It's Not Who You Are, It's What They Know: Wage Gaps and Informational Frictions

Nadav Kunievsky*

January 27, 2024

Abstract

Can informational asymmetries among firms account for all observed wage gaps across social groups? We confirm this through a parsimonious common-value auction model in the labor market with unspecified information structures. Firms with identical characteristics encounter workers with unobserved productivity and extend wage offers based on their information about worker productivity and competing offers. Using 2010 American Community Survey data, we show that wage disparities among both black and white men and women can be explained using a common productivity distribution for all social groups and differences in what firms know, if the mean of this common productivity distribution ranges between \$48,000 and \$132,800. Our results emphasize the importance of understanding what firms know in shaping wage distributions and explaining wage disparities

*This paper benefited from discussions with Stephane Bonhomme, Ben Brooks, Hazen Eckehert, Michael Galperin, Natalie Goldshtein, Francesco Ruggieri, Azeem Shaikh, and Alex Torgovitsky.

Firms can rarely, if ever, hire a worker after obtaining complete information on the worker's productivity and their outside options. These informational frictions can arise from various sources, including an inefficient hiring process, imperfect information on the firm's own production technology, attention costs of the interviewer, or cognitive biases on their part. These frictions have proven to be of great importance in countless theoretical results (e.g., Aigner and Cain (1977), Phelps (1972), Spence (1978), Bergemann et al. (2021b), Bergemann and Morris (2019a)). However, while we would like to take these into account both in modeling firms' decisions and in empirical exercises, this has proven to be fairly difficult. There are many informational environments in which firms operate that are not observable by researchers. Despite the crucial place information holds in theoretical research, most of the empirical literature on wage gaps has focused on other differences between groups. These differences are brought on by the structure of the labor market. The first type of fundamental differences is driven by the workers' productivity distribution. As many papers argue (for example, Altonji and Blank (1999), Blau and Kahn. (2017), Goldin (2014)), differences in workers' abilities between groups can drive differences in observed wages. These differences can stem from various sources, such as pre-market conditions, which generate differences in workers' productivity. Other sources that have been considered extensively include firms' taste-based discrimination, which can affect firms' willingness to pay for workers of different groups, or self-selection of workers into different occupations and industries due to differences in preferences, to name a few. The second type of mechanism that has been thoroughly explored in the literature and can be used as a cause for differences in wages is market frictions, such as search costs, probability of finding a job, differences in bargaining power, and differences in outside options.

In contrast to these explanations, this paper explores whether differences in wage distributions can stem from differences in firms' information about workers' productivity and their potential outside options. These differences can potentially be significant, affecting the wage distributions of different groups in various way, creating different in wage prospect for workers who are productively equivalent.

To gauge the potential importance of information frictions, we construct a static, parsimonious, common-value auction model of the labor market. This model assumes that heterogeneous workers receive job offers from identical firms that differ only in the information available to them. We then explore how this varies across markets with different levels of search frictions, captured by the number of wage offers a single worker receives. Since we are interested in examining the impact of information on the labor market, we leave unspecified the information firms have. We then ask how much of the wage gap between workers can be explained by the correlation between gender and race and the other information firms observe before making a wage offer. To form this test, we leverage an equivalence result from the robust prediction literature and information design (Bergemann and Morris (2013), Bergemann and Morris (2016)). This result shows that the set of distribution outcomes that can arise under a Bayes Nash Equilibrium (BNE) with some information structure corresponds to a set of joint distributions of actions and states, known as Bayes Correlated Equilibrium (BCE). We use this equivalence result to partially identify the set of possible productivity distributions which, with some signal structure possibly correlated with race and gender, can give rise to the observed wage gaps.

We find that information can potentially have a very large effect on the wage distribution, creating a significant divergence between workers' marginal product and their wage. For example, without any assumptions on the information available to firms, and in markets with relatively low search frictions, we can bound the mean productivity of white male workers to be between approximately $48,000 per year, which is roughly their mean wage, and $128,500. Moreover, we find that we can explain all of the wage gaps between white men, white women, and black men and women without needing to assume differences in productivity. More specifically, we find that the entire wage gap in our sample can be attributed to information frictions and can be supported by a productivity distribution with a mean bounded between approximately $48,000 and $132,800 per year in an economy with large search frictions, and between $48,000 and $93,600 in an economy with low search frictions.

This paper contributes to the vast literature on discrimination and, specifically, on statistical discrimination (Arrow (1973), Phelps (1972), Aigner and Cain (1977), Altonji and Pierret (2001), Lange (2007)). These early papers show that different information can give rise to differences in wage distributions but, as we argue in section 1, fail to explain differences in average wage. To address this issue, follow-up papers by Lundberg and Startz (1983) and Coate and Loury (1993) offer models in which minority workers end up investing less in human capital, generating equilibrium differences in workers' productivity available to firms. Unlike previous papers that attempt to explain wage gaps using statistical discrimination, we ask whether gaps can be explained without needing to change the underlying distribution of workers' productivity, but by relaxing the assumptions on the types of information firms have and allowing for firms to act based on private information. While our model does not exclude taste-based discrimination, it assumes that it's another force that affects the productivity distribution of workers as seen from the firms' perspective. A recent paper, by Chambers and Echenique (2021), also spotlight information frictions. They explore whether wage gaps and discriminatory policies can potentially arise from differences in information in an environment where the same worker can generate varying productivity for different firms possessing distinct information. Theoretically, they find that such an environment invariably supports an information structure capable of creating wage differences. In contrast to their work, our paper focuses on an empirical exercise involving homogeneous firms with different information on workers, all of whom can generate the same surplus for these firms. Our findings underscore that a key driver of divergent wage policies is the may be information itself, rather than the underlying productivity distribution, which remains constant in our exercise.

As discussed above, this paper builds on recent results from the robust prediction literature (Bergemann and Morris (2013), Bergemann and Morris (2016), Bergemann et al. (2017)) that explore the range of outcomes that can arise in a game with an unspecified information structure. These results are being used for informationally robust identification in a growing number of papers. Syrgkanis et al. (2021) is the most similar paper to ours; it explores how to achieve identification in a model of general second and first-price auctions without specifying the information available to individuals. They then use their identification results to analyze second-price auctions in the BingeAds sponsored auction marketplace and the OCS auction dataset to infer the underlying valuation distributions. Magnolfi and Roncoroni (2017) uses the BCE in an entry game with binary actions to identify the set of parameters on the utility function that are robust to the information firms have. Gualdani and Sinha (2019) employs the BCE framework we work with in this paper to identify the set of parameters and distributions governing an agent in a discrete choice model without specifying the information structure. Finally, Bergemann et al. (2021a) consider how to perform counterfactual analysis while holding the information fixed.

Additionally, this paper contributes to the recent empirical literature that emphasizes the role of workers' outside options in wage gaps. Caldwell and Danieli (2021) uses a two-sided matching model with transfers, based on Shapley and Shubik (1971), where heterogeneous workers and firms have idiosyncratic preferences for each other. They then calculate their outside option index for workers in Germany and find that it can explain roughly 25% of the wage gap. Black (1995) constructs a search model where some discriminatory employers reduce workers' outside options, thereby generating a wage decrease. In our model, the distribution of outside options is not generated by assuming workers have different preferences or due to monopolistic power, but because of the information firms have on workers and other firms' willingness to pay. Compared to previous papers, we allow for uncertainty over workers' outside options.

The paper proceeds as follows: Section 2 introduces the model. Section 3 discusses identification and how to test for the potential role of information in shaping the wage gap. Section 4 focuses on inference and computation, Section 5 presents our data and results, and Section 6 concludes.

1 A Simple Common-Value Auction Model of the Labour Market

1.1 The model

Let $\mathcal{J}$ be the set of firms in the market. Let $\mathcal{I}$ be the set of workers. There are $|\mathcal{G}|$ groups of workers. Workers have heterogeneous productivity, $v \in \mathcal{V} \subset \mathbb{R}_+$, drawn from a distribution $\mu(v|g) \in \Delta(\mathcal{V})$. A job offer from firm j to worker i consists of a wage $w_j^i \in W$. We assume that workers receive $N \leq |\mathcal{J}|$ jobs offers. We denote as $\boldsymbol{w}_i \in \mathcal{W} = W^N$ the vector of wage offers worker i receives. We further assume that both firms and workers are risk neutral and that the firms' production function is additive in the number workers. Therefore, if a firm succeeds in hiring a worker, that firm's marginal profit is given by $v - w$. We assume firms don't have a cost of making a wage offer. Worker i's utility from a vector of wage offers $\boldsymbol{w}_i$ is $u(\boldsymbol{w}) = \max_j \boldsymbol{w_i}$. Implying workers choose to worker at the firm who offers the highest wage. In the case of a tie, the worker selects at randiom one of the firms who offers the highest wage.

Before extending a wage offer, we assume that all firms observe both the worker's group, $g_i \in \mathcal{G}$, and a public signal, $x_i \in \mathcal{X}$, observed by all firms and by the econometrician. We do not restrict the correlation between the public signal and the workers' productivity. In addition to these signal, firms may observe additional signal, possibly private, $t_j \in \mathcal{T}_j$, prior to making a wage offer to the worker. The signal vector $t = (t_1, ..., t_J)$ can be arbitrarily correlated with both the worker's productivity and the public signals. We also do not put any restriction on the correlations between the different firms' signals. We denote the augmented signal structure $(\mathcal{G}, \mathcal{X}, \mathcal{T}, \mathrm{P}(g, x, t|v))$ by $S \in \mathcal{S}$. Let $k_i(\boldsymbol{w}) = \arg\max_j \boldsymbol{w}$ be the set of firms that offer the highest wage to worker i, then the worker is allocated to firm j with probability

$$q_j^i(\boldsymbol{w}) = \begin{cases} \frac{1}{|k(\boldsymbol{w})|} & \text{if j} \in k(\boldsymbol{w}) \\ 0 & otherwise \end{cases}$$

Finally, firms' j interim-expected marginal profit from offering a wage w_j to worker i, after observing the worker's public signals x_i, g_i and the private signal t_j is

$$\mathrm{E}\Bigg[(v_i - w_j)q(\boldsymbol{w})|t_j, x_i, g_i\Bigg] \propto$$
$$\sum_v \sum_{t_{-j}} \sum_{\boldsymbol{w}_{-j}} (v_i - w_j) q_j^i(\boldsymbol{w}) \Bigg[\prod_{k \neq j} \beta_k(w_k|t_k, x_i, g_i)\Bigg] p(t|v_i, x_i, g_i)\mu(v_i|x_i, g_i)p(g_i, x_i)$$

where $\beta_k(w|.)$ is the wage policy functions of firm k, given the firm's signals. A Bayes Nash Equilibrium (BNE) in this model is a mapping $\beta_k : \mathcal{S} \to \Delta(W_k)$ for each firm j, such that the firm maximizes its expected profit, conditional on their signals, and that workers choose to work at the firm that offered the highest wage.

In the model above, heterogeneity in wage offers stems from firms having access to different information. Specifically, information in the model plays two key roles in determining wage. The first is by affecting the firm's evaluation of worker productivity. Firms having different information structures implies that different firms evaluate the worker productivity differently, affecting their willingness to pay. The second channel through which information affects wages is firms' belief on the worker's outside option. Within the model, firms are asking themselves what other firms know about the worker and try to guess what other firms would be willing to pay for the worker.

To see the importance of these two channels, consider a simple setup with two firms, in which worker productivity is distributed uniformly between 0 and 1. Assume that that the two firms' signals are perfectly correlated. In that case, each firm knows that the other firm observes the same signal, then they would end up conducting a Bertrand competition, where wages would be the expected worker's productivity, given the common signal the average wage would be the worker's average productivity, as discussed in section 1.2. On the other hand, consider the polar opposite case, in which we have two firms, one is uninformed while the other one is perfectly informed.[1] An equilibrium in this setup would be that the informed firm would offer a wage of $\frac{v}{2}$, while the uninformed firm would randomize between 0 and 0.5 and the average wage would be $\frac{1}{3}$. To see this, notice that the uninformed firm would never make an offer higher than $\frac{1}{2}$, as it has negative ex-ante surplus. Next, notice that for any wage offer $w_{UI} \in [0, 0.5]$ the uninformed firm makes, the workers productivity, conditioned on the uniformed making the higher offer, is distributed uniformly between $[0, 2w_{UI}]$, and therefore, the expected surplus of the uninformed firm is $E[v - w_{UI}|\text{UI wins}] = 0$ for any $w_{UI} \in [0, 0.5]$. Finally, the informed firm surplus from offering wage w_I is given by $(v - w_I)P(w_I > w_{UI}) = (v - w_I)\frac{w_I}{0.5}$. which is maximized at $w_I = \frac{v}{2}$. Given these two equilibrium strategies, the

[1] This example is taken from Milgrom and Weber (1982)

worker's average wage would be $\frac{1}{3}$,[2] which is lower than the average wage under first case, or under complete information. Therefore, we can see that simply changing the firms' access to information may have a large effect on the realized wage distribution and on the relation between workers productivity and their wages.

1.2 Relation To Phelps (1972)

Before moving forward and considering a whether a general information structure might be needed to explain wage gaps, we can first ask whether there exists a public signal, available to all firms, which can induce the observed wage distributions of workers from two groups. In his seminal paper on statistical discrimination, Phelps (1972) considers a model similar to ours, but restricts attention to public and normal signals. In his model, there exist incomplete information on the worker productivity, and all firms observe the same public signal. Therefore, due to Bertrand competition, wages are set by the expected productivity of workers. Specifically, let v, the productivity of workers from group g, be distributed normally with mean α_g and variance $\sigma_{v,g}$. Firms cannot observe the worker productivity, but they have access to a public noisy signal

$$y = v + u$$

where the noise distributed normally $u \sim \mathcal{N}(0, \sigma_{u,g})$. Given the signal, the expected value of a worker's productivity is given by

$$E[v|y] = (1 - \gamma)\alpha_g + \gamma y$$

[2]To see this, notice that both firms make a wage offer uniformly on $[0, 0.5]$, and the winning wage offer is distributed with the CDF $F(x) = (\frac{w}{0.5})^2$ and the PDF $f(w) = 8w$. Therefore, the observed average workers wage is

$$\int_0^{0.5} w \times 8w dw = \frac{1}{3}$$

where $\gamma = \frac{\sigma_{v,g}}{\sigma_{v,g}+\sigma_{u,g}}$. As discussed in Phelps (1972) and Aigner and Cain (1977), the resulting wage distribution for two groups of workers would be different if either the noisy signal or the underlying productivity are distributed differently across different groups. Specifically, we can see that as employers get a more precise signal, they will put higher weight on the signal in determining the wage, and rely less on the group mean. As Aigner and Cain (1977) note, with risk-neutral firms, the Phelps model implies that differences in average wages can only be explained by differences in workers' average productivity, which implies that in this statistical discrimination model, information is not enough to induce the observed wage gaps between groups and we need to assume that there exist differences in the underlying productivity distribution to rationalize the observed wage gaps.

As it turns out, this observation is more general than in the case of the normal distribution. Under the assumption that the market is competitive, and that firms are risk neutral, the differences in mean wages must be driven by differences in the underlying distribution, and cannot be explained by public signals, as shown in the claim below

Claim 1. Let g_1 and g_2 be two groups of workers. Assume that firms are risk neutral and observe the worker's group membership and a public signal $t_{g_i} \in \mathcal{T}_{g_i}$, drawn from a conditional distribution $\pi(t|v, g_i) \in \Delta(\mathcal{T})$. Assume that the observed mean wages of workers from group 1 and 2 are different, $\bar{w}_{g_1} \neq \bar{w}_{g_1}$, then it must be the case that $E[v|g_1] \neq E[v|g_2]$

Proof. First, notice that as all firms observe the same signal and are competing for the same worker, they are engaging in a Bertrand competition. As firms are risk neutral, this implies that all firms offer a wage that is equal to the expected value $w(t) = E[v|t, g_i]$. Then, notice

$$E[v|g_i] = E[E[v|t, g_i]|g_i] = E[w|g_i] = \bar{w}_{g_i}$$

Which implies that $\bar{w}_{g_1} \neq \bar{w}_{g_2} \implies E[v|g_1] \neq E[v|g_2]$ $\square$

As we know that averages of wages across gender and race are different, we know that in the setup shown in our model, it is not enough to assume that there is a set of signals, available to all firms, that can explain wage gaps, while holding the underlying productivity distributions the same across groups. Therefore, to examine the potential importance of information in explaining the wage gaps, we make the relaxation in our model that different firms may observe different signals on workers. This introduces an additional component to the strategic wage setting. Namely, firms need to make a guess on the worker's outside option. This additional strategic consideration can create a divergent between workers' productivity and their marginal output and as a result, generate wide wage gaps across groups with otherwise identical productivity distributions.

2 Partial Identification of Productivity Distribution and Inference

Our objective in this paper is to examine how much of the observed wage gap between different groups can be explained by differences in information access. We therefore can ask whether there exists a single distribution of workers productivity, $\mu \in \Delta(\mathcal{V})$, that can induce the observed wage gaps, with some information structure. More formally, let $H(w|g_i)$ be the observed wage cumulative distribution function (CDF) of workers from group i,[3] let $\kappa_k(w|t_k, g_i) = \int_0^w \beta_k(w|t_k, g_i)$ be the CDF of firm k wage offers, conditioned on the firms' signals. Finally, let $\kappa(w|t, g_i) = \prod_{k=1}^J \kappa_k(w|t_k, g_i)$ be the predicted CDF for workers from group g, and some signal t. We want to examine whether the distribution of worker's productivity for workers of the two groups is the same. Specifically, we ask whether there exist two information structures and a distribution of workers productivity, such that, $\mu(v|g_1) = \mu(v|g_2) = \mu(v)$, and can generate the observed wage distributions, i.e.

$$H(w|g_i) = \int_{v,t} \kappa(w|t, g)\mathrm{P}(t|v, g)\mu(v)dvdt \tag{1}$$

As we are interested in the set of possible distributions μ that can generate the observed data, we now turn to explore how we can identify this set, within the basic model in section

[3]From here on, we suppress x for clarity

1. First, throughout our analysis, we assume that the econometrician has access to data on wages, worker demographics and worker characteristics, such as education level or experience.

Assumption 1. *The econometrician observes the joint distribution $H(w, g, x)$, and their induced conditional probabilities.* $H(w|g, x) \in \Delta(\mathcal{W})$.

This assumption on the data available to the researcher is true for a large share of the empirical labor literature, which uses data on workers' wages but does not have access to data on workers' wage offers or productivity. Next, we define the set of model predictions to be the set of wage distributions that can result from the auction game with some information structure.[4]

Definition 2.1. The set of BNE predictions, $H \in \Delta(\mathcal{W})$, for a given information structure S and productivity distribution μ, is the of wage distribution induced by a BNE in the auction game

$$Q(S, \mu) = \{H : H(w) = \kappa(w|s)\mathrm{P}(s|v)\mu(v)\}$$

We also make the following assumption on the data generating process

Assumption 2. *The observed wage distribution is a result of a Bayes Nash Equilibrium in the labor-market auction game*

This assumption is quite strong, as the model we consider here is fairly restrictive. It does not allow workers to choose where to work based on job characteristics, other than wage. The model also assumes that all firms are homogeneous in their production technology and can extract the same output from workers. Although these are restrictive assumptions, they stress how - in an economy with almost no firm heterogeneity - information differences alone can generate a wide range of diverse outcomes. Finally, we can define the identified set of workers productivity distribution as

$$Q^{BNE}(H) = \{\mu : \exists S \in \mathcal{S} \text{ such that } H(w) \in Q(S.\mu)\}$$

[4]For clarity, we omit the group g indicator and add it when needed.

This definition of the identified set may not seem useful because we need to iterate over all productivity distributions in $\Delta(\mathcal{V})$. For each distribution, we must find an information structure that induces the observed wage distribution. However, seminal results by Bergemann and Morris (Bergemann and Morris (2013), Bergemann and Morris (2016), Bergemann and Morris (2019b)) in information design and non-parametric estimation provide methods that transform this into a computationally feasible problem.

Before jumping to the result, it is worth introducing some notation. A game-form is a tuple $G = (\mathcal{W}, \mu)$ of the possible actions and prior distribution over the workers productivity. We define the a game to be the pair $(G, \mathcal{S})$.

Definition 2.2 (Bayes Correlated Equilibrium)**.** A joint distribution $\pi \in \Delta(\mathcal{V} \times \mathcal{W})$ is a Bayes Correlated Equilibrium of the basic form game $\mathcal{G}$, if for each firm j and wage offer w_j and deviation w'_j we have

$$\sum_{v} \sum_{w_{-j}} \left[(v - w_j) q(w_j, \boldsymbol{w_{-j}}) - S(v - w'_j) q(w'_j, \boldsymbol{w_{-j}}) \right] \pi(v, w_j, \boldsymbol{w}_{-j}) \geq 0 \quad \text{(Obedience Constraint)}$$

and the marginal of π with respect to the states is preserved

$$\sum_{w \in \mathcal{W}} \pi(v, \boldsymbol{w}) = \mu(v) \quad \text{(prior consistency)}$$

Bergemann and Morris, shows that the set of distribution of actions and states, $\pi \in \Delta(v, \boldsymbol{w})$, that can be induced by BNE of $(\mathcal{G}, \mathcal{S}')$, under some information structure $\mathcal{S}'$, is equivalent to the set of Bayes Correlated Equilibrium (BCE).

Theorem 1 (Bergemann and Morris (2016))**.** A distribution $\pi \in \Delta(\mathcal{V} \times \mathcal{W})$ that can arise as an outcome of a Bayes-Nash Equilibrium, under some information structure $\mathcal{S}$, if and only if it is a Bayes Correlated Equilibrium of the basic game $\mathcal{G}$

Next, we define the set of BCEs that can induce the observed wage distribution. Let π be a

BCE, and let

$$BCE(H) = \left\{ \pi : \sum_{\max(\boldsymbol{w}) \leq w} \sum_{v} \pi(v, \boldsymbol{w}) = H(w) \right\}$$

Similarly, we define set of productivity distributions, implied from the BCE, as the set of marginals over v

$$Q^{BCE}(H) = \{\mu : \pi \in BCE(H), \sum_{\boldsymbol{w}} \pi(v, \boldsymbol{w}) = \mu(v)\}$$

Using Theorem 1, we have the following proposition

Proposition 1. The set of productivity distributions from a Bayes-Nash Equilibrium in the auction game is equal to the set of productivity distributions from a Bayes-Correlated Equilibrium in the basic game:

$$Q^{BCE} = Q^{BNE}$$

Proof. The proof follows trivially from the fact that that the set of BCEs is a convex set and Theorem 1. □

Therefore, proposition 1 shows us that it's enough to look for all the joint distributions of wage offers and workers productivity that can induce the observed wage distribution and satisfy the obedience and prior consistency constraints. In Appendix A.2 we provide an illustrative example to show the identifying power of BCE in the case of one bidder.

2.1 Testing for the potential distorting effect of informational frictions

As discussed in the previous section, we want to see how much of the differences in the wage distribution can be attributed to information frictions. Following our discussion above, we can test whether a distribution μ can induce the observed wage distribution, with some information structure, by examining all the joint distributions π that have a marginal μ and satisfy the following constraints For every g we have

$$\begin{aligned} \forall j, w, w' : \sum_{\boldsymbol{w}_{-j}, v} \pi(v, \boldsymbol{w}|g) \Big[(v - w) q(w, \boldsymbol{w_{-j}}) - (v - w') q(w, \boldsymbol{w_{-j}}) \Big] \geq 0 \quad \text{(Obedience)} \\ \forall j : \sum_{v} \sum_{\boldsymbol{w} : w = max(\boldsymbol{w})} \pi(v, \boldsymbol{w}|g) = h(w) \quad \text{(Data-Match)} \end{aligned} \tag{2}$$

where $h(w)$ is the density function of H. The first constraint is the obedience constraint, which, together with the third constraint, assures us that the resulting joint distribution of actions and states is a BCE, and therefore, there exists some BNE, with some information structure, that can induce it. The data match constraint, makes sure that the BCEs we consider can induce the observed wage distributions in the data.

As we are interested in the extent in which information, and not other underlying differences across groups, drives the size of the wage gap, we can first check whether there exist $\pi(v, \boldsymbol{w}|g_1)$ and $\pi(v, \boldsymbol{w}|g_2)$, that satisfy the linear constraint in 2 and

$$\sum_{\boldsymbol{w}} \pi(v, \boldsymbol{w}|g_1) = \sum_{\boldsymbol{w}} \pi(v, \boldsymbol{w}|g_2) \quad \forall v \in \mathcal{V} \tag{3}$$

Finding $\pi(v, \boldsymbol{w}|g_1)$ and $\pi(v, \boldsymbol{w}|g_2)$ that satisfies (2) and (3) assures us that there exists a single distribution μ that, with some information structure, can induce the wage distributions of the two groups. If such a distribution exists, then we cannot rule out the possibility that the observed wage gap between the two groups is induced by differences in the information firms have before making a job offer. If we cannot find a distribution that satisfies 2 and 3, then the differences in wages across groups are not driven solely by information frictions, but must be driven also by differences in the underlying productivity distribution.

Further more, we can also quantify the potential distorting effect of informational frictions in the labor market by finding the distribution of workers' productivity, implied by π, satisfying (2) that has the smallest mean and compare it to the observed mean wage. This would give us an upper bound on the potential size of information in shaping the wage distribution.

Specifically, we want to measure

$$\begin{aligned} &\max \sum_{v} v \sum_{\boldsymbol{w}} \pi(v, \boldsymbol{w}|g) - \sum_{w} wh(w|g) \\ &s.t\ (2) \end{aligned} \tag{4}$$

The size of 4 gives us a bound on how much wages can diverge from the workers productivity and how much rents firms can extract from workers by utilizing their information.

2.1.1 Relation to other measures on discrimination

In the economics, discrimination is broadly categorized into statistical and taste-based paradigms, grounded in seminal works by Arrow (1973) and Becker (1957). Statistical discrimination involves decision-makers, commonly employers, using observable characteristics—such as race or gender—as heuristic proxies for unobservable attributes like skill or reliability, thereby generating biased outcomes (Phelps, 1972). Taste-based discrimination, by contrast, is rooted in the decision-maker's intrinsic preferences or prejudices against certain groups (Becker, 1957). Although these forms of discrimination have disparate motivations, both yield equivalently adverse impacts on marginalized populations.

Within the framework of our model, discrimination is entirely subsumed into the productivity distribution, μ. Specifically, if employers possess disutility in hiring from disadvantaged groups, this will manifest as a shift left in the distribution of productivity ν. Our metric for evaluating disparity is designed to answer the following question: In a setting where firms only have access to group membership information—assuming this constraint is also known to be shared by other firms—would wage offers be identical for individuals belonging to different groups?

Consequently, the model serves as a diagnostic tool: if we were to conclusively rule out the existence of a common productivity distribution across groups, it would imply that firms are incorporating race or group identity in their decision-making processes. Conversely, the identification of a parameter μ in line with our assumptions would suggest that the observed disparities between social groups could be attributed, at least partially, to additional information that firms possess either about individual productivity or competitive firms.

3 Computation and Inference

The set of joint distributions that satisfy (2) gives a tractable way to characterize the identified set of productivity distributions. Unfortunately, the size of π, the joint distribution, grows exponentially with the number of firms making a wage offer to the worker. For example, for a grid of size 15 and 10 firms, we need to keep track of 15^{11} variables. Therefore, If we represent the joint distribution as a vector of floats we would need around 35Gb of memory, and if we want to solve for the test for 3, we would need to hold twice as much memory. This clearly makes an analysis for a large number of players infeasible. Instead, we can use certain characteristics of the auction setup in order to reduce the dimensions of the problem.

We start by defining the set of identified productivity means to be

$$M = \{m = E[v; \mu] : \mu \in Q^{BNE}(H)\}$$

In Appendix A.1 we show that this set is convex. This implies that it's enough to identify $\max(M)$ and $\min(M)$ to describe this set. Next, we show that we can restrict attention to a set of bi-mass distributions, that puts a positive mass only on 0, the lower bound of the support of the wage distribution, and $\bar{w} = \max(\mathcal{W}_i)$, the upper bound.

Claim 2. Let $\mu \in Q^{BCE}(H)$, then there exists a $\tilde{\mu}$ with two mass points on 0 and $\bar{w}$ and mean $E[v; \tilde{\mu}] = E[v; \mu]$ such that $\tilde{\mu} \in Q^{BCE}(H)$

The proof of this claim, as all other claims in this section is in Appendix A.1. Next, we show that in order to check whether there exists a single distribution that can induce the wage distributions of two groups, then, it is sufficient to only check whether the set of means that can generate the wage distribution of one group, M_{g_1}, intersects with the set of means that can generate the wage distribution of the second group, M_{g_2}.

Claim 3. Let M_{g_i}, $i \in \{1,2\}$ be the set of identified means that can induce the wage distribution of group g_i. Then, there exists a distribution of worker productivity μ such that $\mu \in Q^{BCE}(H_{g_i})$ for $i \in \{1,2\}$ if and only if $M_{g_1} \cap M_{g_2} \neq \emptyset$. Also, the set of distribution means in $Q^{BCE}(H_{g_1}) \cap Q^{BCE}(H_{g_2})$ is contained in $[\max\{\underline{m}_{g_1}, \underline{m}_{g_2}\}, \min\{\overline{m}_{g_1}, \overline{m}_{g_2}\}]$ where $\overline{m}_{g_i} = \max(M_{g_i})$ and $\underline{m}_{g_i} = \min(M_{g_i})$.

Claim 2 and 3 and the fact that M is convex, implies that instead of characterising the entire set of possible distributions, we can just focus on finding M_{g_1} and M_{g_2} while restricting our search to a family of bi-mass distributions. This reduces the computational burden by, first, reducing the size of the joint distribution we need keep track of, and second, it allows us to solve the linear problem separately for each group and compare the set of identified means instead of solving the two problems together and require that (3) hold.

Finally, notice that the problem is grown exponentially with the number of players. We therefore want to solve a smaller problem, that takes advantage of our setup. We do it by taking advantage of the anonymous game structure of the auction game, and noticing that firms only care about the productivity of the worker, the highest wage offer, and the second highest wage offer. To take advantage of this we start by defining the object $p(w, w^1, n^1, w^2, n^2, v)$ which is the joint probability of a firm making a wage offer w, while the highest wage offer is w^1, the number of people who bid w^1 is n^1, the second highest wage offer is w^2 and n^2 is the number of firms who bid w^2. Notice that $p(w, w^1, n^1, w^2, n^2, v)$ has all the information needed to calculate the obedience and data match constraints. [5]

[5] Notice that by defining $p(w, w^1, n^1, w^2, n^2, v)$ to be a distribution over the order statistics, we have also impose the following trivial constraints

$$\begin{aligned}
& w^1 \geq w \\
& w^1 \geq w^2 \\
& \text{if } w^1 = w^2 \text{ then } n^1 = n^2 > 1 \\
& \text{if } w^1 > w^2 \text{ then } n^1 = 1, n^1 + n^2 \leq N \\
& \text{if } n^1 = n^2 = N \text{ then } w = w^1 \\
& \text{if } n^1 + n^2 = N \text{ then } w \in \{w^1, w^2\} \\
& \text{if } w^1 > w^2 \text{ then } w \notin [w^2, w^1]
\end{aligned} \tag{5}$$

Furthermore, $p(w, w^1, n^1, w^2, n^2, v)$ does not grow exponentially with the number of players and therefore it is easier to work with, for larger set of players. We therefore, want to show that we can express the set of $Q^{BCE}(H)$ in terms of this object.

To do so, we start by requiring that $p(w, w^1, n^1, w^2, n^2, v)$ satisfy the obedience constraint

$$\sum_{w^1,n^1,w^2,n^2,v} p(w, w^1, n^1, w^2, n^2, v)((v-w)q(w, w^1, n^1, w^2, n^2)-(v-w')q(w', w'^1, n'^1, w'^2, n'^2)) \geq 0 \quad \forall w, w' \tag{6}$$

where (w'^1, n'^1, w'^2, n'^2) is the first and second order statistics of the modified distribution, if a firm changes it's action from w to w'. We also require that it satisfy the data match constraint

$$\sum_{w,n^1,w^2,n^2,v} p(w, \tilde{w}, n^1, w^2, n^2, v) = h(\tilde{w}) \tag{7}$$

The next set pf constraints assures that we have enough players to play against w^1 and w^2, in a symmetric BCE[6]. The first constraint considers the case in which $w^1 > w^2$ and $n^1 + n^2 = N$

$$\frac{p(w^1, w^1, n^1, w^2, n^2, v)}{\binom{N-1}{n^1-1}\binom{N-n^1}{n^2}} = \frac{p(w^2, w^1, n^1, w^2, n^2, v)}{\binom{N-1}{n^2-1}\binom{N-n^2}{n^1}} \tag{8}$$

when $w^1 > w^2$, and $n^1 + n^2 < N$ we also require

$$\frac{p(w^1, w^1, n^1, w^2, n^2, v)}{\binom{N-1}{n^1-1}\binom{N-n^1}{n^2}} = \sum_{w<w^2} \frac{p(w, w^1, n^1, w^2, n^2, v)}{\binom{N-1}{n^1}\binom{N-1-n^1}{n^2}} \tag{9}$$

Finally, when $w^1 = w^2$, and $n^1 = n^2 < N$, we require that

$$\frac{p(w^1, w^1, n^1, w^1, n^1, v)}{\binom{N-1}{n^1-1}} = \sum_{w<w^1} \frac{p(w, w^1, n^1, w^1, n^1, v)}{\binom{N-1}{n^1}} \tag{10}$$

Denote by $BCEM(H)$ the set of marginals $p(w, w^1, n^1, w^2, n^2, v)$ that satisfy the above

[6] Claim 6 in the appendix shows that we can symmetrize any BCE when we use data only on the winning bids

constraints for a given observed wage distribution H

$$BCEM(H) = \left\{ p(w, w^1, n^1, w^2, n^2, v) : p(w, w^1, n^1, w^2, n^2, v) \text{ satisfies } (5) - (10) \right\}$$

and let $Q^{BCE_M}(H)$ be the implied set of productivity distributions

$$Q^{BCEM}(H) = \left\{ \mu \in \Delta(\mathcal{V}) : \sum_{w,w^1,n^1,w^2,n^2} p(w, w^1, n^1, w^2, n^2, v) = \mu(v) \text{ and } p \in BCEM(H) \right\}$$

The next claim shows that the set of productivity distributions implied by any marginal in $BCEM(H)$ is the same as the set of productivity distribution we that can be rationalize with a BCE.

Claim 4. $Q^{BCE}(H) = Q^{BCEM}(H)$

Therefore, instead of traversing the space of $BCEs$, we can use the restricted space of $BCEM$. This space does not grow exponentially with the number of players; rather, it grows quadratically, subject to additional constraints.

3.1 Inference

The identification arguments presented above assumed that we know the wage distribution H. However, when doing empirical analysis, we actually observe a an $i.i.d$ sample from the joint distribution $H(w, x, g)$, and therefore the analysis should take into account the sample variation. To do so, we follow the inference method suggested by Fang et al. (2020) for inference on linear systems with known coefficients. In what follows, we briefly describe the statistical test.

Given a $i.i.d$ sample of wages $\{w\}_i^n$ with w distributed according to $P \in \boldsymbol{P}$ Fang et al. (2020) show how to test the following hypothesis

$$H_0 : P \in \mathbf{P}_0 \quad H_1 : P \in \mathbf{P} \backslash \mathbf{P}_0$$

where

$$\mathbf{P}_0 \equiv \{P \in \mathbf{P} : \beta(P) = Ax \text{ for some } x \geq 0\}$$

where $A \in \mathbb{R}^{p \times d}$, with p as the number of constraints and d is the number of variables.[7] Fang et al. (2020) shows that in order to test whether x satisfies the linear problem, we can use the test statistics T_n

$$T_n \equiv \max \left\{ \sup_{s \in \hat{\mathcal{V}}_n^{\mathrm{e}}} \sqrt{n} \left\langle s, \hat{\beta}_n - A\hat{x}_n^{\star} \right\rangle, \sup_{s \in \hat{\mathcal{V}}_n^{\mathrm{i}}} \sqrt{n} \left\langle A^{\dagger} s, \hat{x}_n^{\star} \right\rangle \right\}$$

where $\hat{\beta}_n$ is an estimator for $\beta(P)$, which in our case is the density of the wage distribution and x_n^* is $A^{\dagger}\hat{\beta}_n$, in which $A^{\dagger}$ is the Moore-Penrose pseudoinverse of A. The test statics checks two types of violation - the first is whether $\hat{\beta}_n$ is in the range of A and the second is whether there exists $x \geq 0$, that solves the linear system. Fang et al. (2020) show how to calculate the critical value of the test by bootstrapping the sample $\hat{\beta}$ and solving a linear program at each iteration. The critical value they derive depends on a tuning parameter λ. We choose λ using the data-driven method they suggest.[8]

4 Data and Results

4.1 Data

We use the American Community Survey (ACS) 2010 sample to construct the wage distributions. We restrict our sample to individuals between the ages 21 and 65, who are in the labor force and are employed in the private sector. We remove self employed workers and restrict attention to workers who work full-time. We also remove people who earn at the top 1%. Figure 2 plots the wage distributions for white men, white women, black men and black

[7]It is known that any linear program with inequality constraints can be turned into a linear problem in standard form, in which all the inequalities are be written as equalities, with added slack variables. In our implementation we rewrite the linear problem in section 3 in its standard form

[8]It seems that different values of λ do not change the results by much

women, and table 1 shows some descriptive on the workers from different groups. It is quite apparent from both the figure and the table that the two distributions are very different, where the distributions of women and black workers are more concentrated at low values.

Finally, to solve the linear program in (3) we normalize we normalize the wage distribution to be between $[0, H \times \frac{2}{3}]$ and discretize the set of bids as $\{0, .., H \times 2/3\}$. We let $H = 15$ which implies that we allow the highest value worker to be 1.5 times the maximum wage in our sample ($240,000).

4.2 Results

4.2.1 Market Frictions

Figure 1 shows the upper and lower bound on the average productivity of workers in dollars, per year, by demographic group, and under the assumption that there N firms who make a job offer to the worker.[9] The figure shows that the upper bounds on the productivity of white men is higher than that for the other groups. The lower bound for all groups is given simply by the mean wage (since under complete information, the observed wage distribution is the productivity distribution).[10] The figure also shows that as the number of firms who are making a wage offer increases, and therefore, the competition among firms intensifies, the set of productivity distributions that can induce the wage distribution shrinks. Table 3 shows the difference between the upper and lower bounds for each group and under the assumption that there are N firms offering wage. Notice that as the lower bound is given by the mean wage, this table presents the results to (4) and gives information on the potential distorting effect of information. We can see that as the number of firms who compete for workers is smaller, the potential role information can play is larger. For example, the difference between the mean productivity of white men and their average wage can go up to $114,000, if each worker only receives two wage offers. On the other hand, if there are less search frictions in

[9] The bounds are showing the upper bound and lower bound of the confidence interval construct as described in section 4.1 and were calculated from the value of $N = \{2, 3, 5, 7, 10, 20, 50\}$

[10] The small decline in driven by sample noise and our inference method

the economy and each worker receives 50 wage offers, then the average wage can differ from the average productivity level by roughly \$80,000. These bounds are not very tight, as the 90% of the wage distribution for white men is \$96,000. But these bounds capture the large role information can have in shaping the wage distribution.

The shaded area in figure 1 describes the set of bi-mass productivity distribution that can explain all four wage distributions. As discussed above, the set of these distribution decreases as the number of firms increases. Therefore, we can conclude that without imposing additional assumptions on the information set of the firms, we can't rule out that information frictions alone can explain all of the wage gap in the data.

Next, we turn to make a set of assumptions on the information set of the agent. First, we assume that workers sort into occupations and that it is common knowledge among all firms what is each worker's occupation. Tables 6, 7, 8 show the bounds on the mean productivity across different occupations, for different groups. First, we can see that the set of possible mean productivity for each occupation is wide. For example, the productivity for white men working at management, business, science and arts occupations can generate, on average between \$76,525 and \$208,894 and on the other side, workers in production can generate on average between \$38,514 and \$133,732. Interestingly, we find that we cannot rule out that workers in all occupations have the same average productivity. In our setup, information can give rise to differences in workers' wage across occupations, even if the distribution of workers productivity is the same across all occupations. For example, firms might find it much harder to assess whether a worker is going to be a good manager or not than it would assessing whether a worker would do a good job in the assembly line. These differences in the available information to firms can generate the observed differences in wages, rather than self-selection of different quality workers or the role of each occupation in the production process.

Next, we impose the assumption that all firms observe the workers' experience and education level.[11] We divide the education level to three categories - High school dropouts, high-school graduates/have some college education, and workers with college degree. Similarly, we divide

[11]Following the convention, we define experience to be $Age - 6 - \text{School Years}$

the experience level into three groups - 0-6, 6-12 and more than 12 years. This partition captures the shape of the wage schedule, as discussed in Rubinstein and Weiss (2006).

A common practice in the empirical labor literature is to condition wages on both experience and education. This goes back to Mincer (1958), who rationalized the linear structure of the wage equation using compensation differential arguments. Later papers justify the inclusion of these variables in wage equation based on a human capital rationale (Heckman et al. (2005)), implying that workers' ability changes as they acquire education and experience on-the-job training. In most of these models, workers are being compensated by firms, which are assumed to observe the investments workers are making in human capital. In the framework we present, this amounts to an assumption on the information available to the firms. Specifically, we assume that all firms observe the workers investment in education and the experienced they gained.

Table 4 and 5 estimate the bounds on the mean productivity, under the assumption that firms observe a public signal on the workers education level and experience. Interestingly, we find that for relatively low level of competition, wage disparities between highly educated and experienced white men and uneducated and inexperienced white men cannot be explained by a single productivity distribution and different signals observed by the firms. This implies that, under the assumption that all firms observe workers' education level and experience, workers' ability differs between experienced educated workers and non experienced educated workers. We again cannot rule out that there are no differences between the four groups of workers.

In our latest exercise, we examine whether all the information frictions are driven by different selection patterns. In our model, selection and sorting into different industries can be thought of as components of the signal and information firms possess. For instance, when not conditioning on industry, part of the information firms hold may include the industry in which they operate, and the informational content of the industry is influenced by the varying selection patterns within these sectors. Consider a simplified model where both men and women are know their latent productivity. Further assume that women are discouraged from pursuing STEM fields before entering the labor market. In this scenario, the women who do choose STEM are likely to have higher latent abilities. In our framework, different market structures can be interpreted as signals. Consequently, we may wish to tighten the conditions of our test to exclude cases where selection is not informative, conditional on group membership. This would imply that $\mu(v|g, Industry) = \mu(v|Industry)$. In such a world, agents may choose industries differently, but these variances in selection are not stratified by group.

Table 9, 10, 11 present results on mean average productivity. Our findings suggest that we cannot dismiss the possibility that differences in average wages are influenced by factors other than selection. Thus, we demonstrate that information frictions can account for wage gaps even when selection patterns are consistent across industries.

To further strengthen the test, we investigate whether wage disparities can persist in the absence of selection across all industries and groups. For this, we require that $\mu(v|g, Industry) = \mu(v)$. In scenarios with two competing firms, the bounds on the mean wage distribution remain largely unchanged, falling between $(481266, 1318475)$, implying that other informational factors can continue to influence the observed wage gap.

5 Conclusion

In this paper, we explored the potential role information frictions play in shaping the wage gaps. We found that differences in average wages across white men, white women, black men and black women can be explained only by information frictions. This result differs from previous results in the statistical discrimination literature that argued that incomplete information is not enough to explain average wage gaps between groups of workers. We find that the simple model can generate the observed wage distribution without the need to argue for differences in the underlying productivity distribution of workers. This paper stresses the potential importance that information may have on the wage setting process. It implies that additional research is needed to understand what firms know about their job applicants and the applicants' outside option. Within the framework we use here, it will be interesting to explore further what assumptions we need to impose on the accuracy of the information firms have, to be certain that information frictions are not the sole reason for observed wage gaps. Also, leveraging the results from Bergemann et al. (2017) for the lowest possible revenue, over all information structures, we can try and see what is the largest wage gap possible that can be driven solely by differences in information. Finally, throughout the paper we make a strong assumption that firms know the number of competitive wage offers. Following Bergemann et al. (2021b) we can try and relax this assumption and see how this affects the set of identified distributions.

Tables and Figures

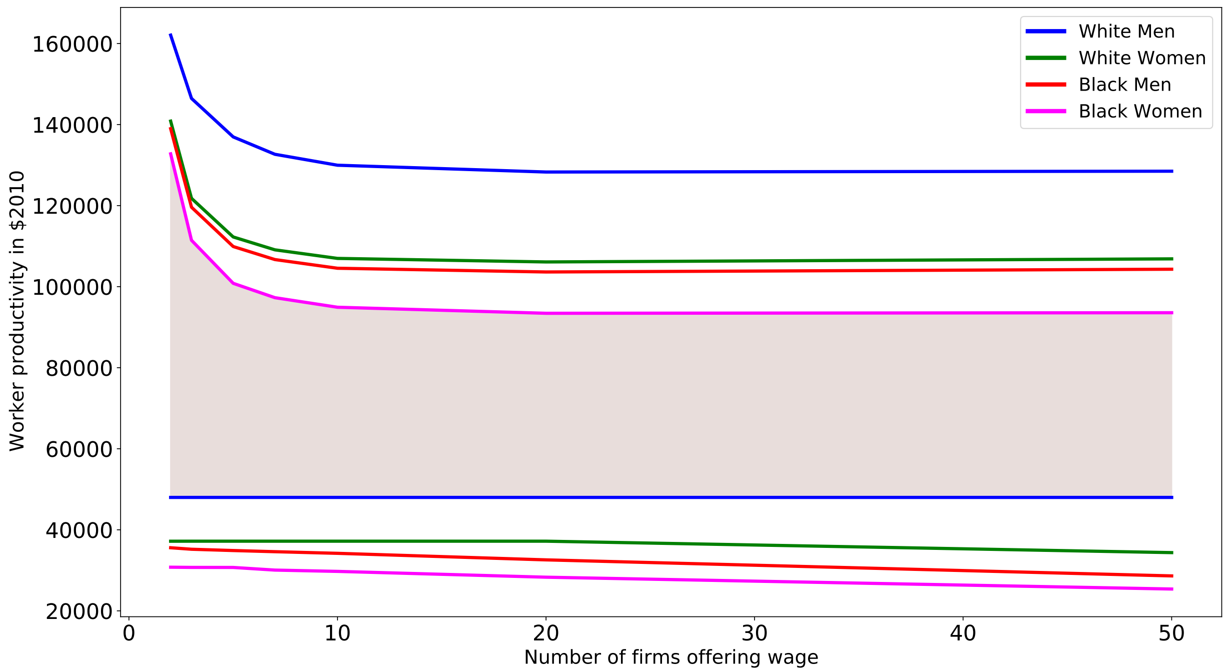

Figure 1: Upper and Lower bound on the average productivity of workers

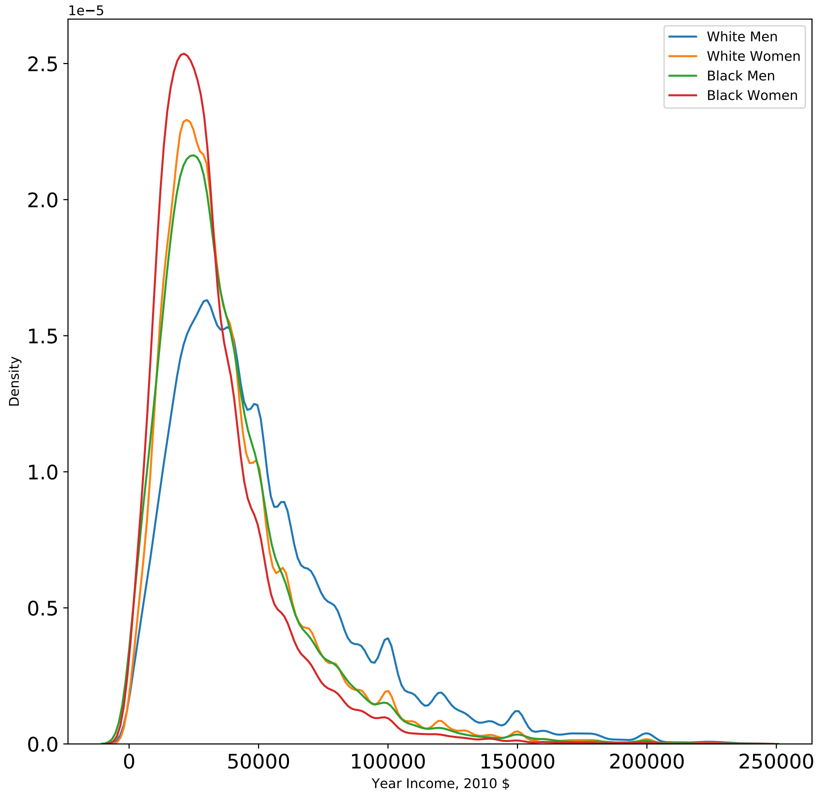

Figure 2: The four groups wage density

	WM	WW	BM	BW
Mean	48563.17	37554.15	36432.41	31341.16
Max	240000	240000	240000	240000
Min	100	10	160	270
5%	10000	8600	6011	6400
10%	14400	12000	11000	10000
25%	24000	20000	19200	16800
50%	40000	30000	30000	26000
75%	63000	48000	46996	40000
90%	96000	71000	70000	60000
95%	120000	90000	86000	74000

Table 1: Descriptive Statistics

Number of firms making wage offer	WM	WW	BM	BW
2	[47987,162073]	[37179,140853]	[35573,138985]	[30766,132775]
3	[47988,146452]	[37188,121794]	[35199,119605]	[30718,111444]
5	[47988,136953]	[37188,112263]	[34881,109901]	[30707,100821]
7	[47988,132665]	[37188,109088]	[34596,106686]	[30056,97281]
10	[47988,129968]	[37188,106971]	[34193,104542]	[29736,94920]
20	[47988,128291]	[37188,106100]	[32577,103624]	[28301,93431]
50	[47988,128493]	[34368,106859]	[28620,104303]	[25384,93569]

Table 2: The potential effect of information frictions - Lower and upper bound on mean productivity

Number of firms making a wage offer						
2	3	5	7	10	20	50
[47987,132775]	[47988,111444]	[47988,100821]	[47988,97281]	[47988,94920]	[47988,93431]	[47988,93569]

Table 3: The potential effect of information frictions - Lower and upper bound on the mean productivity of distribution who can explain the four wage distributions

		2 Firms			
Education Level	Experience	WM	WW	BM	BW
High School Dropout	0-6	[15319,92703]	[11081,87987]	[9011,96062]	[7092,94718]
	7-12	[20342,96624]	[16152,94052]	[16435,103435]	[14517,94652]
	> 12	[28838,122783]	[20210,92011]	[24767,117984]	[19899,95538]
High School Graduate	0-6	[21256,99727]	[18746,90545]	[17669,93798]	[16549,92353]
	7-12	[31603,126274]	[26236,111052]	[25679,119453]	[22478,102406]
	> 12	[45527,148205]	[34188,129723]	[35864,133170]	[29737,125400]
Colledge Degree	0-6	[40168,144552]	[35318,131746]	[31989,136923]	[30233,129410]
	7-12	[62895,181397]	[52576,164796]	[45543,158272]	[45392,148662]
	> 12	[82479,212782]	[62469,184971]	[63189,190784]	[54328,171938]
		7 Firms			
High School Dropout	0-6	[14029,66141]	[10583,57881]	[7281,65979]	[6608,62251]
	7-12	[19533,73363]	[15407,67793]	[14517,79088]	[11659,67656]
	> 12	[27643,89631]	[19345,68390]	[22643,89364]	[18391,71701]
High School Graduate	0-6	[21025,76537]	[18745,65859]	[17669,68383]	[15921,66302]
	7-12	[30797,92048]	[25448,84638]	[24225,88884]	[21381,80770]
	> 12	[45530,120607]	[34196,96999]	[34821,101871]	[28827,89957]
Colledge Degree	0-6	[38983,116040]	[34111,100075]	[28947,104308]	[28086,95252]
	7-12	[62879,148784]	[52584,130671]	[41184,126222]	[42260,123713]
	> 12	[82479,176048]	[62481,151496]	[61468,155905]	[51629,135734]

Table 4: Bounds on workers average productivity, conditional on education level and potential experience

		10 Firms			
Education Level	Experience	WM	WW	BM	BW
High School Dropout	0-6	[13402,64374]	[9210,55929]	[5595,63972]	[6352,60087]
	7-12	[18774,71814]	[14893,66074]	[13406,77559]	[10829,65936]
	> 12	[27146,87215]	[18788,66823]	[21884,86833]	[17799,70146]
High School Graduate	0-6	[20880,74991]	[18745,64213]	[17669,66683]	[15599,64571]
	7-12	[30210,89768]	[25181,82102]	[23488,86375]	[20952,79312]
	> 12	[45530,117837]	[34196,94823]	[34379,99804]	[28420,87588]
Colledge Degree	0-6	[37989,114158]	[33622,97962]	[27790,102131]	[27246,92988]
	7-12	[60372,145800]	[52582,128350]	[38531,123631]	[40364,120912]
	> 12	[82479,174027]	[62481,148403]	[58411,153232]	[50631,133584]
		20 Firms			
High School Dropout	0-6	[11840,63242]	[7294,54182]	[3492,62351]	[4524,58341]
	7-12	[17532,71316]	[12464,65073]	[9429,76783]	[7855,64738]
	> 12	[26031,85318]	[17814,65963]	[18931,84802]	[16250,69312]
High School Graduate	0-6	[19643,74645]	[17464,63243]	[14190,65654]	[13672,63339]
	7-12	[29067,88216]	[24222,79797]	[21368,84287]	[19135,78954]
	> 12	[45530,115852]	[34197,93631]	[33062,98990]	[27361,85758]
Colledge Degree	0-6	[35849,112219]	[32144,97114]	[23457,101523]	[21664,91743]
	7-12	[56893,143332]	[52569,127208]	[31380,121825]	[35756,118823]
	> 12	[82479,171734]	[62481,145958]	[52517,150815]	[45987,132912]

Table 5: Bounds on workers average productivity, conditional on education level and potential experience

	2 Firms			
Occupation Group	WM	WW	BM	BW
Management, Business, Science, and Arts Occupations	[76625,208894]	[57573,181048]	[58336,187206]	[48810,168396]
Business Operations Specialists	[65573,186839]	[52541,164298]	[48725,176692]	[43796,151843]
Financial Specialists	[71936,199559]	[52686,162749]	[52865,178582]	[43615,143603]
Computer and Mathematical Occupations	[74141,196949]	[63720,181718]	[60328,179537]	[52433,166268]
Architecture and Engineering Occupations	[71872,192156]	[55976,172515]	[62691,185354]	[53822,165105]
Life, Physical, and Social Science Occupations	[67111,196682]	[52802,175720]	[44551,171823]	[41623,154256]
Community and Social Services Occupations	[37646,138424]	[37043,130940]	[33372,132875]	[31944,125830]
Legal Occupations	[88543,235467]	[56902,174134]	[60966,225476]	[48260,181602]
Education, Training, and Library Occupations	[47662,163162]	[32599,132008]	[37783,157180]	[28892,132628]
Arts, Design, Entertainment, Sports, and Media Occupations	[53833,176646]	[44932,154469]	[44734,166339]	[41846,176973]
Healthcare Practitioners and Technical Occupations	[66053,197836]	[49731,156613]	[52702,179129]	[45163,153952]
Healthcare Support Occupations	[27874,130456]	[24848,100183]	[24823,121753]	[23327,100108]
Protective Service Occupations	[32309,139368]	[28109,131308]	[28286,133987]	[24715,120371]
Food Preparation and Serving Occupations	[20828,97493]	[17492,88966]	[19136,98924]	[16585,87479]
Building and Grounds Cleaning and Maintenance Occupations	[26093,116161]	[18498,89078]	[22015,105658]	[17688,89980]
Personal Care and Service Occupations	[30939,132934]	[21914,103643]	[23918,119923]	[20125,97484]
Sales and Related Occupations	[50945,168246]	[33413,140368]	[35719,143231]	[23720,115050]
Office and Administrative Support Occupations	[36085,136175]	[31668,123477]	[29733,127330]	[28946,122347]
Farming, Fishing, and Forestry Occupations	[27890,127262]	[39127,152968]	[25718,149189]	[39177,153329]
Construction and Extraction Occupations	[37479,138066]	[30028,136089]	[32630,135689]	[24688,144758]
Extraction Workers	[44358,157283]	[21782,152090]	[26276,163722]	NA
Installation, Maintenance, and Repair Workers	[42828,139650]	[37730,142002]	[37729,136710]	NA
Production Occupations	[38514,133732]	[26918,111167]	[32992,130707]	NA
Transportation and Material Moving Occupations	[33802,131167]	[24511,111238]	[29967,127651]	NA

Table 6: Bounds on workers average productivity, conditional on workers occupation

	7 Firms			
Occupation Group	WM	WW	BM	BW
Management, Business, Science, and Arts Occupations	[76625,169777]	[57577,145076]	[54116,153555]	[45329,135079]
Business Operations Specialists	[65448,154220]	[52099,131696]	[42098,138651]	[39345,122227]
Financial Specialists	[71918,163955]	[32485,130089]	[48050,142737]	[39945,117936]
Computer and Mathematical Occupations	[74051,156648]	[61596,146183]	[56233,147102]	[47310,131507]
Architecture and Engineering Occupations	[70479,154077]	[54610,138489]	[55688,150741]	[46836,132635]
Life, Physical, and Social Science Occupations	[65720,160237]	[51258,138811]	[39839,133847]	[37862,119501]
Community and Social Services Occupations	[33007,109672]	[34161,101851]	[26580,104162]	[29032,93117]
Legal Occupations	[88496,200619]	[56878,144154]	[53748,190271]	[38300,145431]
Education, Training, and Library Occupations	[45377,127193]	[31392,98486]	[33176,124857]	[26541,97419]
Arts, Design, Entertainment, Sports, and Media Occupations	[52800,139921]	[44850,127186]	[36617,131619]	[36937,138408]
Healthcare Practitioners and Technical Occupations	[65886,162302]	[48325,123964]	[45033,150164]	[42621,122434]
Healthcare Support Occupations	[24961,98421]	[23822,79817]	[22091,88241]	[22076,78861]
Protective Service Occupations	[29757,105633]	[24337,97460]	[26597,99728]	[21001,93328]
Food Preparation and Serving Occupations	[19804,74132]	[17088,63444]	[17599,75130]	[15758,61597]
Building and Grounds Cleaning and Maintenance Occupations	[25176,88096]	[17906,64631]	[20791,83971]	[16728,65258]
Personal Care and Service Occupations	[28405,98663]	[20846,81088]	[22266,85450]	[19376,73986]
Sales and Related Occupations	[50911,138401]	[32574,105960]	[32731,111553]	[21909,91027]
Office and Administrative Support Occupations	[36063,104420]	[30972,89282]	[27728,91958]	[27952,86430]
Farming, Fishing, and Forestry Occupations	[25201,99290]	[35471,123528]	[21461,112086]	[30778,126935]
Construction and Extraction Occupations	[36623,107669]	[26634,101684]	[29600,102368]	[24274,108122]
Extraction Workers	[40273,125597]	[18606,120786]	[14447,119779]	NA
Installation, Maintenance, and Repair Workers	[42799,113402]	[37258,113339]	[35077,107657]	NA
Production Occupations	[38518,104185]	[25519,88313]	[30712,97642]	NA
Transportation and Material Moving Occupations	[33793,98113]	[23298,86653]	[28609,92708]	NA

Table 7: Bounds on workers average productivity, conditional on workers occupation

	20 Firms			
Occupation Group	WM	WW	BM	BW
Management, Business, Science,				
and Arts Occupations	[76625,166430]	[57577,141935]	[45345,148821]	[37079,131489]
Business Operations Specialists	[58736,149263]	[46327,128267]	[29731,135836]	[31570,117638]
Financial Specialists	[71885,159846]	[45646,126686]	[33171,140255]	[33936,112715]
Computer and Mathematical Occupations	[68836,153449]	[56905,141257]	[46843,141379]	[38030,129238]
Architecture and Engineering Occupations	[67924,150277]	[46086,136264]	[45065,146401]	[34522,131654]
Life, Physical, and Social Science Occupations	[55480,155727]	[40652,135961]	[28514,130594]	[21564,115947]
Community and Social Services Occupations	[24490,107272]	[28945,99429]	[16578,102073]	[22298,90054]
Legal Occupations	[88496,195748]	[47229,140724]	[53748,184538]	[31877,142224]
Education, Training, and Library Occupations	[38981,123319]	[28460,95193]	[17513,120517]	[22330,94152]
Arts, Design, Entertainment,				
Sports, and Media Occupations	[47738,136981]	[37656,122293]	[24753,128216]	[26116,135459]
Healthcare Practitioners and Technical Occupations	[58297,157688]	[45870,119938]	[32726,146931]	[36208,117997]
Healthcare Support Occupations	[20014,94702]	[22574,78217]	[16147,83810]	[19959,77462]
Protective Service Occupations	[24897,102426]	[18697,93604]	[20706,95901]	[15724,88592]
Food Preparation and Serving Occupations	[18077,71971]	[15925,60509]	[15409,73057]	[14335,58548]
Building and Grounds Cleaning				
and Maintenance Occupations	[23477,83498]	[16375,61965]	[18349,78770]	[14886,62529]
Personal Care and Service Occupations	[24228,95010]	[19357,79344]	[19014,80651]	[17723,71946]
Sales and Related Occupations	[47882,134475]	[30151,102301]	[25232,108178]	[18490,85934]
Office and Administrative Support Occupations	[36060,101409]	[29988,85300]	[24901,88066]	[26061,82214]
Farming, Fishing, and Forestry Occupations	[20928,94898]	[30530,118689]	[11645,108258]	[18845,122541]
Construction and Extraction Occupations	[34674,105381]	[18908,98501]	[24187,99302]	[8308,101214]
Extraction Workers	[31598,120985]	[11504,117622]	[12112,114695]	NA
Installation, Maintenance, and Repair Workers	[39817,108685]	[26796,109516]	[30595,105905]	NA
Production Occupations	[35848,101749]	[23603,83685]	[27117,94170]	NA
Transportation and Material Moving Occupations	[30841,94733]	[20671,81744]	[25467,88790]	NA

Table 8: Bounds on workers average productivity, conditional on workers occupation

	2 Firms			
Industry	WM	WW	BM	BW
Manufacturing	[51603,165744]	[41228,147681]	[37974,140989]	[32236,135136]
Agriculture, Forestry, Fishing and Hunting	[26349,121153]	[20759,114528]	[20276,114465]	[13559,117558]
Mining, Quarrying, and Oil and Gas Extraction	[58573,182287]	[49893,174495]	[42515,171245]	[43006,180193]
Utilities	[68786,188346]	[55564,171634]	[52700,179344]	[45394,164217]
Construction	[41245,145932]	[38722,136342]	[32555,133411]	[34464,161730]
Wholesale Trade	[50023,161877]	[40677,146380]	[34408,134032]	[31644,139270]
Retail Trade	[38702,144191]	[29245,121394]	[29409,132853]	[23726,109194]
Transportation and Warehousing	[44925,151051]	[34796,130588]	[35069,135860]	[31672,128256]
Information	[59357,182527]	[48665,163438]	[47853,164229]	[40560,145176]
Finance and Insurance	[68880,198556]	[45656,150463]	[48799,168783]	[38032,136619]
Real Estate and Rental and Leasing	[45301,157944]	[38286,141775]	[31590,135128]	[28701,126959]
Professional, Scientific, and Technical Services	[73554,204013]	[50989,163977]	[54941,182658]	[45757,160898]
Management of Companies and Enterprises	[76108,221525]	[50109,178051]	[38003,209860]	[30729,161417]
Administrative and Support and Waste				
Management and Remediation Services	[34711,142735]	[30748,135058]	[27053,127743]	[25166,119261]
Educational Services	[44992,155291]	[34927,132773]	[35322,143380]	[33985,132721]
Health Care and Social Assistance	[50806,170414]	[36929,137531]	[34237,138185]	[30027,129365]
Arts, Entertainment, and Recreation	[34991,142072]	[27996,126447]	[26216,129311]	[22568,116629]
Accommodation and Food Services	[25733,115601]	[20555,100725]	[22556,113567]	[17749,97714]
Other Services (except Public Administration)	[35504,134105]	[23456,108539]	[27857,132699]	[20734,99251]

Table 9: Bounds on workers average productivity, conditional on workers industry

	7 Firms			
Industry	WM	WW	BM	BW
Manufacturing	[51612,134608]	[41231,118723]	[36726,110687]	[30235,101586]
Agriculture, Forestry, Fishing and Hunting	[24607,92681]	[18941,90768]	[17736,89932]	[8749,93443]
Mining, Quarrying, and Oil and Gas Extraction	[55265,148578]	[45856,141418]	[32556,129821]	[27710,140728]
Utilities	[66263,149144]	[51783,136749]	[48326,145166]	[38776,126950]
Construction	[41239,117385]	[35609,107135]	[29524,100374]	[29389,127405]
Wholesale Trade	[48910,133800]	[38893,117560]	[31052,102648]	[25541,105761]
Retail Trade	[38710,113214]	[28570,95350]	[27410,97872]	[22440,87595]
Transportation and Warehousing	[44052,123077]	[32990,98875]	[32898,105019]	[29800,95238]
Information	[59334,148093]	[46611,132647]	[43232,127500]	[35926,117606]
Finance and Insurance	[68880,162241]	[45670,124398]	[42958,136879]	[35404,107319]
Real Estate and Rental and Leasing	[43829,131150]	[35568,111745]	[25634,100984]	[24940,91391]
Professional, Scientific, and Technical Services	[73558,165286]	[50963,133384]	[52721,146898]	[42123,128791]
Management of Companies and Enterprises	[68437,184609]	[44744,140143]	[38003,188899]	[18735,114030]
Administrative and Support and Waste Management and Remediation Services	[33106,109539]	[29329,99664]	[24243,96567]	[23813,90303]
Educational Services	[43031,125904]	[33431,101282]	[32062,114458]	[30621,101375]
Health Care and Social Assistance	[50829,139442]	[36207,106364]	[30652,105565]	[28557,96175]
Arts, Entertainment, and Recreation	[34464,109715]	[25202,94023]	[22697,91891]	[19824,88929]
Accommodation and Food Services	[24510,92875]	[19695,76858]	[20192,90335]	[16580,72427]
Other Services (except Public Administration)	[33874,102865]	[22392,86758]	[24614,97725]	[19042,76866]

Table 10: Bounds on workers average productivity, conditional on workers industry

	20 Firms			
Industry	WM	WW	BM	BW
Manufacturing	[51612,130834]	[37980,116419]	[32591,108308]	[26179,98146]
Agriculture, Forestry, Fishing and Hunting	[21109,88130]	[14737,85484]	[13076,84088]	[3480,91745]
Mining, Quarrying, and Oil and Gas Extraction	[48265,145092]	[33985,138430]	[24853,127153]	[27710,134493]
Utilities	[61672,144757]	[44739,134383]	[33646,139736]	[28042,122922]
Construction	[38744,114453]	[31312,104486]	[24550,97369]	[17793,123590]
Wholesale Trade	[45915,129799]	[35197,115140]	[24300,99105]	[17443,102681]
Retail Trade	[36154,110269]	[26704,90931]	[23412,93794]	[18857,84367]
Transportation and Warehousing	[41544,118305]	[29460,95778]	[28707,102418]	[25914,92120]
Information	[53656,144592]	NA	[34904,123931]	NA
Finance and Insurance	[68880,157830]	[42065,119983]	[32769,133746]	[30948,105078]
Real Estate and Rental and Leasing	[37659,127280]	[31054,109052]	[18259,97315]	[18567,87449]
Professional, Scientific, and Technical Services	[73558,161522]	[46587,129410]	[41286,143796]	[35427,124724]
Management of Companies and Enterprises	[57786,180154]	[30070,137348]	[38003,187760]	[16930,106873]
Administrative and Support and Waste Management and Remediation Services	[29293,106325]	[25667,95805]	[19796,92246]	[19712,85836]
Educational Services	[36777,121483]	[30639,98317]	[23872,110801]	[25647,98710]
Health Care and Social Assistance	[43814,135558]	[34379,103685]	[24682,102088]	[26470,92274]
Arts, Entertainment, and Recreation	[26851,106333]	[21969,89930]	[16239,87826]	[14144,83966]
Accommodation and Food Services	[22004,87941]	[17967,74770]	[17136,85375]	[13866,69803]
Other Services (except Public Administration)	[31306,100058]	[19250,83238]	[17398,93708]	[16398,75103]

Table 11: Bounds on workers productivity, conditional on workers industry

A Appendix

A.1 Proofs

Claim 5. The set of identified means,

$$M = \{m = E[v; \mu] : \mu \in Q^{BNE}(H)\}$$

is convex.

Proof. fix $m^*, m^{**} \in M$ and choose $m \in [m^*, m^{**}]$ and λ such that $\lambda m^* + (1-\lambda)m^{**} = m$. We want to show that there exists a joint distribution π with marginal $\sum_w \pi(v, w) = \mu(v)$ and $E[v; \mu] = m$ such that μ is part of the identified set of distributions. Let π^* and π^{**} be two BCEs that induce H and have marginals μ^* and μ^{**} with the corresponding means. We can then define define π to be $\lambda\pi^*(v, \boldsymbol{w}) + (1-\lambda)\pi^{**}(v, \boldsymbol{w})$. Notice that for each v we have

$$\sum_w \pi(v, w) = \sum_w \lambda\pi^*(v, \boldsymbol{w}) + (1-\lambda)\pi^{**}(v, \boldsymbol{w}) = \mu(v)$$

Similarly, π satisfies the data match constraint

$$\begin{aligned}
\sum_v \sum_{\boldsymbol{w}:\max(\boldsymbol{w})=w} \pi(v, \boldsymbol{w}) &= \sum_v \sum_{\boldsymbol{w}:\max(\boldsymbol{w})=w} \lambda\pi^*(v, \boldsymbol{w}) + (1-\lambda)\pi^{**}(v, \boldsymbol{w}) \\
&= \lambda H(w) + (1-\lambda)H(w) \\
&= H(w)
\end{aligned}$$

and also the obedience constraint

$$\sum_v \sum_{w_{-j}} \pi(v, \boldsymbol{w})\Delta(w_j, w', w_{-j}, v) = \sum_v \sum_{w_{-j}} \lambda\pi^*(v, \boldsymbol{w}) + (1-\lambda)\pi^*(v, \boldsymbol{w})\Delta(w_j, w', w_{-j}, v) \geq 0$$

Therefore π is a BCE that induces the wage distribution H and $m \in M$ □

A.1.1 Proof of Claim 2

Proof. Let $\mu \in Q^{BCE}(H)$ and fix a π such that $\sum_w \pi(v, \boldsymbol{w}) = \mu(v)$ and π induces H. Then notice

$$
\begin{aligned}
0 \leq \sum_{v,\boldsymbol{w}_{-i}} \pi(v,\boldsymbol{w}) \Big[(v - w_k) q(w_k, \boldsymbol{w_{-k}}) - (v - w'_k) q(w_k, \boldsymbol{w_{-k}}) \Big] = \\
\sum_{v,\boldsymbol{w}_{-i}} p(\boldsymbol{w}) F(v|\boldsymbol{w}) \Big[v(q(w_k, \boldsymbol{w_{-k}}) - q(w_k, \boldsymbol{w_{-k}})) + (w_k q(w_k, \boldsymbol{w_{-k}}) - w'_k q(w'_k, \boldsymbol{w_{-k}}) \Big] = \\
\sum_{\boldsymbol{w}_{-i}} p(\boldsymbol{w}) \Big[E[v|\boldsymbol{w}] (q(w_k, \boldsymbol{w_{-k}}) - q(w_k, \boldsymbol{w_{-k}})) + (w_k q(w_k, \boldsymbol{w_{-k}}) - w'_k q(w'_k, \boldsymbol{w_{-k}}) \Big]
\end{aligned}
$$

We can therefore construct the following $\tilde{\pi}$ by equating the marginals $\sum_v \pi(v, \boldsymbol{w}) = \sum_v \pi(v, \boldsymbol{w})$ and defining

$$
\begin{aligned}
\tilde{\pi}(\bar{w}|\boldsymbol{w}) &= p \text{ s.t } p \times \bar{w} = E[v|\boldsymbol{w}] \\
\tilde{\pi}(0|\boldsymbol{w}) &= 1 - \pi(\bar{w}|\boldsymbol{w}) \\
\forall v \notin \{0, H\}, \tilde{\pi}(v|\boldsymbol{w}) &= 0.
\end{aligned}
$$

Notice that by construction $\tilde{\pi}$ satisfies both the obedience constraint and data match constraint and therefore $\tilde{\pi} \in Q^{BCE}(H)$. Finaly, let $\tilde{\mu} = \sum_{\boldsymbol{w}} \tilde{\pi}(v, \boldsymbol{w})$, and notice that due to the law of iterated expectations we have that $E[v; \tilde{\mu}] = E[v; \mu]$ as needed. $\square$

A.1.2 Proof of claim 3

Proof. The first direction is easy. If $M_{g_1} \cap M_{g_2} \neq \emptyset$ then we know that $Q^{BCE}(H_{g1}) \cap Q^{BCE}(H_{g2}) = \emptyset$. We prove the reverse direction by construction. Let $\mu_{g_1} \in Q^{BCE}(H_{g1}), \mu_{g_2} \in Q^{BCE}(H_{g2})$ and have the same mean $m_{g_1} = m_{g_2}$. We want to show that there exist at least one distribution that can rationalize both distributions. From claim 2, we know that we can construct a distribution $\tilde{\mu}_{g_i} \in Q^{BCE}(H_{g_i})$, with two mass points on the edges of the support and $m_{g_i} = E[v; \tilde{m}]$. Therefore we can construct two such distributions $\tilde{\mu}_{g_1}$ and $\tilde{\mu}_{g_2}$. But as $E[v; \tilde{mu}_{g_1}] = E[v; \tilde{mu}_{g_2}]$, then it must be $\tilde{\mu}_{g_1} \stackrel{d}{=} \tilde{\mu}_{g_2}$, as needed. Further notice that we can do this for each mean value in the interval $[\max\{\underline{m}_{g_1}, \underline{m}_{g_2}\}, \min\{\overline{m}_{g_1}, \overline{m}_{g_2}\}]$, which concludes the proof. ∎

A.1.3 Proof of claim 4

Before proving claim **??**, we show that if we only have access to wages, and not wage offers, it is without loss to restrict attention only to a symmetric (i.e. exchangeable) BCEs

Claim 6. For any $\pi \in BCE(H)$, there exists a symmetrized $\tilde{\pi}$ that is also in $BCE(H)$.

Proof. We want to show that there exist and exchangable BCE $\tilde{\pi}(v, \boldsymbol{w})$ that can induce the same winning bid. We show this by construction. Let Ξ be the set of permutations of $\{1, ...N\}$ and we associate each permutation with a mapping from $\mathcal{W}^N \to \mathcal{W}^N$ where $\xi(\boldsymbol{w})$ is a permuted profile of wage offers, in which $\xi_i(\boldsymbol{w}) = w_{\xi(i)}$. First, notice that any permuation of the players in a BCE is also a BCE. Then, fix $\pi \in BCE(H)$, and define define $\tilde{\pi}$ to be

$$\tilde{\pi}(v, \boldsymbol{w}) = \frac{1}{N!} \sum_{\xi \in \Xi} \pi(v, \xi(\boldsymbol{w}))$$

and notice that $\tilde{\pi}$ satisfies the obedience constraint and the prior consistency constraint and therefore a BCE. Further notice that it can generate the winning bid distribution

$$\begin{aligned}
\sum_{v} \sum_{\boldsymbol{w}:\max(\boldsymbol{w})=w} \tilde{\pi}(v, \boldsymbol{w}) &= \sum_{v} \sum_{\boldsymbol{w}:\max(\boldsymbol{w})=w} \frac{1}{N!} \sum_{\xi \in \Xi} \pi(v, \xi(\boldsymbol{w})) \\
&= \frac{1}{N!} \sum_{\xi \in \Xi} \sum_{v} \sum_{\xi(\boldsymbol{w}):\max(\xi(\boldsymbol{w}))=w} \pi(v, \xi(\boldsymbol{w})) \\
&= \frac{1}{N!} N! H(w) \\
&= H(w)
\end{aligned}$$

as needed. ∎

We can now show the proof for claim 4.

Proof. We start by showing that $Q^{BCEM}(H) \subseteq Q^{BCE}(H)$. let $\mu \in Q^{BCEM}$ and choose a $p(w, w^1, n^1, w^2, n^2, v) \in BCEM(H)$ that satisfies $\sum_{w,w^1,n^1,w^2,n^2} p(w, w^1, n^1, w^2, n^2, v) = \mu(v)$. We want to show that we can construct a symmetric BCE, π, which satisfies all $i \in \mathcal{N}$

$$\sum_{\substack{\pi: w_i = w, w^1 = \tilde{w}^1, w^2 = \tilde{w}^2, \\ n^1 = \tilde{n}^1, n^2 = \tilde{n}^2}} \pi(v, \boldsymbol{w}) = p(\tilde{w}, \tilde{w}^1, \tilde{n}^1, \tilde{w}^1, \tilde{n}^1, v) \tag{11}$$

Notice that such a BCE would clearly satisfy the obedience constraint and the data match constraint. Let

$$\Pi_i(\tilde{w}, \tilde{w}^1, \tilde{n}^1, \tilde{w}^2, \tilde{n}^2) = \Big\{ \boldsymbol{w} : \boldsymbol{w}_i = \tilde{w}, \boldsymbol{w}^1 = \tilde{w}^1, \boldsymbol{w}^2 = \tilde{w}^2,$$
$$|\{i : \boldsymbol{w}_i = \tilde{w}^1\} = \tilde{n}^1 |\{i : \boldsymbol{w}_i = \tilde{w}^2\} = \tilde{n}^2, \boldsymbol{w}_i \in \{w : p(w, \tilde{w}^1, \tilde{n}^1, \tilde{w}^2, \tilde{n}^2) > 0)\}, \Big\}$$

be the set of wage offers vectors in which firm i offers wage $\tilde{w}$, the other wage offers generate a distribution that satisfy the order statistics and includes only wage offers that are played with positive probability. Consider the case in which $w^1 > w^2$, $n^1 + n^2 = N$. Without loss of generality, we fix firm 1 and set for every joint probability of v, and $\boldsymbol{w} \in \Pi_1(w^1, w^1, n^1, w^2, n^2)$, the following

$$\pi(v, \boldsymbol{w}) = \frac{p(w^1, w^1, n^1, w^2, n^2, v)}{\binom{N-1}{n^1-1}\binom{N-n^1}{n^2}}$$

and for every v, $\boldsymbol{w} \in \Pi_1(w^2, w^1, n^1, w^2, n^2)$ set

$$\pi(v, \boldsymbol{w}) = \frac{p(w^2, w^1, n^1, w^2, n^2, v)}{\binom{N-1}{n^2-1}\binom{N-n^2}{n^1}}$$

The notice that that for each i and v and $w = w^1$ we have

$$\begin{aligned}
&\sum_{\boldsymbol{w}: w_i = w^1, w^1 = \tilde{w}^1 n^1 = \tilde{n}^1 w^2 = \tilde{w}^2 n^2 = \tilde{n}^2} \pi(v, \boldsymbol{w}) \\
&= \sum_{\boldsymbol{w}: w_1 = w^1, w_i = w^1, w^1 = \tilde{w}^1 n^1 = \tilde{n}^1 w^2 = \tilde{w}^2 n^2 = \tilde{n}^2} \pi(v, \boldsymbol{w}) + \sum_{\boldsymbol{w}: w_1 = w^2, w_i = w^1, w^1 = \tilde{w}^1 n^1 = \tilde{n}^1 w^2 = \tilde{w}^2 n^2 = \tilde{n}^2} \pi(v, \boldsymbol{w}) \\
&= \sum_{\boldsymbol{w}: w_1 = w^1, w_i = w^1, w^1 = \tilde{w}^1 n^1 = \tilde{n}^1 w^2 = \tilde{w}^2 n^2 = \tilde{n}^2} \frac{p(w^1, w^1, n^1, w^2, n^2, v)}{\binom{N-1}{n^1-1}\binom{N-n^1}{n^2}} \\
&+ \sum_{\boldsymbol{w}: w_1 = w^2, w_i = w^1, w^1 = \tilde{w}^1 n^1 = \tilde{n}^1 w^2 = \tilde{w}^2 n^2 = \tilde{n}^2} \frac{p(w^2, w^1, n^1, w^2, n^2, v)}{\binom{N-1}{n^2-1}\binom{N-n^2}{n^1}} \\
&= \sum_{\boldsymbol{w}: w_1 = w^1, w_i = w^1, w^1 = \tilde{w}^1 n^1 = \tilde{n}^1 w^2 = \tilde{w}^2 n^2 = \tilde{n}^2} \frac{p(w^1, w^1, n^1, w^2, n^2, v)}{\binom{N-1}{n^1-1}\binom{N-n^1}{n^2}} \\
&+ \sum_{\boldsymbol{w}: w_1 = w^2, w_i = w^1, w^1 = \tilde{w}^1 n^1 = \tilde{n}^1 w^2 = \tilde{w}^2 n^2 = \tilde{n}^2} \frac{p(w^1, w^1, n^1, w^2, n^2, v)}{\binom{N-1}{n^1-1}\binom{N-n^2}{n^1}} \\
&= \frac{p(w^1, w^1, n^1, w^2, n^2, v)}{\binom{N-1}{n^1-1}\binom{N-n^1}{n^2}} \binom{N-1}{n^1-1}\binom{N-n^1}{n^2} = p(w^1, w^1, n^1, w^2, n^2, v)
\end{aligned}$$

where the third equality comes from constraint (8). An equivalent argument shows that this holds for every i and w^2. Next, consider the case in which $w^1 > w^1$ and $n^1 + n^2 < N$. Let $\mathcal{W} = \left\{ w : w < w^2, p(w, w^1, n^1, w^2, n^2) > 0 \right\}$ and define for each v and $\boldsymbol{w} \in \cup_{w \in \mathcal{W}} \Pi_1(w, w^1, n^1, w^2, n^2)$

$$\pi(v, \boldsymbol{w}) = \frac{p(w, w^1, n^1, w^2, n^2, v)}{\binom{N-1}{n^1}\binom{N-1-n^1}{n^2}}$$

Set again, for every v, and $w \in \Pi_1(w^1, w^1, n^1, w^2, n^2)$, the following probability

$$\pi(v, \boldsymbol{w}) = \frac{p(w^1, w^1, n^1, w^2, n^2, v)}{\binom{N-1}{n^1-1}\binom{N-n^1}{n^2}}$$

and for every v, $w \in \Pi_1(w^2, w^1, n^1, w^2, n^2)$ set

$$\pi(v, \boldsymbol{w}) = \frac{p(w^2, w^1, n^1, w^2, n^2, v)}{\binom{N-1}{n^2-1}\binom{N-n^2}{n^1}}$$

Now, consider a firm i, making a wage offer $w \in \mathcal{W}$ and notice that

$$\sum_{\substack{\boldsymbol{w}: w_i = w, w^1 = \tilde{w}^1 n^1 = \tilde{n}^1 \\ w^2 = \tilde{w}^2 n^2 = \tilde{n}^2}} \pi(v, \boldsymbol{w}) =$$

$$\sum_{\substack{\boldsymbol{w}: w_1 = w^1, w_i = w, w^1 = \tilde{w}^1, n^1 = \tilde{n}^1, \\ w^2 = \tilde{w}^2, n^2 = \tilde{n}^2}} \pi(v, \boldsymbol{w}) + \sum_{\substack{\boldsymbol{w}: w_1 = w^2, w_i = w, w^1 = \tilde{w}^1 n^1 = \tilde{n}^1 \\ w^2 = \tilde{w}^2 n^2 = \tilde{n}^2}} \pi(v, \boldsymbol{w})$$

$$+ \sum_{w \in \mathcal{W}} \sum_{\substack{\boldsymbol{w}: w_1 = w, w_i = w, w^1 = \tilde{w}^1 n^1 = \tilde{n}^1 \\ w^2 = \tilde{w}^2 n^2 = \tilde{n}^2}} \pi(v, \boldsymbol{w})$$

$$= \frac{p(w, w^1, n^1, w^2, n^2, v)}{\binom{N-1}{n^1}\binom{N-1-n^1}{n^2}} \binom{N-1}{n^1} \binom{N-1-n^1}{n^2} = p(w, w^1, n^1, w^2, n^2, v)$$

where again we used the (8) and (9). An analogous argument would show that for each i the marginalization of our constructed BCE satisfies 11 for each firm which plays w^1 and w^2.

Next, consider the case in which $w^1 = w^2$ and $n^1 = n^2 < N$. define again $\mathcal{W} = \Big\{ w : w < w^2, p(w, w^1, n^1, w^2, n^2) > 0 \Big\}$ and set for each v and each $w \in \Pi(w^1, w^1, n^1, w^1, n^1)$

$$\pi(v, \boldsymbol{w}) = \frac{p(w^1, w^1, n^1, w^1, n^1, v)}{\binom{N-1}{n^1-1}}$$

similarly for each $\boldsymbol{w} \in \cup_{w \in \mathcal{W}} \Pi_1(w, w^1, n^1, w^2, n^2)$, and for each v define

$$\pi(v, \boldsymbol{w}) = \frac{p(w, w^1, n^1, w^1, n^1, v)}{\binom{N-1}{n^1}}$$

A similar argument to the previous ones would show that marginalizing over these distribution satisfy 11. Finally, for the case in which $w^1 = w^2$ and $n^1 = n^2 = N$ define

$$\pi(v, \boldsymbol{w}) = p(w^1, w^1, N, w^1, N, v)$$

which clearly satisfy 11. Notice that by construction π satisfies data match and the obedience constraints and therefore $\pi \in BCE(H)$. Further notice that by construction $\sum_{\boldsymbol{w}} \pi(v, \boldsymbol{w}) = \mu(v)$ and therefore $\mu \in Q^{BCE}(H)$ which shows $Q^{BCEM}(H) \subseteq Q^{BCE}(H)$

Next, we turn to show that $Q^{BCE}(H) \subset Q^{BCEM}(H)$. The argument are similar to the argument made to show the reverse direction, but we keep the proof here for completion. Fix $\mu \in Q^{BCE}$. Let $Q^{BCESYM}(H) = \{\mu : \exists \pi \in BCE(H), \sum_w \pi(v, w) = \mu(v), \text{ and } \pi \text{ symmetric } \}$. By claim 6 we know that $\mu \in Q^{BCESYM}(H)$ we can then show that $Q^{BCESYM}(H) \subset Q^{BCEM}(H)$. Fix a symmetric BCE $\pi \in BCE(H)$ such that $\sum_w \pi(v, w) = \mu(v) \forall v \in \mathcal{V}$. We can construct a $p(w, w^1, n^1, w^2, n^2)$ by marginalizing over π for a specific player. i.e. we define

$$p(\tilde{w}, \tilde{w}^1, \tilde{n}^1, \tilde{w}^2, \tilde{n}^2, v) = \sum_{\boldsymbol{w}: w_1 = w, w^1 = \tilde{w}^1, n^1 = \tilde{n}^1, w^2 = \tilde{w}^2, n^2 = \tilde{n}^2} \pi(v, \boldsymbol{w})$$

Notice that this construction immediately satisfies (6) and (7) and that the marginal of $\sum_{\tilde{w}, \tilde{w}^1, \tilde{n}^1, \tilde{w}^2, \tilde{n}^2} p(\tilde{w}, \tilde{w}^1, \tilde{n}^1, \tilde{w}^2, \tilde{n}^2, v) = \mu(v)$. To conclude the proof we need to show that $p(w, w^1, n^1, w^2, n^2)$ satisfies (8) - (10). To see that 8 is satisfied, let $X(w^1, n^1, x^2, n^2) \subset \{w^1, w^2, \underline{\mathrm{w}}\}^N$ be the set of vectors indicating which firm make a wage offer w^1, which make wage offer w^2 and who makes lower wage offer $\underline{\mathrm{w}} < w^2$, such that each vector satisfy $|\{i : x_i = w^1\}| = n^1$, $|\{i : x_i = w^2\}| = n^2$ and $|\{i : x_i = \underline{\mathrm{w}}\}| = N - n^1 - n^2$. Consider the case where $w^1 > w^2$ and $n^1 + n^2 = N$. Notice that due to symmetry we have that for each $x \in X(w^1, n^1, x^2, n^2)$ we have that $\sum_{\substack{\boldsymbol{w}: w_i = w^1 \forall i: x_i = w^1 \\ w_i = w^2 \forall i: x_i = w^2}} \pi(v, \boldsymbol{w}) = c$, where c is a constant. Then, notice that

$$p(w^1, w^1, n^1, w^2, n^2, v) = \sum_{\substack{\boldsymbol{w}: w_1 = w^1, \boldsymbol{w}^1 = w^1, \boldsymbol{n}^1 = n^1, \\ \boldsymbol{w}^2 = w^2, \boldsymbol{n}^2 = n^2}} = \binom{N-1}{n^1 - 1} \binom{N - n^1}{n^2} c$$

$$p(w^2, w^1, n^1, w^2, n^2, v) = \sum_{\substack{\boldsymbol{w}: w_1 = w^2, \boldsymbol{w}^1 = w^1, \boldsymbol{n}^1 = n^1, \\ \boldsymbol{w}^2 = w^2, \boldsymbol{n}^2 = n^2}} = \binom{N-1}{n^2 - 1} \binom{N - n^2}{n^1} c$$

which together implies (8). Similar line of arguments can show that this 9 and 10 are satisfied as well. Therefore, we show $p(w, w^1, n^1, w^2, n^2, v) \in BCEM(H)$ and therefore $\mu \in Q^{BCEM}(H)$ which concludes the proof. $\square$

A.2 Illustrative Example

To get a better intuition on the information contained in the observed wage distribution and the obedience constraints, we consider a simple illustrative example, with only a single firm making wage offers to workers. We assume that the worker productivity distribution lies on the finite support $\mathcal{V} = \{5, 10, 15\}$ and that the firm also offers wages from a finite set of wage offers $\mathcal{W} = \{5, 10, 15\}$. The marginal-profit for the firm from hiring a worker of type v, at wage w is $v - w$. Finally, we assume that workers accept the job offer, at wage w, only if the offered wage is $w >= v - 5$. Workers with $v = 5$ are willing to work for the firm at any wage $w \in \mathcal{W}$.[12] Let $p(10)$ and $p(5)$ be the share of workers who earn 10 and 5 in the data.[13] Before extending a wage offer, the firm observes certain signals on the worker productivity, $t \in \mathcal{T}$, which is unobserved by the analyst. Therefore the firm's interim-expected profit, by offering a wage W, is given by $\pi(w) = E[\mathbb{1}\{w > v - 5\}(v - w)|t]$. Let $F(w|t)$ be the wage setting rule for the firm, given the observed signal t, then a BNE satisfies that is F, such that for each w with $F(w|t) > 0$ we have $\pi(w) \geq \pi(w'), \forall w' \in \mathcal{W}$.

Using theorem 1, we can consider the set of possible distributions of v, by looking for a distribution of v, which satisfies the obedience and the data-match constraint. Specifically, let $\mathrm{P}(v, w)$ be the joint probability of observing a wage offer, w and a worker with productivity

[12] This reservation wage assumption assures us that the firm has an incentive to make wage offers higher than 5

[13] Notice that offering a wage of 15 is a dominated strategy, and therefore we don't expect to see workers with wage 15

v, then the obedience constraint gives us the following four inequalities

$$\begin{aligned} \mathrm{P}(15,10) &\geq \mathrm{P}(10,10)+\mathrm{P}(5,10) \quad (10 \to 5) \\ \mathrm{P}(10,5)+\mathrm{P}(5,5) &\geq \mathrm{P}(15,5) \quad (5 \to 10) \\ \mathrm{P}(15,10)+\mathrm{P}(10,10)+\mathrm{P}(5,10) &\geq 0 \quad (10 \to 15) \\ \mathrm{P}(10,5)+\mathrm{P}(5,5) &\geq 0 \quad (5 \to 15) \end{aligned} \tag{12}$$

where only the first two constraints bind. Now, consider that we want to derive bounds on the first moment of the workers productivity distribution. let $\mathrm{P}(v|w)$ be the probability of the state being v, given that the agent received a signal w. Then, using Bayes rule we can re-write these constraints as

$$\begin{aligned} \mathrm{P}(15|10) &\geq \mathrm{P}(10|10)+\mathrm{P}(5|10) \\ \mathrm{P}(10|5)+\mathrm{P}(5|5) &\geq \mathrm{P}(15|5) \end{aligned}$$

To derive the upper bound we can solve for

$$\begin{aligned} \max_{\mu(v)} \bar{E[v]} &= 15 \times \mathrm{P}(15)+10 \times \mathrm{P}(10)+5 \times \mathrm{P}(5) \\ &= 15 \times (\mathrm{P}(15|5)p(5)+\mathrm{P}(15|10)p(10)) \\ &+ 10 \times (\mathrm{P}(10|5)p(5)+\mathrm{P}(10|10)p(10)) \\ &+ 5 \times (\mathrm{P}(5|5)p(5)+\mathrm{P}(5|10)p(10)) \end{aligned}$$

Given the obedience constraint above and $\sum_v \mathrm{P}(v|w)=1$ for each w. Notice that in order to maximize the above expression, we want to push as much weight onto $\mathrm{P}(15|w)$. However, The second obedience constraint constrains us from doing so, while still having the firm bid 10. For the firm to bid 10, the probability of gaining positive profit must be larger then the probability of losing. Therefore, to solve the maximization problem, we can set $\mathrm{P}(15|10)=1$, $\mathrm{P}(15|5)=0.5$ and $\mathrm{P}(10|5)=0.5$, and get the following upper bound

$$\overline{E[v]} = 12.5p(5)+15p(10) = 15-2.5p(5)$$

Using a similar line of reasoning, and the first obedience constraint will give us the lower

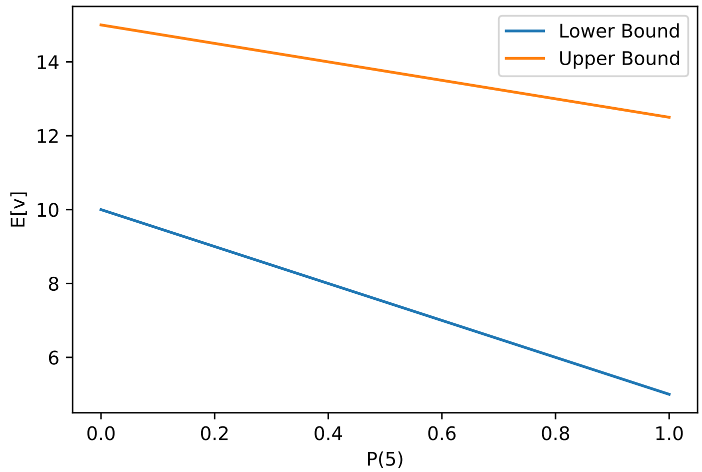

Figure 3: Upper and Lower bounds on the mean worker productivity in the single firm game

bound

$$\underline{\mathrm{E(v)}} = 5p(5) + 10p(10) = 10 - 5p(5)$$

From these bounds we can see that the data shows that only a small share of workers is earning high wages, then the distribution of workers cannot have too much weight on high values. And similarly, if the share of workers earning low wages is small, then it must be that there is a large share of workers with high productivity. Figure 3 below plots the upper and lower bound as a function of the $p(5)$.

Finally, notice that in this example we consider only one firm. In the general model introduced in section 1, the worker reservation wage was set by the other firms. This implies that actions on the part of one firm could not induce a profitable deviation in other firms. For example, consider an extreme case, in which we observe that the wage distribution is a degenerate distribution with point mass on 10. This can only be result of an equilibrium where both firms know that state is 10 with certainty, and therefore Bertrand competition pushes prices to 10. On the other hand, in the single firm example $E[v] \in [10, 15]$. The intuition for this is that in the reservation wage example, the reservation wage does not "react optimally" to the firms actions, and therefore, the set of possible outcomes is large. On the other hand, in the competitive environment, the firm can't only take into consideration the value of the worker but also needs to consider what the other firms will be willing to offer to the worker.

References

Aigner, D. J. and Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *ILR Rview*, 30(2):175–187.

Altonji, J. G. and Blank, R. M. (1999). Race and gender in the labor market. *Handbook of Labor Economics*, Chapter 48:3143–3259.

Altonji, J. G. and Pierret, C. R. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116(1):313–350.

Arrow, K. J. (1973). The theory of discrimination. *O. Ashenfelter and A. Rees (eds.), Discrimination in Labor Markets*.

Bergemann, D., Brooks, B., and Morris, S. (2017). First-price auctions with general information structures. *Econometrica*, 85(1):107–43.

Bergemann, D., Brooks, B., and Morris, S. (2021a). Counterfactuals with latent information. *Working Paper*.

Bergemann, D., Brooks, B., and Morris, S. (2021b). Search, information, and prices. *Journal of Political Economy*, 129(8):2275–2319.

Bergemann, D. and Morris, S. (2013). Robust predictions in games with incomplete information. *Econometrica*, 81(4):1251–1308.

Bergemann, D. and Morris, S. (2016). Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics*, 11(2):487–522.

Bergemann, D. and Morris, S. (2019a). Information design: A unified perspective. *Journal of Economic Literature*, 57(1):44–95.

Bergemann, D. and Morris, S. (2019b). Information design: A unified perspective. *Journal of Economic Literature*, 57(1):44–95.

Black, D. A. (1995). Discrimination in an equilibrium search model. *Journal of Labor Economics*, 12(2):309–24.

Blau, F. D. and Kahn., L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):281–302.

Caldwell, S. and Danieli, O. (2021). Outside options in the labor market. *Working Paper*.

Chambers, C. P. and Echenique, F. (2021). A characterisation of 'phelpsian' statistical discrimination. *The Economic Journal*, 131(637):2018–2032.

Coate, S. and Loury, G. C. (1993). Will affirmative-action policies eliminate negative stereotypes? *The American Economic Review*, 83(5):1220–40.

Fang, Z., Santos, A., Shaikh, A., and Torgovitsky, A. (2020). Inference for large-scale linear systems with known coefficients. *SSRN Electronic Journal*.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):1091–1119.

Gualdani, C. and Sinha, S. (2019). Identification and inference in discrete choice models with imperfect information. *arXiv preprint arXiv:1911.04529*.

Heckman, J. J., Lochner, L. J., and Todd, P. E. (2005). Earnings functions, rates of return and treatment effects: The mincer equation and beyond. *Working Paper Series. National Bureau of Economic Research*.

Lange, F. (2007). The speed of employer learning. *Journal of Labor Economics*, 25(1):1–35.

Lundberg, S. J. and Startz, R. (1983). Private discrimination and social intervention in competitive labor market. *The American Economic Review*, 73(3):340–47.

Magnolfi, L. and Roncoroni, C. (2017). Estimation of discrete games with weak assumptions on information. *Working Paper.*

Milgrom, P. and Weber, R. J. (1982). A theory of auctions and competitive bidding. *Econometrica: Journal of the Econometric Society*, pages 1089–1122.

Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy*, 66(4):281–302.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The american economic review*, 62(4):659–661.

Rubinstein, Y. and Weiss, Y. (2006). Chapter 1 post schooling wage growth: Investment, search and learning. *Handbook of the Economics of Education*, 11(2):1–67.

Shapley, L. S. and Shubik, M. (1971). The assignment game i: The core. *International Journal of Game Theory*, 1(1):30–111.

Spence, M. (1978). Job market signaling. *In Uncertainty in economics*, pages 281–306.

Syrgkanis, V., Tamer, E., and Ziani, J. (2021). Inference on auctions with weak assumptions on information. *arXiv preprint arXiv:1710.03830.*