk-neutral and exit the market at rate $d \in (0, 1)$, being replaced by an unemployed worker of the same skill. Exiting workers experience a disutility cost $c \geq 0$. Workers also discount the future at rate $r > 0$.

The output flow of a worker of skill i in a firm with productivity x is given by $p\left(x, i\right)$. We assume that $\frac{\partial p(x,i)}{\partial x} > 0$ and $\frac{\partial p(x,i)}{\partial i} \geq 0$. The output is obtained by successfully completing a given task; failed tasks generate an output flow of zero. We assume that there is a change in tasks faced by a given worker[4] at a Poisson rate $\delta \in (0, 1)$. The new task is drawn from a discrete distribution $H\left(\cdot\right)$ with support $\{1, ..., J\}$ and with $J > N$. In order to simplify exposition, we follow Garicano and Rossi-Hansberg (2006) and assume that the p.d.f. is strictly decreasing, i.e., $h\left(z\right) < h\left(z'\right)$ if $z < z'$. We assume that a worker of skill level i is able to solve tasks in the support $[0, i]$. We consider that there is no team production in the sense that each worker in the firm either solves a task by herself or not at all. We also assume that workers cannot be reallocated to a different task. Notice that if the worker cannot solve a problem, given task persistence and a bad enough productivity in this case, the optimal decision by the firm as well as the worker is to destroy the job match. Therefore, a worker of skill level i is dismissed with probability $\delta_i = \delta\left[1 - H\left(i\right)\right]$, where $H\left(i\right) = \Pr\left(\text{task} \leq i\right)$. We also assume that a new task can be sampled at the moment that a worker and a firm first match. Consequently, the likelihood that a job match happens at a task in which the worker can successfully execute the job is given by the arrival rate $(1 - \delta_i)$.

The labor market is frictional, and search is random. This means that search frictions in the labor market prevent workers from instantaneously matching with the best job offer in the market. Instead, from time to time workers meet randomly with one of the firms in the market. On the other hand, labor markets are partially segmented, so the arrival rate of a job offer may depend on the worker's skill level. We assume that a worker of skill level i meets a potential employer at rate $\lambda(i) > 0$, irrespective of her employment status. A meeting becomes a job offer at rate $(1 - \Gamma(\underline{x}(i)))(1 - \delta_i)$, where $\underline{x}(i)$ is the lowest TFP-level firm that hires skill level i (which will be determined in equilibrium). A job offer consists of a wage rate w. Worker skills are observable, so firms can condition wage offers to skills. The job offer must be accepted or rejected on the spot, and when rejected, it cannot later be recalled. Firms are able to hire everyone who accepts their wage offers, and they choose the wages offered to differently skilled workers in order to maximize their steady state profits. In setting profit maximizing wages, firms take into account the distribution of wages posted by the other firms in the market. Notice that firms will post one wage for each skill level in the economy. If a firm decides not to hire a given skill level, we will assume the firm is posting a wage rate of 0 for that particular skill.

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 $p(\underline{x}, 1) > b(1)$, which implies that all firms are active at least at the lowest rung of the skill ladder. We will also keep as a maintained assumption that $b'(i) \geq \frac{\partial p(x,i)}{\partial i}$, i.e., the productivity growth at home

[4] Notice that this is an i.i.d. shock.

production with respect to skill level is at least as large as the increase in productivity at the job at any given type x firm.

2.1 Worker's Problem

From the framework outlined above, the expected discounted lifetime income when a type i worker is unemployed, $U(i)$, can be expressed as the solution of the following equation:

$$(r+d)U(i) = b(i) - dc + \lambda(i)[1-\Gamma(\underline{x}(i))](1-\delta_i)\int \max\{J(w',i)-U(i),0\}dF_i(w') \tag{1}$$

where $b(i)$ is the flow value of home production, c is the cost of exit (injury or death) that arrives at rate $d \in (0,1)$, and $\lambda(i)[1-\Gamma(\underline{x}(i))](1-\delta_i)$ is the effective arrival rate of a job offer, i.e., the rate at which the type i unemployed workers meet a firm that would like to hire that skill $(\lambda(i)[1-\Gamma(\underline{x}(i))])$ and offers a position that the worker can fulfill the required task $(1-\delta_i)$.

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

$$(r+d)J(w,i) = w - dc + \lambda(i)[1-\Gamma(\underline{x}(i))](1-\delta_i)\int_w (J(w',i)-J(w,i))dF_i(w') + \delta_i(U(i)-J(w,i)) \tag{2}$$

where δ_i is the rate at which a type i worker faces a change in tasks that make her unproductive at her work. Notice 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 worker of skill level i, $R(i)$, exists and is defined by $J(R(i),i) = U(i)$. Then, from equations (1) and (2), we obtain that

$$R(i) = b(i) \tag{3}$$

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 steady state is given by

$$d(m_i - u_i) + \delta_i(m_i - u_i) = \lambda(i)[1-\Gamma(\underline{x}(i))](1-\delta_i)(1-F_i(R(i)))u_i$$

where the left-hand side (LHS) of the above equality represents the inflow of skill i into the unemployment pool, while the right-hand side (RHS) represents the outflow. Rearranging the above equation, we obtain the following:

$$u_i = \frac{(d+\delta_i)m_i}{d+\delta_i+\lambda(i)[1-\Gamma(\underline{x}(i))](1-\delta_i)(1-F_i(R(i)))} \tag{4}$$

Consequently, the unemployment rate among skill i workers is given by

$$unemp_i = \frac{u_i}{m_i} = \frac{(d+\delta_i)}{d+\delta_i+\lambda(i)[1-\Gamma(\underline{x}(i))](1-\delta_i)(1-F_i(R(i)))} \tag{5}$$

In order to present some concepts that will be used in later sections, let's define γ_i the fraction of skill i workers among the unemployed population. Given equation (4), we obtain

$$\gamma_i = \frac{\frac{(d+\delta_i)m_i}{d+\delta_i+\lambda(i)[1-\Gamma(\underline{x}(i))](1-\delta_i)(1-F_i(R(i)))}}{\sum_{j=1}^{\mathcal{I}} \frac{(d+\delta_j)m_j}{d+\delta_j+\lambda(j)[1-\Gamma(\underline{x}(j))](1-\delta_j)(1-F_j(R(j)))}} \tag{6}$$

2.2 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 ($F_i(w),\ i \in \{1, 2, ..., \mathcal{I}\}$) and the distribution of wages that employed workers are currently earning at other firms, given by $G_i(w),\ i \in \{1, 2, ..., \mathcal{I}\}$. 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 attact 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 $G_i(w), \forall i \in \{1, 2, ..., \mathcal{I}\}$ and the firm's labor force for each skill level i. In order to do that, let's start with the firm's decision of hiring a given skill level or not. In order to hire a type i worker, the firm must pay a wage that is at least as high at the reservation wage $R(i)$. Consequently, a firm with TFP level x will only hire skill i if

$$p(x,i) \geq R(i)$$

Substituting the expression for $R(i)$ presented in (3), we have

$$p(x,i) \geq b(i) \tag{7}$$

Since $p(x,i)$ is strictly increasing on x, we have that $\underline{x}(i)$, the lowest TFP-level firm that hires skill level i, is determined by the following equality:

$$p(\underline{x}(i), i) = b(i) \tag{8}$$

as expected, for any $x > \underline{x}(i)$, the demand for skill i is strictly positive. Notice that

$$\frac{\partial p(x,i)}{\partial x}\Big|_{x=\underline{x}(i)} \times \frac{d\underline{x}(i)}{di} + \frac{\partial p(x,i)}{\partial i}\Big|_{x=\underline{x}(i)} = b'(i)$$

Rearranging it produces the following:

$$\frac{d\underline{x}(i)}{di} = \frac{b'(i) - \left.\frac{\partial p(x,i)}{\partial i}\right|_{x=\underline{x}(i)}}{\left.\frac{\partial p(x,i)}{\partial x}\right|_{x=\underline{x}(i)}}$$

Since $\left.\frac{\partial p(x,i)}{\partial x}\right|_{x=\underline{x}(i)} \geq 0$, the sign of $\frac{d\underline{x}(i)}{di}$ depends on the numerator. We will consider 2 cases:

a. $b'(i) = \frac{\partial p(x,i)}{\partial i}$, $\forall i \in \mathcal{I}$. This case follows the literature in human capital and search (see Fu (2011) and Burdett, Carrillo-Tudela, and Coles (2011)). Consequently, we have $\frac{d\underline{x}(i)}{di} = 0$, and active firms share a support of skills, i.e., $\underline{x}(i) = \underline{x}^*, \forall i \in \mathcal{I}$.

b. $b'(i) > \frac{\partial p(x,i)}{\partial i}$, $\forall i \in \mathcal{I}$. This case considers that autarky productivity increases faster than firm output does with skills. We could consider this the entreprenureal route. As expected, here we have $\frac{d\underline{x}(i)}{di} > 0$.

We now consider the distribution of employed type i workers that earn a wage less than w, $G_i(w)$. Let's initially consider how this distribution evolves over time:

$$\begin{aligned} &\tfrac{dG_i(w,t)}{dt} = \lambda(i)[1-\Gamma(\underline{x}(i))](1-\delta_i)[F_i(w,t) - F_i(R(i),t)]u_i(t) \\ &- \{d + \delta_i + \lambda(i)[1-\Gamma(\underline{x}(i))](1-\delta_i)[1-F_i(w,t)]\}\, G_i(w,t)(m_i - u_i(t)) \end{aligned} \tag{9}$$

The first term on the RHS of (9) 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 flow out into death, unemployment, and higher-paying jobs, respectively. Since in steady state $\frac{dG(w,t|i)}{dt} = 0$, this distribution can be rewritten as

$$G_i(w) = \frac{(d+\delta_i)[F_i(w) - F_i(R(i))]}{\{d + \delta_i + \lambda(i)[1-\Gamma(\underline{x}(i))](1-\delta_i)[1-F_i(w)]\}\,(1 - F_i(R(i)))} \tag{10}$$

Finally, the steady state number of skill i workers earning in the interval $[w, w-\varepsilon]$ is represented by $[G(w) - G(w-\varepsilon)]\,(m_i - u_i)$, while $[1-\Gamma(\underline{x}(i))]\,[F(w) - F(w-\varepsilon)]$ is the measure of firms offering wages in the same interval. Notice that $[1-\Gamma(\underline{x}(i))]$ represents the measure of workers actively hiring skill i workers. Thus, the measure of skill i workers per actively hiring i firms earning a wage w can be expressed as

$$l(w;i) = \frac{(m_i - u_i)}{[1-\Gamma(x_{\min}(i))]} \lim_{\varepsilon\to 0} \frac{G(w\,|i) - G(w-\varepsilon\,|i)}{F(w\,|i) - F(w-\varepsilon\,|i)}$$

Solving it and substituting u_i, we obtain the following:

$$l(w;i) = \frac{\lambda(i)(1-\delta_i)(d+\delta_i)m_i}{\{d + \delta_i + \lambda(i)[1-\Gamma(\underline{x}(i))](1-\delta_i)(1-F_i(w))\}^2} \tag{11}$$

2.3 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 value being employed in this firm more highly 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 profit upward, it comes at the cost of earning less profit per worker, which pushes profits downward. At the optimal wage, the firm balances these two counteracting dimensions, 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 $[\underline{x}, \overline{x}]$. Firms can offer different wages for different skill levels. Therefore, the profit function for a firm that posts wages $\{w(i)\}_{i=1}^{N}$ and has productivity level x is given by

$$\pi(x) = \sum_{i \in A(x)} (p(x,i) - w(i))\, l(w(i); i) \tag{12}$$

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

$$A(x) = \{i \in \{1, ..., N\} \,|p(x,i) \geq R(i)\} \tag{13}$$

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

$$\pi(x; i) = \max_{w} (p(x,i) - w)\, l(w; i) \tag{14}$$

for each $i \in A(x)$.

Then, the optimal wage posted by a firm of productivity x for a skill level $i \in A(x)$ is given by $w = K(x; i)$, given $p(x,i), F(\cdot|i)$. From the first order condition – i.e., $\frac{\partial \pi(x,w(i);i)}{\partial w(i)} = 0$, we obtain

$$-\, l(w; i) + (p(x,i) - w)\frac{dl(w; i)}{dw} = 0 \tag{15}$$

From our expression for $l_i(w)$, we have that

$$\frac{dl_i(w)}{dw} = \frac{2\lambda(i)(1-\Gamma(\underline{x}(i)))(1-\delta_i)f_i(w)}{d + \delta_i + \lambda(i)[1-\Gamma(\underline{x}(i))](1-\delta_i)(1-F_i(w))} l_i(w) \tag{16}$$

Substituting it back in the F.O.C. yields:

$$2\lambda(i)[1-\Gamma(\underline{x}(i))](1-\delta_i)f_i(w)(p(x,i)-w)-\{d+\delta_i+\lambda(i)[1-\Gamma(\underline{x}(i))](1-\delta_i)(1-F_i(w))\}=0 \quad (17)$$

which provides us with an implicit equation for w as a function of $p(x,i)$ and F. At this stage, we proceed as if the second order conditions are satisfied. Below, we then 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:

$$(p(x,i)-K(x;i))\, l(K(x;i);i) \quad (18)$$

Then, a differentiation with respect to x, using the envelope theorem, gives us

$$\frac{\partial \pi(x;i)}{\partial x}=\frac{\partial p(x,i)}{\partial x} l(K(x;i);i) \quad (19)$$

Then, we have that

$$\pi(x;i)=\int_{\underline{x}(i)}^{x}\frac{\partial p(x',i)}{\partial x'} l\left(K\left(x';i\right);i\right)dx' \quad (20)$$

From equation (11) and because of the result that $[1-F_i(K(x';i))]=\frac{[1-\Gamma(x')]}{[1-\Gamma(\underline{x}(i))]}$ (see Bontemps et. al. (2002), Proposition 3), we have that

$$\pi(x;i)=\int_{\underline{x}(i)}^{x}\frac{\partial p(x',i)}{\partial x'}\frac{\lambda(i)(1-\delta_i)(d+\delta_i)m_i}{\{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x')]\}^2}dx' \quad (21)$$

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:

$$K(x;i)=p(x,i)-\int_{\underline{x}(i)}^{x}\frac{\partial p(x',i)}{\partial x'}\left(\frac{d+\delta_i+\lambda_1(1-\delta_i)[1-\Gamma(x)]}{d+\delta_i+\lambda_1(1-\delta_i)[1-\Gamma(x')]}\right)^2 dx' \quad (22)$$

Before we move on, let's present some important basic results.

Lemma 1 *For any skill level, the wage is increasing in firm productivity, i.e.,* $\frac{\partial K(x;i)}{\partial x}>0$.

Moreover, define $l_i(x)\equiv l_i(K(x,i))$. Then, we get the additional results:

Lemma 2 $\frac{\partial l_i(x)}{\partial x}\geq 0$, $\forall i\in\mathcal{I}$. *Consequently, firms with higher TFP hire more of each skill.*

Then, define firm size of a firm with TFP x by

$$S(x)=\sum_{i\in\mathbf{I}(x)} l_i(x) \quad (23)$$

Then, a simple corollary of lemma 2 follows (keeping in mind that $\frac{d\underline{x}(i)}{di}\geq 0$).

Corollary 1 *Firm size is increasing in x, i.e., $S(x) > S(x')$ if $x > x'$.*

Now, let's define the skill distribution at a x-type firm by

$$\Phi_x(i) = \frac{\sum_{i'=1}^{i} l_{i'}(x)}{S(x)} \tag{24}$$

We then want to show the following proposition:

Proposition 1 *If $x > x'$ we have that $\Phi_x(i) \leq \Phi_{x'}(i),\ \forall i \in \mathcal{I}$. Consequently $\Phi_x(i)$ F.O.S.D. $\Phi_{x'}(i)$.*

The next theorem collects all the results we presented previously.

Theorem 1 *In this economy, we are able to show that firms with a higher total factor productivity, in equilibrium*

- *Are bigger;*
- *Hire more at all skill levels*
- *Hire proportionately more at the top*
- *Hire all the skill that firms with lower TFPs hire and may also hire workers with higher skills that their lower TFP counterparts – i.e. the support of skills hired by high TFP firms strictly contain the support of skills hired by low TFP firms*
- *Pay higher wages at all skill levels*

Before we move on to the description of the distributions of posted and earned wages of different skill types, let's present the average wage and standard deviation of wages within firms, as a function of the firm TFP. As we will see in later sections, these expressions will be important in our discussion of the decomposition of wage dispersion in within and between firms components. Based on our previous calculations, we easily can show that the average wage and variance of the within-firm wage distribution as a function of the firm's TFP x is given by

$$E_{\phi_x}[w] = \frac{\sum_{i=1}^{\mathcal{I}} K(x,i) l_x(i)}{S(x)} \tag{25}$$

and

$$Var_{\phi_x}[w] = \frac{\sum_{i=1}^{\mathcal{I}} (K(x,i) - E_{\phi_x}[w])^2 l_x(i)}{S(x)} \tag{26}$$

respectively.

2.3.1 Wage distributions by skill type

In this subsection, we will look how the equilibrium distributions of posted and earned wages vary across worker types. In this sense, we will be able to evaluate how workers with different abilities are faced with different job market opportunities and outcomes. In order to do that, we will initially focus on how firms' wage posting strategies vary with workers' skills. In particular, based on the optimal wage posting strategy presented in equation (22), we obtain the following result.

Lemma 3 $\frac{\partial K(x,i)}{\partial i} > 0$*, i.e., wages at any TFP level increase with skill level.*

Therefore firms with higher productivity offer higher wages at all skill levels, and workers with higher skill levels receive higher wage offers from 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 the Lemma 3, 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 skilled worker must have a lower total factor productivity.

Corollary 2 *If a given wage* w *is offered for both skills* i *and* j*,* $i > j$*, we must have that the firm offering wage* w *for skill level* i *has lower productivity than the firm offering the wage for skill* j*, i.e.,* $w = K(x, i) = K(x', j) \Rightarrow x < x'$.

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 are given by* $[b(i), K(\overline{x}, i)]$ *and* $[b(j), K(\overline{x}, j)]$*, where* $b(i) \geq b(j)$ *and* $K(\overline{x}, i) > K(\overline{x}, j)$.

Then, considering the distribution of posted wages by firms and how the firm's total factor 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, FOSD), we have

Corollary 4 *The average offered wage increases with skill, i.e., if* $i > j$*,* $E_{F_i}[w] \geq E_{F_j}[w]$.

Once we obtained these results for the distribution of posted wages, we are now able to present some key characteristics for the distributions of earned wages by skill levels. From our previous calculations, we have

$$G_i(w) = \frac{(d+\delta_i)F_i(w)}{\{d+\delta_i+\lambda(i)[1-\Gamma(K^{-1}(\underline{w}_i,i))](1-\delta_i)[1-F_i(w)]\}} \tag{27}$$

while the p.d.f. is given by

$$g_i(w) = \begin{cases} \frac{(d+\delta_i)\{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(K^{-1}(\underline{w}_i,i))]\}}{\{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(K^{-1}(\underline{w}_i,i))][1-F_i(w)]\}^2} f_i(w) & \text{if } w \in [\underline{w}_i, \overline{w}_i] \\ 0 & \text{otherwise} \end{cases} \tag{28}$$

We are now able to show the following proposition:

Proposition 3 *If $i > j$, $G_i(w)$ FOSD $G_j(w)$.*

Then, a simple corollary of the previous result is presented next.

Corollary 5 *The average earned wage increases with skill, i.e., if $i > j$, $E_{G_i}[w] \geq E_{G_j}[w]$.*

2.3.2 Economy-wide wage distributions

In this subsection, 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 of all, let's be very precise about the densities for each skill level i:

$$f_i(\tilde{w}) = \begin{cases} \left[\frac{\gamma(K^{-1}(\tilde{w},i))}{[1-\Gamma(K^{-1}(\underline{w}_i,i))]\frac{\partial K(K^{-1}(\tilde{w},i),i)}{\partial K^{-1}(\tilde{w},i)}}\right] & \text{if } \tilde{w} \in [\underline{w}_i, \overline{w}_i] \\ 0 & \text{otherwise} \end{cases} \tag{29}$$

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

$$F(\tilde{w}) = \frac{\sum_{i=1}^{\mathcal{I}}[1-\Gamma(K^{-1}(\underline{w}_i,i))]F_i(\tilde{w})}{\sum_{i=1}^{\mathcal{I}}[1-\Gamma(K^{-1}(\underline{w}_i,i))]} \tag{30}$$

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

$$f(\tilde{w}) = \frac{\sum_{i=1}^{\mathcal{I}}[1-\Gamma(K^{-1}(\underline{w}_i,i))]f_i(\tilde{w})}{\sum_{i=1}^{\mathcal{I}}[1-\Gamma(K^{-1}(\underline{w}_i,i))]} \tag{31}$$

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

$$E_F[w] = \int_{\underline{w}}^{\overline{w}} \tilde{w} f(\tilde{w}) d\tilde{w} = \frac{\sum_{i=1}^{\mathcal{I}} [1 - \Gamma(K^{-1}(\underline{w}_i, i))] \int_{\underline{w}_i}^{\overline{w}_i} \tilde{w} f_i(\tilde{w}) d\tilde{w}}{\sum_{i=1}^{\mathcal{I}} [1 - \Gamma(K^{-1}(\underline{w}_i, i))]} \tag{32}$$

Simplifying the expression and substituting the average offered wage per skill, we have

$$E_F[w] = \frac{\sum_{i=1}^{\mathcal{I}} [1 - \Gamma(K^{-1}(\underline{w}_i, i))] E_{F_i}[w]}{\sum_{i=1}^{\mathcal{I}} [1 - \Gamma(K^{-1}(\underline{w}_i, i))]} \tag{33}$$

Aggregated density of earned wages

In this case, we focus on the wage that workers actually are earning. 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 (10), we have that the aggregated cumulative distribution of earned wages is given by

$$G(\tilde{w}) = \frac{\sum_{i=1}^{\mathcal{I}} G_i(\tilde{w})(m_i - u_i)}{\sum_{i=1}^{\mathcal{I}} (m_i - u_i)} \tag{34}$$

while the corresponding density function is given by

$$g(\tilde{w}) = \frac{\sum_{i=1}^{\mathcal{I}} g_i(\tilde{w})(m_i - u_i)}{\sum_{i=1}^{\mathcal{I}} (m_i - u_i)} \tag{35}$$

Based on these results, the average aggregate wage in this economy is given by

$$E_G[w] = \int_{\underline{w}}^{\overline{w}} w g(w) dw = \int_{\underline{w}}^{\overline{w}} w \frac{\sum_{i=1}^{\mathcal{I}} g_i(w)(m_i - u_i)}{\sum_{i=1}^{\mathcal{I}} (m_i - u_i)} dw = \frac{\sum_{i=1}^{\mathcal{I}} (m_i - u_i) E_{G_i}[w]}{\sum_{i=1}^{\mathcal{I}} (m_i - u_i)} \tag{36}$$

where $E_{G_i}[w]$ can be derived after some manipulations and a change in variables considering $w = K(x, i)$:

$$E_{G_i}[w] = \int_{\underline{x}(i)}^{\overline{x}} K(x', i)\,(d + \delta_i) \times \left[\frac{d + \delta_i + \lambda(i)(1 - \delta_i)[1 - \Gamma(\underline{x}(i))]}{\{d + \delta_i + \lambda(i)(1 - \delta_i)[1 - \Gamma(x')]\}^2} \right] \frac{\gamma(x')}{[1 - \Gamma(\underline{x}(i))]} dx' \tag{37}$$

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

$$Var_G(w) = \int_{\underline{w}}^{\overline{w}} (w - E_G(w))^2 g(w) dw \tag{38}$$

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

$$Var_G(w) = \frac{1}{M - U} \int_{\underline{x}}^{\overline{x}} \sum_{i=1}^{\mathcal{I}} (K(x, i) - E_G(w))^2 \frac{\lambda(i)(1 - \delta_i)(d + \delta_i) m_i}{\{d + \delta_i + \lambda(i)(1 - \delta_i)[1 - \Gamma(x)]\}^2} \gamma(x) dx \tag{39}$$

Substituting equation (11), we can rewrite $Var_G(w)$ as

$$Var_G(w) = \frac{1}{M-U}\int_{\underline{x}}^{\overline{x}}\sum_{i=1}^{\mathcal{I}}(K(x,i)-E_G(w))^2 l_x(i)\gamma(x)dx \tag{40}$$

2.4 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-firms 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 the following:

$$\sigma^2 = \sum_{j=1}^{J} p_j\sigma_j^2 + \sum_{j=1}^{J} p_j(\overline{w}_j - \overline{\overline{w}})^2 \tag{41}$$

The first term on RHS of equation (41) 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 (41) represents the between-firms component of the wage variance. In this expression, $\overline{w}_j$ is the mean wage in firm j, and $\overline{\overline{w}}$ 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 xs, 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 $\Delta \to 0$, we have

$$\begin{array}{c}\text{Within-firm}\\ \text{component}\end{array} = \int_{\underline{x}}^{\overline{x}}\frac{S(x)}{M-U}Var_{\phi_x}(w)\gamma(x)dx \tag{42}$$

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

$$\begin{array}{c}\text{Between-firms}\\ \text{component}\end{array} = \int_{\underline{x}}^{\overline{x}}\frac{S(x)}{M-U}(E_{\phi_x}(w)-E_G(w))^2\gamma(x)dx \tag{43}$$

Therefore, the entire decomposition can be rewritten as

$$Var_G(w) = \int_{\underline{x}}^{\overline{x}}\frac{S(x)}{M-U}Var_{\phi_x}(w)\gamma(x)dx + \int_{\underline{x}}^{\overline{x}}\frac{S(x)}{M-U}(E_{\phi_x}(w)-E_G(w))^2\gamma(x)dx \tag{44}$$

where $Var_{\phi_x}(w)$, $E_{\phi_x}(w)$, and $E_G(w)$ are presented in equations (26), (25), and (36), respectively.

Moreover, we can adapt Lazear and Shaw's (2008) 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:

$$\sigma^2 = \sum_{i=1}^{I} m_i^e (\sigma_i^e)^2 + \sum_{i=1}^{I} m_i^e (\overline{w}_i^e - \overline{\overline{w}})^2 \tag{45}$$

Similar to the decomposition in terms of within and between firms, the first term in the RHS of equation (45) represents the within-skill component of the overall wage inequality, and the second term in the RHS is the between-skills 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, $\overline{w}_i^e$ is the average wage for employed workers with skill level i, and $\overline{\overline{w}}$ is the mean wage in the economy.

Mapping this decomposition to our model, we have $\overline{w}_i^e = E_{G_i}[w]$, where $E_{G_i}[w]$ is given by equation (37) and $(\sigma_i^e)^2$ is given by

$$(\sigma_i^e)^2 \equiv Var_{G_i}(w) = \begin{array}{c} \frac{(d+\delta_i)\{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(\underline{x}(i))]\}}{\lambda(i)(1-\delta_i)[1-\Gamma(\underline{x}(i))]} \times \\ \times \int_{\underline{x}(i)}^{\overline{x}} (K(x,i) - E_{G_i}[w])^2 \frac{\lambda(i)(1-\delta_i)\gamma(x)}{\{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)]\}^2} dx \end{array} \tag{46}$$

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

$$m_i^e = \frac{\frac{\lambda_i(1-\delta_i)[1-\Gamma(\underline{x}(i))]m_i}{d+\delta_i+\lambda_i(1-\delta_i)[1-\Gamma(\underline{x}(i))]}}{\sum_{j=1}^{\mathcal{I}} \frac{\lambda_j(1-\delta_j)[1-\Gamma(\underline{x}(j))]m_j}{d+\delta_j+\lambda_j(1-\delta_j)[1-\Gamma(\underline{x}(j))]}} \tag{47}$$

Consequently, the decomposition of overall wage variance in terms of within- and between-skills is given by

$$\begin{array}{c}\text{Within-skill} \\ \text{component}\end{array} = \sum_{i=1}^{\mathcal{I}} m_i^e Var_{G_i}(w) \quad \text{and} \quad \begin{array}{c}\text{Between-skills} \\ \text{component}\end{array} = \sum_{i=1}^{\mathcal{I}} m_i^e (E_{G_i}[w] - E_G[w])^2 \tag{48}$$

where m_i^e, $Var_{G_i}(w)$, $E_{G_i}[w]$, and $E_G[w]$ are given by equations (47), (46), (37), and (36), respectively. In the next sections we will use both within- and between-firms and within- and between-skills decompositions in order to define the sources of the increase in the wage dispersion between 1985 and 2009.

3 Calibration

In this section, we gauge the performance of the model with some quantitative results, focusing on its performance over time. Our model incorporates two-sided heterogeneity into a search framework and has important implications about the relationship between worker skills and firm productivities.

We assume that worker heterogeneity in terms of skill is well measured by educational attainment. This allows us to calibrate our model in terms of skill endowments directly from the data. However, calibrating firm-level heterogeneity in terms of productivity x is not as straightforward. Estimating a firm-level TFP distribution from establishment-level data is beyond the scope of this paper. Instead, we proceed in two different ways. First, we nonparametrically calibrate the implied output distributions $p(x, i)$ from the observed wage distributions conditional on skill levels. This approach has the benefit of generating output distributions for different educational attainment groups without having to impose a particular functional form on the matched output, as well as not having to impose any constraints on the TFP distribution. The drawback of this approach is that any effects of the changes in the output distribution on wage inequality cannot be clearly attributed to one source. In particular, we cannot disentangle the contributions of changes in the TFP distribution and changes in labor productivity of different skill groups. In our second approach, we impose a functional form for $p(x, i)$ and use estimates of TFP distribution for public companies from Imrohoroglu and Tuzel (2014). While this approach allows us to distinguish between the two possible channels, it has the caveat of relying on estimates of TFP distribution from a selected sample of firms and a particular functional form for $p(x, i)$.

We calibrate the steady state model for the U.S. economy at two different periods (1985 and 2009). We use these different calibrations in order to evaluate the contribution of changes in the productivity distribution, the educational attainment distribution, and labor market frictions to the increase in overall wage inequality and to the change in within-and between-skills groups components.

In order to calibrate our parameters, we use wage and educational data from the 1985 Current Population Survey's Merged Outgoing Rotation Groups (CPS MORG) and the 2009 American Community Survey (ACS). From the 1985 CPS MORG, we classify five different educational-attainment levels based on Jaeger's (1997) method, allowing us to compare the results with the classification presented at the 2009 ACS. We decided to work with the 2009 ACS instead of the 2009 CPS MORG because of the higher values for top-coding in earnings presented at the ACS. We also use data from the CPS MORG's of 1985 and 2009 in order to obtain data for the average employment to unemployment (EU) transition rates and the average transition from unemployment to employment (UE) for each education group.

We calibrate the parameters in our model considering one month as the unit of time. We choose the death rate d in order to match the average overall death probability rate at the median age of 37 throughout the entire period (1985-2009). Moreover, for each time period (1985 and 2009) we calibrate the job finding rate $\lambda(i)(1-\delta_i)(1-\Gamma(\underline{x}(i)))$ to match the average UE transition for each education group from the CPS data. Similarly, we calibrate the job destruction rate δ_i such that $\delta_i + d$ match the average EU transition rate in the data for each education group. The parameters for $b(i)$ are pinned down such that the lowest TFP firm makes zero profits by hiring any of the skills. Table 1 summarizes the calibrated parameters for labor market frictions, educational attainment, and unemployment benefits and/or home production income for both 1985 and 2009.

Table 1: Parameters

A: 1985

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$
m_i	0.1843	0.4039	0.1998	0.1234	0.0886
$\lambda(i)$	0.4782	0.4538	0.5459	0.5290	0.5236
δ_i	0.0584	0.0287	0.0233	0.0125	0.0093
$b(i)$	0.2220	0.3398	0.4359	0.5201	0.5965

B: 2009

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$
m_i	0.0936	0.2665	0.2917	0.2258	0.1223
$\lambda(i)$	0.5464	0.4870	0.5187	0.5031	0.4998
δ_i	0.0530	0.0274	0.0227	0.0123	0.0097
$b(i)$	0.1513	0.2267	0.2871	0.3396	0.3868

As we can observe in table 1, while labor frictions seemed to decrease for low-education groups – both high school dropouts and high school graduates observed an increase in λ_i and a decrease in δ_i between 1985 and 2009 – all other groups experienced an increase in labor market frictions.

3.1 Nonparametric Estimation of $p(x, i)$

We follow Bontemps et al. (2000) and Launov (2005) to estimate $p(x, i)$. In particular, we follow the steps presented below:

1. We nonparametrically (kernel estimation) estimate $G_i(w)$. We use Launov'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 the closest to the estimated ones but still satisfy the model's conditions, i.e., the distribution must be single-peaked and
$$3\kappa_i g_i(w)^2 - g_i'(w)\left[1 + \kappa_i G_i(w)\right] > 0 \tag{49}$$
where $\kappa_i = \frac{\lambda(i)[1-\Gamma(\underline{x}(i))](1-\delta_i)}{d+\delta_i}$.
3. We assume all firms hire all skill levels ($\Gamma(\underline{x}(i)) = 0$, $\forall i$).

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 productivity and productivity distribution as
$$p(x, i) = w + \frac{1 + \kappa_i G_i(w)}{2\kappa_i g_i(w)} \tag{50}$$

$$\gamma(p(x,i)) = \frac{2\kappa_i(1+\kappa_i)g_i(w)^3}{\left\{ \begin{array}{c} 3\kappa_i \left\{1+\kappa_1 G_i(w)\right\}^2 g_i(w)^2 \\ -g_i'(w)\{1+\kappa_i G_i(w)\}^3 \end{array} \right\}} \tag{51}$$

Note that the recovered productivity distributions will be 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 the firm productivity 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 those worker types. The single-peak feature guarantees that for wages higher than the mod-wage, this condition is satisfied just by the 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 productivity distributions so that we can get the wage densities exactly for workers with a college degree or more, both for 1985 and 2009. 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 satisfies conditions 2 and 3 in figures 1 and 2 for 1985 and 2009, respectively. As we can see, the adjusted distributions, serving as the inputs in our nonparametric calibration of the output distributions, are quite close to the empirical counterparts. Our largest "mismatch" is for the 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 2009. 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 nonparametric 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.

Finally, in terms of how much of the overall variation observed in the data our model can capture, table 2 compares the average wage and within-group standard deviation for the model and data in 1985 and 2009. As presented in table 2, the model does a decent job of capturing the wage variation in both time periods and across the educational groups, in particular, for the lower educational groups. Moreover, 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 Schonberg (2007) and use different methods to impute censored wages. All methods deliver results that are qualitatively the same.

Table 2: Comparison Model vs. Data
(Calculations using values in $1000 2009 USD)

A. Average Earned Wage

	Overall	HS Dropout	HS Graduate	Some College	College	Post-Graduate
1985 - Model	3.27	2.38	2.80	3.24	4.36	5.57
1985 - Data	3.27	2.38	2.78	3.20	4.31	5.55
2009 - Model	4.23	2.20	3.02	3.46	5.49	7.69
2009 - Data	4.21	2.23	3.02	3.45	5.48	7.67

B. St. Deviation of earned wages

	Overall	HS Dropout	HS Graduate	Some College	College	Post-Graduate
1985 - Model	2.08	1.10	1.30	1.59	2.52	3.67
1985 - Data	2.09	1.13	1.33	1.63	2.56	3.75
2009 - Model	4.00	.98	1.49	1.76	4.33	7.64
2009 - Data	4.22	1.00	1.51	1.77	4.68	8.41

3.2 Parametric Calibration of $p(x, i)$

In this case, we assume a production function $p(x, i)$ that has a functional form

$$p(x, i) = x^{\alpha} A(i)^{\beta} \tag{52}$$

We use the estimates for total factor productivity for public companies in the U.S. in order to obtain a distribution of TFP at two different time periods: 1985 and 2009. The estimates are from Imrohoroglu and Tuzel (2014) and are obtained by implementing the methodology of Olley and Pakes. We then apply a nonparametric Kernel estimation in order to pin down the distribution. Figure 3(a) presents the two density functions. As one can observe, the density for 2009 is close to a mean-spread preserve of the 1985 distribution. The 1985 distribution has a mean 0.7789 and standard deviation 0.3070, while the 2009 distribution has a mean 0.7725 and standard deviation 0.4372. This result is in line with what Faggio et al. (2010) observed for the U.K.

We pin down α and β such that the β is equal to the Bureau of Labor Statistics (BLS) estimate for the labor share in years 1985 and 2009 accordingly and that $\alpha + \beta = 1$ in both years. Consequently, we have that, in 1985 $\alpha = 0.386$ and $\beta = 0.614$ and in 2009, $\alpha = 0.417$ and $\beta = 0.583$. Note that β may not map into labor share in our model, depending on the dispersion in the earned wages and other endogenous outcomes such as employment distributions. Instead of interpreting β literally as the labor share, we vary it to explore the robustness of the model's quantitative features to changes in the production function.

We calibrate $A(i)$ in order to match the average wage for each education group in each period of the analysis (1985 and 2009). In figure 3(b) we present our results. As one can see, differently from the other skill groups, the labor productivity of high school dropouts decreased between 1985 and 2009, a circumstance which may partially explain the decrease in their average compensation. While all other skill groups have seen an increase in labor productivity, this increase has been more pronounced in higher educational-attainment groups, in particular post-graduates.

Finally, in figure 4 we test the model's performance in matching the data by comparing the model generated standard deviation of wages for each skill group and each time period (1985 and 2009) against their empirical counterparts. As previously mentioned, the calibration of $A(i)$ was made in order to match the average earned wages per skill group, so we use the standard deviation in order to evaluate the model fit. As we can see, the model performs quite well at the lower levels of education. While the model properly detects the increase in wage dispersion across the two time periods, it is unable to generate a large portion of the overall within-skill wage dispersion that is observed in the data for college graduates and post-graduates.

4 Nonparametric Approach

As we discussed before, while the results based on our parametric approach have the advantage of allowing us to discuss within-and between-firms wage inequality, it comes at the cost of depending on the TFP estimates based on a biased sample and the imposition of a particular functional form for the production function. In this section, we consider the results based on the nonparametric calibration of the output flow $p(x, i)$ following Bontemps et. al. (2000). This methodology allows us to use the wage distribution in order to recover the productivity distribution, circumventing the need for firm-level data. As a drawback, we are unable to separately estimate x and $p(x, i)$. Consequently, our estimates are not directly related to a given firm. However, we are still able to discuss within- and between-group wage dispersion in terms of skill groups.

4.1 Benchmark results: 1985 and 2009

We now present our results for the calibrations for 1985 and 2009. First of all, as we can see in figure 5, in terms of changes in the output distribution $p(x, i)$ there is clear evidence of FOSD for high skill levels.

This is more pronounced for skill levels 4 (college graduates) and 5 (post graduates). On the contrary, for lower levels of educational attainment, we see either minor movements (for high school graduates and some college) or a technical change that may have even hindered productivity (for high school dropouts). Moreover, as we can see in figure 6, while the FOSD can be clearly seen in panel (a), with the larger increases in average output for college graduates and post-graduates, panel (b) shows that this increase was accompanied by a large increase in within-group output inequality for high-skill workers. In fact, while college graduates and post-graduates have seen an increase in the standard deviation of output flow of 28.62 and 31.03 percent, respectively, between 1985 and 2009, workers with some college have seen an increase in output standard deviation of only 8.75 percentage points. Interestingly, high school graduates and high school dropouts have actually seen a drop in output dispersion of -0.24 and -10.92 percent, respectively.

In terms of changes in the wages per worker as a function of educational attainment, results are as expected. The model shows that while average wages have increased significantly for college graduates and post-graduates – by 25.72 and 38.01 percent, respectively – it moved by just 7.75 percentage points for high school graduates and 6.65 percent for workers with some college, and it went down by 7.69 percent for high school dropouts. In terms of the within-group standard deviation, we see that wage dispersion has increased significantly for high-skill workers while moving only slightly at the middle of the educational attainment distribution, and, has even gone down among high school dropouts. In particular, we have seen an increase in the standard deviation of earned wages of 71.98 and 108.52 percent between 1985 and 2009 for college graduates and post-graduates, respectively. In contrast, high school graduates and workers with some college have seen an increase in standard deviation wages by only 15.02 and 10.86 percentage points, respectively. Moreover, high school dropouts have actually seen a drop in wage dispersion by -10.92 percent.

Table 3: Variance Decomposition
(Calculations using values in $1000 2009 USD)

Skills - Magnitudes

	Overall	**Within-group**	**Between-group**
1985 - Model	4.32 (100%)	3.45 (80%)	0.87 (20%)
1985 - Data	4.35 (100%)	3.58 (82%)	.77 (18%)
2009 - Model	16.02 (100%)	13.23 (83%)	2.79 (17%)
2009 - Data	16.02 (100%)	13.23 (85%)	2.79 (15%)
% Change 1985-2009 : Model	270.61%	283.06%	221.18%
% Change 1985-2009 : Data	268.44%	269.17%	265.04%

Table 3 presents the overall wage variance in the data and in the model for both time periods, along with the breakdown in terms of its within- and between-group components. The percentages presented between parentheses underneath each number correspond to the fraction of the overall variance that is explained by each component. First of all, we would like to highlight the fact that our nonparametric calibration captures quite well the overall wage dispersion observed in the data and the decomposition of this wage dispersion across its within- and between-group components. 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 nonparametric calibration seems a good starting point to our counterfactual exercises. In terms of the results presented in table 3, we see a large increase in wage dispersion between 1985 and 2009. Moreover, while there is a slightly larger increase in the within-group component, the overall breakdown between within- and between-group components stayed reasonably stable across the two time periods. However, this breakdown is heavily influenced by the tail estimates of censored observations, so caution is necessary when interpreting the results.

4.2 Counterfactual experiments

In this section, we consider a few counterfactual exercises. In principle, our model has three sets of parameters/inputs: labor market frictions, defined by $\lambda_i, \delta_i, d, i \in \{1, 2, ..., 5\}$; educational attainment distributions, characterized by $m_i, i \in \{1, 2, ..., 5\}$; and output per worker-skill $p(x, i)$ distributions, obtained through the nonparametric estimation using earned wage data. In our counterfactual exercises, we will 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 inputs separately. In order to do that, we start from the 1985 benchmark calibration inputs and consider that only one of the three sets of inputs (labor market frictions, educational distribution, and output distributions) is changed to its 2009 counterpart. We then compare the obtained wage distribution against the 1985 and 2009 benchmark results. Moreover, in some cases we also consider the additional contribution of moving another input toward its 2009 values in terms of how much closer we get to the 2009 wage distributions.

Table 4: St. Deviation Decomposition - within-group
(Values in \$1000 2009 USD)

	HS Dropout	**HS Grads**	**Some College**	**College Grads**	**Post Grads**
1985 Benchmark	1.10	1.29	1.59	2.52	3.67
2009 Benchmark	.98	1.49	1.76	4.33	7.64
Difference	-0.12	0.19	0.17	1.81	3.98
Counterfactuals:					
Production	234.25%	38.81%	118.26%	110.30%	117.06%
Labor	-154.47%	49.90%	-15.56%	-4.65%	-6.83%
Education	0.00%	0.00%	0.00%	0.00%	0.00%
Production + Labor	100%	100%	100%	100%	100%

Table 4 and figure 8 summarize the results. Table 4 shows the standard deviation of earned wages for each educational group in 1985 and in 2009 and their difference. Then, it shows how much of this difference can be explained by each counterfactual channel. Notice that the production counterfactual, that keeps labor market frictions and the educational attainment distribution at their 1985 levels while moving the firm-skill output distributions to their 2009 levels, explains more than 100 percent of the difference for all but high school graduates. This result shows that the technological change not only explains the bulk of the result, but also that changes in the labor market frictions helped to attenuate the impact of SBTC. We would also like to highlight that once the change in standard deviation of earned wages was negative for high school dropouts between 1985 and 2009, a percentage above 100 percent in the production scenario implies that we should have expected an even bigger drop in wage dispersion for high school dropouts in this scenario. Panel (b) in figure 8 gives a visual perspective for the result. Moreover, while the scenario in which we update only the educational attainment distribution seems to not affect the within-group dispersion, it actually affects the overall dispersion and the break down between within- and between-groups, as one can see in table 5.

Table 5 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 groups. As in table 4, the percentages in the counterfactual rows show how much of the difference between the 1985 and 2009 values is observed in each counterfactual scenario. As one can see, 79.20

percent of the overall increase in wage dispersion can be explained by just the change in output distributions. Moreover, we observe that there is a clear interaction between education and output parameters that boostw between-group wage dispersion, while labor market frictions again act in the opposite direction.

Table 5: Variance Decomposition - Skills
(Calculations using values in $1000 2009 USD)

	$Var_G(w)$	*Within-group*	*Between-group*
1985	4.32	3.45	0.87
2009	16.02	13.23	2.79
Difference	11.70	9.78	1.92
Counterfactuals:			
Production	79.20%	72.60%	112.75%
Labor	-2.63%	-0.80%	-11.88%
Education	9.25%	10.15%	4.68%
Education + Production	119.89%	117.59%	131.58%

Finally, table 6 shows the impact of the changes in the output distribution, educational attainment, and labor market frictions between 1985 and 2009 on the average wages earned by different skill groups and the average wage in the overall economy. Results are in line with what we found for wage dispersion. Changes in production clearly boosted the wages for high-skill workers while helping to depress the average wage of low-skill workers. Conversely, changes in the labor market frictions seem to partially undo the effect of the technological change. While changes in the education distribution are not relevant for the average wage of particular educational groups, it is quite relevant to explain the increase in overall average wages, in particular once combined with the technological changes.

Table 6: Average Earned Wages
(Values in \$1000 2009 USD)

	HS Dropout	HS Grad	Some College	College	Post Graduate	Overall
1985 Benchmark	2.38	2.80	3.24	4.37	5.57	3.27
2009 Benchmark	2.20	3.02	3.46	5.49	7.69	4.23
Difference	-0.18	0.22	0.22	1.12	2.12	0.95
Counterfactuals:						
Production	196.86%	33.65%	118.02%	111.47%	119.69%	43.36%
Labor	-133.32%	63.67%	-15.88%	-7.23%	-10.66%	5.79%
Education	0.00%	0.00%	0.00%	0.00%	0.00%	34.62%
Education + Labor	-133.32%	63.67%	-15.88%	-7.23%	-10.66%	34.26%
Production + Labor	100.00%	100.00%	100.00%	100.00%	100.00%	44.98%
Education + Production	196.86%	33.65%	118.02%	111.47%	119.69%	105.01%

5 Parametric Approach

In this section, we present our results for the parametric approach. As previously mentioned, the main benefit of following a parametric approach is to allow us to disentangle the overall effect of the changes in output per skill-TFP pair – $p(x, i)$ – over the wage distribution. In particular, we would like to know how much of the increase in average wages and wage inequality can be attributed to changes in the TFP distribution as opposed to changes in the labor productivity of different skill groups, $A(i)$, over time. Again, the main drawback of the parametric approach is not only the fact that we must impose a particular functional form to the output function, but also that we need to rely on estimates for TFP coming from a selected sample of firms that is biased toward large firms.

5.1 Benchmark results: 1985 and 2009

We now present our results for the calibrations targeting 1985 and 2009 data moments. First of all, we calibrate $A(i)$ and use Imrohoroglu and Tuzel (2014) in order to pin down the distribution of x. Our results on $p(x, i)$ were already partially discussed in section 3.2, so here we focus on comparisons of the wage distribution across educational groups and firms across our two time periods. As seen in figure 9, average wages are monotonically increasing with TFP in each period of time, as expected because of lemma 3. However, the changes across the two time periods are strikingly different. While average wages increased quite a lot at the firms with the highest TFPs, they have changed little or decreased at low and average TFP firms. Moreover, the range of wages paid by high-TFP firms increased significantly between 1985 and 2009, while it had changed much less at the low-TFP firms. This pattern is seen clearly in figure 9(c), with the relationship between TFP and within-firm wage distribution becoming steeper in 2009 in comparison to that of 1985.

In terms of the wage dispersion across educational groups, figures 9(b) and 9(d) present our results for the benchmark calibrations of 1985 and 2009. Notice that average wages increase with educational attainment in both periods, but we also see that average wages have decreased for high school dropouts while increasing for all other skill groups. Moreover, the increase has been particularly large among college graduates and post-graduates. However, as discussed in section 3.2, these results come directly from the empirical evidence, as the model is calibrated in order to match the average wage at each skill group. In terms of the within-group wage dispersion, figure 9(d) shows that while within-group wage inequality increases with educational attainment both in 1985 and 2009, the increase in within-group wage dispersion between 1985 and 2009 has been concentrated in the high-education groups. Table 7 presents the model's variance decomposition in terms of within and between-firms, and within and between-groups. While our data does not allow us to compare the between-and within-firm components against their empirical counterparts, as in section 4.1, we can compare the within-and between-group components against their counterparts directly calculated from the data. Different from the nonparametric case, the parametric calibration is unable to match the empirical counterpart in this dimension. This shortcoming is likely to come from the biased TFP distribution that we use to calibrate the model. As demonstrated by Burdett and Mortensen (1998), low-TFP firms do not compete with high-TFP firms in equilibrium, focusing on attracting unemployed workers. Consequently, in equilibrium this lower tail of the TFP distribution has a significant impact on the within-group wage dispersion, in particular for the high-skill groups. Moreover, the fact that we work with the conditional TFP distribution, i.e., conditional on being a public (and consequently large) firm, implies that the TFP input overestimates the competition among high-TFP firms, boosting even more within-skill wage dispersion. Finally, our assumptions about the production function may be too restrictive to capture the possible spread of output across (x, i) pairs. In any case, we would suggest the reader to keep in mind the possible shortcomings of this analysis while we describe the results of our counterfactual exercises.

5.2 Counterfactual experiments

In this section, we consider the counterfactual experiments for the parametric approach. In the parametric case, we are able not only to consider similar counterfactuals in terms of changes in the labor market frictions and educational attainment distribution, but also to break down the counterfactual on output distributions $p(x, i)$ in terms of changes in the TFP distribution $\gamma(x)$ and changes in the labor productivity of different skill levels $A(i)$. Moreover, since we assumed a parametric functional form for $p(x, i) = x^{\alpha} A(i)^{\beta}$, we can also evaluate the impact of changes in α and β on wage averages and dispersion to explore the robustness of our results. Finally, the parametric case allows us to discuss not only the impact of the counterfactual experiments on wage averages and dispersion across skill groups,

Table 7: Variance Decomposition

Firm Decomposition

	Within-group	**Between-group**
1985	0.48	0.52
2009	0.60	0.40

Skill Decomposition

	Within-group	**Between-group**
1985 - Model	0.29	0.71
1985 - Data	0.81	0.19
2009 - Model	0.21	0.79
2009 - Data	0.77	0.23

but also in terms of changes across firms. Our results are graphically presented in figures 10 and 11, while tables 8 and 9 present the results for average earned wages and wage dispersion across skills.

In terms of the average wages across skill groups, we can see in table 8 that, similarly to the nonparametric case, changes in labor market frictions helped high school dropouts, while having only a minor impact over highly skilled workers. Moreover, changes in the education distribution again seem to have no impact on average wages across different educational groups. It does, however, affect the overall wage average and dispersion. In terms of the impact of changes in output, we can now split the effect of changes in TFP and changes in labor productivity on average wages earned by groups with different educational attainment. Interestingly, from table 8 we see that the effect of the changes in TFP was actually negative for groups with high educational attainment. This implies that the mean preserving spread that occurred in the TFP distribution between 1985 and 2009 had on average adversely affected highly skilled workers. On the other hand, we see a large increase in labor productivity for skilled workers that seem to have more than compensated the losses due to changes in TFP.

A similar pattern emerges from the picture of the impact of the counterfactuals over the average wages paid by firms with different levels of TFP, presented in figure 11. While changes in labor frictions and in educational attainment distributions seem to have only a second order effect on average wages paid by firms with different TFP levels, the changes in labor productivity boosted average wage paid by high-TFP firms, while changes in the TFP distribution seem to deeply depress the wage paid by highly productive firms.

Table 8: Average Earned wages
(Values in $1000 2009 USD)

	HS Dropout	HS Graduate	Some College	College	Post-Graduate
1985 - Benchmark	2.34	2.83	3.31	4.52	5.85
2009 - Benchmark	2.19	3.01	3.48	5.46	7.60
Difference	-0.15	0.17	0.17	0.93	1.75
Counterfactuals:					
TFP	58.00%	-108.49%	-174.26%	-73.52%	-64.52%
Labor	-34.09%	33.32%	18.20%	0.12%	0.91%
Education	0.00%	0.00%	0.00%	0.00%	0.00%
Labor Productivity	60.37%	138.68%	164.80%	128.37%	123.87%
Production Fcn	85.18%	-86.11%	-105.56%	-29.42%	-23.20%

In terms of the changes in wage dispersion, Table 9 shows that none of the counterfactuals can single-handedly explain the observed changes, in particular for low-skill workers. As we can see from both table 9 and figure 10, both TFP and labor productivity can explain a large share of the increase in wage dispersion for high-skill groups (college graduates and post graduates). However, the magnitude of the effect of labor productivity is significantly larger. Moreover, once we look at the wage dispersion across firms in figure 11, we again obtain a similar result to the one for average wages, with labor productivity significantly increasing the dispersion of wages paid by high-TFP firms, while changes in the TFP distribution reduce the wage dispersion at high-TFP firms.

Table 9: St. Deviation Decomposition - within-group
(Calculations using values in $1000 2009 USD)

	HS Dropout	HS Graduate	Some College	College	Post-Graduate
1985 - Benchmark	0.54	0.57	0.62	0.77	0.95
2009 - Benchmark	0.55	0.65	0.72	1.02	1.36
Difference	0.01	0.08	0.09	0.26	0.41
Counterfactuals:					
TFP	-75.76%	9.02%	28.62%	52.68%	69.13%
Labor	-42.73%	-5.53%	-2.42%	-0.02%	-0.09%
Education	0.00%	0.00%	0.00%	0.00%	0.00%
Labor Productivity	-387.74%	58.65%	55.99%	79.13%	86.92%
Production Fcn	-44.37%	0.16%	-0.00%	-0.85%	-2.34%

In summary, the results from the counterfactual experiments for the parametric calibration suggest that the pattern of increase in average wages and wage dispersion in highly educated groups has been driven mostly by an increase in the labor productivity of these individual groups. In fact, changes in TFP distribution seem to have a countervailing effect, at least in terms of average wages. Moreover, once we look at average wages and wage dispersion across firms, we observe that increases in labor productivity in particular at high-skill groups have boosted the average wages and the wage dispersion at high-TFP firms. Conversely, changes in TFP seem to have the exact opposite effect.

6 Conclusion

In this paper, we present a model that can 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 in organizations as equilibrium properties. We calibrate the model both parametrically and nonparametrically using wage data from the CPS MORG and ACS for the years 1985 and 2009. Our results show that technological change 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. Changes in labor market frictions helped to partially undo the effect of technology. Changes in the educational attainment, while not affecting the average wage and wage dispersion within particular educational groups, helped to boost overall wage dispersion and average wages. In terms of the components of wage dispersion, we see that the increase in overall wage dispersion was mostly due to an increase in the within-group component.

Overall, our results suggest that SBTC was a major factor for the increase in wage inequality between educational groups. Our results for the parametric calibration using TFP estimates for public companies suggest that most of the increase in wage dispersion and inequality across firms is due to an increase in labor productivity, in particular for highly educated workers. Moreover, changes in the TFP distribution between 1985 and 2009, although contributing to the overall wage dispersion, had actually a negative effect on average wages for all but high school dropouts. Moreover, changes in labor productivity across educational groups and across the two time periods seem to have driven up not only average wages paid by high-TFP firms, but also their wage dispersion, consequently increasing both the within-and between-firm components of wage dispersion.

References

Acemoglu, Daron and David Autor "Skills, Tasks and Technologies: Implications for Employment and Earnings," in *Handbook of Labor Economics*, **Vol. 4B**, Chapter 12, 2011, 1043-1278.

Albrecht, Jim and Susan Vroman, " A Matching Model with Endogenous Skill Requirements," *International Economic Review*, **43**, 2002, 283-305.

Autor, David, Lawrence Katz, and Melissa Kearney, "Rising Wage Inequality: The Role of Composition and Prices," *NBER Working Paper* no. 11628, mimeo.

Barth, Erling, Alex Bryson, James Davis, Richard Freeman, "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)**, 2016, S67-S97.

Bontemps, Christian, Jean-Marc Robin, and Gerard van der Berg , "An empirical equilibrium job search model with search on the job and heterogeneous workers and firms," *International Economic Review,* **40**, 1999, 305-358.

Bontemps, Christian, Jean-Marc Robin, and Gerard van der Berg , "Equilibrium Search with Continuous Productivity Dispersion: Theory and Non-parametric Estimation," *International Economic Review,* **41**, 2000, 1039-1074.

Buchinsky, Moshe, "Changes in the US wage structure 1963E987: application of quantile regression," *Econometrica,* **62**, 1994, 405-458.

Buchinsky, Moshe, "The dynamics of changes in the female wage distribution in the USA: a quantile regression approach," *Journal of Econometrics*, **13**, 1998, 1-30.

Budría, Santiago and Pedro Telhado-Pereira, "Educational Qualifications and Wage Inequality: Evidence for Europe," *Revista de Economía Aplicada*, **56(XIX)**, 2011, 5-34.

Burdett, Kenneth and Dale Mortensen, "Wage Differential, Employer Size, and Unemployment," *International Economic Review,* **39**, 1998, 257-273.

Caliendo, Lorenzo, Ferdinando Monte and Esteban Rossi-Hansberg, "The Anatomy of French Production Hierarchies," *Journal of Political Economy*, **123(4)**, 2015, 809-852.

Card, David and John DiNardo, "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles," *Journal of Labor Economics,* **20**, 2002, 733-783.

Colombo, Massimo and Marco Delmastro, “Some stylized facts on organization and its evolution,” *Journal of Economic Behavior & Organization,* **40**, 1999, 255-274.

Dessein, Woulter and Tano Santos “Adaptive Organizations,” *Journal of Political Economy,* **114(5)**, 2006, 956-995.

Dolado, Juan, Marcel Jansen, and Juan Jimeno, “On-the-job Search in a Matching Model with Heterogeneous Jobs and Workers,” *Economic Journal,* **119(534)**, 2009, 200-228.

Fallick, Bruce and Charles Fleischman, “Employer-to-Employer Flows in the U.S. Labor Market: The Complete Picture of Gross Worker Flows,” 2004, mimeo.

Faggio, Giulia, Kjell Salvanes, and John Van Reenen, “The evolution of inequality in productivity and wages: panel data evidence, ”*Industrial and Corporate Change* **19(6)**, 2010, 1919-1951.

Garicano, Luis,“Hierarchies and the Organization of Knowledge in Production,” *Journal of Political Economy* **108(5)**, 2000, 874-904.

Garicano, Luis, and Esteban Rossi-Hansberg, “Organization and Inequality in a Knowledge Economy,” *Quarterly Journal of Economics,* **121(4)**, 2006, 1383-1435.

Gosling, Amanda, Stephen Machin, and Costas Meghir, “The Changing Distribution of Male Wages in the U.K.,” *Review of Economic Studies*, **67(4)**, 2000, 635-666.

Hartog, Joop, Pedro Pereira, and José Vieira, “Changing returns to education in Portugal during the 1980s and early 1990s: OLS and quantile regression estimators,” *Applied Economics*, **33(8)**, 2001, 1021-1037.

Imrohoroglu, Ayse and Selale Tuzel, “Firm Level Productivity, Risk, and Return,” *Management Science* **60(8)**, 2014, 2073 - 2090.

Lauer, Charlotte, “Education, gender and earnings in France and Germany: level and dispersion effects,” ZEW Discussion paper 04-54, 2004.

Launov, Andrey, “Structural Estimation of Search Equilibrium Models with Wage Posting,” 2005, mimeo.

Lazear, Edward and Kathryn Shaw, “Wage Structure, Raises and Mobility: International Comparisons of the structure of wages within and across firms,” NBER Working Paper 13654, 2007, mimeo.

Lemieux, Thomas, "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skills?" *American Economic Review,* **96**, 2006, 461-498.

Machado, José and José Mata, "Earnings functions in Portugal 1982-1994: Evidence from quantile regressions," *Empirical Economics*, **26**, 115-134.

Maier, Michael, Friedhelm Pfeiffer, and Winfried Pohlmeier, "Returns to Education and Individual Heterogeneity," 2004, mimeo.

Nagypal, Eva, "Worker Reallocation Over the Business Cycle: The Importance of Job-to-Job Transitions," University of Northwestern, mimeo, 2008.

Olley, G. S., and A. Pakes, "The dynamics of productivity in the telecommunications equipment industry," *Econometrica,* **64**, 1996, 1263-1297.

Papageorgiou, Theodore "Large Firms and Within Firm Occupational Reallocation ", mimeo.

Song, Jae, David Price, Fatih Guneven, Nicholas Bloom, and Till von Wachter, "Firming up Inequality," 2016, mimeo.

Tag, Joacin, "Production Hierarchies in Sweden," *Economic Letters,* **121**, 2013, 210-213.

Uren, Lawrence and Gabor Virag, "Skill Requirements, Search Frictions, and Wage Inequality," *International Economic Review,* **52**, 2011, 379-406.

Appendix

Proof of Lemma 1

Proof. Taking the derivative with respect to x, we have

$$\frac{\partial K(x,i)}{\partial x} = 2[d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)]]\lambda(i)(1-\delta_i)\gamma(x)\times$$
$$\times\int_{\underline{x}(i)}^{x}\frac{\partial p(x',i)}{\partial x'}\left(\frac{1}{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x')]}\right)^2 dx'$$

Consequently, $\frac{\partial K(x,i)}{\partial x} > 0$. ■

Proof of Lemma 2

Proof. Taking the derivative with respect to x, we get

$$\frac{\partial l_i(x)}{\partial x} = \frac{2\lambda(i)(1-\delta_i)(d+\delta_i)m_i\{d+\delta_i+\lambda(i)(1-\delta_i)(1-\Gamma(x))\}\lambda(i)(1-\delta_i)\gamma(x)}{\{d+\delta_i+\lambda(i)(1-\delta_i)(1-\Gamma(x))\}^3} > 0.$$

■

Proof of Proposition 1

Proof. Based on the results obtained up to now, we know that the support of skills for a type x firm is $\{1,...,\mathbf{I}(x)\}$ while the support of skills hired by firm x' is $\{1,...,\mathbf{I}(x')\}$ 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 $\Phi_x(i) \leq \Phi_{x'}(i)$. Let's then now consider the case in which $i < \mathbf{I}(x')$. In this case, both firms hire this particular skill. Since the distributions are discrete, let's focus on the p.d.f.s $\phi_x(i)$ and $\phi_{x'}(i)$. Then, notice that

$$\phi_x(i) = \frac{l_x(i)}{S(x)}$$

As we showed before

$$\frac{dl_i(x)}{dx} = \frac{2\lambda(i)(1-\delta_i)\gamma(x)}{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)]}l_i(x)$$

and

$$S'(x) = \sum_{j=1}^{I}\frac{2\lambda(j)(1-\delta_j)\gamma(x)}{d+\delta_j+\lambda(j)(1-\delta_j)[1-\Gamma(x)]}l_j(x)$$

Then

$$\frac{\partial\phi_x(i)}{\partial x} = \frac{1}{S(x)^2}\sum_{j=1}^{I}2\gamma(x)l_i(x)l_j(x)\left[\frac{\lambda(i)(1-\delta_i)}{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)]} - \frac{\lambda(j)(1-\delta_j)}{d+\delta_j+\lambda(j)(1-\delta_j)[1-\Gamma(x)]}\right]$$

Rearranging it, we obtain

$$\frac{\partial\phi_x(i)}{\partial x} = \frac{1}{S(x)^2}\sum_{j=1}^{I}2\gamma(x)l_i(x)l_j(x)\left\{\frac{d[\lambda(i)(1-\delta_i)-\lambda(j)(1-\delta_j)]+\delta_j\lambda(i)(1-\delta_i)-\delta_i\lambda(j)(1-\delta_j)}{[d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)]]\times[d+\delta_j+\lambda(j)(1-\delta_j)[1-\Gamma(x)]]}\right\}$$

So the sign of the derivative depends on the numerator of the terms within curly brackets. Keep in mind that if $i > j$ we have that $\delta_i < \delta_j$, and we are also keeping the assumption that $\lambda(i) \geq \lambda(j)$. Consequently, we have that $\lambda(i)(1-\delta_i) > \lambda(j)(1-\delta_j)$ and $\delta_j\lambda(i)(1-\delta_i) > \delta_i\lambda(j)(1-\delta_j)$. So, the terms in the summation in which $i < j$ are negative, while the terms in which $i > j$ are positive. As expected if $i = j$, the term is zero.

Based on these results, we have that $\frac{\partial\phi_x(1)}{\partial x} < 0$ and $\frac{\partial\phi_x(I)}{\partial x} \geq 0$ (with inequality if $x > \underline{x}(I)$). Now let's consider the C.D.F. Since the variable is discrete, the C.D.F. is a step function. We therefore just need to see how the steps move up or down with x. Since we showed that $\frac{\partial\phi_x(1)}{\partial x} < 0$, we know that the first step moves down with x. Let's then consider the second step:

$$\frac{\partial\phi_x(1)}{\partial x} + \frac{\partial\phi_x(2)}{\partial x} = \frac{1}{S(x)^2}\sum_{i=1}^{2}\sum_{j=3}^{I} 2\gamma(x)l_i(x)l_j(x)\left\{\frac{\begin{bmatrix} d[\lambda(i)(1-\delta_i)-\lambda(j)(1-\delta_j)]+ \\ +\delta_j\lambda(i)(1-\delta_i)-\delta_i\lambda(j)(1-\delta_j)\end{bmatrix}}{\begin{bmatrix} [d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)]]\times \\ \times[d+\delta_j+\lambda(j)(1-\delta_j)[1-\Gamma(x)]]\end{bmatrix}}\right\}$$

where, since the positive term in the summation for $i = 2$ cancels out with a negative term in $i = 1$, we have only negative terms left, and $\frac{\partial\phi_x(1)}{\partial x} + \frac{\partial\phi_x(2)}{\partial x} < 0$. Consequently, the second step also goes down with x. The same argument can be made for the next step. In fact, considering a given $i' < I$, we have that

$$\sum_{i=1}^{i'}\frac{\partial\phi_x(i)}{\partial x} = \frac{1}{S(x)^2}\sum_{i=1}^{i'}\sum_{j=i'+1}^{I} 2\gamma(x)l_i(x)l_j(x)\left\{\frac{\begin{bmatrix} d[\lambda(i)(1-\delta_i)-\lambda(j)(1-\delta_j)]+ \\ +\delta_j\lambda(i)(1-\delta_i)-\delta_i\lambda(j)(1-\delta_j)\end{bmatrix}}{\begin{bmatrix} [d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)]]\times \\ \times[d+\delta_j+\lambda(j)(1-\delta_j)[1-\Gamma(x)]]\end{bmatrix}}\right\} < 0$$

Consequently, increasing x pushes the steps down at any i. Therefore, we have that $\Phi_x(i) \leq \Phi_{x'}(i)$, $\forall i \in \mathcal{I}$, implying that $\Phi_x(i)$ F.O.S.D. $\Phi_{x'}(i)$. ■

Proof of Lemma 3

Proof.

$$\begin{gathered}\frac{\partial K(x,i)}{\partial i} = \frac{\partial p(x,i)}{\partial i} - \int_{\underline{x}(i)}^{x}\frac{\partial^2 p(x',i)}{\partial i\partial x'}\left(\frac{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)]}{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x')]}\right)^2 dx' \\ +2\int_{\underline{x}(i)}^{x}\frac{\partial p(x',i)}{\partial x'}\left(\frac{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)]}{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x')]}\right)\left(\frac{(d+\delta_i)\lambda'(i)(1-\delta_i)-(1+d)\delta_i'\lambda(i)}{(d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x'^2}\right)(\Gamma(x)-\Gamma(x'))dx' \\ +\frac{\partial p(\underline{x}(i),i)}{\partial\underline{x}(i)}\left(\frac{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)]}{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(\underline{x}(i))]}\right)^2\frac{d\underline{x}(i)}{di}.\end{gathered}$$

Notice that

$$\int_{\underline{x}(i)}^{x} \frac{\partial^2 p(x',i)}{\partial i \partial x'} \left(\frac{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)]}{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x')]} \right)^2 dx'$$
$$= \left\{ \begin{array}{c} \frac{\partial p(x',i)}{\partial i} \left(\frac{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)]}{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x')]} \right)^2 |_{\underline{x}(i)}^{x} \\ -2\int_{\underline{x}(i)}^{x} \frac{\partial p(x',i)}{\partial i} \left[\frac{(d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)])^2}{(d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x'^3}\right] \lambda(i)(1-\delta_i)\gamma(x')dx' \end{array} \right\}$$
$$= \left\{ \begin{array}{c} \frac{\partial p(x,i)}{\partial i} - \frac{\partial p(\underline{x}(i),i)}{\partial i} \left(\frac{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)]}{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(\underline{x}(i))]} \right)^2 \\ -2\int_{\underline{x}(i)}^{x} \frac{\partial p(x',i)}{\partial i} \left[\frac{(d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)])^2}{(d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x'^3}\right] \lambda(i)(1-\delta_i)\gamma(x')dx' \end{array} \right\}$$

Substituting it back gives

$$\frac{\partial K(x,i)}{\partial i} = \left\{ \begin{array}{c} \frac{\partial p(\underline{x}(i),i)}{\partial i} \left(\frac{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)]}{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(\underline{x}(i))]} \right)^2 \\ +2\int_{\underline{x}(i)}^{x} \frac{\partial p(x',i)}{\partial i} \left[\frac{(d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)])^2}{(d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x'^3}\right] \lambda(i)(1-\delta_i)\gamma(x')dx' \\ +2\int_{\underline{x}(i)}^{x} \frac{\partial p(x',i)}{\partial x'} \left(\frac{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)]}{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x')]} \right) \left(\frac{(d+\delta_i)\lambda'(i)(1-\delta_i)-(1+d)\delta_i'\lambda(i)}{(d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x'^2} \right) (\Gamma(x)-\Gamma(x'))dx' \\ +\frac{\partial p(\underline{x}(i),i)}{\partial \underline{x}(i)} \left(\frac{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(x)]}{d+\delta_i+\lambda(i)(1-\delta_i)[1-\Gamma(\underline{x}(i))]} \right)^2 \frac{d\underline{x}(i)}{di} \end{array} \right\}.$$

So, if $\lambda'(i) \geq 0$ we have that $\frac{\partial K(x,i)}{\partial i} > 0$. ■

Proof of Proposition 2

Proof. From Lemmas 1 and 3, 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 are $[b(i), K(\overline{x},i)]$ and $[b(j), K(\overline{x},j)]$, where $b(i) \geq b(j)$ and $K(\overline{x},i) > K(\overline{x},j)$. 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 $w \in (K(\overline{x},j), K(\overline{x},i))$, we have that $F_j(w) = 1$ and $F_i(w) < 1$, so $F_i(w) \leq F_j(w)$. Finally, if $w \in [b(i), K(\overline{x},j)]$, so the wage is offered to both type i and j workers. From previous calculations, we have that

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

From Corollary 2, we have that $K^{-1}(w,i)) < K^{-1}(w,j))$, since the firm offering the same wage w for a higher-skill worker must have a lower TFP than the firm offering the same wage for a lower-skilled worker. Consequently, $\Gamma(K^{-1}(w,i)) < \Gamma(K^{-1}(w,j))$. On the other side, since based on our assumptions we have $\frac{d\underline{x}(i)}{di} \geq 0$, 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

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

Manipulating it, we have

$$F_i(w) - F_j(w) = \frac{\left\{\begin{array}{c} \Gamma(K^{-1}(\underline{w}_j, j))[1-\Gamma(K^{-1}(w,i))] \\ -\Gamma(K^{-1}(\underline{w}_i, i))[1-\Gamma(K^{-1}(w,j))] \\ +\Gamma(K^{-1}(w,i)) - \Gamma(K^{-1}(w,j)) \end{array}\right\}}{[1-\Gamma(K^{-1}(\underline{w}_j, j))][1-\Gamma(K^{-1}(\underline{w}_i, i))]}$$

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

$$F_i(w) - F_j(w) = \frac{\left\{\begin{array}{c} [1-\Gamma(K^{-1}(\underline{w}_i, i))][1-\Gamma(K^{-1}(w,j))] \\ -[1-\Gamma(K^{-1}(\underline{w}_j, j))][1-\Gamma(K^{-1}(w,i))] \end{array}\right\}}{[1-\Gamma(K^{-1}(\underline{w}_j, j))][1-\Gamma(K^{-1}(\underline{w}_i, i))]} < 0$$

since $[1-\Gamma(K^{-1}(\underline{w}_i, i))] < [1-\Gamma(K^{-1}(\underline{w}_j, j))]$ and $[1-\Gamma(K^{-1}(w,j))] < [1-\Gamma(K^{-1}(w,i))]$. Consequently, $F_i(w) \leq F_j(w)$, $\forall w$. Therefore, $F_i(w)$ FOSD $F_j(w)$. ■

Proof of Corollary 4

Proof.

$$E_{F_i}[w] = \int_{\underline{x}(i)}^{\overline{x}} K(x', i)\frac{\gamma(x')}{1-\Gamma(\underline{x}(i))}dx' \geq \int_{\underline{x}(i)}^{\overline{x}} K(x', j)\frac{\gamma(x')}{1-\Gamma(\underline{x}(i))}dx' \geq \int_{\underline{x}(j)}^{\overline{x}} K(x', j)\frac{\gamma(x')}{1-\Gamma(\underline{x}(j))}dx' = E_{F_j}[w]$$

where the first inequality comes from the fact that $\frac{\partial K(x,i)}{\partial i} > 0$, while the second inequality comes from $F_i(w)$ FOSD $F_j(w)$. Now simplifying the above expression, we have $E_{F_i}[w] \geq E_{F_j}[w]$. ■

Proof of Proposition 3

Proof. From Lemmas 1 and 3, 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 are $[b(i), K(\overline{x}, i)]$ and $[b(j), K(\overline{x}, j)]$, where $b(i) \geq b(j)$ and $K(\overline{x}, i) > K(\overline{x}, j)$. 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 $w \in (K(\overline{x}, j), K(\overline{x}, i))$ we have that $G_j(w) = 1$ and $G_i(w) < 1$, so $G_i(w) \leq G_j(w)$. Finally, if $w \in [b(i), K(\overline{x}, j)]$, so the wage is earned to both type i and j workers. In this case, we have

$$G_i(w) - G_j(w) = \left\{\begin{array}{c} \frac{(d+\delta_i)F_i(w)}{\{d+\delta_i+\lambda(i)[1-\Gamma(K^{-1}(\underline{w}_i,i))](1-\delta_i)[1-F_i(w)]\}} \\ -\frac{(d+\delta_j)F_j(w)}{\{d+\delta_j+\lambda(j)[1-\Gamma(K^{-1}(\underline{w}_j,j))](1-\delta_j)[1-F_j(w)]\}} \end{array}\right\}$$

From our assumptions, we know that since $i > j$, we have $\delta_i < \delta_j$ and $\lambda(i) > \lambda(j)$. Moreover, from previous calculations, we have that

$$[1-\Gamma(K^{-1}(\underline{w}_i, i))][1-F_i(w)] = [1-\Gamma(K^{-1}(w,i))]$$

But then, based on Lemmas 1 and 3, we have $[1 - \Gamma(K^{-1}(w, i))] > [1 - \Gamma(K^{-1}(w, j))]$. Finally, from Proposition 2, we have $F_i(w) < F_j(w)$. Putting all these results together, we have

$$(d + \delta_i)F_i(w) < (d + \delta_j)F_j(w)$$

and

$$\{d+\delta_i+\lambda(i)[1-\Gamma(K^{-1}(\underline{w}_i, i))](1-\delta_i)[1-F_i(w)]\} > \{d+\delta_j+\lambda(j)[1-\Gamma(K^{-1}(\underline{w}_j, j))](1-\delta_j)[1-F_j(w)]\}$$

and, consequently, $G_i(w) - G_j(w) < 0$. In summary, $G_i(w) \leq G_j(w)$, $\forall w$. Therefore, $G_i(w)$ FOSD $G_j(w)$. ■

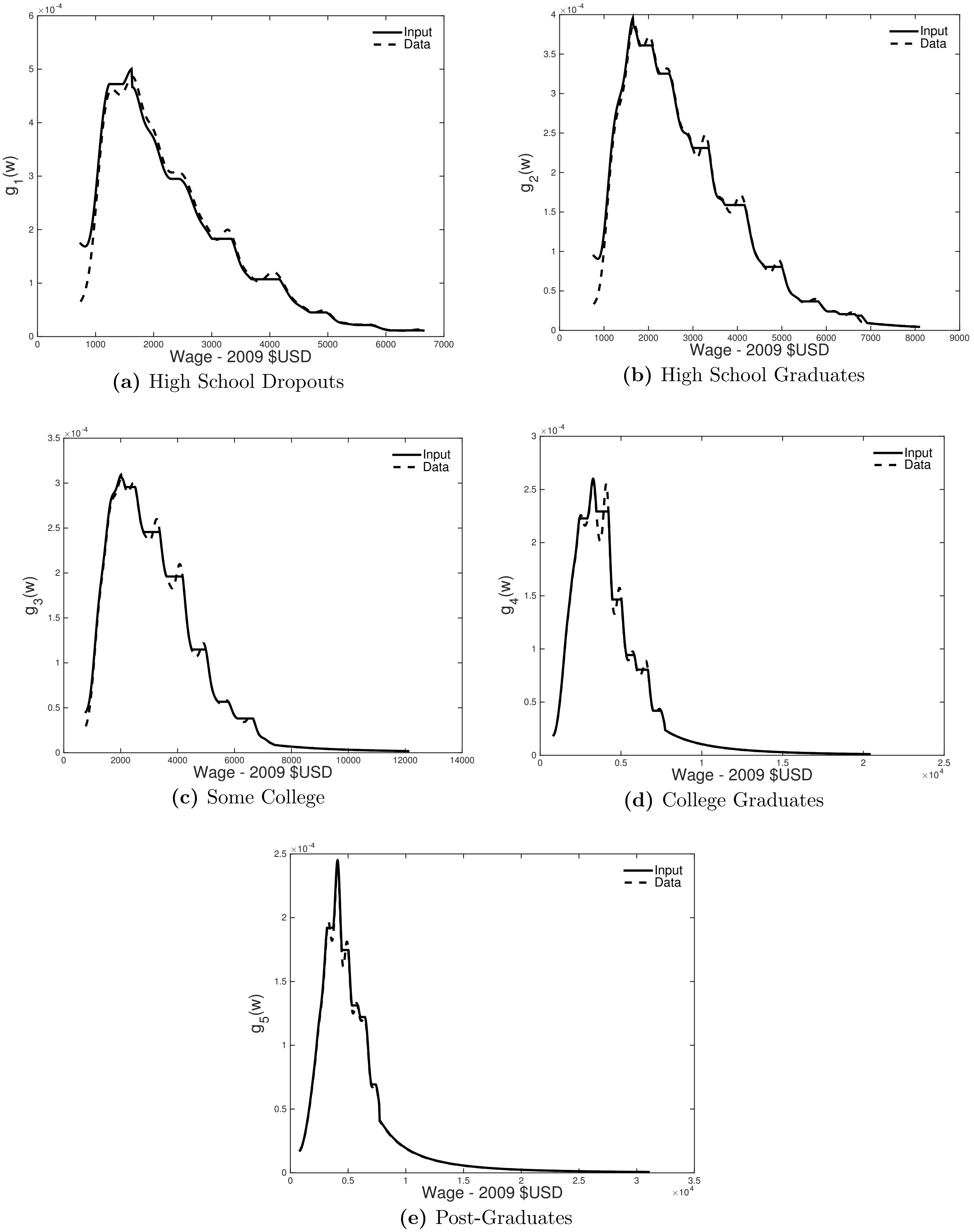
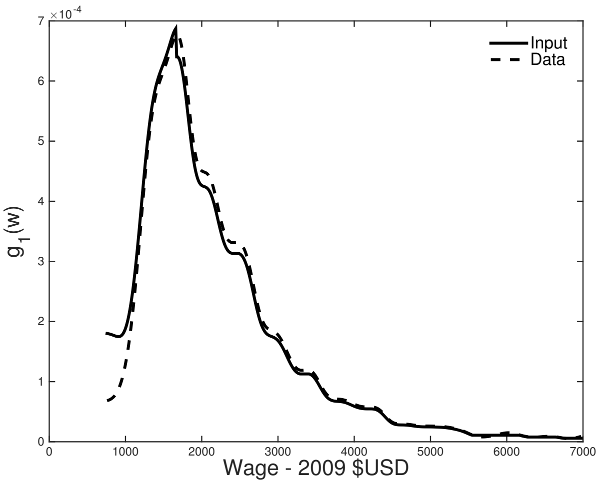
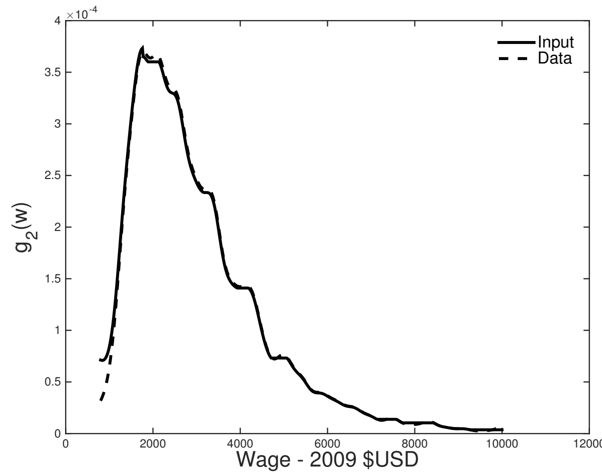


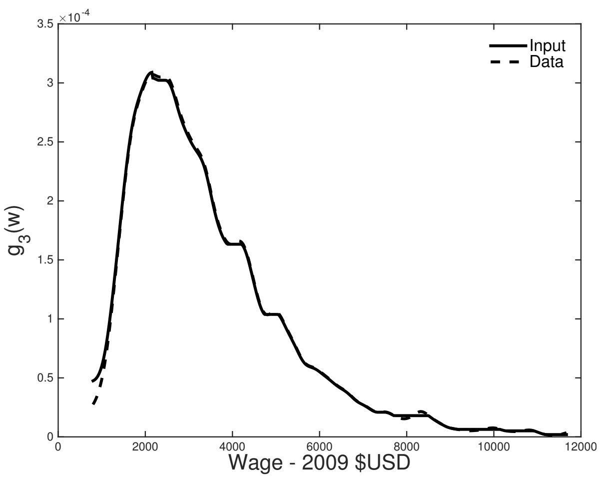
Figure 1: Comparison between Empirical Wage Distributions and Caliobrated Distributions that Fulfill Model Requirements – 1985



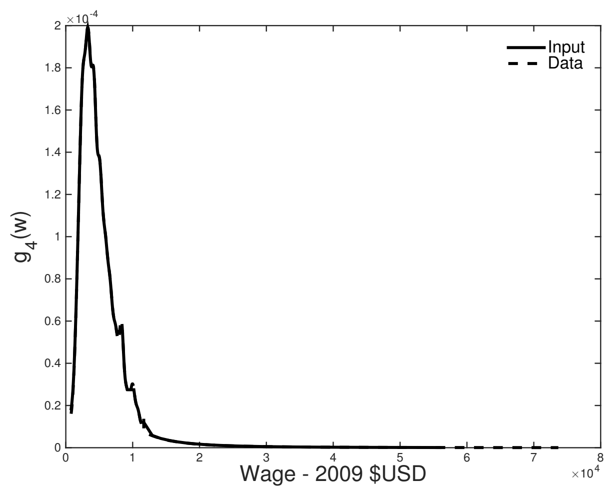
(a) High School Dropouts



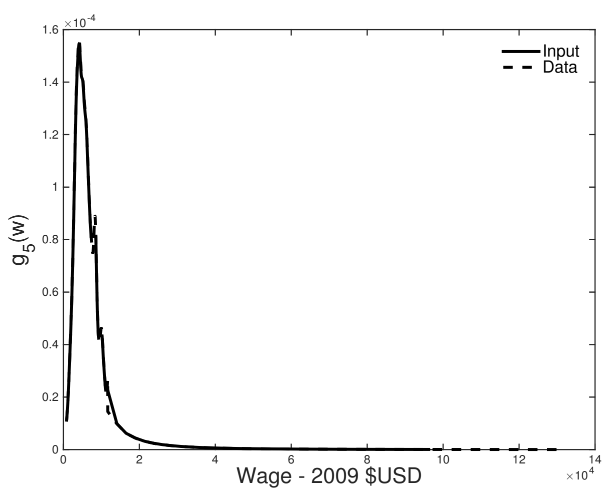
(b) High School Graduates



(c) Some College

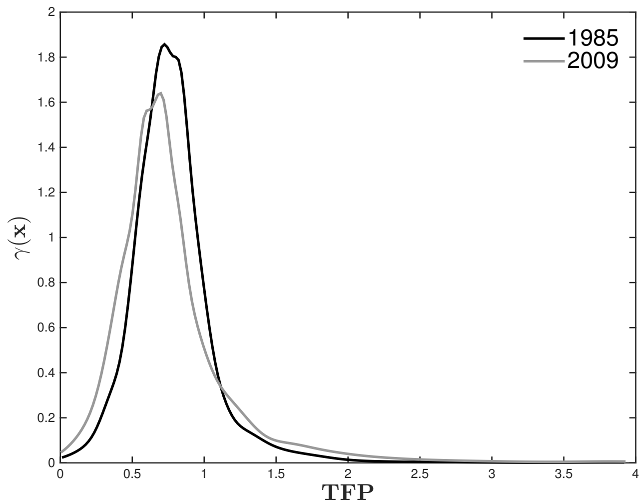


(d) College Graduates

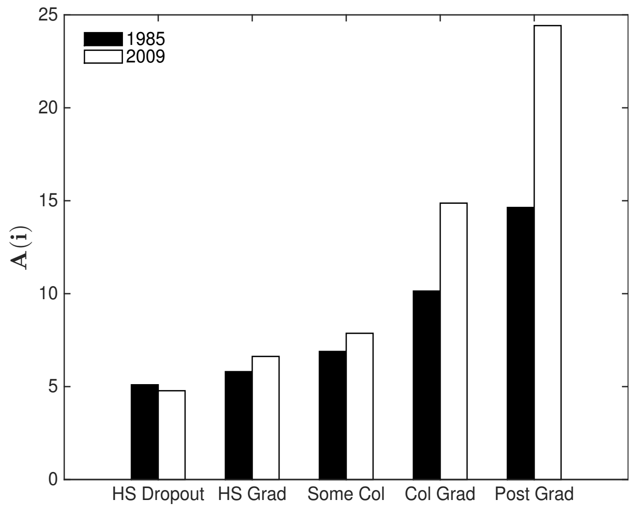


(e) Post-Graduates

Figure 2: Comparison between Empirical Wage Distributions and Calibrated Distributions that Fulfill Model Requirements – 2009



(a) TFP Distributions Based on COMPUSTAT Data



(b) Labor Productivity by Educational Group

Figure 3: Factor Calibrated Productivities: 1985 vs. 2009

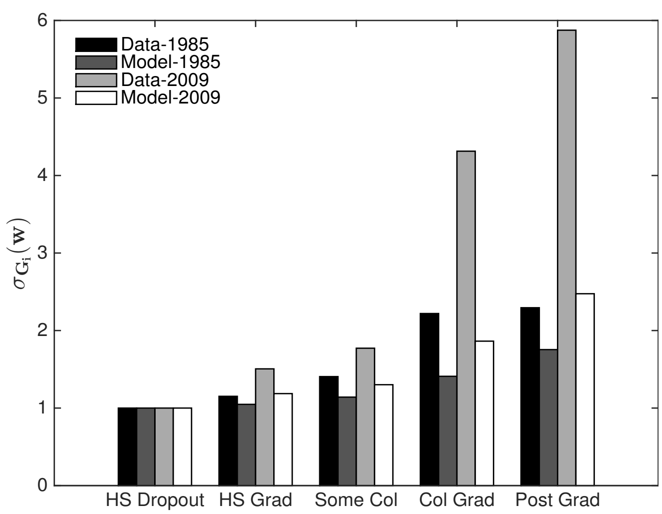
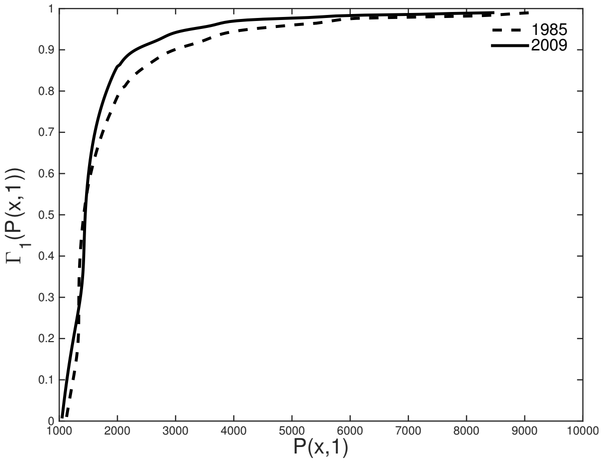
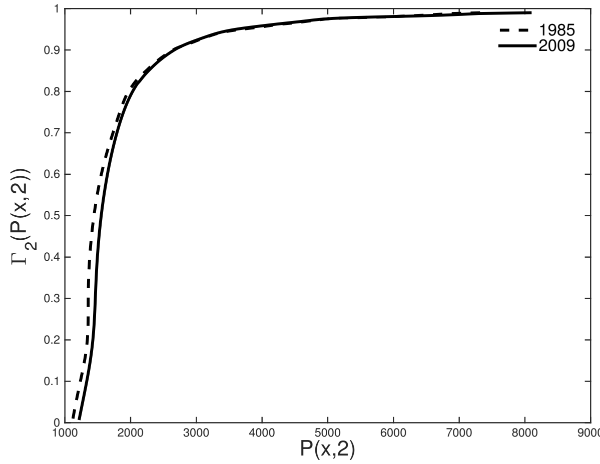


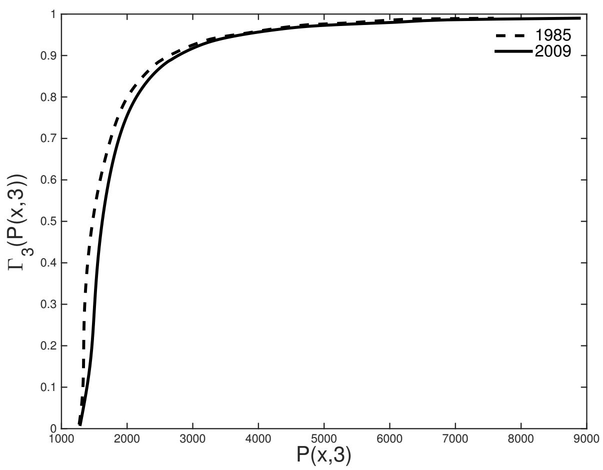
Figure 4: Model Fit: Standard Deviation of Earned Wages



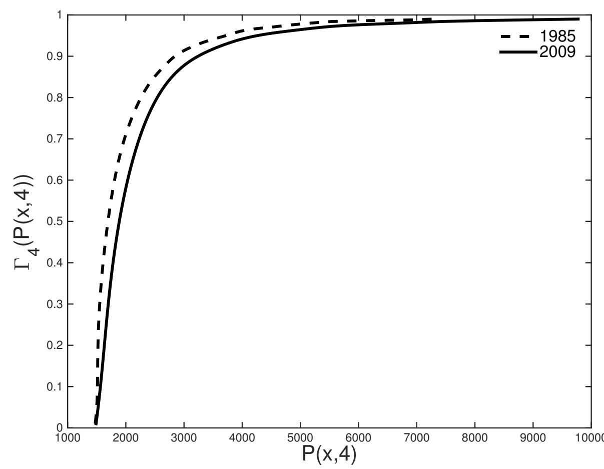
(a) High School Dropouts



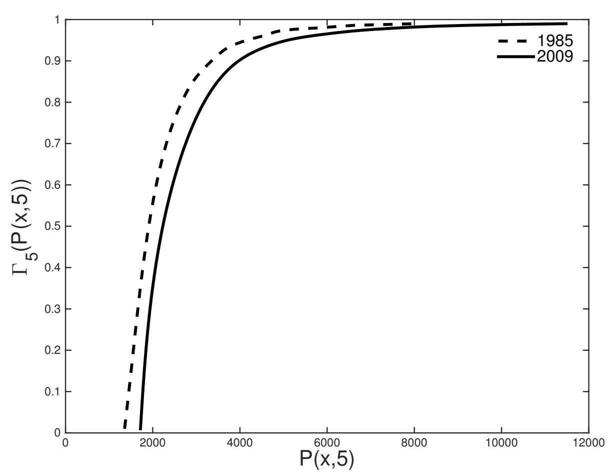
(b) High School Graduates



(c) Some College

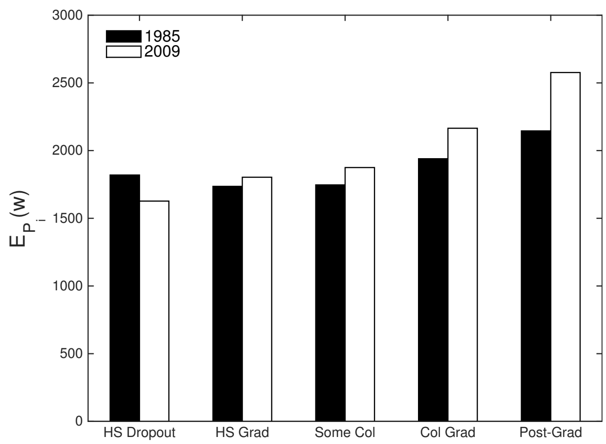


(d) College Graduates

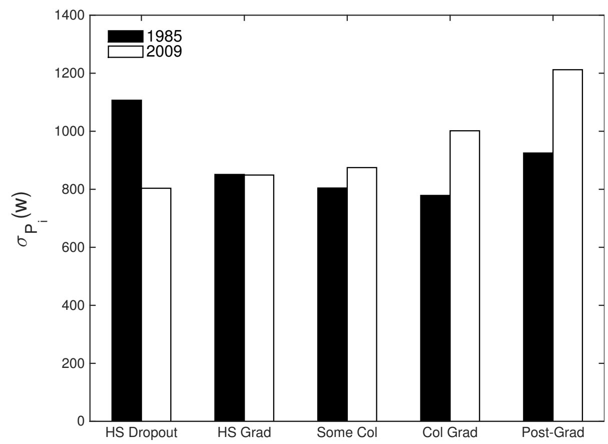


(e) Post-Graduates

Figure 5: Comparison between Cumulative Distributions of Outputs for Different Skill Levels – 1985 vs. 2009

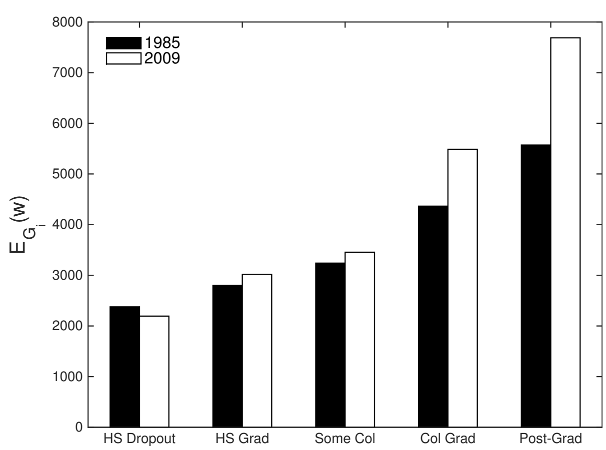


(a) Average Output as a Function of Skill Level

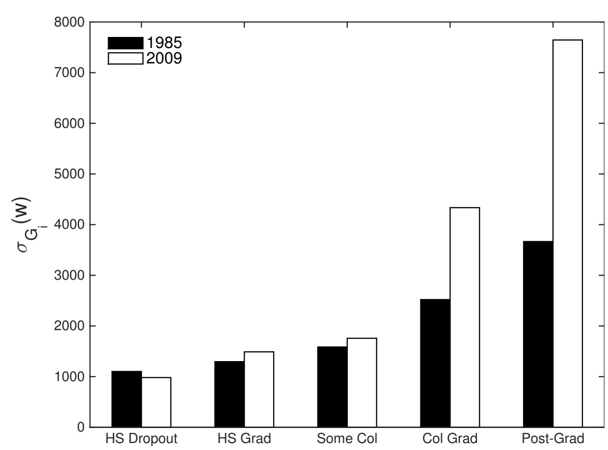


(b) Standard Deviation of Output as a Function of Skill Level

Figure 6: Average and Standard Deviation of Output for Different Skill Levels – 1985 vs. 2009

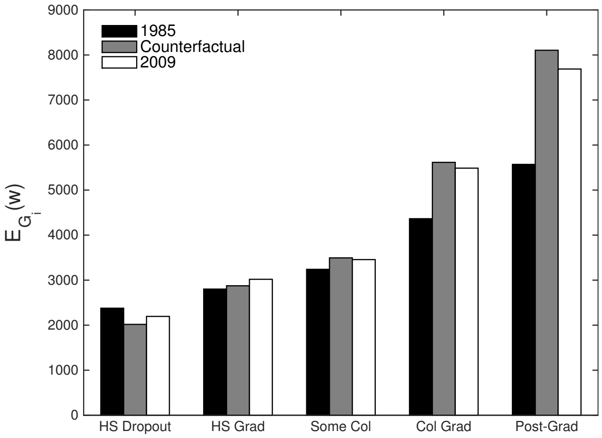


(a) Average Wages as a Function of Skill Level

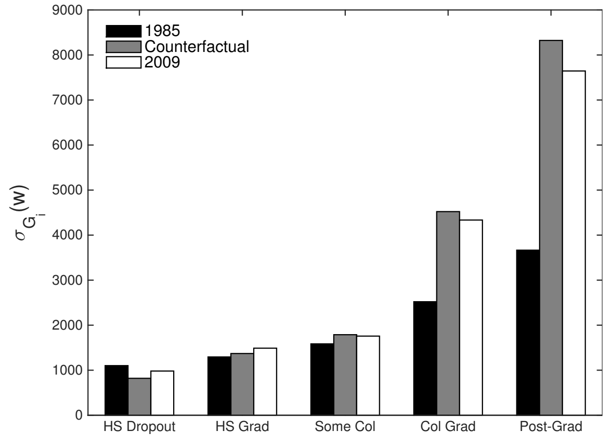


(b) Standard Deviation of Wages as a Function of Skill Level

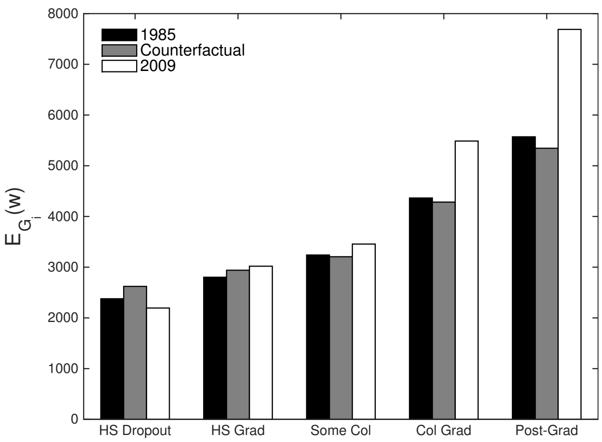
Figure 7: Average and Standard Deviation of Wages for Different Skill Levels – 1985 vs. 2009



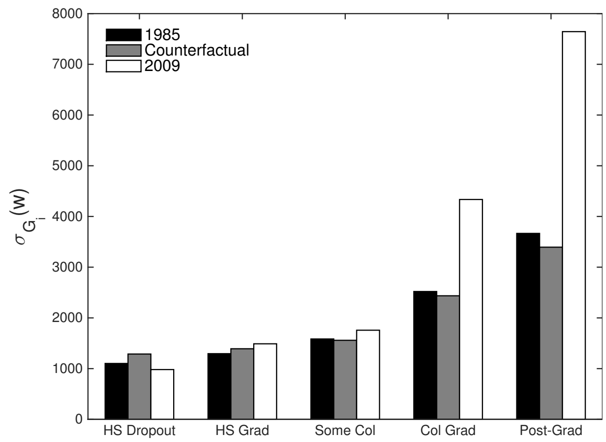
(a) Average Wage - Output Counterfactual



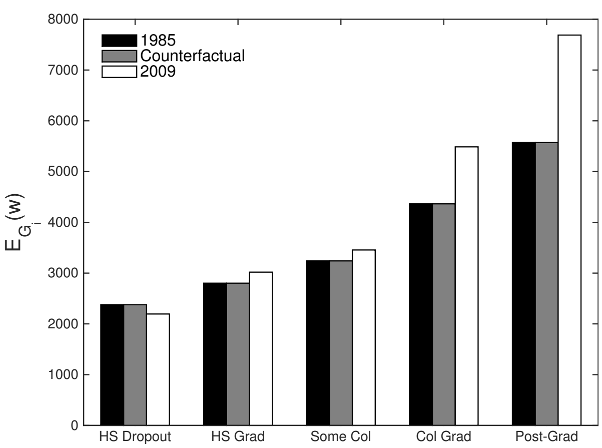
(b) Standard Deviation of Wages - Output Counterfactual



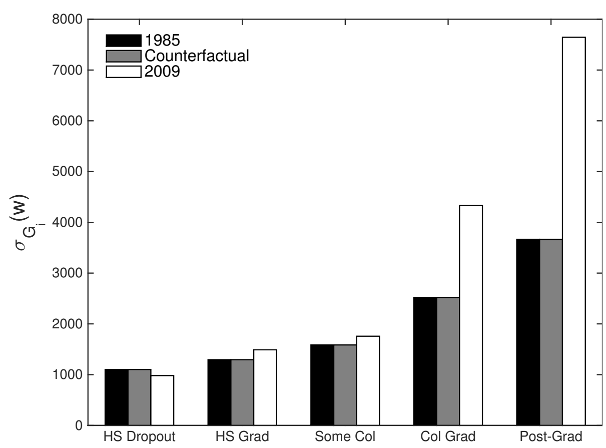
(c) Average Wage - Labor Frictions Counterfactual



(d) Standard Deviation of Wages - Labor Frictions Counterfactual

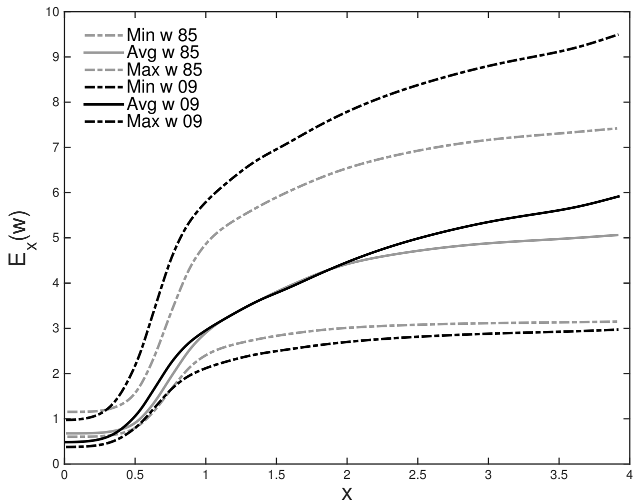


(e) Average Wage - Education Dist. Counterfactual

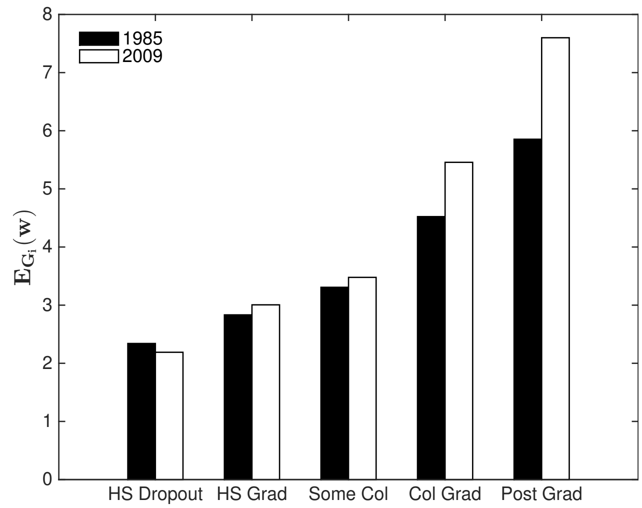


(f) Standard Deviation of Wages - Education Dist. Counterfactual

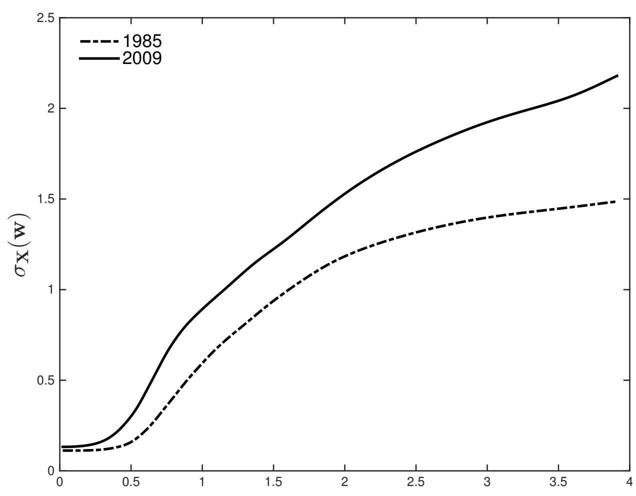
Figure 8: Average and Standard Deviation of Wages - Counterfactuals



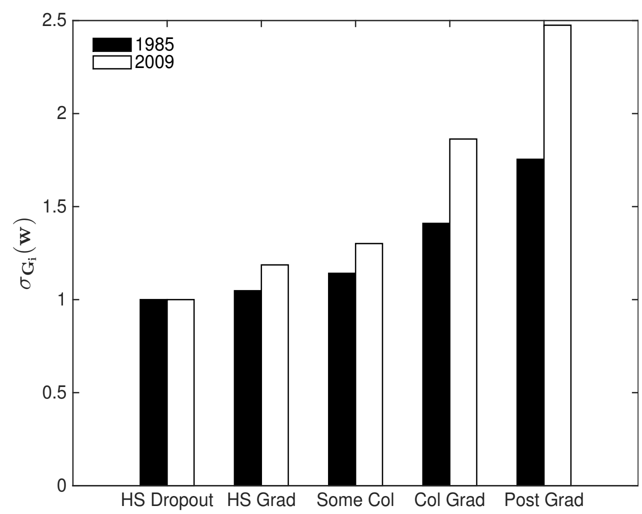
(a) Average Wages as a Function of Firm's TFP



(b) Average Wages as a Function of Worker's Educational Attainment

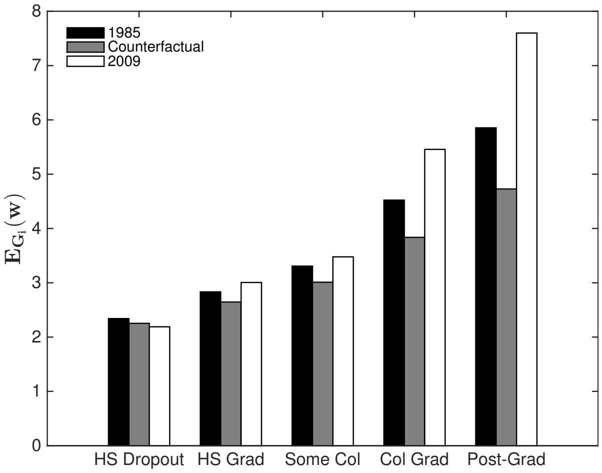


(c) Standard Deviation of Earned Wages as a Function of Firm's TFP

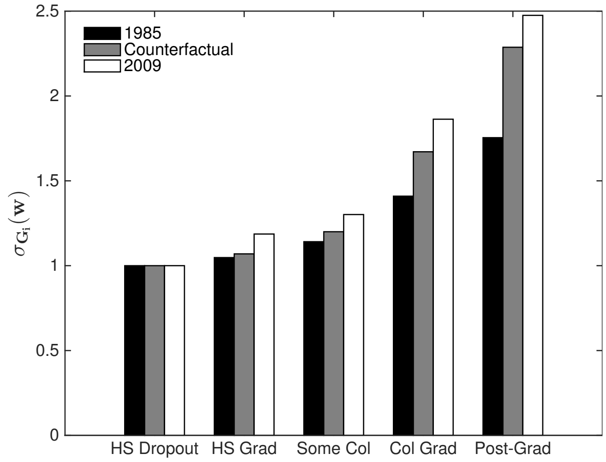


(d) Standard Deviation of Earned Wages as a Function of Educational Attainment

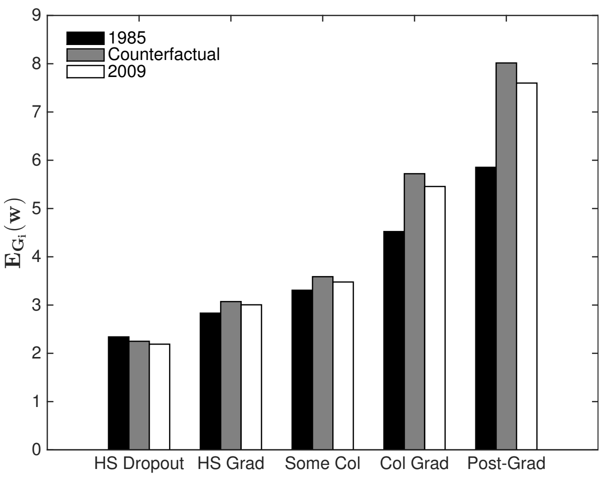
Figure 9: Average and Standard Deviation of Wages as a Function of Educational Attainment and TFP: 1985 vs. 2009



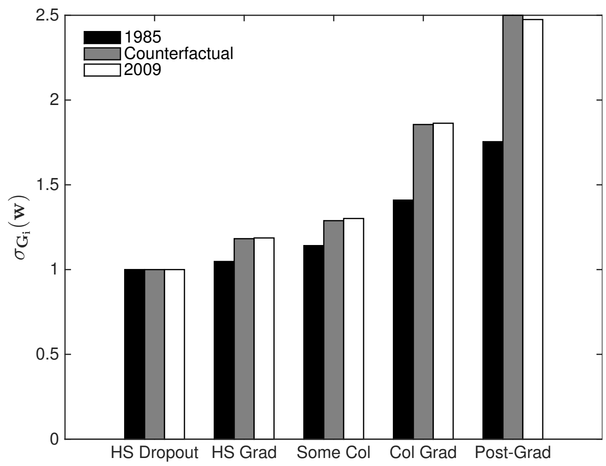
(a) Average Wage - TFP Counterfactual



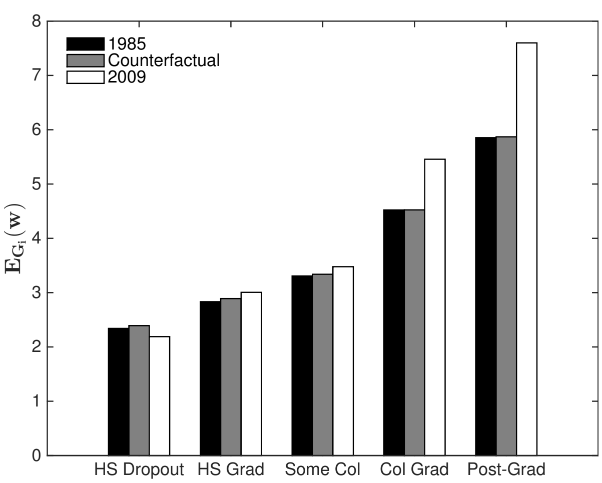
(b) Standard Deviation of Wages - TFP Counterfactual



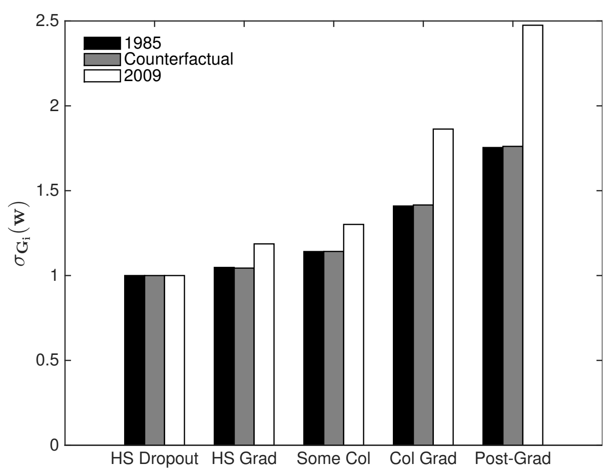
(c) Average Wage - $A(i)$ Counterfactual



(d) Standard Deviation of Wages - $A(i)$ Counterfactual

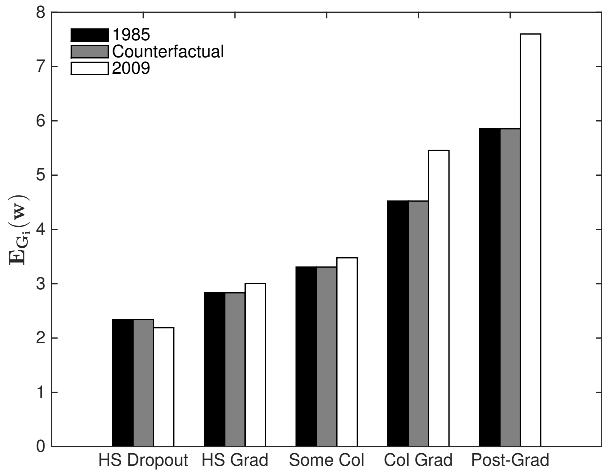


(e) Average Wage - Labor Frictions Counterfactual

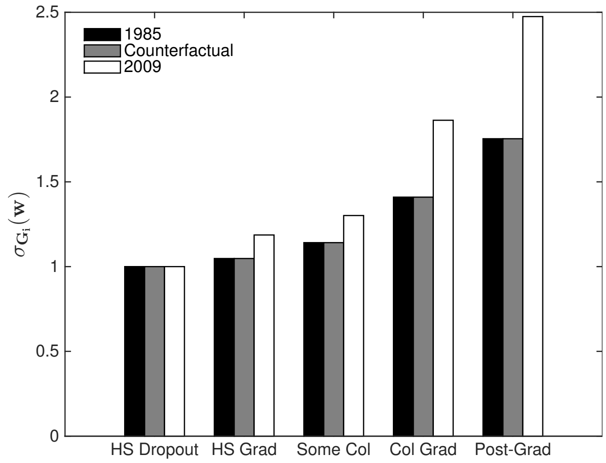


(f) Standard Deviation of Wages - Labor Frictions Counterfactual

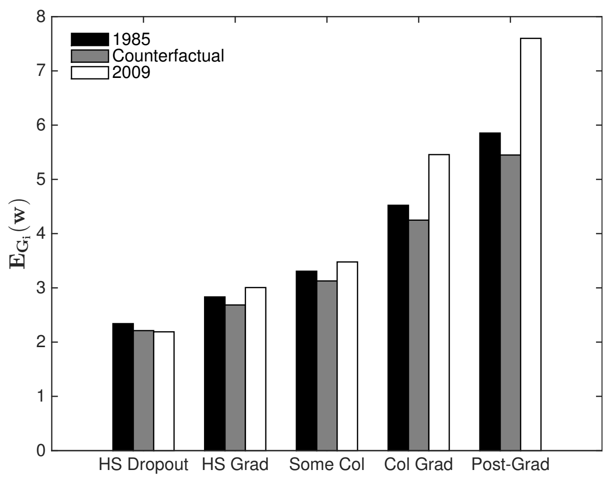
Figure 10: Average and Standard Deviation of Wages by Skill Levels – Counterfactuals



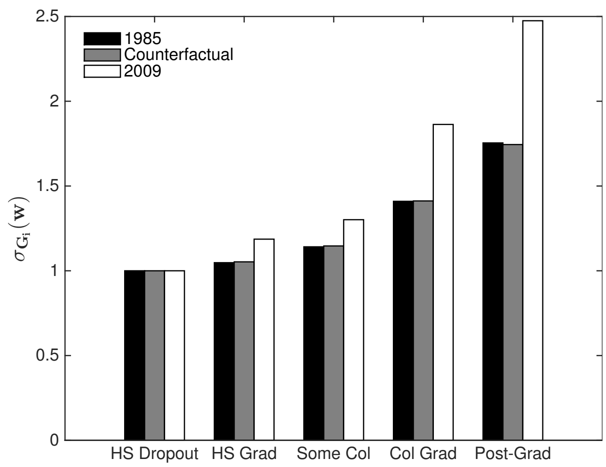
(g) Average Wage - Education Dist. Counterfactual



(h) Standard Deviation of Wages - Education Dist. Counterfactual

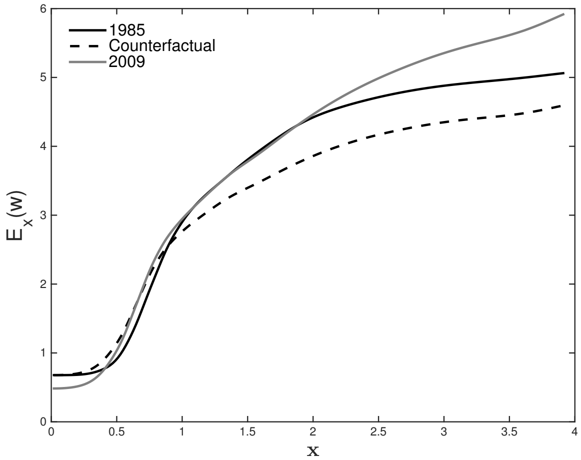


(i) Average Wage - Production Fcn. Counterfactual

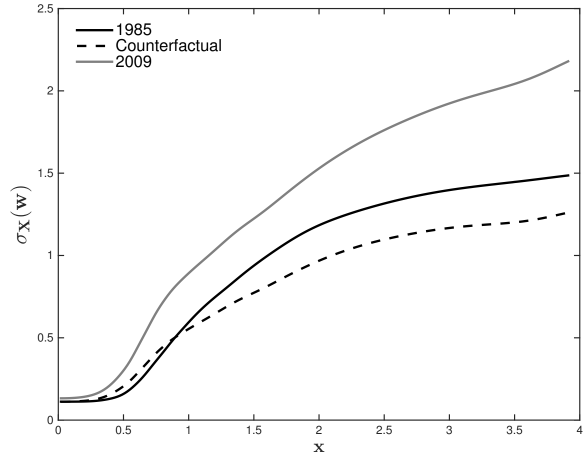


(j) Standard Deviation of Wages - Production Fcn. Counterfactual

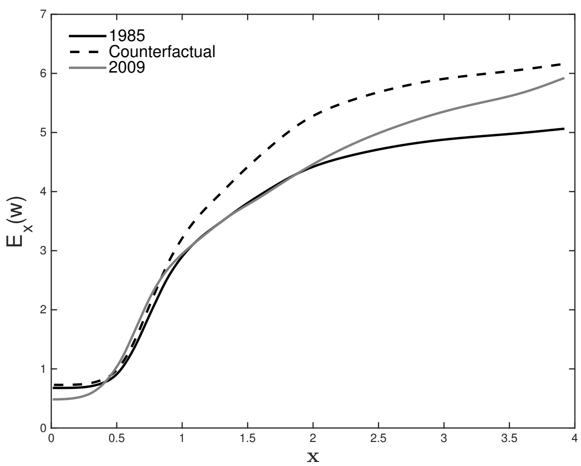
Figure 10: Average and Standard Deviation of Wages by Skill Levels – Counterfactuals



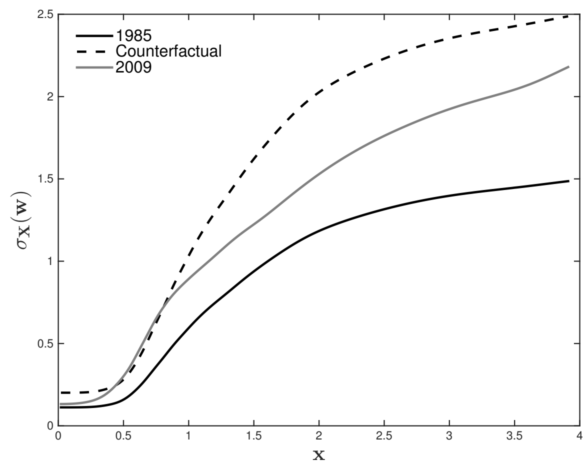
(a) Average Wage - TFP Counterfactual



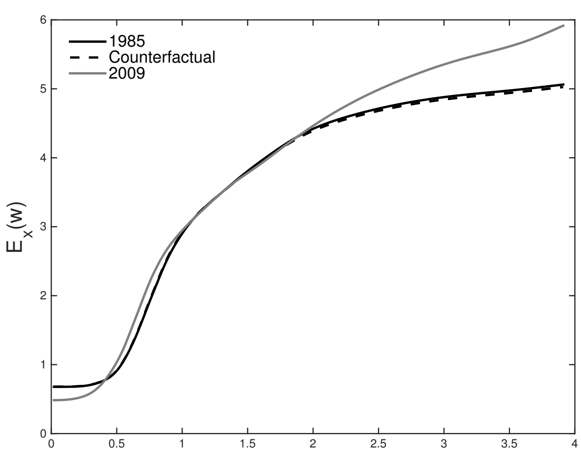
(b) Standard Deviation of Wages - TFP Counterfactual



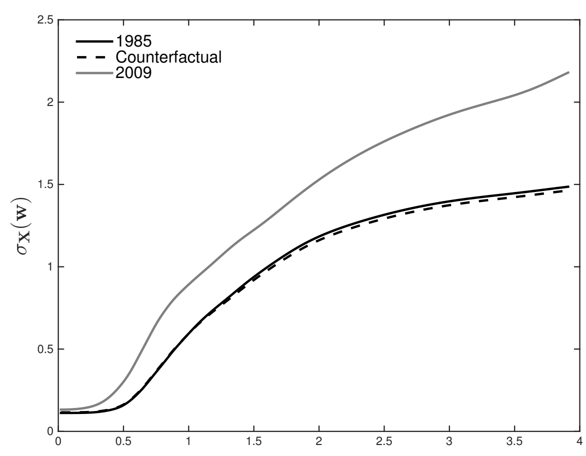
(c) Average Wage - $A(i)$ Counterfactual



(d) Standard Deviation of Wages - $A(i)$ Counterfactual

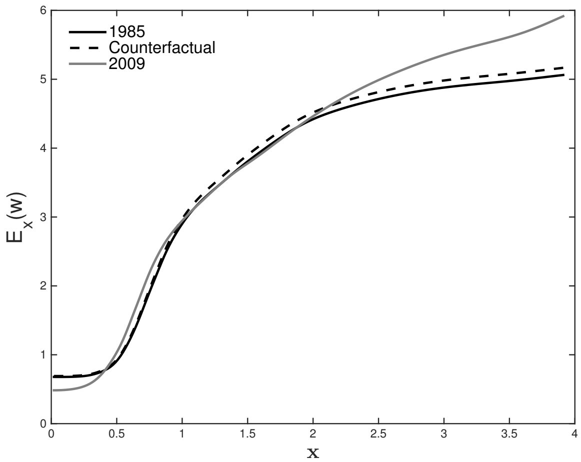


(e) Average Wage - Labor Frictions Counterfactual

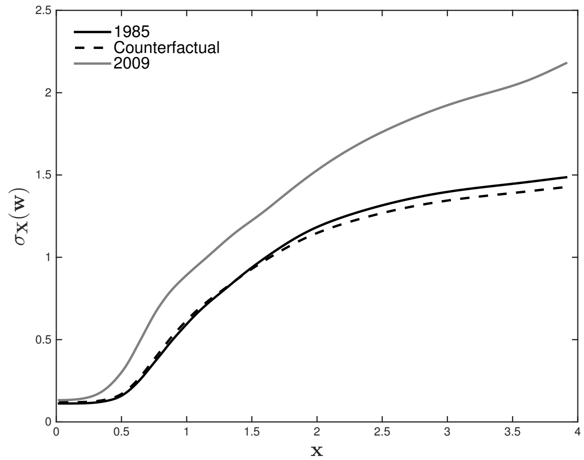


(f) Standard Deviation of Wages - Labor Frictions Counterfactual

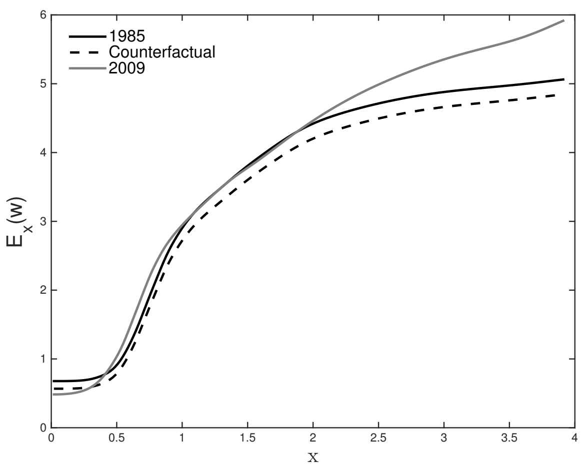
Figure 11: Average and Standard Deviation of Wages by TFP – Counterfactuals



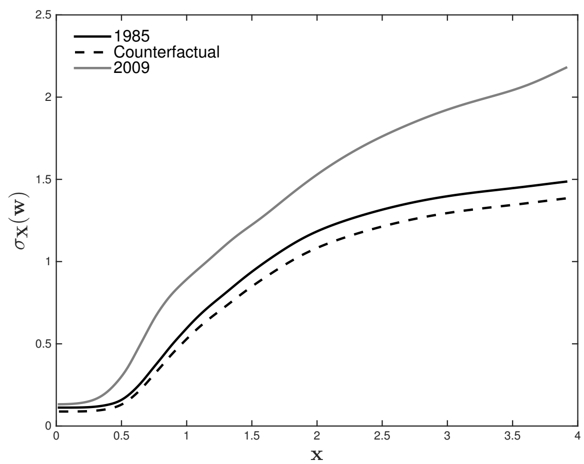
(g) Average Wage - Education Dist. Counterfactual



(h) Standard Deviation of Wages - Education Dist. Counterfactual



(i) Average Wage - Production Fcn. counterfactual



(j) Standard Deviation of Wages - Production Fcn. counterfactual

Figure 11: Average and Standard Deviation of Wages by TFP – Counterfactuals